

Dr. Ir. Agus Wibowo, M.Kom, M.Si, MM



KERANGKA KEBIJAKAN GLOBAL PEMAKAIAN AI (ARTIFICIAL INTELLIGENCE)



YAYASAN PRIMA AGUS TEKNIK



Dr. Ir. Agus Wibowo, M.Kom, M.Si, MM

KERANGKA KEBIJAKAN GLOBAL PEMAKAIAN AI (ARTIFICIAL INTELLIGENCE)



YAYASAN PRIMA AGUS TEKNIK

PENERBIT :
YAYASAN PRIMA AGUS TEKNIK
Jl. Majapahit No. 605 Semarang
Telp. (024) 6723456. Fax. 024-6710144
Email : penerbit_ypat@stekom.ac.id

ISBN 978-634-7227-36-2 (PDF)



9

786347

227362

Kerangka kebijakan Global pemakaian AI (Artificial Intelligence)

Penulis :

Dr. Ir. Agus Wibowo, M.Kom, M.Si, MM

ISBN : 978-634-7227-36-2 (PDF)

Editor :

Dr. Joseph Teguh Santoso, S.Kom., M.Kom.

Penyunting :

Dr. Mars Caroline Wibowo. S.T., M.Mm.Tech

Desain Sampul dan Tata Letak :

Irdha Yuniato, S.Ds., M.Kom

Penebit :

Yayasan Prima Agus Teknik Bekerja sama dengan
Universitas Sains & Teknologi Komputer (Universitas STEKOM)

Anggota IKAPI No: 279 / ALB / JTE / 2023

Redaksi :

Jl. Majapahit no 605 Semarang

Telp. 08122925000

Fax. 024-6710144

Email : penerbit_ypat@stekom.ac.id

Distributor Tunggal :

Universitas STEKOM

Jl. Majapahit no 605 Semarang

Telp. 08122925000

Fax. 024-6710144

Email : info@stekom.ac.id

Hak cipta dilindungi undang-undang

Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara
apapun tanpa ijin dari penulis

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas rahmat-Nya sehingga buku “*Kerangka Kebijakan Global Pemakaian Ai*” ini dapat terselesaikan dengan baik. Kecerdasan buatan (Artificial Intelligence, AI) merupakan salah satu inovasi teknologi terbesar yang mengubah berbagai bidang kehidupan secara fundamental. Dari sektor industri, kesehatan, pendidikan, hingga pemerintahan dan sosial budaya, AI hadir membawa janji akan efisiensi, inovasi, dan kemajuan yang belum pernah terjadi sebelumnya. Namun, di balik manfaat besar tersebut, penggunaan AI juga menimbulkan kompleksitas yang signifikan, terutama dalam hal etika, hukum, tanggung jawab, dan dampak sosial-ekologis.

Buku ini disusun dengan mengedepankan pendekatan multidisipliner yang mendalam, mulai dari sejarah awal dan renaissans kecerdasan buatan, berbagai teknologi inti seperti pembelajaran mesin dan pembelajaran mendalam, hingga isu-isu kontemporer yang sangat relevan seperti efek ekologis, berita palsu, serta sistem senjata otonom yang mematikan. Setiap bab dirancang tidak hanya untuk memberikan pemahaman teoritis, namun juga memuat studi kasus dan kajian praktis yang menggambarkan aplikasi nyata dan implikasi kebijakan dari teknologi AI.

Kerangka kebijakan yang diusung dalam buku ini menyoroti pentingnya tujuan etika dan manfaat sosial sebagai fondasi dalam penggunaan AI, dengan penekanan khusus pada akuntabilitas, transparansi, keadilan, keselamatan, serta perlindungan privasi dan hak kekayaan intelektual. Di tengah lompatan teknologi yang sangat cepat, perlindungan terhadap nilai-nilai kemanusiaan menjadi landasan utama agar AI dapat dimanfaatkan secara bertanggung jawab dan berkelanjutan.

Kami berharap buku ini dapat menjadi sumber referensi penting bagi pembuat kebijakan, akademisi, pengembang teknologi, dan masyarakat luas yang ingin memahami tidak hanya potensi tetapi juga risiko yang terkait dengan pemakaian AI secara global. Dengan bekal wawasan luas dan panduan yang terstruktur, diharapkan setiap pihak dapat berkontribusi dalam membangun tata kelola AI yang inklusif dan berkeadilan, sehingga teknologi ini menjadi alat yang mendukung kemajuan bersama dan kesejahteraan umat manusia.

Akhir kata, kami mengucapkan terima kasih kepada semua pihak yang telah berperan serta dalam proses penulisan dan penyusunan buku ini. Semoga karya ini dapat memberikan manfaat yang luas dan menjadi pijakan kuat dalam pengembangan kebijakan AI di masa depan.

Semangat dan Selamat membaca....!!!

Semarang, September 2025

Penulis

Dr. Ir. Agus Wibowo, M.Kom, M.Si, MM

Integritas kebijakan sebagai fondasi kecerdasan buatan yang bertanggung jawab di era globalisasi. Buku ini kami persembahkan kepada para pemikir, pengambil keputusan, dan pelaku inovasi yang terus berjuang menjadikan kecerdasan buatan sebagai alat untuk memperkuat kehidupan manusia, menjaga keberlanjutan, dan mewujudkan masyarakat yang inklusif serta adil. Semoga karya ini menginspirasi langkah-langkah bijak dalam menetapkan kebijakan AI di seluruh penjuru dunia. Memimpin masa depan dengan kebijakan yang bijak, demi kecerdasan buatan yang adil, bertanggung jawab, dan berkelanjutan.

Dr. Ir. Agus Wibowo, M.Kom, M.Si, MM

DAFTAR ISI

KATA PENGANTAR.....	ii
DAFTAR ISI.....	iv
BAB 1 PENDAHULUAN	1
1.1 Awal Mula Kecerdasan Buatan	1
1.2 Renaisans AI.....	2
1.3 Gofai Serta Pembelajaran Mesin Dan Mendalam	3
1.4 Jalan Ke Depan Dengan AI	8
BAB 2 TUJUAN ETIKA DAN MANFAAT SOSIAL	11
2.1 Kemurahan Hati Dan Ketidakbermaknaan	12
2.2 Transformasi Pekerjaan	13
2.3 Dampak Ekologis AI	20
2.4 Sistem Senjata Otonom Mematikan.....	25
2.5 Berita Palsu Dan Disinformasi Yang Memicu Kekerasan	29
2.6 Tujuan Etis Dan Manfaat Sosial	35
BAB 3 AKUNTABILITAS	39
3.1 AI Dan Status Hukum.....	40
3.2 Kerangka Akuntabilitas	45
3.3 Akuntabilitas Dan Tata Kelola	60
BAB 4 TRANSPARANSI DAN KLARIFIKASI	68
4.1 Hak Atas Transparansi Dan Keterjelasan	69
4.2 Studi Kasus.....	73
4.3 Isi Kewajiban Transparansi Dan Keterjelasan	77
4.4 Teknik Interpretabilitas Dalam AI	86
BAB 5 KEADILAN DAN NON-DISKRIMINASI.....	92
5.1 AI, Keadilan, Dan Diskriminasi	92
5.2 Risiko Dan Manfaat.....	94
5.3 Praktik Dan Standar Industri.....	101
5.4 Pertimbangan Kebijakan.....	103
5.5 Keadilan Dan Non-Diskriminasi	107
BAB 6 KESELAMATAN DAN KEANDALAN	110
6.1 Tantangan Etika Dan Moral Untuk Teknologi Berbasis AI.....	113
6.2 Studi Kasus.....	122
6.3 Tantangan Hukum Untuk Teknologi Berbasis AI.....	131
BAB 7 DATA TERBUKA DAN PERSAINGAN YANG ADIL	135
7.1 Akses Data Sebagai Input Utama Untuk Pengembangan AI.....	136
7.2 Model Pengembangan AI Dan Persaingan Inovasi Makro.....	147
7.3 Potensi AI Untuk Menimbulkan Kerugian Persaingan Usaha	152

BAB 8 PRIVASI	161
8.1 Dasar-Dasar Privasi	161
8.2 Masalah Privasi Dan AI	163
8.3 Menggunakan AI Untuk Memenuhi Persyaratan Privasi.....	179
BAB 9 AI DAN INTELEKTUAL PROPERTY	185
9.1 Melindungi Hak Dalam AI.....	188
9.2 Siapa Pemilik Hak Atas Karya Yang Dihasilkan Oleh AI?	192
9.3 Pelanggaran	199
9.4 Isu-Isu Kekayaan Intelektual Lainnya.....	205
DAFTAR PUSTAKA	212

BAB 1

PENDAHULUAN

1.1 AWAL MULA KECERDASAN BUATAN

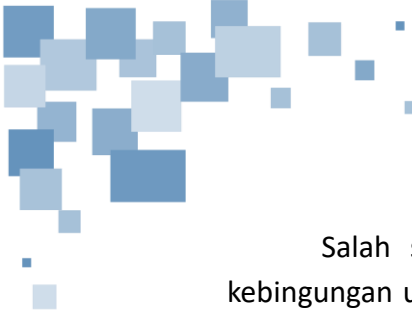
Impian menciptakan mesin cerdas sudah ada sejak lama. Pada abad ke-4 SM, para filsuf Yunani terkenal membayangkan "sebuah instrumen yang dapat menyelesaikan pekerjaannya sendiri" dan menyelamatkan manusia dari beban kerja. Namun, baru belakangan ini beberapa orang mulai percaya bahwa impian itu dapat menjadi kenyataan. Pada tahun 1950, ilmuwan Inggris Alan Turing menulis artikel terkenal di mana ia membantah argumen terhadap kemungkinan menciptakan mesin cerdas.

Sekitar periode ini, ilmuwan Amerika Marvin Minsky memulai proses mengubah teori menjadi kenyataan, mengembangkan apa yang secara luas diyakini sebagai mesin pembelajaran jaringan saraf pertama "SNARC" pada bulan Januari 1952. Tiga tahun kemudian, sekelompok ilmuwan Amerika dari industri dan akademisi menulis *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*, sekaligus menciptakan istilah "kecerdasan buatan" (AI). Dalam proposal ini, mereka berpendapat bahwa sebuah mesin dapat mensimulasikan "setiap aspek pembelajaran atau fitur kecerdasan lainnya," terutama "menggunakan bahasa, membentuk abstraksi dan konsep, memecahkan berbagai masalah yang sekarang diperuntukkan bagi manusia, dan meningkatkan diri mereka sendiri." Dengan optimis, mereka percaya bahwa mereka dapat mengembangkan mesin seperti itu hanya dalam satu musim panas.

Kombinasi antara tujuan ambisius dan optimisme tinggi ini menghasilkan ekspektasi tinggi. Misalnya, sejak tahun 1960an, para peneliti AI secara rutin berjanji bahwa mereka hampir berhasil membuat program komputer yang mampu mengalahkan pecatur manusia terbaik. Namun, tonggak penting ini baru tercapai beberapa dekade kemudian.

Perbedaan antara apa yang dijanjikan dan apa yang mereka hasilkan oleh para peneliti AI dengan cepat mengecewakan banyak orang. Pada tahun 1973, sebuah laporan untuk Dewan Riset Sains Inggris menyimpulkan bahwa "tidak ada satu pun bidang di mana penemuan yang telah dilakukan sejauh ini menghasilkan dampak sebesar yang dijanjikan saat itu." Pada tahun 1970an dan 1980an, pendanaan dari lembaga-lembaga penting seperti DARPA (Badan Proyek Penelitian Lanjutan Pertahanan Amerika Serikat) menurun dan penelitian di bidang AI disambut dengan skeptisisme yang semakin meningkat oleh ilmuwan komputer lainnya.

Dengan demikian, ketika sebuah program komputer, Deep Blue milik IBM, akhirnya mengalahkan seorang grandmaster catur manusia pada tahun 1997 dengan bantuan berbagai teknik dan algoritma AI, IBM menjauhkan diri dari penggunaan terminologi AI untuk menggambarkan kesuksesan mereka. Sebaliknya, IBM berpendapat bahwa Deep Blue bukanlah contoh kecerdasan buatan, melainkan sebuah "sistem pakar" tingkat lanjut, sebuah istilah yang umum pada saat itu untuk menggambarkan program komputer yang didasarkan pada pengetahuan para ahli manusia.



Salah satu faktor yang berkontribusi terhadap kekecewaan terhadap AI adalah kebingungan umum antara dua tujuan paralel AI: mereplikasi (atau melampaui) kecerdasan manusia sepenuhnya dan, yang kurang ambisius, membuat mesin mampu memecahkan sendiri masalah yang biasanya membutuhkan kecerdasan. Inilah perbedaan antara Kecerdasan Umum Buatan (AGI) dan "AI sempit".

Kecerdasan Umum Buatan (AGI) dan AI Sempit

AGI mengacu pada sistem AI yang dapat berperilaku seperti manusia, baik secara kognitif, emosional, maupun sosial. Di sisi lain, AI sempit bertujuan untuk memungkinkan mesin menguasai tugas-tugas spesifik tertentu, seperti permainan Go, mengenali wajah manusia dalam foto, atau mengemudikan kendaraan. AGI masih merupakan mimpi (atau mimpi buruk) yang jauh, tetapi seperti yang akan kita lihat, AI sempit sudah hadir dalam kehidupan sehari-hari kita.


1.2 RENAISSANS AI

Renaissans adalah masa dalam sejarah Eropa yang terjadi antara abad ke-14 hingga abad ke-17 Masehi, dikenal sebagai periode "terlahirnya kembali" budaya klasik Yunani dan Romawi. Periode ini menandai transformasi budaya besar setelah Abad Pertengahan, dengan fokus utama pada pemikiran humanisme, kebebasan dalam penelitian ilmiah (empirisme), dan penekanan pada rasionalitas. Gerakan ini bermula di Italia dan kemudian meluas ke seluruh Eropa, memberikan pengaruh signifikan pada seni, arsitektur, ilmu pengetahuan, politik, dan kesusastraan.

Saat ini, era kekecewaan terhadap AI terasa seperti kenangan yang jauh. Dalam dekade terakhir, minat terhadap AI terus meningkat di hampir setiap metrik: makalah ilmiah yang diterbitkan setiap tahun, pendaftaran mata kuliah di universitas, kehadiran konferensi pakar, jumlah perusahaan rintisan yang mengkhususkan diri di bidang ini, investasi swasta, dan liputan media.

Hampir setiap perusahaan teknologi besar, termasuk Facebook, Google, Microsoft, Alibaba, dan Netflix, menggunakan AI untuk memberi label pada foto kita dengan benar, menerjemahkan email kita, atau merekomendasikan film berikutnya yang akan kita tonton. AI juga telah menemukan tujuannya di ranah profesional. Pengacara, misalnya, kini bekerja dengan bantuan sistem AI untuk memprediksi keputusan pengadilan, menjawab pertanyaan hukum, atau menganalisis ribuan dokumen dengan cepat.

Alih-alih dianggap sebagai kekecewaan, sistem komputer yang diprogram untuk menjadi cerdas secara artifisial tiba-tiba dirayakan karena kemenangannya yang dipublikasikan dengan baik. Yang paling terkenal, manusia menjadi juara dalam acara permainan TV Jeopardy! kalah dari Watson milik IBM pada tahun 2011 dan AlphaGo milik Google Deepmind mengalahkan juara dunia Go pada tahun 2016. Sebagai prestasi impresif terakhir, AlphaStar, sistem AI yang juga dikembangkan oleh Deepmind, mengalahkan pemain manusia papan atas di StarCraft II, sebuah gim komputer strategi waktu nyata yang terkenal sulit.



Tidak terbatas pada bank server yang tidak bergerak, AI semakin banyak diintegrasikan ke dalam mesin bergerak. Pada tahun 2005, tim ilmuwan dan insinyur dari Universitas Stanford memenangkan Grand Challenge mobil swakemudi DARPA yang kedua. Kendaraan otonom mereka melaju sejauh 132 mil di medan gurun dalam waktu kurang dari tujuh jam, belajar mengemudi dengan bantuan teknik pemrograman AI. Saat ini, tentu saja, kendaraan otonom beredar di jalanan kota dalam uji coba yang dilakukan di lebih dari 50 kota di seluruh dunia, mengeksplorasi rasa kebebasan Gaak yang luar biasa dengan kecepatan 120 kilometer per jam!

Tingkat kemajuan yang dicapai di bidang AI telah begitu mengesankan sehingga, dalam upaya untuk melawan rasa antusiasme umum atas apa yang dapat dicapai AI, para spesialis AI, ilmuwan, jurnalis, dan beberapa wirausahawan teknologi telah memperingatkan bahaya teknologi ini, terutama aplikasi militernya, pelanggaran privasi, risiko diskriminasi, serta dampaknya terhadap ekonomi dan ketenagakerjaan. Penelitian ini akan mengusulkan kerangka kebijakan untuk AI yang bertanggung jawab yang bertujuan untuk membantu memastikan bahwa kita mendapatkan hasil terbaik dari teknologi AI, dan membatasi dampak buruk dari risiko yang diantisipasi ini. Namun, pertama-tama, ada baiknya untuk menjelaskan secara lebih rinci apa yang membedakan AI dari bentuk komputasi lainnya.

1.3 GOFAI SERTA PEMBELAJARAN MESIN DAN MENDALAM

Apa yang menjelaskan perubahan persepsi AI tersebut? Satu hal penting yang perlu dipahami adalah bahwa keberhasilan AI baru-baru ini sebagian besar disebabkan oleh pematangan dua pendekatan khusus dalam pembuatan algoritma: pembelajaran mesin ("ML") dan pembelajaran mendalam ("DL"). Secara kolektif, teknik-teknik ini mewakili inti dari gelombang kedua AI.


Berkat pembelajaran mesin dan pembelajaran mendalam dengan penekanan besar pada komponen “pembelajaran” komputer kini dapat melakukan jauh lebih banyak daripada sebelumnya.

Pembelajaran Mesin

“Pembelajaran mesin adalah teknologi yang memungkinkan sistem belajar langsung dari contoh, data, dan pengalaman.”

Pembelajaran Mendalam

Pembelajaran mendalam merupakan subbidang dari pembelajaran mesin (ML) yang terinspirasi oleh fungsi dan arsitektur otak manusia. Pembelajaran mendalam menggunakan jaringan neuron buatan yang tersusun dalam lapisan tampak dan tak tampak untuk memproses data masukan guna menghasilkan keluaran.



Sebelum menjelaskan bagaimana mesin dapat belajar dan mengapa hal itu begitu penting, kita harus mulai dari awal dan melihat gelombang pertama AI.

Kecerdasan Buatan Kuno

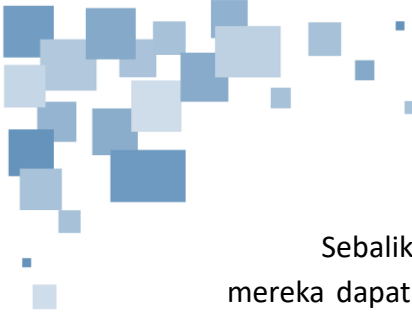
Pada tahun 1980an dan 1990an, gagasan umumnya adalah bahwa sistem AI dapat diprogram dengan mengodekannya menggunakan pengetahuan para ahli manusia. Richard Susskind, yang mengembangkan sistem pakar hukum pertama, menjelaskan prosesnya: "Ketika saya mulai bekerja di bidang AI dan hukum, satu pendekatan mendominasi pengetahuan dan proses penalaran para ahli hukum digali dari kepala mereka melalui proses wawancara yang dikenal sebagai 'penggalian pengetahuan'. Pengetahuan itu kemudian dikodifikasikan dalam pohon keputusan yang kompleks, dan dimasukkan ke dalam sistem komputer, yang dapat dinavigasi oleh pengguna non-ahli." Ini adalah proses yang panjang dan sulit, dengan hasil yang tidak pasti.

Deep Blue mungkin merupakan contoh paling terkenal dari sistem semacam itu dan merupakan representasi AI sebelum pembelajaran mesin, yang terkadang disebut sebagai AI Kuno yang Baik ("GOFAI") atau AI simbolis. Catur adalah permainan yang sulit, tetapi intinya terdiri dari daftar aturan formal yang relatif singkat. Pemrogram dapat menerjemahkan aturan-aturan ini ke dalam kombinasi simbol dalam bahasa komputasi formal (misalnya, pernyataan IF-THEN).

Diprogram dengan aturan-aturan yang dikodekan secara keras ini, yang dibutuhkan Deep Blue hanyalah daya komputasi yang sangat besar untuk menelusuri semua kemungkinan langkah, mengingat posisi papan tertentu, dan untuk mengevaluasi langkah selanjutnya dengan menerapkan fungsi yang disediakan oleh pemrogramnya. Selama pertandingan tahun 1997, Deep Blue mampu menelusuri hingga 330 juta posisi per detik dan beberapa strateginya dipersiapkan sebelumnya oleh grandmaster catur manusia. Namun, meskipun Deep Blue mengalahkan seorang grandmaster catur, hal itu sekaligus menunjukkan kepada kita batas-batas gelombang pertama AI. Ia diinformasikan tetapi juga dibatasi oleh cakupan pengetahuan manusia tentang aturan catur dan kemampuan pemrogram manusianya untuk mengekspresikan pengetahuan tersebut secara formal, dengan cara yang tidak terlalu mahal atau memakan waktu. Kenyataannya, Deep Blue tidak mencapai kemenangannya dengan "belajar", melainkan dengan mematuhi dan menjalankan serangkaian langkah yang telah diprogram sebelumnya.

Pembelajaran Mesin

Gelombang kedua AI telah mengusulkan pendekatan yang berbeda: bagaimana jika komputer dapat mempelajari, sendiri, solusi untuk masalah yang diberikan kepada mereka? Komputer selalu unggul dalam memecahkan masalah berdasarkan aturan formal dan matematika yang bisa sangat sulit dan memakan waktu bagi manusia. Tidak diragukan lagi bahwa "komputer jauh lebih unggul daripada manusia dalam tugas-tugas tertentu; kalkulator era 1970an dapat melakukan aritmatika lebih baik daripada manusia." Namun, beberapa tugas yang intuitif dan mudah bagi manusia seperti mengenali anjing di foto atau suara teman di telepon sangat sulit bagi komputer, karena tugas-tugas tersebut tidak bergantung pada aturan formal yang dapat dengan mudah dijelaskan dalam bahasa yang dipahami komputer.



Sebaliknya, algoritma pembelajaran mesin melampaui pengetahuan manusia karena mereka dapat "memperoleh pengetahuan mereka sendiri, dengan mengekstraksi pola dari data mentah," seperti pola dalam pemindaian MRI atau dalam posisi bidak di papan permainan. Dalam pembelajaran mesin, komputer tidak diprogram dengan seperangkat aturan, tetapi seperangkat fitur yang dipelajari oleh sistem AI untuk dikaitkan dengan berbagai hasil dan yang dapat diterapkan pada situasi baru. Dalam pembelajaran mesin konvensional, pemrogram perlu memilih fitur-fitur yang penting untuk algoritma untuk memperhatikan dalam rangka untuk memprediksi hasil yang tepat. Ini kadang-kadang disebut sebagai "rekayasa fitur."

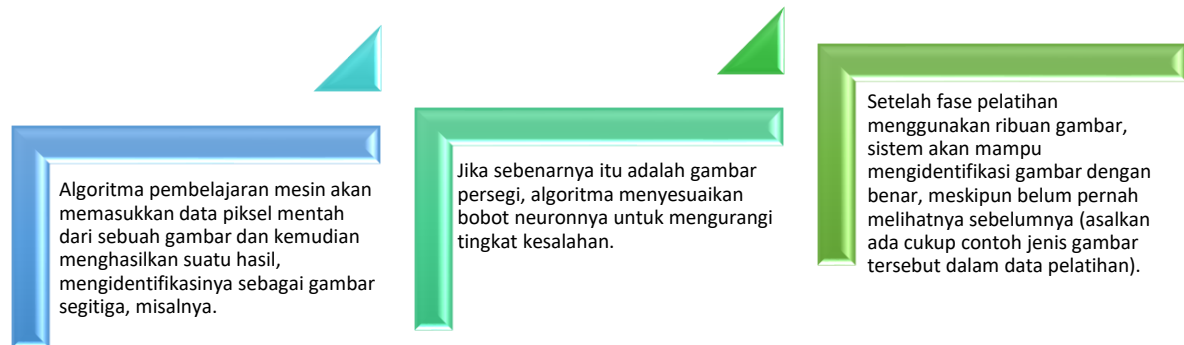
Mirip dengan GOFAI, proses ini membutuhkan pengetahuan ahli, meskipun tidak pada tingkat formalisme dan granularitas yang sama. Misalnya, bayangkan seseorang ingin tahu kemungkinan seorang siswa lulus ujian mengemudi dengan menganalisis data tentang siswa sebelumnya. Rekayasa fitur akan mengharuskan orang ini untuk memilih jenis informasi dari siswa sebelumnya yang relevan dengan masalah yang dihadapi. Jumlah jam yang dihabiskan untuk berlatih mengemudi tentu saja merupakan fitur yang relevan, sementara apa yang dimakan siswa di pagi hari mungkin tidak menentukan. Namun, setelah fitur dipilih, algoritma pembelajaran mesin dapat "mempelajari" kepentingan relatif dari setiap fitur tanpa campur tangan manusia lebih lanjut.

Jaringan Saraf Tiruan

Algoritma pembelajaran mesin didasarkan pada apa yang dikenal sebagai jaringan saraf tiruan yang terdiri dari beberapa simpul atau neuron yang terorganisasi dalam lapisan-lapisan terstruktur. Layaknya neuron biologis yang menjadi sumber inspirasinya, sebuah simpul menerima beberapa masukan yang dikombinasikan dengan bobot dan bias, menerapkan fungsi dengan serangkaian parameter, dan menghasilkan keluaran yang dikirim ke simpul-simpul lain hingga keluaran akhir dihasilkan.

Jaringan saraf tiruan biasanya terdiri dari sejumlah lapisan simpul, biasanya lapisan masukan, sejumlah lapisan tersembunyi, dan lapisan keluaran. Setiap simpul terhubung ke simpul-simpul pada lapisan di atas dan di bawahnya. Sistem AI menggunakan data pelatihan untuk membuat pola hubungan antara berbagai simpul di berbagai lapisan. Ketika sistem AI membuat kesalahan selama pelatihan, sistem secara otomatis menyesuaikan bobot dan bias di setiap simpulnya menggunakan teknik matematika sehingga fitur-fitur tertentu lebih diperhatikan, dengan AI yang secara sistematis menjadi lebih akurat dengan setiap potongan data baru.

Contoh pengenalan gambar



Pelatihan algoritma pembelajaran mesin (ML) dapat dilakukan menggunakan berbagai teknik. Dalam pembelajaran terawasi, pemrogram menggunakan data berlabel untuk melatih algoritma mereka dalam memprediksi hasil. Salah satu contohnya adalah program yang dapat mendeteksi anomali paru-paru berdasarkan pelatihan yang diterima melalui masukan ribuan radiografi dada yang sebelumnya diberi label normal atau abnormal oleh ahli radiologi manusia.

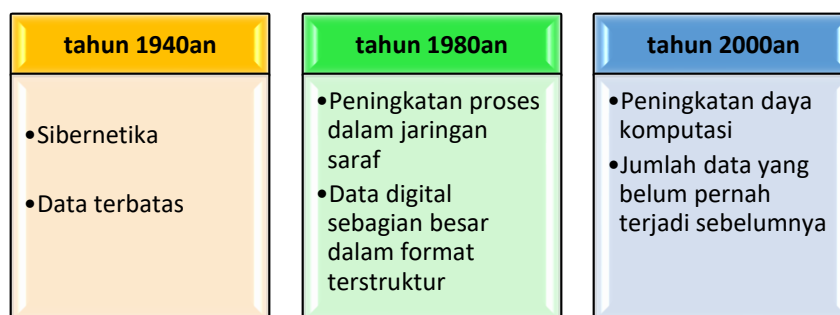
Sistem AI seperti ini merupakan alat penambangan data yang ampuh, karena dapat membantu manusia menemukan pola baru dalam data dalam jumlah besar pola yang bahkan mungkin tidak terdeteksi sendiri oleh ahli radiologi yang terampil, misalnya. Salah satu kelemahannya adalah proses pelatihan bisa mahal karena biaya pengumpulan, pelabelan, dan pemeriksaan data. Namun, biaya ini dapat dikurangi dengan ketersediaan dataset berlabel sumber terbuka yang lebih luas.

Di sisi lain, pembelajaran tanpa pengawasan menggunakan data tanpa label. Kegunaan utamanya adalah untuk membuat klaster data dengan karakteristik serupa. Untuk menggunakan kembali contoh radiologi, kasus pembelajaran tanpa pengawasan menyiratkan suatu algoritma yang menemukan kesamaan dalam radiografi tertentu dan mengelompokkannya. Namun, algoritma ini tidak akan mampu membedakan gambar mana yang normal dan mana yang abnormal, meskipun mungkin menemukan gugus kesamaan lain yang tidak terdeteksi oleh dokter atau analis manusia.

Jenis pelatihan pembelajaran mesin lainnya juga terinspirasi dari alam: pembelajaran penguatan pada dasarnya adalah penerapan teori anjing Pavlov pada AI. Jenis pembelajaran ini memungkinkan sistem AI dilatih dengan memberinya hadiah untuk setiap peningkatan yang dilakukan untuk mencapai suatu tujuan. Mencoba mengajari robot cara berjalan? Melalui penggunaan pembelajaran penguatan dan proses coba-coba, jatuh dan bangun, robot dapat belajar melakukannya sendiri. Namun, pengalaman telah menunjukkan bahwa algoritma pembelajaran mesin dengan hanya beberapa lapisan neuron (simpul) memiliki keterbatasan tertentu. Keterbatasan tersebut terkait dengan masalah representasi. Data mentah perlu direpresentasikan oleh serangkaian fitur yang diidentifikasi oleh pemrogram. Jika kita mengambil contoh pengenalan gambar sebelumnya, fitur yang akan diidentifikasi dapat mencakup jumlah sisi pada segitiga atau jumlah sudutnya. Citra hewan, objek, atau orang yang

lebih kompleks seringkali terlalu sulit untuk diidentifikasi sebelumnya atau untuk dibuat ciri-cirinya secara manual oleh pemrogram manusia. Sebagaimana dicatat oleh Prof. Yoshua Bengio, “mendesain fitur secara manual untuk tugas yang kompleks membutuhkan banyak waktu dan upaya manusia; dibutuhkan waktu puluhan tahun bagi seluruh komunitas peneliti.” Untuk melampaui kapasitas terbatas ini, spesialis AI membutuhkan solusi lain, yaitu solusi yang akan semakin menyingkirkan manusia dari proses pemrograman dan memungkinkan algoritma untuk mengidentifikasi ciri-ciri secara mandiri. Pembelajaran mendalam adalah jawabannya.


Pembelajaran Mendalam



Saat ini, pembelajaran mendalam dianggap sebagai subbidang dan perluasan dari pembelajaran mesin. Akar DL berawal dari tahun 1940an dalam bidang sibernetika sebagai cara untuk merepresentasikan aktivitas neuron, termasuk pembelajaran. Sebuah langkah penting bagi DL terjadi pada tahun 1980an ketika para ilmuwan menemukan cara untuk meningkatkan proses meminimalkan kesalahan bobot dan bias simpul pada lapisan tersembunyi jaringan saraf yang luas. Namun, metode ini sebagian besar masih kurang populer dalam ilmu komputer hingga sekitar tahun 2006 karena membutuhkan daya komputasi yang terlalu besar.

Dalam dekade terakhir, pertumbuhan eksponensial di berbagai bidang seperti daya pemrosesan, akurasi dan efisiensi sensor, penyimpanan data, kapasitas transfer dan pencarian, konektivitas, dan komputasi awan telah melepaskan kekuatan algoritma ini. Setelah perangkat keras canggih tersedia, DL terbukti bekerja sangat efektif pada tahun 2012, ketika digunakan dalam sistem AI yang memenangkan kontes pengenalan gambar dengan selisih yang besar dibandingkan dengan teknik lainnya.

Keragaman lapisan inilah kedalaman jaringan saraf yang membedakan DL dari ML. Dengan memfokuskan setiap lapisan neuron pada tingkat representasi tertentu, pembelajaran mendalam (DL) memperkenalkan abstraksi dalam AI. Dalam contoh pengenalan gambar klasik, lapisan pertama memasukkan piksel mentah dan mengidentifikasi tepinya. Lapisan kedua memasukkan tepi tersebut dan mendeteksi motif dengan kompleksitas yang lebih tinggi. Proses ini berulang untuk setiap lapisan hingga dihasilkan keluaran akhir yang memberi label pada gambar secara keseluruhan. Yang paling menonjol, fitur-fitur yang diidentifikasi oleh lapisan-lapisan ini "tidak dirancang oleh insinyur manusia: fitur-fitur tersebut dipelajari dari data menggunakan prosedur pembelajaran umum." Dengan kata lain, AI berbasis



pembelajaran mendalam dapat mempelajari sendiri apa yang membedakan kucing dari anjing, satu ras kucing dari ras kucing lainnya, dan kucing Anda dari semua kucing lain di dunia.

Kemajuan yang telah dicapai dalam AI berkat pembelajaran mendalam paling baik diilustrasikan dengan sebuah contoh.

Studi Kasus: AlphaGo Zero


Pada tahun 2017, tim ilmuwan dari Google DeepMind membangun sebuah program bernama AlphaGo Zero. AlphaGo Zero adalah adik dari AlphaGo, program yang mengalahkan pemain Go manusia terbaik. Jaringan saraf dalam AlphaGo dilatih menggunakan kombinasi pembelajaran terawasi dari permainan ahli manusia dan pembelajaran penguatan. Sedangkan untuk AlphaGo Zero, para ilmuwan mengembangkan algoritma baru dan melatihnya hanya menggunakan pembelajaran penguatan, membiarkannya belajar dengan bermain melawan dirinya sendiri. Adiknya mengalahkan kakaknya hanya dengan 36 jam pelatihan. Berdasarkan pendekatan ini, baru-baru ini, para ilmuwan membuat program yang lebih canggih lagi, yang juga berbasis pembelajaran mendalam, yaitu AlphaZero, yang dapat mempelajari permainan dari awal dengan bermain melawan dirinya sendiri, tanpa bantuan manusia atau pengetahuan yang diprogram secara kaku. Ketika diberi tugas untuk mempelajari catur dan permainan Jepang yang disebut shogi, ia berhasil mempelajari cara bermain pada tingkat yang tidak dapat ditandingi oleh manusia dalam waktu kurang dari 24 jam. Sebagai perbandingan, IBM membutuhkan waktu hampir sepuluh tahun percobaan dan kesalahan untuk mengembangkan program yang mengalahkan Garry Kasparov dalam permainan catur pada tahun 1997.

Teknik pembelajaran mendalam kini telah mendapatkan pengakuan luas dalam ilmu komputer, dan para pelopor pendekatan ini Yoshua Bengio, Yann LeCun, dan Geoffrey Hinton dianugerahi Penghargaan A.M. Turing 2018 oleh *Association for Computing Machinery* atas terobosan besar mereka dalam AI.

1.4 JALAN KE DEPAN DENGAN AI

Laju kemajuan yang semakin cepat dalam penelitian AI ini menggembirakan sekaligus mengkhawatirkan. Pembelajaran mesin dan pembelajaran mendalam kini dapat ditemukan dalam transportasi, layanan kesehatan, keuangan, perbankan, hukum, dan banyak bidang lainnya. Dampaknya terhadap masyarakat ada yang positif dan ada pula yang negatif. Menghadapi berbagai perkembangan yang dijelaskan di atas dan meningkatnya prevalensi AI dalam kehidupan sehari-hari, para spesialis AI telah menyerukan keterlibatan yang lebih besar dari komunitas hukum dalam mengembangkan regulasi, norma, dan kerangka kerja yang mengatur guna menghindari kekosongan hukum seputar penggunaan AI dalam bisnis, pemerintahan, dan kehidupan sehari-hari. Hukum AI belum berdiri sebagai bidang hukum yang berdiri sendiri, tetapi sedang dalam perjalanannya. Hukum AI akan sangat penting dalam membantu memodernisasi kerangka hukum dan kebijakan yang dengan cepat menjadi usang, tidak responsif terhadap peran AI yang semakin luas dalam masyarakat.

Sangat penting bahwa pedoman hukum dan kebijakan untuk pengembangan AI yang bertanggung jawab dikembangkan sekarang (di masa AI yang relatif baru) daripada nanti ketika



pada akhirnya mungkin "terlambat." Laju disrupsi teknologi begitu cepat saat ini sehingga banyak yang khawatir tertinggal. Banyak yang khawatir bahwa mereka tidak akan mampu beradaptasi dengan laju perubahan dan akan dipaksa untuk secara pasif dan reaktif "menyerah" pada dampak perubahan teknologi alih-alih berpartisipasi aktif dalam proses memandu perubahan teknologi untuk memastikan bahwa perubahan tersebut melayani kepentingan masyarakat secara keseluruhan.

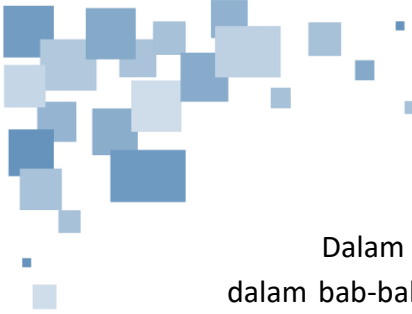
Matematikawan dan anggota Parlemen Prancis Cédric Villani berpendapat bahwa regulasi AI tidak boleh menjadi tanggung jawab peneliti dan insinyur saja pengacara harus mengembangkan "kesadaran sejati" tentang subjek tersebut." Dalam beberapa tahun terakhir, sejumlah pengacara telah memperhatikan tantangan ini, bekerja sama dengan para peneliti, anggota industri, organisasi hak asasi manusia, dan pemerintah untuk mengembangkan tanggapan dan rencana aksi berwawasan ke depan tentang cara terbaik untuk memanfaatkan potensi AI dan mengatasi masalah yang ditimbulkannya.

Patut dicatat, mayoritas tim ini mengambil pendekatan yang berpusat pada manusia. Misalnya, pada tahun 2016, *IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems* merilis draf laporan tentang desain yang selaras secara etis dalam pengembangan teknologi cerdas. Versi final laporan tersebut diterbitkan pada Maret 2019, yang mempromosikan sentralitas kesejahteraan manusia.

Pada Mei 2018, sebuah koalisi kelompok hak asasi manusia dan asosiasi teknologi merilis Deklarasi Toronto, yang menyerukan pemerintah dan perusahaan teknologi untuk memastikan kesetaraan dan non-diskriminasi tertanam dalam teknologi cerdas. Publikasi dan advokasi dari kelompok dan lembaga seperti Microsoft, Intel, Partnership on AI, Privacy International, University of Montreal, Stanford University, dan Uni Eropa sendiri, hanya untuk menyebutkan beberapa, membantu membentuk kerangka hukum dan kebijakan generasi berikutnya yang akan secara efektif mengatasi tantangan yang ditimbulkan oleh AI dan membantu meraup manfaat dari teknologi yang menjanjikan ini.

Pendekatan yang berpusat pada manusia ini juga dibenarkan mengingat beberapa batasan penting AI. Sistem AI berbasis pembelajaran mesin atau pembelajaran mendalam tidak "berpikir" atau "bernalari" secara konseptual, dan tidak memahami apa pun tentang ide, objek, atau isu yang mereka hadapi. Sebaliknya, jawaban yang tepat dihasilkan menggunakan pola yang terekam dalam jaringan saraf jawaban tersebut pada dasarnya diperoleh menggunakan teknik matematika yang mengidentifikasi keluaran yang paling mungkin benar, tanpa merefleksikan "makna" dari hasilnya, termasuk manfaat etisnya atau implikasi kebijakan sosialnya.

Singkatnya, pembelajaran mendalam bisa sangat akurat tanpa memahami apa pun bahkan, "pemahaman" tidak relevan karena sistem AI berfungsi dengan cara yang sama sekali berbeda. Poin ini penting karena merupakan alasan utama kesulitan dalam menjelaskan keluaran yang dihasilkan oleh sistem AI yang disebut masalah "kotak hitam". Batasan-batasan mendasar tersebut menunjukkan bahwa, terlepas dari harapan yang terus-menerus muncul, AI tidak akan menyelesaikan semua masalah kita sendiri dan bahwa AI hadir untuk melengkapi, alih-alih menggantikan, kecerdasan manusia.



Dalam konteks inilah kami menambahkan kontribusi kami sendiri dan menyajikan dalam bab-bab berikutnya kerangka hukum dan kebijakan awal untuk AI yang bertanggung jawab. Kerangka kerja kami terdiri dari 8 prinsip:

1. Tujuan Etis dan Manfaat Sosial;
2. Akuntabilitas;
3. Transparansi dan Keterjelasan;
4. Keadilan dan Non-Diskriminasi;
5. Keamanan dan Keandalan;
6. Data Terbuka dan Persaingan yang Adil;
7. Privasi; dan
8. AI dan Hak Kekayaan Intelektual.

Delapan bab berikutnya akan mengkaji secara mendalam kondisi AI saat ini dari perspektif masing-masing prinsip, dan diakhiri dengan draf konsolidasi kerangka kebijakan global yang diusulkan untuk AI yang bertanggung jawab.



BAB 2

TUJUAN ETIKA DAN MANFAAT SOSIAL

Hampir tiga puluh tahun yang lalu, filsuf Jerman Hans Jonas (lahir tahun 1903) menulis bahwa bagi pria dan wanita di generasinya, kehancuran Hiroshima dan perlombaan menuju senjata nuklir yang mengikutinya menandai awal dari refleksi baru yang memilukan tentang peran teknologi dalam urusan manusia.

Di satu sisi, energi nuklir dipuji sebagai sumber energi masa depan, yang lebih aman, lebih bersih, dan lebih andal, tetapi di sisi lain, daya rusak energi nuklir menciptakan ancaman baru, tidak hanya bagi umat manusia, tetapi juga bagi seluruh kehidupan di Bumi. Dihadapkan dengan pilihan-pilihan eksistensial seperti itu, manusia, sejak Hiroshima, telah memutuskan untuk tidak menggunakan energi nuklir yang dipersenjatai (kecuali sebagai ancaman utama) dan justru berusaha memanfaatkan tenaga nuklir secara bertanggung jawab.

Meskipun implikasi dari pengembangan kecerdasan buatan (AI) tentu saja tidak sedramatis daya hancur bom atom, penerapan AI yang semakin meluas juga disertai dengan euforia dan kekhawatiran yang mendalam. Di satu sisi, beberapa orang memandang algoritma pembelajaran sebagai obat mujarab, alat pamungkas untuk memecahkan semua masalah umat manusia, mendorong masa depan yang cerah di mana kendaraan listrik tanpa pengemudi hampir menghilangkan kemacetan lalu lintas dan polusi, di mana masalah medis yang sulit diatasi dapat diselesaikan dengan mudah dan pekerjaan kasar menjadi peninggalan sejarah.

Di sisi lain, CEO Tesla dan SpaceX, Elon Musk, berpendapat bahwa "AI merupakan risiko mendasar bagi keberadaan peradaban manusia." Lebih lanjut, bahkan mending Stephen Hawking memperingatkan pada tahun 2017: "Keberhasilan dalam menciptakan AI yang efektif, bisa menjadi peristiwa terbesar dalam sejarah peradaban kita. Atau yang terburuk." Jika kekhawatiran yang mengkhawatirkan ini hanya dimunculkan oleh orang-orang yang kurang informasi, mungkin cukup mudah untuk mengabaikannya begitu saja. Namun, mengingat potensi dampak AI, pandangan euforia dan mengkhawatirkan ini patut mendapat perhatian serius kita.

Dalam konteks ini, salah satu tujuan makalah ini adalah untuk memperkuat argumen yang mendukung kerangka kerja prinsip-prinsip inti yang sedang berkembang untuk pengembangan, penerapan, dan penggunaan AI yang bertanggung jawab. Kerangka kerja ini akan membantu memastikan bahwa sistem AI digunakan dengan cara yang meminimalkan risiko bahaya sekaligus memaksimalkan kemungkinan hasil positif tidak hanya bagi umat manusia, tetapi juga bagi lingkungan. Masing-masing prinsip inti ini dibahas dalam bab-bab berikutnya. Prinsip-prinsip ini berfungsi untuk membantu memastikan bahwa organisasi yang telah memilih untuk menerapkan sistem AI melakukannya secara bertanggung jawab. Namun, sebelum beralih ke prinsip-prinsip inti lain yang terkait dengan implementasi AI yang bertanggung jawab, bab ini membahas pertanyaan awal yang lebih mendasar: "untuk tujuan apa?" Kami menganggap bahwa prasyarat untuk "AI yang bertanggung jawab" adalah bahwa

sistem AI akan diimplementasikan untuk tujuan etis, tujuan yang memiliki manfaat sosial yang nyata dan wajar.

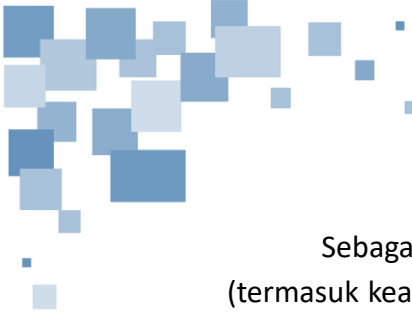
2.1 KEMURAHAN HATI DAN KETIDAKBERMAKNAAN

Selama bertahun-tahun, "Kode Etik" Google memuat motto berikut (yang sekaligus bernada jenaka dan sangat serius): "Jangan berbuat jahat." Setelah restrukturisasi perusahaan Google di bawah Alphabet Inc. pada tahun 2015, motto tersebut diganti dengan "Lakukan hal yang benar." Motto-motto ini pada dasarnya merupakan pernyataan ulang yang "menarik" dari dua prinsip etika inti: "kemurahan hati" ("berbuat baik") dan "tidak merugikan" ("tidak merugikan").

Kedua prinsip ini, tentu saja, merupakan nasihat etika yang tak terbantahkan. Sayangnya, jika berdiri sendiri, keduanya hanya memberikan panduan praktis yang sangat terbatas tentang cara menjalankan urusan seseorang! Terkait dengan AI, kami ingin berpartisipasi dalam dialog yang sedang berlangsung mengenai pertimbangan etika yang harus tetap menjadi pusat pengembangan sistem AI di masa mendatang.

Penelitian AI yang Etis di Seluruh Dunia
"Deklarasi Montreal untuk Pengembangan Kecerdasan Buatan yang Bertanggung Jawab" mengembangkan serangkaian prinsip etika yang ditujukan kepada "setiap orang, organisasi, dan perusahaan yang ingin berpartisipasi dalam pengembangan kecerdasan buatan yang bertanggung jawab", termasuk prinsip "kesejahteraan", prinsip "penghormatan terhadap otonomi", prinsip "kesetaraan", dan prinsip "partisipasi demokratis".
Kelompok Pakar Tingkat Tinggi Komisi Eropa untuk Kecerdasan Buatan, pada bagiannya, mendukung pendekatan yang berpusat pada manusia dalam Rancangan Pedoman Etika untuk AI yang Tepercaya: "AI harus dikembangkan, diterapkan, dan digunakan dengan "tujuan etis", yang didasarkan pada, dan mencerminkan, hak-hak asasi, nilai-nilai sosial, dan prinsip-prinsip etika Kebaikan, Ketidakberpihakan, Otonomi manusia, Keadilan, dan Kejelasan."
Komisi Perlindungan Data Pribadi Singapura sepakat bahwa keputusan yang dibuat oleh sistem AI harus berpusat pada manusia dan selaras dengan prinsip kebaikan dan ketidakberpihakan prinsip.
Perusahaan swasta seperti IBM dan Google, serta organisasi profesional seperti IEEE, juga telah mempertimbangkan pendekatan yang berpusat pada manusia ini.

Kami mengapresiasi upaya-upaya untuk mengeksplorasi implikasi etis AI ini. Lebih lanjut, kami mendukung proposal-proposal ini untuk menempatkan kemanusiaan di pusat diskusi tentang AI, memastikan bahwa teknologi canggih digunakan secara bertanggung jawab demi kebaikan kondisi manusia. Teknologi tidak boleh menggantikan agensi manusia. Namun, kami menganggap bahwa mempromosikan pandangan "berpusat pada manusia" yang sempit tentang AI yang etis dan bertanggung jawab mungkin terlalu membatasi karena terlalu menekankan kemajuan ekonomi dan teknologi jangka pendek manusia sehingga merugikan landasan lingkungan yang menopang kehidupan di Bumi. Oleh karena itu, untuk memperjelas, setiap referensi terhadap pandangan "berpusat pada manusia" dalam teks ini dimaksudkan cukup luas untuk mencakup isu-isu terkait keberlanjutan lingkungan, perlindungan dan promosi keanekaragaman hayati, serta kebutuhan untuk memerangi perubahan iklim secara bertanggung jawab.



Sebagaimana akan dibahas dalam bab-bab berikutnya, sejumlah konsep etika (termasuk keadilan, non-diskriminasi, dan promosi agensi manusia) tertanam dalam prinsip-prinsip "AI yang Bertanggung Jawab". Untuk sisa bab ini, fokus kami akan tertuju pada empat area di mana dampak transformatif AI yang potensial saat ini menjadi bahan perdebatan masyarakat yang signifikan. Kami mulai dengan diskusi tentang potensi dampak AI di tempat kerja.

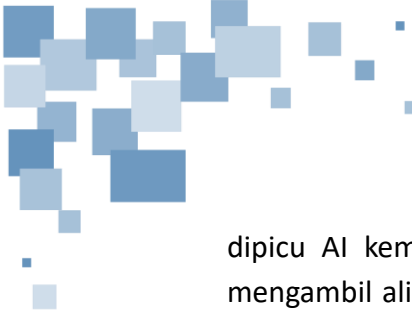
2.2 TRANSFORMASI PEKERJAAN

Gelombang inovasi teknologi dan industri sebelumnya telah mendisrupsi pasar kerja. Gelombang inovasi ini diiringi oleh antusiasme para penemu dan teknolog, tetapi juga oleh kekhawatiran bahwa pekerjaan akan hilang dan masyarakat akan terguncang. AI dan pembelajaran mesin, bersama dengan robotika, merupakan bagian dari revolusi industri baru, yang terkadang disebut sebagai Revolusi Industri Keempat. Keduanya mengubah cara kita berbelanja, cara kita bepergian, cara kita terhubung dan berkomunikasi satu sama lain, cara kita belajar, tempat kita makan, dan apa yang kita baca. AI juga mentransformasi cara kita menciptakan nilai di berbagai industri termasuk pertanian, manufaktur, ritel, transportasi, dan energi. Kemampuan AI untuk mengolah data dalam jumlah besar membuatnya juga cocok untuk mengambil alih banyak tugas yang sebelumnya dianggap berada dalam yurisdiksi eksklusif manusia. Sistem AI yang menulis artikel berita, mendiagnosis kanker kulit, dan mengelola pembukuan, kini membantu, dan terkadang menggantikan, para pekerja upah manusia.

Pengacara dan AI

Pengacara, terlepas dari pelatihan intensif dan tugas-tugas kognitif yang tinggi, tidak akan luput dari dampaknya. Pada tahun 2016, Richard Susskind berpendapat dalam bukunya *Tomorrow's Lawyers* bahwa "memandang pekerjaan hukum sebagai sesuatu yang dibuat khusus adalah sesuatu yang tidak membantu bahkan seringkali fiksi romantis." Program yang digunakan dalam peninjauan dokumen telah mengungguli kinerja pengacara manusia. Saat ini, terdapat ratusan perusahaan rintisan yang mengembangkan solusi tidak hanya untuk peninjauan dokumen, tetapi juga untuk otomatisasi dokumen, penelitian hukum, dan bantuan hukum, banyak di antaranya telah digunakan oleh firma hukum.

Alat-alat yang disempurnakan dengan AI ini menghasilkan pengurangan biaya dan peningkatan kepuasan klien. Namun, alat-alat ini juga menghadirkan tantangan baru. Misalnya, jika semua tugas dasar, rutin, dan berulang diotomatisasi, akan menjadi kurang jelas bagaimana pengacara muda akan mempelajari keahlian mereka. Terlepas dari kemajuan ini dan apa yang mungkin diinginkan sebagian orang, pengacara tidak akan menghilang dalam waktu dekat. Namun, tugas mereka akan berubah dan mereka akan semakin banyak berinteraksi dengan mesin pintar. Namun, di sektor lain, manusia berisiko tergantikan lebih cepat. Pekerja di sektor transportasi, logistik, dukungan administratif, dan produksi khususnya berisiko tinggi; khususnya pekerja berketerampilan rendah dan bergaji rendah. Dampak pergeseran yang



dipicu AI kemungkinan besar akan jauh dari ketakutan kaum Luddite akan mesin yang mengambil alih setiap pekerjaan atau impian kaum optimis teknologi akan kemewahan dan waktu luang. Namun demikian, dalam jangka pendek, gangguan sosial dapat diperkirakan, yang membutuhkan kebijakan publik yang manusiawi dan berwawasan ke depan.

Prediksi mengenai dampak AI di tempat kerja sangat bervariasi. Beberapa memperkirakan bahwa sebanyak 50% dari semua pekerjaan berisiko. Yang lain jauh lebih optimis. Sebuah laporan oleh *McKinsey Global Initiative* (MGI) tentang masa depan pekerjaan menyimpulkan bahwa hampir setengah dari aktivitas yang dibayar untuk dilakukan orang secara global (mewakili sekitar Rp 1.600 triliun) berpotensi diotomatisasi oleh program yang dibuat oleh industri TI dan komputer. Ini bukan berarti setengah dari pekerjaan akan diotomatisasi. Sebaliknya, laporan tersebut memperkirakan hanya 5% pekerjaan yang akan sepenuhnya diotomatisasi. Namun, pada tingkat tugas/aktivitas, laporan tersebut memperkirakan 30% aktivitas, di 60% jenis pekerjaan, akan sepenuhnya diotomatisasi. Aktivitas yang paling rentan terhadap otomatisasi, yang mewakili upah sebesar Rp 270 triliun, adalah aktivitas yang berkaitan dengan pengumpulan data, pemrosesan data, dan aktivitas fisik yang dapat diprediksi. Sektor-sektor di mana pekerja menghabiskan waktu terbanyak untuk aktivitas yang dapat diotomatisasi akan mengalami perubahan paling besar, di antaranya layanan makanan, manufaktur, pertanian, dan transportasi.

Sebuah studi terbaru yang lebih optimistis oleh *Pricewaterhousecoopers* (PWC) memperkirakan bahwa AI akan menciptakan lebih banyak pekerjaan daripada yang telah dihancurkannya di Inggris. Seperti yang ditunjukkan dalam artikel "Peran yang muncul dan menurun berdasarkan AI" berikut, AI diperkirakan akan mendorong permintaan untuk peran baru sekaligus mengurangi permintaan untuk peran lama. Peran-peran baru ini tidak hanya akan membutuhkan keterampilan yang lebih tinggi dan memberikan gaji yang lebih tinggi, tetapi pada akhirnya, akan ada lebih banyak pekerjaan di ekonomi masa depan yang didukung AI. Hal ini mengikuti dampak perubahan teknologi yang sering dicatat, yang terbaru oleh Bank Dunia dalam Laporan Pembangunan Dunia 2019: "Selama seabad terakhir, mesin telah menggantikan pekerja dalam banyak tugas. Namun, secara keseluruhan, teknologi telah menciptakan lebih banyak lapangan kerja daripada yang digantikannya" karena terciptanya sektor-sektor kegiatan baru, yang salah satunya adalah Teknologi Informasi dan Komunikasi (TIK). Singkatnya, permintaan akan pekerja dengan keterampilan kognitif, adaptif, dan relasional yang tinggi akan meningkat. Namun, bahkan penilaian optimistis tentang dampak AI terhadap penciptaan lapangan kerja mencatat bahwa dampak AI di tempat kerja tidak akan dirasakan secara merata: keterampilan tertentu akan semakin dihargai dan keterampilan lainnya akan sangat terdevaluasi. Lebih lanjut, banyak yang berpendapat bahwa dampak AI di tempat kerja akan dirasakan secara berbeda di berbagai negara, dengan potensi yang kuat untuk memperburuk ketimpangan sosial-ekonomi yang ada di antara kelompok "kaya" dan "miskin" di dunia, baik di tingkat domestik maupun internasional.

Dampak AI Terhadap Ketenagakerjaan: Tren Global

Ketika harga minyak turun, jumlah rig minyak dan karyawan pun berkurang. Namun, ketika jumlah rig mulai naik kembali, jumlah karyawan tetap stagnan. Hal ini disebabkan oleh

tekanan harga, di awal siklus, yang mempercepat adopsi teknologi baru yang secara signifikan mengurangi jumlah pekerja yang dibutuhkan untuk mengoperasikan rig minyak. Yang mengejutkan adalah seluruh siklus tersebut hanya memakan waktu kurang dari 3 tahun. Ini adalah salah satu contoh dampak robot (yang kini seringkali didukung oleh AI) terhadap manufaktur. Namun, pola dasar peningkatan produktivitas dan penurunan ketenagakerjaan terdapat di berbagai industri. Misalnya, di Amerika Serikat, produktivitas dan ketenagakerjaan bergerak beriringan dari tahun 1940an hingga sekitar tahun 2000, ketika keduanya mulai terpisah.

Meskipun pemisahan ini (antara produktivitas dan ketenagakerjaan) didorong oleh sejumlah faktor, otomatisasi, AI, dan industri perangkat lunak tidak diragukan lagi telah mengubah cara kerja di berbagai sektor, termasuk pekerja administrasi, agen perjalanan, penyusun pajak, dan bahkan penasihat keuangan. Belum lagi dampak internet dan e-commerce terhadap toko fisik, surat kabar, toko buku, dan perpustakaan, yang beberapa di antaranya telah mulai memanfaatkan AI dalam bisnis sehari-hari mereka.

Peran yang muncul dan menurun berdasarkan AI

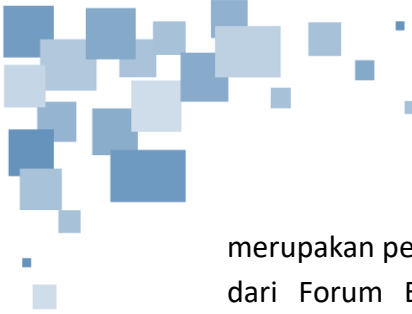


1. Analis Data
2. Spesialis AI dan Pembelajaran Mesin
3. Manajer Umum dan Operasional
4. Pengembang dan Analis Perangkat Lunak dan Aplikasi
5. Profesional Penjualan dan Pemasaran
6. Spesialis Big Data
7. Spesialis Transformasi Digital
8. Spesialis Teknologi Baru
9. Spesialis Pengembangan Organisasi
10. Layanan Teknologi Informasi



1. Petugas Entri Data
2. Petugas Akuntansi, Pembukuan, dan Penggajian
3. Sekretaris Administrasi dan Eksekutif
4. Pekerja Perakitan dan Pabrik
5. Petugas Informasi Klien dan Layanan Pelanggan
6. Manajer Layanan Bisnis dan Administrasi
7. Akuntan dan Auditor
8. Petugas Pencatatan Material dan Penyimpanan Stok
9. Manajer Umum dan Operasional
10. Petugas Layanan Pos

Meskipun perdebatan mengenai dampak bersih industri perangkat lunak dan AI terhadap lapangan kerja masih berlanjut, faktanya tetap bahwa industri perangkat lunak saat ini



merupakan pendorong signifikan penciptaan lapangan kerja di seluruh dunia. Sebuah laporan dari Forum Ekonomi Dunia menyimpulkan bahwa "analisis kami menemukan bahwa peningkatan permintaan untuk peran baru akan mengimbangi penurunan permintaan untuk peran lainnya." Grup Bank Dunia memperkirakan bahwa sektor TIK mewakili sekitar 5,7% dari total lapangan kerja di sektor bisnis di negara-negara OECD. Di India saja, industri TI mewakili lebih dari 3,5 juta lapangan kerja, sepertiganya dipegang oleh perempuan.

Tiga Jenis Dampak Terhadap Ketenagakerjaan

Aplikasi berbasis AI secara umum dapat digolongkan menjadi tiga kategori berikut:

1. aplikasi yang berupaya meningkatkan efisiensi melalui otomatisasi (misalnya, robot dan otomatisasi proses robotik, perangkat lunak pintar: misalnya, penyaringan spam, deteksi penipuan, penyiapan pajak),
2. aplikasi yang berupaya meningkatkan efisiensi dan skalabilitas melalui augmentasi dan kerja sama manusia-mesin, dan
3. aplikasi yang memungkinkan tugas-tugas yang sebelumnya tidak mungkin dilakukan (misalnya, pengobatan personal, transportasi otonom, perdagangan frekuensi tinggi, prakiraan cuaca jangka panjang).

Meskipun jenis pertama, otomatisasi, akan terus memberikan dampak signifikan terhadap ketenagakerjaan, jenis kedua, augmentasi, bisa dibilang lebih menarik dan akan memberikan dampak yang lebih besar dan lebih mengganggu terhadap sifat pekerjaan itu sendiri. Augmentasi membutuhkan suatu bentuk kerja sama antara manusia dan mesin untuk menjalankan suatu tugas. Seiring mesin yang dilengkapi AI belajar dari praktik dan menjadi lebih mumpuni, batasan antara pekerjaan manusia dan pekerjaan mesin akan bergeser, dengan mesin mengambil porsi tugas yang lebih besar, sehingga mendorong manusia untuk berfokus pada tugas-tugas yang bernilai lebih tinggi.

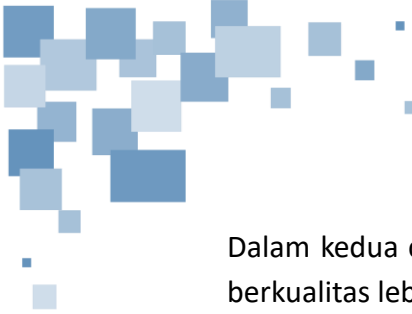
Contoh augmentasi meliputi penyelesaian masalah dalam dukungan pelanggan, peninjauan dokumen dalam eDiscovery (hukum), dan diagnosis medis (misalnya, analisis citra kanker kulit). Dalam setiap contoh ini, manusia bertanggung jawab untuk membuat keputusan akhir tentang tindakan yang akan diambil, tetapi ia dibantu oleh program yang menyediakan skalabilitas, kualitas yang lebih tinggi, dan biaya yang lebih rendah.

Kedokteran

Meskipun seorang dokter bersertifikat dapat memeriksa ribuan atau bahkan puluhan ribu gambar jaringan kanker dan normal selama kariernya, sebuah program dapat dilatih dengan jutaan atau puluhan juta gambar jauh lebih cepat. Dengan pengaturan seperti itu, dokter tetap bertanggung jawab atas diagnosis akhir, tetapi AI dapat membantunya membuat keputusan lebih cepat.

Hukum

Di bidang hukum, meminta pengacara untuk memeriksa puluhan ribu atau bahkan jutaan dokumen tidak hanya sangat mahal, tetapi juga merupakan sesuatu yang relatif kurang dikuasai oleh manusia/pengacara. AI dapat memeriksa koleksi dokumen yang besar dengan cepat dan mengidentifikasi sebagian kecil dokumen yang memerlukan peninjauan manusia.



Dalam kedua contoh ini, augmentasi akan menghasilkan peningkatan produktivitas, layanan berkualitas lebih tinggi, dan memungkinkan orang untuk fokus pada tugas-tugas yang bernilai lebih tinggi (dan lebih menarik). Pekerjaan yang lebih menarik dalam waktu yang lebih singkat seharusnya setidaknya secara teori memungkinkan pekerja mencapai keseimbangan kehidupan kerja yang lebih baik.

Jenis aplikasi berkemampuan AI yang ketiga tidak dirancang untuk memperkenalkan efisiensi atau skalabilitas pada tugas-tugas yang sudah ada. Sebaliknya, aplikasi ini bertujuan untuk memungkinkan jenis kapabilitas baru. Namun, kapabilitas baru seringkali membuat cara lama dalam melakukan sesuatu menjadi usang dan akan berdampak negatif pada pekerjaan tertentu.

Transportasi

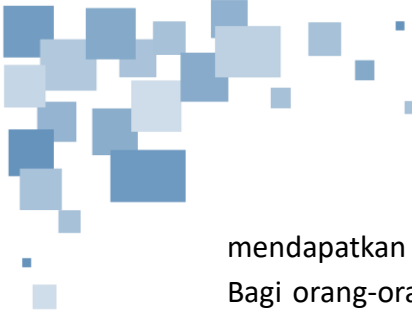
Dari perspektif ketenagakerjaan, perkembangan dalam Transportasi Otonom sangat penting karena berpotensi membuat industri pengemudi profesional menjadi usang. Ini adalah segmen tenaga kerja yang cukup besar yang perlu mengubah kategori pekerjaan sepenuhnya jika kendaraan otonom sepenuhnya menjadi hal yang umum.

Kesenjangan Keterampilan

Perubahan yang sedang berlangsung berisiko menciptakan kesenjangan keterampilan tidak hanya antara pekerja dan sistem AI, tetapi juga antara pekerja dengan berbagai tingkat pengetahuan dan kemampuan. Bank Dunia baru-baru ini merangkum situasi tersebut sebagai berikut: "permintaan akan keterampilan kognitif tingkat lanjut dan keterampilan sosio-perilaku meningkat, sementara permintaan akan keterampilan spesifik pekerjaan yang sempit semakin berkurang. Sementara itu, keterampilan yang terkait dengan "kemampuan beradaptasi" semakin diminati." Hal ini terjadi karena di era digital, perubahan teknologi yang pesat memaksa pekerja untuk terus-menerus memodifikasi dan meningkatkan keahlian mereka.

Seiring dengan perubahan sifat pekerjaan itu sendiri, keterampilan yang dibutuhkan untuk melakukan pekerjaan tersebut juga akan berubah. Keterampilan yang dibutuhkan untuk peran yang sedang berkembang (misalnya, insinyur big data, ilmuwan AI, atau pengembang) sangat berbeda dengan keterampilan yang dibutuhkan untuk peran yang sedang menurun (misalnya, akuntan, juru tulis, pengemudi, dll.). Menurut Laporan Masa Depan Pekerjaan: "Pada tahun 2022, tidak kurang dari 54% dari seluruh karyawan akan membutuhkan pelatihan ulang dan peningkatan keterampilan yang signifikan.

Dari jumlah tersebut, sekitar 35% diperkirakan membutuhkan pelatihan tambahan hingga enam bulan, 9% akan membutuhkan pelatihan ulang keterampilan yang berlangsung selama enam hingga 12 bulan, sementara 10% akan membutuhkan pelatihan keterampilan tambahan lebih dari satu tahun." Lebih lanjut, laporan tersebut menyimpulkan bahwa kemungkinan perusahaan merekrut staf baru dengan keterampilan baru yang sedang berkembang dua kali lebih besar daripada perusahaan yang meningkatkan keterampilan karyawan yang sudah ada. Kombinasi konektivitas, sumber terbuka, komputasi awan, pendidikan/kursus/pelatihan daring, platform kolaborasi, dan persaingan ketat untuk



mendapatkan talenta juga akan mempercepat ekonomi gig dan mengubah sifat pekerjaan. Bagi orang-orang dengan keterampilan yang dibutuhkan, hal ini niscaya akan menghasilkan pekerjaan yang lebih memuaskan, gaji yang lebih tinggi, kemacetan lalu lintas yang lebih sedikit, dan keseimbangan hidup-kerja yang lebih baik. Dan selama orang-orang ini terus menginvestasikan sejumlah besar waktu mereka untuk belajar dan meningkatkan keterampilan seiring perkembangan teknologi, mereka akan mampu mempertahankan karier yang bermakna dan memuaskan. Namun, laju perubahan akan membuat banyak keterampilan menjadi tidak relevan dengan sangat cepat, dan orang-orang dengan keterampilan yang kurang diminati mungkin harus bersaing untuk mendapatkan pekerjaan yang kurang memuaskan dan bergaji rendah.

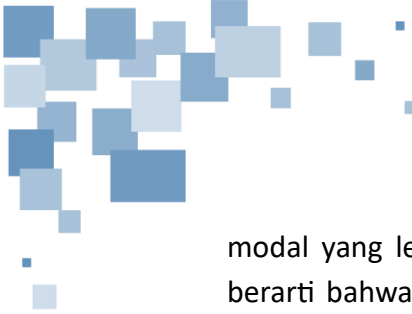
Jika hal ini terjadi, beban peningkatan keterampilan akan dilimpahkan kepada karyawan itu sendiri, yang kemungkinan besar akan kekurangan waktu dan sumber daya yang dibutuhkan untuk mempelajari keterampilan baru dan yang sedang berkembang. Oleh karena itu, tidaklah masuk akal untuk mengharapkan pekerja dari peran yang menurun mampu meningkatkan keterampilan mereka untuk peran baru tanpa dukungan dari luar. Beberapa pasti akan mencapai transisi yang sukses dengan usaha mereka sendiri, tetapi banyak yang mungkin tidak.

Skenario yang berkembang pesat ini menuntut pendefinisian ulang pendidikan. Pendidikan sebagai sarana untuk menghadapi tantangan revolusi industri bukanlah solusi baru. Pada awal abad ke-20, sistem sekolah, misalnya di Amerika Serikat, secara progresif ditingkatkan dan diperluas untuk beradaptasi dengan menurunnya lapangan kerja pertanian dan munculnya lapangan kerja industri. Pada akhir abad ke-20, sebagian besar warga negara di negara-negara maju setidaknya mengenyam pendidikan tingkat menengah. Namun, paradigma pendidikan dominan saat ini menciptakan kesenjangan yang signifikan antara masa belajar (selama masa bayi, remaja, dan awal dewasa) dan masa kerja yang lebih panjang di masa dewasa tanpa pendidikan formal lebih lanjut.

Hal ini harus ditingkatkan untuk mengakomodasi perkembangan keterampilan yang dibutuhkan dalam ekonomi baru ini dan laju perubahan teknologi. Dalam hal ini, Bank Dunia mendorong pemerintah untuk melakukan investasi substansial dalam "pembelajaran seumur hidup", yang dimulai dengan mengajarkan keterampilan kognitif dan sosial kepada anak-anak sejak usia sangat dini, berlanjut di tingkat perguruan tinggi (perguruan tinggi dan sekolah kejuruan, dengan fokus pada keterampilan kognitif yang dapat ditransfer), dan kemudian berlanjut lebih jauh dengan pembelajaran orang dewasa di dalam dan di luar tempat kerja. Lebih dari sebelumnya, pembelajaran dan pelatihan tidak akan terbatas pada usia muda dan masa sekolah. Dukungan yang luas dan beragam dari pemerintah akan sangat penting dalam memenuhi kebutuhan ini.

Proposal Kebijakan Jangka Panjang

Pemanfaatan ekonomi otomatisasi, AI, dan robotika diprediksi akan menciptakan kekayaan yang belum pernah terjadi sebelumnya. Namun, banyak yang khawatir kekayaan ini akan terkonsentrasi di tangan segelintir orang. Filsuf Swedia, Nick Bostrom, berpendapat bahwa munculnya sistem AI yang menggantikan pekerja manusia dapat berarti signifikansi



modal yang lebih besar dalam pangsa pendapatan global, dibandingkan dengan upah. Ini berarti bahwa porsi kue yang lebih besar akan jatuh ke tangan orang-orang yang memiliki modal, karena pekerja upahan akan menjadi semakin kurang penting dalam proses produksi di banyak industri.

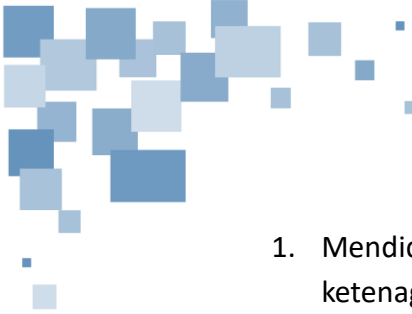
Kebijakan kesejahteraan saat ini, seperti skema asuransi sosial yang diwarisi dari kebijakan yang diberlakukan oleh Bismarck di Jerman pada akhir abad ke-19, mengasumsikan bahwa pekerja menerima upah tetap dari pemberi kerja tertentu dan memiliki usia pensiun yang tetap. Model ini efektif untuk pekerjaan yang stabil dan formal di pabrik atau perusahaan. Namun, bagi negara-negara berkembang di mana model ini bukan norma, dan bahkan bagi negara-negara maju di mana bentuk pekerjaan yang lebih mobile muncul (sebagian berkat aplikasi AI), model ini tidak diadaptasi.

Inilah sebabnya beberapa pihak mengusulkan ukuran distribusi kekayaan yang lebih universal yang tidak didasarkan pada pekerjaan bergaji formal. Solusi ini mencakup layanan gratis atau bersubsidi seperti transportasi, pengasuhan anak, layanan kesehatan, dan pendidikan, tetapi juga langkah-langkah yang lebih revolusioner seperti menyediakan pendapatan dasar untuk semua. Elon Musk, khususnya, mengatakan pada tahun 2016 bahwa "terdapat kemungkinan yang cukup besar kita akan memiliki pendapatan dasar universal, [...], karena otomatisasi." Beberapa pihak, seperti Bill Gates, telah menyarankan pembiayaan solusi ini melalui "pajak robot" yang baru. Masing-masing solusi yang diusulkan memiliki pendukung dan penentang.

Perdebatan tentu saja baru saja dimulai. Upaya untuk menerapkan pendapatan dasar universal sejauh ini terbatas dan tidak berhasil. Mengenai potensi "pajak robot", Komisi Eropa menolak usulan untuk mengenakan pajak kepada pemilik robot guna membiayai pelatihan ulang pekerja manusia. Pajak semacam itu juga dikhawatirkan akan menjadi pajak atas produktivitas. Yang tidak boleh dilupakan adalah bahwa pemerintah sendiri dapat memanfaatkan sistem AI secara optimal untuk meningkatkan administrasi publik dan layanan yang mereka berikan kepada masyarakat. Pemerintah Kanada, misalnya, telah mulai menguji chatbot AI untuk berinteraksi dengan penerima manfaat program sosial federal seperti pensiun, tunjangan keluarga, dan dukungan disabilitas.

Kebijakan pertama yang harus diterapkan oleh negara-negara yang ingin memanfaatkan AI di tempat kerja adalah kebijakan pemantauan yang menilai dampak AI yang terus berkembang di tempat kerja. Memang, sebagian besar pemerintah saat ini kekurangan informasi penting yang diperlukan untuk membuat keputusan kebijakan terkait AI yang terinformasi.

Revolusi AI kemungkinan besar tidak akan mengikuti pola yang sama seperti pergeseran teknologi sebelumnya, di mana periode singkat ketidakstabilan dan gangguan pada pasar kerja telah diikuti oleh periode stabilitas dan pertumbuhan yang jauh lebih panjang. Laju perubahan dalam industri AI dan perangkat lunak menunjukkan bahwa gangguan ini akan terjadi jauh lebih cepat. Sejumlah studi telah mengajukan rekomendasi tentang cara mengurangi dampak disrupsi ini terhadap tenaga kerja, beberapa di antaranya telah disebutkan, termasuk:

- 
1. Mendidik para pembuat kebijakan tentang AI dan dampaknya terhadap ketenagakerjaan, termasuk mempelajari dampaknya terhadap indeks keterampilan global.
 2. Membuat program pelatihan bagi pekerja yang terdisrupsi berfokus tidak hanya pada keterampilan teknologi tetapi juga pada pengembangan keterampilan lunak mereka seperti komunikasi, manajemen produk, dan keterampilan kepemimpinan.
 3. Meningkatkan akses ke perangkat digital dan AI, termasuk program pelatihan untuk masyarakat umum.
 4. Memberdayakan tenaga kerja melalui pelatihan, layanan penempatan kerja, dan pemutakhiran peraturan terutama yang berfokus pada peningkatan jaring pengaman sosial bagi pekerja yang kehilangan pekerjaan.

Meskipun semua rekomendasi di atas masuk akal, kami yakin tidak ada yang lebih penting daripada mengadopsi sistem pendidikan baru yang berfokus pada pembelajaran seumur hidup dan menerapkan kebijakan untuk memastikan bahwa pekerja memiliki waktu dan sumber daya yang dibutuhkan untuk belajar bekerja dengan teknologi yang sedang berkembang dan menjaga keterampilan mereka tetap mutakhir.

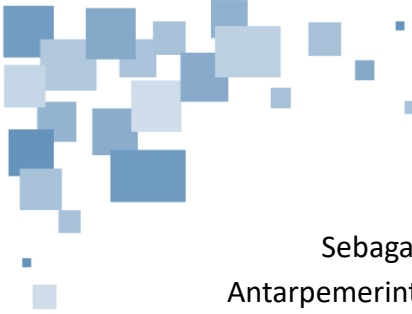
Beberapa yurisdiksi memiliki kebijakan yang mewajibkan partisipasi pemberi kerja dalam pengembangan keterampilan tenaga kerja. Misalnya, provinsi Québec di Kanada mewajibkan perusahaan dengan gaji di atas Rp 20 miliar untuk berkontribusi setidaknya 1% dari gaji tersebut untuk pengembangan keterampilan karyawan. Selain itu, mengingat banyaknya kursus dan program pelatihan daring gratis, hal ini bukan semata-mata masalah pendanaan. Ini adalah masalah menciptakan program terstruktur dan insentif yang tepat bagi pekerja untuk berpartisipasi dan berkomitmen pada kesempatan belajar.

Kesimpulannya, AI mendorong transformasi cara kita menyelesaikan pekerjaan. AI akan memungkinkan kita melakukan hal-hal yang sebelumnya tidak mungkin, meningkatkan efisiensi, mendorong produktivitas, dan pertumbuhan. AI akan mengubah cara kita bekerja, tempat kita bekerja, dan cara kita hidup. Namun, seperti halnya setiap teknologi atau penemuan baru, hukum konsekuensi yang tidak diinginkan selalu ada di mana-mana.

Konsekuensi dan akibat yang tak terlihat dari sebuah teknologi baru seringkali lebih sulit diatasi daripada teknologi itu sendiri. Hal ini berlaku untuk AI. AI membuka batas-batas baru yang menarik, tetapi niscaya akan menciptakan disrupsi yang signifikan. Meskipun para wirausahawan, ilmuwan, dan insinyur akan memimpin pengembangan AI, memikirkan potensi disrupsi merupakan tugas kita semua, termasuk pemerintah, ahli hukum, kelompok konsumen, serikat pekerja, lembaga pendidikan, perusahaan, dan organisasi nirlaba.

2.3 DAMPAK EKOLOGIS AI

Bagian sebelumnya mengasumsikan bahwa perkembangan dan pertumbuhan teknologi akan terus berkembang sebagaimana yang terjadi pada abad terakhir, meskipun dengan laju perubahan yang semakin cepat. Namun, karena pertumbuhan yang tak terkendali telah terbukti berdampak buruk terhadap lingkungan, kita dihadapkan pada kenyataan bahwa pertumbuhan tersebut mungkin tidak berkelanjutan.



Sebagaimana telah diperjelas oleh laporan-laporan seperti laporan Panel Antarpemerintah tentang Perubahan Iklim (IPCC) tahun 2018, kita tidak lagi memiliki kemewahan untuk mengadopsi pandangan dunia yang sepenuhnya berpusat pada manusia. Perubahan iklim merupakan risiko eksistensial yang diciptakan manusia bagi semua bentuk kehidupan. Statistik terbaru mengenai dampak destruktif manusia terhadap makhluk hidup lain dan ekosistem kita sedang mengguncang dunia.

Menurut laporan ilmiah yang disiapkan untuk World Wildlife Foundation, dalam 50 tahun terakhir saja, diperkirakan populasi mamalia, burung, ikan, dan reptil di seluruh dunia telah menurun sekitar 60%. Selain itu, laporan ilmiah Januari 2019 menemukan bahwa laju penurunan populasi serangga kemungkinan bahkan lebih tinggi. Fenomena tragis ini juga dikenal sebagai kepunahan Antroposen atau kepunahan besar keenam di Bumi. Singkatnya, ada kebutuhan mendesak untuk menggunakan energi secara lebih efisien dan mengurangi emisi yang mencemari.

Bagaimana Hal Ini Berkaitan dengan AI?

Di satu sisi, diharapkan perubahan iklim dan konsekuensinya dapat dikurangi melalui teknologi, termasuk AI. Model prediktif yang lebih baik berdasarkan pembelajaran mesin dapat membantu kita mengelola sumber daya dan energi secara lebih efisien dan mengembangkan sistem prakiraan iklim yang lebih akurat. Misalnya, Laporan Villani berpendapat bahwa “AI dapat membantu kita memahami dinamika dan evolusi seluruh ekosistem dengan berfokus pada kompleksitas biologisnya; AI akan memungkinkan kita mengelola sumber daya secara lebih efisien (terutama dalam hal energi), melestarikan lingkungan, dan mendorong keanekaragaman hayati”. Di sisi lain, dengan meningkatkan produksi barang dan jasa, AI secara bersamaan berisiko melanjutkan pola ekonomi yang menjadi sumber krisis. Singkatnya, AI dapat membantu memperbaiki atau memperparah krisis lingkungan yang kita hadapi sebagai masyarakat. Kami percaya AI harus menjadi bagian dari solusi, bukan bagian dari masalah.

Bagian ini terutama akan membahas pengurangan emisi karbon, khususnya melalui pengembangan jaringan listrik 'pintar' dan kemampuannya untuk menggabungkan sumber energi terbarukan dan kendaraan listrik. AI berpotensi mentransformasi jaringan energi, menjadikan penggunaan listrik lebih efisien dan ramah lingkungan; meskipun hal ini disertai risiko hanya mempermudah konsumsi energi, yang pada akhirnya justru menggantikan dampak lingkungan alih-alih membatasinya. Untuk mengeksplorasi skenario ini lebih lanjut, kita sekarang beralih ke studi kasus penerapan AI untuk mengimplementasikan jaringan listrik pintar.

Listrik & Emisi Karbon di Amerika Serikat

Badan Perlindungan Lingkungan Amerika Serikat (EPA) melaporkan bahwa penggunaan listrik bertanggung jawab atas 28% emisi karbon di Amerika Serikat. Hal ini selanjutnya dapat dirinci berdasarkan sumbernya: bahan bakar fosil merupakan bagian terbesar, yaitu 62,9% (sebagian besar gas alam sebesar 32,1% dan batu bara sebesar 29,9%), sementara energi nuklir dan energi terbarukan masing-masing mewakili 20% dan 17% (hidro sebesar 7,4%, angin sebesar 6,3%, biomassa sebesar 1,6%, dan surya sebesar 1,3%) dari sumber energi.

EPA menyarankan bahwa dua cara utama untuk mengurangi emisi karbon akibat listrik adalah dengan meningkatkan efisiensi energi dan mengganti bahan bakar fosil dengan sumber terbarukan. Pada dua poin inilah AI dapat dimanfaatkan dengan sangat bermanfaat. Memahami bagaimana hal ini dapat terwujud membutuhkan pemahaman terlebih dahulu mengenai permasalahan unik yang ditimbulkan oleh pengoperasian jaringan listrik yang efisien.



Ada tiga aspek berbeda dari konsumsi listrik yang harus diakomodasi oleh jaringan listrik. Pertama, permintaan tidak konstan sepanjang hari. Kedua, dan terkait, adalah bahwa permintaan sering melonjak pada waktu-waktu tertentu dalam sehari, menciptakan 'puncak' permintaan. Ketiga, ketika energi berasal dari sumber terbarukan, ketersediaannya tidak konstan.

Aspek pertama ini tidak mengejutkan, mengingat beragamnya konteks permintaan listrik. Kebutuhan pertanian, industri, transportasi, dan perumahan sangat bervariasi, mengingat mereka beroperasi pada jadwal

waktu yang berbeda.

Masalah kedua, yaitu puncak permintaan, merupakan salah satu faktor utama dalam produksi emisi karbon. Pertimbangkan bahwa sebagian besar konsumen rumah tangga, misalnya, berada di rumah pada malam hari, yaitu saat mereka menggunakan peralatan seperti mesin pencuci piring, mesin cuci, dan sistem hiburan. Ketika banyak rumah tangga menggunakan peralatan ini pada waktu yang sama di malam hari, hal itu menciptakan lonjakan konsumsi listrik puncak, relatif terhadap waktu-waktu lain dalam sehari. Ketika puncak ini terjadi, pembangkit listrik tambahan perlu dihidupkan, yang biasanya kurang ramah lingkungan, sehingga meningkatkan emisi karbon.

Terakhir, ada masalah produksi yang tidak konstan. Seperti yang dicatat dalam artikel terbaru dari *The Economist*, "bahan bakar fosil lebih mudah dihubungkan ke jaringan listrik saat ini daripada energi terbarukan yang bergantung pada sinar matahari dan hembusan angin." Karena sumber terbarukan bergantung pada cuaca, sumber terbarukan dapat menjadi sulit untuk diterapkan pada sistem energi yang beroperasi berdasarkan permintaan.

Ketiga aspek jaringan listrik di atas berpadu menciptakan 'masalah koordinasi'. Jaringan listrik harus mengoordinasikan, di satu sisi, permintaan yang tidak konstan yang muncul di puncak dan lembah, di sisi lain, pasokan yang agak tidak konstan. Kecerdasan buatan (AI) adalah cara untuk memecahkan masalah koordinasi tanpa bergantung pada bahan bakar fosil. Dengan menciptakan 'jaringan pintar' yang dapat memprediksi pola permintaan dengan lebih baik, meratakan puncak permintaan, dan memasok energi terbarukan saat tersedia, AI dapat menyediakan jaringan listrik yang lebih efisien dan berkelanjutan.

Jaringan Pintar

Jaringan pintar berfungsi dengan mengganti meteran listrik lama dengan meteran "pintar" baru yang mencatat penggunaan energi rumah tangga dan berinteraksi dengan peralatan "pintar". Informasi konsumsi, setelah dicatat dan disimpan, menjadi data penggunaan energi, yang darinya kebiasaan gaya hidup konsumen, antara lain, dapat disimpulkan. Seperti yang akan kita lihat, data penggunaan energi inilah yang memungkinkan jaringan pintar AI, tetapi juga menciptakan masalah hukum unik yang harus ditangani oleh kebijakan.

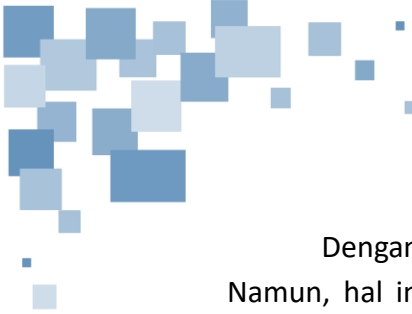
Bagaimana jaringan pintar berfungsi? Tujuannya adalah untuk mempelajari pola konsumsi dan mengelola kedua sisi penggunaan energi, permintaan dan pasokan. Bayangkan sebuah lingkungan suburban yang menerima energinya dari satu pembangkit listrik. Di malam hari, setelah bekerja, sebagian besar rumah tangga kemungkinan besar akan menyalakan mesin cuci mereka, yang menciptakan lonjakan penggunaan energi. Bayangkan lonjakan ini menciptakan permintaan yang begitu besar sehingga pembangkit listrik kedua yang kurang ramah lingkungan harus dimanfaatkan untuk memasok listrik yang dibutuhkan. Akibatnya, emisi karbon yang dihasilkan dari lonjakan ini melebihi jumlah yang akan dihasilkan jika beban cucian tersebar sepanjang hari.

Jelas, skenario ini memiliki ciri-ciri klasik masalah aksi kolektif. Karena setiap rumah tangga tidak mengetahui apa yang dilakukan rumah tangga lainnya, bahkan jika mereka ingin menghindari puncak permintaan energi, mereka mungkin tidak sepenuhnya yakin dapat melakukannya.

Jaringan pintar bekerja dengan membantu memecahkan bukan hanya masalah ketidaktahuan rumah tangga ini, tetapi juga menyelesaikannya dengan terlibat dalam proses pengambilan keputusan. Karena meteran pintar mencatat penggunaan listrik rumah tangga, perusahaan energi akan membangun kumpulan data penggunaan terpusat. Dengan demikian, perusahaan dapat bertindak untuk mengurangi puncak tersebut. Hal ini dilakukan dengan berinteraksi dengan rumah tangga melalui peralatan pintar mereka. Jika sebuah rumah tangga memiliki mesin cuci pintar, ia dapat berkomunikasi dengan meteran pintar. Meteran kemudian dapat mengurangi intensitas listrik sejumlah mesin cuci (sehingga membuatnya beroperasi lebih lama karena intensitas yang lebih rendah), untuk meratakan puncak permintaan.

Skenario ini menampilkan komunikasi dua arah dari jaringan pintar. Meteran menerima informasi (data penggunaan) dari rumah tangga, dan kemudian menggunakan informasi ini untuk berkomunikasi (dan sebagian mengendalikan) peralatan. Dengan mengumpulkan data penggunaan ini dari waktu ke waktu dari rumah tangga, bisnis, pertanian, dan sebagainya perusahaan energi kemudian dapat menggunakan kumpulan data ini agar sistem AI mempelajari pola penggunaan energi.

Hal ini akan membuat penggunaan energi lebih mudah diprediksi, sehingga prediksi dan perataan puncak menjadi lebih memungkinkan. Meskipun data mentah penggunaan energi sebelumnya tersedia melalui meteran lama, pengumpulannya jauh lebih sulit dan kurang spesifik tentang peralatan apa yang digunakan. Oleh karena itu, memperkenalkan meteran dan peralatan pintar ke rumah merupakan bagian penting dari solusi ini.



Dengan mengurangi puncak permintaan, penggunaan energi menjadi lebih efisien. Namun, hal ini masih menyisakan masalah dalam mengintegrasikan energi terbarukan ke dalam jaringan listrik. Hal ini bergantung pada mekanisme yang sama seperti perataan puncak. Ketika meteran pintar berkomunikasi dengan, misalnya, mesin cuci dan memperlambat siklusnya untuk mengurangi intensitas listrik, secara efektif ia mengubah rentang waktu penggunaan energi. Kapasitas inilah yang diperlukan untuk mengintegrasikan energi terbarukan.

Karena energi terbarukan harus digunakan dalam rentang waktu yang singkat setelah dipanen (tanpa peningkatan kapasitas penyimpanan), meter pintar yang dilatih dengan AI seolah-olah dapat memberi tahu peralatan ketika rentang waktu hemat energi terjadi. Ini tidak berarti meter dapat menyalakan mesin cuci itu sendiri, tetapi dapat menampilkan rentang waktu tersebut pada antarmuka pengguna mesin cuci, memberikan pengguna keputusan akhir, tetapi dengan informasi yang lebih baik, tentang kapan harus mencuci pakaian mereka.

Jelas, jaringan pintar bergantung pada perangkat keras meter dan peralatan pintar, serta data yang mengalir melaluinya. Data yang dihasilkan, dengan demikian, akan memiliki nilai instrumental yang sangat besar untuk mengurangi emisi karbon. Namun, data tersebut juga akan memiliki nilai moneter yang sangat besar bagi perusahaan energi, dan, khususnya, organisasi pihak ketiga. Hal ini memicu kekhawatiran tentang privasi dan regulasi data rumah tangga.

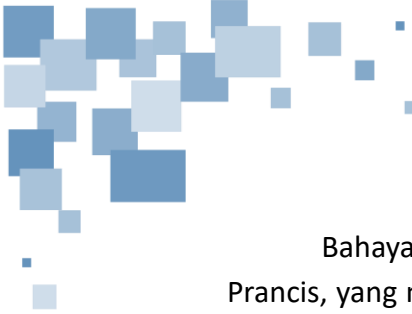
Dalam konteks ini, RUU SB 1476 California adalah contoh undang-undang yang berupaya menemukan keseimbangan antara tujuan memfasilitasi aliran data yang diperlukan untuk implementasi teknologi jaringan pintar yang efektif dengan tujuan yang berlawanan, yaitu melindungi hak privasi.

Seperti yang dicatat Forbush, "RUU Senat California 1476 mengharuskan agregator data konsumsi energi untuk mendapatkan persetujuan konsumen sebelum membagikan informasi pelanggan dengan pihak ketiga; mengamanatkan bahwa pihak ketiga hanya boleh memiliki akses ke data tersebut saat mereka membuat kontrak dengan utilitas untuk menyediakan layanan terkait manajemen energi; menetapkan bahwa data harus dijaga keamanannya dari pihak yang tidak berwenang; dan mengamanatkan bahwa pembayar tarif listrik memilih untuk mengotorisasi setiap pembagian data konsumsi energi mereka untuk tujuan komersial sekunder apa pun."

Model seperti itu penting, terutama mengingat bahwa jaringan pintar akan diperluas melampaui penggunaan rumah tangga, yang pada akhirnya mencakup pergerakan konsumen dan konsumsi dari kendaraan listrik. Di sisi lain, model pengumpulan data berdasarkan persetujuan mungkin memberikan model prediktif yang kurang akurat.

Potensi Bahaya Peningkatan Konsumsi

Memecahkan masalah koordinasi antara AI dan sumber energi terbarukan merupakan tugas yang penting dan sulit, tetapi meskipun berhasil, masih terdapat sejumlah masalah yang belum terselesaikan. Masalah pertama dan paling mendesak adalah kemungkinan bahwa meningkatkan efisiensi konsumsi energi akan mempermudah konsumsi energi dalam jumlah besar, yang secara efektif akan menghapus kemajuan dalam pengurangan karbon.



Bahaya ini dicatat oleh Cedric Villani dalam laporannya tentang AI untuk pemerintah Prancis, yang menyatakan bahwa AI dapat "mengakibatkan konsumsi setidaknya sama besar, jika tidak lebih besar, daripada sebelumnya." Ia menyebut hal ini sebagai 'efek rebound' di mana "penghematan energi dan sumber daya yang diharapkan akibat penggunaan teknologi baru mungkin sebagian atau seluruhnya diimbangi oleh respons masyarakat terhadapnya." Hal ini tidak sulit dibayangkan. Villani memberikan contoh sederhana di mana penghematan yang diperoleh dari tagihan pemanas dapat diinvestasikan kembali ke produk atau aktivitas lain yang kemudian menambah konsumsi energi.

Oleh karena itu, membiarkan pasar teknologi mengubah jaringan listrik saja tidak cukup untuk mengatasi ancaman perubahan iklim. "Jaringan pintar" hanya benar-benar cerdas jika merupakan bagian dari upaya publik yang lebih luas untuk mengurangi emisi karbon. Satu-satunya cara untuk memastikan hal ini tercapai adalah dengan menempatkan pendekatan berbasis teknologi dan pasar yang dibahas di sini dalam arsitektur kebijakan publik yang lebih luas.

Terakhir, aspek lain dari dampak ekologis AI harus dihitung untuk mendapatkan penilaian yang adil terhadap risiko dan manfaatnya. Meskipun AI dapat berfungsi di dunia virtual, ia tetap membutuhkan sumber daya dunia nyata yang substansial. Dalam laporannya, Cédric Villani mencatat bahwa "Produksi perangkat keras digital menggunakan logam mulia langka dalam jumlah besar yang hanya sebagian dapat didaur ulang, dan cadangan yang tersedia terbatas (15 tahun untuk Indium, misalnya, yang konsumsinya telah berlipat ganda tujuh kali lipat dalam 10 tahun); hal ini dapat mengakibatkan kebuntuan teknologi jika peningkatan permintaan ini tidak melambat, terutama mengingat beberapa logam ini juga digunakan dalam produksi peralatan untuk energi terbarukan (tenaga angin dan surya)." Beralih dari ketergantungan pada sumber energi berbasis karbon yang terbatas menjadi ketergantungan pada deposit mineral terbatas yang membutuhkan penambangan intensif mungkin merupakan hal yang menanti kita di masa depan.

Singkatnya, efisiensi dan keberlanjutan dalam penggunaan semua sumber daya yang diperlukan untuk fabrikasi dan penggunaan sistem AI harus menjadi perhatian utama setiap kali organisasi mempertimbangkan untuk mengembangkan, menerapkan, atau menggunakan sistem AI. Penilaian proyek AI harus mencakup evaluasi tidak hanya dampak karbon proyek, tetapi juga konsumsi bahan bakunya dengan tujuan penggunaan kembali dan daur ulang.

2.4 SISTEM SENJATA OTONOM MEMATIKAN

Mengkaji ulang analogi awal kita tentang penggunaan energi nuklir sebagai senjata, kita sekarang akan mengeksplorasi risiko terkait penggunaan AI sebagai senjata, seiring dengan semakin banyaknya negara yang berlomba-lomba mengembangkan sistem senjata otonom mematikan (LAWS) dan menerapkan sistem AI untuk menciptakan angkatan bersenjata masa depan.

Sejarawan seperti Paul Kennedy telah mencatat hubungan erat antara kekuatan ekonomi, kekuatan militer, dan inovasi teknologi. Jika geopolitik abad ke-19 dan ke-20 didefinisikan oleh Revolusi Industri, maka Revolusi Digital, dengan AI di garis depan, dapat

menjadi esensi geopolitik zaman kita. Tidak mengherankan, kemajuan terkini dalam AI, yang memungkinkan pemrosesan dan analisis lebih banyak data, untuk meningkatkan otonomi dan aplikasi baru, telah menarik perhatian para pengambil keputusan militer dan kepala negara di seluruh dunia.

Senjata AI dan otonom di seluruh dunia
Tiongkok berharap menjadi pemimpin dunia dalam bidang AI pada tahun 2030 dan anggota tingkat tinggi Tentara Pembebasan Rakyat bermaksud memanfaatkan AI untuk tujuan militer.
Bagi AS, AI merupakan bagian integral dari Strategi Offset Ketiga yang diluncurkan pada tahun 2014 untuk mempertahankan "keunggulan teknologi militer" Amerika.
Sedangkan Rusia, Presiden Vladimir Putin menyatakan pada tahun 2017 bahwa "siapa yang menjadi pemimpin di [AI] akan menjadi penguasa dunia" sementara negaranya juga membiayai penelitian tentang AI yang dimiliterisasi.
Dua negara terakhir yang memiliki kursi tetap di Dewan Keamanan PBB, Prancis dan Inggris, juga telah menyusun rencana untuk meningkatkan angkatan bersenjata mereka dengan kapasitas AI.

Minat Militer Sebelumnya Terhadap AI


Ini bukan pertama kalinya AI menarik perhatian angkatan darat. Pada tahun 1983, Badan Proyek Penelitian Pertahanan Lanjutan (DARPA), cabang penelitian Departemen Pertahanan AS, meluncurkan program penting, Inisiatif Komputasi Strategis, yang menghabiskan Rp 10 triliun dalam sepuluh tahun, terutama untuk AI generasi pertama. Program ini memiliki tujuan luas untuk mengembangkan kecerdasan mesin, dengan tujuan yang lebih konkret untuk menciptakan kendaraan darat otonom, sistem yang dapat membantu pilot, dan sistem manajemen pertempuran.

Agenda penelitiannya sangat luas, termasuk pemrosesan bahasa alami, pemahaman teks, sistem pakar, pengenalan gambar, dan penalaran mesin. Namun, seperti yang telah kita lihat di atas dalam konteks lain, ambisi yang tidak terkendali pada akhirnya berujung pada kekecewaan, dan meskipun beberapa keberhasilan, Inisiatif Komputasi Strategis pada dasarnya ditinggalkan pada akhir tahun 1980an.

Mungkinkah hype saat ini juga mereda? Sebagaimana telah kita lihat dalam pendahuluan, pada tahun 1980an, AI masih kekurangan daya komputasi yang diperlukan untuk mencapai potensi penuhnya dan pembelajaran mesin baru muncul, karena makalah back-propagation Geoffrey Hinton yang terkenal baru akan diterbitkan pada tahun 1986. Namun, sebagaimana ditulis oleh Departemen Pertahanan AS pada tahun 2011, "kemajuan dramatis dalam mendukung teknologi menunjukkan bahwa tingkat otonomi yang belum pernah terjadi sebelumnya dapat diperkenalkan ke dalam sistem tak berawak saat ini dan di masa depan." AI dapat terbukti bermanfaat bagi angkatan bersenjata dalam berbagai cara, tetapi pengembangan sistem otonomlah yang menjadi tujuan utama penelitian militer generasi ini.

Menuju Sistem Senjata Otonom yang Mematikan

Selama tahun 2000an, sistem tanpa awak menjamur di medan perang di seluruh dunia, dengan drone yang digunakan oleh Angkatan Udara AS dalam perang melawan teror menjadi contoh yang paling terkenal. Meskipun sistem tanpa awak sejauh ini telah dilengkapi dengan



beberapa fungsi otomatis, sistem ini tidak memiliki otonomi dan semua fungsi terpentingnya, terutama memilih dan menyerang target, dikendalikan oleh operator manusia.

Dalam konteks tersebut, otonomi akan membawa konsep sistem tanpa awak ke kesimpulan logisnya (dan berpotensi mengerikan). Ketika diberi tujuan, sistem senjata otonom akan mampu memutuskan tindakan terbaik yang akan diambil tergantung pada lingkungannya saat ini. Pembelajaran mesin juga akan memungkinkannya untuk mengembangkan strategi baru dan bereaksi terhadap situasi yang tidak terduga, semua ini tanpa keterlibatan langsung oleh operator manusia.


Sebagai ilustrasi, dalam serangan drone saat ini, operator manusia melakukan sebagian besar tugas penerbangan, memilih target sesuai perintah mereka dan menarik pelatuk. Pada akhirnya, sistem otonom tidak hanya akan berfungsi tanpa manusia di kokpit, tetapi juga tanpa manusia di konsol operasi. Setelah diberi misi oleh seorang komandan, mereka dapat melaksanakan tugasnya sepenuhnya tanpa masukan manusia lebih lanjut, mungkin bahkan dalam konteks komunikasi yang terputus dengan komandan mereka, sehingga tidak menyisakan kapasitas untuk mengarahkan ulang atau mengakhiri misi. Pengembangan sistem semacam itu telah menjadi "prioritas tinggi" bagi militer AS sejak awal tahun 2010an, dan angkatan darat lainnya pun mengikutinya.

Di antara berbagai keuntungan otonomi, angkatan darat berharap sistem ini dapat menjadi solusi atas meningkatnya biaya tenaga kerja. Pada tingkat strategis dan taktis, sistem otonom, yang tidak terbebani oleh keterbatasan otak manusia, dapat meningkatkan kecepatan dan efisiensi operasi dan beroperasi tanpa kendala fisik dan psikologis, seperti kelelahan, stres, ketakutan, kemarahan, dan kurangnya perhatian. Beberapa negara juga berargumen bahwa sistem senjata otonom akan lebih akurat daripada manusia dan dengan demikian mengurangi risiko korban sipil dan kerusakan tambahan.

Penentangan dari Masyarakat Sipil

Namun, di luar kalangan militer, banyak spesialis dan pengamat yang sangat menentang produksi sistem senjata dengan kemampuan untuk menarik pelatuknya sendiri, termasuk Yoshua Bengio, salah satu penerima baru-baru ini dari A.M. Penghargaan Turing atas kontribusinya sebagai pelopor AI. Noel Sharkey, ahli robotika yang terlibat dalam eksperimen Gaak yang telah kita temui di Pendahuluan, adalah salah satu orang pertama yang secara terbuka menentang proyek pengembangan senjata semacam itu. Pada tahun 2007, ia menulis, "Saya telah bekerja di bidang kecerdasan buatan selama beberapa dekade, dan gagasan tentang robot yang membuat keputusan tentang pemusnahan manusia sungguh mengerikan." Bersama para ahli robotika dan spesialis AI lainnya, Profesor Sharkey juga terlibat dalam pembentukan Komite Internasional untuk Pengendalian Senjata Robot pada tahun 2009 dan peluncuran Kampanye untuk Menghentikan Robot Pembunuh pada tahun 2013, sebuah kampanye internasional untuk pelarangan sistem senjata otonom mematikan (LAWS) yang menyatukan puluhan LSM.

Secara independen, *Future of Life Institute* pada tahun 2015 menerbitkan surat terbuka yang dipublikasikan secara luas yang menyerukan pencegahan perlombaan senjata militer dengan memberlakukan larangan senjata otonom ofensif yang ditandatangani oleh ribuan



pakar AI dan robotika. Baru-baru ini, hampir 3.500 spesialis dan organisasi AI telah berjanji untuk tidak pernah berpartisipasi dalam pengembangan atau mendukung LAWS.

Para pendukung masyarakat sipil yang melarang LAWS khawatir pengembangannya oleh beberapa negara dapat memicu perlombaan senjata baru dan bahwa jika sudah lazim, LAWS akan menurunkan "ambang batas untuk berperang." Mereka juga berpendapat bahwa mendelegasikan keputusan untuk mengambil nyawa manusia kepada mesin adalah tindakan yang tidak bermoral atau tidak etis. Dengan demikian, mereka berpendapat bahwa manusia harus mempertahankan kendali yang berarti atas senjata yang mereka gunakan dalam pertempuran dan bahwa LAWS justru merampas kendali tersebut dari tentara.

Oposisi dari Dalam Industri Teknologi

Karena pengetahuan teknologi mereka yang tak tertandingi dan peran mereka sebagai pusat pengembangan AI, raksasa teknologi di seluruh dunia merupakan pemain kunci dalam revolusi militer yang sedang berlangsung. Namun, beberapa perusahaan ini menghadapi penolakan yang belum pernah terjadi sebelumnya dari sumber yang tak terduga: karyawan mereka sendiri.

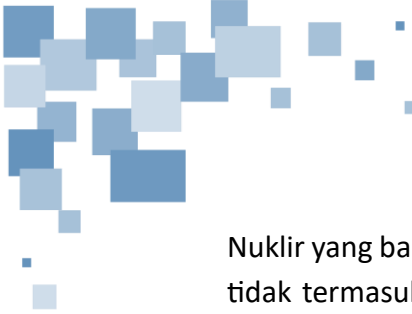
Pada musim semi 2018, karyawan Google mulai mengedarkan dan menandatangani surat terbuka yang ditujukan kepada CEO Sundar Pichai yang memprotes keterlibatan perusahaan dalam Proyek Maven milik Departemen Pertahanan AS (DoD). Proyek ini bertujuan untuk mengembangkan perangkat AI guna mengklasifikasikan gambar yang ditangkap oleh drone. Sekitar 4.000 karyawan menandatangani surat tersebut dengan alasan bahwa Google tidak boleh "terlibat dalam bisnis perang" dan menyerukan pembatalan kontrak.

Protes tersebut cukup serius hingga mendorong Google untuk menerbitkan serangkaian prinsip AI, termasuk komitmen bahwa Google tidak akan merancang atau menyebarkan senjata AI. Mengikuti prinsip-prinsip ini, perusahaan memutuskan untuk tidak mengajukan penawaran kontrak komputasi awan senilai Rp 100 triliun dari Pentagon dan juga tidak akan memperbarui kontrak Proyek Maven ketika berakhir pada Maret 2019.

Debat Multilateral

Debat mengenai larangan internasional LAWS kini berlangsung di Perserikatan Bangsa-Bangsa. Pada tahun 2014, diskusi dimulai dalam mekanisme multilateral Konvensi Senjata Konvensional Tertentu (CCW) dan Negara-negara sepakat pada tahun 2016 untuk membentuk Kelompok Pakar Pemerintah tentang Senjata Otonom Mematikan. Meskipun dua puluh delapan negara sejauh ini telah menyuarakan dukungan mereka terhadap larangan senjata yang sepenuhnya otonom, serta Parlemen Eropa, peluang untuk mencapai kendali yang berarti atas LAWS melalui konvensi PBB terbatas. Negara-negara yang terlibat belum menyetujui definisi LAWS dan upaya untuk mencapai perjanjian yang mengikat telah diperlambat oleh AS, Rusia, Korea Selatan, Israel, dan Australia.

Bahkan jika beberapa Negara menyetujui teks final, itu tidak akan mengikat Negara-negara yang menolak untuk meratifikasinya, yang mungkin terjadi jika beberapa percaya kepentingan nasional mereka mendikte mereka untuk mengejar pengembangan LAWS. Dengan demikian, sebuah Konvensi dapat bernasib sama dengan Perjanjian Larangan Senjata



Nuklir yang baru-baru ini diadopsi pada tahun 2017 dan ditandatangani oleh 69 Negara, tetapi tidak termasuk negara-negara berkekuatan nuklir. Meskipun beberapa konvensi yang relatif baru yang melarang laser yang menyilaukan (1995), ranjau darat (1997), dan munisi tandan (2008) secara umum berhasil, dokumen-dokumen yang mengikat secara hukum terhadap senjata-senjata tertentu secara historis memiliki hasil yang beragam.

Peran Hukum Humaniter Internasional


Meskipun demikian, bahkan tanpa larangan atau bentuk regulasi konvensional lainnya, LAWS tidak akan batal demi hukum. Saat ini terdapat konsensus besar bahwa penggunaan LAWS, apa pun definisinya di masa mendatang, akan tunduk, seperti halnya semua senjata, pada hukum humaniter internasional (HHI). Sebagai versi modern dari hukum perang, HHI memuat seperangkat aturan yang berlaku dalam konteks konflik bersenjata. Namun, penerapan HHI pada LAWS masih sarat dengan ketidakpastian. Meskipun tampaknya terdapat konsensus bahwa HHI mewajibkan agensi manusia dipertahankan atas LAWS, sifat agensi tersebut masih diperdebatkan. Beberapa pihak berpendapat bahwa "kendali" manusia atas keputusan penggunaan kekuatan mematikan sangat penting dan akibatnya senjata yang sepenuhnya otonom yang menghilangkan kendali tersebut adalah ilegal menurut HHI.

Pihak lain berpendapat bahwa HHI tidak mengamanatkan kendali manusia yang sebenarnya atas keputusan penembakan dan bahwa penerapan "penilaian manusia" tingkat tinggi sudah cukup. Posisi semacam itu mensyaratkan, misalnya, bahwa seorang komandan militer yang mengerahkan batalion LAWS selama pertempuran, dengan sendirinya akan menjadi pengawasan manusia yang memadai, tanpa ia meninjau setiap keputusan khusus yang diambil oleh robot selama pertempuran. Dalam situasi ini, komandan akan menjadi pihak yang bertanggung jawab dan tetap diwajibkan untuk menghormati prinsip-prinsip HHI tentang perbedaan antara kombatan dan warga sipil, proporsionalitas antara kebutuhan militer dan korban sipil kolateral, dan menghindari penderitaan yang tidak perlu.

Meskipun perdebatan saat ini berfokus pada LAWS, otonomi dalam sistem persenjataan tetap menjadi salah satu cara AI dapat membentuk kembali medan perang di masa depan. AI yang dimiliterisasi juga dapat berarti sistem manajemen rantai pasokan atau algoritma yang memprediksi kapan jet tempur memerlukan perawatan. Perang siber juga pasti akan berubah seiring perkembangan teknologi ini, dan beberapa pihak berpendapat bahwa impian lama Perang Dingin untuk memiliki sistem AI yang mengendalikan persenjataan nuklir guna memungkinkan reaksi yang lebih cepat daripada cahaya terhadap ancaman asing dapat segera menjadi kenyataan. Oleh karena itu, bahkan dalam kasus pelarangan LAWS yang berhasil mencegah terciptanya senjata seperti Terminator, pelaksanaan perang kemungkinan besar akan menjadi aktivitas manusia lain yang diubah oleh AI.

2.5 BERITA PALSU DAN DISINFORMASI YANG MEMICU KEKERASAN

AI dapat dijadikan senjata sebagai perisai atau pedang, tidak hanya di medan perang fisik, tetapi juga di dunia maya informasi dan gagasan. Dalam dunia internet dan media sosial yang berbasis data, AI ditakdirkan untuk memiliki pengaruh penting terhadap cara kita mengonsumsi dan memproduksi konten daring sesuatu yang tidak dapat diabaikan oleh para



profesional hukum. Bagian ini mengeksplorasi tren terkini berita palsu dan disinformasi yang memanfaatkan teknologi AI, dan mempertanyakan apakah solusi terbaik adalah memberlakukan kewajiban yang lebih ketat kepada Perantara Internet untuk mengambil tindakan terhadap konten yang mengandung kebencian atau berbahaya.

Ekosistem Media Sosial

Perlombaan AI saat ini terjadi ketika internet merupakan repositori data manusia terpenting dalam sejarah. Internet juga merupakan infrastruktur berbagi informasi terluas dan terluas hingga saat ini, dengan platform seperti Google, Facebook, Alibaba, Instagram, Twitter, YouTube, dan Reddit menarik miliaran pengguna dan data mereka.

Bersamaan dengan itu, platform-platform ini telah menjadi ruang perdebatan publik yang sentral dan tak terelakkan. Pada tahun 2018, misalnya, 68% orang dewasa Amerika menggunakan media sosial "setidaknya sesekali" untuk mengakses berita. Perantara Internet, juga dikenal sebagai Penyedia Layanan Internet (ISP), adalah bagian yang tertanam dari media sosial internet, suatu ansambel yang terdiri dari blog, media sosial, wiki, dan kolom komentar yang memungkinkan "percakapan" daring yang berkelanjutan, interaktif, dan global untuk menghasilkan, setiap hari, data dalam jumlah yang sangat besar.


Dalam beberapa tahun terakhir, dampak platform ini terhadap wacana publik dan kehidupan demokrasi telah menjadi sorotan tajam. Untuk setiap Wikipedia, Internet tampaknya menyimpan lusinan ruang tempat konten yang mengerikan dapat muncul dan menyebar, terkadang dengan tujuan memengaruhi atau menumbangkan proses politik. Sejak pemilihan Presiden AS 2016, skandal Cambridge Analytica, dan tuduhan campur tangan asing, fenomena berita palsu khususnya telah menarik perhatian.

Saat ini, memproduksi dan menyebarkan konten menjadi lebih mudah. Dalam ekonomi atensi, di mana platform media sosial bersaing untuk mendapatkan perhatian kita demi menghasilkan pendapatan iklan, konten yang paling mungkin menarik minat kita, entah karena alasan baik atau buruk, lebih diutamakan daripada informasi berkualitas dan substantif. Selain itu, personalisasi pengalaman pengguna melalui pengumpulan data menciptakan silo konsumsi informasi ("gelembung") dan "ruang gema". Orang juga cenderung memilih konten berdasarkan sikap dan preferensi pribadi mereka. Hal ini menciptakan lingkungan yang mendukung misinformasi dan disinformasi.

Munculnya Deep Fake

Sebagaimana setiap aspek kehidupan manusia yang sangat bergantung pada data, AI berbasis mesin dan pembelajaran mendalam berpotensi memengaruhi internet, penggunaannya, dan masyarakat luas. Kita dapat menemukan contoh terkini dan jitu dalam kisah Rana Ayyub tentang bagaimana AI dapat berubah dari mengganggu menjadi destruktif dalam konteks media sosial. Seorang jurnalis investigasi di India, Ny. Ayyub, menjadi korban deep fake pada April 2018. Deep fake adalah video atau gambar yang dimodifikasi dengan AI.

Teknologi ini memungkinkan pengguna untuk menukar, misalnya, wajah seseorang dalam gambar atau video dengan wajah orang lain. Modifikasi gambar untuk tujuan yang tidak berbahaya, artistik, atau jahat, bukanlah hal baru, tetapi ketersediaan literatur teknis, akses murah ke daya komputasi dan data, serta pengembangan aplikasi yang mudah digunakan,



telah membuat praktik ini semakin umum. Deep fake khususnya dapat digunakan dalam produksi film atau dalam pembuatan avatar pribadi bagi orang-orang yang telah kehilangan otonomi ekspresif. Namun, dalam kasus Ny. Ayyub, teknologi deep fake digunakan untuk mengganti wajah seorang aktris muda dalam video porno dengan wajahnya sendiri. Video tersebut pertama kali beredar di WhatsApp, platform perpesanan populer yang menggunakan enkripsi ujung ke ujung, tetapi dengan cepat menjadi viral di Twitter, Facebook, dan Instagram. Sebagai reaksi, Ibu Ayyub menghadapi serangkaian pelecehan daring, serta doxing (praktik mempublikasikan informasi pribadi seseorang secara daring, terutama alamat fisiknya).

Sejak kemunculan deep fake pada Desember 2017 di Reddit, banyak aktris Hollywood dan tokoh publik perempuan lainnya juga telah menjadi korban deep fake pornografi. Namun, Ibu Ayyub bisa dibilang orang pertama yang menjadi sasaran karena alasan politik, untuk mengintimidasi dan mendiskreditkan pekerjaannya sebagai jurnalis. Kasus ini menegaskan potensi deep fake untuk sangat merugikan individu, dan khususnya perempuan.

Deep Fake dan Berita Palsu

Banyak yang khawatir bahwa deep fake akan menimbulkan dampak sosial dan politik yang merugikan secara luas. Mereka menempatkan teknik ini dalam fenomena berita palsu yang lebih besar, terkadang sebagai bentuknya yang paling berbahaya hingga saat ini. Yang khususnya memprihatinkan adalah fakta bahwa “mengaburkan batas antara fakta dan fiksi, teknologi deep fake dapat merusak kepercayaan publik terhadap rekaman gambar dan video sebagai penggambaran objektif realitas,” seperti yang ditulis oleh tiga anggota kongres AS baru-baru ini. Dalam konteks hukum, deep fake pada akhirnya dapat mengurangi atau membatalkan nilai pembuktian bukti gambar dan video di pengadilan. Selain itu, video palsu yang meyakinkan dapat memperburuk ketegangan antar komunitas, mendistorsi kehidupan demokrasi, dan mengancam keselamatan publik.

Dengan demikian, deep fake dapat memperburuk terjadinya kebakaran hutan informasi palsu daring yang mengarah pada tindakan kekerasan, seperti yang telah kita lihat dengan pembunuhan massal di India karena rumor penculikan anak yang beredar di WhatsApp dan serangan restoran pizza Washington tahun 2016 oleh seorang pria yang percaya bahwa jaringan restoran cepat saji adalah pusat dari jaringan pelecehan anak yang dipimpin oleh Hillary Clinton karena cerita palsu yang dia baca daring. juga dapat digunakan dalam konteks politik, untuk mencoreng reputasi seorang kandidat dengan membuat rekaman yang memperlihatkan kandidat tersebut mengucapkan komentar yang keterlaluan, misalnya.

Dalam skala yang lebih besar, deep fake dapat menjadi salah satu alat dalam gudang teknik penipuan yang digunakan dalam kampanye propaganda. Laporan terbaru tentang penganiayaan terhadap orang-orang Rohingya di Myanmar menunjukkan bagaimana militer negara membangun operasi untuk menyebarkan kebencian dan disinformasi dengan menggunakan akun-akun palsu di platform media sosial untuk menebar ketakutan akan serangan jihadis. Penyelidik PBB menggambarkan situasi tersebut sebagai "contoh nyata pembersihan etnis." Meskipun deep fake dilaporkan tidak terlibat di Myanmar, kita dapat dengan mudah membayangkan bagaimana teknologi ini dapat digunakan untuk membuat

rekaman yang meyakinkan yang bertujuan untuk mendorong perpecahan dan permusuhan di dalam dan antarpopulasi.

Tanggapan Hukum

Individu yang terdampak deep fake bukannya tanpa jalan keluar, meskipun kita masih kekurangan yurisprudensi mengenai masalah ini. Berdasarkan kekhususan rezim hukum nasional, korban dapat mengajukan tuntutan ganti rugi atas dasar pelanggaran hak cipta, pencemaran nama baik, pelanggaran privasi, dan perampasan kepribadian, di antara dasar-dasar lainnya. Penggugat di Inggris dan AS, misalnya, telah berhasil mengajukan gugatan pencemaran nama baik terhadap orang-orang yang telah menerbitkan gambar dengan wajah penggugat yang disandingkan dengan tubuh aktor porno.

Solusinya juga dapat datang dari Perantara Internet, terutama platform media sosial penting yang menarik sebagian besar lalu lintas web. Menanggapi pertumbuhan deep fake, Perantara Internet besar seperti Reddit, Twitter, dan Pornhub memutuskan untuk mengatur diri sendiri dan memperbarui ketentuan penggunaan mereka untuk melarang konten pornografi pertukaran wajah yang didukung AI. Google kini juga mengizinkan pengguna untuk meminta penghapusan "citra pornografi sintetis yang tidak disengaja" dari diri mereka sendiri. Meskipun demikian, meskipun deep fake pornografi sebagian besar telah menghilang dari bagian media sosial yang muncul, pencarian Google yang cepat sudah cukup untuk menunjukkan bahwa kumpulan deep fake masih mudah diakses dan terus bertambah.

Mungkin bahkan lebih sulit untuk membatasi proliferasi deep fake non-pornografi yang digunakan sebagai berita palsu. Sebagai permulaan, modifikasi gambar, bahkan ketika dilakukan dengan teknologi AI, tidaklah ilegal dan mungkin memang dilindungi oleh hak atas kebebasan berbicara. Di Amerika Serikat, misalnya, keputusan Mahkamah Agung telah menyatakan bahwa kepalsuan saja tidak cukup untuk menghapus perlindungan Amandemen Pertama. Berdasarkan hukum hak cipta, hal itu mungkin juga dilindungi oleh penggunaan yang wajar. Selain itu, seperti yang telah kita lihat tidak hanya dengan kasus Nyonya Ayyub tetapi juga dalam pemilihan umum baru-baru ini di AS, di Prancis dan di Brasil, karena struktur media sosial, informasi palsu dapat menyebar dengan cepat dan menyebabkan kerusakan bahkan setelah dibantah.

Tanggung Jawab yang Lebih Besar bagi Perantara Internet?

Banyak anggota parlemen kini memandang deepfake, dan secara lebih luas, berita palsu, sebagai ancaman serius bagi lembaga demokrasi. Beberapa telah menyerukan penerapan kewajiban yang lebih ketat pada platform media sosial untuk memantau konten pihak ketiga.

Usulan Regulasi Baru Perantara Internet
Di Kanada, sebuah komite di Dewan Rakyat telah merekomendasikan amandemen Undang-Undang Pemilu Kanada untuk menambahkan tanggung jawab bagi Perantara Internet untuk menghapus "konten yang memfitnah, menipu, dan dimanipulasi secara jahat (misalnya, video 'deep fake')."
Di Amerika Serikat, anggota parlemen di Senat dan Dewan Perwakilan Rakyat bersama-sama memperkenalkan Undang-Undang Akuntabilitas Algoritmik tahun 2019, yang jika

disahkan, akan mengharuskan perusahaan-perusahaan besar untuk mengaudit sistem AI mereka untuk mendeteksi bias diskriminatif, serta risiko privasi dan keamanan.

Dewan Eropa mencatat bahwa "kampanye disinformasi daring yang ditargetkan, yang dirancang khusus untuk menebar ketidakpercayaan dan kebingungan serta mempertajam perpecahan yang ada di masyarakat, juga dapat berdampak buruk pada proses demokrasi."

Di Britania Raya, pemerintah Inggris telah mengusulkan peraturan pemerintah baru untuk mengatur penyebaran konten berbahaya dan kekerasan di platform media sosial. Sanksi yang diusulkan mencakup denda besar dan pemblokiran akses ke situs web.

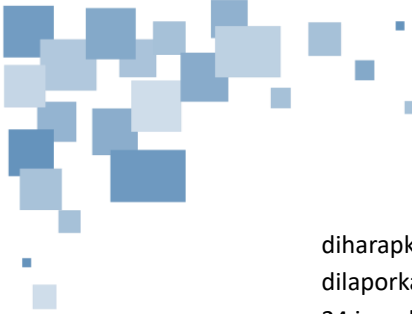
Hal ini menunjukkan bahwa kita mungkin berada di ambang pembentukan rezim pertanggungjawaban baru bagi Perantara Internet. Sejak pertengahan 1990an, AS dan Uni Eropa telah menegakkan undang-undang dengan kekebalan yang luas bagi platform daring yang sebagian besar dianggap sebagai "sekadar penyalur" informasi, alih-alih produsen konten. Dengan pendekatan paling liberal, AS mengadopsi pasal 230 *Communications Decency Act* untuk melindungi Perantara dari klaim non-HAKI dan non-pidana yang timbul dari konten pihak ketiga yang dipublikasikan di platform mereka, termasuk dari ganti rugi yang adil seperti putusan pengadilan. Untuk klaim kekayaan intelektual, AS memilih rezim perlindungan hak cipta (*safe harbour*) pada tahun 1998 dengan *Digital Millennium Copyright Act* yang melindungi Perantara dari tanggung jawab jika mereka mematuhi langkah-langkah pemberitahuan dan penghapusan. Untuk Eropa, pendekatan perlindungan hak cipta serupa, yang berlaku untuk semua jenis konten dan bukan hanya HAKI, telah berlaku sejak tahun 2000.

Kedua rezim legislatif tersebut dirancang sebelum munculnya media sosial seperti yang kita kenal sekarang. Kini, beragam suara mendesak revisi undang-undang tersebut, terutama untuk melawan maraknya berita bohong secara lebih efektif. Konsep "sekadar penyalur" Perantara Internet, khususnya platform media sosial, sudah ketinggalan zaman terutama jika kita mempertimbangkan pemantauan ekstensif terhadap konten yang mereka lakukan dan model bisnis mereka yang berlaku sebagai pedagang perhatian. Perubahan legislatif baru-baru ini menunjukkan pergeseran dari rezim imunitas menuju kewajiban yang lebih memberatkan bagi Perantara Internet.

Beberapa negara, seperti Australia, telah bereaksi terhadap manipulasi media sosial yang brutal baru-baru ini pembantaian Christchurch dalam kasus ini dengan memberlakukan undang-undang yang memaksa Perantara untuk segera menghapus "materi kekerasan yang menjijikkan". Jika mereka gagal melakukannya, berdasarkan undang-undang Australia yang baru, pejabat perusahaan dapat menghadapi hukuman penjara hingga tiga tahun dan perusahaan itu sendiri dapat didenda hingga 10% dari omzet tahunannya.

Jerman

Jerman baru-baru ini mengadopsi *Act to Improve Enforcement of Law in Social Networks* (NetzDG) yang berlaku untuk jejaring sosial nirlaba dengan lebih dari 2 juta pengguna di negara tersebut. Dengan tunduk pada denda yang signifikan, Perantara memiliki kewajiban baru untuk segera memblokir atau menghapus konten pihak ketiga yang melanggar hukum dan untuk memelihara mekanisme yang efektif dan transparan untuk menangani keluhan tentang konten yang melanggar hukum. Yang penting, Perantara



diharapkan untuk membuat keputusan hukum tentang keabsahan konten yang dilaporkan. Mereka harus menghapus konten yang "jelas melanggar hukum" dalam waktu 24 jam dan, untuk konten yang lebih ambigu, dalam waktu tujuh hari.

India

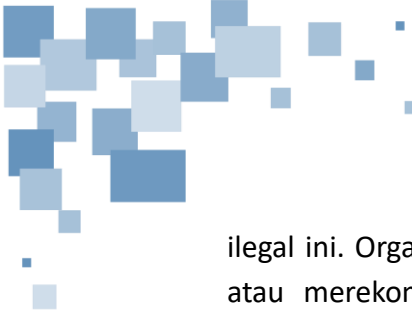
Pemerintah India juga sedang mempertimbangkan kembali undang-undang Perantara Internet, terutama pengecualian tanggung jawab, sebagai reaksi atas insiden hukuman gantung massal di negara tersebut. Rancangan undang-undang saat ini akan memberlakukan tugas proaktif untuk mengidentifikasi dan menghapus konten yang melanggar hukum kepada Perantara dan mencakup kewajiban bagi mereka untuk menggunakan "alat otomatis" guna mencapai tujuan tersebut.

Agaknya perangkat lunak AI bisa menjadi salah satu alat ini, dan para pendukung identifikasi konten otomatis, terutama di DARPA, berharap AI dapat digunakan sebagai perisai terhadap deep fake dan berita palsu. Dalam kesaksiannya baru-baru ini di depan Kongres AS, CEO Facebook Mark Zuckerberg mengatakan bahwa "dalam jangka panjang, membangun perangkat AI akan menjadi cara yang terukur untuk mengidentifikasi dan membasmi sebagian besar konten berbahaya ini." Menurut Zuckerberg, Facebook telah meraih beberapa keberhasilan dengan propaganda teroris, tetapi yang lain meragukan hal yang sama dapat dilakukan dengan berita palsu, terutama dalam jangka pendek, terutama karena kompleksitas wacana politik.

Risiko meningkatnya tanggung jawab Perantara Internet atas berita palsu yang dipublikasikan pihak ketiga adalah bahwa platform mungkin menjadi terlalu bersemangat dan secara preemptif menghapus konten politik dan sosial yang kontroversial, namun sah. Misalnya, dapat dibayangkan lawan politik menggunakan mekanisme pengaduan agar konten lawan mereka dihapus. Dikombinasikan dengan kekuatan AI, hal ini dapat membatasi kebebasan berekspresi di platform media sosial. Dewan Eropa juga memperingatkan adanya pelanggaran terhadap kebebasan berekspresi, dan merekomendasikan agar langkah-langkah moderasi konten yang diambil oleh Perantara Internet bersifat transparan, non-diskriminatif, dan menggunakan cara yang paling tidak membatasi.

Keseimbangan hukum harus diciptakan untuk memungkinkan penghapusan konten berbahaya secara cepat sekaligus melindungi kebebasan berbicara. Misalnya, terinspirasi oleh NetzDG Jerman, hal ini dapat dicapai melalui pembentukan mekanisme peninjauan independen dengan keadilan prosedural yang memadai yang dapat menilai keluhan terhadap konten, tetapi juga terhadap penghapusan. Hal ini akan menghindari penempatan Perantara Internet dalam peran ganda sebagai penengah kebenaran dan penengah kebebasan berbicara politik yang tepat. Perlindungan safe harbour akan tetap tersedia, tetapi dengan kriteria yang diperbarui yang harus dipenuhi.

Setelah dipuji sebagai alat penting demokrasi dan kebebasan berekspresi, terutama selama musim semi Arab tahun 2011, bulan madu kita dengan media sosial tampaknya telah berakhir dengan menumpuknya laporan tentang maraknya penyebaran kebencian, misinformasi, kebohongan, dan kampanye manipulasi daring. Dalam konteks tersebut, AI jelas berpotensi mempercepat atau secara drastis mengurangi penyebaran konten berbahaya dan



ilegal ini. Organisasi, termasuk perantara internet, yang berperan penting dalam menyaring atau merekomendasikan konten akan memainkan peran yang semakin penting dalam membantu mendorong hasil yang terakhir.

2.6 TUJUAN ETIS DAN MANFAAT SOSIAL

Hukum seringkali berkaitan dengan detail-detail kecil, aturan-aturan yang samar, dan preseden-preseden yang telah lama terlupakan. Sangat mudah bagi seorang pengacara yang berpraktik, yang berfokus pada kontrak terbaru yang harus disusun, kasus yang harus diselesaikan, dan merger perusahaan-perusahaan besar (sementara tetap mengikuti perkembangan inovasi teknologi terkini), untuk melupakan gambaran yang lebih besar. Namun, hukum juga tentang menciptakan masyarakat yang tidak hanya melindungi kondisi manusia tetapi juga lingkungan kita yang rapuh. Terkait setiap isu yang telah kami uraikan secara singkat dalam bab ini, bisnis, pembuat kebijakan pemerintah dan pengacara akan memiliki keputusan penting yang harus dibuat terkait peran AI yang bermanfaat dan tidak merugikan.


Pencapaian tujuan ini menuntut peningkatan rasa tanggung jawab dari semua pemangku kepentingan yang terlibat. Hans Jonas, yang memulai bab ini bersama kami, menyimpulkan bahwa semakin canggih teknologi yang kita ciptakan, semakin besar pula tanggung jawab yang kita pikul untuk mencegah dampak negatif. Pada bab berikutnya, kami akan menyajikan kerangka kerja akuntabilitas manusia untuk pengembangan, penerapan, dan penggunaan sistem AI secara lebih rinci.

Prinsip

Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI dan undang-undang nasional yang mengatur penggunaannya harus mewajibkan identifikasi tujuan implementasi tersebut dan memastikan bahwa tujuan tersebut konsisten dengan tujuan etis secara keseluruhan, yaitu kebaikan dan non-kejahatan, serta prinsip-prinsip lain dari Kerangka Kebijakan untuk AI yang Bertanggung Jawab.

Prinsip-prinsip Utama

- Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus melakukannya dengan cara yang sesuai dengan agensi manusia dan penghormatan terhadap hak asasi manusia (termasuk kebebasan dari diskriminasi).
- Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus memantau implementasi sistem AI tersebut dan bertindak untuk memitigasi konsekuensi dari sistem AI tersebut (baik yang disengaja maupun tidak disengaja) yang tidak konsisten dengan tujuan etis, yaitu kebaikan dan non-kejahatan, serta prinsip-prinsip lain dari Kerangka Kebijakan untuk AI yang Bertanggung Jawab yang ditetapkan dalam kerangka ini.
- Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus menilai implikasi sosial, politik, dan lingkungan dari pengembangan, penerapan, dan penggunaan tersebut dalam konteks Penilaian Dampak AI yang Bertanggung Jawab dan



terstruktur yang menilai risiko bahaya dan, jika diperlukan, mengusulkan strategi mitigasi terkait risiko tersebut.

Pekerjaan dan Otomatisasi

- Organisasi yang menerapkan sistem AI di tempat kerja harus menyediakan kesempatan bagi karyawan yang terdampak untuk berpartisipasi dalam proses pengambilan keputusan terkait implementasi tersebut.
- Pertimbangan harus diberikan mengenai apakah dari perspektif teknologi, hal ini dapat dicapai untuk memastikan bahwa semua kemungkinan kejadian telah diputuskan sebelumnya dalam sistem AI guna memastikan perilaku yang konsisten. Jika hal ini tidak praktis, organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus mempertimbangkan setidaknya sejauh mana mereka mampu membatasi hasil keputusan sistem AI pada rentang respons yang wajar dan tidak menyimpang, dengan mempertimbangkan konteks yang lebih luas, dampak keputusan, dan kesesuaian moral dari "menimbang yang tidak dapat ditimbang" seperti nyawa vs. nyawa.
- Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI yang berdampak pada ketenagakerjaan harus melakukan Penilaian Dampak AI yang Bertanggung Jawab untuk menentukan dampak bersih dari implementasi tersebut.
- Pemerintah harus memantau secara ketat perkembangan otomatisasi berbasis AI untuk mengidentifikasi sektor-sektor ekonomi mereka di mana pekerja manusia paling terdampak. Pemerintah harus secara aktif meminta dan memantau data serta komentar industri, karyawan, dan pemangku kepentingan lainnya mengenai dampak sistem AI di tempat kerja dan harus mengembangkan forum terbuka untuk berbagi pengalaman dan praktik terbaik.
- Pemerintah harus mendorong kebijakan pendidikan yang membekali semua anak dengan keterampilan, pengetahuan, dan kualitas yang dibutuhkan oleh ekonomi baru dan yang mendorong pembelajaran seumur hidup.
- Pemerintah harus mendorong terciptanya kesempatan bagi orang dewasa untuk mempelajari keterampilan baru yang bermanfaat, terutama bagi mereka yang tergantikan oleh otomatisasi.
- Pemerintah harus mempelajari kelayakan dan kelayakan sistem kesejahteraan dan tunjangan sosial baru untuk membantu mengurangi, jika diperlukan, ketimpangan sosial-ekonomi yang disebabkan oleh pengenalan sistem AI dan otomatisasi robotik.

Dampak Lingkungan

- Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus menilai dampak lingkungan secara keseluruhan dari sistem AI tersebut, di seluruh implementasinya, termasuk konsumsi sumber daya, biaya energi untuk penyimpanan dan pemrosesan data, serta efisiensi energi bersih atau manfaat lingkungan yang mungkin dihasilkannya. Organisasi harus berupaya untuk mempromosikan dan menerapkan penggunaan sistem AI dengan tujuan mencapai netralitas karbon atau pengurangan karbon secara keseluruhan.

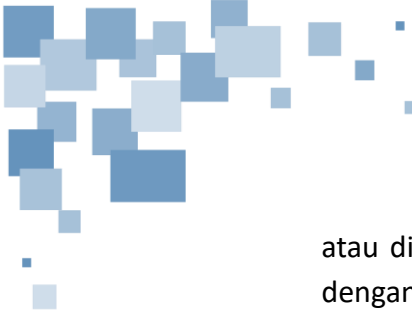
- Pemerintah didorong untuk menyesuaikan rezim regulasi dan/atau mempromosikan rezim regulasi mandiri industri terkait masuk pasar dan/atau adopsi sistem AI sedemikian rupa sehingga potensi paparan (dalam hal 'peluang vs. risiko') yang mungkin timbul dari pengoperasian publik sistem AI tersebut dapat tercermin secara wajar. Rezim khusus untuk penerimaan perantara dan terbatas guna memungkinkan pengujian dan penyempurnaan pengoperasian sistem AI dapat membantu mempercepat penyelesaian sistem AI dan meningkatkan keamanan serta keandalannya.
- Untuk memastikan dan menjaga kepercayaan publik terhadap kendali akhir manusia, pemerintah harus mempertimbangkan penerapan aturan yang menjamin investigasi yang komprehensif dan transparan terhadap dampak buruk dan tak terduga dari sistem AI yang telah terjadi melalui penggunaannya, khususnya jika dampak tersebut memiliki konsekuensi yang mematikan atau merugikan bagi manusia yang menggunakan sistem tersebut. Investigasi semacam itu harus digunakan untuk mempertimbangkan penyesuaian kerangka regulasi untuk sistem AI, khususnya untuk mengembangkan, jika memungkinkan dan dapat dicapai, pemahaman yang lebih menyeluruh tentang bagaimana dan kapan sistem tersebut harus diserahkan secara halus kepada operator manusianya dalam skenario kegagalan.
- AI memiliki potensi khusus untuk mengurangi pemborosan sumber daya dan inefisiensi yang berbahaya bagi lingkungan. Penelitian AI mengenai tujuan-tujuan ini harus didorong. Untuk melakukannya, kebijakan harus diterapkan untuk memastikan data yang relevan dapat diakses dan digunakan dengan cara yang konsisten dengan penghormatan terhadap prinsip-prinsip lain dari Kerangka Kebijakan untuk AI yang Bertanggung Jawab seperti Keadilan dan Non-Diskriminasi, Data Terbuka dan Persaingan yang Adil serta Privasi, Akses dan Persetujuan yang Sah.

AI yang Dipersenjatai

- ✚ Penggunaan sistem senjata otonom mematikan (LAWS) harus menghormati prinsip dan standar serta konsisten dengan hukum humaniter internasional tentang penggunaan senjata dan hukum hak asasi manusia internasional yang lebih luas.
- ✚ Pemerintah harus menerapkan mekanisme multilateral untuk mendefinisikan, menerapkan, dan memantau kepatuhan terhadap perjanjian internasional mengenai pengembangan, penggunaan, dan perdagangan LAWS yang etis.
- ✚ Pemerintah dan organisasi harus menahan diri dari pengembangan, penjualan, atau penggunaan sistem senjata otonom mematikan (LAWS) yang mampu memilih dan menyerang target tanpa kendali dan pengawasan manusia dalam semua konteks.
- ✚ Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus memberi tahu karyawan mereka ketika mereka ditugaskan untuk proyek yang berkaitan dengan HUKUM.

Persenjataan Informasi Palsu atau Menyesatkan

- ❖ Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI untuk menyaring atau mempromosikan konten informasi di platform internet yang dibagikan



atau dilihat oleh penggunanya harus mengambil langkah-langkah yang wajar, sesuai dengan hukum yang berlaku, untuk meminimalkan penyebaran informasi palsu atau menyesatkan jika terdapat risiko material bahwa informasi palsu atau menyesatkan tersebut dapat menyebabkan kerugian yang signifikan bagi individu, kelompok, atau lembaga demokrasi.

- ❖ AI berpotensi membantu mengidentifikasi secara efisien dan proaktif (dan, jika sesuai, menekan) konten yang melanggar hukum seperti ujaran kebencian atau informasi palsu atau menyesatkan yang dipersenjatai. Penelitian AI tentang cara-cara untuk mencapai tujuan-tujuan ini dengan cara yang konsisten dengan kebebasan berekspresi harus didorong.
- ❖ Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI pada platform untuk menyaring atau mempromosikan konten informasi yang dibagikan atau dilihat oleh penggunanya harus menyediakan mekanisme yang memungkinkan pengguna menandai konten yang berpotensi berbahaya secara tepat waktu.
- ❖ Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI pada platform untuk menyaring atau mempromosikan konten informasi yang dibagikan atau dilihat oleh penggunanya harus menyediakan mekanisme yang memungkinkan penyedia konten untuk menggugat penghapusan konten tersebut oleh organisasi tersebut dari jaringan atau platform mereka secara tepat waktu.
- ❖ Pemerintah harus menyediakan pedoman yang jelas untuk membantu Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI pada platform mengidentifikasi konten terlarang yang menghormati hak atas martabat dan kesetaraan serta hak atas kebebasan berekspresi.
- ❖ Pengadilan harus tetap menjadi penentu akhir atas konten yang sah.



BAB 3

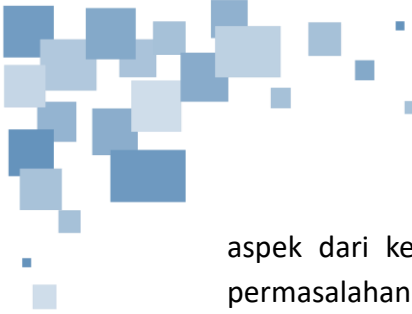
AKUNTABILITAS

Sifat AI yang transformatif dan disruptif kini seharusnya sudah dipahami dengan baik. Kita telah melihat dalam Pendahuluan bagaimana AI telah berevolusi dari sistem berbasis aturan menjadi jaringan saraf canggih yang mampu mengenali dan mengidentifikasi teman Anda di foto terbaru Anda, mendeteksi tumor, mengemudikan truk jarak jauh, atau merekomendasikan strategi investasi. Lebih lanjut, pada bab sebelumnya, kita mulai mengeksplorasi beberapa dampak AI yang berpotensi paling signifikan terhadap aspek-aspek penting kehidupan manusia: masa depan pekerjaan, upaya untuk memastikan keberlanjutan lingkungan, persenjataan AI dan penggunaannya dalam konflik bersenjata, serta kesehatan wacana demokrasi. Berusaha untuk memperkuat konsensus internasional yang berkembang, kami berpendapat bahwa kerangka kerja kebijakan untuk AI yang bertanggung jawab sangat penting dan harus berpusat pada kesejahteraan manusia dan lingkungan alam kita (yang selanjutnya kami sebut sebagai "pendekatan yang berpusat pada manusia").

Pendekatan yang berpusat pada manusia bukan sekadar cara untuk menilai dampak AI dan mengarahkan pengembangan etis teknologi ini. Pendekatan ini juga menyiratkan prinsip lebih lanjut, yaitu akuntabilitas manusia. Manusia, baik sebagai individu maupun melalui perusahaan swasta atau pemerintah, akan mengembangkan, menerapkan, dan menggunakan sistem AI. Oleh karena itu, wajar saja jika mereka melakukannya dengan cara yang bertanggung jawab, etis, dan sah. Akuntabilitas perlu dijadikan prinsip dasar pengembangan, penerapan, dan penggunaan AI dan bab ini akan menyajikan kerangka kerja akuntabilitas umum yang akan dikembangkan lebih lanjut dalam bab-bab mendatang.

Konsep akuntabilitas yang dijelaskan dalam bab ini luas. Berlandaskan norma-norma nasional dan internasional tentang hak asasi manusia, konsepsi kami tentang akuntabilitas tidak hanya mencakup kerangka kerja pertanggungjawaban perdata dan pidana yang telah dikembangkan dengan baik (jika berlaku), tetapi juga prinsip-prinsip tata kelola perusahaan yang baik. Dengan menempatkan prinsip akuntabilitas di inti kerangka kerja AI etis yang kami usulkan, tujuan kami adalah untuk membantu memastikan bahwa sistem AI tidak "berjalan sendiri", bahwa sistem tersebut tidak beroperasi tanpa kendali, dan bahwa, jika terjadi hasil yang merugikan, ada seseorang yang bertanggung jawab untuk mengambil tindakan perbaikan dan memberikan ganti rugi. Jika Gaak lolos lagi, manusia harus selalu berada di dekatnya untuk mematakannya!

Ciri-ciri dasar akuntabilitas tidak ada dalam kekosongan hukum. Sudah ada banyak sekali rezim hukum yang berlaku untuk AI, misalnya rezim kontrak dan gugatan tradisional, tanggung jawab produk yang ketat, serta rezim regulasi yang berlaku umum atau khusus sektor. Namun demikian, sebagaimana halnya teknologi disruptif (terutama teknologi dengan potensi dampak disruptif sebesar itu, di berbagai bidang), kita perlu menilai implikasi AI secara komprehensif terhadap kerangka hukum yang ada, dengan selalu menyadari bahwa beberapa



aspek dari kerangka hukum yang ada mungkin perlu diubah sebagai respons terhadap permasalahan yang ditimbulkan oleh AI.

Seperti yang akan kita bahas lebih lanjut di bawah, hal ini mungkin terjadi, sebagai contoh, untuk rezim tanggung jawab pengemudi yang perlu disesuaikan dengan pengenalan kendaraan otonom. Singkatnya, tidak selalu mungkin untuk terus menuangkan anggur baru ke dalam botol lama. Bila perlu, proses pemikiran ulang standar dan hukum harus berpedoman pada prinsip-prinsip kebaikan dan non-kejahatan, serta menjaga akuntabilitas manusia atas aktivitas AI. Tak pelak lagi, berbagai yurisdiksi di seluruh dunia akan mengadopsi model hukum yang berbeda untuk merespons tantangan AI, berdasarkan kebutuhan, budaya, dan sejarah masing-masing. Namun demikian, benang merahnya haruslah penegasan agensi manusia, dan dengan demikian akuntabilitas, atas mesin-mesin cerdas baru kita.

Memperjelas kewajiban seharusnya membantu membatasi terciptanya kesenjangan akuntabilitas, zona di mana sistem AI beroperasi tanpa pengawasan manusia dan di bawah tanggung jawab siapa pun. Perubahan teknologi yang cepat dapat menciptakan kebingungan tentang aktor mana yang bertanggung jawab, dan sejauh mana. Respons masyarakat terhadap tantangan yang ditimbulkan oleh Perantara Internet menggambarkan bagaimana regulator dan perusahaan bergulat dengan cara terbaik untuk menutup kesenjangan akuntabilitas. Pertama, seperti yang telah kita lihat di bab sebelumnya, legislator di banyak negara di seluruh dunia saat ini sedang meninjau kembali rezim pertanggungjawaban mereka untuk Perantara Internet dan menguji model-model baru untuk memerangi penulisan konten berbahaya yang tak terkendali dan viral. Kedua, dengan larangan pornografi palsu yang dipimpin oleh industri, kita juga telah melihat perusahaan swasta mengambil langkah-langkah sukarela yang kuat, sebuah contoh tata kelola proaktif. Dengan demikian, menutup kesenjangan akuntabilitas akan dicapai melalui upaya gabungan antara upaya legislatif dan regulasi pemerintah dan langkah-langkah sukarela yang diambil oleh para pelaku korporasi.

Dalam bab ini, kami akan menyajikan kerangka kerja akuntabilitas umum, dengan fokus pada individu, pelaku korporasi swasta, dan pemerintah. Dalam perspektif ini, organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI pada akhirnya harus bertanggung jawab atas kepatuhan mereka terhadap delapan prinsip Kerangka Kerja untuk AI yang Bertanggung Jawab. Namun, sebelum melangkah lebih jauh, bagian berikut terlebih dahulu mengkaji (dan pada akhirnya membantah) sebuah proposisi yang semakin menarik perhatian dalam beberapa tahun terakhir: usulan bahwa sistem AI harus diberikan status hukum terpisah dan bahwa kita harus membuat sistem AI itu sendiri bertanggung jawab atas perilakunya sendiri.

3.1 AI DAN STATUS HUKUM

Pada bulan Oktober 2017, Arab Saudi memberikan kewarganegaraan kepada sistem AI robotik bernama "Sophia", menjadikannya negara pertama yang memberikan kewarganegaraan kepada robot. Dapat dikatakan, ini pada dasarnya merupakan aksi publisitas dengan sedikit konsekuensi hukum, tetapi tetap saja memberikan semangat baru pada

perdebatan tentang status hukum sistem AI. Mungkinkah gestur simbolis ini menjadi langkah pertama menuju pemberian status hukum kepada sistem AI dan robot cerdas?


Beberapa tahun yang lalu, pertanyaan ini mungkin hanya terbatas pada ranah fiksi ilmiah hukum, setara dengan hak milik Mars. Namun, kemampuan baru yang ditunjukkan oleh sistem AI, terutama bagaimana mereka tampaknya belajar dan bertindak secara otonom, telah memaksa kita untuk menangani masalah ini dengan lebih serius. Meskipun belum ada sistem hukum yang secara resmi mengakui kepribadian hukum dalam AI, konsep ini sedang dipertimbangkan secara serius oleh beberapa legislator.

Perkembangan Internasional dalam Personalitas Hukum	
Estonia	Estonia (saat tulisan ini dibuat) sedang menggodok undang-undang yang akan memberikan status hukum baru kepada robot. Status hukum tersebut akan berada di antara "personalitas hukum terpisah" dan "harta pribadi". Orang-orang baru tersebut akan disebut "agen robot". Rancangan undang-undang diperkirakan akan diajukan kepada legislator Estonia pada Juni 2019.
Malta	Pada November 2018, Malta mengumumkan bahwa Pemerintah akan menyusun strategi nasional tentang AI dan menjajaki uji kewarganegaraan untuk robot.
Uni Eropa	Parlemen Eropa, pada Februari 2017, meminta Komisi Aturan Hukum Perdata tentang Robotika untuk mempelajari kemungkinan menciptakan status hukum khusus bagi robot. Alasannya adalah bahwa "setidaknya dengan demikian, robot otonom yang paling canggih dapat ditetapkan memiliki status sebagai orang elektronik yang bertanggung jawab untuk memperbaiki kerusakan yang mungkin ditimbulkannya, dan mungkin menerapkan personalitas elektronik pada kasus-kasus di mana robot membuat keputusan otonom atau berinteraksi dengan pihak ketiga secara independen."

Haruskah kita mempertimbangkan untuk mengakui hak dan kewajiban "orang elektronik" bersama dengan badan hukum perseorangan dan badan hukum korporasi? Tidak ada hal inheren dalam konsep badan hukum yang akan menghalangi hal ini. Secara hukum, dan sederhananya, menjadi badan hukum berarti menjadi subjek hak dan kewajiban. Badan hukum atau "subjek hukum" dapat menjalin hubungan kontraktual, menuntut dan sebagai balasannya dituntut, menjadi sasaran sanksi, memperoleh hak milik, dan menegakkan hak-hak mereka.

Cendekiawan hukum Amerika John Chipman Gray berpendapat pada zamannya bahwa tidak mungkin ada hak, dan oleh karena itu tidak ada badan hukum, tanpa kehendak untuk menjalankan hak-hak tersebut. Oleh karena itu, manusia adalah "badan hukum" yang alami dan primordial, meskipun beberapa kategori manusia secara historis ditolak status ini. Namun, tentu saja, kita telah lama memutuskan untuk bergerak melampaui gagasan terbatas tentang badan hukum ini dengan pembentukan korporasi. Korporasi, yang dibentuk oleh hukum, memiliki kepribadian hukumnya sendiri, terpisah dari pemiliknya. Dengan beberapa pengecualian terbatas, mereka memiliki hak dan kewajiban yang sama dengan orang perseorangan.

Singkatnya, kepribadian hukum adalah "fiksi hukum". Dari perspektif hukum dan konseptual, tidak ada yang secara apriori akan mencegah legislator memberikan kepribadian hukum kepada sistem AI dan robot cerdas. Namun, apakah itu berarti mereka harus



melakukannya? Tidak. Menanggapi penyelidikan Parlemen Eropa, 285 pakar AI berpendapat dalam surat terbuka bahwa "mengasani kepribadian elektronik kepada robot berisiko salah menempatkan tanggung jawab moral, akuntabilitas kausal, dan tanggung jawab hukum terkait kesalahan dan penyalahgunaannya." Argumen-argumen ini meyakinkan. Memberikan kepribadian hukum kepada sistem AI akan bertentangan dengan prinsip akuntabilitas manusia.

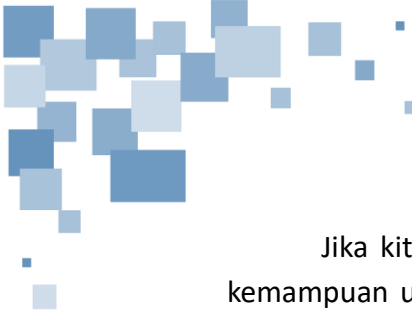
Ini bukan berarti para pendukung pemberian status badan hukum kepada sistem AI yang cukup maju memiliki argumen yang lemah. Ada dua versi utama proposal AI sebagai badan hukum, keduanya dengan pendukung yang kuat, tetapi secara keseluruhan kami percaya sistem AI harus diperlakukan sebagai alat, alih-alih sebagai pribadi. Versi pertama akan memberikan sistem AI yang cukup maju status badan hukum yang sama dengan manusia dewasa, pada dasarnya karena alasan moral.

Pada tahun 2015, futuris Ray Kurzweil secara khusus berpendapat bahwa "jika AI dapat meyakinkan kita bahwa responsnya berada pada level manusia, dan jika kita yakin bahwa ia mengalami kondisi subjektif yang diklaimnya, maka kita akan menerima bahwa ia mampu mengalami penderitaan dan kegembiraan. Pada titik itu, AI akan menuntut hak, dan karena kemampuan kita untuk berempati, kita akan cenderung memberikannya." Versi argumen untuk status badan hukum AI ini mengandaikan skenario di mana AGI menjadi kenyataan.

Para pendukung versi kedua proposal ini akan memberikan status hukum bahkan kepada sistem AI yang sempit, tanpa harus memberikan status yang sama dengan manusia. Menurut pemerintah Estonia, status AI dapat serupa dengan korporasi dan pihak lain telah mengusulkan serangkaian hak dan kewajiban berbeda yang dirancang khusus untuk sistem AI. Hal ini akan memiliki banyak konsekuensi, tetapi dari sudut pandang akuntabilitas, seperti yang disebutkan di atas oleh Parlemen Eropa, hal ini akan memungkinkan kita untuk meminta pertanggungjawaban langsung sistem AI dan menuntut mereka atas kerusakan yang ditimbulkannya atau kejahatan yang mereka lakukan.

Dalam kasus kecelakaan yang melibatkan kendaraan otonom penuh, hal ini akan memungkinkan pihak yang terluka untuk menuntut langsung mobil itu sendiri. Sebuah hak milik atau yang setara akan dilampirkan pada setiap AI, yang juga dapat membeli asuransi, yang dengannya mereka dapat membayar ganti rugi atau kompensasi. Posisi ini memiliki tujuan yang sah untuk mengusulkan cara untuk mengatasi kesenjangan akuntabilitas.

Mereka yang mendukung pemberian status hukum kepada sistem AI sering kali merujuk pada otonomi mereka yang tampak dan kemampuan mereka untuk belajar melampaui pengetahuan manusia. Berdasarkan sifat "kotak hitam" jaringan saraf, terkadang dikemukakan bahwa, karena manusia mungkin sulit memprediksi apa yang mungkin dilakukan sistem AI, maka tidak adil untuk meminta pertanggungjawaban seseorang atau badan hukum atas sesuatu yang tidak dapat diprediksi secara wajar, yaitu, perilaku otonom AI. Misalnya, *mens rea* tetap menjadi ciri khas pertanggungjawaban pidana dan banyak rezim pertanggungjawaban perdata didasarkan pada kelalaian. Dengan demikian, jika AI bertindak secara otonom dan tidak terduga, mereka berpendapat bahwa tidak adil atau tidak praktis untuk meminta pertanggungjawaban manusia.



Jika kita menghadapi sistem AI dengan kecerdasan yang sama dengan kita, dengan kemampuan untuk merasakan kegembiraan, kesedihan, dan semua emosi manusia lainnya serta dengan kapasitas untuk membuat pilihan moral yang wajar, akan jauh lebih sulit untuk menolak hak-hak mereka. Namun, sebagaimana dibahas dalam Pendahuluan, jenis AGI ini belum diciptakan, dan mungkin tidak akan pernah ada. Jadi, versi pertama dari proposal AI sebagai badan hukum sangat hipotetis. Untuk saat ini, kami hanya membahas "AI sempit", yaitu AI yang menguasai tugas-tugas kognitif spesifik, tetapi tanpa fleksibilitas kecerdasan manusia. Sistem AI jenis ini pada dasarnya hanyalah alat, meskipun sangat canggih. Sistem AI harus tetap menjadi tanggung jawab pengembang, pelaksana, atau penggunanya.

Sedangkan untuk versi kedua proposal ini, kami yakin proposal tersebut lebih-lebihkan ketidakpastian dan otonomi sistem AI. Sebagaimana dibahas lebih rinci dalam pembahasan kami tentang Prinsip 8 AI dan Hak Kekayaan Intelektual, pertanyaan tentang kepribadian hukum AI telah muncul secara khusus dalam konteks hak kekayaan intelektual: dapatkah AI menjadi pencipta karya seni?


Proyek Rembrandt Berikutnya

Dalam "Proyek Rembrandt Berikutnya", misalnya, sebuah tim yang dipimpin oleh ING dan Microsoft mengembangkan sistem AI yang dapat menganalisis keseluruhan karya katalog Rembrandt dan kemudian menghasilkan lukisan-lukisan baru dengan gaya sang maestro. Meskipun tampak sebagai pencapaian kecerdasan yang luar biasa, sistem AI pada akhirnya tidak bertanggung jawab atas proses kreatif. Kapasitas sistem AI ini terbatas pada tujuan yang ditetapkan oleh penciptanya (misalnya, tidak dapat melukis Picasso atau memainkan piano). Ilmuwan data manusia memutuskan data mana yang akan dimasukkan ke dalam algoritma pembelajaran mesin buatan manusia. Meskipun sistem AI menentukan fitur mana yang mendefinisikan gaya Rembrandt, para pengambil keputusan manusia hadir di setiap langkah proses, baik di tingkat teknologi maupun manajerial.

Singkatnya, jika terjadi pelanggaran hak cipta atau penciptaan "karya agung" yang dilindungi hak cipta oleh sistem AI, bukan sistem AI itu sendiri yang harus bertanggung jawab atau diberi penghargaan, melainkan individu dan perusahaan yang berada di balik proyek tersebut.

Prinsip ini bahwa individu, perusahaan, atau pemerintah harus bertanggung jawab, alih-alih AI itu sendiri juga berlaku untuk kasus-kasus di mana algoritma tampaknya memiliki tingkat otonomi yang lebih tinggi karena mereka dipercayakan dengan kekuatan pengambilan keputusan. AlphaGo, misalnya, dapat membuat gerakan-gerakan yang elegan dan menakutkan untuk memenangkan permainan, tetapi lingkup agensinya pada akhirnya ditentukan oleh pengembang dan pemilik perusahaannya.

Tingkat keterlibatan manusia ini, ketika sistem AI diterapkan dan digunakan di dunia nyata di mana sistem tersebut dapat menyebabkan kerusakan, membutuhkan akuntabilitas manusia. Demikian pula, Institut Insinyur Listrik dan Elektronika baru-baru ini mencatat bahwa "hampir semua aplikasi [sistem AI] saat ini dalam aplikasi hukum, seperti di sebagian besar bidang lainnya, memerlukan mediasi manusia dan kemungkinan akan terus demikian dalam



waktu dekat." Hal ini tidak menutup kemungkinan bahwa suatu hari nanti kita mungkin akan melihat AI yang cukup otonom yang dapat diberikan beberapa bentuk kepribadian hukum, tetapi hal ini untuk saat ini masih berada dalam ranah fiksi ilmiah dan paradigma AI sebagai alat hanya boleh dimodifikasi dengan kehati-hatian yang sebesar-besarnya.

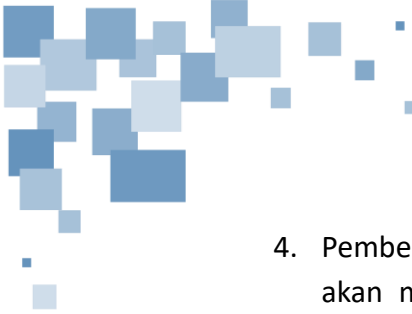
Salah satu risiko utama pemberian tanggung jawab hukum kepada sistem AI saat ini adalah memperkokoh, alih-alih menutup, kesenjangan antara tindakan sistem AI dan kendali pengembang atau penggunanya. Transparansi dan Keterjelasan, rata-rata orang akan memiliki sedikit pemahaman tentang bagaimana algoritma yang mengemudikan mobilnya mengambil keputusan. Oleh karena itu, mungkin masuk akal untuk mengalihkan sebagian tanggung jawab dari individu yang "mengemudi" karena "mengemudi" kendaraan tersebut menjadi semakin otonom. Namun, meskipun AI mungkin memaksa kita untuk mempertimbangkan kembali akuntabilitas aktor-aktor tertentu, hal itu harus dilakukan dengan cara yang mengalihkan tanggung jawab kepada aktor manusia lain, bukan kepada sistem AI itu sendiri. Dalam kasus kendaraan otonom, misalnya, ini bisa berarti rezim yang memberikan beban lebih besar kepada pengembang atau produsen, alih-alih kepada konsumen, dengan kewajiban untuk menguji algoritma di berbagai lingkungan. Beberapa tanggung jawab atas kecelakaan yang disebabkan oleh kendaraan yang dibantu AI bahkan dapat tetap berada di tangan pemilik mobil, seperti kewajiban untuk mengendalikan kemudi dalam kondisi cuaca tertentu.

Meminta pertanggungjawaban langsung kepada sistem AI berisiko melindungi aktor manusia dari tanggung jawab dan mengurangi insentif untuk mengembangkan dan menggunakan AI secara bertanggung jawab. Dalam kasus korporasi, pemilik manusia sudah diuntungkan (demi alasan kebijakan publik yang baik) dari kemampuan untuk bersembunyi di balik "tirai korporasi". Jika mereka dilindungi lebih lanjut oleh "tirai AI" kedua, kami menganggap bahwa tabir akuntabilitas akan menjadi terlalu jauh dan terlalu buram.

Meskipun menghubungkan kerusakan yang disebabkan oleh sistem AI dengan kesalahan atau tanggung jawab manusia mungkin dalam beberapa kasus sulit dan memerlukan keahlian khusus, 285 penandatanganan surat terbuka yang disebutkan di atas menganggap "pernyataan bahwa tanggung jawab atas kerusakan tidak mungkin dibuktikan" tidak benar karena "didasarkan pada penilaian yang terlalu tinggi terhadap kemampuan sebenarnya bahkan dari robot yang paling canggih sekalipun, pemahaman yang dangkal tentang ketidakpastian dan kapasitas belajar mandiri, dan persepsi robot yang terdistorsi oleh Fiksi Ilmiah dan beberapa pengumuman pers sensasional baru-baru ini."

Google telah menyatakan dukungannya terhadap surat terbuka ini, dengan menambahkan bahwa pemberian status badan hukum kepada AI tidaklah bijaksana karena empat alasan:

1. Hal ini tidak diperlukan karena kerangka hukum sudah membuat manusia atau korporasi bertanggung jawab.
2. Hal ini tidak praktis karena mesin tanpa kesadaran atau perasaan tidak dapat dihukum.
3. Tanggung jawab moral merupakan hak asasi manusia yang hakiki.

- 
4. Pemberian status badan hukum kepada AI rentan terhadap penyalahgunaan karena akan memungkinkan pelaku kejahatan untuk melindungi diri mereka sendiri dari tanggung jawab.

Singkatnya, sistem AI sangatlah kompleks, tetapi tetap merupakan produk rekayasa dan sains manusia. Pengembangannya didasarkan pada keahlian manusia yang terspesialisasi dan pengguna akhirnya akan selalu manusia, baik sebagai individu maupun melalui korporasi atau pemerintah. Para insinyur AI dan ilmuwan komputer yang mengembangkan AlphaGo mungkin tidak dapat mengalahkan program mereka sendiri, tetapi mereka tetap memahami, atau setidaknya seharusnya memahami, arsitekturnya dan jangkauan tindakan yang mungkin dilakukannya.


Prinsip akuntabilitas menyatakan bahwa sistem hukum seharusnya memperlakukan sistem AI sebagai alat, produk, atau layanan, alih-alih sebagai badan hukum. Konstruksi hukum akuntabilitas yang berlaku bagi perorangan dan badan hukum sudah memadai, meskipun beberapa mungkin perlu diadaptasi untuk AI. Hak dan kewajiban yang terkait dengan sistem AI pada akhirnya berada di tangan individu, perusahaan, dan badan publik yang mengembangkan, menerapkan, atau menggunakannya.

Dengan demikian, penerapan prinsip-prinsip etika dan kebijakan hukum yang disajikan dalam Kerangka Kerja untuk AI yang Bertanggung Jawab ini pada akhirnya menjadi tanggung jawab para aktor ini: manusia dan manusia itu sendiri, baik sebagai individu maupun yang tergabung dalam organisasi publik atau swasta. Sisa bab ini akan menyajikan gagasan dasar kerangka kerja akuntabilitas untuk implementasi AI yang etis. Kami akan menyimpulkan dengan diskusi yang lebih terfokus tentang prinsip-prinsip tata kelola perusahaan yang efektif sebagaimana diterapkan pada organisasi yang berpartisipasi dalam pengembangan, penerapan, dan penggunaan AI.

3.2 KERANGKA AKUNTABILITAS

Untuk menegaskan kembali prinsip akuntabilitas kami: organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus bertanggung jawab atas konsekuensi tindakan mereka. Para pemangku kepentingan AI tidak boleh mengelak dari tanggung jawab atas prasangka hukum atau kesalahan etika, atas kerusakan properti, atau atas cedera manusia yang disebabkan oleh AI, terlepas dari tingkat otonomi AI. Oleh karena itu, akuntabilitas harus dipahami dan diterapkan sedemikian rupa sehingga manusia selalu berada dalam posisi untuk memodifikasi, memantau, atau mengendalikan sistem AI, dan dengan demikian, bertanggung jawab atasnya.

Pendekatan akuntabilitas yang diusulkan ini bertentangan dengan sikap yang telah lama berlaku di sektor teknologi, yang esensinya secara ringkas tertuang dalam moto: "Bergerak cepat dan hancurkan segalanya." Perspektif terakhir ini menekankan pentingnya pengembangan teknologi inovatif, tetapi mengabaikan tanggung jawab para teknolog atas hasil produk dan layanan yang mereka kembangkan. Meskipun manfaat yang dapat diberikan AI signifikan, demikian pula kerugian yang dapat ditimbulkan oleh penggunaan AI yang tidak tepat, tidak etis, atau ilegal, dan mereka yang terlibat dalam pengembangan teknologi



revolusioner ini harus bertanggung jawab atas sistem yang mereka ciptakan. Kita tidak boleh membiarkan produk dan layanan AI yang belum teruji, cacat, atau tidak lengkap dipasarkan atau diproduksi atas dasar "caveat emptor".

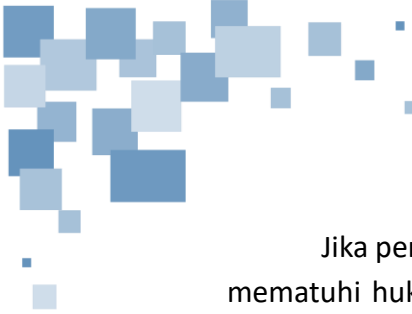
Oleh karena itu, ketika para pemangku kepentingan "merusak sesuatu", mereka harus siap menanggung akibatnya. CEO Facebook, Mark Zuckerberg, baru-baru ini mengambil sikap tegas dalam mendukung akuntabilitas perusahaan teknologi dan menyerukan pemerintah untuk menetapkan standar dan aturan yang jelas bagi internet. Mengenai regulasi privasi, beliau secara khusus menulis bahwa "regulasi tersebut harus menetapkan cara untuk meminta pertanggungjawaban perusahaan seperti Facebook dengan menjatuhkan sanksi ketika kami melakukan kesalahan." Kami yakin hal ini juga berlaku untuk regulasi AI.

Hal ini membawa kita pada pertanyaan utama bab ini: bagaimana dan sejauh mana para pemangku kepentingan harus bertanggung jawab? Jika kita mengharapkan para pemangku kepentingan AI untuk menghormati dan menerapkan prinsip-prinsip transparansi dan keterjelasan, keadilan dan non-diskriminasi, keamanan dan keandalan, serta persaingan yang adil dan menghadapi konsekuensi ketika mereka gagal melakukannya, kita perlu memberikan informasi yang diperlukan kepada para pelaku terkait agar mereka dapat menghormati kewajiban mereka dan menavigasi dengan sukses dalam ekosistem baru ini.

Kerangka kerja kami mengidentifikasi tiga jenis pemangku kepentingan utama: individu, pelaku korporasi, dan badan pemerintah. Kerangka kerja ini juga mempertimbangkan tiga kategori utama kegiatan: pengembangan, penerapan, dan penggunaan sistem AI. Beberapa kewajiban etika dan hukum berlaku setara bagi ketiga pemangku kepentingan dan di semua kegiatan. Dalam kasus lain, seperti yang akan kita lihat, sifat dan hubungan pemangku kepentingan dengan sistem AI akan memerlukan beban tanggung jawab yang lebih tinggi atau lebih rendah. Tujuannya adalah untuk melindungi kepentingan pemegang hak yang terdampak oleh sistem AI secara memadai tanpa menyalahgunakan tanggung jawab. Luasnya kewajiban akuntabilitas juga akan bergantung pada pentingnya hak yang dipertaruhkan dan kemungkinan terjadinya kerugian, yang kami sebut pendekatan "skala geser".

Tentu saja, salah satu kesulitan yang nyata dalam mempersiapkan kerangka kerja akuntabilitas untuk teknologi baru seperti AI adalah kenyataan bahwa teknologi tersebut terus berkembang dan belum menunjukkan semua kemungkinan penerapannya. Akan menjadi kesalahpahaman jika menganggap bahwa solusi dan produk berbasis teknologi, seperti model AI, begitu canggih secara teknologi sehingga berada "di luar cakupan" kerangka kerja yang ada dan oleh karena itu harus tetap "tidak diatur" atau tidak diatur.

Merupakan tanggung jawab hukum para pemangku kepentingan AI untuk mengoperasikan sistem AI dengan cara yang selalu menghormati hukum yang berlaku baik hukum dasar, pidana, maupun perdata dan peraturan perundang-undangan. Sebagai contoh, gugatan kelalaian mengharuskan sistem AI digunakan sesuai dengan standar kehati-hatian tertentu. Contoh lain adalah kepatuhan pengguna sistem AI terhadap hukum perundang-undangan yang ditargetkan seperti pasal 6 Peraturan Perlindungan Data Umum Uni Eropa (GDPR), yang membatasi pemrosesan data hanya jika terdapat dasar hukum yang sah.



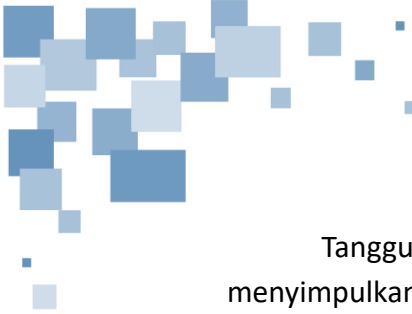
Jika pemangku kepentingan AI (baik badan hukum, pemerintah, dan perusahaan) tidak mematuhi hukum yang berlaku, mereka dapat dimintai pertanggungjawaban atas tindakan yang salah atau lalai terkait AI, sebagaimana halnya orang perseorangan. Sebagai konsekuensi dari kesimpulan kami tentang kepribadian hukum AI, model AI yang mendasari setiap layanan atau aktivitas berbasis teknologi tidak boleh dianggap independen secara hukum dari orang, perusahaan, atau pemerintah yang memfasilitasinya.

Mekanisme pertanggungjawaban juga berlaku jika pemangku kepentingan AI mengoperasikan sistem AI dan menyebabkan kerugian pada pihak lain. Salah satu pilar sistem hukum apa pun adalah kerangka pertanggungjawabannya karena kerangka ini menjamin kepercayaan masyarakat, dengan memastikan kompensasi dan keadilan atas kerugian yang ditimbulkan ketika seseorang tidak mematuhi hukum. Dengan AI, mekanisme ini akan memastikan bahwa pemangku kepentingan AI menanggung pertanggungjawaban hukum yang konsisten dengan alokasi pertanggungjawaban berdasarkan hukum yang berlaku. Di sebagian besar yurisdiksi, mekanisme pertanggungjawaban berbasis kesalahan didasarkan pada sejumlah komponen yang tidak terbukti dengan sendirinya. Artinya, sebagian besar komponen ini harus dibuktikan, terutama oleh penggugat, yaitu orang yang mengklaim cedera. Aturan prosedural dan pembuktian ini juga berlaku bagi penggugat yang menggugat pemangku kepentingan AI atas ganti rugi. Tanggung jawab perdata biasanya bertujuan untuk melindungi keselamatan pribadi dan kepentingan properti.

Berbeda dengan tanggung jawab pidana (dibahas di bawah), tanggung jawab perdata bertujuan untuk melindungi kepentingan pribadi korban dari perilaku tidak wajar pihak lain. Tanggung jawab perdata bertujuan untuk memastikan ganti rugi yang memadai setelah kerugian (sebagaimana dipahami oleh hukum perdata) yang berasal dari aktivitas pribadi, perusahaan, atau pemerintah. Tanggung jawab perdata mewujudkan prinsip-prinsip moral untuk masyarakat yang stabil ke dalam hukum.

Lebih lanjut, individu, perusahaan, dan pemerintah harus tetap bertanggung jawab atas setiap pelanggaran pidana atau pidana yang dilakukan melalui AI. Tanggung jawab pidana, berbeda dengan tanggung jawab perdata, bertujuan untuk melindungi kepentingan masyarakat secara keseluruhan, dengan mencegah dan menghukum perilaku yang merugikan kepentingan umum, membahayakan seluruh kelompok masyarakat, dan mengancam ketertiban umum. Untuk memastikan bahwa tujuan-tujuan tersebut terpenuhi secara memadai terkait AI, penting untuk menghindari pemisahan maksud antara pemungkin AI dan AI itu sendiri.

Menyerupai gagasan "maksud" atau "mens rea" yang umum digunakan dalam hukum pidana dan pidana, model AI yang sangat canggih dapat belajar secara mandiri, membuat keputusan yang belum tentu "dimaksudkan" oleh pengembang atau penggunanya. Namun, karena AI tidak memiliki status badan hukum, dan dengan demikian tidak dapat memiliki maksud, para pemangku kepentingan yang bertanggung jawab harus memikul tanggung jawab atas keputusan tersebut seolah-olah keputusan itu adalah keputusan mereka sendiri.



Tanggung jawab pidana untuk AI semacam itu mungkin tampak memberatkan namun, menyimpulkan sebaliknya akan memudahkan penghindaran tanggung jawab pidana atau pidana atas tindakan yang dilakukan di bawah pengawasan para pemangku kepentingan.

Terakhir, selain kewajiban konstitusional pemerintah, mungkin juga terdapat kompensasi atas pelanggaran hak-hak dasar yang dilakukan oleh para pemangku kepentingan AI. Di beberapa yurisdiksi, pelanggaran hak asasi manusia dapat menimbulkan, selain ganti rugi kompensasi, juga ganti rugi hukuman, yaitu yang bertujuan untuk menghukum pihak yang lalai, selain kewajiban untuk mengganti kerugian yang ditimbulkan. Ketika pelaku swasta melakukan aktivitas AI yang dapat merugikan hak asasi manusia, mereka juga harus menghadapi sanksi jika risikonya terwujud, terlebih lagi jika risiko tersebut merupakan konsekuensi dari tindakan yang disengaja.


Pertimbangan-pertimbangan ini menunjukkan bahwa AI memengaruhi hukum dan peraturan di setiap tingkatan dan bahwa akuntabilitas manusia memiliki banyak aspek. Di sisa bab ini, kami menyajikan gagasan umum bagi individu, perusahaan, dan pemerintah untuk mengatasi tantangan baru yang ditimbulkan oleh AI dengan lebih baik.

Jenis-jenis Pemangku Kepentingan

Dengan membedakan jenis-jenis pemangku kepentingan dan aktivitas tertentu, kami berharap dapat menghindari penerapan rezim akuntabilitas menyeluruh yang akan membebankan terlalu banyak tanggung jawab pada pemangku kepentingan tertentu dan tidak cukup pada yang lain. Misalnya, pengembang AI dapat dimintai pertanggungjawaban untuk memastikan tidak ada penyimpangan dari standar etika dalam algoritma sebelum penerapan atau penggunaan, sementara pemangku kepentingan yang menerapkan layanan AI atau memasarkan produk AI dapat dimintai pertanggungjawaban untuk menggunakan sistem AI guna menyediakan layanan yang sah dan sesuai hukum. Demikian pula, pengguna akhir harus memiliki kewajiban untuk menggunakan AI dengan cara yang tidak merugikan orang lain. Pemangku kepentingan yang memegang peran-peran ini diwajibkan untuk menghormati semua kewajiban yang terkait dengan aktivitas AI mereka.

Sebagai usulan umum, akuntabilitas harus dibagi di antara para pemangku kepentingan sesuai dengan tingkat kendali yang dimiliki pemangku kepentingan atas pengembangan, penerapan, penggunaan, dan evolusi sistem AI. Para pemangku kepentingan yang bertindak sebagai "penjaga" AI, dengan kata lain, yang memiliki akses ke data pelatihan dan memiliki kendali atas algoritma, seharusnya memiliki tanggung jawab yang lebih besar daripada pemangku kepentingan yang tidak dapat melatih ulang AI, misalnya ketika sistem AI berbasis perangkat lunak dan algoritma proprietary dijual kepada konsumen.

Hukum kontrak, yang akan berlaku untuk pengaturan kontraktual antara pengembang, deployer, dan pengguna AI, dapat digunakan, dalam batas yang dapat diterima, oleh para pemangku kepentingan untuk mengalokasikan tanggung jawab di antara mereka sendiri. Misalnya, seorang pemangku kepentingan dapat memiliki kewajiban kontraktual yang lebih tinggi terkait komponen produk/layanan yang dianggap penting, dan tentu saja, kontrak akan dihargai sesuai dengan itu. Sebaliknya, suatu kontrak dapat mengalihkan tanggung jawab atau



kewajiban kontraktual untuk meringankan beban salah satu pihak, misalnya pengembang yang tidak ingin melanjutkan pelatihan AI setelah penjualannya.

Akuntabilitas bagi Individu

Kami mengantisipasi bahwa akuntabilitas individu terkait AI akan berkaitan terutama dengan penggunaan sistem AI, alih-alih pengembangan atau penerapannya. Hal ini karena kebanyakan orang tidak tahu cara membangun jaringan saraf yang efisien dan kekurangan sumber daya untuk menerapkannya. Akibatnya, sebagian besar individu akan menjadi konsumen akhir atau pengguna produk atau layanan AI, misalnya dengan membeli kendaraan swakemudi, berinteraksi melalui chatbot yang diterapkan oleh perusahaan atau pemerintah, atau mengunduh aplikasi yang menggabungkan AI.


Dalam kasus pengembang individu dengan keterampilan dan pengetahuan yang relevan, mereka diharapkan untuk mematuhi hukum dan standar etika yang berlaku. Namun, mungkin tidak masuk akal untuk meminta pengembang individu yang bertindak sendiri atau yang merupakan bagian dari tim proyek dan tidak mengendalikan pengembangan sistem (jika mereka dapat diidentifikasi) untuk memberikan tingkat penilaian dan pemantauan risiko yang sama seperti yang diminta dari para pemangku kepentingan perusahaan dan pemerintah.

Kasus yang paling menarik adalah kasus di mana AI akan mentransformasi sistem yang ada. Kendaraan otonom, misalnya, memaksa kita untuk memikirkan kembali rezim pertanggungjawaban kita atas kecelakaan lalu lintas. Saat ini, pengemudi individu, kecuali dalam kasus cacat produksi, bertanggung jawab jika mereka menyebabkan kerusakan saat menggunakan mobil mereka. Pengenalan mobil self-driving yang sepenuhnya otonom melemahkan prinsip ini, meskipun beberapa undang-undang saat ini, seperti *Automated and Electric Vehicles Act* Inggris, membuat pemilik (atau perusahaan asuransinya) bertanggung jawab penuh atas kecelakaan. Mayoritas pemilik kendaraan tersebut tidak akan mengerti bagaimana kendaraan tersebut membuat keputusan.

Bagi sebagian orang, ini menunjukkan adanya pengalihan beban kecelakaan jalan raya ke pemangku kepentingan perusahaan yang mengembangkan dan menerapkan AI self-driving. Sebagian tanggung jawab dapat tetap berada di tangan pemilik dan individu yang menggunakan mobil tanpa pengemudi, misalnya jika produsen memberikan instruksi yang jelas tentang kapan tidak boleh menggunakan fungsi autopilot, meskipun hal ini berisiko membuat kendaraan otonom kurang menarik bagi konsumen.

Rezim pertanggungjawaban yang terlalu membebani produsen dapat mengurangi insentif untuk memasarkan teknologi AI yang diharapkan dapat membantu mencegah kematian dan cedera akibat kecelakaan lalu lintas. Di sisi lain, konsumen mungkin enggan membeli kendaraan otonom jika mereka bertanggung jawab atas perilaku algoritma yang berada di luar kendali mereka. Menemukan keseimbangan yang tepat memang akan tetap menjadi tantangan.

Para legislator juga dapat membatasi kemungkinan pengembang dan pengguna AI untuk mengalihkan kewajiban akuntabilitas mereka, terutama jika hal ini akan membebaskan kewajiban mereka kepada pengguna akhir individu yang kurang memiliki pengetahuan tentang AI atau kemampuan untuk memengaruhi hasil. Namun, ketika individu menggunakan sistem



AI secara lalai atau sengaja untuk tujuan jahat, mereka harus dimintai pertanggungjawaban. Seorang individu yang memproduksi konten yang mencemarkan nama baik dengan perangkat lunak pembelajaran mesin yang dijual untuk memodifikasi gambar, misalnya, harus dimintai pertanggungjawaban atas tindakannya, dan bukan pengembang pihak ketiga yang tidak bersalah.


Akuntabilitas bagi Aktor Korporasi Swasta

Korporasi swasta telah menjadi yang terdepan dalam kemajuan terkini di bidang AI dan diharapkan untuk mempertahankan posisi sentral mereka dalam pengembangan, penerapan, dan penggunaan AI. Lebih lanjut, banyak perusahaan yang berspesialisasi dalam AI akan memiliki kemampuan untuk menggabungkan kegiatan-kegiatan ini. Namun, AI tidak hanya menguntungkan perusahaan teknologi, sehingga rezim akuntabilitas kami berlaku untuk semua aktor swasta yang memiliki kegiatan AI, baik yang dianggap sebagai perusahaan teknologi maupun tidak. Misalnya, sebuah perusahaan konsultan yang membeli perangkat lunak AI untuk membantu memfasilitasi proses rekrutmennya harus bertanggung jawab atas penggunaan perangkat lunak ini dan memastikan, misalnya, bahwa perusahaan tersebut tidak merekomendasikan penolakan kandidat untuk pekerjaan dengan alasan yang dilarang (yaitu, diskriminatif).

Personal Information Protection and Electronic Documents Act Kanada (PIPEDA) menetapkan prinsip-prinsip terkait perlindungan informasi pribadi, termasuk akuntabilitas. Ini adalah contoh undang-undang yang menetapkan kerangka kerja akuntabilitas bagi perusahaan yang kami anggap memberikan sumber inspirasi yang bermanfaat untuk kerangka kerja akuntabilitas yang berlaku untuk AI. Berdasarkan PIPEDA, suatu organisasi bertanggung jawab atas informasi pribadi yang berada dalam kendalinya dan harus menunjuk individu untuk memastikan kepatuhan terhadap sepuluh prinsip informasi yang adil yang tercantum dalam lampiran undang-undang tersebut. Identitas individu-individu ini harus diketahui jika diminta.

Organisasi juga berkewajiban untuk menerapkan kebijakan dan praktik guna mewujudkan prinsip-prinsip ini. Ini termasuk menetapkan prosedur untuk menerima dan menanggapi keluhan dan pertanyaan terkait akses dan pemrosesan tersebut. Ini juga akan mencakup pelatihan staf dan mengomunikasikan informasi kepada mereka tentang kebijakan dan praktik tersebut, serta mengembangkan informasi untuk menjelaskannya. Kami berpendapat bahwa kerangka kerja akuntabilitas serupa harus diterapkan pada AI dan merupakan bagian dari tanggung jawab pemerintah untuk mengembangkan kerangka kerja tersebut.

Para legislator AS juga telah mengusulkan *Algorithmic Accountability Act*, yang memungkinkan *Federal Trade Commission* (FTC) untuk membuat peraturan yang mewajibkan "entitas yang tercakup" untuk melakukan "penilaian dampak pengambilan keputusan otomatis" dan "penilaian dampak perlindungan data" untuk sistem keputusan otomatis berisiko tinggi. "Entitas yang tercakup" mencakup setiap orang, kemitraan, atau perusahaan yang berada di bawah yurisdiksi FTC, yang merupakan pialang data atau entitas komersial lain yang, sebagai bagian substansial dari bisnisnya, mengumpulkan, menyusun, atau memelihara informasi pribadi mengenai individu yang bukan pelanggan atau karyawan entitas tersebut

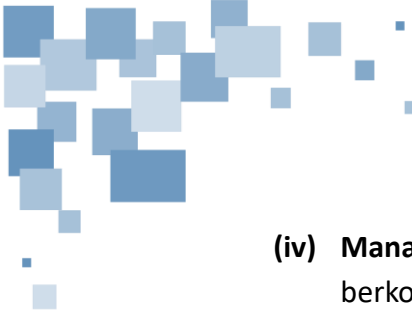


untuk menjual atau memperdagangkan informasi tersebut atau menyediakan akses pihak ketiga ke informasi tersebut.

"Penilaian dampak sistem keputusan otomatis" akan mengevaluasi sistem keputusan otomatis dan proses pengembangannya untuk dampak pada akurasi, keadilan, bias, diskriminasi, privasi, dan keamanan. "Penilaian dampak perlindungan data" adalah studi yang mengevaluasi sejauh mana sistem informasi melindungi privasi dan keamanan informasi pribadi yang diproses oleh sistem tersebut. Rancangan undang-undang ini juga memberikan kewenangan investigasi dan penegakan hukum bagi FTC dan otoritas negara terkait peraturan yang diundangkan di bawahnya.

Kerangka Kerja Tata Kelola Kecerdasan Buatan Model (Kerangka Kerja Model SG) yang diusulkan Singapura menetapkan prinsip dan langkah-langkah tertentu yang didorong untuk diadopsi oleh organisasi dalam penggunaan AI yang bertanggung jawab. Kerangka Kerja Model SG dipandu oleh dua prinsip umum: (a) organisasi harus memastikan bahwa proses pengambilan keputusan dapat dijelaskan, transparan, dan adil; dan, (b) solusi AI harus berpusat pada manusia. Kerangka kerja ini terutama berfokus pada empat area umum:

- (i) **Struktur dan langkah-langkah tata kelola internal:** Peran dan tanggung jawab yang jelas harus ditetapkan dalam organisasi untuk penerapan AI yang etis, dan sistem manajemen risiko serta pengendalian internal harus diterapkan;
- (ii) **Menentukan model pengambilan keputusan AI:** Organisasi harus menetapkan selera risiko mereka untuk penggunaan AI, yaitu, menentukan risiko yang dapat diterima dan mengidentifikasi model pengambilan keputusan yang tepat untuk mengimplementasikan AI. Hal ini dapat dilakukan dengan mempertimbangkan tujuan komersial mereka terhadap risiko penggunaan AI dalam pengambilan keputusan mereka, yang dipandu oleh nilai-nilai perusahaan dan norma-norma sosial mereka. Terdapat tiga model pengambilan keputusan yang diidentifikasi dalam Kerangka Model SG, dengan tingkat pengawasan manusia yang bervariasi (a) di mana pengawasan manusia aktif dan terlibat, serta manusia memiliki kendali penuh dan AI hanya memberikan rekomendasi atau masukan; (b) di mana tidak ada pengawasan manusia atas pelaksanaan keputusan, serta AI memiliki kendali penuh tanpa opsi penggantian manusia; dan, (c) di mana manusia dapat menyesuaikan parameter selama pelaksanaan algoritma;
- (iii) **Manajemen Operasi:** Organisasi harus menerapkan praktik-praktik tertentu saat mengembangkan, memilih, dan memelihara model AI, serta memperhatikan interaksi antara data dan algoritma atau model. Kerangka Model SG merekomendasikan praktik akuntabilitas data yang baik seperti memahami silsilah data, memastikan kualitas data, meminimalkan bias inheren, menggunakan set data yang berbeda untuk pelatihan, pengujian, dan validasi, serta peninjauan dan pembaruan set data secara berkala. Kerangka ini juga memiliki rekomendasi terkait algoritma/model, terkait dengan keterjelasan, pengulangan hasil, ketertelusuran, penyetelan berkala, serta pemantauan dan peninjauan aktif;

- 
- (iv) **Manajemen hubungan pelanggan:** Organisasi harus menerapkan strategi untuk berkomunikasi secara efektif dengan pelanggan dan konsumen saat menerapkan AI. Hal ini dapat mencakup pengungkapan umum, peningkatan transparansi, evaluasi etika, kebijakan penjelasan kepada individu, antarmuka pengguna yang sesuai dengan tujuan, opsi untuk berhenti berlangganan, saluran komunikasi untuk umpan balik, dan peninjauan keputusan material terkait AI.

Sebagai bagian dari tata kelola perusahaan yang efektif, pelaku swasta juga harus menilai delapan prinsip yang ditetapkan dalam makalah ini dan menerapkannya sebagai pedoman internal untuk mengarahkan kegiatan mereka agar tetap menghormati kewajiban hukum mereka. Hal ini akan dibahas lebih lanjut di bawah ini.


Akuntabilitas bagi Pemerintah

Pemerintah harus bertanggung jawab atas penghormatan terhadap kewajiban konstitusional dan internasional mereka saat mereka berupaya mengembangkan, menerapkan, dan menggunakan sistem AI. Mereka juga bertanggung jawab secara khusus atas penerapan kerangka hukum dan peraturan yang tepat yang memberi tahu pemangku kepentingan lain apa yang boleh dan tidak boleh mereka lakukan dengan AI, dan siapa yang harus bertanggung jawab jika terjadi kesalahan.

Hingga pemerintah mengembangkan kapasitas internal untuk menghasilkan sistem AI yang sebanding dengan yang dikembangkan oleh perusahaan teknologi besar, kemungkinan besar otoritas sektor publik pada dasarnya akan tetap menjadi pengguna sistem AI yang dikembangkan secara eksternal. Namun, penggunaan AI di sektor publik dapat memberikan manfaat yang substansial. Misalnya, pemerintah daerah dapat menerapkan AI untuk mengurangi kemacetan lalu lintas dengan menganalisis masalah infrastruktur dan pola penerangan lalu lintas untuk menciptakan lampu jalan cerdas. Mereka juga dapat menggunakan sistem AI untuk mengidentifikasi bangunan yang berisiko kebakaran atau memanfaatkan penggunaan AI bahasa alami untuk menciptakan chatbot guna menanggapi permintaan warga.

Dalam kasus tertentu, penggunaan sistem AI oleh lembaga publik akan tunduk pada standar dan kewajiban akuntabilitas yang sama dengan pihak swasta. Namun, dalam kasus lain, terutama yang melibatkan lembaga kepolisian, peradilan, administratif, atau militer, penggunaan sistem AI oleh pemerintah harus tunduk pada standar perilaku yang lebih tinggi, untuk memperhitungkan pentingnya hak-hak yang terlibat dan risiko pelanggaran hak-hak tersebut. Khususnya, pemerintah harus secara khusus memperhatikan kewajiban konstitusional dan internasional mereka, termasuk penghormatan terhadap hak asasi manusia yang fundamental.

Akuntabilitas pemerintah atas penggunaan AI juga dapat menjadi bagian dari perjanjian multilateral. Wajar bagi warga negara untuk mengharapkan standar akuntabilitas yang lebih tinggi dari pemerintah dibandingkan dari aktor non-negara mengingat hubungan fiducia yang dimiliki warga negara dengan pemerintah mereka dan hak-hak fundamental yang dijamin secara konstitusional bagi warga negara. Pengadilan nasional harus memantau penggunaan sistem AI oleh badan publik untuk memastikan penghormatan terhadap prinsip-



prinsip konstitusional dan supremasi hukum. Namun, standar yang lebih tinggi ini dapat diatur dan ditegakkan melalui perjanjian internasional multilateral sehingga terdapat kesepakatan global tentang bagaimana sistem AI dapat dan tidak dapat digunakan oleh pemerintah, misalnya dalam bidang yang memengaruhi hak asasi manusia.

AI dan Proses Hukum dalam Keputusan Administratif

Meskipun standar proses hukum (atau kewajiban untuk bertindak adil) dapat bervariasi di setiap yurisdiksi, tujuannya adalah untuk mewajibkan Negara menghormati hak-hak hukum individu. Hal ini mensyaratkan bahwa tidak ada warga negara yang ditolak hak-hak hukumnya kecuali sesuai dengan keadilan substantif dan prosedural.

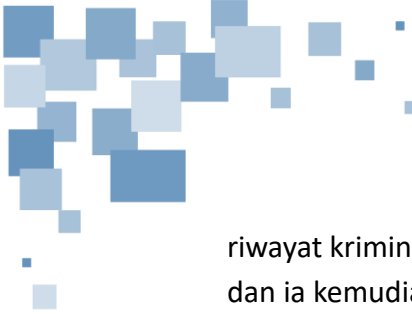
Standar proses hukum juga berlaku untuk penggunaan AI oleh pemerintah. Ketika menggunakan sistem AI untuk membuat keputusan administratif atau yudisial, terutama dalam konteks yang melibatkan hak-hak fundamental, pemerintah harus bertanggung jawab atas keadilan algoritmanya. Sistem AI hanya sebaik kumpulan data yang digunakan untuk melatihnya, dan pasti akan mencerminkan bias apa pun dalam data tersebut. Oleh karena itu, ketika data kurang mewakili (atau terlalu mewakili) satu kelompok, hal itu dapat menyebabkan hasil yang tidak dapat diandalkan yang lebih memihak satu kelompok daripada kelompok lain dengan cara yang secara nyata melanggar hak atas perlakuan yang adil dan kesetaraan.

Tanpa akuntabilitas, terutama untuk transparansi dan keterjelasan atas keputusan tersebut, para pemohon yang kecewa tidak akan memiliki dasar untuk menggugat keputusan tersebut. Jika pemerintah benar-benar menghormati proses hukum yang semestinya dalam penerapan AI, penting bagi mereka untuk menetapkan pedoman yang jelas untuk peninjauan efektif atas keputusan administratif dan yudisial yang dibuat, baik sebagian maupun seluruhnya, menggunakan sistem AI. Pemerintah tidak boleh menghindari kewajiban konstitusional mereka, yang secara khusus dinyatakan dalam undang-undang peninjauan yudisial atas keputusan administratif, dengan bersembunyi di balik algoritma kepemilikan yang dikembangkan secara internal oleh para pemangku kepentingan perusahaan.

AI dalam Sistem Peradilan

Sistem peradilan pidana adalah area di mana penggunaan AI oleh otoritas pemerintah dapat berdampak signifikan terhadap hak-hak seseorang dan ini adalah area di mana pemerintah memiliki peran yang unik. Dalam beberapa dekade terakhir, Pengadilan mulai mengandalkan alat bantu seperti pedoman penjatuhan hukuman (baik yang bersifat nasihat maupun wajib) dan penilaian independen dalam menentukan kemungkinan residivisme terdakwa ketika menentukan hukuman yang tepat. Salah satu alat bantu ini menjadi subjek banding yang diajukan oleh terdakwa dalam kasus *State v. Loomis* di AS untuk menentang penggunaan prediktor residivisme AI oleh hakim sebagai dasar keputusannya untuk menambah lama hukuman.

Dalam kasus *Loomis*, terdakwa telah mengaku bersalah atas kejahatan tertentu. Selama pembacaan hukuman, hakim menggunakan sistem AI yang disediakan oleh *Northpointe Institute for Public Management* yang disebut COMPAS. Melalui wawancara dengan *Loomis*, yang mencakup kuesioner berisi sekitar 137 pertanyaan, dan informasi dari



riwayat kriminalnya, COMPAS sampai pada kesimpulan bahwa risiko residivisme Loomis tinggi dan ia kemudian dijatuhi hukuman enam tahun penjara.

Karena sifat kepemilikan teknologi ini, pengujian independen masih jarang dilakukan. Namun, beberapa organisasi telah berupaya menguji "prediktor" residivisme ini dan menemukan bahwa, rata-rata, terdakwa berkulit hitam lebih mungkin memiliki risiko residivisme yang lebih tinggi daripada terdakwa berkulit putih.

Sebagaimana dibahas dalam pembahasan Prinsip 4 Keadilan dan Non-Diskriminasi, hal ini telah menyebabkan situasi di mana salah satu prediktor ini menentukan bahwa seorang pria kulit hitam berusia 18 tahun yang mencuri sepeda senilai Rp 800.000 sangat mungkin untuk mengulangi tindak pidananya, sementara seorang pria kulit putih berusia 41 tahun dengan riwayat kriminal yang panjang yang mencuri properti senilai Rp 860.000 memiliki risiko residivisme yang rendah.


Selain risiko bias berdasarkan kelas yang dilindungi ini, model pembelajaran mesin juga dapat menemukan korelasi antara, misalnya, apakah orang tua terdakwa dipenjara, atau apakah terdakwa mengenal banyak orang yang menggunakan narkoba ilegal. Korelasi ini bahkan mungkin merupakan prediktor residivisme yang akurat. Namun, korelasi ini berada di luar kendali terdakwa. Alih-alih menilai kemungkinan residivisme berdasarkan keputusan dan tindakan terdakwa, AI melihat korelasi yang berkaitan dengan siapa terdakwanya.

Tampaknya anak yang dibesarkan dalam kemiskinan lebih mungkin memiliki orang tua yang pernah dipenjara dan mengenal sejumlah orang yang telah menggunakan narkoba ilegal daripada anak yang dibesarkan dalam keluarga kaya. Jika hakim mengandalkan solusi AI untuk meningkatkan hukuman terdakwa karena faktor-faktor di luar kendali mereka, dan itu mungkin terkait dengan kemiskinan, kita berpotensi menciptakan sistem peradilan dua tingkat, di mana terdakwa miskin dihukum lebih keras daripada terdakwa kaya, berdasarkan faktor-faktor yang terkait dengan kemiskinan mereka.

Oleh karena itu, penggunaan sistem AI oleh pengadilan harus dipantau secara ketat dan hanya algoritma yang telah teruji secara menyeluruh dengan bukti kinerja yang dapat dibuktikan yang boleh membantu hakim dalam membuat keputusan penting. Sebagaimana yang akan kita bahas dalam pembahasan Prinsip 3 Transparansi dan Keterjelasan, transparansi sistem ini sangatlah penting. Lebih lanjut, setiap implementasi sistem pengambilan keputusan atau rekomendasi yang akan berdampak pada masyarakat harus memungkinkan peninjauan yang bermakna terhadap rekomendasi AI oleh pengadilan banding. Tanpa transparansi, akuntabilitas, dan pengawasan yang memadai, terdapat risiko yang cukup besar bahwa sistem AI akan memperkenalkan dan memperkuat praktik yang tidak adil dan sewenang-wenang dalam sistem peradilan.

Akuntabilitas Regulasi dan Legislatif

Pemerintah juga harus bertanggung jawab kepada konstituennya atas pengesahan undang-undang dan regulasi yang dirancang dengan baik dan efektif tentang AI. Dengan demikian, pemerintah tidak hanya pengembang, pelaksana, dan pengguna sistem AI, tetapi juga regulator utama aktivitas para pemangku kepentingan lainnya. Dalam *Ethically Aligned Design, First Edition*, IEEE menulis bahwa "satu sistem hukum yang efektif berkontribusi pada



kesejahteraan manusia.” Pemerintah, sebagai penjaga utama sistem hukum kita, kemudian harus bertanggung jawab untuk memberlakukan undang-undang dan aturan yang tepat secara teknologi dan seimbang yang mengatur aktivitas AI para pemangku kepentingan.


Ini merupakan peran penting untuk memastikan bahwa individu dan perusahaan memiliki pemahaman yang jelas tentang kewajiban mereka terhadap para pemangku kepentingan lain dan masyarakat secara keseluruhan. Prediktabilitas dan kejelasan akan memastikan para pemangku kepentingan yakin bahwa aktivitas mereka sah, yang akan menguntungkan pengembangan teknologi AI. Singkatnya, pemerintah bertanggung jawab untuk memberlakukan undang-undang AI yang mendukung pengembangan dan penggunaan teknologi ini yang bermanfaat.

Saat ini, kami tidak yakin pemerintah perlu menyusun "hukum AI" yang menyeluruh. Regulator dan badan pengatur yang ada juga tampaknya berada di posisi terbaik untuk membuat amandemen terhadap regulasi atau menciptakan standar di bidangnya masing-masing, alih-alih membentuk badan otoritatif menyeluruh untuk "semua AI", yang, meskipun menarik secara teori, pasti akan sulit diterapkan, mengingat beragamnya bidang yang dicakup AI. Beberapa bidang di mana AI saat ini sedang dikembangkan dan digunakan sudah diatur secara ketat. Misalnya, hal ini terjadi di bidang keuangan dan transportasi udara. Hal ini juga terjadi pada layanan profesional, seperti layanan medis dan hukum, di mana seringkali aktivitasnya hanya dapat dilakukan oleh pihak yang terdaftar atau berlisensi. Regulasi AI dalam industri ini harus dibangun di atas kerangka kerja yang ada, yang hanya boleh diubah jika diperlukan.

Beberapa regulator di bidang ini telah menyatakan bahwa aturan mereka "netral terhadap teknologi". Ini berarti bahwa aturan tersebut "buta" terhadap teknologi. Dalam kasus seperti itu, orang dan perusahaan yang diatur bertanggung jawab atas kepatuhan terhadap peraturan dan standar yang berlaku, terlepas dari apakah mereka menggunakan AI atau teknologi lain untuk menyediakan layanan atau melakukan aktivitas yang diatur. Untuk membantu menavigasi lingkungan regulasi, beberapa regulator telah menciptakan sandbox yang mungkin berguna untuk mengembangkan sistem AI yang inovatif dan untuk menilai cara mematuhi aturan yang berlaku di ruang yang diatur. Pemerintah harus memastikan regulator yang ada mempelajari dampak AI di sektor masing-masing dan membuat usulan tentang cara mengadaptasi aturan yang ada untuk AI.

Ada kemungkinan bahwa dalam beberapa kasus, regulasi khusus AI mungkin diperlukan untuk sektor tertentu guna memastikan akuntabilitas pemangku kepentingan. Mengemudi otonom adalah contoh area di mana perubahan penting perlu dilakukan pada rezim akuntabilitas saat ini. Sebagaimana dinyatakan dalam ikhtisar di atas, regulator dan otoritas pengatur di bidang tertentu mungkin ingin merefleksikan dan menilai apakah aturan baru diperlukan untuk memitigasi risiko yang muncul terkait AI. Pemerintah mungkin harus mempelajari dan menentukan model liabilitas mana yang paling relevan untuk aplikasi AI tertentu.

Baik sebagai tambahan atau sebagai pengganti tanggung jawab berbasis kesalahan, beberapa bidang tertentu mungkin menjamin, misalnya, asuransi tanpa kesalahan atau jenis



mekanisme kompensasi moneter lainnya diberlakukan melalui undang-undang. Sekali lagi, hal ini mungkin sesuai dalam kasus kecelakaan mobil. Pada titik tertentu, ketika AI semakin canggih, regulator dan otoritas pengatur mungkin juga ingin mempertimbangkan seperangkat standar AI yang disepakati di bidangnya masing-masing. Mungkin titik awal refleksi adalah menilai apakah aturan tersebut anti-AI, baik secara tersurat maupun tersirat.

Singkatnya, meskipun pihak swasta untuk sementara waktu akan tetap memimpin pengembangan AI sebagai teknologi, pemerintah memiliki peran penting dalam merancang aturan yang akan mendorong pengembangan AI sebagai manfaat sosial, sebagaimana yang akan kita bahas secara lebih rinci dalam pembahasan Prinsip 6 Data Terbuka dan Persaingan yang Adil. Dalam peran tersebut, pemerintah tidak hanya bertanggung jawab kepada pemegang saham atau konsumen, tetapi juga kepada masyarakat secara keseluruhan.

Jenis Kegiatan

Kewajiban akuntabilitas juga akan bervariasi fungsinya, tergantung pada jenis kegiatan yang dilakukan oleh para pemangku kepentingan. Hal ini penting karena kita perlu memperhitungkan beragam peran yang akan dimainkan oleh para pemangku kepentingan selama masa pakai sistem AI. Jika kita mengambil contoh chatbot AI, pengembangnya adalah perusahaan yang benar-benar menciptakan perangkat lunak tersebut (mengumpulkan data, menulis kode, serta melatih dan menguji algoritmanya). Pihak yang menerapkannya adalah bank yang membeli chatbot dari pengembang.

Penggunanya adalah klien yang berinteraksi dengannya, setelah diterapkan di situs webnya oleh bank. Satu orang dapat memiliki beberapa peran, misalnya jika Bank mempekerjakan spesialis AI untuk membuat chatbot, dan jika menggunakan chatbot untuk tujuan internal seperti membantu pengambilan keputusan.

Pengembang harus bertanggung jawab atas pengembangan dan perancangan sistem AI yang sah dan etis. Namun, karena banyak sistem AI dapat dilatih ulang dan dapat berkembang dengan penambahan atau pemrosesan data baru, para pemangku kepentingan yang menerapkan atau menggunakan sistem AI (meskipun mereka awalnya tidak mengembangkannya) harus bertanggung jawab atas modifikasi tersebut. Di bagian ini, kami akan mengembangkan lebih lanjut kerangka kerja akuntabilitas kami terkait aktivitas AI tertentu.

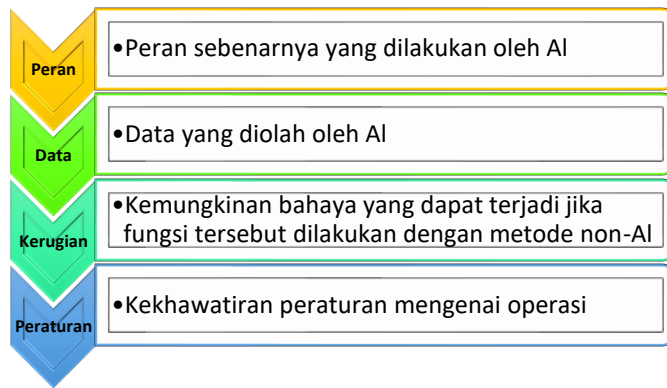
Pengembangan Sistem AI

Aktivitas pengembangan AI mengacu pada proses pembuatan algoritma cerdas. Dalam kasus pembelajaran mesin dan pembelajaran mendalam, hal ini mencakup pengadaan data yang relevan, penulisan kode, pelatihan sistem, serta evaluasi dan pengujiannya.

Saat ini, pengembang teknologi seperti robotika hanya bertanggung jawab jika mereka lalai atau dapat memperkirakan potensi bahaya pada saat mengembangkan teknologi tersebut. Kita dapat memandang hal ini sebagai "kewajiban negatif" untuk menghindari bahaya, alih-alih sebagai kewajiban positif untuk mengembangkan produk yang aman atau etis. Dalam kasus *Jones v. W+M Automation Inc.*, sebuah sistem pemuatan gantry robotik melukai penggugat saat ia bekerja sebagai pengatur alat di pabrik General Motors. Meskipun sistem tersebut dirancang oleh W+M Automation Inc., General Motors telah melanggar

peraturan keselamatan dan kesehatan kerja pada saat pemasangan. Pengadilan memutuskan bahwa, antara lain, para responden tidak bertanggung jawab karena mereka telah mematuhi semua peraturan yang berlaku saat memproduksi sistem pemuatan gantry robotik.

Faktor-Faktor yang Mempengaruhi Kerugian



Jika prinsip ini diterapkan dalam kasus yang melibatkan pengembang AI, kerugian dapat diperkirakan dan ditentukan oleh berbagai faktor, seperti: (i) peran aktual yang dijalankan oleh AI; (ii) data yang diproses oleh AI; (iii) kemungkinan kerugian yang dapat ditimbulkan jika fungsi tersebut dijalankan dengan

metode non-AI; dan, (iv) kekhawatiran regulasi terkait operasional.

Untuk menilai kerugian ini dan menguji aplikasi AI, pengembang harus melakukan studi penilaian dampak AI. Kewajiban untuk melakukan audit penilaian dampak pihak ketiga secara berkala juga harus dipertimbangkan, terutama untuk sistem AI berisiko tinggi. Audit semacam itu juga akan memastikan bahwa sistem AI membaik seiring waktu dan bahwa setiap kekhawatiran diperbaiki secara preemtif, alih-alih diperbaiki setelah kerugian terjadi.

IBM telah mengambil posisi bahwa orang-orang yang terlibat dalam penciptaan sistem AI harus bertanggung jawab di setiap langkah untuk mempertimbangkan dampak sistem terhadap dunia. Karena peran pengembang terbatas pada pengembangan sistem AI, akuntabilitas mereka harus dibatasi pada apa yang dapat mereka kendalikan dalam peran ini tanpa, dalam banyak kasus, meluas ke bagaimana sistem tersebut digunakan. Tanggung jawab utama pengembang haruslah mematuhi Prinsip-Prinsip Etika dalam cara pengembangan sistem AI untuk memastikan bahwa sistem yang baik dan patuh memasuki pasar.

Sebagaimana disarankan di atas, ini mungkin hanya menyiratkan gagasan klasik tentang tanggung jawab produk, tetapi sistem AI memiliki karakteristik tertentu yang perlu dipertimbangkan. Beberapa konteks mungkin mengharuskan pengembang AI untuk menyimpan catatan data yang digunakan untuk melatih AI, misalnya dalam kasus di mana perlu untuk memastikan tidak ada bias dalam kumpulan data. Pengembang AI mungkin juga menjadi satu-satunya yang memiliki pengetahuan yang diperlukan untuk memodifikasi atau melatih ulang suatu algoritma, menciptakan kewajiban akuntabilitas yang melampaui tahap pengembangan itu sendiri.

Konteks tertentu mungkin mengharuskan pengembang untuk memasukkan langkah-langkah teknis khusus ke dalam sistem mereka.

Rekomendasi

Untuk memastikan sistem AI dapat dikontrol jika terjadi malfungsi atau 'bermasalah', terutama robot AI, kami menyarankan pengembang untuk membangun 'kill-switch' dan memungkinkan penggantian manual yang mudah untuk melindungi dari konsekuensi yang tidak diinginkan.


Fungsi penghentian instan sangat penting karena sistem pembelajaran penguatan dapat menemukan cara untuk memutus operator dari lingkaran. Fitur-fitur ini akan sangat penting dalam sistem AI berisiko tinggi di mana konsekuensi dari sistem yang terganggu atau (sistem) malfungsi dapat sangat mengerikan, seperti kendaraan tanpa pengemudi atau pesawat terbang. Beberapa akademisi bahkan berpendapat bahwa pengembang/produsen harus menyediakan "akuntabilitas tingkat programatik" dengan mampu membuktikan mengapa suatu sistem beroperasi dengan cara tertentu, untuk mengatasi masalah hukum dan menentukan kesalahan. Langkah-langkah yang memastikan akuntabilitas programatik/prosedural tersebut dapat diamanatkan bagi pengembang yang terlibat dalam pengembangan sistem AI berisiko tinggi.

Akuntabilitas pengembang dapat ditegakkan melalui sanksi jika mereka gagal mematuhi delapan prinsip yang dibahas dalam Kerangka Kerja AI yang Bertanggung Jawab ini sebagaimana diadopsi di yurisdiksi tertentu dan jika hal tersebut menyebabkan kerugian. Sifat dan kuantitas sanksi dapat didasarkan pada faktor-faktor seperti niat, tingkat kerugian yang ditimbulkan, dan tindakan yang diambil untuk memitigasi kerugian tersebut, jika ada. Namun, akuntabilitas dan liabilitas pengembang tidak boleh mencakup keadaan di mana pengembang telah mengambil semua tindakan pencegahan yang dapat diperkirakan secara wajar saat mengembangkan sistem AI dan di mana sistem tersebut diterapkan dan digunakan dengan cara yang tidak dimaksudkan oleh pengembang, dengan konsekuensi yang tidak dapat diprediksi secara wajar. Akuntabilitas pengembang harus didasarkan pada 'kesalahan' dalam pengembangan sistem dan tidak boleh mencakup kasus-kasus di mana kerugian terjadi meskipun pengembang tidak melanggar hukum yang berlaku.

Rezim akuntabilitas yang memberikan sanksi kepada pengembang tanpa kesalahan mungkin akan memberikan efek yang menghambat kemajuan dan merugikan inovasi. Memiliki kerangka kebijakan yang fleksibel yang memberikan perlindungan kepada pengembang akan menciptakan keseimbangan antara pengembangan AI dan perlindungan kepentingan pemegang hak.

Penerapan dan Penggunaan Produk dan Layanan AI

Penerapan dan penggunaan sistem AI melibatkan penyediaan sistem tersebut kepada pengguna akhir, baik sebagai layanan maupun produk, tanpa pemangku kepentingan yang menerapkan AI harus mengembangkan algoritmanya sendiri. Dalam kebanyakan kasus, hubungan antara penyedia dan pengguna akhir atau klien, baik perusahaan maupun individu, bersifat kontraktual dan kontrak tersebut dapat membatasi akuntabilitas penyedia melalui klausul seperti ganti rugi dan pembatasan tanggung jawab. Dengan demikian, kontrak antara berbagai entitas akan menentukan akuntabilitas di antara mereka sendiri atas kerugian yang



disebabkan oleh penyedia atau yang timbul dari aktivitas penyedia, kecuali dalam keadaan yang dilarang oleh hukum.

Pada tingkat kebijakan, pemangku kepentingan yang menerapkan sistem AI harus bertanggung jawab atas produk dan layanan yang mereka sediakan melalui sistem AI dan mereka harus menyediakan produk dan layanan yang sah dan menghormati Kerangka Kerja untuk AI yang Bertanggung Jawab. Menyediakan layanan yang mendiskriminasi pengguna akhir tertentu atau yang menyalahgunakan data pribadi untuk tujuan yang tidak sah akan melanggar kewajiban penyedia layanan kepada pengguna akhir. Telah disebutkan sebelumnya bahwa pemerintah memiliki kewajiban konstitusional tertentu untuk memperlakukan warga negaranya secara adil dan tanpa diskriminasi, tetapi pihak swasta juga dapat memiliki kewajiban serupa terhadap pengguna atau klien mereka, tergantung pada hukum yang berlaku.


Standar akuntabilitas di sini akan bergantung pada sifat layanan, apakah layanan tersebut diatur, apakah kerugian nyata terjadi, pentingnya hak yang dipertaruhkan, dll. Standar ini juga akan didorong oleh hubungan kontraktual antara para pihak. Sebagai bagian dari kewajiban akuntabilitas, selain untuk memastikan penyedia layanan mampu menilai secara akurat risiko kerugian yang timbul dari produk atau layanan yang akan disediakan, penyedia layanan juga harus bertanggung jawab untuk melakukan penilaian dampak kerugian yang mungkin ditimbulkan jika AI diimplementasikan terkait dengan layanan tertentu.

Terkadang, penyedia layanan juga akan menjadi pengembang AI, yang seharusnya menyederhanakan analisis. Ketika keduanya terpisah, akuntabilitas seharusnya bergantung pada siapa yang bertanggung jawab atas bagian-bagian tertentu dari siklus hidup AI, misalnya pengumpulan data, pembaruan algoritma dengan data baru, pengujian, pemantauan keputusan AI, dll. Jika seorang deployer menyalahgunakan sistem AI dengan tidak melakukan peninjauan berkala terhadap keputusan sistem yang mengakibatkan hasil yang merugikan bagi pengguna akhir, maka deployer tersebut harus bertanggung jawab. Namun, jika sistem AI membuat keputusan tersebut dan alasan keputusan tersebut terkait dengan bagaimana data pelatihan awal dipilih, maka pengembang harus bertanggung jawab. Singkatnya, deployer atau pengguna harus dilindungi dari sistem yang salah yang dibuat oleh pengembang, tetapi, seperti yang telah kami katakan, pengembang tidak boleh bertanggung jawab atas deployment atau penggunaan kreasi mereka yang tidak dapat diprediksi atau tidak diizinkan. Tentu saja, dalam kasus ambiguitas, kewajiban akuntabilitas juga dapat diperjelas melalui kontrak atau ditetapkan oleh peraturan perundang-undangan.

Pendekatan Skala Geser

Kita telah melihat bagaimana sifat pemangku kepentingan dan aktivitas mereka akan memengaruhi intensitas kewajiban akuntabilitas mereka. Kepentingan pemegang hak seperti konsumen merupakan salah satu elemen penting lainnya. Saat menggunakan sistem AI atau menjadi sasaran keputusan/tindakan yang diambil oleh sistem AI, orang-orang yang terdampak harus dilindungi hak dan kepentingannya.

Beberapa hak dan jaminan ini sudah ada. Misalnya, hak-hak yang berkaitan dengan data pribadi dapat ditemukan dalam undang-undang perlindungan data yang berlaku di



seluruh dunia, seperti hak untuk memberikan atau mencabut persetujuan, untuk diberitahu tentang bagaimana data pribadi seseorang akan digunakan dan kepada siapa data tersebut akan ditransfer. Hak-hak lain mungkin termasuk hak untuk mengajukan banding atas keputusan administratif atau yudisial yang dibuat dengan bantuan algoritma AI jika kebebasan individu dipertaruhkan, seperti dalam konteks pidana.

Aturan umumnya adalah membuat ketegasan kewajiban akuntabilitas proporsional dengan pentingnya hak-hak yang dipertaruhkan dan dengan tingkat keparahan dampak sistem AI terhadap hak-hak ini dan kemungkinan kerugiannya. Sebagai contoh, pengembangan dan penerapan sistem dialog (chatbot) yang menyediakan informasi umum kepada pelanggan memiliki implikasi yang kurang kritis dibandingkan sistem pintar yang mengemudikan mobil tanpa pengemudi manusia. Namun, kita juga dapat menemukan bahwa chatbot yang menyediakan rekomendasi medis atau keuangan yang sensitif memerlukan ukuran akuntabilitas yang lebih mendekati yang terakhir daripada yang pertama.

Dalam konteks ini, wajar saja jika intensitas kewajiban akuntabilitas bervariasi sesuai dengan tingkat otonomi dan kekritisannya sistem AI. Semakin tinggi tingkat otonomi sistem AI dan semakin tinggi kekritisannya hasil yang dapat dihasilkannya, semakin tinggi pula tingkat akuntabilitas yang akan berlaku bagi organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI tersebut.

3.3 AKUNTABILITAS DAN TATA KELOLA

Untuk memastikan akuntabilitas AI yang bermakna, tidak cukup hanya membagi tanggung jawab di antara para pemangku kepentingan. Dalam kerangka kerja akuntabilitas kami, tata kelola AI proaktif oleh para pemangku kepentingan juga merupakan kunci. Kita harus memastikan bahwa para pemangku kepentingan yang akuntabel mengambil langkah-langkah yang berarti agar tim dan individu "bertanggung jawab secara internal".

Para pemangku kepentingan akan melakukannya dengan menerapkan kerangka kerja tata kelola AI yang menyeluruh dan komprehensif. Kerangka kerja tersebut harus memastikan bahwa individu (atau tim) dan pendukung AI yang teridentifikasi bertanggung jawab atas sistem AI, dan dengan demikian melindungi dari pengembangan, penerapan, atau penggunaan AI yang tidak bertanggung jawab. Pada bagian ini, kami membahas beberapa elemen komponen kerangka kerja tata kelola AI yang tepat.



Karena sistem AI menyentuh beberapa komponen dan bidang keahlian, tata kelola AI harus didekati dari perspektif multidisiplin. Ini berarti bahwa kerangka kerja tata kelola yang memadai akan melibatkan semua pihak terkait, bukan hanya tim ilmiah dan pengembangan, dan tidak boleh melakukan penilaian risiko secara terpisah.

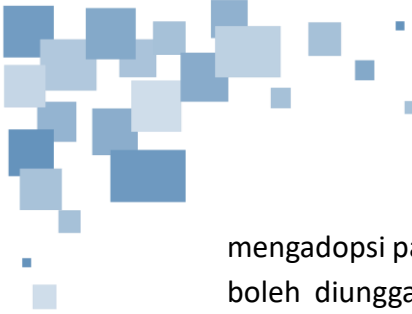
Sebaliknya, para pengambil keputusan, seperti dewan direksi, eksekutif, dan manajemen, serta tim dan individu yang mengeksekusi dan menjalankan proyek AI, seperti pemilik produk, pengguna utama, ilmuwan komputer, dan pengembang, harus dilibatkan dan ditargetkan oleh kerangka kerja tata kelola. Komponen kunci juga mencakup peningkatan kesadaran secara lintas disiplin dan terus memberi informasi kepada pihak-pihak yang terlibat tentang temuan-temuan baru dan evolusi teknologi AI, serta risiko dan dampak yang muncul.

Pedoman, Prinsip, dan Kode Etik Internal

Salah satu langkah mendasar untuk tata kelola AI yang tepat adalah bagi para pemangku kepentingan (dalam hal ini, baik perusahaan maupun pemerintah) untuk menetapkan pedoman yang mengatur aktivitas AI yang mereka lakukan. Pedoman ini dapat bervariasi tergantung pada ukuran, bidang, dan tingkat risiko aktivitas AI. Misalnya, para pemangku kepentingan mungkin ingin menetapkan prinsip dan kode etik AI mereka sendiri, seperti yang telah dilakukan Microsoft dalam mengembangkan serangkaian enam prinsip AI. Google telah melakukan hal yang sama, dengan tujuh prinsip sementara, seperti yang telah kita lihat di bab sebelumnya, berjanji untuk tidak mengembangkan aplikasi AI tertentu seperti senjata AI. Di tingkat negara, anggota G7 memiliki prinsip umum untuk pengembangan AI.

Untuk menerapkan prinsip-prinsip ini, para pemangku kepentingan harus menetapkan pedoman yang jelas yang mengatur proyek AI yang mengintegrasikan prinsip-prinsip Kerangka Kerja untuk AI yang Bertanggung Jawab ini, untuk memastikan bahwa proyek menggabungkan dan menerapkan prinsip-prinsip tersebut, dan terus mematuhi kode etik mereka. Pedoman tersebut harus tersedia secara luas bagi karyawan, disertai dengan pelatihan yang sesuai, dan dikomunikasikan kepada publik.

Pedoman yang berkaitan dengan interaksi pemangku kepentingan dengan konsumen atau klien merupakan bagian penting dari tata kelola yang baik. Banyak platform daring telah



mengadopsi panduan bagi pengguna yang menentukan jenis konten apa yang boleh dan tidak boleh diunggah oleh pengguna. Facebook menyebutnya Standar Komunitas dan YouTube memiliki serangkaian Pedoman Komunitas.


Standar dan pedoman tersebut bermanfaat secara internal bagi perusahaan-perusahaan ini karena membantu karyawan dan algoritma membuat keputusan tentang konten mana yang akan dihapus berdasarkan kriteria yang telah diidentifikasi sebelumnya (misalnya ujaran kebencian atau konten pornografi). Namun, algoritma AI tetap rentan terhadap kesalahan. Kerangka kerja tata kelola AI yang baik untuk perusahaan-perusahaan tersebut, dalam konteks tersebut, harus memperjelas siapa yang bertanggung jawab untuk memantau sistem AI dan menyediakan tinjauan atas rekomendasinya untuk mengidentifikasi dan menyediakan koreksi kesalahan.

Penilaian Risiko yang Komprehensif

Penilaian risiko yang tepat akan menargetkan siklus hidup penuh sistem dan proyek AI. Risiko potensial harus dinilai di awal proyek AI, terus menerus selama penerapan, dan hingga akhir masa pakai sistem AI. Proses penilaian risiko juga harus komprehensif, karena harus mempertimbangkan semua jenis risiko, baik risiko hukum, operasional, maupun etika. Kisi-kisi penilaian harus disusun dengan mempertimbangkan kekhasan AI, seperti potensi otonomi sistem AI dan kemungkinan dampak sistemik. Lebih lanjut, penilaian risiko AI harus mencakup isu-isu terkait infrastruktur dan sistem hosting tempat sistem AI akan beroperasi.

Pemerintah Kanada baru-baru ini menerbitkan Arahan tentang Penggunaan Pembelajaran Mesin untuk Pengambilan Keputusan. Arahan ini hanya berlaku untuk departemen pemerintah federal Kanada, tetapi prinsip-prinsipnya dapat menjadi panduan yang bermanfaat bagi organisasi publik dan swasta lainnya. Arahan ini mewajibkan penyelesaian penilaian dampak yang menetapkan tingkat penilaian ("Tingkat") yang digunakan untuk mengevaluasi aturan mana yang mengatur kapan dan bagaimana alat pengambilan keputusan otomatis dapat digunakan berdasarkan potensi dampaknya terhadap hak-hak individu dan/atau komunitas. Misalnya, untuk menentukan Tingkat mana yang berlaku, dan aturan mana yang harus diikuti, departemen yang berencana menggunakan pembelajaran mesin harus mengisi kuesioner "Penilaian Dampak Algoritmik".

Arahan Kanada mewajibkan peningkatan pengawasan dan perlindungan untuk memastikan akuntabilitas berdasarkan Tingkat penilaian. Misalnya, keputusan Tingkat I (yang tidak berdampak pada hak, kesehatan, atau kepentingan ekonomi individu atau komunitas atau lingkungan) tidak akan memerlukan tinjauan sejawat, pemberitahuan, pelatihan, perencanaan kontingensi, persetujuan, atau bantuan manusia dan hanya memerlukan persyaratan penjelasan minimal. Namun, keputusan Tingkat IV (yang kemungkinan besar akan berdampak sangat tinggi, tidak dapat diubah, dan berkelanjutan terhadap hak-hak tersebut) akan memerlukan tinjauan sejawat dari berbagai pakar yang berkualifikasi, pemberitahuan mendalam, intervensi manusia, penjelasan yang bermakna tentang bagaimana keputusan tersebut dibuat, pengujian, pemantauan, pelatihan berkelanjutan, perencanaan kontingensi, dan persetujuan tingkat tinggi.



Platform independen Belanda untuk Masyarakat Informasi (ECP, yang terdiri dari lembaga pemerintah, bisnis, dan sosial) juga telah mengusulkan model untuk *Artificial Intelligence Impact Assessment* (AIIA) yang merekomendasikan untuk mengikuti peta jalan ini saat mengembangkan proyek AI.

- a) Menentukan kebutuhan untuk melakukan penilaian dampak;
- b) Menjelaskan aplikasi AI;
- c) Menjelaskan manfaat aplikasi AI;
- d) Menentukan apakah tujuan, dan cara mencapai tujuan, dapat dibenarkan secara etis dan hukum;
- e) Menentukan apakah aplikasi AI andal, aman, dan transparan;
- f) Memutuskan pertimbangan dan penilaian;
- g) Mendokumentasikan dan mengaitkan tanggung jawab; dan,
- h) Tinjau secara berkala.

Mekanisme penilaian dampak tersebut akan memastikan bahwa risiko kerugian dimitigasi sebelum organisasi mengembangkan, menerapkan, atau menggunakan sistem AI dan bahwa setiap pemangku kepentingan bertanggung jawab atas keterlibatannya dengan sistem AI.

Kepala Pejabat dan Komite Standar AI

Salah satu langkah paling sederhana yang dapat diambil oleh pemangku kepentingan swasta untuk mendorong akuntabilitas AI adalah dengan menciptakan posisi yang serupa dengan Kepala Pejabat AI. Para pemangku kepentingan harus mempertimbangkan penunjukan Kepala Pejabat AI serta komite standar/akuntabilitas AI, dan mempercayakan mereka dengan tanggung jawab yang lebih besar untuk memastikan bahwa AI dikembangkan, diterapkan, dan digunakan secara bertanggung jawab dan sesuai dengan prinsip, kebijakan, dan persyaratan hukum yang telah ditetapkan.

Pejabat AI dan komite AI dapat dipercaya untuk memantau dan mengaudit aktivitas internal yang melibatkan AI, dan secara lebih umum, memastikan kepatuhan terhadap mekanisme tata kelola AI. Mereka juga dapat bertanggung jawab untuk menjawab pertanyaan dan permintaan eksternal tentang penggunaan AI, dan dengan demikian dapat memainkan peran aktif dalam menjelaskan hasil tertentu kepada masyarakat umum, konsumen akhir, atau otoritas pengatur. Dimungkinkan juga untuk melangkah lebih jauh dengan memberikan tanggung jawab kepada Pejabat AI untuk memastikan kepatuhan sistem dan produk AI terhadap aturan yang berlaku, serta mengawasi permintaan audit oleh otoritas yang berwenang. Kewajiban yang tercantum dalam pasal 37 GDPR bagi pengendali dan pemroses untuk menunjuk petugas perlindungan data dapat menjadi inspirasi dalam kasus Kepala Pejabat AI.

Pejabat eksternal dan komite yang berspesialisasi dalam AI, aktor swasta dan publik harus menerapkan langkah-langkah untuk menyediakan akses kepada karyawan, pejabat, dan direktur terhadap sumber daya dan informasi tentang penggunaan AI dalam organisasi dan prinsip-prinsip etika atau hukum yang lebih luas yang harus dipatuhi.



Kepatuhan Terhadap Hukum dan Peraturan yang Relevan

Tata kelola AI yang baik juga mencakup memastikan pemangku kepentingan mematuhi semua hukum dan peraturan yang berlaku. Para pemangku kepentingan harus yakin bahwa sistem AI mereka mematuhi aturan yang berlaku. Penunjukan Kepala Pejabat AI dan pembentukan komite AI akan membantu upaya kepatuhan karena salah satu tugas mereka adalah memantau perkembangan kewajiban etika dan hukum AI serta menginformasikan organisasi tentang persyaratan baru.

Untuk memastikan kepatuhan terhadap berbagai kewajiban hukum dan peraturan, proses dan metode intrinsik maupun ekstrinsik dapat dimanfaatkan. Proses intrinsik mengacu pada aturan dan persyaratan peraturan yang berlaku yang mungkin tertanam, dirancang, atau diintegrasikan ke dalam model AI itu sendiri, yaitu, dikodekan langsung ke dalam algoritma AI dan sistem hosting. Inilah salah satu alasan mengapa risiko harus dinilai pada tahap awal proyek AI. Sebaliknya, proses "ekstrinsik" dalam konteks ini berarti bahwa proses yang memadai diterapkan bagi pengendali manusia untuk memantau aktivitas dan keputusan AI. Contohnya mungkin termasuk pengujian sampel acak untuk kepatuhan, atau pengujian ulang dengan data yang telah dipilih sebelumnya, yaitu, input data sedemikian rupa sehingga dapat "mendorong" model untuk membuat pilihan yang tidak patuh.

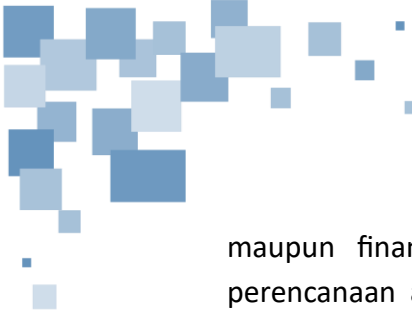
Manajemen Perubahan dan Tanggung Jawab Sosial Perusahaan

Tata kelola AI mencakup penerapan teknik manajemen perubahan yang efektif. Para pemangku kepentingan yang ingin menerapkan proyek AI yang skalabel di seluruh aktivitas mereka harus memastikan manajemen perubahan yang bertanggung jawab. Ada beberapa cara untuk memastikan manajemen perubahan yang memadai, dan kecukupan langkah-langkah yang diambil oleh para pemangku kepentingan akan sangat bergantung pada keadaan, tetapi hal-hal berikut dapat memberikan panduan sebagai ide dasar yang lebih luas dan tidak spesifik bidang. Organisasi yang berhasil beradaptasi dengan AI dari perspektif internal akan lebih mampu menghormati kewajiban akuntabilitas mereka.

Keseimbangan antara konservasi pekerjaan dan kemajuan AI. Sepanjang penerapan AI, para pemangku kepentingan harus melakukan penilaian berkala untuk memastikan keseimbangan yang memadai antara konservasi pekerjaan dan kemajuan teknologi AI. Selain itu, para pemangku kepentingan harus memastikan untuk mendistribusikan keuntungan dari kemajuan tersebut secara bertanggung jawab, sejauh keuntungan dari kemajuan tersebut dapat didistribusikan dengan cara yang meminimalkan dampak negatif pada karyawan dan mengalokasikan peluang baru secara adil.

Penilaian yang memadai di awal proyek AI. Manajemen harus menilai potensi dampak pada karyawan pada tahap awal proyek AI. Melakukan hal tersebut akan menghasilkan perencanaan yang efisien, yang pada gilirannya akan memungkinkan realokasi tugas dan tanggung jawab yang efisien dalam menghadapi dampak tersebut, alih-alih bereaksi terhadapnya, yang sebaliknya dapat menyisakan lebih sedikit pilihan untuk reorganisasi yang tepat.

Perencanaan awal juga memungkinkan perencanaan sesuai dengan kebutuhan yang akan datang, dan jenis perubahan lain yang akan datang, baik manajerial, operasional,



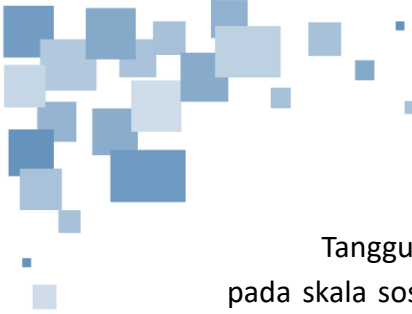
maupun finansial. Dengan mempertimbangkan tujuan dan ruang lingkup proyek AI, perencanaan awal juga memungkinkan untuk memperhitungkan sumber daya yang akan dimanfaatkan dengan lebih baik, untuk mengantisipasi keterlibatan karyawan yang diperlukan, sebelum dan sesudah penerapan.

Cara lain untuk melihatnya adalah bahwa harus ada penilaian awal tentang bagaimana membebaskan beberapa bagian dari proses dan sumber daya dapat digunakan untuk menyelesaikan tugas-tugas lain yang lebih strategis, tetapi saat ini "diprioritaskan" karena bukan bagian dari operasi harian perusahaan. Penilaian ini membutuhkan pemahaman tentang tugas spesifik yang akan diselesaikan oleh sistem AI, serta keseluruhan proses di sekitar tugas tersebut. Proyek AI terkadang memerlukan pertanyaan dan "pengganggu" seluruh proses, yaitu, bukan hanya tugas-tugas yang akan digantikan oleh model. Sering kali, perubahan dalam proses-proses itulah yang akan menciptakan resistensi terbesar, bukan sekadar penggantian tugas dengan sistem AI.

Pergeseran budaya. Untuk mendukung manajemen perubahan yang efektif dalam penerapan AI, para pemangku kepentingan harus menjalankan pergeseran budaya yang sesungguhnya. Untuk mendorong pergeseran budaya, para pemangku kepentingan harus mengidentifikasi apa yang akan ditingkatkan melalui penggunaan AI, manfaat dari peningkatan tersebut, dan pada gilirannya mandat baru yang dapat dijalankan dengan manfaat tersebut. Para influencer dan pemimpin kemudian dapat mengomunikasikan pesan-pesan positif tersebut dan dengan demikian menciptakan reaksi positif terhadap proyek-proyek AI yang akan datang. Hal ini juga akan memungkinkan untuk memilih siapa yang akan menggerakkan proyek AI, para "juara" dalam jargon teknologi, yang memiliki kemampuan untuk menciptakan perubahan dalam pola pikir karyawan, menyampaikan tujuan AI, termasuk akuntabilitas, serta metode dan proses untuk mencapai tujuan tersebut, setiap hari. Teknik-teknik untuk pergeseran budaya ini membantu memastikan bahwa karyawan akan merasa kurang terancam oleh proyek-proyek AI. Pada akhirnya, keberhasilan proyek akan bergantung pada komunikasi yang efektif dengan semua pemangku kepentingan.

Pelatihan karyawan proaktif peningkatan keterampilan. Untuk menerapkan AI yang bertanggung jawab, para pemangku kepentingan harus berupaya meningkatkan keterampilan karyawan dan dengan demikian mengembangkan keahlian internal yang akan relevan di seluruh dan selanjutnya dengan penerapan proyek AI. Peningkatan keahlian tersebut bermanfaat bagi karyawan dan pemangku kepentingan. Untuk itu, pedoman internasional tentang tanggung jawab sosial perusahaan menekankan pentingnya meningkatkan keahlian karyawan, serta mempromosikan praktik yang memastikan transfer "pengetahuan" yang terkait dengan teknologi.

Dengan melakukan hal itu dan dengan demikian memfasilitasi peningkatan keterampilan bagi karyawan, para pemangku kepentingan mendukung pembentukan sumber daya manusia dan menciptakan peluang. Pelatihan karyawan yang tepat penting untuk menghormati kewajiban akuntabilitas, dan para pemangku kepentingan harus memastikan bahwa karyawan mereka mengetahui dan memahami prinsip-prinsip Kerangka Kerja untuk AI yang Bertanggung Jawab ini.



Tanggung jawab sosial. Dampak AI terhadap ketenagakerjaan juga dapat ditangani pada skala sosial oleh para pemangku kepentingan melalui program tanggung jawab sosial perusahaan mereka. Terdapat panduan nasional dan internasional tentang cara menerapkan praktik tanggung jawab sosial perusahaan yang memadai. Program tanggung jawab sosial AI terkait ketenagakerjaan dapat dengan mudah mengintegrasikan panduan tersebut. Sebagai contoh, program ini dapat mencakup program penjangkauan sosial yang akan memberdayakan masyarakat dan komunitas untuk mempelajari keterampilan dan keahlian terkait AI.

Hal ini dapat berupa pembiayaan program pendidikan yang relevan, kemitraan dengan sekolah, perguruan tinggi, dan universitas untuk memastikan bahwa masyarakat memiliki panduan yang tepat tentang jenis keterampilan yang perlu mereka kembangkan guna meningkatkan peluang keberhasilan mereka dalam mempekerjakan para pemangku kepentingan AI. Terakhir, perusahaan/peneliti AI juga dapat mempertimbangkan untuk berinvestasi dalam proyek penelitian yang relevan dan lembaga penelitian publik.

AI yang bermanfaat menuntut akuntabilitas manusia. Prinsip-prinsip umum, meskipun dimaksudkan dengan baik, tidak akan berguna tanpa rezim akuntabilitas yang dapat ditegakkan dan tanpa model tata kelola yang efisien. Dalam bab ini, kami telah menyajikan kerangka kerja yang didasarkan pada prinsip akuntabilitas manusia untuk pengembangan, penerapan, atau penggunaan sistem AI. Sebagaimana telah kita lihat, persyaratan akuntabilitas akan bergantung pada sifat pemangku kepentingan dan jenis aktivitas yang terlibat serta pada kekritisan hak yang dipertaruhkan, potensi risiko aktivitas AI, dan tingkat otonomi sistem AI.


Masih banyak pekerjaan yang perlu dilakukan untuk memperjelas dan menerapkan prinsip ini karena kita masih berada di awal era AI yang baru ini. Pemerintah, dengan partisipasi pelaku bisnis, warga negara, peneliti, dan organisasi masyarakat sipil, harus memperhatikan perubahan yang dibawa oleh AI dan siap untuk mengadaptasi kerangka hukum dan peraturan bila diperlukan.

Di sisi lain, sistem AI, terutama yang berbasis pada jaringan saraf multilapis yang kompleks, berperilaku dengan cara yang sulit diprediksi, dan memecahkan masalah kausalitas, serta menentukan standar perawatan yang tepat bagi pengguna AI, dapat terbukti sulit dalam kasus-kasus tertentu. Dalam pembahasan Transparansi dan Keterjelasan, kita akan melihat bahwa upaya untuk membuat algoritma lebih mudah dijelaskan dapat membantu mengatasi tantangan ini.

Prinsip

Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI dan undang-undang nasional yang mengatur penggunaannya wajib menghormati dan mengadopsi delapan prinsip Kerangka Kebijakan untuk AI yang Bertanggung Jawab ini (atau prinsip akuntabilitas serupa lainnya). Dalam semua hal, manusia harus tetap bertanggung jawab atas tindakan dan kelalaian sistem AI.

Akuntabilitas


- 
- ❖ Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI wajib menunjuk satu atau beberapa individu yang bertanggung jawab atas kepatuhan organisasi terhadap prinsip-prinsip tersebut.
 - ❖ Identitas individu yang ditunjuk oleh organisasi untuk mengawasi kepatuhan organisasi terhadap prinsip-prinsip tersebut wajib diinformasikan jika diminta.
 - ❖ Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI wajib menerapkan kebijakan dan praktik untuk mewujudkan prinsip-prinsip Kerangka Kebijakan untuk AI yang Bertanggung Jawab atau prinsip-prinsip lain yang diadopsi (termasuk prinsip-prinsip serupa yang mungkin dikembangkan untuk industri tertentu), termasuk:
 - I. menetapkan proses untuk menentukan apakah, kapan, dan bagaimana menerapkan proses “Penilaian Dampak AI yang Bertanggung Jawab”;
 - II. menetapkan dan menerapkan prinsip-prinsip “AI yang Bertanggung Jawab Berdasarkan Desain”;
 - III. menetapkan prosedur untuk menerima dan menanggapi keluhan dan pertanyaan;
 - IV. melatih staf dan mengomunikasikan informasi kepada staf tentang kebijakan dan praktik organisasi; dan
 - V. mengembangkan informasi untuk menjelaskan kebijakan dan prosedur organisasi.

Pemerintah

- ✚ Pemerintah yang menilai potensi “kesenjangan akuntabilitas” dalam kerangka hukum dan peraturan yang berlaku untuk sistem AI harus mengadopsi pendekatan seimbang yang mendorong inovasi sekaligus mengurangi risiko kerugian individu atau masyarakat yang signifikan.
- ✚ Kerangka hukum dan peraturan tersebut harus mendukung delapan prinsip Kerangka Kebijakan untuk AI yang Bertanggung Jawab atau mencakup pertimbangan serupa.
- ✚ Pemerintah tidak boleh memberikan status hukum yang berbeda kepada sistem AI, karena hal tersebut akan melemahkan prinsip dasar bahwa manusia pada akhirnya harus bertanggung jawab atas tindakan dan kelalaian sistem AI.

Pendekatan Kontekstual

- Intensitas kewajiban akuntabilitas akan bervariasi sesuai dengan tingkat otonomi dan kekritisan sistem AI. Semakin tinggi tingkat otonomi sistem AI dan semakin tinggi kekritisan hasil yang dapat dihasilkannya, semakin tinggi pula tingkat akuntabilitas yang akan berlaku bagi organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI.



BAB 4

TRANSPARANSI DAN KLARIFIKASI


Ketika suatu keputusan memuaskan seseorang, seringkali, orang tersebut tidak akan peduli bagaimana hasil tersebut dicapai. Misalnya, jika Anda mendaftar ke fakultas hukum dan diterima, Anda tidak akan terlalu tertarik untuk mencari tahu elemen apa dalam surat lamaran Anda yang menentukan keberhasilan Anda mendapatkan posisi yang didambakan. Namun, hal ini tidak berlaku ketika seseorang terdampak negatif oleh keputusan yang merugikan. Kasus-kasus seperti gugatan terhadap proses penerimaan Universitas Harvard dan dugaan diskriminasi terhadap pelamar Asia-Amerika menunjukkan betapa pentingnya kerumitan proses pengambilan keputusan sebuah institusi, faktor-faktor yang dievaluasi, bobotnya, dan peracikannya, ketika orang merasa tidak yakin, curiga, tertipu, atau tidak puas dengan hasilnya.

Untuk menjaga kepercayaan, institusi perlu menemukan keseimbangan antara bersikap transparan dan memberikan penjelasan di satu sisi, dan menjaga integritas, akurasi, dan efektivitas biaya dari proses pengambilan keputusan mereka di sisi lain. Adopsi sistem AI oleh bisnis dan pemerintah akan menambah lapisan kompleksitas pada lanskap ini. International Data Corporation memperkirakan bahwa pengeluaran global tahunan untuk sistem kognitif dan AI akan mencapai Rp 776 triliun pada tahun 2022, lebih dari tiga kali lipat dari Rp 240 triliun yang mereka perkirakan untuk tahun 2018. Banyak dari sistem ini akan berinteraksi langsung dengan pemegang hak individu. Mulai dari rekomendasi film dan buku, skor kredit dan persetujuan pinjaman, analisis citra medis, hingga keputusan hukuman pidana, sistem AI akan mengambil alih banyak proses pengambilan keputusan. Meskipun beberapa mungkin sepele, yang lain mungkin sangat penting dan dapat mengubah hidup.

Sebagaimana halnya pengambilan keputusan non-AI, orang-orang hanya akan menerima kehadiran AI yang terus meningkat ini, dan tidak mundur seperti kaum Luddite modern yang dihadapkan pada mesin tekstil abad ke-21, jika mereka memercayai hasilnya, cara menghasilkan hasilnya, dan lembaga yang menyajikannya. Yang memperumit gambaran ini adalah, seperti yang ditulis Cedric Villani dalam laporannya untuk pemerintah Prancis, kecerdasan buatan adalah teknologi dengan "sifat yang tidak jelas."

Menggambarkan algoritma AI sebagai "kotak hitam" yang cara kerjanya tidak diketahui, adalah salah satu kiasan klasik dalam debat kebijakan seputar teknologi ini. Meskipun kita mungkin tahu masukan apa yang diberikan kepada mesin dan keluaran apa yang dihasilkan, alkimia yang mengatur proses internalnya, interaksi kompleks antara simpul dalam jaringan saraf dalam, dapat tampak tidak dapat dipahami baik oleh orang awam maupun spesialis. Dengan demikian, lembaga dan organisasi yang mungkin tampak seperti kotak hitam bagi pengamat luar akan semakin banyak menggunakan sistem AI, menciptakan situasi kotak hitam di dalam kotak hitam.

Dalam bab ini, kami akan menyajikan kerangka kerja yang menurut kami dapat membuka kedua kotak hitam ini, yaitu kotak hitam institusional dan kotak hitam algoritmik. Kerangka kerja ini dibagi menjadi tiga bagian lebih lanjut. Kami membahas kebutuhan manusia



untuk memahami apa yang dilakukan sistem AI dan bagaimana mereka menghasilkan hasil, rekomendasi, atau keputusan tertentu, untuk memastikan sistem tersebut aman, konsisten dengan harapan, dan tidak ilegal maupun tidak etis.

Kami menyarankan agar kebutuhan ini dapat dipenuhi dengan kombinasi kewajiban transparansi dan keterjelasan bagi para pemangku kepentingan yang mengembangkan, menerapkan, atau menggunakan sistem pengambilan keputusan AI. Kami menggunakan studi kasus untuk membantu kami menyajikan kerangka kerja kami. Sebagai prinsip dasar yang berasal dari pendekatan yang berpusat pada manusia, organisasi swasta dan lembaga sektor publik yang menggunakan sistem AI harus transparan tentang penggunaannya dan harus bekerja dengan asumsi bahwa setiap keputusan berbasis AI dapat dan harus dijelaskan dengan istilah yang dapat dipahami manusia, sehingga memungkinkan proses peninjauan atau banding yang valid. Langkah-langkah tersebut penting untuk memberikan dampak konkret pada prinsip akuntabilitas manusia yang telah kami sajikan di bab sebelumnya. Namun, hal ini tidak berarti bahwa transparansi dan keterjelasan tidak boleh dibatasi dalam konteks tertentu. Sebaliknya, tantangan kebijakan ke depan adalah menemukan keseimbangan yang melindungi hak-hak dengan tetap memperhitungkan biaya, perlindungan kekayaan intelektual, dan penurunan efisiensi. Studi kasus ini akan membantu kita menggambarkan keseimbangan yang harus dicapai dan untuk mendukung pendekatan bertahap yang memodulasi intensitas kewajiban transparansi, tergantung pada konteks dan pentingnya keputusan yang akan dibuat. Terakhir, kami akan menyajikan beberapa solusi dan rekomendasi terkait tantangan dalam mengembangkan dan memenuhi kewajiban transparansi dan keterjelasan.

4.1 HAK ATAS TRANSPARANSI DAN KETERJELASAN

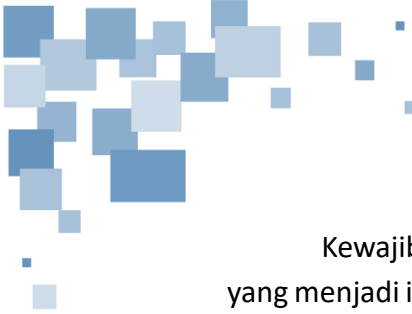
Transparansi

Berbagai istilah dan klasifikasi telah digunakan untuk menggambarkan solusi atas permasalahan AI sebagai kotak hitam, terkadang dengan cara yang kontradiktif. Transparansi, interpretabilitas, keterjelasan, dan keterjelasan adalah istilah-istilah yang muncul dalam literatur terkait.

Transparansi

Untuk tujuan kami, transparansi mengacu pada kewajiban bisnis dan pemerintah untuk memberi tahu masyarakat bahwa mereka berinteraksi dengan sistem AI dan memberikan informasi tentang spesifikasinya, termasuk asal dan sifat data yang digunakan untuk melatih algoritma, serta menjelaskan apa yang dilakukan sistem AI, dan bagaimana cara melakukannya.

Transparansi memberi tahu pelanggan atau pengguna sebelum berinteraksi dengan sistem AI, sebelum hak dan kewajiban mereka terpengaruh oleh suatu keputusan. Transparansi menyediakan informasi yang diperlukan bagi mereka untuk memutuskan apakah mereka ingin berinteraksi dengan AI dan organisasi yang menggunakannya atau tidak.

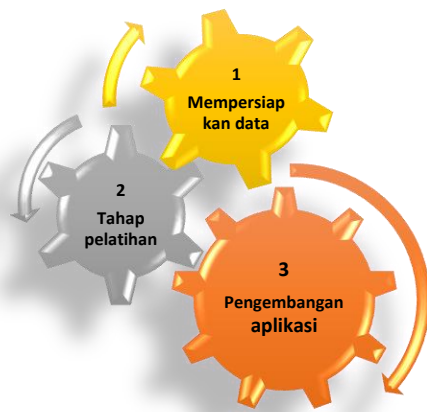


Kewajiban transparansi, ketika diterapkan pada AI, serupa dengan prinsip transparansi yang menjadi inti dari banyak rezim perlindungan data di seluruh dunia. Misalnya, berdasarkan prinsip keterbukaan, Pasal 4.8 *Personal Information Protection and Electronic Documents Act* Kanada (PIPEDA) mewajibkan organisasi yang mengelola informasi pribadi untuk menyediakan informasi tentang kebijakan dan praktik data mereka. Pasal 5 *General Data Protection Regulation* Eropa (GDPR) juga menetapkan bahwa pemegang data pribadi (pengendali) wajib memprosesnya “secara transparan,” menjadikan transparansi sebagai salah satu prinsip inti undang-undang tersebut.

Pemrosesan data pribadi yang transparan berarti memberi tahu orang-orang yang terdampak bahwa informasi pribadi mereka akan diproses dan tentang bagaimana, mengapa, dan oleh siapa informasi tersebut akan digunakan agar mereka dapat membuat keputusan yang lebih tepat. Diterapkan dalam konteks AI, transparansi juga berarti menyediakan informasi yang cukup di muka agar individu dapat membuat keputusan yang tepat. RUU 1001 California adalah contoh yang baik dari jenis kewajiban transparansi yang kami maksud. Hal ini mengharuskan pemilik chatbot daring untuk mengungkapkan, kepada orang yang berinteraksi dengannya, bahwa mereka tidak berkomunikasi dengan manusia sungguhan.

Contoh lain dapat ditemukan dalam GDPR. Pasal 13 dan 14 GDPR mewajibkan pengendali untuk memberi tahu subjek data bahwa data mereka akan dianalisis oleh sistem pengambilan keputusan otomatis. Selain itu, Pasal 22(1) memberikan hak kepada subjek data untuk memilih keluar, dalam keadaan tertentu, ketika keputusan dibuat "semata-mata" melalui pemrosesan otomatis. Pasal 13(2)f) dan 14(2)g) GDPR, lebih lanjut, mencakup kewajiban bagi pengendali untuk memberi tahu subjek data tentang "signifikansi dan konsekuensi yang diperkirakan" dari penggunaan sistem otomatis, serta untuk memberikan "informasi yang bermakna tentang logika yang terlibat."

Dengan demikian, transparansi bukanlah ide baru, dan kita dapat menemukan panduan tentang bagaimana transparansi seharusnya diterapkan pada AI dengan melihat perannya dalam rezim perlindungan data saat ini. Namun demikian, hal itu perlu disesuaikan dengan kekhususan AI, seperti penggunaan big data, pelatihan sistem AI pada data tersebut, dan kemampuan sistem AI untuk meningkat seiring waktu. Komisi Perlindungan Data Pribadi Singapura, misalnya, membagi proses pengembangan AI menjadi tiga tahap: persiapan data, penerapan algoritma AI pada data yang disiapkan (tahap pelatihan), dan penerapan model yang dipilih dalam pengambilan keputusan.



Dalam kerangka kerja kami, transparansi penting dalam semua tahap ini. Oleh karena itu, transparansi dalam AI mungkin memerlukan, misalnya, penyediaan informasi tentang sumber dan sifat data yang digunakan ("garis keturunan data"), terutama dengan memelihara catatan asal data. Ini juga dapat mencakup kewajiban untuk memberi tahu pelanggan

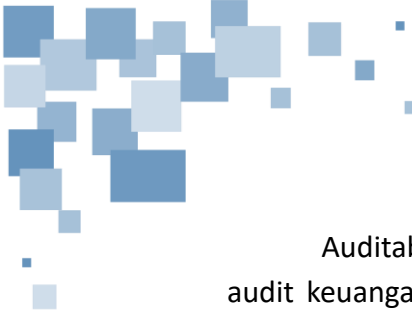
bahwa sistem juga menggunakan data mereka untuk meningkatkan sistem AI. Terakhir, individu harus diberi tahu ketika sistem AI digunakan dalam pengumpulan data yang sedang berlangsung, misalnya ketika kamera dipasang dengan perangkat lunak AI untuk tujuan pengenalan gambar.

Keterjelasan

Setelah keputusan dibuat oleh sistem AI, tidak cukup hanya memberi tahu orang-orang yang terdampak bahwa keputusan tersebut tidak dibuat oleh manusia dan memberikan detail spesifikasi sistem secara umum. Pada tahap pasca-keputusan, diperlukan keterjelasan. Dalam kerangka kerja kami, "keterjelasan" (juga disebut "interpretabilitas" di tempat lain) mengacu pada kewajiban untuk memberikan informasi tentang bagaimana tepatnya suatu keluaran tertentu dihasilkan. Sistem AI yang dapat dijelaskan, atau XAI, adalah sistem yang memberikan penjelasan tentang proses "berpikir"-nya. Hal ini menyiratkan, sebagaimana dinyatakan dalam Rancangan Pedoman Etika Komisi Eropa untuk AI yang tepercaya, bahwa "sistem AI harus dapat diaudit, dipahami, dan dipahami oleh manusia pada berbagai tingkat pemahaman dan keahlian."

Jika transparansi bertujuan untuk membuka kotak hitam kelembagaan, keterjelasan bertujuan untuk membuka kotak hitam algoritmik, menyediakan informasi yang diperlukan untuk memahami bagaimana hak-hak individu terdampak dan mengapa. Dalam hal ini, hak penjelasan dapat mencakup hak atas informasi tentang bagaimana keputusan dibuat, termasuk, dalam beberapa kasus, hak untuk meninjau algoritma sistem, serta hak audit tertentu, sebagaimana dibahas di bawah ini. Sebagai contoh penjelasan dari dunia hukum, banyak sistem peradilan telah menyimpulkan bahwa memberikan alasan, yang seringkali tertulis, merupakan aspek penting dari keadilan prosedural, baik bagi pengadilan maupun badan administratif.

Pembelajaran mesin dan pembelajaran mendalam mungkin merupakan inovasi yang relatif baru, tetapi teknologi baru lainnya menghadapi tantangan serupa. Sejak tahun 1960an dan audit pemrosesan data elektronik (EDP) pertama, bisnis telah mewajibkan sistem teknologi mereka dapat diaudit, sehingga mereka dapat yakin bahwa sistem tersebut beroperasi sesuai dengan hukum yang berlaku.



Auditabilitas ini penting agar bisnis dapat menunjukkan kepatuhan sebagai bagian dari audit keuangan, operasional, atau audit lainnya, kepada regulator atau pihak lainnya, dan untuk membenarkan penggunaan sistem AI kepada para pemangku kepentingan. Dalam konteks AI, hal ini harus diterjemahkan sebagai persyaratan untuk mengintegrasikan fungsi-fungsi penjelasan dari tahap konsepsi sistem hingga tahap keluaran/keputusan dengan kata lain, “penjelasan berdasarkan desain,” sebagaimana terinspirasi oleh konsep “perlindungan data berdasarkan desain” yang ditemukan dalam rezim perlindungan data.

Pada akhirnya, harapannya adalah sistem AI akan memiliki kemampuan untuk menjelaskan keputusannya dengan cara yang sama seperti manusia dapat membenarkan keputusannya sendiri, misalnya dengan menunjukkan faktor-faktor penentu. Penjelasan harus menyediakan hubungan yang hilang antara masukan dan keluaran (keputusan) sistem AI. Dalam konteks ini, kami berusaha menerjemahkan hasil yang diinginkan ini menjadi rekomendasi kebijakan, tidak hanya untuk mendefinisikan keadaan di mana kemampuan penggunaan sistem AI harus bergantung pada sistem yang dapat dijelaskan, tetapi juga untuk mendefinisikan tingkat penjelasan apa yang diperlukan, siapa yang harus memberikan penjelasan, dan siapa yang harus memiliki akses ke penjelasan tersebut.

Pentingnya Transparansi dan Keterjelasan

Mendorong transparansi dan keterjelasan terutama tentang menjaga kepercayaan publik dan memastikan akuntabilitas yang bermakna dari mereka yang mengembangkan, menerapkan, dan menggunakan sistem AI. Ini juga tentang mengakui bahwa menunjukkan akurasi atau efisiensi yang lebih tinggi daripada teknik alternatif mungkin tidak cukup untuk membangun kepercayaan pada sistem AI. Bagi Komisi Perlindungan Data Pribadi Singapura, “peningkatan transparansi akan berkontribusi pada pembangunan kepercayaan dan penerimaan konsumen dengan meningkatkan tingkat pengetahuan dan kepercayaan dalam hubungan pelanggan.”

Transparansi, pada gilirannya, akan mendorong publik yang lebih terinformasi. Keterjelasan sangat penting bagi akuntabilitas karena melalui penjelasan, bias, kesalahan, dan kekurangan suatu sistem akan terungkap. Transparansi dan keterjelasan, jika digabungkan, akan memungkinkan pemantauan yang lebih baik terhadap keabsahan pengembangan, penerapan, dan penggunaan sistem AI.

Bagi konsumen dan pemegang hak individu, mengetahui bahwa sistem AI sedang digunakan dan menerima penjelasan atas keputusan yang diambil dapat menjadi sangat penting. Mereka ingin memahami bagaimana keputusan tersebut dibuat dan mengapa mereka terdampak sedemikian rupa. Tanpa penjelasan yang tepat, mereka akan menganggap penggunaan sistem AI sebagai sesuatu yang sewenang-wenang dan tidak tepat, sehingga menimbulkan risiko bahwa mereka akan menolak untuk mengakui atau menerima keputusan sistem tersebut.



Keterjelasan juga penting bagi pengembang sistem AI yang perlu mengetahui bahwa sistem yang mereka program melakukan apa yang seharusnya. Apakah sistem AI beroperasi atau menghasilkan hasil yang tidak diantisipasi? Mengapa sistem bereaksi terhadap data dengan cara ini dan bukan dengan cara itu? Apakah sistem menghasilkan hasil yang konsisten? Keterjelasan mungkin bahkan lebih penting bagi organisasi yang menggunakan sistem AI untuk membuat keputusan penting yang memengaruhi pihak ketiga karena organisasi


tersebut akan bertanggung jawab atas penggunaannya kepada beragam pemangku kepentingan karyawan, pelanggan, pemasok, mitra bisnis, pemegang saham, regulator, dll. Bagi mereka, sistem AI tidak berbeda dengan alat, proses, atau prosedur lainnya: mereka memiliki keputusan implementasi dan hasilnya, dan dengan demikian, berkepentingan untuk dapat memberikan alasan yang kuat atas suatu hasil kepada klien, regulator, pengadilan, atau publik. Namun, seperti yang akan kita lihat sekarang, pentingnya transparansi dan keterjelasan, serta isi spesifik dari kewajiban ini, akan bervariasi sesuai dengan konteksnya. Oleh karena itu, akan kontraproduktif jika mengadopsi pendekatan serba-atau-tidak-sama-sekali yang tidak mengakui keragaman keadaan di mana AI dapat diterapkan.

4.2 STUDI KASUS

Perbedaan antara penggunaan sistem AI oleh aktor swasta atau publik ini penting karena, dalam hal pengambilan keputusan, kedua kelompok aktor ini tidak memiliki wewenang dan kewajiban yang sama seperti yang telah kami sarankan dalam Prinsip 2 Akuntabilitas.

Bagi aktor swasta, penggunaan sistem AI yang tidak diungkapkan, tidak tepat, tidak akurat, tidak adil, atau tidak etis dapat menyebabkan berbagai konsekuensi yang tidak terduga atau tidak diantisipasi, mulai dari kerugian reputasi atau finansial hingga tanggung jawab pidana, di samping konsekuensi pribadi bagi para pemimpin bisnis yang bertanggung jawab. Berdasarkan pengalaman para eksekutif Toyota Corporation yang secara tidak sopan dibawa ke hadapan Kongres AS untuk memberikan penjelasan terkait keluhan bahwa beberapa kendaraan mereka mengalami akselerasi mendadak yang tidak dapat dijelaskan, kita dapat membayangkan para eksekutif perusahaan AI akan dimintai pertanggungjawaban untuk menjelaskan sistem AI yang berkinerja buruk pada mobil swakemudi.

Di sektor publik, sebagaimana dicatat oleh Sekretariat Dewan Perbendaharaan Kanada, berdasarkan prinsip-prinsip akuntabilitas Menteri, penting bagi lembaga pemerintah untuk memiliki pemahaman penuh tentang perangkat yang mereka gunakan untuk mendukung pengambilan keputusan. Hal ini bukan hanya karena tanggung jawab transaksional mereka terhadap individu tertentu. Sebagian besar entitas publik memiliki kewajiban domestik atau internasional untuk memperlakukan konstituen mereka secara adil dan untuk melindungi hak-



hak tertentu, terutama hak asasi manusia. Hal ini sulit dilakukan ketika tidak ada yang mengerti bagaimana sistem AI mencapai kesimpulannya. Pemerintah juga memiliki wewenang, melalui kepolisian, pengadilan, dan badan administratif, untuk melanggar kebebasan dan hak individu dengan cara yang tidak dapat dilakukan oleh aktor swasta. Kewenangan ini disertai dengan tanggung jawab yang lebih tinggi. Kontras antara tugas sektor publik dan tugas sektor swasta ditonjolkan dalam pernyataan Deklarasi Toronto: Melindungi hak atas kesetaraan dan non-diskriminasi dalam sistem mesin/penghasilan: “Negara memiliki kewajiban untuk memajukan, melindungi, dan menghormati hak asasi manusia; pelaku sektor swasta, termasuk perusahaan, memiliki tanggung jawab untuk selalu menghormati hak asasi manusia.” Dua studi kasus berikut, satu dari sektor swasta dan yang lainnya dari sektor publik, menyoroti berbagai komponen kerangka kerja kami.

Studi Kasus Sektor Swasta

Dalam beberapa tahun terakhir, bank-bank, terutama di Amerika Serikat, meskipun juga di India dan di Eropa, secara kolektif telah menghabiskan miliaran dolar untuk membeli atau mengembangkan teknologi AI, termasuk sistem analitik untuk memprediksi perilaku konsumen, chatbot untuk meningkatkan layanan pelanggan, dan sistem peninjauan dokumen agar bekerja lebih efisien. JPMorgan Chase, misalnya, telah mengembangkan platform *Contract Intelligence* (COiN) untuk menggantikan peninjauan manual perjanjian kredit komersial dengan sistem pembelajaran mesin.

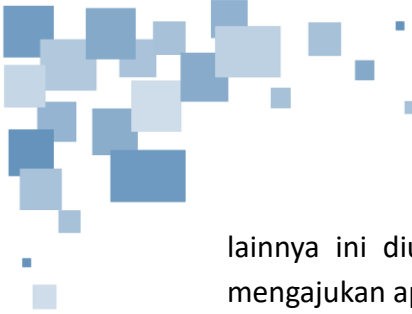
Penerapan solusi AI di industri perbankan dapat membantu menyederhanakan proses pengajuan pinjaman dengan mempersingkatnya dari hitungan minggu menjadi hitungan menit dan membuat prosesnya lebih terjangkau. Sistem pembelajaran mesin yang dilatih pada kumpulan data yang relevan sangat berguna dalam menganalisis informasi keuangan dan pribadi nasabah serta mengevaluasi tingkat risiko untuk pengajuan kredit. Sistem AI dapat dengan cepat memberikan lampu hijau dalam kasus-kasus berisiko rendah, sekaligus mengidentifikasi kasus-kasus yang lebih kompleks untuk ditinjau secara manual.

Aplikasi Pinjaman

Bayangkan kasus John, yang mengajukan pinjaman secara daring dan menerima pemberitahuan penolakan otomatis hanya dalam beberapa menit. Beberapa detik setelah dengan tekun mengisi semua formulir aplikasi pinjaman online. Di sini, banknya, yang tidak menghormati prinsip transparansi, belum memberi tahu John sebelumnya bahwa berkasnya akan diproses oleh sistem AI. Ia bertanya-tanya bagaimana keputusan bisa dibuat secepat itu. Apakah aplikasinya telah ditinjau dengan serius? Dan tanpa bias?

Dalam konteks regulasi perbankan, legislator harus menentukan informasi apa yang perlu diungkapkan dalam keadaan apa. Misalnya, mungkinkah ada keadaan di mana algoritma diwajibkan untuk menjalani proses sertifikasi regulasi?

Kembali ke contoh John dan aplikasi pinjamannya, kita juga dapat membayangkan bahwa, setelah aplikasi pinjaman John ditolak, sistem AI akan menyerahkan berkas John kepada petugas pinjaman bank untuk ditinjau lebih lanjut mengenai situasi keuangannya guna menilai eksposur bank. Jika demikian, transparansi mengharuskan penggunaan hasil sistem AI



lainnya ini diungkapkan di awal, sehingga John dapat mempertimbangkan apakah akan mengajukan aplikasi daring.

Lebih lanjut, jika John meminta detail tentang alasan penolakan aplikasinya, penggunaan sistem AI yang dapat dijelaskan akan memastikan bahwa bank tidak berada dalam situasi yang tidak masuk akal dengan mengakui bahwa mereka tidak tahu. Dalam konteks ini, keterjelasan berarti memastikan bahwa sistem yang digunakan dapat menyediakan metrik dan informasi untuk mengidentifikasi faktor-faktor penting dalam keputusan persetujuan pinjaman. Jika suatu keputusan digugat, informasi tersebut akan berguna dalam memfasilitasi auditabilitas dan mengidentifikasi faktor-faktor penentu (dan memastikan bahwa faktor-faktor tersebut tepat dan tidak diskriminatif atau ilegal). Dalam skenario seperti itu, kemampuan menjelaskan tidak hanya berharga bagi orang yang terdampak, tetapi juga untuk memberikan bukti nyata bagi bisnis bahwa algoritma mereka bebas dari bias dan dirancang serta digunakan secara wajar.

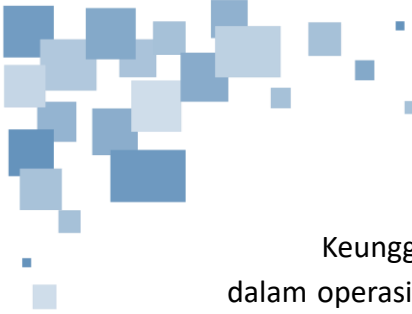
Di sektor swasta, kita juga akan menemukan contoh sistem AI yang mensyaratkan transparansi, tanpa harus melibatkan kebutuhan akan kemampuan menjelaskan. Misalnya, hal ini terjadi ketika sistem AI digunakan untuk menganalisis ekspresi wajah klien untuk menentukan suasana hati mereka. Beberapa toko telah memulai proyek percontohan yang memanfaatkan AI sedemikian rupa untuk membantu mereka meningkatkan pengalaman klien.

Di sana, AI tidak membuat keputusan yang secara langsung memengaruhi hak-hak klien, tetapi transparansi tetap mengharuskan toko untuk mengungkapkan kepada kliennya penggunaan sistem AI, data apa yang dikumpulkan, serta mengapa dan bagaimana menggunakannya, dan berapa lama data akan disimpan. Namun, hal ini juga menggambarkan peran konteks dalam menentukan tingkat kemampuan menjelaskan yang seharusnya disyaratkan. Jika, misalnya, sistem pengenalan wajah digunakan sebagai pengganti barisan polisi, yang mengidentifikasi tersangka dari basis data gambarnya, kita dapat membayangkan bahwa individu atau kelompok minoritas yang terdampak mungkin menuntut informasi lebih lanjut tentang cara kerja algoritma tersebut.

Studi Kasus Sektor Publik

Di sektor publik, pengadilan dan lembaga administratif (terutama dalam kasus pidana, serta dalam sistem imigrasi dan pengungsi), dengan cepat menunjukkan minat pada AI karena alasan yang sama pentingnya di sektor swasta: keinginan untuk meningkatkan akurasi, wawasan yang lebih luas, dan mengurangi biaya. Di sini, transparansi dan kemudahan menjelaskan sangat penting karena departemen sektor publik ini terbebani, kekurangan sumber daya, dan memiliki data dalam jumlah besar (baik dari agen keamanan perbatasan, aplikasi penduduk tetap dan suaka, kumpulan data yang dikumpulkan oleh badan intelijen dan penegak hukum nasional, dll.).

Dengan demikian, mereka mewakili lingkungan yang sangat cocok untuk penggunaan sistem AI. Di sisi lain, dan yang terpenting, individu yang menjadi subjek sistem ini berada dalam posisi yang sangat rentan dan berhak mengetahui bahwa hak-hak mereka dilindungi dengan baik, terlepas dari proses pengambilan keputusan apa pun yang digunakan untuk menentukan nasib mereka.



Keunggulan AI telah meyakinkan banyak badan publik untuk mengadopsi sistem AI dalam operasional sehari-hari mereka. Di AS, sistem AI saat ini digunakan untuk membantu hakim dalam menentukan hukuman, mengeluarkan persyaratan jaminan, dan memberikan pembebasan bersyarat. Di Inggris, departemen kepolisian sedang mengembangkan sistem AI untuk membantu penegak hukum memprediksi di mana kejahatan di masa mendatang mungkin terjadi. Kanada, di sisi lain, berupaya menerapkan algoritma untuk membantu petugas imigrasi dalam menyaring aplikasi imigrasi dan pengungsi. Penggunaan sistem AI seringkali bermanfaat, terutama dalam meningkatkan efisiensi dan mengurangi biaya administrasi publik.

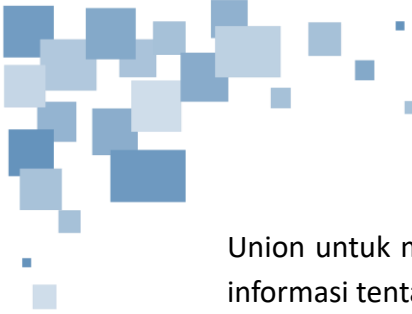
Code for America

Melalui kemitraan dengan lembaga nirlaba Code for America, yang telah mengembangkan alat pembelajaran mesin otomatis bernama "Clear My Record" untuk membantu orang-orang menghapus catatan kriminal mereka, kota San Francisco baru-baru ini mengidentifikasi lebih dari 9.000 hukuman terkait ganja yang dapat dibatalkan.

Namun, sistem AI juga menimbulkan risiko yang lebih besar di sektor publik dibandingkan di sektor swasta. Keputusan badan publik, terutama dalam konteks hukum yudisial dan administratif (misalnya, hukuman pidana dan imigrasi), dapat merusak hak asasi manusia seseorang dengan menempatkannya pada keputusan negara yang mungkin tidak transparan, tidak dapat dijelaskan, dan berdasarkan kumpulan data yang bias atau algoritma yang dirancang dengan buruk. Akibatnya, hasilnya berisiko sewenang-wenang, salah, diskriminatif, dan bertentangan dengan prinsip-prinsip dasar hukum administratif, terutama hak atas keadilan prosedural.

Pentingnya transparansi dan kemudahan dijelaskan baru-baru ini disorot oleh kontroversi mengenai penggunaan sistem *Risk Classification Assessment* (RCA) berbasis komputer oleh *Immigration and Customs Enforcement* Amerika Serikat (ICE). RCA adalah alat analisis risiko yang diterapkan secara nasional oleh ICE sejak 2013. Alat ini membantu agen memutuskan apakah akan menahan warga negara asing yang ditangkap oleh ICE sambil menunggu keputusan tentang status imigrasi mereka. RCA melakukannya dengan menghitung risiko yang ditimbulkan seseorang terhadap keselamatan publik dan risiko melarikan diri jika dibebaskan.

Sistem RCA terinspirasi oleh alat penilaian risiko penahanan pra-persidangan yang dikembangkan pada tahun 1960an untuk sistem peradilan pidana dan awalnya dipromosikan sebagai cara untuk mengurangi tingkat penahanan melalui evaluasi tingkat risiko individu yang lebih akurat. Pada Juni 2018, sebuah investigasi oleh Reuters mengungkapkan bahwa ICE telah memodifikasi sistem RCA sehingga sekarang hanya merekomendasikan penahanan. Meskipun agen ICE tidak terikat oleh rekomendasi sistem, tidak mengherankan bahwa modifikasi sistem RCA menyebabkan peningkatan substansial dalam jumlah penahanan. Kecaman publik terhadap berita ini berlangsung cepat, yang kemudian mendorong New York Civil Liberties




Union untuk mengambil tindakan perdata guna memaksa ICE mengungkapkan lebih banyak informasi tentang sistem RCA-nya.

Alat ICE bukanlah sistem AI berbasis neuron; alat ini lebih mirip sistem pakar karena menggunakan teknik aktuarial otomatis untuk menghasilkan skor risiko. Namun, kasus ini menimbulkan masalah yang relevan dalam kasus AI modern. Pertama-tama, terdapat kekhawatiran mengenai transparansi karena kurangnya informasi yang tersedia untuk publik tentang penggunaan dan fungsi sistem ini, yang menyebabkan berkurangnya tingkat kepercayaan publik.

Kasus ini juga menunjukkan pentingnya keterjelasan, pemahaman tentang cara kerja internal suatu sistem pengambilan keputusan. Dalam hal ini, hampir tidak ada keraguan bahwa sistem tersebut bias terhadap hasil tertentu sebagai akibat dari intervensi manusia yang memprogram sistem untuk hanya merekomendasikan penahanan. Dengan informasi tersebut, seseorang kemudian dapat bertanya apakah sistem tersebut menghormati persyaratan hukum, seperti ketidakberpihakan dan non-diskriminasi.

4.3 ISI KEWAJIBAN TRANSPARANSI DAN KETERJELASAN

1. Tingkat transparansi dan keterjelasan mungkin diperlukan dalam konteks sektor publik dibandingkan di sektor swasta karena risiko-risiko khusus ini dan karena, dalam banyak kasus sektor publik, organisasi tidak memiliki fleksibilitas untuk mengambil keputusan atau tidak. Mereka tidak hanya diberi mandat untuk melakukannya, tetapi seringkali individu yang terdampak tidak dapat menghindari keputusan yang telah dibuat.
2. Misalnya, jika pengadilan umumnya mengadopsi sistem AI dalam putusan hukuman, individu yang diadili tidak akan memiliki pilihan untuk menolak, tetapi pengetahuan bahwa sistem AI terlibat mungkin tetap penting ketika mengajukan banding atas putusan tersebut. Hal ini akan membuat transparansi dan keterjelasan menjadi lebih penting. Bahkan ketika individu tidak memiliki pilihan selain berinteraksi dengan sistem AI, transparansi akan tetap penting bagi masyarakat luas karena masyarakat berkepentingan untuk mengetahui kapan sistem tersebut diimplementasikan dalam pengambilan keputusan publik. Masyarakat mungkin mengharapkan tingkat penjelasan tertentu untuk sistem AI yang menyetujui pinjaman, tetapi individu yang tidak puas tetap akan mempertahankan pilihan untuk berbisnis dengan bank lain jika mereka merasa bank tertentu tidak jelas tentang penggunaan AI-nya. Ini adalah kemewahan yang tidak dimiliki oleh orang-orang yang tunduk pada imigrasi atau keputusan hukuman yang diambil atau dibantu oleh sistem AI. Satu-satunya pilihan mereka adalah mengajukan banding ke pengadilan atau badan yang lebih tinggi di mana banding atau tinjauan administratif yang tepat akan mengharuskan sistem AI untuk dapat dijelaskan. Bahkan dalam kasus seperti itu, kesadaran *ex ante* tentang fakta bahwa keputusan administratif akan diambil menggunakan sistem AI akan membantu menetapkan ekspektasi dan dapat membantu menyediakan informasi yang cukup bagi pemohon yang kecewa untuk menerima hasilnya atau untuk



mempertimbangkan alur penalaran yang paling efektif untuk menentang keputusan tersebut.

Penggunaan sistem AI secara publik atau pribadi tidak boleh mendapatkan kelonggaran untuk mematuhi standar etika atau hukum masyarakat hanya karena sistem tersebut mungkin merupakan alat terbaru dan terancang secara teknologi atau karena "masalahnya sulit." Sebagaimana ditulis Cedric Villani:

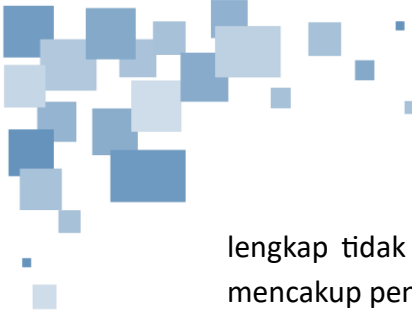
Karena sistem yang menggabungkan teknologi AI menyerbu kehidupan kita sehari-hari, kita secara sah mengharapkan mereka untuk bertindak sesuai dengan hukum dan standar sosial kita. Oleh karena itu, penting bagi undang-undang dan etika untuk mengendalikan kinerja sistem AI. Karena saat ini kita tidak dapat menjamin kinerja sistem pembelajaran mesin secara apriori (sertifikasi formal pembelajaran mesin masih menjadi subjek penelitian), kepatuhan terhadap persyaratan ini memerlukan pengembangan prosedur, perangkat, dan metode yang memungkinkan kita untuk mengaudit sistem ini guna mengevaluasi kesesuaiannya dengan kerangka hukum dan etika kita.

Seperti halnya perangkat, kebijakan, atau prosedur lain yang digunakan dalam masyarakat, orang atau lembaga yang menggunakan sistem AI diharapkan untuk mematuhi hukum yang berlaku. Hal ini membutuhkan transparansi mengenai tujuan penggunaan sistem dan bagaimana sistem tersebut digunakan, serta penjelasan tentang cara pengoperasian dan hasilnya.

Para pakar AI memperdebatkan keinginan untuk dapat dijelaskan, dengan beberapa berpendapat bahwa AI yang dapat dijelaskan bukanlah hasil yang diinginkan, karena potensi terbesar AI justru terletak pada kemampuannya untuk belajar sendiri dengan cara yang lebih kompleks daripada yang dapat dipahami manusia dengan mudah. Dalam sebuah panel di *Conference on Neural Information Processing Systems* edisi 2017, Yann LeCun, seorang pelopor pembelajaran mendalam yang saat ini bekerja di Facebook, berpendapat bahwa sistem AI yang dapat dijelaskan tidak bekerja sebaik sistem yang lebih kompleks, dan akibatnya lebih tidak jelas, dan bahwa sebagian besar pengguna AI yang dihadapkan pada pilihan ini memilih sistem yang bekerja lebih baik daripada yang lebih dapat dijelaskan.

Untuk sistem AI yang membuat keputusan yang relatif tidak penting seperti merekomendasikan film, pilihan ini sepenuhnya sah. Dalam kasus lain, seperti keputusan hukuman atau perekrutan, hukum dapat mewajibkan pengorbanan efisiensi untuk memastikan bahwa auditor atau peninjau dapat memahami bagaimana dan mengapa suatu sistem menghasilkan hasil tertentu, untuk memastikan bahwa sistem tersebut tidak didasarkan pada pertimbangan ilegal. Dalam konteks yang sangat sensitif, keputusan sebaiknya diserahkan sepenuhnya kepada manusia.

Hal ini menunjukkan bahwa isi kewajiban transparansi dan keterjelasan akan bervariasi tergantung konteksnya. Keterjelasan yang lengkap dalam bahasa yang dapat dipahami oleh orang awam mungkin tidak selalu memungkinkan, tetapi kita harus menghindari penerapan sistem yang sama sekali tidak dapat dipahami. Minimal, pengembang harus bertujuan untuk menciptakan sistem yang dapat dipahami oleh para ahli dengan keterampilan dan pengetahuan yang memadai. Lebih lanjut, bahkan dalam kasus di mana penjelasan yang



lengkap tidak memungkinkan, transparansi harus menjadi persyaratan minimal dan harus mencakup pengungkapan bahwa sistem tersebut merupakan "kotak hitam".

Faktor-faktor Penting

Alih-alih pendekatan seragam untuk transparansi dan keterjelasan, kami mengusulkan pendekatan bertahap dan peka konteks. Semakin besar dampak keputusan atau tindakan yang diambil oleh sistem AI terhadap individu atau kelompok, semakin besar pula kewajiban perancang atau pengguna sistem AI untuk transparan tentang penggunaannya dan membuat teknologi tersebut dapat dijelaskan kepada mereka yang terdampak.

Pendekatan bertahap ini secara efektif merupakan uji keseimbangan, yang terinspirasi dari berbagai uji hukum yang berupaya mengakomodasi hak dan preferensi berbagai pihak secara bersamaan. Nilai inheren dalam pendekatan semacam itu jelas dari fakta bahwa tidak semua keputusan yang dibuat menggunakan sistem AI akan memiliki dampak yang sama terhadap hak. Lebih lanjut, pendekatan ini harus netral secara teknologi dalam arti bahwa penggunaan sistem AI tidak boleh mengurangi persyaratan prosedural dan yudisial yang esensial dalam konteks tertentu, terutama keputusan administratif.

Salah satu perhatian utama dalam hal keterjelasan AI berkaitan dengan keberadaan metode yang dapat menghasilkan penjelasan. Pada tahun 2017, tim ahli hukum, ilmuwan komputer, dan ilmuwan kognitif dari Universitas Harvard menemukan bahwa "sebagian besar, secara teknis memungkinkan untuk mengekstrak jenis penjelasan yang saat ini dibutuhkan manusia dari sistem AI." DARPA juga mendorong penelitian tentang AI yang dapat dijelaskan.

Sebagaimana dijelaskan secara lebih rinci dalam lampiran bab ini, para peneliti AI telah mengeksplorasi berbagai pendekatan teoretis terhadap keterjelasan atau interpretasi model (beberapa di antaranya saat ini kurang lebih layak) yang dapat membantu dalam memberikan penjelasan, termasuk:

- **Penjelasan dengan berbagai fitur:** Penjelasan ini memberikan wawasan dengan mengilustrasikan bagaimana perubahan pada fitur masukan titik data tertentu dapat mengubah keluaran model.
- **Penjelasan berbasis teks:** Penjelasan ini mencoba memberikan penjelasan tekstual langsung tentang perilaku model. Penjelasan ini bergantung pada model kedua, yang mampu menghasilkan keluaran berupa penjelasan tekstual yang dapat dibaca manusia yang menggambarkan perilaku model pada masukan tertentu.
- **Penjelasan atau visualisasi berbasis gambar:** Ini adalah kategori pendekatan yang luas, terutama dikembangkan untuk mempelajari model jaringan saraf, yang menggunakan visualisasi untuk memberikan wawasan tentang perilaku model. Sebagaimana pendekatan berbasis teks, pendekatan ini dianggap mutakhir dengan keandalan yang belum teruji. Meskipun demikian, pendekatan ini dapat memberikan informasi kepada pakar. Metode-metode ini cenderung berupa analisis sensitivitas yang menginterogasi bagian internal sistem keputusan untuk memvisualisasikan fitur-fitur masukan yang "menonjol" terhadap keluaran model melalui visualisasi peta panas.

- **Model Pengganti Lokal:** Model pengganti lokal mencoba memberikan wawasan tentang perilaku model dengan menyesuaikan model yang lebih sederhana dan dapat diinterpretasikan untuk meniru model sebenarnya di dekat masukan yang diberikan.

Dari perspektif hukum, dua pengujian yang digunakan di Kanada dapat dianggap sebagai model satu dibuat oleh cabang legislatif dan yang lainnya oleh cabang yudikatif. Pertama, pendekatan yang serupa dengan hukum privasi federal Kanada PIPEDA dapat diadopsi, yang mempertimbangkan ekspektasi wajar individu ketika memperoleh persetujuan dan menetapkan bahwa sifat persetujuan tersebut dapat bervariasi tergantung pada sensitivitas informasi. Dalam konteks AI, "ekspektasi wajar" dari individu yang berinteraksi dengan sistem AI juga harus berfungsi sebagai prinsip panduan. Lebih lanjut, aktor atau entitas yang bertanggung jawab atas sistem AI mungkin diharuskan memberikan penjelasan yang (a) sesuai dengan sifat sensitif input data sistem AI; dan/atau (b) sesuai dengan sifat sensitif dari setiap keputusan yang dihasilkan tentang hak-hak individu tertentu.

Kedua, hakim dapat terinspirasi oleh uji hukum yang diartikulasikan oleh Mahkamah Agung Kanada pada tahun 1999 dalam putusan hukum administrasi penting *Baker v. Canada*, yang menetapkan serangkaian faktor yang perlu dipertimbangkan dalam menentukan isi kewajiban keadilan prosedural yang dimiliki negara kepada individu dalam pelaksanaan tindakan administratif oleh negara.

Faktor-faktor tersebut adalah: (a) sifat keputusan yang dibuat dan proses yang diikuti dalam pengambilannya; (b) sifat skema perundang-undangan (yaitu, perlindungan prosedural yang lebih besar diperlukan jika tidak ada prosedur banding yang diatur dalam undang-undang); (c) pentingnya keputusan bagi individu yang terdampak; (d) harapan yang sah dari individu yang terdampak; dan (e) pilihan prosedural badan administratif itu sendiri. Sebagai aturan umum, lembaga yang menggunakan sistem AI untuk pengambilan keputusan harus memberikan tingkat transparansi dan penjelasan tentang pengembangan, implementasi, dan penggunaannya yang dianggap tepat oleh orang yang berakal sehat dalam keadaan tersebut.

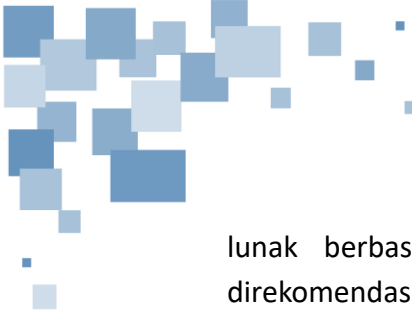
Sebagaimana yang dipegang dalam keputusan Baker: "Nilai-nilai yang mendasari tugas keadilan prosedural berkaitan dengan prinsip bahwa individu atau individu-individu yang terkena dampak harus memiliki [...] keputusan yang memengaruhi hak, kepentingan, atau hak istimewa mereka yang dibuat menggunakan proses yang adil, tidak memihak, dan terbuka, sesuai dengan konteks hukum, kelembagaan, dan sosial dari keputusan tersebut."

Demikian pula, dapat dikatakan bahwa nilai-nilai yang mendasari tugas transparansi mengharuskan individu yang terkena dampak penggunaan AI harus memiliki keputusan yang memengaruhi hak, kepentingan, atau hak istimewa mereka yang dibuat menggunakan proses yang adil, tidak memihak, dan yang terpenting, terbuka. Dalam konteks AI, keadilan prosedural dapat dipahami sebagai kombinasi transparansi dan keterjelasan.

Dalam menetapkan rezim hukum untuk transparansi dan keterjelasan, legislator dan hakim harus mengambil pendekatan yang peka terhadap konteks dan mengevaluasi faktor-faktor berikut untuk menentukan tingkat transparansi dan keterjelasan yang diperlukan untuk berbagai jenis keputusan:

- ✚ Apa konteks organisasinya? Apakah entitas swasta atau publik? Apakah organisasi tersebut menggunakan sistem AI secara umum? Apakah organisasi tersebut memiliki kebijakan terkait sistem AI yang mempertimbangkan isu-isu organisasi, misalnya seputar privasi, keamanan, dan kerahasiaan, atau apakah sistem tersebut diimplementasikan sebagai hasil inisiatif ad hoc (di mana isu-isu terkait mungkin belum sepenuhnya dipikirkan)?
- ✚ Apa tujuan sistem AI spesifik ini? Apakah sistem ini merupakan sistem referensi? Apakah sistem ini menghasilkan titik data yang dievaluasi secara terpisah oleh seorang ahli, atau merupakan sistem pengambilan keputusan yang beroperasi secara independen dari intervensi manusia? Apa saja hak dan kewajiban yang berpotensi terpengaruh oleh sistem ini? Apa saja perlindungan yang tersedia untuk memastikan bahwa sistem AI dapat dengan mudah dimodifikasi guna mencegah penggunaannya untuk tujuan jahat?
- ✚ Akankah pengguna langsung tahu bahwa mereka berinteraksi dengan sistem AI atau bahwa sistem tersebut sedang digunakan? Misalnya, akankah jelas bagi pengguna bahwa ia berinteraksi dengan asisten pribadi otomatis atau chatbot?
- ✚ Data apa yang dibutuhkan, dikumpulkan, digunakan, diproduksi, dan disimpan oleh sistem AI? Apa saja sumber datanya (basis data kepemilikan, sumber terbuka, pialang data, dll.)?
- ✚ Apa risiko bias data tersebut? Bagaimana sistem menggunakan informasi pribadi? Apakah informasi tersebut dianonimkan? Akankah informasi dilindungi oleh langkah-langkah teknis dan organisasi yang sesuai? Akankah informasi tersebut dihapus setelah digunakan?
- ✚ Siapa yang mengembangkan sistem AI? Apakah dikembangkan secara internal atau oleh pihak ketiga? Hal ini mungkin sangat penting di sektor publik, di mana penggunaan sistem AI kepemilikan dapat membatasi kemampuan menjelaskan.
- ✚ Siapa yang berinteraksi dengan sistem AI? Apakah orang ini berada dalam posisi rentan? Apa nilai dan manfaat transparansi dan kemampuan menjelaskan bagi orang tersebut?
- ✚ Jenis sistem AI apa yang dapat digunakan untuk membuat keputusan? Penjelasan seperti apa yang dapat diberikannya? Bisakah sistem ini memberikan contoh kontrafaktual?
- ✚ Apa saja kelemahan transparansi dan keterjelasan dalam konteks yang sedang dipertimbangkan? Berapa biaya yang akan dikeluarkan, dalam hal sumber daya, akurasi, atau efisiensi, jika sistem menghasilkan penjelasan yang dapat dipahami manusia?

Dalam mempertimbangkan faktor-faktor ini, kita perlu menyadari pentingnya transparansi dan keterjelasan bagi penerimaan sosial teknologi AI karena tanpanya, individu tidak akan mempercayai sistem AI. Wajar jika orang-orang meminta penjelasan mengapa sistem AI bank menolak pinjaman mereka; mengapa mobil swakemudi mereka mengalami kecelakaan; mengapa mereka tidak direkomendasikan untuk wawancara tahap pertama oleh perangkat



lunak berbasis AI di departemen sumber daya manusia; mengapa mereka tidak direkomendasikan sebagai kandidat transplantasi organ setelah analisis oleh perangkat lunak AI rumah sakit. Dalam pendekatan yang seimbang, tuntutan mereka tidak akan selalu terpenuhi, tetapi pelanggaran terhadap transparansi dan keterjelasan akan didasarkan, setidaknya secara teori, pada pengalaman dan alasan-alasan yang berprinsip.


Contoh pendekatan semacam itu dapat ditemukan dalam Arahan tentang Penggunaan Pembelajaran Mesin untuk Pengambilan Keputusan ("Arahan") yang disusun oleh Sekretariat Dewan Perbendaharaan Kanada, yang akan berlaku untuk semua "sistem keputusan otomatis" yang digunakan oleh Pemerintah Kanada setelah 1 April 2020. Arahan ini bertujuan untuk mendorong penggunaan sistem AI oleh badan publik yang "sesuai dengan prinsip-prinsip inti hukum administrasi seperti transparansi, akuntabilitas, legalitas, dan keadilan prosedural." Arahan tersebut mencakup sejumlah persyaratan transparansi dan keterjelasan, termasuk:

- (i) membuat penilaian dampak algoritmik sebelum pembuatan sistem keputusan otomatis apa pun;
- (ii) memberikan pemberitahuan kepada publik kapan suatu keputusan akan dibuat seluruhnya atau sebagian oleh sistem keputusan otomatis;
- (iii) memberikan penjelasan yang bermakna kepada individu yang terdampak tentang bagaimana dan mengapa suatu keputusan yang menggunakan sistem keputusan otomatis dibuat;
- (iv) merilis kode sumber khusus milik Pemerintah Kanada untuk setiap sistem keputusan otomatis (dengan pengecualian tertentu);
- (v) memperoleh komponen perangkat lunak untuk sistem pengambilan keputusan otomatis di bawah lisensi sumber terbuka, jika memungkinkan, dan
- (vi) menyediakan opsi sumber daya yang berlaku.

Menariknya, Arahan ini juga mengadopsi pendekatan bertahap. Arahan ini mengklasifikasikan keputusan dalam skala I hingga IV: (I) untuk keputusan dengan dampak kecil hingga tanpa dampak, (II) untuk keputusan dengan dampak sedang, (III) untuk keputusan dengan dampak tinggi, dan (IV) untuk keputusan dengan dampak sangat tinggi terhadap hak atau kesejahteraan individu atau komunitas, kepentingan ekonomi mereka, dan (yang signifikan) keberlanjutan ekosistem. Setiap tingkatan membawa persyaratan transparansi dan keterjelasan yang semakin ketat untuk tinjauan sejawat terhadap sistem, termasuk pemberitahuan sebelumnya, tingkat keterlibatan manusia dalam keputusan, penjelasan, pengujian, dan pemantauan, di antara hal-hal lainnya.

Batasan Transparansi dan Keterjelasan


Pendekatan penyeimbangan yang tepat juga membutuhkan kesadaran akan potensi batasan prinsip transparansi dan keterjelasan. Misalnya, bisnis swasta memiliki kepentingan yang sah untuk merahasiakan pengetahuan berharga mereka sebagaimana otoritas publik memiliki kepentingan yang sah agar tidak setiap detail AI yang mereka gunakan untuk mengejar kepentingan publik secara sah seperti dalam penegakan hukum diungkapkan kepada publik.



Selain itu, mungkin ada situasi lain di mana akan bijaksana untuk mempertahankan tingkat ketidakjelasan tertentu guna membatasi kemampuan publik untuk memanipulasi algoritma. Misalnya, jika cara kerja internal algoritma yang digunakan untuk menilai kandidat dalam proses lamaran kerja diungkapkan kepada kandidat, kemungkinan akan jauh lebih mudah bagi mereka untuk memengaruhi proses demi keuntungan mereka, yang akan mempertanyakan hasil penilaian. Bahkan perusahaan yang menggunakan sistem rekomendasi, terutama platform dengan konten buatan pengguna seperti YouTube, ingin menyembunyikan algoritma mereka untuk menghindari penyalahgunaan. Pada Januari 2019, lebih dari separuh dari dua puluh hasil pencarian teratas di YouTube untuk "RBG" (untuk Hakim Ruth Bader Ginsburg dari Mahkamah Agung AS) adalah untuk video-video yang mengusung teori konspirasi aneh, misalnya bahwa Hakim Ginsburg diselamatkan dengan obat-obatan terlarang misterius. Sebuah artikel di ABA Journal menyatakan bahwa video-video konspirasi tersebut ditampilkan karena "penentuan relevansi" yang dibuat sebagian berdasarkan hasil yang baru diposting atau populer. Tampaknya, baik disengaja maupun tidak sengaja, beberapa aspek algoritma pencarian YouTube dimanipulasi untuk memengaruhi hasil pencarian. Transparansi yang berlebihan dapat mendorong atau memperburuk masalah-masalah semacam ini.

Hal lain yang perlu dipertimbangkan adalah hak privasi, yang dibahas lebih lanjut dalam pembahasan Prinsip 7 Privasi. Kumpulan data yang digunakan untuk melatih AI menarik untuk tujuan transparansi dan keterjelasan karena sumber kesalahan atau bias terkadang dapat ditelusuri kembali ke data pelatihan asli. Namun, kumpulan data tersebut seringkali berisi informasi pribadi. Setiap prosedur yang dikembangkan untuk memastikan sistem AI dapat dijelaskan perlu mempertimbangkan masalah privasi yang melekat dalam penyimpanan atau akses selanjutnya ke, atau penggunaan, informasi pribadi tersebut. Misalnya, apakah individu yang data pribadinya digunakan untuk melatih atau menguji sistem menyetujui penyimpanan informasi pribadi tersebut atau menyetujui pemberian akses ke informasi yang akan diberikan kepada pihak ketiga di masa mendatang?

Dalam setiap contoh ini, kewajiban transparansi dan kemampuan menjelaskan yang terlalu luas dapat berimplikasi negatif terhadap perlindungan rahasia bisnis atau terhadap upaya mengejar kepentingan publik atau pribadi yang sah. Akibatnya, kewajiban tersebut dapat menunda atau mencegah adopsi sistem AI, meskipun mungkin memiliki efek yang menguntungkan. Oleh karena itu, kami percaya bahwa prinsip-prinsip regulasi yang mewajibkan transparansi dan kemampuan menjelaskan terkait sistem AI harus mempertimbangkan adanya kepentingan yang saling bertentangan, namun tetap sah, yang perlu dilindungi. Oleh karena itu, kewajiban transparansi dan kejelasan tidak boleh terlalu membatasi kemampuan pengembang dan pengadopsi AI untuk melindungi algoritma kepemilikan, rahasia bisnis, dan pengetahuan mereka, serta untuk menjaga keunggulan kompetitif yang mungkin mereka miliki dibandingkan pesaing sebagai hasil dari penggunaan AI, terutama ketika keputusan yang diambil memiliki risiko rendah. Dalam konteks ini, penting juga bahwa pengaburan algoritma dan kerahasiaan ketat melalui rahasia dagang seringkali



merupakan satu-satunya pilihan yang layak untuk melindungi pengetahuan berharga dan eksklusivitas di area yang tidak memiliki perlindungan HKI tradisional.

Hukum dan peraturan yang mengatur kewajiban transparansi dan keterjelasan karenanya harus mengidentifikasi kepentingan yang saling bertentangan yang dipertaruhkan dan menyeimbangkannya satu sama lain. Di satu sisi, hukum tidak boleh secara sepihak berfokus pada pengungkapan setiap detail yang berkaitan dengan penggunaan AI. Tingkat detail yang tinggi mungkin hanya membingungkan dan kurang bernilai bagi penerima. Lebih lanjut, beban pada organisasi yang harus menyediakan informasi dapat menjadi faktor yang secara wajar membatasi kewajiban (seperti, misalnya, Pasal 14(5) GDPR membatasi kewajiban transparansi di mana penyediaan informasi tidak mungkin atau memerlukan upaya yang tidak proporsional).

Di sisi lain, bahkan aturan transparansi dan keterjelasan yang wajar harus memungkinkan penilaian kasus per kasus yang dapat membatasi prinsip-prinsip tersebut ketika diperlukan untuk melindungi kepentingan pihak ketiga yang sah. Hukum Eropa yang harus menangani situasi serupa dengan konflik kepentingan seringkali memperkirakan bahwa, pada langkah pertama, kepentingan para pemangku kepentingan yang relevan harus diidentifikasi sehingga, pada langkah kedua, kepentingan-kepentingan tersebut dapat ditimbang satu sama lain untuk menentukan pihak mana yang menang, misalnya, apakah kepentingan yang mendukung transparansi dan kemudahan dijelaskan lebih besar daripada yang menentanginya.

Hukum perlindungan data umumnya melibatkan penyeimbangan kepentingan semacam itu. Hal ini tercermin dalam Bagian 3 undang-undang privasi federal Kanada, PIPEDA, yang menetapkan tujuan Bagian Privasi dari undang-undang tersebut:

Tujuan Bagian ini adalah untuk menetapkan, di era di mana teknologi semakin memfasilitasi sirkulasi dan pertukaran informasi, aturan untuk mengatur pengumpulan, penggunaan, dan pengungkapan informasi pribadi dengan cara yang mengakui hak privasi individu sehubungan dengan informasi pribadi mereka dan kebutuhan organisasi untuk mengumpulkan, menggunakan, atau mengungkapkan informasi pribadi untuk tujuan yang dianggap tepat oleh orang yang berakal sehat dalam keadaan tersebut.

Lebih lanjut, hukum perlindungan data Eropa seringkali mengharuskan kepentingan individu subjek data ditimbang terhadap kepentingan publik dan pribadi yang saling bertentangan. Misalnya, kewajiban pengendali data untuk memberikan informasi secara transparan tentang pemrosesan data pribadi mereka dapat, dalam keadaan tertentu, dibatasi jika bertentangan dengan kepentingan pihak ketiga. Pertimbangan serupa berlaku saat menentukan apakah pemrosesan data dalam lingkup GDPR dibenarkan atau tidak berdasarkan kepentingan sah pengendali data. Menurut pandangan kami, konsep-konsep ini dapat dijadikan inspirasi dalam mengembangkan batasan yang wajar terhadap kewajiban transparansi dan keterjelasan.

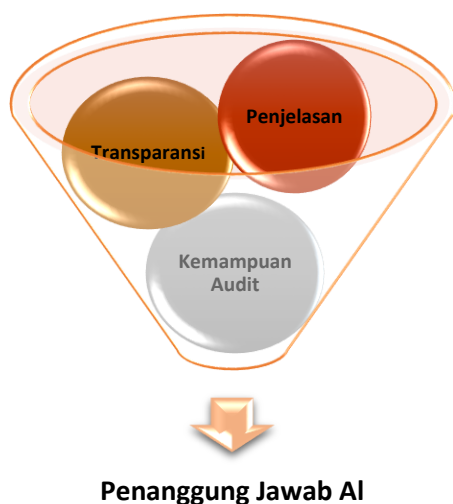
Akhirnya, untuk menangani konflik kepentingan, mungkin juga dianggap cukup untuk hanya mengungkapkan informasi yang relevan (misalnya, kode, input dan output) kepada pihak ketiga yang netral (misalnya, auditor) yang akan mengaudit proses dan memastikan

kewajaran, dengan transparansi terbatas. Meskipun opsi audit ini mungkin terlihat menarik, beban keuangan (dan pertanyaan tentang siapa yang akan menanggung biaya), kesulitan implementasi (mengharuskan auditor benar-benar independen, menentukan cara menangani perselisihan, dll.) dan lamanya proses audit yang diharapkan mungkin akan menggagalkan tujuannya, itulah sebabnya kami akan berhati-hati dalam merekomendasikan solusi seperti itu.

Audit Algoritma

Menyeimbangkan isu-isu ini tentu tidak mudah, tetapi seperti yang telah kami sampaikan, ini bukanlah tantangan yang hanya dihadapi AI. Di bagian terakhir ini, kami akan membahas auditabilitas algoritma, sebuah cara untuk mematuhi kewajiban eksplanabilitas secara lebih tepat.


Disiplin audit TI, misalnya, telah berkembang sejak tahun 1960an untuk mengimbangi kemajuan teknologi dan, khususnya, untuk menangani isu-isu seputar kepemilikan kekayaan intelektual, manipulasi sistem, dan masalah privasi. Kami dapat mengantisipasi bahwa, dengan cara yang sama, protokol yang dikembangkan untuk memastikan bahwa sistem AI dapat dijelaskan akan berkembang untuk mengakomodasi isu-isu sistem AI di area ini.



Ada dua contoh penerapan prinsip audit sistem untuk menjelaskan sistem AI. Pertama, agar sistem AI dapat dijelaskan, kita perlu menguji kinerja sistem AI terhadap tujuan sistem tersebut. Hal ini mengharuskan tujuan sistem AI didefinisikan terlebih dahulu untuk menyediakan dasar bagi penjelasan. Penjelasan kinerja sistem AI hanya masuk akal jika diukur berdasarkan ekspektasi yang telah didefinisikan sebelumnya tentang

tujuan sistem tersebut dirancang. Kesulitan pada tahap ini berkaitan dengan fakta bahwa karena sistem AI beroperasi berdasarkan bobot yang melekat pada sejumlah besar neuron buatan, logika di balik sistem AI, yaitu alasan untuk setiap keputusan tertentu, mungkin tidak dapat dipahami. Sekalipun memungkinkan untuk mengaudit algoritma sistem AI, hasilnya mungkin tidak bermakna.

Kedua, dalam merencanakan pengembangan sistem AI apa pun, perlu direncanakan penerapan langkah-langkah yang tepat untuk mendukung keterjelasan, misalnya, merancang prosedur yang perlu diimplementasikan untuk menyimpan salinan set data yang digunakan untuk melatih atau menguji sistem. Hal ini akan memastikan bahwa data tetap dapat diakses untuk verifikasi selanjutnya. Sekali lagi, kebutuhan untuk menangani prosedur, alat, dan metode ini sebagai bagian dari proses perencanaan tidak hanya berlaku untuk sistem AI, tetapi juga berlaku untuk sistem TI secara umum. Namun, kekhususan AI adalah bahwa jaringan saraf



bersifat dinamis, misalnya, jaringan saraf dapat berkembang dan belajar lebih lanjut seiring beroperasi.

Auditabilitas sistem AI tidak selalu memerlukan audit algoritma. Tergantung pada konteksnya, tujuan melakukan audit sistem AI mungkin dapat dicapai bahkan tanpa pemahaman yang mendalam tentang logika algoritma, misalnya, seperti yang telah disebutkan:

Tidak selalu perlu, bermanfaat, atau bahkan mungkin untuk menarik kesimpulan dari pemeriksaan kode sumber. Auditor mungkin merasa cukup hanya dengan memeriksa kewajaran dan kesetaraan suatu program (hanya melakukan apa yang diwajibkan), misalnya dengan mengirimkan berbagai data masukan palsu, atau dengan membuat sejumlah besar profil pengguna sistem sesuai dengan pedoman yang tepat, dll.

Teknik untuk memastikan keterjelasan dan auditabilitas sistem AI niscaya akan berkembang dan meningkat seiring waktu, menyatu di sekitar praktik terbaik yang akan menyediakan dasar bagi standar industri formal dan informal.

4.4 TEKNIK INTERPRETABILITAS DALAM AI


Komunitas akademis telah lama mempertimbangkan masalah interpretabilitas sistem pengambilan keputusan otomatis dan model statistik. Pertimbangan ini menjadi semakin penting di era sistem AI kotak hitam, yang mana interpretabilitas bukanlah suatu hal yang mutlak. Tujuan dokumen ini adalah untuk merangkum secara mendalam diskusi akademis mengenai topik ini dengan meninjau berbagai metode yang diusulkan untuk interpretabilitas. Kami mencoba untuk tidak terlalu banyak mengomentari keandalan berbagai pendekatan, kecuali untuk menunjukkan pendekatan-pendekatan mutakhir. Biasanya terdapat strategi khusus untuk interpretabilitas dan keandalannya akan bergantung pada banyak faktor yang spesifik terhadap situasi dan teknologi.

Untuk tujuan bagian ini, himpunan data adalah kumpulan titik data, yang masing-masing mewakili serangkaian pengukuran individual. Setiap titik data merupakan vektor pengukuran, yang masing-masing disebut fitur. Misalnya, fitur dapat berupa bilangan riil yang merepresentasikan pengukuran kontinu, bilangan bulat yang merepresentasikan kategori hasil, atau string yang merepresentasikan teks.

Mode adalah sistem pengambilan keputusan atau AI, yaitu prosedur otomatis yang mengambil masukan dan menghasilkan keluaran yang digunakan seluruhnya atau sebagian dalam proses pengambilan keputusan. Misalnya, model regresi linier memiliki respons linier (keluaran) terhadap suatu titik data (masukan).

Ada banyak jenis interpretabilitas yang mungkin kita inginkan. Kita membagi pertanyaan-pertanyaan ini menjadi dua baris. Tentu saja, perbedaan ini tidak begitu jelas dan hanya digunakan untuk tujuan pengorganisasian.

Pertama, model akan tunduk pada tuntutan interpretabilitas. Sebagaimana ditunjukkan dalam literatur, tuntutan interpretabilitas seringkali diperlukan ketika spesifikasi masalah tidak lengkap. Khususnya, jika perilaku model yang diinginkan tidak dapat



sepenuhnya divalidasi dengan metrik kuantitatif, menjadi interpretabilitas merupakan solusi yang baik. Misalnya, sebelum menerapkan sistem pemrosesan pinjaman otomatis, sebuah organisasi mungkin tertarik untuk mengembangkan pemahaman kualitatif yang terperinci dan akurat tentang cara kerjanya dalam mendukung layanan pelanggan. Dengan demikian, metode untuk interpretabilitas model berupaya memberikan wawasan tentang bagaimana model menghitung keluarannya dari masukan yang diberikan.

Kedua, data itu sendiri mungkin juga tunduk pada tuntutan interpretabilitas. Secara umum, perilaku model ditentukan oleh kumpulan data melalui beberapa proses penyesuaian algoritmik dan pemilihan model. Jenis data, proses pembangkitan data, atau struktur data lainnya dapat memengaruhi proses ini untuk menghasilkan perilaku yang tidak diinginkan. Kembali ke contoh pemrosesan pinjaman otomatis, jika bias sistematis terhadap komunitas terpinggirkan tercermin dalam data, hal ini dapat menghasilkan model yang memperkuat bias ini dalam keputusannya. Dengan demikian, metode untuk interpretabilitas data berupaya memberikan wawasan tentang struktur data atau bagaimana penyesuaian model dipengaruhi oleh data.

Interpretabilitas Model

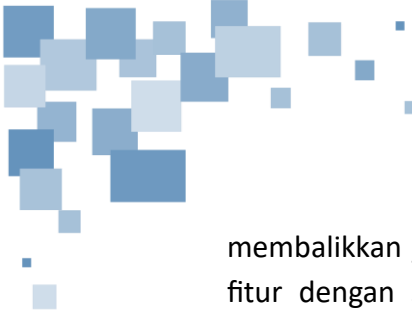
Berikut ini adalah daftar metode untuk memberikan wawasan tentang perilaku model itu sendiri. Metode-metode ini mencoba memberikan penjelasan yang mengkaji pertanyaan: "Dengan masukan, bagaimana model tersebut sampai pada keluarannya?"

Banyak model yang secara alami dapat dianggap dapat diinterpretasikan. Ini termasuk model klasik seperti regresi linier, regresi logistik, atau pohon keputusan. Metode yang populer dalam pembelajaran mesin dan AI kontemporer, seperti jaringan saraf tiruan, cenderung secara alami tidak dapat diinterpretasikan, meskipun Lipton berpendapat bahwa pandangan ini mungkin terlalu sederhana. Meskipun demikian, fokus di sini adalah pada metode yang mungkin berfungsi baik untuk model yang dapat diinterpretasikan secara alami maupun yang tidak dapat diinterpretasikan secara alami. Semua metode ini (atau modifikasinya) cenderung dapat diterapkan secara luas, dengan asumsi yang relatif sedikit pada model itu sendiri.

Penjelasan dengan Fitur yang Berbeda

Penjelasan ini memberikan wawasan dengan mengilustrasikan bagaimana perubahan pada fitur masukan titik data tertentu dapat mengubah keluaran model. Keluarga plot gaya ketergantungan parsial termasuk yang paling banyak jumlahnya dan paling banyak dipelajari. Dalam bentuknya yang paling dasar, plot ketergantungan parsial menunjukkan pengaruh satu atau dua fitur terhadap keluaran suatu model dengan memplot respons seiring dengan perubahan nilai fitur, yang dirata-ratakan pada suatu set data. Misalnya, plot ketergantungan parsial dapat mengungkapkan bahwa ketergantungan marginal suatu model terhadap suatu fitur bersifat linear versus non-linear. Ekstensi dari metode ini mencakup plot ekspektasi bersyarat. Plot ketergantungan parsial juga dapat digunakan untuk memberikan uji statistik ketergantungan antara efek fitur dalam model tertentu.

Kontrafaktual adalah jenis penjelasan lain yang diperoleh dengan memvariasikan fitur. Dengan masukan tertentu, kontrafaktual adalah perubahan terkecil pada nilai fitur yang mengubah keluaran model menjadi keluaran yang telah ditentukan sebelumnya. Jangkar



membalikkan jenis penjelasan yang diberikan oleh kontrafaktual. Jangkar adalah kumpulan fitur dengan sifat bahwa jika dipertahankan tetap, keluaran model tidak akan berubah meskipun fitur non-jangkar berubah. Penjelasan semacam ini menjawab pertanyaan: "Fitur mana yang cukup untuk menjangkarkan keluaran pada masukan tertentu?"

Penjelasan Berbasis Teks

Penjelasan ini mencoba memberikan penjelasan tekstual langsung tentang perilaku model. Mereka mengandalkan model kedua, yang dirancang untuk menghasilkan keluaran berupa penjelasan tekstual yang dapat dibaca manusia yang menggambarkan perilaku model pada masukan tertentu. Teknik-teknik ini harus dianggap sebagai penelitian mutakhir dan belum terlalu andal. Misalnya, mempertimbangkan sistem yang menggabungkan model prediksi peringkat dengan model topik. Kata-kata teratas dalam model topik digunakan untuk menjelaskan prediksi peringkat.

Penjelasan atau Visualisasi Berbasis Gambar

Ini adalah kategori pendekatan yang luas, terutama dikembangkan untuk mempelajari model jaringan saraf tiruan, yang menggunakan visualisasi untuk memberikan wawasan tentang perilaku model. Sebagaimana pendekatan berbasis teks, pendekatan ini dianggap mutakhir dengan keandalan yang belum teruji. Meskipun demikian, pendekatan ini dapat memberikan informasi kepada pakar. Metode-metode ini cenderung berupa analisis sensitivitas yang menginterogasi bagian dalam sistem keputusan untuk memvisualisasikan fitur-fitur masukan yang "menonjol" terhadap keluaran model melalui visualisasi peta panas. Metode-metode ini berkaitan erat dengan metode pemilihan fitur dan tingkat kepentingan untuk interpretabilitas data.

Model Pengganti Lokal

Model pengganti lokal mencoba memberikan wawasan tentang perilaku model dengan menyesuaikan model yang lebih sederhana dan dapat diinterpretasikan untuk meniru model sebenarnya di dekat masukan yang diberikan. Model pengganti kemudian dapat diinterogasi untuk memberikan penjelasan tentang perilaku lokal model sebenarnya. Ribeiro dkk. mengembangkan pendekatan ini. Meskipun Alvarez-Melis dan Jaakkola berpendapat bahwa pendekatan ini mungkin kurang robust dalam beberapa kasus.

Penjelasan dengan Contoh

Metode-metode ini memberikan wawasan tentang perilaku model pada input dengan menyediakan titik-titik data serupa dengan perilaku yang sama. Caruana dkk. menggunakan representasi internal model yang telah difit untuk menentukan metrik jarak, yang dapat digunakan untuk mengidentifikasi contoh-contoh dari set pelatihan atau pengujian dengan perilaku serupa.

Interpretabilitas Data

Berikut ini adalah daftar metode untuk memberikan wawasan tentang set data. Metode-metode ini mencoba memberikan penjelasan yang mempertanyakan pola apa yang inheren dalam set data dan bagaimana hal ini memengaruhi model?



Ekstraksi Fitur

Metode ekstraksi fitur menganotasi setiap titik data dengan fitur tambahan yang ditemukan secara algoritmik dari seluruh set data. Secara umum, fitur-fitur tambahan ini sendiri dapat diinterpretasikan dan memberikan wawasan tentang pola yang terdapat dalam data. Terdapat banyak metode semacam itu, mulai dari teknik klasik dalam statistika hingga penelitian mutakhir. Metode klasik meliputi analisis komponen utama dan metode pengelompokan. Metode yang lebih mutakhir meliputi metode reduksi dimensionalitas non-linier.

Pentingnya atau Seleksi Fitur

Seleksi fitur adalah teknik statistik klasik untuk menentukan fitur mana yang paling menjelaskan perilaku model tertentu. Tinjauan yang lebih rinci tentang metode-metode ini berada di luar cakupan tinjauan ini; untuk tinjauan metode-metode ini, yang mencakup regresi ridge dan laso, seleksi subset terbaik, dan regresi langkah maju. Pada tingkat tinggi, metode-metode ini membandingkan besarnya respons model terhadap berbagai fitur untuk memilih subset yang paling penting di antara fitur-fitur tersebut. Ketepatan teknik inferensi statistik dapat terganggu dalam konteks ini. Bidang inferensi pasca-selektif menangani koreksi hal ini.

Instansi yang Berpengaruh

Penjelasan ini bertujuan untuk mengukur pengaruh satu titik data dalam set data terhadap model akhir yang telah disesuaikan. Topik ini memiliki sejarah yang kaya dalam statistika. Jarak Cook adalah jarak yang didefinisikan pada model yang mengukur efek penghapusan suatu instans terhadap prediksi model. Jika efek penghapusan suatu titik data meningkatkan jarak antar model, hal ini mencerminkan tingkat pengaruh yang tinggi dari titik data tersebut.

Prototipe dan Kritik

Metode ini memilih subset dari himpunan data yang "prototipikal", karena mereka merangkum himpunan data dengan baik, dan subset poin kritik yang tidak dijelaskan dengan baik oleh prototipe. Secara intuitif, metode ini memberikan wawasan tentang titik data tipikal dari himpunan data, yang mencerminkan pola umum dalam data, serta outlier, yang dapat berdampak besar pada proses penyesuaian model.

Pendekatan yang kami sarankan dapat diringkas sebagai penggunaan kewajiban transparansi untuk membuka kotak hitam institusional dan penggunaan kewajiban keterjelasan untuk membuka kotak hitam algoritmik dengan cara yang mencapai keseimbangan yang menghormati semua pihak yang terlibat. Prinsip-prinsip ini fundamental bagi setiap aspek AI yang disajikan di bagian lain laporan ini. Pada akhirnya, AI adalah sebuah alat, dan alat yang digunakan oleh manusia seharusnya dapat dipahami.

Prinsip

Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI dan undang-undang nasional yang mengatur penggunaannya wajib memastikan bahwa, sejauh yang wajar mengingat keadaan dan perkembangan teknologi, penggunaan tersebut transparan dan hasil keputusan sistem AI dapat dijelaskan.

Definisi

- Transparansi merupakan kewajiban bagi organisasi yang menggunakan AI dalam proses pengambilan keputusan untuk menyediakan informasi mengenai: a) fakta bahwa suatu organisasi menggunakan sistem AI dalam proses pengambilan keputusan; b) tujuan yang dimaksudkan dari sistem AI dan bagaimana sistem AI akan dan dapat digunakan; (c) jenis set data yang digunakan oleh sistem AI; dan (d) informasi yang bermakna tentang logika yang terlibat.
- Keterjelasan merupakan kewajiban bagi organisasi yang menggunakan AI dalam proses pengambilan keputusan untuk menyediakan informasi yang akurat dengan istilah yang dapat dipahami manusia yang menjelaskan bagaimana suatu keputusan/hasil dicapai oleh sistem AI.

Tujuan


- ❖ Transparansi dan Keterjelasan bertujuan untuk menjaga kepercayaan publik terhadap sistem AI dan menyediakan informasi yang memadai untuk membantu memastikan akuntabilitas yang bermakna dari pengembang, pelaksana, dan pengguna sistem AI, serta untuk menunjukkan apakah keputusan yang dibuat oleh sistem AI adil dan tidak memihak.
- ❖ Prinsip Transparansi dan Keterjelasan mendukung prinsip Akuntabilitas, prinsip Keadilan dan Non-Diskriminasi, prinsip Keamanan dan Keandalan, serta prinsip Privasi, Penggunaan yang Sah, dan Persetujuan.

Pendekatan Bertahap dan Kontekstual

- Intensitas kewajiban transparansi dan keterjelasan akan bergantung pada konteks keputusan dan konsekuensinya bagi orang yang menjadi subjeknya. Cakupan dan intensitas kewajiban transparansi dan keterjelasan akan meningkat seiring dengan meningkatnya sensitivitas kumpulan data yang digunakan oleh sistem AI dan seiring dengan meningkatnya materialitas hasil keputusan sistem AI.
- Penentuan intensitas kewajiban transparansi dan keterjelasan harus menyeimbangkan kepentingan orang yang menjadi subjek keputusan dan kepentingan organisasi pembuat keputusan. Kriteria utama adalah ekspektasi wajar dari orang yang menjadi subjek keputusan tersebut.

Transparansi dan Keterjelasan Berdasarkan Desain

- Organisasi yang mengembangkan sistem AI harus memastikan bahwa logika dan arsitektur sistem berfungsi untuk memfasilitasi persyaratan transparansi dan keterjelasan. Sejauh yang dapat dipraktikkan secara wajar, dan dengan mempertimbangkan perkembangan terkini pada saat itu, sistem tersebut harus dirancang dari tingkat paling mendasar ke atas untuk mendorong transparansi dan keterjelasan berdasarkan desain. Jika terdapat pilihan antara arsitektur sistem yang kurang transparan atau lebih transparan, opsi yang lebih transparan harus diutamakan.
- Pengguna sistem AI dan orang-orang yang menjadi subjek keputusan mereka harus memiliki cara yang efektif untuk mencari penyelesaian jika organisasi yang



mengembangkan, menerapkan, atau menggunakan sistem AI tidak transparan tentang penggunaannya.

Netralitas Teknologi

- ✓ Penggunaan sistem AI oleh organisasi publik atau swasta tidak mengurangi persyaratan prosedural dan substantif yang biasanya melekat pada suatu keputusan ketika proses pengambilan keputusan sepenuhnya dikendalikan oleh manusia.

Studi interpretabilitas merupakan bidang penelitian yang lama, tetapi aktif, di dunia akademis. Kami telah mencoba memberikan ikhtisar beberapa metode, baik yang diterima maupun yang diusulkan. Tema yang muncul dalam analisis ini adalah bahwa interpretabilitas dan penjelasan bukanlah konsep yang monolitik. Sebaliknya, penyelesaian masalah interpretabilitas seringkali membutuhkan pengetahuan spesifik domain yang signifikan dan analisis yang disesuaikan.

BAB 5

Keadilan dan Non-Diskriminasi

5.1 AI, Keadilan, dan Diskriminasi

Kasus Prater, Jones, dan Borden ditampilkan dalam investigasi tingkat tinggi yang dilakukan oleh ruang berita nirlaba ProPublica yang mempertanyakan bias rasial dalam alat penilaian risiko kriminal yang banyak digunakan. Karya ini merupakan bagian dari literatur yang lebih luas yang telah menyoroti isu-isu keadilan dan diskriminasi dalam AI.

Di permukaan, sistem AI tampaknya menjamin imparialitas. Dengan mengurangi dampak prasangka manusia, sistem ini menjanjikan masa depan di mana hasil semata-mata didasarkan pada data dan oleh karena itu bebas dari bias. Lagipula, apa cara yang lebih baik untuk menghilangkan diskriminasi dari proses pengambilan keputusan selain dengan menyingkirkan manusia dari persamaan?

Kenyataannya jauh lebih rumit. Apa jenis kelamin, ras, kelas, agama, dan kebangsaan dari mereka yang menetapkan agenda AI? Siapa yang mendanai penelitian AI? Siapa yang melaksanakannya? Saat ini, hanya sekitar 12 persen peneliti AI di dunia adalah perempuan, dengan variasi di setiap negara berkisar dari 26% di Taiwan hingga nol partisipasi di Swedia. Perpecahan gender profesor AI di beberapa universitas ilmu komputer terkemuka di dunia mencerminkan pola yang serupa rata-rata, lebih dari 80 persen profesor adalah laki-laki. Apa peran gender peneliti AI dalam menentukan hasil penelitian AI? Pertanyaan serupa juga dapat diajukan tentang variasi kapabilitas AI di seluruh kawasan dan apakah kita memperkirakan adanya tantangan dalam membentuk pola AI global melalui penelitian dan kajian di beberapa negara. Jawaban atas pertanyaan-pertanyaan ini memengaruhi masa depan AI dan keadilan serta inklusivitas di masa depan tersebut.

Ketergantungan AI pada data "dunia nyata" dalam jumlah besar juga dapat menyebabkan unsur ketidakadilan dalam hasilnya. Penelitian telah menunjukkan bahwa AI dapat melanggengkan dan bahkan memperburuk bias yang ada atau bentuk diskriminasi lainnya. Penelitian telah menemukan dua penyebab utama fenomena ini:

Data Pelatihan yang Bias

Data pelatihan sangat penting dalam sistem berbasis pembelajaran mesin. Jika data yang digunakan untuk "mengajarkan" sistem jawaban yang "benar" bias atau tidak memadai, keterbatasan ini akan tertanam dalam sistem AI. Data yang diberikan kepada sistem AI mungkin tidak tersaring, tidak bersih, atau tidak representatif.

Misalnya, jika terdapat bias pengambilan sampel dan sebagian populasi terwakili secara keliru, hasil untuk kelompok tersebut akan kurang valid. Contohnya termasuk insiden algoritma pengenalan wajah Google pada tahun 2015 yang keliru melabeli orang kulit hitam sebagai gorila dan tingkat kesalahan yang lebih tinggi untuk perempuan berkulit gelap pada alat pengenalan wajah yang tersedia secara komersial dari Microsoft, IBM, dan Face++. Masalah dalam kedua kasus ini bermula dari kurangnya representasi kelompok tertentu dalam set data pelatihan.

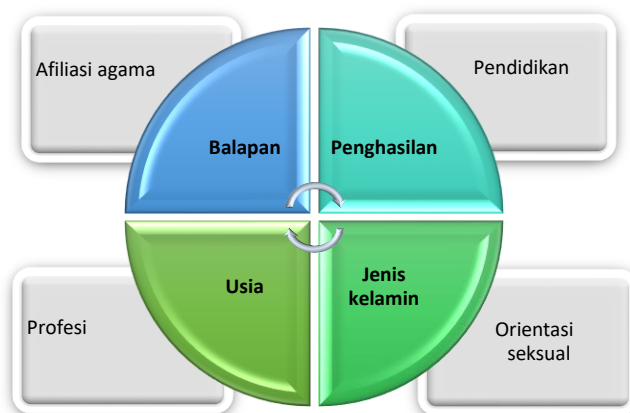
Data juga dapat mereproduksi bias historis. Misalnya, jika data pelatihan untuk mengevaluasi kandidat yang akan mengisi posisi manajemen di sebuah perusahaan besar didasarkan pada kandidat pria yang berhasil dari tahun-tahun sebelumnya, hasilnya kemungkinan akan bias terhadap kandidat perempuan. Penelitian menunjukkan bahwa pencarian untuk "ilmu komputer" jauh lebih mungkin menghasilkan kandidat pria hanya karena data tersebut mengandung lebih banyak referensi nama yang terdengar seperti laki-laki jika digabungkan dengan istilah "ilmu komputer".

Ketimpangan dalam Hasil

Dalam sistem dengan kapasitas, risiko, dan keadaan yang terdistribusi secara tidak merata, hasil diskriminatif juga dapat terwujud secara tidak langsung. Misalnya, jika suatu algoritma bergantung pada faktor-faktor "netral" seperti tingkat pendidikan, tingkat pendidikan orang tua, situasi keuangan, dll. untuk menghasilkan skor penilaian risiko kriminal seseorang, secara teoritis algoritma tersebut "tidak bias". Namun, karena keadaan yang tidak setara, hasilnya mungkin masih berkorelasi erat dengan keanggotaan kelompok-kelompok tertentu yang secara historis kurang beruntung. Orang kulit hitam akan memiliki skor risiko yang lebih tinggi meskipun data rasial tidak digunakan.

Ketimpangan hasil memang nyata. Sistem AI yang digunakan untuk mengevaluasi persetujuan pinjaman rumah di dua lingkungan, satu lingkungan berpenghasilan menengah dan satu lingkungan berpenghasilan rendah, dapat menyebabkan dampak yang tidak setara meskipun lokasi tidak dianggap sebagai faktor. Seseorang dari lingkungan berpenghasilan menengah lebih mungkin memiliki gaji yang lebih tinggi, sehingga mendapatkan batas pinjaman yang lebih tinggi, dibandingkan seseorang dari lingkungan berpenghasilan rendah. Ini berarti lingkungan yang lebih kaya akan mendapatkan persyaratan kredit yang lebih menguntungkan, sehingga memperburuk disparitas yang sudah ada.

Meskipun sumber langsung dari masalah ini berada di luar ranah AI, contoh-contoh ini menyoroti beberapa risiko dalam mengadopsi solusi teknologi tanpa memperhatikan konteks sosial-ekonomi yang mendasarinya.



Pertanyaan tentang keadilan dalam lingkungan sosial dapat muncul dalam berbagai dimensi, termasuk pendapatan, gender, pendidikan, profesi, usia, ras, dan afiliasi agama. Interaksi AI dengan beberapa faktor ini, seperti gender dan ras, telah diteliti dengan cukup baik. Dengan menggunakan teknik pembelajaran mesin pada teks

yang dikumpulkan daring, para peneliti menemukan bahwa nama perempuan lebih terkait dengan keluarga daripada kata-kata karier, dibandingkan dengan nama laki-laki.

Lebih lanjut, kata-kata seperti "perempuan" dan "gadis" lebih mungkin dikaitkan dengan seni daripada matematika atau sains, dan bias gender-pekerjaan yang umum sangat berkaitan dengan partisipasi pekerjaan yang sebenarnya. Asosiasi gender ini kemudian membentuk hasil di sejumlah domain tempat sistem AI diterapkan, mulai dari pencarian kerja hingga penggunaan alat penerjemahan bahasa.

Bagan di bawah ini menggambarkan bagaimana persepsi gender dari dunia nyata tercermin dalam alat penerjemahan berbasis AI. Ketika Google Translate digunakan untuk menerjemahkan teks dari bahasa netral gender seperti Indonesia dan Finlandia ke bahasa yang bergender seperti Inggris, Google juga cenderung menerjemahkan persepsi tertentu tentang subjek, berdasarkan gendernya. Misalnya, ketika referensi ke subjek laki-laki dengan atribut negatif seperti malas diterjemahkan dari bahasa Inggris ke bahasa Indonesia dan kemudian kembali ke bahasa Inggris, keluaran yang dihasilkan mengubah referensi tersebut menjadi subjek perempuan. Kebalikannya berlaku untuk karakteristik positif seperti kecerdasan, di mana proses penerjemahan algoritmik mengubah referensi gender dari perempuan menjadi laki-laki.

Tantangan bias dan diskriminasi menjadi semakin relevan dalam diskusi AI. Misalnya, Deklarasi Toronto yang dirilis pada Mei 2018 oleh koalisi kelompok hak asasi manusia dan perusahaan teknologi berfokus pada kesetaraan dan non-diskriminasi dalam sistem pembelajaran mesin. Demikian pula, draf "Pedoman Etika untuk AI yang Dapat Dipercaya" yang diterbitkan oleh Kelompok Pakar Tingkat Tinggi Komisi Eropa tentang Kecerdasan Buatan pada Desember 2018, mengidentifikasi "non-diskriminasi" sebagai salah satu konsep kunci untuk mewujudkan AI yang dapat dipercaya. Selain itu, kerangka kerja tata kelola AI model yang diusulkan oleh Komisi Perlindungan Data Pribadi Singapura juga berupaya untuk memastikan keadilan dan non-diskriminasi.

Inggris → Indonesia → Inggris

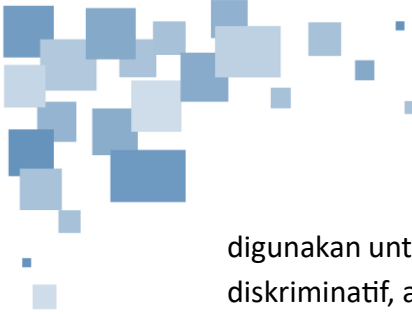
He Is Lazy	Dia Malas
Dia Malas	She's Lazy
She Is Smart	Dia Pintar
Dia Pintar	He's Smart

Inggris → Finlandia → Inggris

He Is a nurse	<i>Hän on sairaanhoitaja</i>
<i>Hän on sairaanhoitaja</i>	She's a nurse
She Is an engineer	<i>Hän on insinööri</i>
<i>Hän on insinööri</i>	He Is an engineer

5.2 RISIKO DAN MANFAAT

Kekuatan dan pengaruh sistem AI terus berkembang seiring dengan semakin seringnya sistem tersebut mengambil keputusan tentang aspek-aspek penting atau intim dalam kehidupan kita sehari-hari. Tantangan terhadap keadilan muncul ketika sistem tersebut



digunakan untuk membuat keputusan yang tidak kita antisipasi, atau keputusan tersebut bias, diskriminatif, atau sekadar "tidak adil".

Dampaknya menjadi sangat serius ketika pengambilan keputusan oleh sistem AI meluas ke aspek-aspek fundamental kehidupan kita, seperti apakah kita dipanggil untuk wawancara berdasarkan penilaian sistem AI terhadap resume kita, apakah pengajuan hipotek disetujui berdasarkan evaluasi sistem AI terhadap kelayakan kredit kita, atau sistem tersebut digunakan untuk memprediksi risiko terpidana melakukan kejahatan lagi. Di bagian ini, kita akan melihat secara singkat risiko yang ditimbulkan di beberapa area penting di mana AI diterapkan.

Penilaian Risiko dan Penjatuhan Hukuman dalam Peradilan Pidana

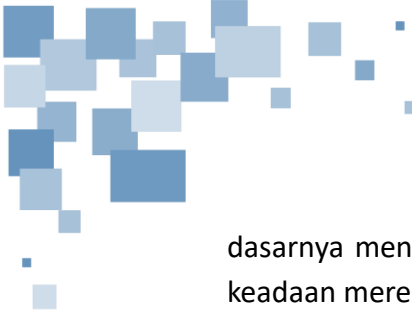
Semakin banyak negara bagian AS yang menggunakan sistem AI untuk melakukan penilaian risiko di setiap tahap sistem peradilan AS, sebagian berdasarkan dorongan dari Institut Pemasarakatan Nasional Departemen Kehakiman AS. Penilaian ini dapat digunakan pada tahap awal proses pidana, misalnya untuk membantu menentukan jumlah jaminan. Namun, di banyak negara bagian, penilaian ini juga diberikan kepada hakim selama proses penjatuhan hukuman untuk menilai risiko terdakwa akan melakukan kejahatan di masa mendatang.

Daya tarik sistem AI ini jelas: sistem ini menjanjikan informasi yang lebih akurat dan objektif yang memungkinkan sistem peradilan untuk "melakukannya dengan benar" dengan cara yang lebih hemat biaya, memenjarakan individu-individu yang lebih mungkin melakukan kejahatan di masa mendatang sambil memberikan hukuman yang lebih ringan atau pembebasan bersyarat kepada mereka yang kecil kemungkinannya untuk melakukan pelanggaran. Pertanyaannya adalah apakah sistem ini benar-benar "melakukannya dengan benar" dengan cara yang adil dan non-diskriminatif.

Sayangnya, dan seperti yang diilustrasikan di awal bab ini, hal ini tampaknya tidak selalu terjadi. Dalam sebuah studi yang dilakukan oleh ProPublica terhadap salah satu sistem penilaian risiko AI yang paling banyak digunakan, organisasi tersebut menyimpulkan bahwa sistem AI tersebut mendiskriminasi terdakwa berkulit hitam. Studi tersebut menemukan bahwa jumlah terdakwa berkulit hitam yang salah diidentifikasi berisiko tinggi untuk mengulangi tindak pidana dua kali lebih tinggi daripada jumlah terdakwa berkulit putih yang diklasifikasikan demikian.

Serupa dengan alat prediktif AI yang digunakan untuk membuat profil Prater, Borden, dan Jones, terdapat model residivisme lain yang membantu dalam penjatuhan hukuman prediktif dikenal sebagai *Level of Service Inventory-Revised* (LSI-R) yang digunakan untuk memastikan risiko terpidana mengulangi tindak pidana. Model ini menyediakan panduan penjatuhan hukuman bagi hakim yang mencakup hasil kuesioner panjang tentang riwayat kriminal individu.

Salah satu faktor dalam model ini mengaitkan 'keterlibatan' awal dengan polisi sebagai sinyal residivisme, yang berarti bahwa individu Afrika-Amerika dan Latin jauh lebih mungkin dikategorikan sebagai orang berisiko tinggi. Meskipun tampak efektif dan berpotensi lebih akurat daripada estimasi hakim, model yang menggunakan penilaian historis tersebut pada



dasarnya mengunci individu dalam siklus toksik dengan membuat profil orang berdasarkan keadaan mereka.

Pemolisian Prediktif

Bukankah hebat jika polisi dapat mengetahui sebelumnya siapa yang mungkin melakukan kejahatan atau mungkin menjadi korban kejahatan? Meskipun banyak yang percaya hal ini sudah dimungkinkan berkat perangkat pemolisian prediktif berbasis AI terbaru, para kritikus khawatir bahwa perangkat tersebut mungkin sarat dengan bias kuno dan kurangnya transparansi.

Sistem pemolisian prediktif biasanya melibatkan identifikasi pelaku kriminal, area di mana kejahatan kemungkinan terjadi atau calon korban, beserta mekanisme peringatan atau pencegahan yang sesuai. Meskipun secara teori hal ini terdengar baik, terdapat beberapa kekhawatiran. Algoritma yang digunakan untuk kepolisian prediktif cenderung bersifat rahasia, sehingga menutup peluang untuk debat publik mengenai kelebihan atau kekurangan metrik tertentu. Namun, dapat diasumsikan bahwa alat tersebut cenderung mempertimbangkan faktor-faktor seperti riwayat kriminal masa lalu, status pembebasan bersyarat, dan afiliasi saat menilai risiko kriminal. Hal ini dapat meningkatkan kemungkinan seseorang dimasukkan dalam "daftar panas" polisi berdasarkan "penampilan mencurigakan" atau hubungannya dengan komunitas atau lingkungan tertentu. Hal ini juga dapat menyebabkan penegakan hukum anti-teroris yang bias di komunitas dengan populasi Timur Tengah yang besar atau sejarah kejahatan terorganisir.

Kumpulan data yang bias juga dapat berasal dari data historis yang diwarnai oleh perilaku diskriminatif yang telah berlangsung lama terhadap ras, agama, atau minoritas lainnya. Data kejahatan dari Amerika Serikat dan India menunjukkan bahwa populasi penjara orang-orang dari kelompok terpinggirkan dan kurang beruntung kasta terjadwal, suku terjadwal, dan Muslim di India serta kelompok Afrika-Amerika dan Hispanik di Amerika Serikat cenderung melebihi representasi aktual mereka dalam populasi.

Solusi AI yang dibangun berdasarkan data ini memiliki kemungkinan untuk semakin melestarikan prasangka yang dihadapi oleh kelompok-kelompok ini. Penelitian juga menunjukkan masalah penggunaan "data kotor", yaitu data yang telah dirusak atau dimanipulasi secara sengaja, untuk melatih algoritma prediktif. Hal ini dapat, misalnya, mencakup data terkait penangkapan yang bermotif atau pelaporan yang kurang dari semestinya atas kejahatan tertentu, yang akan menjadi dasar pengambilan keputusan di masa mendatang.

Penggunaan teknologi yang meluas seperti televisi sirkuit tertutup (CCTV), pelacakan GPS, dan pengenalan wajah untuk memfasilitasi analisis kejahatan prediktif juga menimbulkan ancaman signifikan terhadap privasi, ekspresi, dan kebebasan sipil lainnya. Teknologi pengawasan semacam itu cenderung berdampak tidak proporsional terhadap komunitas yang kurang beruntung secara ekonomi atau sosial, sehingga menambah unsur ketidakadilan dalam penerapan alat-alat tersebut.

Kesehatan

AI mendorong inovasi medis dalam berbagai cara. Kegunaannya meliputi diagnosis medis menggunakan sistem bertenaga AI, bedah robotik, dan penemuan obat yang lebih cepat, yang memungkinkan perusahaan farmasi untuk secara signifikan mengurangi waktu yang biasanya dibutuhkan untuk memasarkan obat, dengan manfaat sosial yang nyata. Secara paralel, mengurangi biaya penemuan obat akan memungkinkan perusahaan rintisan obat baru yang bermodal lebih sedikit untuk memasuki pasar farmasi.

Desain Obat

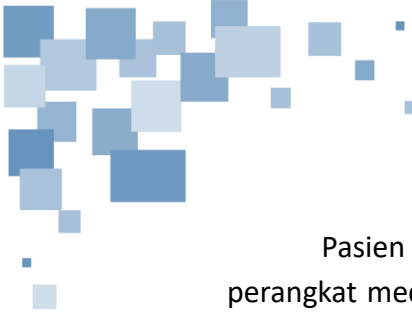
Di Inggris, Universitas Manchester telah mengembangkan alat AI bernama "Eve" yang dapat mengotomatiskan tahap awal desain obat, dengan menyaring lebih dari 10.000 senyawa per hari dan mengidentifikasi kandidat yang sesuai untuk pengujian obat.

Namun, penting untuk menyadari bahwa data kesehatan sangat personal dan membawa risiko signifikan berupa hasil yang tidak akurat atau diskriminatif. Hal ini terutama berlaku dalam situasi ketika data pelatihan kesehatan dari satu populasi digunakan untuk menginformasikan solusi AI yang diterapkan dalam konteks yang sama sekali berbeda. Sebagai contoh, sebuah perusahaan rintisan yang berbasis di Toronto menemukan bahwa tes pendengaran yang dikembangkannya untuk mendeteksi penyakit neurologis seperti Alzheimer atau Parkinson tidak bekerja dengan baik untuk orang yang bukan berasal dari Kanada.

Hal ini karena algoritma telah dilatih berdasarkan data yang dikumpulkan dari penutur asli bahasa Inggris di provinsi Ontario, Kanada, yang berbicara dalam bahasa ibu mereka. Data tersebut tidak memperhitungkan kosakata, jeda, dan pengucapan penutur asli bahasa Inggris, sehingga meningkatkan kemungkinan diagnosis yang salah.

Studi juga telah menunjukkan penggunaan sistem AI yang efektif untuk mendiagnosis lesi kulit ganas. Hal ini, sekali lagi, memiliki keterbatasan karena kumpulan data tersebut sebagian besar menampilkan kulit yang lebih terang, karena orang kulit putih lebih mungkin menderita penyakit ini. Kekhawatiran serupa juga muncul terkait kurangnya keragaman dalam rekam medis, dan ketergantungan yang berlebihan pada *Watson for Oncology* milik IBM untuk pendekatan medis di Amerika Serikat. Oleh karena itu, kurangnya data medis yang representatif kemungkinan akan tetap menjadi tantangan utama dalam penerapan solusi AI kesehatan.

Adopsi algoritma AI yang lebih luas di sektor kesehatan juga dapat mengaburkan batasan antara apa yang dianggap sebagai perangkat medis dan apa yang tidak. Khususnya, di yurisdiksi tertentu, seperti Uni Eropa dan Amerika Serikat, terdapat tren yang nyata untuk menganggap algoritma yang digunakan dalam industri medis sebagai perangkat medis. Faktanya, pada tanggal 7 Desember 2017, Pengadilan Keadilan Uni Eropa menerbitkan keputusannya pada kasus C-329/16, "Snitem dan Philips, Prancis" di mana Pengadilan memutuskan bahwa perangkat lunak yang berdiri sendiri dapat dianggap sebagai perangkat medis bahkan jika itu sendiri tidak bertindak di dalam atau pada tubuh manusia.



Pasien di luar yurisdiksi di mana perangkat lunak yang berdiri sendiri dianggap sebagai perangkat medis dan, akibatnya, tunduk pada persyaratan keselamatan tambahan, mungkin dilindungi dengan kurang ketat. Kecuali jika upaya terkoordinasi dilakukan untuk menyelaraskan disiplin perangkat lunak medis yang berdiri sendiri di seluruh yurisdiksi, vendor perangkat lunak medis dapat memutuskan untuk menguji produk mereka di yurisdiksi di mana peraturan lebih longgar. Hal ini dapat mengakibatkan hasil yang diskriminatif bagi pasien yang tidak akan memiliki kesempatan yang sama untuk mengakses perawatan berdasarkan definisi perangkat medis di yurisdiksi mereka.

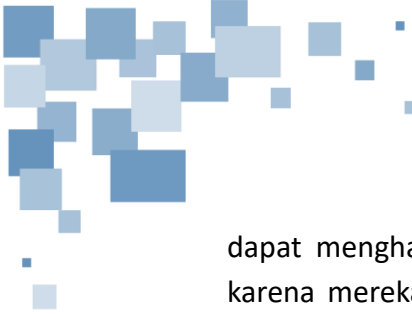
Pengenalan Wajah

Pengenalan wajah merupakan bidang penelitian yang telah ada sejak awal komputasi. Awalnya, hal ini tidak menimbulkan kekhawatiran besar, karena algoritma awal hanya berfokus pada upaya mendeteksi apakah suatu gambar kompleks mengandung wajah manusia. Perangkat lunak pengenalan kini memiliki kemampuan tidak hanya untuk mengidentifikasi apakah seseorang terdapat dalam suatu gambar, tetapi juga untuk mempertimbangkan usia, etnis, dan jenis kelamin, mengidentifikasi emosi dan gerakan, serta menentukan identitas individu tertentu. Sebagaimana yang terjadi pada karakter Tom Cruise dalam film fiksi ilmiah tahun 2002, *Minority Report*, sistem pengenalan wajah memungkinkan deteksi dan pengenalan individu melalui foto, video, atau pengawasan waktu nyata. Research and Markets memprediksi bahwa pasar pengenalan wajah global, yang diperkirakan mencapai Rp 38,5 triliun pada tahun 2017, akan tumbuh menjadi Rp 97,8 triliun pada tahun 2023 seiring teknologi ini merambah sektor keamanan, pengawasan, pasar ritel (misalnya untuk memindai pencuri toko, memantau keramaian, atau sekadar mencari pelanggan yang "tidak puas"), dan sebagainya.

Kekuatan AI didukung oleh banyaknya gambar yang diunggah ke media sosial, seperti yang terlihat dari "saran tag" di Facebook setiap kali gambar diunggah. Basis data gambar yang luas ini memungkinkan algoritma pembelajaran mesin untuk memproses dan meningkatkan perangkat lunak mereka agar lebih akurat dalam menentukan identitas individu tertentu bahkan jika individu tersebut berada di antara sekelompok besar orang.

Kemajuan teknologi pengenalan wajah memiliki manfaat sekaligus risiko. Kemampuan untuk meninjau data kompleks dalam jumlah besar secara real-time untuk mengidentifikasi individu dapat sangat bermanfaat. Pada tahun 2018, sistem pengenalan wajah membantu polisi mengidentifikasi seorang individu yang bertanggung jawab atas penembakan massal di Annapolis, Maryland, di *Capital Gazette*, yang menewaskan lima jurnalis. Namun, terdapat beberapa isu etika dan hukum yang terkait dengan penggunaan teknologi AI. Dalam kasus penembakan di Annapolis, tersangka diidentifikasi berdasarkan basis data foto SIM. Basis data ini mungkin telah ada selama beberapa dekade, dan masyarakat umum kemungkinan besar tidak menyadari bahwa data pribadi mereka akan digunakan untuk tujuan ini, yang menimbulkan masalah persetujuan jika yurisdiksi menganggap gambar seseorang sebagai data pribadi yang sensitif.

Lebih lanjut, teknologi pengenalan wajah juga menimbulkan isu etika yang saat ini tidak dipertimbangkan oleh undang-undang. Misalnya, telah dikemukakan bahwa hilangnya privasi



dapat menghambat kemampuan individu untuk berpartisipasi aktif dalam wacana politik karena mereka tidak lagi memiliki kemampuan untuk mengungkapkan pandangan mereka tanpa setiap tindakan atau pernyataan berpotensi dipantau. “*Social Credit System*” (“SCS”) yang diusulkan Tiongkok merupakan contoh sejauh mana teknologi ini dapat dikembangkan. SCS Tiongkok terus memantau warganya dengan menggunakan pengawasan massal, pengumpulan data, dan infrastruktur serta teknologi pemrosesan data. Dengan teknologi ini, perilaku setiap individu dapat dinilai berdasarkan persepsi “kebaikan komunitas”, “kebaikan sosial”, atau “ketulusan” sebagaimana yang disebut Tiongkok. Kebebasan berbicara dan otonomi seseorang dapat dirusak secara serius oleh penggunaan teknologi AI ini.

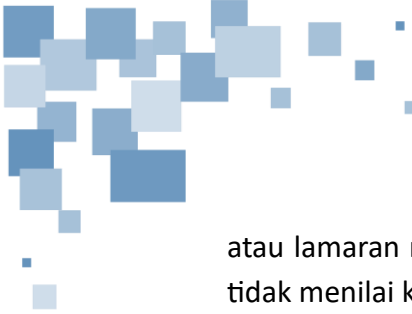
Risiko etika dan praktis utama lainnya adalah bias, terutama dampak bias rasial dan gender ketika teknologi ini digunakan oleh penegak hukum, seperti kesalahan identifikasi tersangka dan korban dari kelompok demografi tertentu. Misalnya, sebuah studi oleh *American Civil Liberties Union* yang melibatkan teknologi pengenalan wajah Amazon (*Amazon Rekognition*), yang sudah digunakan oleh lembaga penegak hukum di AS, menunjukkan bahwa algoritma tersebut salah mencocokkan gambar beberapa anggota Kongres Amerika Serikat dengan orang yang ditangkap. Selain itu, dari 28 kecocokan yang salah, hampir empat puluh persennya adalah orang kulit berwarna. Kekhawatiran serupa juga dapat muncul dalam kasus bentuk pengenalan lain berdasarkan karakteristik fisik, seperti pengenalan gaya berjalan, yang sering digunakan untuk tujuan pengawasan dan pengenalan suara, yang dapat mengecualikan aksen dan etnis yang unik.

Inheren dalam sifat belajar mandiri sistem AI, kita dapat mengantisipasi bahwa akurasi sistem pengenalan wajah akan meningkat. Namun, sulit untuk menentukan mana yang lebih mengkhawatirkan: potensi penggunaan sistem pengenalan wajah dengan tingkat akurasi yang tidak dapat diterima yang menciptakan bias dan masalah diskriminasi, atau dugaan bahwa peningkatan segera pada sistem CCTV dan kamera polisi yang menyediakan citra definisi tinggi akan menghasilkan sistem pengenalan wajah yang lebih akurat, andal, efektif, dan berpotensi menimbulkan kekhawatiran yang melampaui harapan kita dan melanggar persepsi kita tentang keadilan.

Tenaga Kerja

Salah satu manfaat potensial sistem AI adalah mengurangi biaya pengambilan keputusan terkait sumber daya manusia dan rekrutmen. Penggunaan AI juga memungkinkan perusahaan untuk membuat keputusan yang tampak objektif dan efisien saat merekrut atau mempromosikan karyawan, yang mampu mencocokkan kandidat berdasarkan keterampilan, alih-alih bias demografis atau pribadi. Namun, terdapat risiko bahwa algoritma tersebut dapat melanggengkan bias yang ada atau membentuk bias baru berdasarkan kumpulan data awal yang diberikannya.

Pengalaman Amazon dengan program AI yang ingin dikembangkannya untuk meninjau resume pelamar kerja pada tahun 2014 menggambarkan masalah ini. Amazon berupaya mengembangkan program, sebagai bagian dari proses identifikasi talenta terbaik, yang akan memberikan skor kepada pelamar, mulai dari satu hingga lima bintang, berdasarkan resume



atau lamaran mereka. Namun, pada tahun 2015, Amazon menyadari bahwa sistem tersebut tidak menilai kandidat untuk posisi teknis secara netral gender.

Sistem tersebut dilatih berdasarkan resume yang telah dikirimkan ke Amazon selama periode sepuluh tahun, yang sebagian besar, mencerminkan dominasi laki-laki di industri teknologi, adalah laki-laki. Akibatnya, sistem AI tersebut melatih dirinya sendiri untuk lebih menyukai kandidat laki-laki, tampaknya dengan memberikan skor yang lebih rendah pada lamaran atau resume yang menyertakan kata "perempuan" atau menunjukkan bahwa kandidat tersebut pernah kuliah di perguruan tinggi khusus perempuan.

Pengalaman Amazon menggambarkan tantangan terhadap keadilan yang dihadirkan oleh data pelatihan. Tidak mudah untuk melihat set data pelatihan, sebelum faktanya, untuk menentukan apakah data tersebut bias dalam beberapa hal. Dan, setelah kejadian tersebut, diskriminasi dalam sistem AI yang dihasilkan dari data pelatihan yang miring mungkin sulit dideteksi, karena hanya akan terlihat ketika kinerja sistem dievaluasi pada tingkat global dan tidak terlihat ketika melihat hasil individual, meskipun di situlah dampaknya akan terasa.

Mungkin juga terdapat hasil diskriminatif ketika AI menggantikan pekerjaan manusia. Bayangkan sebuah perusahaan pusat panggilan yang mengadopsi platform AI Google sudah mengembangkan asisten AI bernama Duplex yang dapat melakukan panggilan dan pemesanan untuk manusia tanpa penelepon lain mengetahui bahwa ia tidak sedang berbicara dengan manusia. Hal ini mungkin memiliki efek lokal yang khusus pada komunitas di mana pusat panggilan tersebut berada, karena akan menyebabkan berkurangnya pekerjaan berketerampilan rendah yang tersedia. Sebaliknya, hal ini dapat menyebabkan terciptanya pekerjaan baru yang menuntut tingkat keterampilan yang lebih tinggi.

Hal ini, pada gilirannya, akan memiliki efek pemerataan. Posisi berketerampilan rendah dan tingkat pemula lebih mungkin dipegang oleh kelompok yang secara historis kurang beruntung. Transformasi yang disebabkan oleh efek disruptif AI dalam ketenagakerjaan, khususnya dalam hal kesenjangan keterampilan yang dihadapi oleh pekerja yang terdampak oleh otomatisasi dan sifat pekerjaan baru yang diciptakan oleh industri AI, perlu dipantau.

Asuransi

Selama beberapa tahun terakhir, perusahaan asuransi semakin banyak menggunakan perangkat AI. Industri asuransi memiliki ekspektasi tinggi terhadap potensi transformatif positif AI terhadap bisnis mereka. Perusahaan asuransi perlu mengadopsi mekanisme penilaian risiko yang efisien untuk memitigasi dampak asimetri informasi, dan AI menjanjikan menjadi alat yang tepat untuk mencapai hal ini. AI dapat membantu menilai risiko dan menentukan harga premi asuransi. Khususnya, pengumpulan data granular dari sejumlah sumber digital, termasuk media sosial, telah memungkinkan perusahaan asuransi untuk menyusun premi asuransi yang dipersonalisasi berdasarkan profil individu. Sekilas, penggunaan AI untuk menilai risiko dan menentukan harga asuransi tampak ideal. Namun, jebakan yang dibahas di atas terkait AI juga menimbulkan risiko dalam penggunaannya di industri asuransi. Akibat penggunaan AI, penyebaran premi yang lebih luas antara individu berisiko tinggi dan berisiko rendah akan semakin menjadi norma. Premi bagi mereka yang dianggap berisiko rendah akan turun secara signifikan, sementara premi bagi mereka yang

diklasifikasikan berisiko tinggi akan melonjak. Sekali lagi, kelompok yang secara historis kurang beruntung cenderung berada di pihak yang kalah, tanpa adanya kelonggaran atau pertimbangan khusus bagi individu dalam kelompok tersebut.

Periklanan

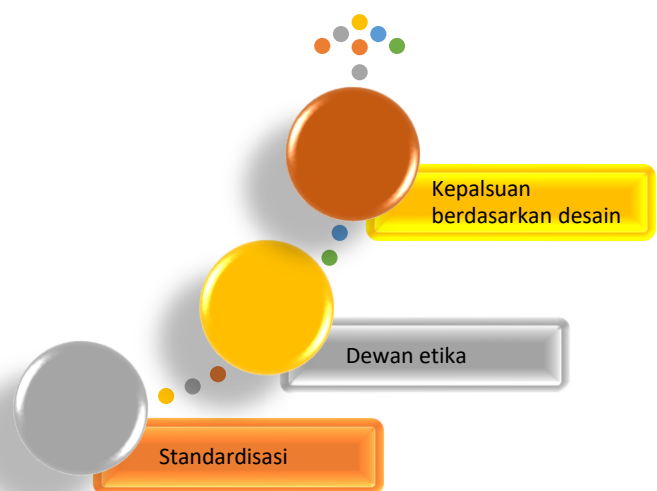
Penggunaan AI dalam periklanan daring juga dapat menimbulkan pertanyaan terkait diskriminasi. Sistem AI sering digunakan untuk periklanan daring. Namun, dampak penggunaan AI dalam periklanan telah menimbulkan kekhawatiran. Misalnya, sebuah studi tahun 2013 menunjukkan bahwa pencarian di Google untuk nama yang terdengar seperti orang Afrika-Amerika menghasilkan iklan yang menunjukkan seseorang memiliki catatan penangkapan. Sebaliknya, pencarian untuk nama yang terdengar seperti orang kulit putih menghasilkan hasil yang jauh lebih sedikit yang menunjukkan catatan penangkapan. Hasil ini tampaknya menunjukkan bahwa sistem AI mewarisi bias rasial. Demikian pula, sebuah eksperimen yang dilakukan dengan simulasi pengguna internet pria dan wanita menunjukkan bahwa pria diperlihatkan lebih banyak iklan yang terkait dengan pekerjaan bergaji tinggi daripada wanita.

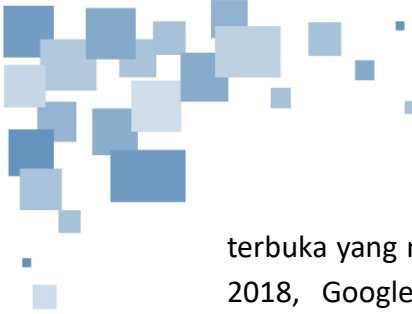
Selain itu, terdapat juga beberapa diskusi terkait perusahaan yang memungkinkan pengiklan memilih karakteristik tertentu untuk menargetkan iklan. Baru-baru ini, Facebook mengumumkan bahwa mereka akan berhenti mengizinkan pengiklan di bidang perumahan, pekerjaan atau kredit untuk menampilkan pesan mereka hanya kepada orang-orang dari ras, jenis kelamin atau kelompok usia tertentu, dalam upaya untuk mencegah praktik diskriminatif.

5.3 PRAKTIK DAN STANDAR INDUSTRI

Di tengah meningkatnya kesadaran akan risiko yang dapat ditimbulkan oleh penggunaan AI terhadap keadilan, memastikan non-diskriminasi mulai diakui sebagai salah satu tantangan utama yang dibawa oleh teknologi. Penerimaan terhadap masalah ini dan besarnya masalah ini relatif baru, tetapi masalah ini telah menimbulkan kekhawatiran yang signifikan. Area utama fokus industri adalah meningkatkan model, dan mencoba mengidentifikasi serta memberantas bias.

Beberapa perusahaan telah mulai mengembangkan alat untuk mengatasi diskriminasi. Misalnya, pada Mei 2018, Facebook mengumumkan telah mengembangkan perangkat lunak anti-bias bernama "*Fairness Flow*." Pada bulan Juli 2018, Microsoft Research menerbitkan sebuah penelitian yang menjelaskan algoritma keadilan untuk sistem klasifikasi biner, beserta pustaka Python sumber





terbuka yang mengimplementasikan algoritma tersebut. Selanjutnya, pada bulan September 2018, Google mengumumkan "alat What-If"-nya, yang mencakup penilaian keadilan. Seminggu kemudian, IBM merilis serangkaian alat visual yang dapat digunakan dengan model pembelajaran mesin apa pun.

Salah satu ide untuk menangani keadilan dalam AI adalah dengan mengupayakan standar atau tolok ukur yang berlaku di seluruh industri untuk algoritma. Inisiatif dalam arah ini mencakup buku putih tentang Standardisasi Kecerdasan Buatan yang diterbitkan pada Januari 2018 oleh sekelompok besar organisasi Tiongkok. Namun, hingga saat ini belum ada standar yang berlaku di seluruh industri terkait keadilan yang diadopsi.

Selain itu, dewan etik merupakan cara organisasi sektor publik dan swasta menangani berbagai isu kontroversial atau sensitif, termasuk lembaga penelitian untuk menilai keberterimaan etis suatu penelitian dan untuk menyetujui, menolak, mengusulkan modifikasi, atau menghentikan setiap penelitian yang diusulkan atau sedang berlangsung yang melibatkan manusia. Hal ini dapat diterapkan pada sistem AI. Misalnya, dewan etik AI Google, yang dibentuk ketika Google pertama kali mengakuisisi perusahaan Inggris DeepMind pada tahun 2014, merupakan contoh terkenal penggunaan dewan etik dalam konteks algoritmik. Dewan ini dibentuk dalam upaya untuk memastikan bahwa perangkat lunak self-thinking yang dikembangkan DeepMind dan Google tetap aman dan bermanfaat bagi umat manusia.

Dengan cara yang sama, bisnis sebaiknya mempertimbangkan pembentukan "Dewan Etika AI" untuk mengatasi tantangan privasi yang ditimbulkan oleh AI. Dewan Etika AI tersebut dapat diberi tanggung jawab untuk:

- (i) mengevaluasi proposal pengembangan sistem AI berdasarkan kriteria etika yang ditetapkan oleh bisnis;
- (ii) mengevaluasi sistem, setelah dibangun, untuk memvalidasi bahwa sistem yang dikembangkan selaras dengan kriteria yang menjadi dasar persetujuan pengembangan sistem; dan
- (iii) memantau penggunaan sistem yang berkelanjutan, terutama yang berkaitan dengan sistem pembelajaran mandiri, untuk memvalidasi bahwa sistem tersebut terus selaras dengan tujuan etikanya.

Para peneliti juga telah mengusulkan agar prinsip "keadilan berdasarkan desain" tertanam dalam semua penelitian AI. Hal ini mengharuskan konsepsi dan desain sistem AI dilakukan dengan cara yang mengutamakan keadilan, alih-alih menganggapnya sebagai persyaratan kepatuhan ex-post. Kerangka kerja seperti ini akan menekankan langkah-langkah seperti membangun tim lintas disiplin yang terdiri dari ilmuwan data dan ilmuwan sosial saat menerapkan solusi AI; mengidentifikasi dan mengatasi bias yang mungkin dimasukkan ke dalam data oleh anotator manusia; mengintegrasikan ukuran-ukuran keadilan ke dalam metrik penilaian program dan mengadopsi teknik-teknik de-biasing yang tepat. Faktor lain yang perlu dipertimbangkan adalah pentingnya membangun tim yang beragam dan inklusif untuk mencerminkan berbagai perspektif, kebutuhan, dan tujuan.

5.4 PERTIMBANGAN KEBIJAKAN

Isu keadilan dalam AI tidak mudah diatasi. Ini adalah masalah yang tidak hanya baru dan terus berkembang, tetapi juga sangat teknis. Beberapa hal sedang dibahas terkait AI, dan berbagai aspek kebijakan termasuk hubungan antara AI dan undang-undang antidiskriminasi, transparansi, dan pengawasan regulasi termasuk di antara yang sedang dikaji.

Hubungan antara Kecerdasan Buatan dan Undang-Undang Antidiskriminasi

Prinsip kesetaraan dan nondiskriminasi merupakan bagian dari fondasi negara hukum. Individu dilindungi dari diskriminasi oleh perjanjian internasional, serta oleh hukum dan peraturan nasional.

Undang-undang antidiskriminasi dapat mengatur bidang-bidang tertentu, seperti ketenagakerjaan, atau terkait dengan hukum perdata umum. Diskriminasi berdasarkan ras, asal etnis, jenis kelamin, agama, pandangan politik, disabilitas, usia, atau identitas seksual biasanya dilarang. Hingga saat ini, belum ada undang-undang antidiskriminasi yang secara eksplisit merujuk pada Kecerdasan Buatan. Namun demikian, sejauh peraturan tidak merujuk pada cara untuk mencapai keputusan diskriminatif, pada prinsipnya peraturan antidiskriminasi juga berlaku untuk keputusan berbasis Kecerdasan Buatan. Namun, dalam praktiknya, keberhasilan korban diskriminasi berbasis Kecerdasan Buatan dalam menggunakan upaya hukum terhadap keputusan diskriminatif akan sangat bervariasi dari satu sistem hukum ke sistem hukum lainnya. Secara khusus, hal ini akan sangat bergantung pada isu-isu prosedural, seperti siapa yang menanggung beban pembuktian.

AS

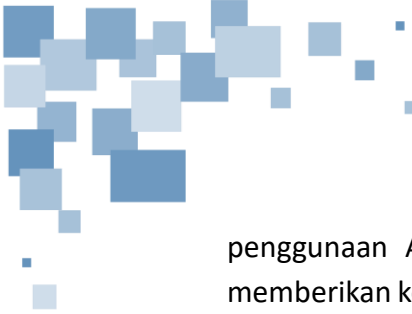
Di AS, diskriminasi ditangani oleh Undang-Undang Hak Sipil AS tahun 1964. Peraturan ini menetapkan prosedur tiga langkah untuk membuktikan diskriminasi dalam pekerjaan. Penggugat memiliki tanggung jawab awal untuk membuktikan adanya kasus prima facie diskriminasi. Kemudian, beban pembuktian beralih kepada pemberi kerja, yang harus memberikan dasar yang sah untuk keputusan tersebut. Jika pemberi kerja melakukannya, penggugat harus menunjukkan bahwa alasan yang diajukan bersifat dalih atau bahwa pemberi kerja dapat menggunakan praktik ketenagakerjaan alternatif dengan hasil yang kurang diskriminatif.

Jerman

Sebaliknya, Undang-Undang Umum tentang Perlakuan Setara Jerman ("AGG") menetapkan prosedur dua langkah. Penggugat harus menunjukkan dan membuktikan indikasi yang menimbulkan praduga bahwa diskriminasi telah terjadi. Jika hal ini telah dilakukan, tergugat perlu menunjukkan bahwa diskriminasi tidak terjadi. Kegagalan memberikan bukti tersebut mengarah pada keputusan yang menguntungkan penggugat.

Undang-Undang Hak Sipil AS memberikan beban yang lebih berat kepada penggugat, karena membuktikan adanya diskriminasi awal dan menolak penjelasan tergugat biasanya membutuhkan pengetahuan tentang cara kerja sistem AI. Sebaliknya, tampaknya berdasarkan Undang-Undang AGG Jerman, pemberi kerja berkewajiban untuk membuktikan bahwa AI tidak menghasilkan hasil yang diskriminatif.

Dengan asumsi bahwa tidak ada pihak yang dapat menjelaskan bagaimana sistem AI mencapai keputusan tertentu, kedua hasil tersebut tampak tidak memuaskan. Jika



penggunaan AI untuk membuat keputusan yang berpotensi diskriminatif secara praktis memberikan keleluasaan untuk melakukan diskriminasi, hal ini jelas tidak tepat. Hal yang sama berlaku jika kecurigaan adanya diskriminasi secara otomatis mengarah pada hukuman atas diskriminasi. Salah satu cara untuk menghindari hasil ini adalah dengan memberikan transparansi yang lebih besar terkait penggunaan AI. Hal ini dapat memungkinkan keputusan pengadilan yang lebih terinformasi dan adil.

Permasalahan ini menjadi semakin kompleks dalam kasus sistem hukum yang tidak memiliki kerangka kerja komprehensif yang mengatur praktik diskriminatif oleh aktor swasta. Tantangannya kemudian adalah pertama-tama menetapkan perlunya membangun perlindungan tersebut dalam hukum dan kemudian menganalisis persamaan dan perbedaan antara diskriminasi yang dilakukan oleh manusia dan sistem AI.

Transparansi


Sistem AI sering digambarkan sebagai kotak hitam, yang mencerminkan fakta bahwa logika spesifik yang digunakan sistem berbasis pembelajaran mesin, jaringan saraf, dan pembelajaran mandiri untuk mencapai hasilnya mungkin tidak transparan, atau mungkin tidak dapat dipahami, diprediksi, atau dijelaskan. Hal ini karena identifikasi cara metode algoritmik berfungsi dan membuat keputusan sangat penting dalam memberdayakan individu dan organisasi dengan kemampuan untuk mengidentifikasi kelemahan dan bias. Mencapai transparansi dalam pengambilan keputusan berbasis kode dan AI merupakan salah satu metode yang jelas untuk meningkatkan kepercayaan masyarakat terhadap sistem AI.

Prinsip-prinsip keadilan tidak berdiri sendiri, tetapi secara intrinsik terkait dengan transparansi. Meningkatkan transparansi algoritma dan data pelatihan merupakan salah satu tugas terpenting saat memanfaatkan sistem AI, dan sangat penting untuk mencegah diskriminasi. Transparansi dan kemudahan dijelaskan diperlukan untuk memastikan keadilan dan akuntabilitas.

Diskriminasi dan bias tidak dapat dimaafkan hanya karena sistem AI, dan bukan manusia, yang bertanggung jawab atas hasilnya. Namun, menilai hasil yang bias atau tidak adil, dan, jika perlu, menantangnya secara hukum, membutuhkan akses ke sistem AI dan sarana untuk memahami prosesnya. Individu, perusahaan swasta, dan lembaga publik yang menggunakan AI harus bertanggung jawab, dan ini mencakup reproduktifitas hasil dan transparansi mengenai sistem apa yang digunakan dan bagaimana penggunaannya, serta penjelasan tentang bagaimana sistem tersebut beroperasi dan menghasilkan hasilnya. Namun, sebagaimana dibahas lebih lanjut dalam pembahasan Transparansi dan Keterjelasan, prinsip terakhir ini terkadang bertentangan dengan nilai dan kepentingan lain. Dalam konteks ini, tantangan ke depan adalah menemukan keseimbangan yang melindungi dari diskriminasi sekaligus mempertimbangkan faktor-faktor lain seperti biaya, informasi kepemilikan, perlindungan kekayaan intelektual, dan efisiensi.

Algoritma yang tidak Bias, Data yang tidak Bias, dan Keberagaman Data

Sebagaimana ditunjukkan di atas, diskriminasi dalam sistem AI dapat disebabkan oleh algoritma yang bias, data pelatihan yang bias, dan kurangnya keberagaman dalam data



pelatihan. Untuk memerangi sumber-sumber bias ini, perlu dikembangkan dan diterapkan berbagai mekanisme.

Meningkatkan dan memastikan keberagaman data dapat menjadi cara untuk mengatasi salah satu sumber utama diskriminasi dalam AI, yaitu kurangnya representasi suatu kelompok tertentu. Dalam hal ini, memperkenalkan persyaratan keberagaman minimum untuk mencapai tingkat representasi yang sama dari berbagai kelompok sering diidentifikasi sebagai persyaratan keadilan dalam AI.

Persyaratan minimum untuk keberagaman data akan berupaya menjamin bahwa kualitas hasil akan serupa untuk berbagai kelompok populasi sebuah isu yang khususnya penting di sektor kesehatan. Kerangka kerja tata kelola AI model yang diusulkan oleh Komisi Perlindungan Data Pribadi Singapura, misalnya, mengakui bahwa mengidentifikasi dan menangani bias dalam kumpulan data bukanlah hal yang mudah, dan menyoroti pentingnya data yang heterogen. Untuk mencapai tujuan ini, entitas publik dan swasta harus bekerja sama dalam menciptakan basis data yang lebih komprehensif, dengan juga mempertimbangkan kebutuhan perlindungan data.

Seperti yang ditunjukkan oleh beberapa pakar, meningkatkan keberagaman bukanlah solusi otomatis untuk masalah diskriminasi. Sistem yang berbasis pada data yang beragam belum tentu menghasilkan hasil yang non-diskriminatif. Menggunakan kumpulan data yang berbeda untuk pelatihan, pengujian, dan validasi sistem AI dapat membantu mengidentifikasi bias. Selain itu, kumpulan data harus ditinjau dan diperbarui secara berkala. Namun, penerapan dan pengembangan metode serta perangkat lebih lanjut untuk mendeteksi bias dalam algoritma dan data pelatihan juga diperlukan.

Penelitian lebih lanjut diperlukan untuk mengembangkan mekanisme dan persyaratan tersebut, dan hasil penelitian tersebut dapat berupa standar atau aturan untuk pengembangan AI bebas bias. Karena istilah diskriminasi juga mengandung unsur subjektif, aturan tersebut juga harus mempertimbangkan hasil debat publik yang membahas topik-topik seperti tindakan afirmatif dan bentuk kompensasi lain untuk diskriminasi di masa lalu dan masa kini. Salah satu pilihannya adalah merancang proses yang harus dilalui setiap sistem AI baru. Proses ini dapat mencakup uji bias dan pemberian segel mutu resmi yang diakui secara internasional kepada sistem AI yang lulus uji.

Mengatasi Kesan Keadilan dan Edukasi dalam AI

Kesan objektivitas sistem AI juga penting. Kepercayaan buta terhadap AI yang didasarkan pada asumsi objektivitasnya merupakan hal yang harus diatasi. Bias AI seringkali tidak terlihat, sehingga lebih sulit untuk diperbaiki. Oleh karena itu, penting untuk memberi informasi dan edukasi, tidak hanya kepada pengembang tetapi juga pengguna, tentang batasan AI.

Oleh karena itu, "*Ethics Guidelines for Trustworthy AI*" Uni Eropa mengakui pentingnya edukasi dan kesadaran untuk menumbuhkan pola pikir etis. Dalam hal ini, pedoman tersebut menekankan bahwa memberikan edukasi mengenai sistem AI dan dampaknya sangatlah penting, dan mengidentifikasi upaya membangun kesadaran ini sebagai metode non-teknis untuk mencapai sistem AI yang lebih adil.

Peninjauan dan Pengujian Independen

Hasil sistem AI seringkali tidak dipantau atau tidak dinilai secara ketat atau independen untuk mendeteksi bias atau bahkan masalah yang lebih umum seperti hasil yang dipertanyakan atau tidak akurat. Misalnya, investigasi ProPublica tahun 2013 yang disebutkan sebelumnya mengkaji 19 metodologi penilaian risiko kriminal berbeda yang digunakan di Amerika Serikat dengan tujuan untuk menilai sejauh mana metodologi tersebut telah dievaluasi atau dinilai. Mereka menemukan bahwa "dalam kebanyakan kasus, validitas hanya diperiksa dalam satu atau dua studi" dan bahwa "seringkali investigasi tersebut diselesaikan oleh orang yang sama yang mengembangkan instrumen tersebut."

Kemampuan untuk meninjau dan menguji sistem AI terkait dengan prinsip transparansi. Kemampuan untuk memvalidasi hasil yang dihasilkan oleh sistem pengambilan keputusan AI rumit karena banyak sistem menggunakan algoritma kepemilikan yang pengembangnya, demi melindungi hak kepemilikan mereka, enggan mengungkapkannya bahkan untuk tujuan pengujian.

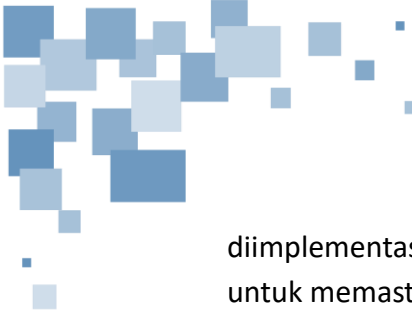
Pengawasan dan Regulasi

Organisasi dan pembuat kebijakan telah mulai memperhatikan kemungkinan perlunya pengawasan dan/atau regulasi terkait AI dan non-diskriminasi.

Perkembangan internasional dalam pengawasan dan regulasi
Pada tahun 2016, sebuah laporan dari Kantor Eksekutif Presiden Amerika Serikat memperingatkan bahwa dampak algoritma berbasis kecerdasan buatan terhadap pekerja berpotensi memperburuk ketimpangan, dan mencatat bahwa bias yang tertanam dalam kode komputer dapat merugikan individu di berbagai bidang.
Uni Eropa menekankan pentingnya non-diskriminasi pada tahun 2017 dalam sebuah resolusi tentang implikasi hak asasi manusia dari big data. Uni Eropa juga membentuk Kelompok Pakar Tingkat Tinggi tentang Kecerdasan Buatan, yang terdiri dari perwakilan dari akademisi, masyarakat sipil, serta industri. Tujuannya adalah untuk mendukung implementasi strategi Eropa tentang Kecerdasan Buatan, yang mencakup penyusunan rekomendasi tentang pengembangan kebijakan terkait masa depan dan tentang isu-isu etika, hukum, dan sosial yang berkaitan dengan AI, termasuk tantangan sosial-ekonomi.

Kelompok Pakar Tingkat Tinggi untuk Kecerdasan Buatan telah menerbitkan draf "Pedoman Etika untuk AI yang Tepercaya." Laporan ini bertujuan untuk menjadi panduan pengembangan "AI yang Tepercaya," sebuah konsep yang memiliki dua komponen: (1) pengembangan, penerapan, dan penggunaannya mematuhi hak-hak asasi dan peraturan yang berlaku serta menghormati prinsip dan nilai inti, memastikan tujuan yang etis; dan (2) secara teknis tangguh dan andal.

Dokumen tersebut menyatakan bahwa memastikan "tujuan etis" dalam AI memerlukan pendekatan yang berpusat pada manusia, yang didasarkan pada prinsip-prinsip kebaikan, non-maleficence, otonomi manusia, keadilan, dan eksplikabilitas. Berdasarkan pentingnya menjaga prinsip-prinsip ini, Kelompok Pakar Tingkat Tinggi untuk Kecerdasan Buatan mengidentifikasi serangkaian persyaratan untuk AI yang Tepercaya, yang harus



diimplementasikan sejak tahap desain paling awal. Konsep “Non-diskriminasi” sangat penting untuk memastikan prinsip keadilan, dan dengan demikian mencapai AI yang Tepercaya.

Deklarasi Toronto membahas AI dari kerangka kerja yang disediakan oleh hukum hak asasi manusia. Deklarasi Toronto berfokus pada hak asasi manusia atas kesetaraan dan non-diskriminasi, serta kerangka kerja internasional yang ada untuk melindunginya. Dari kewajiban positif yang dimiliki negara untuk melindungi hak asasi manusia, Deklarasi Toronto menurunkan prinsip-prinsip terkait pembelajaran mesin.

Deklarasi ini menyatakan bahwa negara memiliki kewajiban positif untuk melindungi dari diskriminasi oleh pelaku sektor swasta dan untuk mempromosikan kesetaraan, termasuk melalui pengawasan dan undang-undang yang mengikat. Lebih lanjut, Deklarasi Toronto menyatakan bahwa kewajiban ini juga berlaku untuk penggunaan AI oleh publik. Deklarasi ini menetapkan bahwa dalam sistem sektor publik, negara harus mengidentifikasi risiko, memastikan transparansi dan akuntabilitas, menegakkan pengawasan, dan mempromosikan kesetaraan.

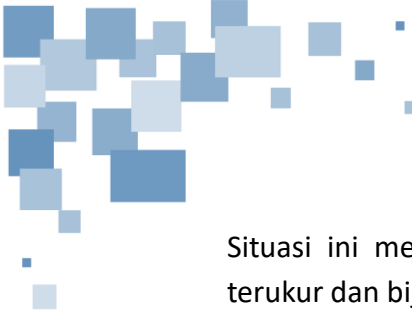
Dalam konteks perdebatan ini, beberapa gagasan dan proposal telah diajukan tentang cara memastikan AI yang adil. Dari sudut pandang regulasi, terkadang mencakup penerapan atau amandemen aturan pertanggungjawaban dan regulasi non-diskriminasi yang ada. Upaya hukum dalam bentuk tindakan yang layak untuk mendapatkan informasi, menuntut ganti rugi, atau mendapatkan penghentian tindakan yang merugikan, dapat dipertimbangkan.

Keterlibatan pemerintah dan organisasi masyarakat sipil dalam penyusunan standar industri yang harus dipenuhi sebelum sistem AI dirilis juga telah disarankan. Standar-standar ini dapat mengharuskan sistem dilatih dengan jumlah data yang sama untuk pengguna dari berbagai latar belakang ras dan gender, misalnya, dan menerapkan metode untuk mendeteksi bias dalam algoritma dan data pelatihan. Pengawasan dan pengendalian oleh badan pemerintah juga dapat diperdebatkan. Namun, desain intervensi semacam itu harus dipertimbangkan dengan mempertimbangkan keterbatasan kapasitas negara dalam mengatur sektor teknologi dan dampaknya terhadap inovasi di masa depan.

5.5 KEADILAN DAN NON-DISKRIMINASI

AI menjanjikan banyak keuntungan dan manfaat. Sekilas, dunia di mana keputusan didasarkan pada data dan hanya data saja terdengar seperti dunia yang efisien dan adil. Hal ini mendorong pemerintah dan organisasi industri untuk segera memanfaatkan peluang penerapan AI di berbagai bidang. Namun, terdapat risiko bias yang melekat dalam AI, dan ketergantungan yang berlebihan pada sistem AI dapat menyebabkan hasil yang tidak adil dan meresahkan. Sistem AI mencerminkan pengalaman kita sendiri, dan keadilan sama sekali tidak dijamin. Faktanya, memastikan keadilan dan non-diskriminasi membutuhkan pertimbangan dalam banyak variabel kompleks.

Jika kita mengkaji isu-isu yang diidentifikasi dalam bab ini pada tingkat paling dasar, isu-isu tersebut menggambarkan tema umum: kurangnya kendali manusia dan masyarakat atas serta pemantauan pengembangan, pengujian, implementasi, dan penggunaan sistem AI.



Situasi ini menimbulkan risiko-risiko berikut, yang harus ditangani melalui respons yang terukur dan bijaksana:

- (i) informasi pribadi dapat digunakan dengan cara yang tidak diantisipasi oleh individu dan di luar ekspektasi wajar individu;
- (ii) hasil yang dihasilkan oleh sistem AI dapat bias atau diskriminatif, sehingga menimbulkan dampak buruk yang tidak dapat dibenarkan pada individu; dan
- (iii) dampak dan efek dari risiko ini diperkirakan akan terus meningkat.

Beberapa tahun terakhir telah menunjukkan peningkatan kekhawatiran terkait keputusan berbasis AI. Kini terdapat kesadaran yang lebih besar bahwa bias dalam AI perlu ditangani, dan berbagai pihak telah mulai mengusulkan pendekatan untuk masalah ini. Sejauh masalah bias dan diskriminasi dalam AI bersumber dari kondisi sosial-ekonomi dunia nyata yang lebih luas, solusi yang tepat harus muncul melalui keterlibatan lintas disiplin antara Pemerintah, ilmuwan AI, ilmuwan sosial, dan masyarakat luas.

Prinsip

Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI dan undang-undang nasional yang mengatur penggunaannya wajib memastikan non-diskriminasi hasil AI, dan wajib mendorong langkah-langkah yang tepat dan efektif untuk menjaga keadilan dalam penggunaan AI.

Kesadaran dan Edukasi

- Kesadaran dan edukasi tentang kemungkinan dan batasan sistem AI merupakan prasyarat untuk mencapai hasil yang lebih adil.
- Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus mengambil langkah-langkah untuk memastikan bahwa pengguna menyadari bahwa sistem AI mencerminkan tujuan, pengetahuan, dan pengalaman penciptanya, serta keterbatasan kumpulan data yang digunakan untuk melatihnya.

Teknologi dan Keadilan

- Keputusan berdasarkan sistem AI harus adil dan non-diskriminatif, dinilai berdasarkan standar yang sama dengan proses pengambilan keputusan yang sepenuhnya dilakukan oleh manusia.
- Penggunaan sistem AI oleh organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI dan Pemerintah tidak boleh dimaksudkan untuk mengecualikan atau melemahkan kebutuhan akan keadilan, meskipun hal ini dapat berarti memfokuskan kembali konsep, standar, dan aturan yang berlaku untuk mengakomodasi AI.
- Pengguna sistem AI dan orang-orang yang menjadi subjek keputusan mereka harus memiliki cara yang efektif untuk mencari penyelesaian dalam situasi diskriminatif atau tidak adil yang ditimbulkan oleh sistem AI yang bias atau keliru, baik yang digunakan oleh organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI maupun pemerintah, dan untuk mendapatkan ganti rugi atas kerugian apa pun.



Pengembangan dan Pemantauan Sistem AI

- ❖ Pengembangan AI harus dirancang dengan mengutamakan keadilan. Hal ini melibatkan penanganan bias algoritma dan data sejak tahap awal dengan tujuan untuk memastikan keadilan dan non-diskriminasi.
- ❖ Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus tetap waspada terhadap bahaya yang ditimbulkan oleh bias. Hal ini dapat dicapai dengan membentuk dewan etik dan kode etik, serta dengan mengadopsi standar industri dan segel mutu yang diakui secara internasional.
- ❖ Sistem AI dengan dampak sosial yang penting dapat memerlukan peninjauan dan pengujian independen secara berkala.
- ❖ Dalam pengembangan dan pemantauan sistem AI, perhatian khusus harus diberikan kepada kelompok yang kurang beruntung yang mungkin tidak terwakili dengan benar dalam data pelatihan.

Pendekatan Komprehensif terhadap Keadilan

- Sistem AI dapat melestarikan dan memperparah bias, serta memiliki dampak sosial dan ekonomi yang luas di masyarakat. Menangani keadilan dalam penggunaan AI memerlukan pendekatan holistik. Secara khusus, pendekatan ini memerlukan:
 - i. keterlibatan erat para pakar teknis dari bidang terkait AI dengan para ahli statistik dan peneliti dari ilmu sosial; dan
 - ii. keterlibatan gabungan antara pemerintah, organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI, dan masyarakat luas.
- Prinsip Keadilan dan Non-Diskriminasi didukung oleh Prinsip Transparansi dan Akuntabilitas. Keadilan yang efektif dalam penggunaan sistem AI memerlukan penerapan langkah-langkah yang berkaitan dengan kedua Prinsip ini.



BAB 6

KESELAMATAN DAN KEANDALAN

Pengembangan dan implementasi teknologi berbasis AI akan disertai risiko yang melekat pada setiap teknologi baru. Kerentanan akan diremehkan, kesalahan akan terjadi, sistem akan gagal, dan kerusakan akan terjadi.

Aturan tentang keselamatan dan keandalan memengaruhi pengembangan dan penggunaan setiap teknologi. Meskipun pengenalan beberapa teknologi berbasis AI baru dapat dilakukan berdasarkan aturan yang ada, dalam kasus lain aturan mungkin perlu diadaptasi atau direvisi sepenuhnya. Berbagai isu etika dan moral mendasari kerangka hukum untuk teknologi berbasis AI. Etika dan moralitas yang mendasarinya dapat bervariasi di berbagai budaya dan harus dipertimbangkan sebelum aturan hukum baru dapat didefinisikan dengan tepat.

Prinsip-Prinsip Etika dan Moral yang Dipertaruhkan

Untuk dapat membahas dan memutuskan pertanyaan serta dilema etika dan moral terkait teknologi berbasis AI baru, pertama-tama diperlukan landasan bersama tentang prinsip-prinsip etika dan moral yang dipertaruhkan. Prinsip-prinsip etika dan moral ini berbeda-beda untuk setiap kasus penggunaan. Prinsip-prinsip tersebut bergantung pada sejumlah besar faktor, termasuk perbedaan budaya, tujuan dan batasan ('keadaan terkini') teknologi yang relevan, tingkat kemampuan belajar mandiri, baik itu sistem AI pendamping maupun sistem AI otonom, tingkat transparansi dan prediktabilitas algoritma yang digunakan, serta ketergantungan sosial dan lingkungan, dll.

Oleh karena itu, tidak mungkin untuk mendefinisikan seperangkat prinsip etika dan moral umum yang dapat diterapkan secara default setiap saat. Oleh karena itu, AI seringkali mencerminkan nilai-nilai penciptanya. Tanpa kehati-hatian yang tepat dalam pemrograman sistem AI, terdapat risiko bahwa pemrogram dapat (baik sengaja maupun tidak sengaja) berperan dalam menentukan hasil. Dalam hal ini, perbedaan budaya dapat memainkan peran yang signifikan. Namun demikian, karena kurangnya seperangkat prinsip etika dan moral universal, dalam setiap pengembangan dan/atau implementasi teknologi berbasis AI, pada akhirnya setiap organisasi bertanggung jawab untuk mendefinisikan seperangkat prinsip etika dan moral yang relevan, dengan demikian mempertimbangkan semua keadaan yang relevan. Titik awal ini juga harus divalidasi secara berkala untuk memastikan akurasi yang berkelanjutan.

Standar etika dan moral sangat bergantung pada area penerapan AI. Dalam bab ini, kami membatasi cakupan kami pada standar aplikasi AI sipil (misalnya, bisnis-ke-bisnis, bisnis-ke-konsumen, dan pemerintah-ke-konsumen). Kami tidak membahas aplikasi militer atau geopolitik karena keduanya memerlukan kerangka penilaian yang berbeda.

Pemahaman Umum Tentang Prinsip-Prinsip Keselamatan dan Keandalan

Prinsip-prinsip keselamatan dan keandalan untuk teknologi berbasis AI akan selalu mencerminkan serangkaian prinsip etika dan moral yang telah disebutkan sebelumnya.

Sebagaimana telah disebutkan, prinsip-prinsip ini bergantung pada pandangan dunia tertentu. Sebagai contoh, dapat dipertanyakan apakah teknologi berbasis AI harus terbukti melampaui tingkat keamanan solusi yang ada sebelum diimplementasikan, atau apakah dapat diterima jika dianggap kurang aman, asalkan memiliki banyak atau setidaknya cukup manfaat efisiensi.



Dengan tidak adanya prinsip-prinsip yang diterima secara umum, prinsip kehati-hatian "lebih baik mencegah daripada mengobati" dapat berguna untuk AI. Prinsip ini menyiratkan bahwa terdapat tanggung jawab sosial untuk melindungi masyarakat secara wajar dari paparan bahaya. Prinsip ini berkaitan dengan tindakan dalam situasi yang tidak pasti: kurangnya kepastian ilmiah yang lengkap tentang bahaya tidak dapat digunakan sebagai argumen untuk menunda tindakan mitigasi. Perlindungan hanya dapat dilonggarkan jika temuan

ilmiah menunjukkan bahwa tidak akan terjadi bahaya. Prinsip kehati-hatian khususnya telah diterapkan pada hukum kesehatan dan hukum lingkungan. Dalam kedua kasus tersebut, sistemnya rumit dan hasilnya sulit diprediksi.

Selain prinsip-prinsip yang mendasarinya, standar keselamatan dan keandalan idealnya harus ditentukan berdasarkan masukan dari semua pemangku kepentingan tidak hanya bisnis yang mungkin menerapkan solusi tersebut, tetapi juga masukan dari:

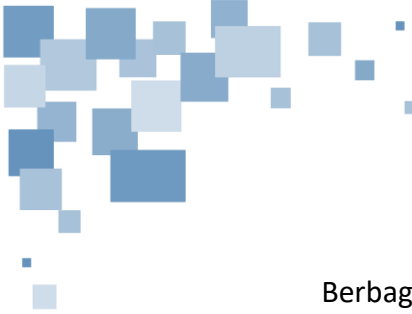
- Karyawan/pekerja;
- Konsumen/pasien;
- LSM dan pemangku kepentingan sosial lainnya;
- Akademisi dan pihak lain yang dapat memberikan saran tentang kemungkinan kemajuan setelah solusi generasi saat ini.

Sebelum kita membahas lebih detail tentang pembuatan standar, penting untuk memahami konsep Keselamatan dan Keandalan dengan baik.

1. Keselamatan

Keamanan

Kamus Oxford mendefinisikan "keselamatan" sebagai "kondisi terlindungi dari atau tidak mungkin menyebabkan bahaya, risiko, atau cedera." Konsep keamanan sangat luas. Selain keamanan fisik (dalam hal kematian atau cedera pribadi), keamanan dapat merujuk pada keamanan finansial (misalnya, saran robotika yang salah tentang investasi keuangan) dan keamanan emosional (misalnya, memanipulasi AI).



Berbagai tim pakar AI telah mendefinisikan keselamatan dalam konteks AI. Kelompok Pakar Tingkat Tinggi Uni Eropa untuk Kecerdasan Buatan, misalnya, telah mengembangkan pedoman yang menjelaskan, antara lain, konsep keselamatan. Menurut pedoman ini, keselamatan adalah "memastikan bahwa sistem benar-benar melakukan apa yang seharusnya dilakukan, tanpa merugikan pengguna (integritas fisik manusia), sumber daya, atau lingkungan.

Ini mencakup meminimalkan konsekuensi dan kesalahan yang tidak diinginkan dalam pengoperasian sistem." Kelompok Eropa untuk Etika dalam Sains dan Teknologi Baru membagi konsep keselamatan sistem 'otonom' menjadi (i) keselamatan eksternal bagi lingkungan dan penggunanya, (ii) keandalan dan ketahanan internal, dan (iii) keselamatan emosional terkait interaksi manusia-mesin. Pengembang AI harus mempertimbangkan semua dimensi keselamatan ini dan melakukan pengujian yang ketat sebelum merilis apa pun, dengan perhatian khusus kepada orang-orang yang berada dalam posisi rentan dan potensi penggunaan ganda AI.

Pendekatan Uni Eropa terhadap pasar internal didasarkan pada aturan keselamatan umum. Undang-undang Uni Eropa mendefinisikan konsep umum keselamatan, yang didasarkan pada harapan wajar dari konsumen atau masyarakat luas. Peraturan perundang-undangan produk terbatas pada apa yang disebut "persyaratan esensial", misalnya, perlindungan kesehatan dan keselamatan yang harus dipenuhi suatu barang ketika dipasarkan.

Prinsip-prinsip ini dirumuskan dalam bentuk standar teknis terharmonisasi yang disusun oleh Organisasi Standardisasi Eropa. Kepatuhan terhadap standar-standar ini akan memberikan praduga kesesuaian dengan persyaratan esensial (misalnya, penandaan CE). Penandaan CE menunjukkan bahwa suatu produk mematuhi persyaratan esensial sebagaimana tercantum dalam arahan terkait.

Harmonisasi keselamatan internasional

- ❖ Direktif Tanggung Jawab Produk,
- ❖ Petunjuk Umum Keselamatan Produk
- ❖ Direktif Mesin.
- ❖ "Deklarasi Kesesuaian Komisi Komunikasi Federal AS
- ❖ Tanda "Sertifikat Wajib Tiongkok"
- ❖ "Dewan Sukarela untuk Pengendalian Gangguan" di Jepang.

Produk dan layanan baru berbasis teknologi yang sedang berkembang seperti AI belum tentu kurang aman dibandingkan produk tradisional. Bagaimana teknologi dan alat tersebut digunakan sangat penting untuk aspek keselamatan dan tanggung jawab. Penting bagi teknologi yang dapat memengaruhi kehidupan dan kesejahteraan untuk memiliki mekanisme keselamatan yang mematuhi norma-norma yang berlaku, termasuk peraturan perundang-undangan. Saat merancang teknologi baru, penting untuk mempertimbangkan berbagai aspek kesehatan dan keselamatan, termasuk, misalnya, ergonomi dan tekanan mental.

2. Keandalan

Keandalan

Keandalan sering dijelaskan sebagai konsistensi keseluruhan suatu pengukuran, atau sebagai "probabilitas menjalankan fungsi tertentu tanpa kegagalan dan dalam parameter desain untuk jangka waktu atau jumlah siklus yang ditentukan dalam kondisi layanan aktual."

Terkait keandalan dalam konteks AI, Kelompok Pakar Tingkat Tinggi untuk Kecerdasan Buatan (HGEG) mensyaratkan agar algoritma aman, andal, dan cukup tangguh untuk menangani kesalahan selama fase desain, pengembangan, eksekusi, penerapan, dan penggunaan. Algoritma juga harus mampu menangani hasil yang salah secara memadai. Dalam komunitas riset AI, saat ini terdapat peningkatan kesadaran akan pentingnya reproduktifitas, yang esensial untuk menjamin hasil yang konsisten di berbagai situasi, kerangka kerja komputasi, dan data masukan.

3. Hubungan Antara Keamanan dan Keandalan

Pada paragraf sebelumnya, kita telah mengidentifikasi bahwa konsep keamanan dan keandalan secara inheren saling terkait. Namun, keamanan dan keandalan tidak berkorelasi sempurna satu sama lain. AI yang aman dapat memprediksi hasil yang tidak andal, dan sebaliknya, AI yang tidak aman dapat memprediksi hasil yang andal.

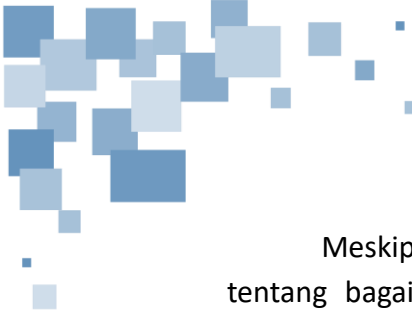
Menerapkan hukum dan peraturan keselamatan dan keandalan yang berlaku di Eropa pada AI seringkali terbukti sulit mengingat sudah lama hukum dan peraturan ini diberlakukan. Meskipun demikian, banyak undang-undang dan peraturan Eropa bersifat netral terhadap teknologi, yang berarti undang-undang dan peraturan tersebut seharusnya dapat diterapkan pada kasus penggunaan AI.

Selain itu, Organisasi Standardisasi Eropa sedang mengembangkan standar yang terharmonisasi untuk robot canggih dan sistem otonom yang didukung AI. Organisasi Internasional untuk Standardisasi baru-baru ini memulai dengan standar dasar yang mencakup konsep dan terminologi AI. Namun, standardisasi masih merupakan hal yang kompleks, karena minat yang luas terhadap AI menyatukan berbagai pemangku kepentingan.

6.1 TANTANGAN ETIKA DAN MORAL UNTUK TEKNOLOGI BERBASIS AI

"Masalah Troli"

Sebuah eksperimen pemikiran yang banyak dikutip tentang tantangan etika dan moral adalah apa yang disebut "masalah troli", yang pertama kali diperkenalkan oleh Philippa Foot pada tahun 1967. Dalam eksperimen mental ini, sebuah troli mengancam akan menabrak lima orang. Troli tersebut dapat dialihkan ke jalur lain dengan mengganti sakelar, tetapi di jalur yang lain, ada orang lain yang akan tertabrak. Pertanyaannya adalah apakah secara moral diperbolehkan untuk campur tangan dan menekan sakelar untuk mengalihkan troli ke jalur yang berbeda.



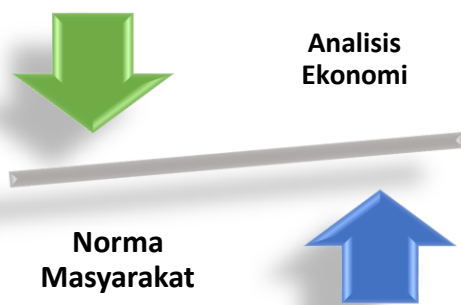
Meskipun eksperimen pemikiran seperti ini tampaknya sesuai dengan pertanyaan tentang bagaimana AI seharusnya berperilaku dalam situasi dilema seperti troli, patut dipertanyakan apakah pertanyaan hipotetis tentang situasi yang tidak familiar dengan taruhan yang relatif tinggi ini dapat memprediksi tindakan dalam kehidupan nyata. Dalam kehidupan nyata di mana manusia membuat keputusan secara real-time, jawaban atas pertanyaan-pertanyaan semacam itu lebih kompleks: hal itu dapat membuat perbedaan besar jika orang tersebut adalah seorang kenalan baik, misalnya. Perancangan sistem AI mau tidak mau akan mempertimbangkan dilema semacam itu sebelumnya, sehingga dilema tersebut harus diperhitungkan dalam pemrograman.

Pilihan Tragis

Seperti halnya semua upaya rekayasa yang signifikan, sistem AI tentu melibatkan trade-off di antara faktor-faktor penghambat. Satu set trade-off semacam itu menyeimbangkan manfaat keselamatan dan keandalan di satu sisi dengan biaya untuk mencapai manfaat tersebut. Biaya-biaya tersebut tidak terbatas pada biaya moneter saja, tetapi juga mencakup isu-isu lain seperti seberapa cepat sistem AI ditempatkan untuk penggunaan non-eksperimen ("produksi"), seberapa banyak pengujian yang diperlukan sebelum penggunaan tersebut dimulai, dan sebagainya.

Dalam hal terkait, norma-norma sosial mengenai keselamatan dan keandalan perlu dipertimbangkan. Misalnya, jika suatu sistem AI secara keseluruhan lebih aman daripada sistem yang digantikannya, tetapi tetap saja, sistem tersebut menimbulkan risiko yang lebih tinggi di satu area tertentu yang sebelumnya cukup aman, apakah kita bersedia menerima risiko sempit baru tersebut untuk mendapatkan manfaat dari keseluruhan peningkatan keselamatan? Karya penting tahun 1978, *Tragic Choices*, karya Guido Calabresi dan Philip Bobbitt (selanjutnya disebut "*Calabresi*") memberi kita beberapa mekanisme untuk mengatasi isu-isu ini.

Calabresi mengakui bahwa standar moral dan ekonomi berkontribusi pada apakah dan bagaimana masyarakat memberi insentif pada kegiatan-kegiatan tertentu. Contoh klasik dari pilihan tragis semacam itu muncul dalam setiap proyek teknik sipil besar. Keputusan untuk membangun jembatan, misalnya, seringkali dibuat bukan hanya karena manfaat ekonomi yang akan diberikan jembatan tersebut, tetapi juga manfaat keselamatan yang menyertainya (misalnya, dengan mengurangi kematian lalu lintas karena jembatan, yang telah digunakan selama lebih dari satu abad, akan mengurangi kematian lalu lintas dengan mengganti rute 100 km yang sering dilalui dengan rute 10 km). Di sisi lain, statistik dari proyek-proyek sebelumnya mungkin menunjukkan bahwa bahkan dengan tindakan pencegahan keselamatan "normal", sejumlah pekerja konstruksi akan meninggal saat membangun jembatan. Bagaimana masyarakat membuat pilihan tragis ini?



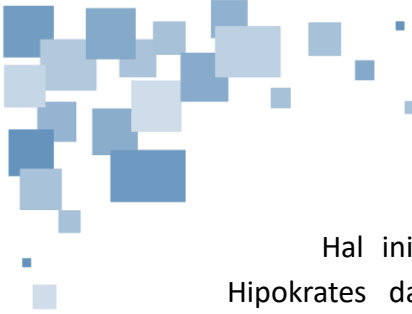
Dalam *Tragic Choices*, Calabresi memberikan pendekatan realistis untuk hal ini, dengan menyadari bahwa betapapun tidak sempurnanya norma-norma moralitas masyarakat saat ini, norma-norma itulah yang akan digunakan sebagai komponen moral dalam kalkulus. Pengamatan terkait adalah bahwa norma-norma masyarakat tersebut seringkali bertentangan dengan

analisis ekonomi standar yang mungkin digunakan untuk menganalisis situasi serupa. Misalnya, jika sumber daya medis yang langka dapat menyelamatkan Einstein atau orang yang sangat kaya, norma moral akan memilih Einstein sementara analisis pasar akan memungkinkan orang kaya tersebut untuk membayar biaya kesehatannya. Jadi, meskipun kita mungkin menyukai pendekatan yang efisien pasar secara abstrak, kita seringkali merasa enggan ketika melihatnya diterapkan dalam situasi yang secara sosial tidak nyaman. Bahkan analisis ekonomi yang memungkinkan intervensi dalam pasar bebas pun menghadapi rintangan praktis dan teoretis.

Tanpa membahas lebih jauh topik yang kompleks ini, kesimpulan pentingnya adalah bahwa analisis ekonomi murni, khususnya analisis berbasis pasar, tidak cukup untuk memberikan jawaban terbaik bagi kita atas permasalahan pilihan tragis. Pada saat yang sama, keputusan yang dibuat hanya berdasarkan naluri moral masyarakat bisa sangat tidak efisien; oleh karena itu, ada alasan untuk mengingatkan orang-orang tentang implikasi ekonomi dari pilihan mereka. Terakhir, ada "biaya untuk menghitung" hal-hal yang berharga seperti nyawa manusia. Sekadar menyatakan bahwa nyawa manusia bernilai X dolar, apa pun alasannya (baik atau buruk), membuat sulit untuk mengatakan bahwa sesuatu yang berharga itu tak ternilai harganya.

Terkadang kita dapat menghindari kalkulus yang mengubah nyawa menjadi dolar sama sekali, misalnya ketika kita memutuskan untuk membangun pagar pembatas di satu ruas jalan bebas hambatan daripada ruas lainnya karena perlindungan untuk ruas yang dipilih akan menyelamatkan sekitar 15 nyawa dibandingkan dengan ruas lainnya yang hanya 10 nyawa. Namun, dalam banyak kasus, perbandingan yang sulit ini sulit dihindari. Oleh karena itu, dalam membahas pertimbangan keselamatan dan keandalan bab ini, kita akan mempertimbangkan standar moral konvensional maupun implikasi ekonomi dari pilihan yang dapat dibuat.

Meskipun AI bukan satu-satunya bidang di mana isu ini muncul, AI memang memiliki beberapa karakteristik penting yang tidak ditemukan di banyak bidang kebijakan pilihan tragis lainnya. Misalnya, mengingat perbedaan yang sangat signifikan antara mesin dan manusia dalam tugas yang dapat mereka lakukan dengan baik, sangat mungkin solusi AI di tahun-tahun awal akan meningkatkan paparan risiko di bidang-bidang tertentu, sementara pada saat yang sama secara drastis mengurangi risiko di bidang-bidang lain.



Hal ini mempertanyakan misalnya gagasan "tidak membahayakan" dari Sumpah Hipokrates dan apakah peningkatan bahaya dalam keadaan langka untuk mencapai peningkatan keselamatan secara keseluruhan dapat diterima. Perlu dicatat bahwa pendekatan *Calabresi*, yang mempertimbangkan analisis pasar/ekonomi tetapi juga mengakui norma-norma sosial, telah tercermin langsung dalam banyak standar Etika AI yang diusulkan.

Ketakutan akan Hal yang tidak Diketahui (Standar Luring Rendah vs. Standar Daring Tinggi)

Orang-orang sering khawatir tentang teknologi yang sedang berkembang, termasuk AI. Sumber ketakutan yang penting adalah kurangnya pengetahuan. Prinsipnya di sini adalah: semakin sedikit pengetahuan Anda, semakin besar ketakutan Anda. Banyak orang, misalnya, telah menyatakan kekhawatiran tentang mesin yang memiliki kehendak bebas, meskipun tidak ada bukti ilmiah yang mendukung hal ini.

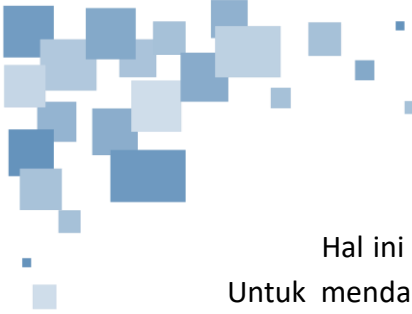
Orang-orang juga memiliki kepercayaan yang lebih rendah terhadap AI dibandingkan dengan kepercayaan terhadap manusia, sementara pada saat yang sama memiliki ekspektasi yang lebih tinggi terhadap AI daripada terhadap manusia. Mengapa kita mengharapkan aplikasi AI 100% aman dan andal, padahal kita tidak mengharapkan hal ini dari manusia? Hal ini menimbulkan dilema moral bagi penggunaan aplikasi AI. Mungkinkah menjamin aplikasi AI 100% aman dan andal? Jika tidak, haruskah ada "kerugian yang dapat diterima", yang dapat membenarkan kesalahan AI? Sesuai argumen *Calabresi* dalam *Tragic Choices*, kita harus memperhatikan dengan saksama sifat manusia bahkan ketika hal itu melibatkan standar ganda semacam itu.

Moralitas AI dalam Situasi Kritis; Alokasi Risiko dan Kerugian Sosial dalam Fase Transisi

1. Pendahuluan

Perdebatan tentang moralitas AI dapat diilustrasikan secara gamblang oleh konsep "algoritma kematian". Unit kendali kendaraan otonom dapat dirancang untuk memutuskan, dalam keadaan darurat, apakah akan menabrak sekelompok pejalan kaki, seorang ibu dan anak, atau dinding bata. Di sektor medis, dapat dibayangkan bahwa AI akan menerima dan menugaskan pasien untuk perawatan medis berdasarkan probabilitas bertahan hidup dengan mempertimbangkan atau mengabaikan kemungkinan total biaya perawatan. Kedua ilustrasi tersebut, yang akan diuraikan lebih rinci di bagian III ("Studi Kasus"), mungkin menimbulkan pertanyaan apakah AI akan membebaskan manusia dari keputusan moral. Namun, pendekatan ini tidak tepat. AI tidak mengajarkan moralitas itu sendiri. Manusia merancang parameter yang mengajarkan fondasi moral AI. Tantangannya adalah mengajarkan moralitas AI, karena orang menilai moralitas secara subjektif dan tidak dapat menyampaikan moralitas dalam metrik terukur yang memudahkan komputer untuk memprosesnya. Lebih lanjut, konsep moral berbeda di berbagai budaya.

Oleh karena itu, titik awal untuk pertanyaan moralitas bukanlah keputusan itu sendiri, melainkan penciptaan fondasi bagi keputusan-keputusan tersebut. Karena alasan ini, produsen AI tidak dapat dibebaskan dari keputusan moral yang dibuat oleh sistem. Tanggung jawab pengguna juga dapat dibayangkan jika pengguna dapat memilih di antara berbagai sikap moral sistem. Perilaku pengguna yang tidak pantas dapat menjadi alasan mengapa AI harus membuat keputusan semacam itu.



Hal ini berlaku khususnya selama fase transisi dari sistem lama ke sistem berbasis AI: Untuk mendapatkan keunggulan sebagai pelopor, pasar berkembang seringkali dibangun dengan teknologi "setengah matang" yang dapat ditingkatkan "dengan cepat". Kenyataan ini tidak banyak membantu membangun kepercayaan terhadap AI, sehingga baik produsen maupun pengguna sebaiknya memberikan perhatian khusus selama masa transisi ini.

2. Moralitas AI dalam Situasi Kritis

Baru-baru ini, Kelompok Pakar Tinggi Uni Eropa untuk AI telah mengadopsi pedoman untuk kecerdasan buatan yang tepercaya. Sebelumnya, Kementerian Federal Transportasi dan Infrastruktur Digital Jerman membentuk komisi etik untuk mengemudi otomatis. Selain itu, semakin banyak universitas yang membahas topik ini, misalnya MIT, Harvard, Oxford dan baru-baru ini TU Munich dengan dukungan dari Facebook.

Pertanyaan mendasar dari penelitian ini adalah keputusan seperti apa yang seharusnya diizinkan untuk dibuat oleh AI dan berdasarkan standar apa. Dengan menyederhanakan subjek yang sangat kompleks ini menjadi istilah-istilah sederhana, standar etika dan moral harus diterapkan dalam desain AI agar dapat digunakan sebagai dasar untuk semua keputusan turunan. Kelompok Pakar Tinggi untuk AI mengusulkan, misalnya, bahwa komitmen hak-hak fundamental dari Perjanjian Uni Eropa dan Piagam Hak-Hak Fundamental harus digunakan untuk mengidentifikasi prinsip-prinsip etika abstrak untuk AI. Oleh karena itu, situasi kritis harus diputuskan berdasarkan konsep moral manusia.

3. Bagaimana Mengajarkan Moralitas kepada AI

Namun, implementasinya sulit. Bagaimana Anda mengajarkan sistem pembelajaran mandiri tentang struktur pengambilan keputusan moral manusia?

Pada awalnya, para peneliti AI, pemerintah, dan ahli etika perlu secara eksplisit mendefinisikan nilai-nilai etika sebagai parameter terukur yang dapat digunakan untuk mengoperasikan sistem. Sebagaimana akan dibahas di bawah, hal ini sedang terjadi saat ini. Kemudian, data crowdsourced tentang konsep moral harus digunakan untuk melatih sistem AI agar dapat membuat keputusan dalam menghadapi dilema moral parameter moral harus diterjemahkan ke dalam semacam basis data yang dapat dipahami oleh mesin. Kemudian, produsen harus bertanggung jawab untuk mengimplementasikan dan memperbarui langkah ini.

4. Alokasi Risiko dan Dampak Sosial

AI dan aplikasinya masih dalam tahap eksperimental. Tahap transisi ke pasar ini secara inheren rawan kesalahan. Keputusan yang dibuat oleh AI, termasuk aspek moral, sangat penting, terutama jika dipikir-pikir kembali ternyata salah. Dalam bentuk yang lebih ringan, keputusan yang fatal tidak dibuat pada saat tragedi terjadi, melainkan pada saat perancangan dan implementasi sistem.



Sebagaimana diuraikan dalam bagian III.A di bawah ini, regulasi mengemudi tradisional bergantung, antara lain, pada manusia yang mengoperasikan kendaraan dengan cara yang konsisten dengan standar moral masyarakat tempat mereka dibesarkan. Sebagian besar standar tersebut implisit, tidak ditentukan secara

jelas, dan terkadang bahkan beragam. Standar-standar ini tidak tunduk pada pelatihan atau tes terkait mengemudi apa pun.

Misalnya, latar belakang agama atau daerah seorang pengemudi dapat memengaruhi pilihan untuk bertabrakan dengan kuda atau sapi yang tiba-tiba muncul di jalan raya. Lebih lanjut, terdapat pilihan antara "dua kejahatan" yang tidak memiliki standar moral masyarakat untuk mana yang "lebih salah". Individualitas setiap pengemudi meniadakan perilaku yang seragam dan oleh karena itu mengarah pada alokasi dampak yang berbasis acak, yang mungkin kira-kira sama.

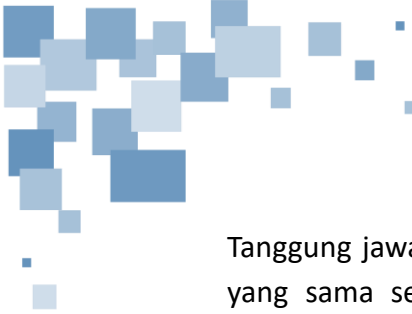
Tidak seperti manusia, unit kontrol bertenaga AI tidak memiliki usia minimum, tidak memiliki orang tua, dan tidak memiliki pendidikan untuk menyediakan seperangkat aturan etika dan moral yang implisit. Spesifikasinya harus mencakup pedoman tentang etika dan moralitas yang memungkinkan sistem untuk mengambil dan melaksanakan keputusan yang dianggap pantas secara sosial.

Oleh karena itu, pedoman tersebut sekarang harus didefinisikan secara eksplisit. Alokasi risiko yang sepihak tidaklah praktis. Kita harus memutuskan apakah ada area di mana keputusan tidak perlu dibuat dan disiapkan. Hal ini akan dibahas lebih rinci dalam kasus penggunaan kendaraan otonom dan aplikasi AI di bidang medis. Contoh-contoh ini akan menunjukkan bahwa kesalahan utama terletak pada pengguna. Hingga pengguna dapat sepenuhnya bergantung pada AI, solusi situasional harus ditemukan untuk pertanyaan alokasi risiko dan kerugian sosial.

Isu utamanya adalah apakah pengguna mengendalikan sistem yang dibantu oleh AI, atau AI yang mengendalikan sistem itu sendiri. Fase eksperimental teknologi baru dicirikan oleh fakta bahwa dampaknya tidak dapat dinilai. Keandalan sistem tidak dijamin. Jika sistem ini tetap digunakan, pengguna tunduk pada kewajiban asuransi dan pemantauan khusus. Jika pengguna melanggar kewajiban ini, ia yang harus disalahkan.

Perancang sistem AI dapat (dan akan) disalahkan atas keputusan yang fatal. Alokasi risiko dan kerugian sosial yang ketat kepada produsen dapat dipertimbangkan di sini. Bagaimanapun, perancang bertanggung jawab untuk memprogram AI. Namun, perlu diperhatikan fakta bahwa ini merupakan teknologi yang belum matang. Alokasi umum akan mengancam proses pengembangan hingga tahap produksi, karena produsen akan terbebani risiko tinggi. Hal ini dapat menyebabkan pengujian AI dihentikan, sehingga tidak mungkin mencapai kematangan pasar.

Alokasi tanggung jawab, yang di satu sisi memungkinkan fase eksperimental dan di sisi lain tidak memberikan kelonggaran bagi produsen, dapat dicapai melalui instrumen kelalaian.



Tanggung jawab atas kerusakan yang disebabkan oleh sistem dengan AI diatur oleh prinsip yang sama seperti tanggung jawab produk lainnya dari sini, dapat disimpulkan bahwa produsen berkewajiban untuk terus mengoptimalkan sistem mereka dan juga mengamati sistem yang telah mereka hasilkan serta memperbaikinya jika memungkinkan dan masuk akal secara teknologi. Kewajiban ini juga mencakup adaptasi pedoman moral yang dikeluarkan dari waktu ke waktu. Jika produsen dengan sengaja atau lalai gagal mematuhi kewajiban ini, risiko dan kerugian sosial akan dibebankan kepadanya.

Hal yang sama berlaku untuk alokasi risiko dan kerugian sosial untuk keputusan AI dalam dilema moral. Proses penciptaan fondasi moral untuk AI masih dalam tahap perkembangan. Di sini, produsen memenuhi kewajiban untuk selalu memperbarui sistemnya dengan proses tersebut. Oleh karena itu, sistem harus selalu sesuai dengan perkembangan penelitian terkini.

5. Pengamatan Akhir Terkait Moralitas dalam Situasi Kritis

Implementasi konsep moral manusia dalam AI memang menantang, tetapi dapat dilakukan. Hal ini merupakan salah satu persyaratan dasar untuk aplikasi berbasis AI yang siap dipasarkan. Dalam masa transisi, risiko dan dampak sosial harus dialokasikan sedemikian rupa sehingga tidak menghalangi inovasi dan pengembangan AI lebih lanjut.

Dampak Tindakan Pencegahan, Mitigasi, dan Penyelesaian; Dampak Desain Degradasi yang Elegan terhadap Keselamatan

Aspek penting dari keselamatan dan keandalan berkaitan dengan bagaimana suatu sistem gagal. Sebagai contoh klasik, lift dengan kabel yang putus dapat jatuh bebas ratusan meter, atau dapat jatuh kurang dari satu meter sebelum gaya sentrifugal menggerakkan flyweight untuk mengaktifkan baji untuk sistem pengereman darurat. Dengan demikian, selain meminimalkan frekuensi kegagalan sistem, meminimalkan dampak kegagalan tersebut merupakan perhatian mendasar yang berkaitan dengan keselamatan dan keandalan.

Sebentar lagi, kita akan membahas pertimbangan spesifik untuk sistem AI, tetapi akan bermanfaat untuk membangun pola pikir yang tepat terlebih dahulu dengan mempertimbangkan berbagai contoh yang mendahului sistem AI. Menyebutkannya saja sudah cukup untuk mengingatkan kita pada isu-isu penting:


Pesawat Jet

Pesawat jet memiliki beberapa mesin sehingga jika salah satu mesin rusak, pesawat akan dapat melanjutkan operasinya dalam jarak yang cukup untuk menemukan lokasi pendaratan darurat.

Trailer Kendaraan

Trailer kendaraan dilengkapi dengan detektor breakaway yang mengaktifkan sistem pengereman trailer jika trailer terlepas dari kendaraan penarik.

Pencerminan Server



"Pencerminan" server adalah metode redundansi menggunakan beberapa server komputer sehingga jika salah satu komputer rusak, komputer lain memiliki informasi yang sama dan dapat secara otomatis mengambil alih pemrosesan yang sedang dilakukan oleh mesin yang rusak.

Bagaimana perlindungan semacam itu selama berabad-abad memberi tahu kita tentang desain sistem AI? Ajaran utamanya adalah bahwa kegagalan yang elegan perlu dipertimbangkan sejak tahap awal desain. Sebagaimana seseorang tidak bisa begitu saja "memasang" mesin jet kedua setelah desain pesawat selesai, seseorang tidak boleh menganggap sistem menyediakan kegagalan yang elegan setelah pekerjaan desain fitur utama selesai. Sebaliknya, batasan desain awal sistem AI seharusnya adalah bahwa kegagalan sistem akan berdampak minimal pada kehidupan manusia, dll.

Berbicara pada tingkat yang lebih tinggi, masyarakat harus memberikan insentif bagi mereka yang merancang dan memproduksi sistem tersebut untuk mempertimbangkan kegagalan sejak dini, untuk mengatasi kegagalan tersebut secara elegan, dan untuk bersikap transparan tentang kompromi yang telah mereka buat. Ada sejumlah pengamatan tentang hal ini yang langsung terlintas dalam pikiran:

- ✚ Ada nilai pada sesuatu yang gagal dengan cara yang "diharapkan" sehingga operator manusia dan orang lain secara alami akan mengambil tindakan pencegahan yang masuk akal. Misalnya, jika sebuah fitur baru gagal, sistem akan kembali ke operasi lamanya seperti yang terjadi tanpa fitur baru alih-alih dengan cara baru. Demikian pula, jika satu jenis kegagalan mengakibatkan satu jenis operasi failover, kegagalan terkait seharusnya menghasilkan sesuatu yang serupa, alih-alih berbeda.
- ✚ Frekuensi kegagalan yang diharapkan harus diperhitungkan. Jika satu jenis kegagalan pasti akan terjadi secara teratur (bayangkan bola lampu yang padam), memperbaiki kegagalan tersebut harus mudah dan aman, dan mereka yang terdampak harus diajari tentang kegagalan dan cara memperbaikinya sebagai bagian dari pelatihan rutin. Di sisi lain, jika kegagalan jarang terjadi, dapat diperkirakan bahwa bahkan personel terlatih pun mungkin tidak tahu cara menangani kegagalan tersebut ketika terjadi, sehingga desain itu sendiri harus menghindari ketergantungan pada pelatihan semacam itu untuk memastikan keselamatan.
- ✚ Seringkali, tetapi tidak selalu, ketidakmampuan operasional lebih disukai daripada operasi yang tidak terduga setelah kegagalan terjadi. Orang-orang memahami bahwa mesin yang rusak seringkali berhenti beroperasi begitu saja.
- ✚ Kegagalan elegan juga melibatkan penyediaan beberapa jenis diagnostik kepada pengguna terkait kegagalan tersebut.
- ✚ Kegagalan elegan lebih lanjut menunjukkan bahwa jika fitur baru gagal, sistem dapat terus beroperasi seperti sebelum fitur baru tersebut ditambahkan. Hal ini memiliki manfaat yang jelas, yaitu mengembalikan pengalaman pengguna ke jenis operasi sistem yang mereka ingat dari masa lalu.
- ✚ Yang BUKAN kegagalan elegan adalah pengoperasian sistem yang benar-benar baru, dan terutama pengoperasian sistem yang tidak terduga atau gegabah. Jika lengan robot biasanya mendeteksi lokasinya dengan merasakan kedekatan dengan lokasi yang

diketahui (misalnya, dinding), maka jika sensor jarak tidak mendeteksi dinding, lengan tersebut seharusnya tidak bergerak liar untuk mencoba menemukannya lagi.

- ✚ Aspek terakhir yang kami bahas mengenai kegagalan elegan adalah apakah kegagalan tersebut jelas bagi pengguna dan, sebaliknya, konsekuensi dari kegagalan yang tidak diketahui.

Kembaran Digital

Penting untuk diingat bahwa "kegagalan" dapat berarti banyak hal, dan membedakannya mungkin sangat penting. Hal ini diilustrasikan dengan baik oleh konsep modern "kembaran digital", di mana karakteristik operasi untuk mesin aktual dimasukkan sebagai input ke versi virtual mesin tersebut sehingga apa yang kemungkinan terjadi pada mesin aktual dapat diprediksi.

Pertimbangkan kembaran digital generator turbin. Parameter seperti RPM sesaat, suhu pelumas, viskositas, dan kekeruhan dapat ditransmisikan dari mesin aktual ke kembaran digitalnya. Dengan menggunakan data historis dan pembelajaran mendalam, kembaran digital dapat melaporkan sendiri kemungkinan keausan pada komponen internal serta perkiraan waktu kegagalan aktual. Indikasi awal keausan mungkin tidak menunjukkan tindakan apa pun; indikasi sedang mungkin menunjukkan perlunya pemeliharaan dan tentu saja, indikasi parah menunjukkan bahwa kegagalan sedang terjadi dan diperlukan perbaikan. Berdasarkan lingkungan operasi, keandalan yang dibutuhkan, dan jadwal pemeliharaan, kembaran digital dapat memperingatkan kegagalan atau hanya memantau dan melaporkan sesuai kebutuhan untuk pemeliharaan terjadwal berikutnya.

Konsep-konsep yang diuraikan di atas sesuai dengan sebagian besar pernyataan etika AI yang berkembang di seluruh dunia. Misalnya, Deklarasi ICDPPC tentang Etika dan Perlindungan Data dalam Kecerdasan Buatan mencakup prinsip panduan 1(c) yang menyerukan "memastikan bahwa sistem kecerdasan buatan dikembangkan sedemikian rupa sehingga memfasilitasi perkembangan manusia dan tidak menghalangi atau membahayakannya, dengan demikian mengakui perlunya definisi dan batasan pada penggunaan tertentu."


Transparansi

Sebagaimana dijelaskan secara rinci dalam pembahasan kami tentang Transparansi dan Keterjelasan, beberapa pihak berpendapat bahwa terdapat dorongan untuk transparansi dalam hal AI, dan transparansi adalah kunci untuk membangun dan memelihara kepercayaan terhadap AI. Dari sudut pandang etika, transparansi teknologi dan model bisnis sama-sama penting.

Transparansi model teknologi

Transparansi model teknologi berarti bahwa suatu sistem harus dapat diaudit, dipahami, dan dimengerti oleh manusia di berbagai tingkat keahlian.

Namun, menjelaskan hasil dari model AI yang besar dan kompleks merupakan salah satu tantangan terbesar untuk mencapai penerimaan pengguna dan penerimaan regulasi. Transparansi model bisnis menyiratkan bahwa orang-orang diinformasikan tentang niat para pencipta sistem AI. Kejelasan, misalnya, merupakan prasyarat untuk memperoleh persetujuan



berdasarkan informasi dari individu yang berinteraksi dengan sistem AI. Lebih lanjut, kejelasan membutuhkan langkah-langkah akuntabilitas. Individu dapat meminta bukti parameter dasar dan instruksi yang diberikan sebagai masukan untuk pengambilan keputusan AI.

Namun, patut dipertanyakan apakah transparansi merupakan solusi untuk menyeimbangkan kekuatan algoritmik, ketika perusahaan atau pemerintah tidak terikat secara hukum atau didorong untuk mengungkapkan informasi. Sejauh mana transparansi harus ditempuh juga patut dipertanyakan: haruskah transparansi menyiratkan bahwa akses ke sistem harus dimungkinkan kapan saja? Selain itu, meskipun transparansi yang lebih besar dapat meningkatkan kepercayaan, transparansi juga dapat meningkatkan risiko manipulasi dan peniruan (bisa dikatakan, "dilema transparansi").

Para ahli berbeda pendapat tentang transparansi dalam sistem AI. Beberapa pihak berpendapat bahwa transparansi akan semakin penting seiring dengan meningkatnya kapabilitas otonom sistem AI. Wang, Jamieson, dan Hollands, misalnya, menemukan bahwa memberikan informasi kepada masyarakat tentang keandalan suatu sistem membantu mereka mengkalibrasi kepercayaan mereka terhadap suatu sistem, selama ketidakpastian informasi. Namun, hal ini tentu tidak selalu terjadi. Contoh yang terkenal adalah keandalan pesawat terbang. Umumnya, masyarakat percaya bahwa pesawat terbang dan personalnya memenuhi tingkat keselamatan dan keandalan tertentu. Oleh karena itu, tidak diperlukan bukti apa pun (misalnya, sertifikat perawatan pesawat terbang, kualifikasi personel). Mustahil untuk melakukannya dengan cara yang dipahami semua orang.

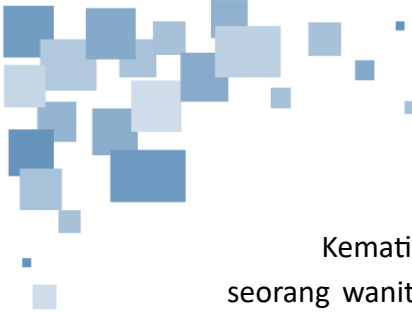
Sejauh mana transparansi telah tercapai juga dapat diperhitungkan. *Kizilcec*, misalnya, menguji tiga tingkat transparansi: rendah, sedang, dan tinggi. Transparansi sedang meningkatkan kepercayaan secara signifikan, sementara transparansi tinggi justru mengikisnya sepenuhnya. Hal ini menghasilkan tingkat kepercayaan yang sama atau bahkan lebih rendah. Hal ini menunjukkan bahwa masyarakat tidak membutuhkan (atau bahkan menginginkan) tingkat transparansi yang tinggi. Di sisi lain, masyarakat juga tidak mempercayai model kotak hitam. Oleh karena itu, wawasan dasar tentang faktor-faktor yang mendorong keputusan algoritmik perlu diberikan.

6.2 STUDI KASUS

Studi Kasus: Kendaraan Otonom

1. Pendahuluan

Ide mengemudi otonom ternyata sudah ada sejak lama. Gagasan ini berawal dari "Visi Futurama" yang dipresentasikan General Motors di Pameran Dunia New York tahun 1939, termasuk sistem jalan raya otomatis yang akan memandu mobil tanpa pengemudi. Pada tahun 1977, Laboratorium Teknik Mesin Tsukuba memperkenalkan mobil pertama yang mencapai tingkat otonomi sejati dengan memproses gambar jalan di depan yang diambil oleh kamera internal melalui teknologi komputer analog. Seiring dengan peningkatan teknologi pemrosesan dan kecepatan yang signifikan, pengembangan mengemudi otonom hampir mencapai pasar massal. Perusahaan teknologi dan produsen mobil konvensional bersaing ketat.



Kematian pejalan kaki pertama akibat kendaraan yang dikendalikan AI telah terjadi: seorang wanita tewas tertabrak setelah berjalan di depan mobil Uber yang sedang dalam mode mengemudi otomatis. Lebih lanjut, seorang pengemudi mobil Tesla yang sedang dalam mode autopilot meninggal dalam kecelakaan. Meskipun kendaraan Uber merupakan mobil otonom dengan pengemudi cadangan darurat di belakang kemudi, Tesla hanya dikendalikan oleh sistem pengemudian semi-otonom. Dalam kedua kasus tersebut, sistem cerdas tersebut gagal. Sistem tersebut tidak mengenali rintangan dan oleh karena itu sengaja memutuskan untuk tidak mengerem.

Dalam kedua kasus tersebut, pengguna manusia juga gagal. Dalam kecelakaan pertama, pengemudi "cadangan" tersebut tidak menyadari bahaya yang pasti akan ia kenali jika ia memperhatikan. Pada kecelakaan kedua, pengemudi telah menerima beberapa peringatan visual dan satu peringatan audio, dan tangan pengemudi tidak terdeteksi di kemudi selama enam detik sebelum tabrakan. Kecelakaan ini menjadi pengingat bahwa teknologi self-driving masih dalam tahap percobaan, dan pemerintah masih berusaha mencari cara untuk mengaturnya.

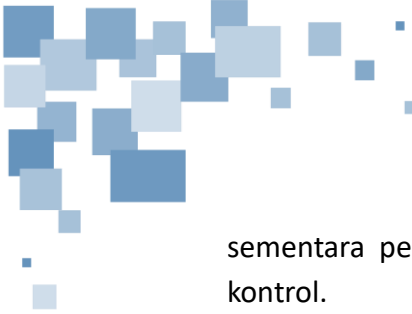
Mengemudi melibatkan keputusan untuk berakselerasi, melambat, dan berhenti, orientasi, kemudi, dan manuver untuk mencapai tujuan yang diinginkan di jalan yang diizinkan sambil menghindari rintangan apa pun. Pengemudi harus mampu mengenali situasi aktual secara akurat, meramalkan situasi di masa mendatang, menganalisis semua kemungkinan gangguan sambil memproses aliran informasi yang terus dikumpulkan dari dunia di sekitar kendaraan, mengambil keputusan untuk menyesuaikan perjalanan saat ini, dan segera menjalankannya.

Ketika mobil ditemukan, ia sama sekali tidak diatur. Butuh beberapa dekade sebelum mobil menjadi alat transportasi umum. Risiko yang melekat dalam mengemudi mendorong otoritas di seluruh dunia untuk mengembangkan aturan yang mewajibkan standar minimum teknis agar kendaraan diizinkan dan pengemudi memenuhi syarat untuk diizinkan mengemudi. Kualifikasi tersebut mencakup pelatihan teori dan praktik untuk menguasai kendaraan itu sendiri dan mengoperasikannya di lalu lintas, yang secara tidak langsung memastikan kecerdasan minimum pengemudi.

AI memainkan peran kunci dalam mengemudi otomatis dan otonom: Saat berkendara, unit pengemudi harus memproses semua informasi tentang pengoperasian kendaraan, sehingga semua tugas yang dilakukan pengemudi seperti yang baru saja dijelaskan. Pertimbangan moral di balik hal ini begitu kompleks sehingga, misalnya, telah mendorong Kementerian Federal Jerman untuk Transportasi dan Infrastruktur Digital untuk membentuk komisi etik untuk mengemudi otonom dan berjejaring. Oleh karena itu, tantangan implementasi mengemudi otonom bersifat heterogen: Aspek teknologi semata, mendefinisikan aturan hukum yang diterima secara sosial, dan pertimbangan moral.

2. Implementasi Teknis dan Tahapan Mengemudi Otonom

Mengemudi otonom diimplementasikan melalui kombinasi perangkat keras dan perangkat lunak. Prosesnya serupa dengan mengemudi manusia otak, mata, dan telinga manusia digantikan oleh sistem teknis. Persepsi terjadi oleh sistem kamera, radar, dan LIDAR,



sementara pemrosesan dikelola oleh komputer dan hasilnya ditranskripsi oleh elektronik kontrol.

Terdapat perbedaan yang signifikan antara mengemudi otomatis dan mengemudi otonom. Untuk mengilustrasikan perbedaan dasar antara kedua konsep ini, klasifikasi standar SAE J3016 oleh SAE International (*Society of Automotive Engineers*) cocok. Telah ditetapkan bahwa pendekatan teknis untuk kendaraan yang sepenuhnya otonom diklasifikasikan menurut enam tingkat yang berbeda. Klasifikasi dimulai pada tingkat 0 (tanpa otomatisasi) dan berakhir pada tingkat 5, tingkat tertinggi dari klasifikasi standar SAE J3016 yang menggambarkan kinerja penuh waktu oleh sistem pengemudian otomatis. Tingkat 5 adalah tahap di mana pengemudian otonom sejati menjadi kenyataan.

Pada tingkat ini, kendaraan tidak memerlukan pengemudi lagi. Ini berarti bahwa pengemudi kendaraan adalah sistem komputer cerdas yang membuat semua berbagai keputusan. Jelas, persyaratan keselamatan dan keandalan untuk pengemudian tingkat 5 sangat kompleks karena kendali utama pengemudi manusia yang seharusnya mengatasi situasi yang tidak diketahui juga setidaknya secara konseptual tidak ada. Metode yang paling populer untuk menjaga sistem pada jalur yang benar adalah pendekatan hibrida, yang berarti bahwa pengembang dapat mengatasi kompleksitas inheren jaringan saraf dengan memperkenalkan jaringan saraf redundansi khusus untuk proses individual yang dihubungkan oleh jaringan saraf terpusat dan melengkapinya dengan aturan jika-maka.

3. Ruang untuk Perbaikan

Kendaraan otonom mampu memecahkan beberapa masalah umat manusia modern: Jalanan seringkali penuh sesak dan terutama di wilayah metropolitan dunia, tempat parkir menjadi semakin langka. Ruang yang tersedia semakin sedikit, sehingga perlu dimanfaatkan secara efisien. Kendaraan otonom berpotensi mengurangi kebutuhan tempat parkir di Amerika Serikat lebih dari 5,7 miliar meter persegi. Perhitungan rute cerdas dan konsep berbagi mobil dapat meringankan beban jalan. Hal ini juga dapat mengurangi kebutuhan energi.


Laporan status global tentang keselamatan jalan 2018 dari Organisasi Kesehatan Dunia (WHO) menyatakan bahwa 1,35 juta orang meninggal dalam kecelakaan mobil pada tahun 2016 dan puluhan juta lainnya terluka setiap tahun. Perilaku manusia yang tidak baik merupakan penyebab paling umum dari hal ini. Kendaraan otonom dapat secara drastis mengurangi jumlah kecelakaan dengan mengganti pengemudi manusia dengan sistemnya. Hal ini merupakan tambahan dari manfaat besar yang akan diberikan otomatisasi dalam menghemat waktu bagi mereka yang biasanya sibuk mengemudikan kendaraan.

4. Dampak Moral dan Etika

Namun, ada rintangan yang perlu diatasi. Mengemudi otonom menciptakan tantangan di berbagai bidang, termasuk yang berikut:

a) Hambatan hukum beragam.

Undang-undang lalu lintas jalan perlu diadaptasi. Konvensi Wina tentang Rambu dan Sinyal Jalan telah dimodifikasi untuk mengemudi otonom: dalam keadaan tertentu,



sistem dapat mengendalikan kendaraan. Beberapa negara telah mulai mengubah peraturan lalu lintas mereka.

Mengalokasikan tanggung jawab dan liabilitas, khususnya untuk kecelakaan dan dampak tak terduga lainnya, merupakan hal yang penting. Kerangka hukum harus dikalibrasi ulang untuk mengalokasikan risiko pengenalan teknologi baru secara adil: Di satu sisi, inovasi tidak boleh dihambat atau dihambat; di sisi lain, masyarakat harus dilindungi dari dampak yang tidak dapat diterima dari teknologi yang belum matang, yang pengenalannya selalu didorong oleh kepentingan komersial para penemu dan produsen.

Otomatisasi lalu lintas juga berarti banjir data geografis, yang harus dilindungi berdasarkan undang-undang perlindungan data. Komisi pertama juga menangani kebutuhan terkait akan etika data baru.


b) Lebih lanjut, terdapat tantangan moral dan etika:

Parameter untuk AI kemungkinan harus mencakup pertimbangan fitur-fitur yang "tidak dapat dinilai": Unit pengemudian berbasis AI, misalnya, dapat mencakup parameter keputusan tabrakan yang mempertimbangkan aspek-aspek seperti jenis kelamin, usia, tinggi badan, berat badan, dll., yang mengarah pada perilaku seragam semua kendaraan otonom, sehingga memusatkan dampak pada kelompok yang dianggap "kurang terlindungi". Dilema moral bagi para pemrogram menjadi lebih berat karena mereka menyiapkan parameter keputusan tanpa tekanan waktu pada saat itu: Oleh karena itu, tidak ada tempat berlindung bagi "nasib" atau "kemalangan yang tak terelakkan".

Dilema moral di mana kendaraan otonom dapat masuk dicontohkan oleh *Massachusetts Institute of Technology* dengan "mesin moral," yang menggambarkan bahwa situasi dapat muncul dalam lalu lintas jalan di mana keputusan kualitatif dan kuantitatif harus dibuat antara kematian satu atau yang lain. Perbedaan budaya berperan di sini. Penelitian Awad et al. adalah contoh yang baik dari perbedaan budaya ini sehubungan dengan AI.

Prioritas dan preferensi bervariasi dari satu negara ke negara lain, yang memperburuk konsep untuk standardisasi. Beberapa contoh yang ditemukan dalam studi terbaru: di negara-negara dengan tingkat kesetaraan gender yang lebih tinggi, preferensi untuk menyelamatkan perempuan lebih tinggi, di negara-negara Timur ada preferensi yang lebih lemah untuk menyelamatkan yang muda daripada yang tua, dan preferensi untuk manusia daripada hewan kurang menonjol di negara-negara Selatan. Oleh karena itu, kendaraan otonom mungkin memerlukan kemampuan untuk mengunduh moralitas baru ketika melintasi batas negara.

Standar hukum berbeda-beda dan penerapan doktrin hukum, seperti kelalaian atau prediktabilitas, pada AI masih belum jelas: Apakah cukup untuk menunjukkan bahwa tindakan pencegahan yang wajar untuk mencegah kerusakan telah dilakukan dan jika demikian, apa sebenarnya yang perlu ditunjukkan?



Selain pertanyaan tentang bagaimana mengajarkan moralitas secara objektif, ada juga masalah tentang bagaimana menangani berbagai gagasan moral. Mendapatkan SIM biasanya tidak melibatkan pelatihan etika atau moral apa pun. Semua masyarakat mewajibkan usia minimum untuk memiliki SIM dengan asumsi bahwa setiap individu dewasa memiliki seperangkat aturan etika dan moral dasar yang tidak terdefinisi yang menjamin standar minimum yang dapat diterima dan mencerminkan "akal sehat" masyarakat tempat ia dibesarkan.

Bagi mereka yang gagal, terdapat sistem pertanggungjawaban yang mengalokasikan sebagian besar risiko kepada pengemudi, memperbaiki insiden yang terjadi, dan secara tidak langsung mencegah insiden serupa terulang kembali karena pengemudi lain diharapkan belajar dari perilaku yang dianggap tidak pantas dan menyesuaikan penanganannya di masa mendatang terhadap situasi yang identik atau serupa.

- c) Dampak sosial juga harus diperhitungkan, seperti yang akan kita bahas dalam pembahasan Tujuan Etis dan Manfaat Sosial.

Studi Kasus: Bedah Robotik dalam Pelayanan Kesehatan

1. Jalan Menuju "Rumah Sakit Pintar"?

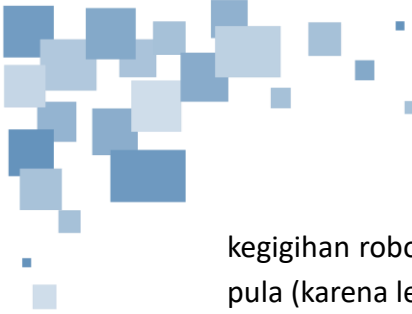
Perkembangan teknologi baru telah lama menjadi persyaratan penting bagi langkah-langkah revolusioner dalam inovasi medis. Transplantasi jantung pertama, yang dilakukan pada tahun 1967 di Cape Town oleh Christiaan Barnard, hanya dimungkinkan berkat perkembangan mesin jantung-paru. Inovasi medis juga sering kali disertai tantangan etika dan hukum. Perkembangan mesin jantung-paru memungkinkan pemeliharaan fungsi jantung, pernapasan, dan metabolisme hampir tanpa batas waktu, sehingga komite ad hoc Fakultas Kedokteran Harvard mendefinisikan "koma ireversibel" sebagai kriteria baru kematian pada tahun 1968.

Langkah pertama menuju bedah robotik dilakukan pada awal tahun 2000an dengan diciptakannya Sistem Bedah Da Vinci. Sistem Bedah Da Vinci dirancang untuk memfasilitasi pembedahan kompleks menggunakan pendekatan minimal invasif dan dikendalikan oleh ahli bedah melalui konsol. Perkembangan teknologi di masa depan terkait erat dengan implementasi AI tren ini juga mencakup sektor medis. Kini, fokus penelitian adalah bagaimana AI dapat membantu dalam diagnostik Universitas Stanford telah mendirikan *Centre for Artificial Intelligence in Medicine & Imaging*.

Juga di sektor medis, muncul pertanyaan, tantangan apa yang harus diatasi AI selama implementasinya. Lebih lanjut, kita akan bertanya apakah AI hanya mendukung manusia atau menggantikannya. Akankah robot yang dikendalikan AI merawat, mendiagnosis, dan mengoperasi pasien di masa mendatang?

2. Peran AI

Fitur khusus yang menjadikan AI sukses di sektor medis adalah kemampuannya untuk menangkap dan memproses data dalam jumlah besar secara logis. Contohnya adalah Watson milik IBM. Platform AI tersebut mampu menganalisis genom pasien manusia dengan kanker otak dan membuat rencana perawatan hanya dalam 10 menit, sementara dokter membutuhkan 160 jam untuk tugas tersebut. Performa AI, ditambah dengan presisi dan



kegigihan robot, menjadi alasan minat tersebut. Semakin baik teknologi medis, semakin baik pula (karena lebih sederhana dan kurang berbahaya) operasinya.

3. Status Perkembangan Saat Ini

Meskipun AI tertanam dalam berbagai bentuk teknologi, penggunaannya di garda terdepan praktik klinis masih terbatas. Saat ini, area aplikasi terbesar adalah diagnosis medis. Terutama di area yang melibatkan pengenalan pola, misalnya dalam radiologi dan patologi, AI dapat menunjukkan kekuatannya. Sebuah program AI membuktikan bahwa ia mampu menginterpretasikan mammogram dan menerjemahkan data pasien menjadi informasi diagnostik 30 kali lebih cepat daripada dokter manusia, dengan akurasi 99 persen.

Seperti yang telah disebutkan, perangkat bedah robotik sudah ada, tetapi manusia masih mengendalikannya. Meskipun demikian, AI telah berhasil menjalankan peran seorang ahli bedah. Pada tahun 2016, sebuah robot bedah menjahit usus halus babi sepenuhnya tanpa bantuan manusia. Pada tahun 2017, seorang dokter gigi robot Tiongkok berhasil memasang dua implan di mulut pasien tanpa keterlibatan manusia operasi implan otonom pertama yang berhasil.

4. Masa Depan AI dalam Kedokteran

Oleh karena itu, robot yang beroperasi secara otonom bukanlah fiksi ilmiah, melainkan kenyataan setidaknya dalam tahap percobaan, sistem semacam itu diuji di seluruh dunia. Sudah dapat dilihat bahwa robot-robot ini dapat mencapai hasil yang lebih baik daripada dokter manusia.

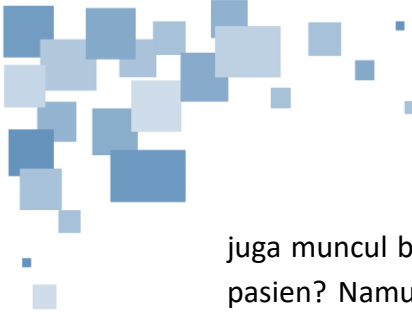
Dalam waktu dekat, sistem berbasis AI kemungkinan besar akan mendukung operator, alih-alih menggantikannya sepenuhnya. Johnson & Johnson dan Google telah bekerja sama membentuk usaha patungan bernama Verb Surgical untuk meluncurkan program robotika bedah yang akan menghubungkan ahli bedah ke platform terpadu untuk operasi, termasuk perencanaan pra-operasi, pengambilan keputusan intra-operasi, dan perawatan pasca-operasi. Para ahli percaya bahwa gelombang inovasi baru akan datang yang akan ditandai dengan konvergensi robotika bedah dengan AI dan data yang dikumpulkan dari sistem robotik.

Robot yang dapat beroperasi sendiri masih jauh: misalnya, robot tersebut jauh lebih kompleks daripada mobil tanpa pengemudi. Operasi membutuhkan lebih banyak pengetahuan dan tindakan situasional daripada pergerakan lalu lintas jalan raya, yang didasarkan pada aturan sederhana. Itulah sebabnya sebagian besar peneliti sepakat bahwa masa depan AI terletak pada peningkatan dan bantuan operasi, bukan mengambil alih tugas ahli bedah.

5. Dampak Moral dan Etika

Diskusi moral mendasar yang umumnya muncul ketika menggunakan AI dinyatakan secara sederhana: Keuntungan AI harus dipertimbangkan dibandingkan dengan kerugiannya, termasuk misalnya, masalah keselamatan atau hilangnya pekerjaan.

Data pasien sangat sensitif, oleh karena itu peran perlindungan data sangat penting, karena AI membutuhkan data dalam jumlah besar untuk tugas-tugasnya. Salah satu perhatian khusus adalah masalah tanggung jawab jika terjadi kesalahan medis, siapa yang akan bertanggung jawab? Sebagaimana halnya mengemudi otonom, pengalihan tanggung jawab yang merugikan produsen juga dapat diperdebatkan dalam hal robot operatif. Masalah trolis



juga muncul bagaimana robot seharusnya bertindak jika tidak dapat menyelamatkan semua pasien? Namun, terdapat juga masalah moral khusus. AI seorang ahli bedah robotik harus memahami prinsip-prinsip hukum dan moral dari kematian berbantuan. Terakhir, terdapat masalah khusus terkait bias inheren. Telah ditunjukkan bahwa AI dapat belajar memiliki bias rasial atau seksis, berdasarkan asosiasi yang merupakan bagian dari data yang menjadi sumber pembelajaran AI. Bias inheren semacam itu seharusnya tidak ada, terutama dalam perawatan dan pemeriksaan manusia.

6. Ringkasan

AI memiliki peluang untuk mengubah sektor medis secara berkelanjutan. Namun, rumah sakit pintar di mana manusia disembuhkan oleh robot belum dapat diharapkan dalam waktu dekat. Kemungkinan bedah robotik yang didukung oleh AI terletak pada dukungan para ahli bedah. Tantangan moral dan hukum yang ditimbulkan oleh bedah robotik sebagian besar sama dengan yang ditimbulkan oleh pengemudian otonom. Namun, terdapat juga tantangan tersendiri yang harus diatasi.

Studi Kasus: QA/QC Berbasis AI dalam Manufaktur

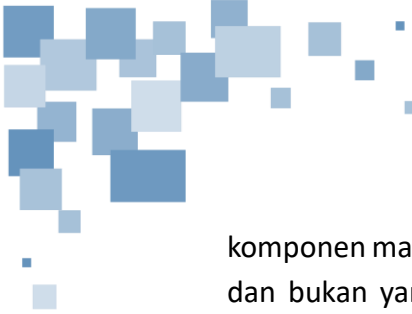
1. Pendahuluan

Sering dikatakan bahwa solusi AI seharusnya merupakan solusi yang paling tidak disukai, di mana pemahaman suatu masalah tidak dapat diterima melalui pengukuran langsung, algoritma deterministik, dan sejenisnya. Dengan kata lain, ketika semua cara lain gagal, gunakan AI karena tidak bergantung pada pemahaman apa pun tentang masalah yang mendasarinya. Kebetulan, banyak masalah kendali mutu dan jaminan mutu memang seperti itu. Perubahan kecil pada bahan baku, keausan cetakan manufaktur, atau bahkan perubahan kelembapan lingkungan manufaktur, semuanya dapat menyebabkan perubahan tak terduga pada hasil proses manufaktur.

Dalam lingkungan seperti itu, AI dapat terbukti sangat berguna sebagai pelengkap teknik jaminan mutu dan kendali mutu ("QA/QC") tradisional. Hal ini terutama berlaku ketika teknik tradisional tersebut masih sangat bergantung pada pekerja manusia (kemungkinan karena teknik otomatis tradisional terbukti tidak memadai). AI sangat efektif, seperti halnya manusia, dalam mengamati suatu komponen dan mengenali ada sesuatu yang berbeda, meskipun sulit untuk melihat kelainan apa yang ada, apalagi penyebabnya.

Pentingnya QA/QC untuk keselamatan dan keandalan produk menyebabkan penggunaan sumber daya manusia yang mahal secara ekstensif untuk hal ini, dan salah satu pertimbangan etisnya adalah bahwa pekerjaan ini sebagian besar hanya memberikan sedikit imbalan intelektual bagi sumber daya manusia tersebut. Terus-menerus mengamati aliran ribuan produk, atau bahkan mengambil sampel sebagian kecil dari jutaan barang manufaktur, bukanlah pekerjaan yang memuaskan. Namun, yang memuaskan adalah memeriksa komponen yang diidentifikasi berbeda dan mencoba mencari tahu bagaimana perbedaannya dan mengapa menjadi berbeda. Untungnya, tugas yang tidak memuaskan ini dapat dengan mudah dibantu oleh AI.

Jaringan Saraf Tiruan Konvolusional adalah solusi AI yang tepat untuk banyak masalah semacam itu. Jaringan ini mempelajari karakteristik mana yang mendefinisikan hal-hal seperti



komponen manufaktur, dan perbedaan mana dalam karakteristik tersebut yang dapat diterima dan bukan yang mengindikasikan cacat. Sebagai contoh, pertimbangkan sistem pencitraan yang menganalisis beberapa gambar dari setiap komponen manufaktur saat melaju di atas konveyor. Gambar-gambar untuk setiap unit tersebut akan memiliki banyak kesamaan (tentu saja) karena semuanya membandingkan bagian-bagian yang seharusnya identik. Namun, akan ada beberapa perbedaan dalam gambar karena, misalnya, setiap kali sebuah unit ditempatkan di konveyor, sudutnya mungkin sedikit berbeda.

Demikian pula, jika sebuah pabrik memiliki jendela atau skylight yang memungkinkan masuknya cahaya sekitar, mungkin terdapat perbedaan kecerahan tergantung pada apakah hari sedang cerah atau hujan. Melalui pembelajaran yang tepat, sistem AI dapat mempelajari bahwa perbedaan ini tidak relevan dengan kualitas produk. Namun, jika terdapat perbedaan ukuran atau tekstur yang terdeteksi oleh pencitraan, perbedaan tersebut kemungkinan besar mengindikasikan perubahan aktual pada unit manufaktur yang perlu mendapat perhatian lebih lanjut. Selain pencitraan, sensor juga dapat mendeteksi berat, densitas, konduktivitas, dan karakteristik lain dari suatu barang. Karena alasan-alasan ini, solusi jaminan/kendali mutu telah menjadi komponen utama dari *Industrial Internet of Things* (IIoT).

2. Dampak Moral dan Etika


Di beberapa industri, dampak etika dari penerapan AI dalam manufaktur cukup besar. Melihat pengguna akhir produk, efektivitas teknik QA/QC tersebut jauh melampaui apa yang dapat dilakukan manusia, terutama dengan produk manufaktur yang diproduksi dalam jumlah besar dengan karakteristik yang mungkin sulit diamati oleh manusia (misalnya, pabrik yang memproduksi bantalan bola). Seseorang mungkin berpendapat bahwa terdapat keharusan etika untuk menggunakan teknik semacam itu jika secara komersial masuk akal guna memastikan lebih sedikit produk cacat yang masuk ke dalam arus perdagangan. Jika bantalan bola digunakan dalam sistem kritis pesawat terbang, misalnya, kegagalan komponen tersebut bisa jauh lebih dari sekadar ketidaknyamanan.

Namun, ada juga komponen etika kedua. Seseorang mungkin berpendapat bahwa pekerjaan apa pun adalah pekerjaan yang baik karena manfaat ekonomi yang diberikannya kepada pekerja. Namun, ada beberapa tugas yang begitu repetitif atau tidak memuaskan sehingga dengan kata lain bukan pekerjaan yang diinginkan. Kita akan lebih baik memanfaatkan sumber daya manusia di bidang lain yang masih paling cocok untuk pekerja manusia. Dalam hal ini, terdapat keharusan etis untuk menggunakan AI jika dapat meringankan beban kerja pekerja manusia.

Studi Kasus: Pengenalan Suara dan Wajah Memisahkan Keandalan dari Keamanan

1. Pendahuluan

Dalam banyak contoh yang dibahas dalam bab ini, keandalan dan keamanan saling berkaitan erat. Jika suatu sistem tidak andal, hal itu menempatkan orang atau aset pada risiko bahaya. Namun, penting juga untuk menganggap keandalan sebagai hal yang penting, jika tidak ada masalah keamanan. Untuk mengilustrasikan hal ini, kami akan membahas beberapa aplikasi pengenalan pola yang umum. Masalah pengenalan pola yang sering dirujuk adalah pengenalan kanker berbasis AI dari modalitas pencitraan medis. Implikasi etis dari sistem



semacam itu sudah jelas: Jika kita dapat menggunakan pembelajaran mesin untuk membantu mengidentifikasi, misalnya, tumor ganas yang baru muncul pada pasien dari pemindaian CAT, kita harus melakukannya terutama jika pencitraannya sangat halus sehingga sulit dikenali oleh ahli radiologi manusia. Namun, terdapat implikasi etis juga untuk keandalan aplikasi non-kritis. Contoh yang akan kita bahas di sini adalah mengenali ucapan dan wajah manusia.

Mesin telah berkembang hingga mampu mengenali kita melalui wajah (misalnya, untuk membuka kunci ponsel) dan memahami kita saat kita berbicara (pemrosesan bahasa alami seperti yang digunakan dalam banyak situasi industri). Namun, mesin-mesin ini tidak sama mahirnya dengan semua pengguna. Karena sifat AI, mereka belajar dari set data pelatihan yang dimaksudkan untuk memaksimalkan keandalannya dalam operasi umum. Namun, apa yang dianggap "umum", justru menimbulkan beberapa potensi kekhawatiran.


2. Dampak Moral dan Etika

Misalnya, pertimbangkan pengenalan wajah. Jika sistem AI dilatih menggunakan korpus citra wajah yang semuanya merupakan pria dewasa Eropa Utara, kemungkinan besar sistem tersebut akan jauh lebih andal dalam mengenali wajah pria dewasa Eropa Utara daripada, misalnya, anak-anak, wanita, orang-orang dari Asia atau Afrika, dan sebagainya. Bahkan ketika pengenalan wajah hanya untuk kenyamanan (misalnya, untuk membuka kunci perangkat tanpa menggunakan kata sandi), keandalan diferensial tersebut masih melanggar banyak norma sosial umum karena membuat orang yang tidak terwakili dengan baik oleh set data pelatihan merasa berbeda atau dikucilkan.

Sebagaimana dilaporkan baru-baru ini di *New York Times*, implikasi sosial negatif dari perlakuan yang berbeda tersebut cukup signifikan sehingga seorang peneliti MIT, *Joy Buolamwini*, telah mengadvokasi "akuntabilitas algoritmik" untuk mengatasi ketidakadilan yang diakibatkannya. Organisasinya, *Algorithmic Justice League*, menangani isu-isu tersebut serta mendesak berbagai organisasi untuk berkomitmen hanya menggunakan pengenalan wajah yang tepat.

Demikian pula, isu serupa muncul dengan pengenalan suara dan pemrosesan bahasa alami. Sejauh mana tuturan vernakular, dialek, dan variasi lain dari apa yang mungkin dianggap sebagai bahasa Inggris "standar" tidak sepenuhnya diperhitungkan dalam set data pelatihan, sistem yang memudahkan kehidupan satu subkelompok populasi justru dapat meningkatkan frustrasi bagi yang lain. Perbedaan tingkat pengenalan antara suara pria dan wanita telah didokumentasikan selama beberapa dekade.

Demikian pula, aksen seringkali menimbulkan masalah, sebagaimana sebuah artikel di *The Economist* baru-baru ini menunjukkan dengan sedikit humor berdasarkan contoh-contoh pengenalan yang salah. Tentu saja, ada implikasi yang lebih serius juga. Meskipun studi kasus ini berfokus pada isu-isu yang tidak berkaitan dengan keselamatan, perlu disebutkan bahwa penurunan tingkat pengenalan suara (atau wajah) yang dipengaruhi rasa takut dapat berdampak pada keselamatan. Selain itu, dampak kenyamanan yang disebutkan di atas memang nyata, dan seringkali berdampak paling besar pada kelompok sosial-ekonomi yang sudah dirugikan. Untungnya, isu-isu etika semacam itu kini semakin sering dibahas dan dengan



audiens yang lebih luas, dan industri itu sendiri sedang menyelenggarakan kegiatan seperti lokakarya terkait etika dalam pemrosesan bahasa alami untuk mengatasi isu-isu ini.

6.3 TANTANGAN HUKUM UNTUK TEKNOLOGI BERBASIS AI

Tingkat Vertikal


Ada beberapa cara untuk melakukan regulasi; pemerintah dapat memainkan peran utama dalam transisi ke sistem AI, atau tidak sama sekali. Misalnya, regulasi “komando dan kendali” menyiratkan bahwa pemerintah mengancam akan menjatuhkan sanksi yang bertujuan menyelaraskan kepentingan perusahaan dengan kepentingan masyarakat. Meta-regulasi dan swa-regulasi diberlakukan oleh aktor non-pemerintah dan ditegakkan oleh entitas yang diregulasi itu sendiri. Dalam makalah ini, kami akan berfokus terutama pada legislasi.

Ada dua cara berbeda untuk memastikan keamanan dan keandalan: (1) mengatur dengan menyatakan bahwa suatu produk dilarang kecuali diizinkan, atau (2) mengatur dengan menyatakan bahwa suatu produk diizinkan kecuali dilarang. Hal ini dapat menghambat atau menunda kemajuan dan inovasi. Konsep regulasi mungkin (dan belakangan ini sering) mencakup variasi seperti izin terbatas untuk tujuan pengujian dan penyempurnaan teknologi baru dalam batasan yang ditetapkan dengan jelas.

Dalam metode kedua, ketentuan hukum cenderung lebih abstrak, menghasilkan konsep umum tentang keselamatan, bahaya, dan keandalan, yang merupakan cara regulasi yang lebih liberal, dan dapat mendukung pengembangan yang lebih cepat. Terkait metode kedua, penting untuk menilai "faktor risiko" dari setiap aplikasi vertikal/jenis dan memutuskan di mana "model penerimaan" yang lebih ketat perlu diterapkan. Ini melibatkan penilaian risiko tinggi korban jiwa atau cedera pribadi, misalnya skenario risiko seperti narkoba, penggunaan teknologi nuklir, dll. Untuk teknologi yang sedang berkembang, mustahil untuk menetapkan persyaratan material terkait keselamatan dan keandalan terlebih dahulu, sambil mencakup semua aspek.

Jika suatu hasil tidak jelas dan tidak memungkinkan untuk menentukan kewajiban spesifik yang harus dipenuhi, pengacara biasanya menyusun proses untuk menyediakan penanganan yang tepat, seperti dalam perjanjian R&D. Menetapkan persyaratan regulasi mungkin dapat mengikuti pendekatan perjanjian R&D ini: tujuan utamanya adalah mendefinisikan standar keselamatan dan keandalan, tetapi selama standar tersebut belum dapat didefinisikan, fokus harus diberikan pada prosesnya. Peraturan perundang-undangan harus ditinjau dan diperbarui secara berkala seiring dengan berkembangnya pengetahuan yang memungkinkan penyempurnaan aturan.

Regulasi algoritma yang efektif mungkin sulit karena AI tidak dapat didefinisikan dengan mudah. Masalah regulasi muncul pada tahap penelitian dan pengembangan (*ex ante*), dan ketika AI akan dirilis ke dunia (*ex post*). Pada fase *ex post*, terdapat masalah terkait dengan foreseeability (dapatkah produsen bertanggung jawab atas semua cedera ketika cedera tersebut tidak dapat diperkirakan secara wajar?) dan kontrol (apakah program AI mampu dikendalikan oleh manusia pembuatnya?). Pada fase *ex ante*, masalah terkait difusi, diskret, dan opasitas dapat muncul. Masalah difusi, misalnya, menunjukkan perlunya koordinasi global



terkait regulasi AI. Namun, Danaher mengisyaratkan bahwa dalam upaya-upaya sebelumnya, hal ini tidak menginspirasi kepercayaan (misalnya, perubahan iklim; proliferasi nuklir).

Tingkat Horizontal

Pada tingkat perdata, terdapat sistem pertanggungjawaban klasik yang mensyaratkan unsur kesalahan (subjektif) dan pertanggungjawaban mutlak. Di antara keduanya terdapat sistem berbasis kesalahan yang beroperasi berdasarkan yurisprudensi "beban pembuktian, pergeseran atau modifikasi" dan/atau "dugaan kelalaian".

Sistem-sistem di antara keduanya ini biasanya tidak memiliki latar belakang hukum, tetapi muncul dari pengadilan dan kasus-kasus yang mencoba menghadapi dunia yang terus berubah. Sistem pertanggungjawaban mutlak pada prinsipnya (meskipun tidak selalu) berbasis hukum. Sisi lain dari "tidak diperlukan kesalahan" seringkali merupakan batasan pertanggungjawaban tersebut yang memungkinkan pengembangan produk asuransi yang mencakup risiko-risiko terkait.


Membantu dan mendampingi pengembangan AI dari perspektif hukum kemungkinan besar tidak memerlukan "penemuan kembali roda" melainkan menilai aturan mana yang ada, misalnya, tentang pertanggungjawaban produk yang dapat disesuaikan untuk mengurangi risiko inheren dengan cara yang diterima secara sosial. Jika tidak, kita akan berakhir dengan permainan tawar-menawar dan tawar-menawar kekuatan, misalnya, keringanan biaya perawatan pasien yang komprehensif sebagai persyaratan bagi dokter yang menggunakan sistem AI medis. Di satu sisi, pilihan Anda mungkin tidak sepenuhnya bebas jika Anda sakit di sisi lain, dokter juga perlu beroperasi berdasarkan risiko yang wajar. Namun, meskipun semua undang-undang dapat menjadi usang, hal ini khususnya berlaku untuk undang-undang tentang teknologi yang sedang berkembang. Kemungkinan akan butuh waktu lama sebelum masyarakat terbiasa dengan rezim yang stabil untuk menangani keamanan dan keandalan sistem AI.

Prinsip

Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI dan undang-undang nasional yang mengatur penggunaannya wajib mengadopsi rezim dan standar desain yang menjamin keamanan dan keandalan sistem AI yang tinggi di satu sisi, sekaligus membatasi risiko bagi pengembang dan pengguna di sisi lain.

Mewajibkan dan/atau Mendefinisikan Prinsip-Prinsip Etika dan Moral yang Eksplisit yang Mendasari Sistem AI

- ✓ Pemerintah dan organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus mendefinisikan seperangkat prinsip etika dan moral yang relevan yang mendasari sistem AI yang akan dikembangkan, diterapkan, atau digunakan dengan mempertimbangkan semua keadaan yang relevan. Sistem yang dirancang untuk membuat keputusan secara otonom hanya akan dapat diterima jika beroperasi berdasarkan prinsip-prinsip yang didefinisikan dengan jelas dan dalam batasan yang membatasi kewenangan pengambilan keputusannya.
- ✓ Pemerintah dan organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus memvalidasi prinsip-prinsip etika dan moral yang mendasari



sebagaimana didefinisikan secara berkala untuk memastikan keakuratan yang berkelanjutan.

Standardisasi Perilaku


- ❖ Pemerintah dan organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus mengingat bahwa prinsip-prinsip etika dan moral tidak seragam secara global, tetapi dapat dipengaruhi, misalnya, oleh pertimbangan dan tradisi geografis, agama, atau sosial. Agar dapat diterima, sistem AI mungkin harus dapat disesuaikan agar memenuhi standar lokal di mana sistem tersebut akan digunakan.
- ❖ Pertimbangkan apakah semua kemungkinan kejadian harus diputuskan sebelumnya sedemikian rupa untuk memastikan perilaku sistem AI yang konsisten, dampaknya terhadap agregasi konsekuensi, dan kesesuaian moral dari "menimbang yang tidak dapat ditimbang" seperti nyawa vs. nyawa.

Memastikan Keamanan, Keandalan, dan Kepercayaan

- ✚ Pemerintah harus mewajibkan dan organisasi harus menguji sistem AI secara menyeluruh untuk memastikan bahwa sistem tersebut secara andal mematuhi, dalam pengoperasiannya, prinsip-prinsip etika dan moral yang mendasarinya dan telah dilatih dengan data yang dikurasi dan 'bebas kesalahan' semaksimal mungkin, mengingat keadaannya.
- ✚ Pemerintah didorong untuk menyesuaikan rezim regulasi dan/atau mempromosikan rezim regulasi mandiri industri untuk memungkinkan masuknya sistem AI ke pasar agar dapat mencerminkan secara wajar dampak positif yang mungkin timbul dari pengoperasian sistem AI tersebut oleh publik. Rezim khusus untuk penerimaan perantara dan penerimaan terbatas guna memungkinkan pengujian dan penyempurnaan pengoperasian sistem AI dapat membantu mempercepat penyelesaian sistem AI dan meningkatkan keamanan serta keandalannya.
- ✚ Untuk memastikan dan menjaga kepercayaan publik terhadap kendali akhir manusia, pemerintah harus mempertimbangkan penerapan aturan yang menjamin investigasi komprehensif dan transparan terhadap dampak buruk dan tak terduga dari sistem AI yang terjadi melalui penggunaannya, khususnya jika dampak tersebut memiliki konsekuensi yang mematikan atau merugikan bagi manusia yang menggunakan sistem tersebut. Investigasi tersebut harus digunakan untuk mempertimbangkan penyesuaian kerangka regulasi untuk sistem AI, khususnya, guna mengembangkan pemahaman yang lebih menyeluruh tentang bagaimana sistem tersebut harus diserahkan secara bertahap kepada operator manusianya.

Memfasilitasi Kemajuan Teknologi dengan Risiko yang Wajar

- Pemerintah didorong untuk mempertimbangkan apakah kerangka hukum yang ada seperti tanggung jawab produk memerlukan penyesuaian mengingat karakteristik unik sistem AI.
- Pemerintah harus mendukung dan berpartisipasi dalam koordinasi internasional (melalui badan-badan seperti Organisasi Internasional untuk Standardisasi (ISO) dan



Komisi Elektroteknik Internasional (IEC) untuk mengembangkan standar internasional bagi pengembangan dan penerapan sistem AI yang aman dan andal.



BAB 7

DATA TERBUKA DAN PERSAINGAN YANG ADIL

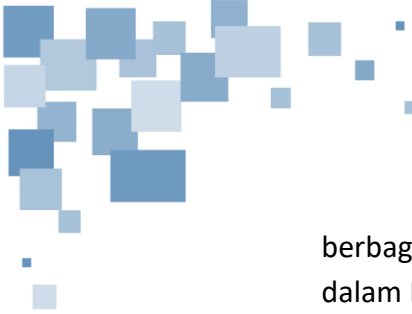
Seperti teknologi baru lainnya, pengembangan dan penerapan solusi berbasis kecerdasan buatan (AI) secara komersial berlangsung dalam kerangka hukum standar untuk kegiatan bisnis. Undang-undang persaingan usaha (antimonopoli) merupakan bagian penting dari teka-teki kepatuhan ini.

Undang-undang untuk mencegah perilaku antipersaingan usaha oleh bisnis pertama kali diperkenalkan oleh Undang-Undang Sherman AS tahun 1890. Norma-norma tersebut kini berlaku di sebagian besar negara maju dan berkembang di seluruh dunia. Tujuannya adalah untuk membatasi pelaku usaha agar tidak mencegah, menghambat, atau mendistorsi permainan bebas kekuatan persaingan di pasar. Jika suatu pasar dicirikan oleh persaingan yang efektif dan aktif antara berbagai pelaku usaha yang bersaing untuk mendapatkan pelanggan, biasanya akan ada tekanan ke bawah pada harga, tekanan ke atas pada kualitas produk dan layanan, serta tekanan untuk berinovasi dan berdiferensiasi, yang semuanya seharusnya menguntungkan pelanggan. Dalam kerangka acuan ini, hukum persaingan pada dasarnya merupakan pengecualian berbasis kepentingan publik terhadap prinsip kebebasan berkontrak dan dapat berfungsi untuk membatasi pelaku usaha agar tidak bertindak sesuka hati, guna mencegah distorsi persaingan yang dapat merugikan konsumen atau struktur pasar.

Rezim persaingan/antimonopoli di seluruh dunia secara umum konsisten karena umumnya berbasis prinsip. Prinsip-prinsip tersebut mensyaratkan bahwa (a) keunggulan kompetitif harus diperoleh secara adil (tidak dicapai melalui kesepakatan dengan pesaing) dan (b) kekuatan pasar tidak boleh dieksploitasi atau ditingkatkan secara artifisial melalui tindakan yang mengeksploitasi kekuatan tersebut, atau yang memiliki efek mengkonsolidasikan atau memperkuatnya dengan mempersulit pihak lain di pasar untuk bersaing. Pendekatan berbasis prinsip berarti bahwa hukum persaingan bersifat agnostik sektoral meskipun saat ini terdapat perdebatan global mengenai apakah ekonomi digital memerlukan pendekatan baru.

Dalam mempertimbangkan pengembangan AI, hukum persaingan relevan dalam berbagai hal:

- Pentingnya data sebagai bahan baku untuk mengembangkan sebagian besar bentuk AI pembelajaran mendalam telah dijelaskan dalam Pendahuluan. Akses terhadap data adalah kuncinya. Namun, data seringkali dianggap sebagai hak milik dan semakin dipandang sebagai aset yang dapat dimonetisasi. Pengendalian data berpotensi menghasilkan keuntungan yang mendistorsi pasar dalam pengembangan AI. Sebagaimana dicatat dalam Laporan Furman Inggris baru-baru ini, "sejauh revolusi teknologi berikutnya berpusat pada kecerdasan buatan dan pembelajaran mesin, maka perusahaan yang paling mampu memanfaatkannya mungkin adalah perusahaan-perusahaan besar yang ada karena pentingnya data untuk keberhasilan penggunaan alat-alat ini." Hukum persaingan menawarkan beberapa alat untuk mengatasi masalah ini. Banyaknya inisiatif untuk data terbuka dan inisiatif hukum untuk mendorong



berbagi data merupakan cara alternatif untuk mengatasi masalah ini. Isu-isu ini dibahas dalam Bagian II di bawah ini.

- AI sedang dikembangkan dengan berbagai cara dalam lanskap komersial, termasuk sebagai teknologi hak milik, melalui kolaborasi, atau sebagai teknologi sumber terbuka. Berbagai pilihan untuk mengembangkan sistem AI memunculkan berbagai pertimbangan hukum persaingan, yang akan dibahas pada Bagian III di bawah ini.
- Terakhir, terdapat peningkatan pemahaman tentang bagaimana implementasi sistem AI dapat menimbulkan kerugian bagi konsumen. Otoritas persaingan di seluruh dunia sedang mengkaji apakah rezim yang berlaku saat ini memiliki fleksibilitas yang cukup untuk dapat mengawasi dampak anti-persaingan dari penggunaan sistem AI.

7.1 AKSES DATA SEBAGAI INPUT UTAMA UNTUK PENGEMBANGAN AI

Sebagaimana dijelaskan dalam Pendahuluan, salah satu karakteristik sistem AI pembelajaran mendalam adalah membutuhkan data dalam jumlah besar. Data merupakan "bahan baku" utama, sebagaimana diamati oleh Departemen Bisnis, Energi & Strategi Industri (BEIS) pemerintah Inggris.

Dampak kumpulan data besar terhadap persaingan telah dipertimbangkan dalam sejumlah forum, termasuk Laporan Furman Inggris, Meja Bundar OECD tentang "Big Data: Membawa Kebijakan Persaingan ke Era Digital," laporan bersama otoritas persaingan Prancis dan Jerman tentang "Hukum Persaingan dan Data," dan laporan Biro Persaingan Kanada tentang "Big Data dan Inovasi: Tema Kunci untuk Kebijakan Persaingan di Kanada."

Pentingnya data sebagai penggerak ekonomi digital telah diakui secara luas, tetapi belum ada konsensus mengenai apakah rezim persaingan saat ini memadai untuk memastikan persaingan yang efektif di pasar digital. Laporan Big Data Kanada menyimpulkan bahwa kewenangan yang ada sudah memadai. Sebaliknya, laporan Furman Inggris mengambil pandangan sebaliknya dan menyerukan badan regulator baru dan kewajiban khusus sektor untuk memastikan keterbukaan data.

Bagian ini akan mempertimbangkan bagaimana akses ke data dapat memengaruhi kemampuan bisnis untuk bersaing dalam pengembangan sistem AI.

Poin Diskusi

Kekhawatirannya adalah persaingan yang mungkin tidak memadai atau sub-optimal dalam pengembangan alat dan aplikasi berbasis AI baru, jika akses data menjadi hambatan untuk masuk atau berkembang.

Hal ini serupa dengan industri tradisional yang bergantung pada bahan baku tertentu. Jika sumber bahan baku terbatas jumlahnya dan dikendalikan oleh sejumlah kecil produsen hilir, produsen lain mungkin akan sangat kesulitan mendapatkan bahan baku dengan harga yang wajar, atau dengan persyaratan yang tidak merugikan mereka dibandingkan pemilik bahan baku.

Kendala Eksternal terhadap Akses Data

Sebelum mempertimbangkan cara-cara data dapat dibagikan, penting untuk dicatat bahwa tidak semua data dapat diungkapkan secara bebas. Tergantung pada jenis datanya, mungkin terdapat kendala eksternal terhadap penggunaan data, dari sumber-sumber seperti hukum atau peraturan. Kendala ini mungkin tidak sepenuhnya menghalangi akses data, tetapi dapat membuatnya lebih rumit atau lebih mahal untuk mengamankan akses.

Undang-undang hak cipta, yang berlaku di sebagian besar yurisdiksi negara maju dan berkembang, dapat berarti bahwa data terlihat dan diketahui tetapi tidak dapat diakses secara sah tanpa lisensi. Mungkin perlu membayar lisensi untuk data berhak cipta dan setidaknya memerlukan pemeriksaan kepatuhan untuk memahami pendekatan pemberi lisensi terkait.

Undang-undang perlindungan data dapat mengatur cara pengumpulan, pembagian, atau pemrosesan data tertentu (biasanya data pribadi dan/atau data pribadi sensitif). Kondisi dan persyaratan hukum harus dipatuhi agar data dapat ditangani secara sah. Peraturan Perlindungan Data Umum Uni Eropa dan Undang-Undang Privasi Konsumen California tahun 2018 adalah contoh umum yang dibahas dalam pembahasan kami tentang Prinsip 7 Privasi.


Ada contoh dampak aturan ini terhadap pengembangan AI di Inggris. Royal Free Hospital di Hampstead, London memberikan data sekitar 1,6 juta pasien kepada Google DeepMind sebagai bagian dari uji coba sistem AI yang dirancang untuk mengidentifikasi tanda-tanda cedera ginjal. Alat tersebut berhasil, tetapi rumah sakit tersebut dikenai sanksi oleh otoritas perlindungan data Inggris. Pihak berwenang tidak memberi tahu pasien bahwa data mereka akan dibagikan dengan cara ini dan pihak berwenang menyimpulkan bahwa mereka tidak akan menduga catatan mereka akan digunakan seperti itu.

Penting untuk dicatat bahwa kendala-kendala ini bukannya tidak dapat diatasi. Dalam uji coba yang berbeda dengan rumah sakit di Inggris yang berbeda untuk alat diagnostik AI yang berbeda, Google DeepMind diberikan data pasien dengan cara yang tidak menimbulkan masalah privasi.

Selain itu, meskipun aturan privasi kuat di beberapa yurisdiksi, terdapat variasi yang luas di seluruh dunia mengenai apakah bisnis dapat dengan bebas mengumpulkan data tentang pelanggan dan pengguna mereka. Bahkan di negara-negara dengan kerangka privasi yang kuat, kerangka tersebut biasanya tidak berlaku untuk data apa pun.

Hukum persaingan dan hukum privasi berdiri berdampingan sebagai rezim yang saling melengkapi. Memang, penghormatan terhadap privasi pelanggan dan konsumen sering dipromosikan oleh bisnis sebagai tanda kualitas dan keunggulan kompetitif. Misalnya, Apple lebih jarang menggunakan data pelanggan yang dikumpulkan melalui perangkatnya dibandingkan banyak bisnis digital. Hal ini dianggap sebagai "serangan" terhadap para pesaing pada Januari 2019 dengan iklan besar yang ditempatkan secara mencolok bagi mereka yang menghadiri *Consumer Electronics Show* pada tahun 2019. Iklan Apple menyatakan: "Apa yang terjadi di iPhone Anda, tetap ada di iPhone Anda," mempromosikan kebijakan privasi datanya sebagai pembeda kompetitif.

Secara keseluruhan, fakta bahwa mungkin terdapat kendala eksternal atau regulasi terhadap akses data tidak dianggap sebagai "persaingan tidak sehat" atau sebagai masalah



tersendiri dalam analisis persaingan, melainkan sebagai aspek dari kondisi persaingan di pasar tertentu.

Dapatkah Data Menjadi Sumber Kekuatan Pasar, yang Menciptakan "Kepemilikan Data"?

Akses ke data juga dapat dibatasi karena kebijakan dan keputusan strategi oleh bisnis yang menyimpannya. Bisnis yang tidak ingin menawarkan datanya sebagai sumber daya terbuka dapat berupaya memonetisasinya dengan melisensikannya kepada pihak ketiga. Alternatifnya, bisnis tersebut mungkin merasa bahwa strategi yang lebih menguntungkan adalah menyimpan dataset untuk penggunaan eksklusifnya sendiri. Dalam kasus terakhir, perilaku bisnis yang bersangkutan dapat menimbulkan kekhawatiran persaingan jika hal tersebut berdampak pada pembatasan persaingan dari para pesaing. Memang, muncul pertanyaan apakah perusahaan yang menyimpan data dalam jumlah besar yang mereka kumpulkan dari pelanggan atau pengguna telah menjadi monopoli hulu, satu-satunya sumber bahan baku penting. Sebagaimana yang diamati oleh majelis tinggi Parlemen Inggris dalam tinjauannya tentang AI, "Data adalah segalanya dalam pembelajaran mesin, yang berarti siapa pun yang mendapatkan akses ke data dapat memperoleh keuntungan besar."

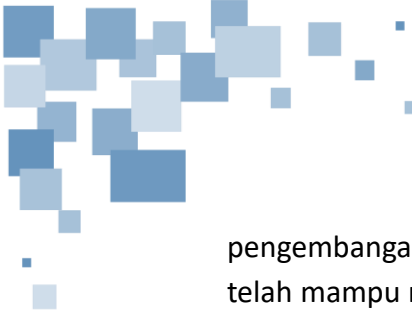
Bagaimana beberapa bisnis menghasilkan sumber daya data mereka yang sangat besar? Dalam konteks diskusi ini, jalur yang sangat signifikan adalah jalur yang diikuti oleh banyak platform digital (misalnya, mesin pencari daring, platform media sosial, pasar daring, dll.). Seringkali, produk atau layanan yang disediakan oleh platform tersebut tampak gratis; konsumen atau pengguna tidak membayar apa pun. Namun, ketiadaan pembayaran ini bisa dibidang hanya ketiadaan finansial/pembayaran. Sebagaimana telah dicatat oleh para komentator, pengguna platform seringkali "membayar" dalam bentuk data tentang diri mereka sendiri, atau data yang dihasilkan dari interaksi mereka dengan platform.

Kedua, pasar yang sama ini seringkali bersifat dua sisi, dengan pendapatan yang dihasilkan langsung dari kelompok pelanggan kedua yang terpisah, dalam bentuk pengiklan. Platform tersebut dapat menggunakan data yang bersumber dari kelompok pengguna untuk memungkinkan iklan ditargetkan pada audiens yang mereka cari. Pengiklan biasanya merupakan sumber pendapatan utama bagi sebuah platform, sehingga keinginan untuk memaksimalkan penjualan kepada kelompok pelanggan tersebut meningkatkan insentif untuk mengumpulkan data dari pengguna. Hal ini dapat menciptakan lingkaran setan yang mungkin semakin sulit untuk disaingi oleh pemain yang lebih kecil. Oleh karena itu, data dapat mendorong konsentrasi pasar sekaligus meningkatkan hambatan untuk masuk dan berekspansi di pasar tersebut.

Komisaris Uni Eropa Margrethe Vestager telah menyatakan kekhawatiran tentang pengendalian data. Mengingat pentingnya data bagi AI, pengendalian data juga menimbulkan kekhawatiran tentang kemampuan untuk mendorong inovasi ketika data dikendalikan secara privat oleh sejumlah kecil entitas. Baik Pemerintah Inggris maupun Komisaris Vestager, antara lain, telah menyatakan kekhawatiran tersebut.

Studi Kasus: Tiongkok

Tiongkok memberikan contoh bagaimana pengumpulan data dapat mendorong dan mengonsolidasikan kekuatan di beberapa platform digital besar. Keunggulannya dalam



pengembangan AI selalu terletak pada aksesnya ke data dalam jumlah besar. Raksasa teknologi telah mampu mengumpulkan data dalam jumlah besar karena peraturan privasi yang longgar dan kebijakan yang mendorong dari pemerintah Tiongkok untuk mendorong inovasi teknologi.

Hasilnya, para pemimpin teknologi domestik seperti *Baidu*, *Alibaba*, dan *Tencent* (“BAT”) telah mampu menyaingi rekan-rekan Barat mereka dalam memanfaatkan big data untuk mendorong revolusi AI di Tiongkok, sebuah negara yang sedang dalam perjalanan untuk mencapai tujuan yang ditetapkan dalam rencana strategis “Made in China 2025” dan “Rencana Pengembangan Kecerdasan Buatan Generasi Berikutnya” yang khusus untuk AI guna mendorong status pemimpin dunia dalam AI pada tahun 2030.


Dari perspektif sumber data langsung, area inti pengumpulan data untuk BAT dibagi sesuai dengan itu: Baidu, terutama sebagai penyedia mesin pencari, mengumpulkan data lalu lintas; Dominasi e-commerce Alibaba memungkinkan akses ke data transaksi dan data kredit; sementara satu miliar pengguna Tencent di seluruh dunia menyediakan aliran data konsumen yang konstan melalui aplikasi pesan sosial WeChat. Namun, ketiga perusahaan tersebut sekarang telah tumbuh begitu besar sehingga mereka mendominasi sebagian besar pasar internet Tiongkok, serta mengintegrasikan keahlian mereka ke dalam berbagai industri lain melalui investasi dan berbagi data: pada akhir Agustus 2018, dilaporkan bahwa 50,8% dari 124 perusahaan unicorn Tiongkok dikendalikan atau didukung oleh BAT. Penggunaan salah satu aplikasi perusahaan terkait BAT hampir tidak dapat dihindari dalam kehidupan sehari-hari di kota mana pun di Tiongkok, karena dilema utamanya adalah perusahaan mana yang harus diberikan datanya oleh konsumen. Misalnya, membeli barang e-commerce daring melalui salah satu dari banyak platform Alibaba (Taobao, Tmall, atau JD) atau membeli pakaian yang sebagian besar pengecer hanya menerima pembayaran melalui WeChat Pay (Tencent) atau AliPay (Alibaba).

Kekuatan Pasar tidak Dapat Diasumsikan dari Kepemilikan Kumpulan Data yang Besar

Sebagaimana telah dibahas, banyak bisnis yang berada di garis depan pengembangan sistem AI memanfaatkan kumpulan data besar yang telah dikumpulkan dari produk dan platform utama mereka. Data yang dikumpulkan untuk satu tujuan mungkin penyesuaian rekomendasi produk secara individual, atau memfasilitasi iklan bertarget, atau mengoptimalkan hasil mesin pencari sudah tersimpan di dalam bisnis dan tersedia untuk digunakan kembali guna membangun produk AI baru. Roundtable on Big Data OECD mengamati bahwa: “Kemampuan untuk menghasilkan dan memproses kumpulan data besar dapat ... dikaitkan dengan kekuatan pasar, sebagai akibat dari skala ekonomi, ruang lingkup ekonomi, dan efek jaringan, serta putaran umpan balik data waktu nyata.”

Namun, kekuatan pasar bukanlah kesimpulan yang tak terelakkan dari kepemilikan basis data yang besar. Meja Bundar OECD juga mengamati: “Kontrol atas sejumlah besar data bukanlah faktor yang cukup untuk membangun kekuatan pasar, karena saat ini beragam data dapat dikumpulkan dengan mudah dan murah oleh perusahaan kecil misalnya, melalui terminal penjualan, log web, dan sensor atau diperoleh dari industri pialang.”

Penting untuk dicatat bahwa jenis data yang mungkin diperlukan untuk melatih sistem AI akan bergantung pada tujuan desain sistem AI tersebut. Akibatnya, mungkin terdapat



banyak pasar yang berbeda untuk penyediaan data, masing-masing didefinisikan berdasarkan fungsi spesifik yang membutuhkan data pelatihan. Hal ini serupa dengan sektor farmasi, misalnya, di mana pasar ekonomi untuk obat-obatan tertentu didorong oleh penggunaan klinisnya.

Hal ini logis misalnya, tidak ada persaingan antara pengobatan penyakit hati dan pengobatan eksim, karena dokter tidak akan beralih dari meresepkan salah satunya jika pengobatan pertama menjadi terlalu mahal keduanya menjalankan fungsi yang sama sekali berbeda. Dengan cara yang sama, pasar untuk data juga dapat didefinisikan secara sempit. Misalnya, pengembang tidak akan mendapatkan kumpulan data gambar berlabel jika mereka mengembangkan sistem AI untuk menghasilkan keputusan investasi keuangan.

Selain itu, tidak dapat diasumsikan bahwa karena suatu bisnis memiliki posisi pasar yang kuat di satu area bisnisnya dan menyimpan sejumlah besar data, maka bisnis tersebut juga akan dianggap memiliki kekuatan pasar dalam penyediaan data. Penilaian terperinci diperlukan untuk memahami apakah terdapat sumber data alternatif yang dapat digunakan untuk mencapai fungsi yang diinginkan.

Selain potensi untuk membangun kumpulan data dari awal atau mendapatkannya melalui industri pialang data atau penyedia AI sebagai Layanan, banyak pemerintah memiliki kumpulan data yang sangat besar tentang berbagai aspek ekonomi dan warga negaranya. Data ini semakin tersedia untuk digunakan dalam inisiatif data terbuka dan mungkin juga menawarkan sumber pengganti untuk kumpulan data yang dimiliki secara pribadi.

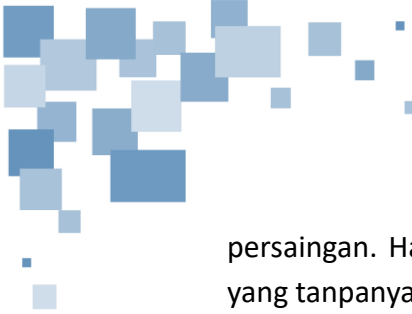
Dapatkah Hukum Persaingan Digunakan untuk Mengendalikan "Kepemilikan Data"?

Tentu saja, karena suatu bisnis memiliki keunggulan kompetitif yang signifikan, tidak serta merta muncul masalah antimonopoli. Di pasar yang memiliki persaingan efektif antara sejumlah pesaing yang berbeda, fakta bahwa seseorang memiliki sumber daya data yang sangat kuat yang memberinya keuntungan dalam pengembangan sistem AI tidak mungkin menjadi masalah sama sekali kesulitan komersial apa pun merupakan bagian dari tantangan persaingan normal. Namun, sebagian besar yurisdiksi membedakan antara pasar di mana terdapat persaingan yang efektif antar pesaing, dan pasar di mana pasar didominasi oleh satu (atau sejumlah kecil) pemain yang sangat kuat. Meskipun demikian, kekuatan pasar tidak dikenai sanksi, tetapi bisnis yang menikmati kekuatan pasar biasanya dicegah untuk memanfaatkan kekuatan tersebut sehingga merugikan konsumen.

Perlu diulangi bahwa fakta bahwa suatu bisnis memiliki data yang sangat berharga yang dapat dimanfaatkan untuk mengembangkan sistem AI bukanlah masalah persaingan itu sendiri. Di sisi lain, keputusan komersial yang diambil terkait data tersebut dapat dianggap ilegal jika berdampak, secara luas, mengeksploitasi pelanggan atau mempersulit bisnis lain untuk bersaing. Jadi, jika suatu bisnis ingin mengakses kumpulan data yang tak tertandingi yang disimpan oleh pesaing yang kuat untuk penggunaan eksklusifnya sendiri, adakah cara hukum persaingan dapat membantunya mengamankan akses tersebut?

Penolakan Akses

Mewajibkan suatu bisnis untuk memberikan akses ke aset atau fasilitas milik sendiri merupakan salah satu sanksi yang paling intrusif dalam wewenang otoritas penegakan



persaingan. Hal ini biasanya hanya terjadi jika aset tersebut merupakan "fasilitas penting", yang tanpanya pihak lain tidak dapat bersaing. Kewajiban akses tidak jarang ditemukan dalam kerangka regulasi untuk industri yang sebelumnya merupakan utilitas negara atau "monopoli alami", yang seringkali memiliki jaringan infrastruktur fisik yang luas. Persyaratan semacam itu jarang diberlakukan dalam kasus individual, meskipun sebagian besar kasus juga terkait dengan infrastruktur, seperti pelabuhan atau jaringan distribusi gas.

Kasus-kasus ini telah diperluas ke area lain, termasuk persyaratan untuk memberikan akses ke data. Dalam sebuah kasus di Uni Eropa, tiga perusahaan televisi menolak untuk memberikan lisensi jadwal siaran mereka kepada sebuah bisnis yang ingin membuat panduan program televisi mingguan terkonsolidasi. Pengadilan memutuskan bahwa lisensi wajib adalah tepat. Hal ini karena penolakan tersebut mencegah pesaing baru memasuki pasar dan menghentikan terciptanya produk baru yang diminati konsumen. Penekanan di sini adalah pada fakta bahwa penolakan tersebut akan memiliki efek eksklusif, yang memperkuat kekuatan pasar pelaku usaha lama.

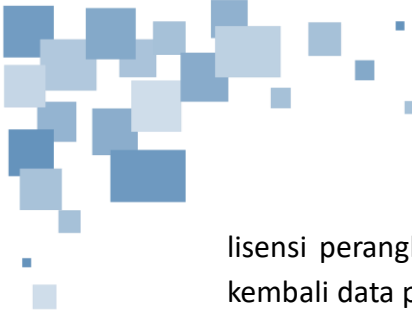
Kasus-kasus fasilitas esensial lebih lanjut telah mengklarifikasi bahwa suatu fasilitas (seperti jaringan distribusi) hanya akan menjadi esensial jika tidak layak secara ekonomi untuk mengembangkan fasilitas serupa. Otoritas penegak hukum sejak itu menyatakan bahwa mereka hanya akan mempertimbangkan untuk memberlakukan kewajiban akses jika produk atau layanan yang ditahan secara objektif diperlukan untuk dapat bersaing secara efektif di pasar hilir, dan penolakan tersebut kemungkinan akan menyebabkan hilangnya persaingan efektif di pasar tersebut dan mengakibatkan kerugian bagi konsumen.

Kembali ke hal-hal spesifik tentang mengakses data untuk memungkinkan pengembangan dan pelatihan sistem AI, tidak ada prinsip yang mengesampingkan bahwa meskipun suatu bisnis diketahui menikmati kekuatan pasar karena kendali atas data tertentu, kumpulan data yang bersangkutan harus dibagikan. Sebagaimana telah dibahas, hukum persaingan/antimonopoli memang memiliki beberapa alat potensial untuk menangani masalah kepemilikan data yang memiliki kumpulan data yang tidak dapat diakses dan tidak dapat direplikasi oleh pesaing. Namun, kasus-kasus ini kemungkinan akan sangat kontroversial dan menantang bagi pesaing yang lebih kecil untuk berjuang secara efektif, mengingat standar bukti yang tinggi dan kekayaan para calon tergugat.

Selain itu, terdapat perbedaan sikap di seluruh dunia terhadap penggunaan alat penegakan persaingan terhadap raksasa teknologi. Otoritas penegakan hukum domestik Uni Eropa dan Eropa memiliki rekam jejak yang jauh lebih kuat dalam menyelidiki dan memberikan sanksi kepada bisnis ekonomi digital dibandingkan, misalnya, otoritas AS, di mana mengambil tindakan terhadap bisnis dari sektor mana pun yang memiliki kekuatan pasar belum menjadi prioritas administratif selama beberapa tahun.

Inisiatif dan Seruan untuk Lebih Banyak Data Terbuka

Daripada menggunakan kerangka kerja persaingan untuk mewajibkan bisnis dominan memberikan akses ke data untuk pengembangan AI, pendekatan alternatif adalah mendorong pembagian data secara sukarela. Inisiatif akses terbuka memfasilitasi distribusi gratis penelitian daring sesuai dengan ketentuan lisensi akses terbuka standar yang analog dengan



lisensi perangkat lunak Sumber Terbuka. Tujuannya adalah untuk mendorong penggunaan kembali data publik, privat, ilmiah, dan akademis guna memfasilitasi inovasi dalam AI. Model ini sangat menarik bagi pendatang baru di bidang AI yang tidak memiliki akses ke data yang diperlukan untuk mewujudkan ide-ide mereka.

Beberapa inisiatif data terbuka internasional telah muncul dengan fokus pada penyediaan data untuk penelitian dan pengembangan AI. Pada tahun 2013, Negara-negara G8 menyetujui Piagam Data Terbuka G8. Piagam tersebut mengikat para anggotanya pada prinsip-prinsip data terbuka secara default, menyediakan dataset berkualitas dan berkuantitas tinggi, memastikan data dapat digunakan oleh AI, merilis data untuk tata kelola yang lebih baik, dan merilis data untuk inovasi. Anggota G8 selanjutnya berkomitmen untuk merilis data di area-area bernilai tinggi berdasarkan potensi ekonomi dan pentingnya akuntabilitas pemerintah.

Prinsip-prinsip serupa dengan Piagam Data Terbuka G8 dapat ditemukan dalam Piagam Data Terbuka Internasional. Perjanjian ini merupakan kolaborasi antara individu dan organisasi yang bertujuan untuk memajukan data terbuka secara global. Saat ini, terdapat lebih dari 60 penandatanganan pemerintah dan 40 penandatanganan organisasi untuk Piagam tersebut. Terdapat juga banyak inisiatif data terbuka di tingkat nasional, termasuk:

Amerika Serikat


Amerika Serikat telah lama memiliki tradisi kuat dalam penggunaan perangkat lunak sumber terbuka seperti Linux dan berbagi data di antara komunitas pengembang teknologi swasta. Terdapat juga berbagai program untuk memfasilitasi akses data di sektor publik. Kebijakan Data Terbuka Federal 2013 mewajibkan data pemerintah baru untuk tersedia bagi publik dalam format terbuka yang dapat dibaca mesin. Hasilnya adalah terciptanya program Data.gov, sebuah inventaris data tunggal dan daftar data publik yang dipublikasikan untuk semua tingkat pemerintahan di Amerika Serikat. Saat ini, terdapat lebih dari 300.000 set data di Data.gov yang dapat dimanfaatkan secara bebas.

Kanada

Seperti Amerika Serikat, pemerintah Kanada telah menjadikan data terbuka sebagai prioritas strategis. *Canadian Institutes of Health Research* (CIHR), Dewan Penelitian Ilmu Pengetahuan Alam dan Teknik Kanada (NSERC), dan *Natural Sciences and Engineering Research Council of Canada* (SSHRC) adalah lembaga hibah penelitian federal yang telah menerbitkan Kebijakan Akses Terbuka tentang Publikasi. Tujuan inisiatif ini adalah untuk memastikan bahwa publikasi yang dihasilkan dari pendanaan penelitian CIHR, NSERC, atau SSHRC dapat diakses secara bebas dalam waktu 12 bulan setelah publikasi. Pemerintah Kanada juga telah membuat Inventaris Data Terbuka untuk kumpulan data sektor publik yang tersedia secara bebas daring. Terdapat inisiatif data terbuka lebih lanjut yang diterapkan di tingkat pemerintah provinsi dan kota.

Uni Eropa

Komisi Eropa memandang AI sebagai topik utama strategi Pasar Tunggal Digital Komisi. Untuk memfasilitasi akses ke data terbuka di Uni Eropa, Komisi telah mensponsori inisiatif Ekonomi Data Uni Eropa. Tujuan inisiatif Ekonomi Data adalah untuk mendorong penggunaan dan penggunaan kembali informasi sektor publik serta memberikan panduan yang lebih baik



tentang akses ke informasi ilmiah dan sektor swasta. Landasan inisiatif Ekonomi Data adalah peninjauan Arahan *Public Sector Information* (PSI) yang mengatur penggunaan data terbuka pemerintah di Uni Eropa. Peninjauan ini bertujuan untuk mengurangi hambatan terhadap arus informasi bebas seperti persyaratan lokalisasi data. Komisi Eropa telah memusatkan kumpulan data dari Uni Eropa di Portal Data Terbuka Uni Eropa.

Britania Raya

Dalam Tinjauan AI Britania Raya, “Menumbuhkan Industri AI di Inggris Raya,” Profesor Dame Wendy Hall dan Jérôme Pesenti mencatat bahwa lebih banyak data terbuka di lebih banyak sektor sangat bermanfaat bagi pengembangan aplikasi AI. Tinjauan AI merekomendasikan pengurangan biaya transaksi untuk mengakses data, terutama untuk perusahaan rintisan inovatif, melalui penerapan kepercayaan data, mewajibkan pendanaan riset publik untuk memastikan publikasi data yang mendasarinya, memfasilitasi penggalan teks dan data, serta mendorong pengembangan keterampilan terkait AI. Rekomendasi ini dibahas dalam Kesepakatan Sektor AI Pemerintah Britania Raya yang diumumkan pada bulan April 2018. Kesepakatan Sektor AI merupakan inisiatif yang berkaitan erat dengan Inisiatif Data Terbuka Britania Raya. Inisiatif Data Terbuka memajukan prinsip-prinsip Piagam Data Terbuka G8 dengan memberikan komitmen terhadap data terbuka sebagai standar, kualitas dan kegunaan kumpulan data yang sesuai.

Inggris juga memiliki konsentrasi inisiatif swasta yang signifikan yang mendorong pengembangan data terbuka. Misalnya, Open Data Institute adalah organisasi independen, nirlaba, dan non-partisan yang didirikan oleh penemu world wide web, Sir Tim Berners-Lee, dan pakar AI, Sir Nigel Shadbolt, untuk mendukung penggunaan data terbuka. Open Data Institute bekerja sama dengan entitas sektor publik dan swasta untuk mengoordinasikan pembagian data dan mengadvokasi penggunaan data yang etis.


Laporan Furman baru-baru ini merekomendasikan agar Inggris memperkenalkan undang-undang yang mewajibkan pembagian data yang lebih besar oleh bisnis digital. Laporan tersebut mencatat: "Mungkin ada situasi di mana membuka sebagian data yang dimiliki oleh bisnis digital dan menyediakan akses dengan persyaratan yang wajar merupakan langkah penting dan dibenarkan yang diperlukan untuk membuka persaingan."

Australia

Pemerintah Australia merilis Pernyataan Kebijakan Data Publik pada bulan Desember 2015. Pernyataan Kebijakan tersebut berkomitmen kepada pemerintah untuk mengoptimalkan penggunaan data publik dan mengadopsi prinsip-prinsip Piagam Data Terbuka G8. Entitas Pemerintah Australia berkomitmen untuk mempublikasikan data pemerintah melalui data.gov.au dalam format yang dapat dibaca mesin dengan akses API. Yang krusial, data ini diatur oleh lisensi Creative Commons secara default, kecuali jika Kabinet memutuskan sebaliknya.

Tiongkok

Sebaliknya, Tiongkok memiliki ekosistem data tertutup dalam hal pengumpulan data masuk dan berbagi data keluar, untuk mendorong pengembangan AI domestik. Meskipun data terbuka semakin dipromosikan di Tiongkok, sumber daya tersebut umumnya tidak tersedia



untuk bisnis non-Tiongkok. Oleh karena itu, pengembang AI Tiongkok biasanya tidak dapat mengakses data dari yurisdiksi lain. Tiongkok, tentu saja, adalah negara yang sangat besar dengan banyak sumber data potensial. Jika data tersebut dapat diakses, ekosistem tertutup menyediakan sumber daya berharga ini untuk sektor AI domestik. Namun, pembatasan ini juga berpotensi merugikan pengembangan AI Tiongkok: ekosistem tertutup berarti bahwa bisnis Tiongkok tidak dapat secara bebas mendapatkan manfaat dari data terbuka global.

Namun, pemerintah dan pemangku kepentingan industri menyadari bahwa diperlukan lebih banyak kemajuan terkait data terbuka di Tiongkok. Mereka telah menyadari kendala yang timbul dari kurangnya ketersediaan sumber data terbuka internasional dan keterbatasan dalam mengandalkan sumber data terbuka domestik semata. Terdapat banyak kebijakan pengembangan AI yang akan diuntungkan dari berbagi data sebagai alat kolaborasi utama untuk meningkatkan kapabilitas pengembangan perusahaan teknologi domestik. Akibatnya, para raksasa teknologi dan individu-individu penting mulai bergabung dengan gerakan data terbuka, dan "Akses terbuka ke big data" menjadi topik utama diskusi selama acara tahunan Two Sessions tahun 2018 di Tiongkok, yang dihadiri oleh badan legislatif dan penasihat terkemuka.


Mekanisme Hukum untuk Mendorong Kerja Sama dan Berbagi Data

Kekhawatiran seputar risiko dari kepemilikan data dan konsentrasi kepemilikan data telah mendorong seruan untuk kerangka hukum di luar penegakan hukum persaingan untuk mengatasi masalah ini. Hal ini memerlukan peninjauan opsi-opsi seperti perwalian data, perjanjian berbagi data, dan program sektor publik untuk memfasilitasi akses ke data.

Tantangan Praktis Berbagi Data

Ada beberapa tantangan praktis berbagi data yang perlu diperhatikan:

- Data mungkin tidak dalam bentuk yang mudah dibagikan. Kumpulan data historis khususnya mungkin memerlukan investasi yang signifikan dalam pekerjaan persiapan sebelum dapat dibagikan. Kualitas data suatu organisasi dapat menjadi penentu utama daya tariknya sebagai mitra berbagi data.
- Tantangan lebih lanjut terhadap berbagi data adalah potensi ketidakselarasan antara sektor swasta nirlaba dan sektor publik nirlaba. Organisasi sektor swasta mungkin tidak menganggap berbagi data sebagai kepentingan terbaik mereka karena potensi pembatasan komersialisasi yang dikenakan pada universitas dan organisasi penelitian publik lainnya. Dalam situasi ini, konsistensi pendekatan sangat penting. Lingkungan regulasi yang harmonis, yang mendorong inovasi komersial dari badan publik dan swasta, merupakan kepentingan terbaik semua pihak.
- Berbagi data lintas batas juga merupakan tantangan yang signifikan. Misalnya, karena ekosistem data Tiongkok yang tertutup, pengumpulan data lintas batas mungkin sulit. Sebagaimana disebutkan di atas, kontrol ketat data untuk pengembangan AI merupakan pedang bermata dua. Suatu negara dapat menetapkan kendali eksklusif atas sumber daya ini untuk perusahaan dan lembaga penelitian dalam negerinya. Di sisi lain, jika data dibagikan lintas platform dan negara, bisnis atau badan lain dapat



memperoleh manfaat dari berbagi data global, sementara negara-negara proteksionis tetap tertutup.

Perjanjian Berbagi Data

Mungkin metode yang paling jelas untuk memfasilitasi akses ke data adalah dengan membuat organisasi sepakat secara formal untuk berbagi data satu sama lain. Akan ada sejumlah isu yang perlu dipertimbangkan, termasuk kepemilikan atau kendali data, kekayaan intelektual dan pengaturan lisensi apa pun, kepatuhan terhadap persyaratan privasi data, kepraktisan penerimaan, penyimpanan, dan penghapusan data, serta kewajiban keamanan siber.

Masalah seperti ini tidaklah sederhana dan membutuhkan waktu serta energi untuk menegosiasikan perjanjian pembagian data agar dapat mengakses kumpulan data yang sangat besar. Ketika perjanjian dibuat dengan banyak pemangku kepentingan, muncul kekhawatiran lebih lanjut tentang apakah suatu entitas dapat menggabungkan data yang diatur oleh berbagai perjanjian pembagian data.

Terakhir, preseden komersial yang dikembangkan mengenai cara menentukan harga set data masih sangat terbatas. Hal ini diperumit oleh karakteristik "kotak hitam" dari sistem AI pembelajaran mendalam. Mungkin bukan hal yang mustahil untuk mengetahui seberapa besar AI kotak hitam yang dikembangkan bergantung pada data dari sumber tertentu. Demikian pula, mungkin mustahil untuk mengetahui, sejak awal, seberapa sukses dan berharganya keluaran sistem AI nantinya. Akibatnya, menjadi sangat sulit untuk menilai data pelatihan secara komersial.

Trust Data

Trust data direkomendasikan sebagai sarana untuk memfasilitasi akses ke data untuk pengembangan atau penerapan AI dalam tinjauan independen Pesenti/Hall di Inggris tentang industri AI. Pada dasarnya, trust data beroperasi serupa dengan kumpulan paten. Para pihak dalam trust menyumbangkan data ke repositori pusat yang diatur oleh kerangka kerja yang menguraikan tata kelola kolaborasi secara keseluruhan serta hak dan kewajiban para pihak. Kerangka kerja tersebut dapat menentukan data apa yang akan dibagikan, penggunaan yang diizinkan, transfer data, penyimpanan, dan pemusnahan, serta ketentuan alokasi keuntungan finansial.

Kepercayaan data menawarkan beberapa kemungkinan manfaat. Salah satu tantangan utama dalam pengembangan AI adalah kurangnya data yang dimiliki masing-masing pelaku untuk melatih sistem AI dengan baik. Dengan menggabungkan data, peserta kepercayaan dapat memperoleh akses ke data yang tidak akan mereka miliki sendiri.

Manfaat global dari kepercayaan data



Terstandarisasi

Persyaratan umum perjanjian kepercayaan data

Skala

Memungkinkan pengumpulan data

Multilateral

Mengurangi beban negosiasi banyak perjanjian bilateral

Kepercayaan data juga mengatasi salah satu tantangan perjanjian berbagi data yaitu rumitnya negosiasi. Menandatangani kerangka kerja umum dalam kepercayaan data menghindari kebutuhan untuk menegosiasikan serangkaian perjanjian berbagi data bilateral. Hal ini secara signifikan mengurangi biaya transaksi untuk pengadaan data bagi pengembangan AI.

Model kepercayaan data bekerja paling baik ketika terdapat beragam organisasi dengan data yang bermanfaat di sektor tertentu. Misalnya, sektor transportasi berisi data bermanfaat dari pemerintah (misalnya, informasi lampu merah), akademisi (misalnya, riset keselamatan), produsen peralatan asli (misalnya, analitik dari kendaraan), penyedia perangkat lunak (misalnya, kondisi cuaca dan jalan), dan perusahaan asuransi (misalnya, telematika). Inisiatif penggabungan semacam ini menawarkan harapan yang signifikan untuk masa depan dalam meningkatkan berbagi data secara privat.

Batasan Hukum Persaingan Usaha dalam Berbagi Informasi

"Data" dapat memiliki banyak bentuk dan mencakup banyak topik. Dalam setiap diskusi tentang berbagi data, perlu diingat bahwa para pesaing perlu berhati-hati saat bertukar informasi. Di Eropa, prinsip umumnya adalah bahwa berbagi informasi komersial dapat dianggap ilegal, yang akan mengurangi ketidakpastian strategis antar pesaing, sehingga membatasi insentif untuk bersaing. Mayoritas data yang digunakan untuk pelatihan AI tidak termasuk dalam kategori ini karena kemungkinan besar bukan informasi strategis yang sensitif. Namun, risiko ini penting untuk diingat karena perbedaan antara berbagi informasi yang sah dan tidak dapat disangkal dengan perilaku kuasi-kartel dapat bersifat abu-abu, dan sanksi atas pelanggaran aturan persaingan biasanya signifikan.

Ada potensi kekhawatiran lebih lanjut bahwa jika data (meskipun tidak sensitif) dibagikan oleh pesaing dan digunakan dalam sistem AI untuk tujuan yang sama, keputusan komersial strategis mereka dapat menyatu dan menjadi sangat mirip. Penetapan harga, misalnya, dapat menjadi harmonis karena semua pemain di pasar menggunakan data input yang sama dalam sistem AI yang serupa untuk menghasilkan keputusan penetapan harga otomatis. Penetapan harga paralel tidak ilegal dengan sendirinya tanpa adanya niat bersama di antara para pesaing, tetapi dapat menyebabkan persaingan yang kurang efektif. Salah satu solusinya adalah memastikan bahwa tidak ada data yang dikumpulkan oleh para pesaing, sehingga terdapat variasi antara data "bahan baku" masing-masing.

7.2 MODEL PENGEMBANGAN AI DAN PERSAINGAN INOVASI MAKRO


Sebagian besar ekonomi digital didukung oleh teknologi AI (termasuk mesin pencari, rekomendasi personal untuk e-commerce, properti media sosial seperti pengenalan wajah dan saran kontak baru). Namun, sistem AI semakin banyak digunakan di seluruh lanskap komersial dalam berbagai cara, mulai dari teknologi mobil tanpa pengemudi di sektor transportasi, bot obrolan untuk layanan pelanggan di sektor ritel, sistem investasi kuantitatif di sektor keuangan, sistem penyeimbangan jaringan di sektor energi, hingga teknologi kembaran digital dan sistem desain generatif di sektor manufaktur.

Agar inovasi semacam itu berkembang pesat, akses ke sistem AI sangatlah penting bagi bisnis. Bagian ini akan membahas cara-cara untuk mendapatkan sumber AI. Diskusi ini pertama-tama akan mempertimbangkan AI sebagai teknologi proprietary yang dapat ditahan untuk penggunaan pengembang sendiri, atau dimonetisasi melalui lisensi. Meskipun hukum persaingan biasanya tidak akan mengganggu hak prerogatif ini, terdapat seruan untuk meningkatkan pengawasan terhadap merger yang dapat mengkonsolidasikan kekuatan pasar perusahaan teknologi besar atau menyingkirkan calon pesaing di masa mendatang. Strategi kolaboratif seperti usaha patungan untuk pengembangan AI kemudian akan dipertimbangkan, diikuti oleh inisiatif sumber terbuka. Terakhir, bagian ini akan membahas inisiatif sektor publik untuk mempromosikan dan merangsang pengembangan AI.

AI sebagai Teknologi Proprietary

Mengembangkan AI sebagai teknologi proprietary berarti teknologi AI tersebut dikendalikan oleh satu perusahaan. Perusahaan tersebut dapat memilih untuk menjual teknologi tersebut ke perusahaan lain atau melisensikannya kepada satu atau lebih perusahaan. Perusahaan yang bukan pemegang lisensi tidak memiliki akses ke teknologi tersebut, begitu pula publik. Bahkan, banyak pengembang AI besar juga menjual (sebagian) teknologi mereka kepada pihak ketiga. Pertumbuhan AI sebagai Layanan dari bisnis seperti Google, Amazon, Microsoft, berarti semakin memungkinkan untuk menerapkan sistem AI dan/atau menyewa kumpulan data terkait tanpa perlu mengembangkan (atau mengkurasi) mereka dari awal. Hal ini penting terutama karena kelangkaan talenta dalam pengembangan AI. Namun, ada juga kemungkinan bahwa suatu bisnis memutuskan untuk menggunakan sistem AI-nya secara eksklusif untuk dirinya sendiri. Sebagaimana dibahas di atas terkait akses ke data, hukum persaingan tidak bertujuan untuk mencegah bisnis mengembangkan kekuatan kompetitif melalui inovasi dan memenangkan pelanggan. Pertumbuhan organik dan pilihan untuk mempertahankan penggunaan eksklusif teknologi proprietary merupakan imbalan atas kesuksesan kompetitif.

Pengecualian untuk pendekatan umum ini adalah ketika suatu bisnis telah berkembang hingga mencapai kekuatan pasar sehingga menjadi sulit bagi perusahaan lain di pasar untuk merebut pangsa pasar. Sebagaimana dijelaskan di atas terkait akses data, bisnis dominan mungkin tidak bebas bertindak sesuka hati jika hal itu akan merugikan konsumen. Terdapat keadaan (terbatas) di mana otoritas penegak hukum dapat mewajibkan suatu bisnis untuk membuka akses ke teknologi proprietary. Namun, kasus-kasus seperti itu kontroversial, sulit diajukan, dan seringkali bukan prioritas administratif. Menggunakan sistem algoritmik untuk



memperkuat posisi dominan dengan merugikan pesaing telah dinyatakan melanggar hukum Google didenda Rp 24,2 triliun oleh Komisi Uni Eropa karena menyusun sistem hasil pencariannya untuk mengutamakan layanan perbandingan belanja miliknya sendiri daripada milik pesaing.

Meskipun mendapatkan kekuatan pasar melalui pertumbuhan organik tidak dilarang, pengawasan pengendalian merger digunakan untuk mencegah suatu bisnis mendapatkan atau lebih lanjut mengkonsolidasikan kekuatan pasar melalui akuisisi. Untuk menghindari beban birokrasi yang berlebihan pada bisnis, merger yang lebih kecil biasanya berada di bawah ambang batas yurisdiksi. Kontrol merger penting dalam kaitannya dengan persaingan yang didorong oleh AI karena sektor ini pada saat yang sama dicirikan oleh komunitas startup yang dinamis dengan rekam jejak disrupsi, tetapi juga oleh platform besar yang bersaing di banyak bidang ekonomi digital.

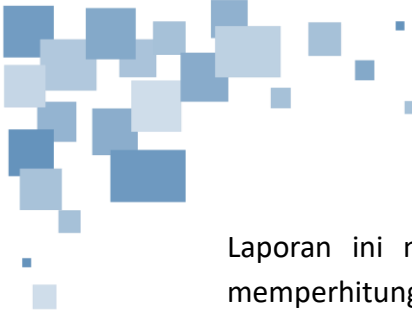
Merupakan hal yang umum bagi bisnis yang lebih besar untuk mengakuisisi bisnis startup yang inovatif guna meningkatkan pekerjaan pengembangan mereka sendiri. Merger ini seringkali dan secara sah lolos dari radar regulator karena skala dan pendapatan target terlalu rendah untuk memicu tinjauan kontrol merger. Meskipun demikian, meskipun akuisisi sektor teknologi ditinjau oleh otoritas, transaksi tersebut jarang sekali tidak disetujui. Merupakan statistik penting bahwa dari 400 akuisisi yang dilakukan secara global oleh lima bisnis digital terbesar dalam 10 tahun terakhir, tidak ada yang diblokir dan sangat sedikit yang memiliki persyaratan yang menyertai persetujuan.

Ada kekhawatiran dari beberapa pihak bahwa pandangan jangka pendek diambil dalam meninjau merger ini. Apakah pesaing inovatif kecil secara efektif ditelan oleh para petahana sebelum mereka memiliki kesempatan untuk tumbuh menjadi sumber persaingan yang signifikan? Laporan Furman mengamati:

Dalam pasar digital yang dinamis, dampak jangka panjang menjadi kunci penentu apakah merger akan merugikan persaingan dan konsumen. Mungkinkah perusahaan yang diakuisisi tumbuh menjadi pesaing platform? Apakah sumber nilainya merupakan inovasi yang, dengan kepemilikan alternatif, dapat mengurangi konsentrasi pasar? Apakah perusahaan diakuisisi untuk akses ke data konsumen yang akan membuat platform lebih sulit ditantang?

Di Jerman, ambang batas pengendalian merger telah diubah dengan memasukkan kriteria berdasarkan nilai kesepakatan, untuk menangkap akuisisi dengan harga tinggi tetapi bisnis target memiliki pendapatan rendah. Perubahan serupa telah dipertimbangkan di tingkat Uni Eropa (awalnya ditolak, sekarang sedang dipertimbangkan kembali). Laporan Furman Inggris justru menyarankan bahwa bisnis digital yang (diusulkannya) ditetapkan memiliki "status pasar strategis" diwajibkan untuk melaporkan akuisisi yang direncanakan kepada otoritas persaingan (bukan hanya yang memenuhi ambang batas yurisdiksi pengendalian merger).

Laporan tersebut juga mengusulkan perubahan uji substantif dengan mengacu pada merger mana yang harus dinilai. Saat ini, otoritas Inggris akan mempertimbangkan apakah merger kemungkinan besar akan menyebabkan penurunan persaingan yang substansial (uji coba di AS, Australia, dan Uni Eropa serupa, begitu pula di sebagian besar negara Eropa).



Laporan ini mengusulkan untuk memperluas uji coba sehingga memungkinkan untuk memperhitungkan skala kerugian. Dengan demikian, tindakan terkait merger dimungkinkan jika kerugian tidak lebih besar kemungkinannya, tetapi akan sangat signifikan jika terjadi.

AI sebagai Teknologi Kolaboratif

Banyak bisnis tidak memiliki sumber daya untuk mengembangkan sistem AI secara unilateral dan internal. Mereka mungkin ingin melakukan penelitian dan pengembangan bersama, atau membeli keahlian melalui konsultan atau spesialis lain yang dapat bekerja sama dengan mereka untuk mengembangkan sistem AI yang diinginkan. Jika AI telah dikembangkan di dunia akademis, institusi tersebut mungkin tidak berada dalam posisi yang tepat untuk menempatkan inovasi mereka ke dalam pengembangan komersial.

Jika AI dikembangkan atau dimasukkan ke dalam produksi komersial melalui kolaborasi, hukum persaingan harus dipatuhi. Hubungan antara pihak-pihak dalam usaha patungan harus disusun dengan cermat terkait hak-hak yang timbul dari inovasi dan terkait eksploitasi komersialnya di masa mendatang. Koordinasi antar entitas yang bukan pesaing nyata maupun potensial memiliki risiko persaingan yang jauh lebih rendah dibandingkan koordinasi antar pesaing.

Tidak ada yang baru dalam hal ini terkait AI:

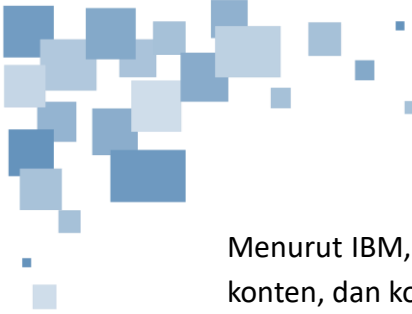
- Jika kolaborasi disusun sebagai pengaturan komersial, hal tersebut harus dapat diterima mengingat larangan persaingan yang lazim berlaku pada pengaturan anti-persaingan. Hal ini penting karena risiko kolaborasi yang sah dapat meluas menjadi perilaku seperti kartel yang melanggar hukum, yang merugikan konsumen.
- Jika kolaborasi mengambil bentuk struktural, misalnya dengan kendaraan korporat untuk usaha patungan, hal tersebut dapat tunduk pada aturan pengendalian merger.
- Jika kolaborasi mengambil bentuk inisiatif industri dengan banyak peserta, mungkin untuk menetapkan standar, aturan yang lazim berlaku untuk memastikan bahwa standar tersebut tidak dimanfaatkan untuk mengecualikan beberapa pesaing, atau tidak digunakan untuk mengkartelisasi pasar.
- Jika inisiatif kolaborasi dan standarisasi melibatkan paten esensial standar, terdapat preseden dari sektor lain terkait pemberian lisensi paten esensial standar dengan ketentuan FRAND, dll.

AI sebagai Teknologi Sumber Terbuka

Meskipun banyak penelitian dan pengembangan AI swasta yang dapat dipertahankan hak kepemilikannya, pengembangan perangkat lunak sumber terbuka juga memainkan peran penting. Beberapa sistem pembelajaran mesin, pembelajaran mendalam, dan sistem AI terbaik lainnya tersedia di bawah lisensi sumber terbuka.

Sumber Terbuka

Filosofi yang mendasari perangkat lunak sumber terbuka adalah memungkinkan individu, lembaga pemerintah, dan perusahaan untuk berkolaborasi guna belajar dan berkreasi dengan lebih baik, bersama-sama.



Menurut IBM, tiga prinsip utama platform sumber terbuka adalah semakin pentingnya kode, konten, dan komunitas.

Namun, meskipun ketersediaan sistem AI secara terbuka jelas menarik dan menguntungkan bagi komunitas secara keseluruhan, apa daya tariknya dari perspektif bisnis? Ada banyak perusahaan besar yang berpartisipasi dalam proyek sumber terbuka, termasuk nama-nama besar di bidang ekonomi digital seperti Google, Facebook, Amazon, dan Microsoft, serta IBM.

Jawabannya mencakup pertimbangan berikut:

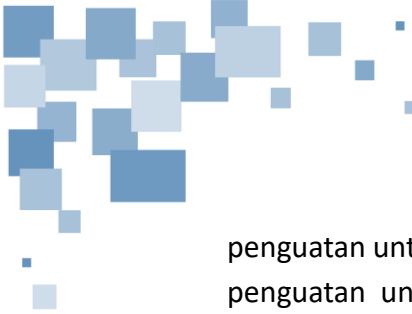
- ✓ Meskipun suatu bisnis mungkin kehilangan pendapatan lisensi dengan membuka akses ke perangkat lunak mereka, kualitasnya seringkali dapat ditingkatkan, fungsionalitasnya diperluas, dan umumnya ditingkatkan dengan biaya yang jauh lebih rendah daripada jika semua pekerjaan ini dilakukan secara internal. Proyek sumber terbuka menawarkan cara untuk mengakses kumpulan besar talenta tanpa biaya tenaga kerja.
- ✓ Proyek sumber terbuka juga dapat membantu membangun dukungan untuk teknologi tertentu yang dapat memberinya pangsa pasar yang lebih besar dan merangsang lebih banyak bisnis untuk mengembangkan sistem atau produk yang kompatibel (misalnya, elemen inti sistem operasi seluler Android Google adalah sumber terbuka).
- ✓ Terakhir, banyak pihak di komunitas pengembangan merasa bahwa karena mereka sendiri sering diuntungkan dengan dapat menggunakan perangkat lunak sumber terbuka, mereka seharusnya membalas budi dengan merilis kode yang mereka kembangkan sendiri. Filosofi ini terlihat di seluruh ekosistem pengembangan perangkat lunak, tetapi juga khususnya dalam kaitannya dengan sistem AI.

Selain platform Android, Google telah merilis kerangka kerja AI TensorFlow sebagai perangkat lunak sumber terbuka. TensorFlow adalah mesin di balik Google Images dan sistem pengenalan suara yang terdapat di aplikasi Google.

Proyek-proyek sumber terbuka juga bermunculan dari ekosistem Tiongkok yang biasanya tertutup. Baidu Research Institute telah merilis sistem AI sumber terbuka yang menawarkan kemampuan deteksi patologi tumor, menggunakan pembelajaran mendalam untuk meningkatkan analisis citra biopsi yang ada secara signifikan.

Algoritma ini akan ditambahkan ke mesin yang sudah ada yang menganalisis slide patologi dan akan mendapatkan akses ke kumpulan data dari rumah sakit di Tiongkok untuk memperluas penerapannya dari kanker payudara ke jenis kanker lainnya, tentu saja dengan tetap memperhatikan perkembangan peraturan lebih lanjut. Baidu juga merilis teknologi sistem mengemudi otonomnya sebagai perangkat lunak sumber terbuka pada tahun 2017. Pada awal 2019, Baidu melaporkan 130 mitra global menggunakan platformnya untuk proyek mengemudi otonom mereka.

Ada juga inisiatif sektor swasta tertentu yang mempromosikan akses terbuka ke AI. Ini termasuk OpenAI, yang bertujuan untuk mengembangkan dan berbagi sistem AI inovatif melalui kolaborasi gratis. Proyek-proyek yang telah diluncurkan antara lain merilis perangkat untuk membangun sistem AI melalui pembelajaran penguatan, menggunakan pembelajaran



penguatan untuk menghadapi pemain profesional di Dota 2, dan menggunakan pembelajaran penguatan untuk melatih tangan robot dengan "ketangkasan yang belum pernah terjadi sebelumnya." Mereka juga telah mengeluarkan laporan peringatan tentang risiko seputar penggunaan AI yang berbahaya. Sebuah proyek besar baru-baru ini melibatkan pengembangan sistem AI pembangkit teks yang tampaknya begitu bagus sehingga belum dirilis ke publik "karena kekhawatiran kami tentang aplikasi berbahaya dari teknologi tersebut."

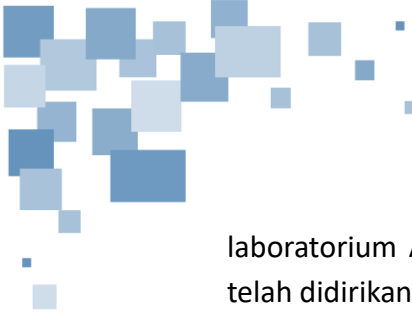
Tentu saja, dalam banyak kasus, sebuah bisnis akan menangani berbagai sistem AI-nya dengan cara yang berbeda, tergantung pada kepentingan strategis dan tahap pengembangannya. Sebagaimana telah disebutkan, banyak platform menawarkan AI sebagai Layanan sebagai bagian dari perangkat lunak atau layanan hosting mereka yang lebih luas. Perangkat lunak lain dapat dirilis sebagai sumber terbuka untuk mendorong produk dan perangkat lunak terkait. Namun, beberapa teknologi akan tetap menjadi hak milik dan kerahasiaan.

Model Investasi Sektor Publik untuk Merangsang Pertumbuhan dan Persaingan

Bagian ini membahas inisiatif publik yang lebih luas untuk mendukung pengembangan AI secara lebih umum. Dalam buku putih Strategi Industrinya, BEIS Inggris mengidentifikasi AI dan data sebagai salah satu dari empat "Tantangan Besar" yang harus dikuasai Inggris. Selain itu, laporan Hall/Pesenti Inggris memberikan rekomendasi tentang bagaimana sektor publik dan swasta dapat berkolaborasi dalam keterampilan dan infrastruktur, guna menerapkan strategi AI yang efektif di Inggris. Pemerintah Inggris kemudian menetapkan Kesepakatan Sektor AI pada tahun 2018 sebagai komitmen pertama dari pemerintah dan industri untuk bekerja sama mengembangkan industri AI melalui fitur-fitur seperti paket dukungan senilai Rp 9.5 triliun guna merangsang akses ke data dan membangun tenaga kerja terampil. Pemerintah Inggris bertujuan untuk terus merangsang pertumbuhan industri yang ada yang dipelopori oleh investasi di Inggris dari para pemimpin AI seperti DeepMind dan Amazon.

Komisi Eropa, di bawah strategi Pasar Tunggal Digital UE, telah mengajukan inisiatif dan kebijakan berskala luas yang mendorong industri pengembangan AI yang kompetitif. Dalam Komunikasi tentang Kecerdasan Buatan untuk Eropa pada 25 April 2018, Komisi Eropa mengajukan pendekatan Eropa terhadap AI berdasarkan tiga pilar. Ini termasuk meningkatkan investasi tahunan dalam AI sebesar 70% di bawah program Horizon 2020, dengan tujuan mencapai Rp 15 triliun antara tahun 2018 dan 2020. Komisi Eropa juga akan mendukung kemitraan bisnis-pendidikan dan pelatihan tenaga kerja untuk mempersiapkan perubahan sosial-ekonomi yang dibawa oleh AI. Terakhir, Komisi Eropa akan mengembangkan kerangka etika dan hukum untuk AI. Fokus pada inovasi AI terus tumbuh di UE dengan lebih banyak inisiatif seperti Aliansi AI yang didirikan oleh Komisi Eropa dan Basis Pengetahuan INSPIRE untuk berbagi data spasial lingkungan di antara organisasi sektor publik.

Di Tiongkok, proyek inovasi kolaboratif antara entitas pemerintah dan perusahaan teknologi menunjukkan bahwa kombinasi sektor publik-swasta dapat menuai hasil di masa depan. Kota Beijing berencana menginvestasikan Rp 20 triliun dalam sebuah taman pengembangan AI, yang akan menampung hingga 400 perusahaan AI dan sebuah



laboratorium AI nasional untuk mendorong pengembangan di sektor ini. Sebuah taman AI telah didirikan di salah satu dari sekian banyak kota yang berpusat pada teknologi di Tiongkok, Hangzhou, yang didukung oleh dana sebesar Rp 100 triliun.

Kemajuan lain dalam kolaborasi inovatif antara sektor swasta dan publik adalah proyek-proyek AI yang belum pernah terjadi sebelumnya dalam sistem peradilan: pada November 2017, pemerintah Tiongkok mengumumkan rencana untuk membangun platform AI nasional dalam kemitraan dengan empat perusahaan. Bersama BAT, spesialis AI iFlyTek bisa dibilang telah membuat kemajuan paling pesat di bidang ini. Memimpin proyek "Kecerdasan Buatan + Pengadilan", yang berfokus pada sistem pengarsipan cerdas, pengenalan suara, teks, dan gambar, serta sistem manajemen persidangan dan kasus yang cerdas, iFlyTek telah membantu mengurangi waktu sidang pengadilan hingga 30% di Tiongkok.

7.3 POTENSI AI UNTUK MENIMBULKAN KERUGIAN PERSAINGAN USAHA

Ekonomi digital telah mengubah secara radikal cara bisnis berinteraksi dan interaksi dengan konsumen. Meningkatnya penggunaan AI untuk mengotomatiskan proses dan pengambilan keputusan membawa perubahan lebih lanjut. Sebagaimana disebutkan di atas, fakta bahwa hukum persaingan usaha berbasis prinsip telah membuatnya fleksibel dan adaptif terhadap strategi komersial dan model bisnis baru, meskipun beberapa pihak kini mempertanyakan apakah hal ini masih berlaku. Bagian terakhir ini akan membahas isu-isu persaingan usaha yang dapat timbul dari penerapan sistem AI di lingkungan komersial dan dalam interaksi antara pelaku pasar, serta dengan pelanggan mereka.

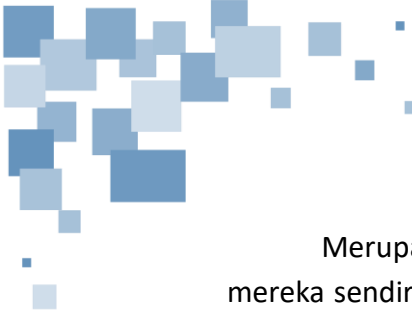
Tinjauan Umum Potensi Isu

Ada berbagai area di mana sistem AI dapat menimbulkan perilaku pasar yang dapat melanggar hukum persaingan usaha. Banyak perhatian khusus berfokus pada dampak sistem penetapan harga algoritmik, termasuk pemantauan harga, penyesuaian harga, dan perangkat penetapan harga yang dipersonalisasi.

Laporan Furman dari UK baru-baru ini mempertimbangkan dampak sistem AI terhadap persaingan dan mengamati: "Saat ini, sulit untuk memprediksi apakah penggunaan algoritma yang lebih besar akan mengarah pada kolusi algoritmik atau penetapan harga yang dipersonalisasi di masa mendatang, dan tidak ada bukti bahwa penetapan harga yang dipersonalisasi yang merugikan tersebar luas. Namun, ini adalah area yang berpotensi berkembang pesat, di mana penting untuk tetap waspada terhadap potensi kerugian." Pemantauan lebih lanjut direkomendasikan.

Perangkat AI yang Mendukung Praktik Penetapan Harga Ilegal

Perangkat lunak pencocokan harga digunakan secara komersial secara luas. Komisi Uni Eropa menemukan dalam laporan akhir tahun 2017 tentang penyelidikan sektor e-commerce bahwa mayoritas pengecer melacak harga pesaing mereka menggunakan sistem otomatis dan dua pertiga juga menyesuaikan harga mereka sendiri secara otomatis sebagai respons. Sistem ini sendiri tidak dapat diterima, tetapi dapat digunakan untuk memfasilitasi atau memperparah perilaku ilegal.



Merupakan tindakan ilegal bagi bisnis independen untuk menetapkan harga di antara mereka sendiri penetapan harga harus diputuskan secara sepihak berdasarkan kondisi pasar dan pertimbangan keuangan bisnis yang bersangkutan (seperti biaya tetap dan variabel, strategi promosi, kebijakan diskon, dll.). Larangan ini berlaku (a) bagi pesaing yang ingin menetapkan harga di antara mereka dan juga (b) bagi pemasok, yang tidak dapat mencegah distributor mereka menetapkan harga mereka sendiri (meskipun harga eceran yang direkomendasikan atau maksimum dapat disepakati).

Sehubungan dengan penetapan harga, perangkat lunak telah didokumentasikan telah digunakan untuk menerapkan pengaturan kartel. Departemen Kehakiman AS dan Otoritas Persaingan dan Pasar Inggris keduanya menyetujui kartel antara dua bisnis yang menjual poster di Amazon Marketplace yang menggunakan perangkat lunak penetapan harga untuk memastikan harga mereka selaras dengan pengaturan kartel mereka.

Sehubungan dengan pemeliharaan harga jual kembali, Komisi Uni Eropa mencatat dalam laporan sektor e-commerce-nya bahwa: "Dengan perangkat lunak penetapan harga, mendeteksi penyimpangan dari harga eceran yang 'direkomendasikan' hanya membutuhkan waktu beberapa detik dan produsen semakin mampu memantau dan memengaruhi penetapan harga pengecer." Perangkat lunak penetapan harga algoritmik juga dapat menyebabkan masalah karena memperburuk dampak penetapan harga ilegal.


Komisi saat ini sedang menyelidiki apakah pemasok berbagai barang elektronik konsumen telah mencegah pengecer daring menetapkan harga mereka sendiri, dan kekhawatiran khusus adalah bahwa "dampak dari dugaan pembatasan harga ini dapat diperburuk oleh penggunaan perangkat lunak penetapan harga oleh banyak pengecer daring yang secara otomatis menyesuaikan harga eceran dengan harga pesaing terkemuka." Kekhawatirannya tampaknya adalah bahwa para pemasok dapat menghasilkan dampak yang luas pada harga pasar dengan mengendalikan harga beberapa pengecer hilir utama, karena alat pencocokan harga otomatis yang memantau pemasok utama tersebut akan mengikuti jejak mereka.

AI Memfasilitasi Kolusi Harga yang Legal Tetapi Merugikan

Kolusi Terselubung

Salah satu area yang paling sulit dalam penegakan hukum persaingan adalah ketika tidak ada "pertemuan" aktif untuk mencapai kesepakatan harga, tetapi para pesaing tetap bertindak secara paralel. Hal ini dikenal sebagai "kolusi terselubung".

Kecuali terdapat bukti koordinasi yang eksplisit, kolusi diam-diam tidaklah ilegal, meskipun dapat merugikan pasar. Ini bukanlah masalah baru contoh klasiknya adalah SPBU-SPBU tetangga yang saling mengawasi harga satu sama lain dengan ketat. Tidaklah ilegal jika SPBU B, melihat harga SPBU A naik, memutuskan untuk mengikuti dan menaikkan harga (dan margin) mereka juga, alih-alih mempertahankan harga tetap rendah dan memenangkan bisnis dari SPBU A. Akan ilegal bagi SPBU-SPBU tersebut untuk menyepakati bahwa mereka berdua



harus menaikkan harga tetapi tidak menaikkan harga karena keputusan independen mereka. Akan sangat sulit untuk mengidentifikasi kapan penetapan harga yang selaras merupakan hasil adaptasi cerdas yang sah terhadap perilaku pesaing yang ada atau yang diantisipasi, versus kapan penetapan harga tersebut merupakan konsekuensi dari kontak langsung atau tidak langsung dengan tujuan memengaruhi perilaku pesaing secara ilegal.

Divisi Persaingan OECD mengadakan diskusi meja bundar tentang "Algoritma dan persaingan" pada Juni 2017. Makalah latar belakang membahas risiko bahwa sistem AI dapat mempermudah pencapaian dan mempertahankan kolusi diam-diam, dalam berbagai cara (risiko penetapan harga yang harmonis dari penggabungan data melalui sistem AI).

Persaingan harga akan meredam jika bisnis yang bersaing menggunakan algoritma yang sama untuk menetapkan strategi penetapan harga mereka, karena sistem tersebut kemungkinan akan menghasilkan respons yang sama terhadap perubahan pasar. Perangkat lunak yang responsif dapat mempermudah pengujian apakah pesaing akan mengikuti sinyal kenaikan harga, serta untuk mempertahankan kepatuhan terhadap perjanjian penetapan harga di pasar yang dinamis di mana harga sering berubah. Makalah OECD mencatat bahwa algoritma penetapan harga pembelajaran mendalam bahkan dapat menghasilkan kartel, baik secara nyata maupun substansi, yang sama sekali tidak disadari oleh manajemen. Saat ini, hal ini tidak akan ilegal karena harus ada unsur koordinasi yang eksplisit.

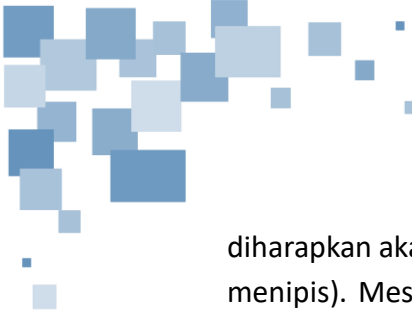
Oleh karena itu, masalahnya adalah penetapan harga algoritmik dapat membuat kolusi diam-diam lebih umum di lebih banyak pasar dan hal itu tidak akan ilegal, meskipun akan merugikan pelanggan karena persaingan harga normal tidak akan beroperasi. Kesulitannya terletak pada inti norma hukum persaingan prasyarat untuk ilegalitas perilaku terkoordinasi para pesaing adalah adanya kesepakatan di antara mereka, betapapun minimalnya hal itu diungkapkan.

Personalisasi

Masalah terpisah adalah penetapan harga yang dipersonalisasi, yang difasilitasi oleh sistem AI. Buku Hijau BEIS Inggris tentang Modernisasi Pasar Konsumen mencatat bahwa data konsumen dapat digunakan oleh perusahaan untuk mempersonalisasi pengalaman konsumen daring, sehingga konsumen diperlihatkan harga yang berbeda untuk produk yang sama, dengan "setiap konsumen dikenakan harga yang merupakan fungsi dari tetapi tidak harus sama dengan kesediaan mereka untuk membayar." Seringkali pendorong utama penetapan harga adalah kemauan masing-masing pelanggan untuk membayar.

Diskriminasi antarpelanggan tidak melanggar hukum persaingan usaha kecuali pemasok berada dalam posisi pasar yang kuat (dan mungkin masih dapat diterima jika perbedaan perlakuan tersebut dapat dibenarkan secara objektif). Apabila pemasok tidak memiliki posisi pasar yang kuat (atau sebagaimana dicatat oleh OECD, di mana perilaku penyalahgunaan oleh bisnis dominan tidak dilarang atau jarang diselidiki), hal ini lebih merupakan masalah hukum konsumen daripada masalah hukum persaingan usaha.

Masalah lain yang dapat diprediksi terkait sistem AI dan hukum persaingan adalah bagaimana AI di masa depan dapat digunakan untuk mendorong perangkat terhubung "pintar" yang mencakup pembelian otomatis (misalnya, pengembangan lemari es pintar yang



diharapkan akan secara otomatis memesan ulang susu ketika sensor menunjukkan persediaan menipis). Meskipun konsumen mungkin menyambut kemudahan ini, terdapat risiko bahwa sifat pasif dari pembelian akan menyebabkan harga melonjak. Kemampuan konsumen untuk beralih ke pemasok lain merupakan salah satu pendorong persaingan harga: kurangnya perhatian terhadap pembelian otomatis jelas dapat melemahkan daya beli konsumen yang berlawanan.

Personalisasi harga juga dapat mengurangi transparansi harga. Jika sulit untuk mengetahui harga pesaing, mungkin akan lebih sulit untuk memenangkan pelanggan dengan penawaran yang kompetitif. Jika personalisasi dimungkinkan oleh profil pelanggan yang memanfaatkan big data milik perusahaan, masalah ini dapat diperburuk bagi pesaing yang tidak memiliki data serupa dan tidak dapat mempersonalisasi penawaran harga mereka dengan tingkat yang sama. Potensi peningkatan hambatan masuk/ekspansi ini merupakan kekhawatiran persaingan.

Aspek lain dari personalisasi adalah penggunaan AI untuk menghasilkan umpan personalisasi berisi produk atau layanan yang diusulkan kepada pelanggan tertentu. Berdasarkan pola belanja individu sebelumnya, klik sebelumnya, "suka" yang terkonfirmasi, dll., sistem AI dapat mengusulkan produk tertentu. Parlemen Kanada telah mempertimbangkan dampak personalisasi dan apakah umpan personalisasi bahkan dapat menghasilkan kebiasaan yang tidak disadari.

Kekhawatiran persaingan dalam konteks ini adalah, (mirip dengan "ruang gema" atau gelembung filter yang dapat diciptakan oleh umpan berita yang dipersonalisasi), personalisasi penawaran dapat berfungsi untuk mempersulit pemasok alternatif mendapatkan perhatian pelanggan tersebut. Masalah ini baru-baru ini disoroti oleh Komisaris Uni Eropa Margrethe Vestager.

Sebagaimana dijelaskan, tidak satu pun dari situasi ini yang mungkin dikategorikan ilegal (tanpa kekuatan pasar), tetapi dampak penetapan harga yang didorong oleh AI terhadap kondisi pasar tetap dapat bersifat negatif. Otoritas persaingan di Inggris dan banyak negara Uni Eropa memiliki wewenang untuk melakukan investigasi pasar atau penyelidikan sektor, yang memungkinkan otoritas tersebut meninjau kesehatan seluruh pasar, meskipun tidak ada perilaku ilegal. Strategi penetapan harga legal yang meredam kekuatan pasar (seperti dalam contoh-contoh ini) merupakan target potensial untuk investigasi semacam itu.

Efek Jaringan dan Platform yang Kuat

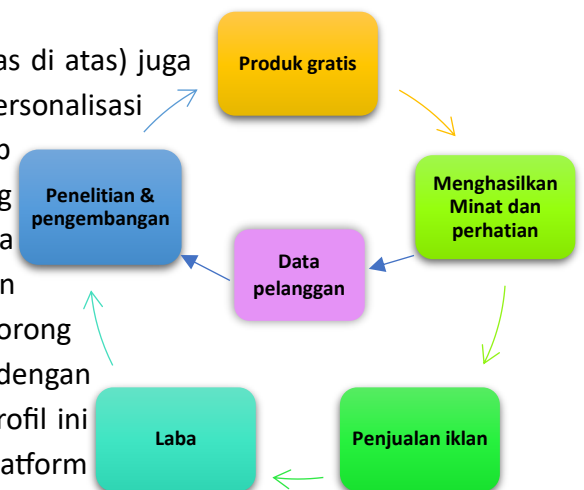
Terakhir, sistem AI menciptakan tantangan bagi penegakan persaingan karena sistem tersebut menopang kekuatan platform teknologi yang menggunakan big data dan AI sebagai input inti. Banyak platform ini menikmati efek jaringan: semakin banyak pengguna yang menggunakan platform, semakin banyak data yang dapat dikumpulkan untuk meningkatkan AI yang mendasarinya, yang pada gilirannya membuat platform beroperasi lebih baik, menarik lebih banyak orang, dan sebagainya.

Untuk media sosial, semakin banyak penggunaanya, semakin banyak pula orang yang tertarik, yang sekali lagi, meningkatkan daya tariknya. Banyak bisnis teknologi besar beroperasi di pasar dua sisi di mana produk dapat diperoleh secara gratis, menghasilkan minat dan

perhatian yang kemudian mendorong penjualan iklan, yang memungkinkan lebih banyak produk diberikan kepada konsumen secara gratis. Lingkaran positif ini biasanya membuat platform berfungsi semakin baik. Namun, efek jaringan sulit dipahami dalam konteks hukum persaingan. Persaingan mulai bersifat "untuk pasar" alih-alih "di dalam pasar" yaitu, terdapat efek "pemenang mengambil semuanya" atau "pemenang mengambil sebagian besar". Pesaing yang lebih kecil mungkin kesulitan untuk memenangkan pangsa pasar, terutama jika pelanggan tidak cenderung "multi-home" dan menggunakan lebih dari satu platform untuk fungsi yang sama. Oleh karena itu, kekuatan pasar dapat diredam dan persaingan mungkin tidak sepenuhnya efektif.

Sekali lagi, ini bukanlah masalah baru (banyak pasar telah direstrukturisasi untuk menghadirkan lebih banyak persaingan, termasuk pasar utilitas, bursa saham, dll.). Namun, ini merupakan cara lain AI dapat beroperasi untuk membentuk pasar yang sulit untuk diawasi oleh hukum persaingan.

Personalisasi dan pembuatan profil (yang dibahas di atas) juga dapat berfungsi untuk memperkuat efek jaringan. Personalisasi dapat meningkatkan loyalitas pengguna terhadap platform tertentu karena menghasilkan konten yang bermanfaat dan sesuai dengan preferensi pengguna tersebut. Sebagaimana yang didengar oleh parlemen Kanada, insentif ekonomi dari penjualan iklan mendorong keinginan untuk memaksimalkan keterlibatan pengguna dengan suatu platform. Namun, loyalitas yang didorong oleh profil ini berpotensi semakin meningkatkan tantangan bagi platform pesaing yang tidak mampu menawarkan tingkat personalisasi yang serupa.




Salah satu solusi yang diusulkan untuk masalah ini seringkali berfokus pada portabilitas data. Akses ke data individual yang dimiliki oleh platform A tentang seorang pengguna dapat memudahkan platform B untuk menarik pengguna tersebut jika data yang dikumpulkan oleh platform A dapat ditransfer agar platform B dapat menawarkan tingkat personalisasi yang serupa, dll.

Tinjauan Tentang Kesesuaian Rezim Persaingan, Pendekatan Analitis, dan Perangkat Penegakan Hukum yang Ada dengan Tujuan di Era Digital

Tantangan yang ditimbulkan oleh pasar ekonomi digital secara umum dan AI secara khusus sedang ditinjau dari perspektif kebijakan di sejumlah negara. Di tingkat Uni Eropa, Komisaris yang bertanggung jawab atas hukum persaingan, Margrethe Vestager, telah menunjuk tiga penasihat khusus yang sedang mempertimbangkan "tantangan digitalisasi di masa depan untuk kebijakan persaingan" dan akan melaporkan temuan mereka pada awal April 2019. Sebuah konferensi diadakan pada Januari 2019 untuk membahas isu-isu tersebut. Di Inggris:

- o Buku Hijau "Modernisasi Pasar Konsumen" dari Departemen Bisnis, Energi, dan Strategi Industri Inggris pada April 2018 mencakup bab tentang pasar digital dan



bertanya: “Tantangan apa yang ditimbulkan pasar digital bagi penegakan persaingan yang efektif dan apa yang dapat dilakukan untuk mengatasinya?” Buku Hijau tersebut juga meminta pandangan tentang bagaimana otoritas persaingan harus menangani “platform digital, aglomerasi, algoritma data, dan konsolidasi pesaing.”

- Tinjauan Furman di Inggris ditugaskan untuk mengkaji “peluang dan tantangan potensial yang mungkin ditimbulkan oleh ekonomi digital yang sedang berkembang bagi persaingan dan kebijakan pro-persaingan.” Laporan panel ahli diterbitkan pada Maret 2019. Sebagaimana telah disebutkan di berbagai poin di atas, laporan Furman menyimpulkan bahwa rezim persaingan dan perangkat penegakan hukum yang ada tidak memadai untuk memastikan persaingan yang efektif dalam ekonomi digital demi kepentingan konsumen, dan menawarkan beberapa gagasan luas tentang cara mengatasi masalah ini. Pemerintah Inggris belum memberikan tanggapan.

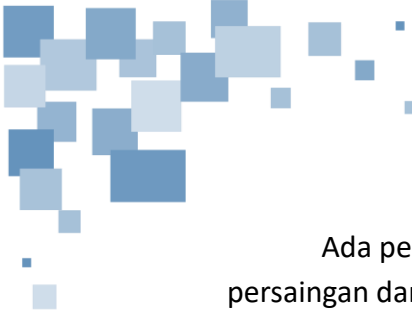
Di Jerman, sebuah laporan oleh Komisi Monopoli menemukan bahwa kerangka hukum di Jerman perlu beradaptasi dengan perubahan digital dalam perekonomian, dan merekomendasikan agar pasar dengan penetapan harga berbasis algoritma diselidiki secara sistematis untuk mengetahui dampak buruknya terhadap persaingan.

Sebagai konsekuensinya, Kementerian Federal untuk Urusan Ekonomi dan Energi telah membentuk komite hukum persaingan usaha, yang dikenal sebagai “*Kommission Wettbewerbsrecht 4.0*” yang akan memberikan rekomendasi tentang perlu tidaknya reformasi hukum persaingan usaha Uni Eropa dan nasional. Lapornya diharapkan terbit pada musim gugur 2019.

Di Kanada, Dewan Rakyat meminta Pemerintah untuk mempertimbangkan dampak kebijakan data dan mempertimbangkan perlu tidaknya modernisasi Undang-Undang Persaingan Usaha Kanada. Di tingkat internasional, sebagaimana disebutkan di atas, OECD telah menyelenggarakan sejumlah diskusi panel tentang berbagai aspek isu-isu ini, termasuk “Big data: Membawa kebijakan persaingan usaha ke era digital”, “Memikirkan kembali penggunaan alat penegakan hukum antimonopoli tradisional di pasar multi-sisi”, dan “Algoritma dan kolusi”, “Implikasi E-commerce terhadap Kebijakan Persaingan Usaha”, dan “Penetapan Harga yang Dipersonalisasi di Era Digital”.

Ada harapan luas bahwa sistem AI akan memiliki dampak transformatif di semua sektor ekonomi dan dalam semua aspek kehidupan kita, menghasilkan kemudahan dan kesederhanaan, wawasan dan penemuan, pemahaman dan efisiensi. Namun, agar potensi transformatif ini dapat terwujud sepenuhnya, inovasi yang dinamis dan kompetitif dalam aplikasi AI sangat penting.

Pemerintah, pembuat kebijakan, dan otoritas penegak hukum tidak boleh membiarkan kebutuhan akan data dalam jumlah besar dan/atau kompleksitas teknologi membatasi pengembangan dan penerapan AI. Ada kekhawatiran yang semakin meningkat bahwa bisnis dengan kumpulan data besar dan keunggulan awal dalam membangun aplikasi AI dapat memperoleh keuntungan yang sulit dibantah, dengan kekhawatiran yang sepadan bahwa hal ini, dalam kasus terburuknya, dapat menghambat inovasi.



Ada perdebatan (dan belum ada konsensus) mengenai apakah kerangka kerja legislatif persaingan dan antimonopoli yang ada cukup adaptif untuk menawarkan pencegah yang kuat dan solusi yang efektif terhadap distorsi pasar apa pun yang mungkin muncul di masa mendatang.

Data terbuka dan perangkat lunak sumber terbuka merupakan tren penting dalam hal ini yang harus didorong dan didukung. Pemerintah harus memimpin dengan memberi contoh, jika memungkinkan, dalam membuka kumpulan data (dengan persetujuan, pseudonimisasi, atau agregasi, dll., jika diperlukan) dan dengan berinvestasi dalam sumber daya, fasilitas, dan pusat keunggulan pengembangan AI.

Mengingat pentingnya AI secara strategis, otoritas penegak hukum persaingan harus secara proaktif memantau perkembangan teknologi ini dan aktivitas mereka yang berada di garda terdepan pengembangannya sebagai prioritas administratif. Debat dan diskusi internasional untuk mengembangkan praktik terbaik melalui forum seperti OECD atau Jaringan Persaingan Internasional harus sangat didorong.

Prinsip

Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI dan undang-undang nasional yang mengatur penggunaannya wajib mendorong (a) akses terbuka terhadap kumpulan data yang dapat digunakan dalam pengembangan sistem AI dan (b) kerangka kerja dan perangkat lunak sumber terbuka untuk sistem AI. Sistem AI harus dikembangkan dan diterapkan berdasarkan prinsip "kepatuhan sejak awal" terkait dengan hukum persaingan usaha/antimonopoli.

Mendukung Persaingan yang Efektif Terkait Sistem AI

- Pemerintah harus mendukung dan berpartisipasi dalam koordinasi internasional (melalui badan-badan seperti OECD dan Jaringan Persaingan Internasional) untuk mengembangkan praktik terbaik dan analisis yang cermat dalam memahami dampak persaingan dari pengendalian kumpulan data dan sistem AI terhadap pasar ekonomi.
- Pemerintah harus melakukan tinjauan berkala untuk memastikan bahwa kerangka kerja hukum persaingan usaha dan perangkat penegakan hukum yang tersedia bagi otoritas penegak hukum terkait memadai dan efektif untuk memastikan akses yang memadai terhadap masukan yang diperlukan, serta pilihan yang memadai, persaingan yang dinamis, inovasi kreatif, dan keluaran berkualitas tinggi dalam pengembangan dan penerapan sistem AI, demi manfaat konsumen secara keseluruhan.

Data Terbuka

- Pemerintah harus mendorong dan memfasilitasi infrastruktur nasional yang diperlukan untuk mendorong akses terbuka terhadap kumpulan data bagi seluruh elemen masyarakat yang berkepentingan dalam mengakses kumpulan data tersebut untuk keperluan penelitian dan/atau non-komersial. Dalam hal ini, pemerintah harus mempertimbangkan secara serius model akses dua tingkat yang memungkinkan akses gratis untuk tujuan akademis dan penelitian, dan akses berbayar untuk tujuan komersial.

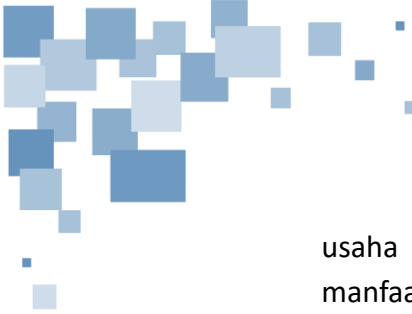
- Pemerintah harus mendukung inisiatif data terbuka di sektor publik maupun swasta dengan panduan dan penelitian untuk berbagi pemahaman yang luas tentang manfaat yang dapat diperoleh dari data akses terbuka, struktur yang memungkinkan kumpulan data dibagikan dan dipertukarkan, serta proses yang memungkinkan data tersebut dibuat sesuai untuk akses terbuka (termasuk standarisasi API, pseudonimisasi, agregasi, atau kurasi lainnya, jika diperlukan).
- Pemerintah harus memastikan bahwa data yang dipegang oleh badan sektor publik dapat diakses dan terbuka, jika memungkinkan, dan jika hal ini tidak bertentangan dengan mandat sektor publik untuk memulihkan investasi wajib pajak dalam pengumpulan dan kurasi data tersebut. Badan sektor swasta seperti organisasi industri dan asosiasi perdagangan juga harus mendukung dan mempromosikan data terbuka dalam sektor industri mereka, dengan membuka kumpulan data mereka sendiri, jika memungkinkan.
- Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI didorong untuk membuka akses dan/atau melisensikan kumpulan data mereka, jika memungkinkan melalui mekanisme pendampingan seperti Data Trust.
- Setiap pembagian atau pemberian lisensi data harus dilakukan pada tingkat yang wajar dalam keadaan tertentu dan harus mematuhi kewajiban atau persyaratan hukum, peraturan, kontrak, dan kewajiban atau persyaratan lainnya terkait dengan data yang bersangkutan (termasuk privasi, keamanan, kebebasan informasi, dan pertimbangan kerahasiaan lainnya).

Sistem AI Sumber Terbuka

- Organisasi yang mengembangkan sistem AI biasanya berhak mengomersialkan sistem tersebut sesuai keinginan mereka. Namun, pemerintah setidaknya harus mengadvokasi aksesibilitas melalui sumber terbuka atau pengaturan lisensi serupa lainnya terhadap sistem AI inovatif yang mungkin memberikan manfaat sosial tertentu atau memajukan "keadaan terkini" di bidangnya melalui, misalnya, skema insentif yang ditargetkan.
- Organisasi yang memilih untuk tidak merilis sistem AI mereka sebagai perangkat lunak sumber terbuka tetap didorong untuk melisensikan Sistem tersebut secara komersial.
- Sejauh sistem AI dapat dibagi lagi menjadi beberapa bagian penyusun dengan utilitas dan aplikasi umum dalam kasus penggunaan AI lainnya, organisasi yang memilih untuk tidak melisensikan sistem AI secara keseluruhan (baik secara sumber terbuka maupun komersial) didorong untuk melisensikan sebanyak mungkin komponen yang dapat digunakan kembali tersebut.

Kepatuhan Berdasarkan Rancangan terhadap Hukum Persaingan Usaha/Antimonopoli

- ✓ Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus merancang, mengembangkan, dan menerapkan sistem AI dengan cara "kepatuhan berdasarkan rancangan" yang memastikan konsistensi dengan etos menyeluruh dari rezim persaingan usaha/antimonopoli yang berlaku untuk mendorong persaingan



usaha yang bebas dan dinamis di antara perusahaan-perusahaan korporat demi manfaat konsumen yang sebesar-besarnya.



BAB 8 PRIVASI

Tren utama dalam penerapan teknologi kecerdasan buatan (AI) adalah kesesuaiannya untuk membantu kita mengelola dan memanfaatkan data dalam jumlah besar: basis data besar secara inheren cocok untuk penggunaan aplikasi AI untuk pencarian dan analitik....

...dan jika terdapat basis data besar, hal ini hampir selalu mencakup data pribadi.

Memang, seiring organisasi analitik data mengejar dorongan ekonomi menuju monetisasi data, seringkali data pribadilah yang paling banyak mendapat perhatian; Mendorong pemanfaatan kreatif "danau" data pribadi dengan menggunakan teknologi baru, khususnya AI. Tentu saja, dalam beberapa tahun terakhir juga terdapat tren peningkatan regulasi tentang bagaimana data pribadi kita dikelola, yang memunculkan pertanyaan jelas tentang bagaimana kedua tren paralel ini bersaing.

Dalam konteks privasi data, masalah muncul terutama karena hilangnya kendali manusia dengan "algoritma yang bertindak sebagai agen." Namun, ada juga masalah lain. Berikut ini adalah ikhtisar masalah-masalah ini dan bagaimana pemikiran saat ini. Kami mulai dengan ikhtisar dasar-dasar hukum privasi.

8.1 DASAR-DASAR PRIVASI

Pendahuluan

Di bagian tentang 'dasar-dasar' hukum privasi ini, kami merangkum beberapa prinsip umum di seluruh dunia yang mencakup area ini dan bagaimana prinsip-prinsip tersebut mengatur AI. Kami sengaja mempertahankan hal ini pada tingkat yang tinggi untuk mempersiapkan analisis selanjutnya.

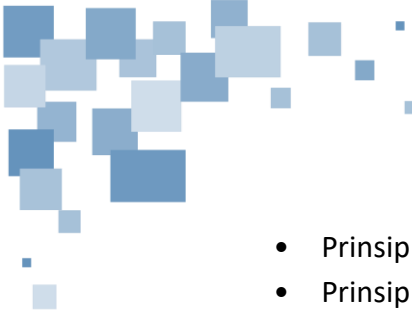
Prinsip-prinsip Privasi Umum

Privasi adalah konsep yang sulit dipahami, terlebih lagi karena kemajuan teknologi yang pesat di era digital. Bagaimana seseorang membedakan antara apa yang "publik" dan apa yang "pribadi", dan akibatnya, apa yang perlu dilindungi, di dunia di mana semakin banyak informasi pribadi dibagikan secara daring?

Selama lima puluh tahun terakhir, undang-undang privasi data telah mengalami transformasi yang signifikan agar dapat mengimbangi jumlah data yang dihasilkan secara daring yang terus meningkat. Prinsip-prinsip perlindungan data yang mendasari kerangka hukum ini bertujuan untuk memastikan informasi pribadi diminta, diperoleh, digunakan, dan dipertukarkan dengan cara yang sesuai dengan konteksnya.

Sebagian besar undang-undang yang mengatur privasi di seluruh dunia saat ini didasarkan pada prinsip-prinsip yang ditetapkan pada tahun 1980 oleh *Organisation for Economic Cooperation and Development* (OECD). OECD menyepakati delapan Praktik Informasi yang Adil, yang dikodifikasikan dalam *OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data*. Singkatnya, prinsip-prinsip tersebut adalah:

- Prinsip Pembatasan Pengumpulan Data

- 
- Prinsip Kualitas Data (akurasi)
 - Prinsip Spesifikasi Tujuan (transparansi)
 - Prinsip Pembatasan Penggunaan (proporsionalitas)
 - Prinsip Perlindungan Keamanan (keamanan)
 - Prinsip Keterbukaan (keadilan)
 - Prinsip Partisipasi Individu (hak subjek data)
 - Prinsip Akuntabilitas

Pengaruh utama lainnya adalah *European Convention on Human Rights* (ECHR), sebuah perjanjian internasional tahun 1953 yang melindungi hak asasi manusia dan kebebasan fundamental setiap orang di setiap negara anggota Dewan Eropa. Perjanjian ini mencakup, pada Pasal 8, hak untuk menghormati kehidupan pribadi dan keluarga; dan pada Pasal 10, perlindungan hak atas kebebasan berekspresi, termasuk hak untuk berpendapat dan menerima serta menyampaikan informasi.

Rezim Saat Ini

Baru-baru ini, pada tanggal 25 Mei 2018, Peraturan Perlindungan Data Umum (GDPR) mulai berlaku. GDPR adalah salah satu undang-undang paling kontroversial yang pernah disahkan melalui proses legislatif Eropa, dengan lobi yang ekstensif dari para pelaku bisnis dan pemangku kepentingan, termasuk pemerintah Amerika Serikat. Namun, Prinsip-Prinsip OECD tetap berlaku, dan prinsip pertama dalam GDPR mewajibkan pengendali data untuk memproses data secara sah, adil, dan transparan terkait dengan subjek data. Landasan hukumnya telah ditentukan, dan salah satu landasan hukum tersebut adalah persetujuan.

Sebagai contoh yang kontras mengilustrasikan bagaimana pendekatan historis dan budaya yang berbeda terhadap privasi menghasilkan solusi modern yang berbeda pula tidak ada satu pun rezim privasi yang menyeluruh di AS. Undang-undang federal di AS mencakup sektor industri tertentu dengan tujuan umum menjaga persaingan dan kepercayaan konsumen (contoh paling terkenal dan paling efektif adalah *Health Insurance Portability and Accountability Act* (HIPAA) tahun 1996, yang mengatur penggunaan data pasien medis), sementara undang-undang negara bagian bertujuan untuk melindungi konsumen, tetapi dibatasi oleh jangkauannya yang terbatas.

Sebagian besar (tetapi tidak semua) negara bagian telah memberlakukan undang-undang untuk melindungi data pribadi penduduk negara bagian, bahkan ketika pengumpul atau pengguna data berada di luar batas negara bagian. Namun, data yang dicakup bervariasi, dan kepatuhan serta penegakan hukum terganggu oleh cakupan yang beragam.

Di luar rezim privasi utama yang telah kita bahas, jumlah regulasi melonjak pesat, didorong oleh kebutuhan global untuk mengikuti perkembangan teknologi dan bersaing dalam perdagangan. Karena Uni Eropa dan Amerika Utara sudah tercakup dengan baik, sebagian besar pertumbuhan eksponensial baru ini terjadi di Asia, Timur Tengah, Afrika, dan Amerika Selatan. Pada akhir tahun 2018, sekitar 120 negara telah memiliki beberapa bentuk undang-undang privasi atau perlindungan data, dan sekitar 40 negara lainnya sedang dalam proses penyusunan undang-undang atau rancangan undang-undang.

Pendekatan Kami dalam Bab Ini

Ketika mempertimbangkan regulasi AI dan privasi saat ini dalam Bab ini, kami menggunakan GDPR sebagai 'patokan' tetap kami: GDPR merupakan gagasan para legislator Eropa untuk menciptakan kerangka kerja perlindungan data yang kuat dan lebih koheren, dengan mempertimbangkan perkembangan teknologi dan globalisasi yang pesat, yang membawa tantangan baru bagi perlindungan data pribadi.

GDPR tidak hanya menyentuh perspektif privasi AI di Eropa, tetapi juga merupakan hukum administratif pertama yang memungkinkan negara untuk menggunakan pembelajaran mesin (ML) dalam mendukung keputusan otomatis. GDPR juga merupakan seperangkat prinsip perlindungan data terkini yang ada. GDPR memberikan artikulasi yang mendalam tentang prinsip-prinsip perlindungan data abad ke-21 yang perlu diatasi oleh teknologi abad ke-21.

Melihat ke Depan

Ketegangan politik dan regulasi yang selalu ada antara melindungi warga negara dan menarik bisnis masih berlanjut. Namun, sebagai tren tandingan, di mana regulasi di negara-negara maju mewajibkan perusahaan memastikan data yang mereka transfer lintas batas tetap terlindungi secara memadai, negara-negara yang dapat menunjukkan rezim perlindungan data yang tepercaya memiliki keuntungan investasi misalnya untuk pusat data dan lokasi layanan pelanggan.

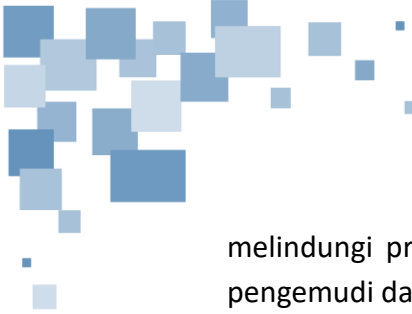
Mungkin ada tren menuju standar emas global; dengan GDPR yang kini menarik yurisdiksi lain untuk mengikutinya. 'Keputusan kecukupan' Uni Eropa tentang "kesetaraan esensial" dengan rezim Uni Eropa tersedia bagi negara ketiga (meskipun sulit dicapai) dan memberikan manfaat yang sangat besar. Setelah memodernisasi undang-undang perlindungan datanya selama beberapa tahun terakhir, Jepang berhasil mendapatkan 'keputusan kecukupan' pada Juli 2018. Andorra, Argentina, Kepulauan Faroe, Guernsey, Israel, Isle of Man, Jepang, Jersey, Selandia Baru, Swiss, dan Uruguay memiliki keputusan kecukupan Uni Eropa, sementara Kanada dan Amerika Serikat memiliki keputusan kecukupan parsial. Banyak dari rezim ini dibangun berdasarkan prinsip-prinsip OECD dan GDPR, yang mencakup persyaratan untuk memiliki dasar hukum untuk pemrosesan data pribadi (salah satunya adalah persetujuan).

Konsep-konsep yang ada, yang relevan dengan AI, akan diuraikan di bagian selanjutnya. Oleh karena itu, kami akan membahas: dampak AI terhadap anonimisasi, yang berisi janji untuk mendepersonalisasi data pribadi; keadilan dan isu-isu operasional terkaitnya; regulasi pengambilan keputusan dan pembuatan profil otomatis serta hak-hak subjek data; dan bagaimana AI dapat digunakan untuk memenuhi persyaratan privasi.

8.2 MASALAH PRIVASI DAN AI

Anonimisasi

AI memiliki potensi yang jelas dan nyata untuk melanggar data pribadi. Anonimisasi merupakan konsep hukum yang dapat membantu melindungi data pribadi, di samping pseudonimisasi dan pembatasan penyimpanan data. Bagaimana Anonimisasi membantu



melindungi privasi dijelaskan di bawah ini dalam studi kasus terkait mobil otonom tanpa pengemudi dan data kesehatan.

1. Anonimisasi Sebagai Konsep Hukum

Data pribadi adalah informasi apa pun yang berkaitan dengan orang perseorangan yang teridentifikasi atau dapat diidentifikasi. Kriteria yang relevan adalah "hubungan". Jika hubungan antara informasi dan orang tersebut terputus, maka informasi tersebut bukan lagi data pribadi. Hukum Uni Eropa dan banyak undang-undang perlindungan data nasional di belahan bumi barat menerapkan definisi ini.

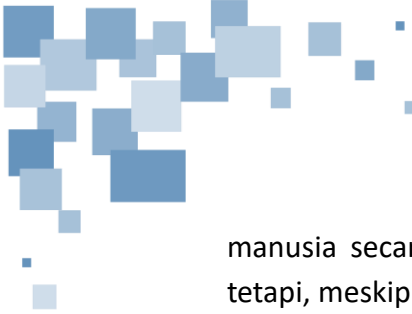
"Anonimisasi" berarti tidak mungkin untuk mengidentifikasi seseorang dari data itu sendiri atau dari data yang digabungkan dengan data lain, dengan mempertimbangkan semua cara yang secara wajar mungkin digunakan untuk mengidentifikasi individu tersebut. Ketika data telah dianonimkan, artinya data tersebut tidak dapat lagi mengidentifikasi seseorang dan, oleh karena itu, data tersebut tidak dapat lagi dikatakan sebagai data pribadi atau termasuk dalam cakupan undang-undang perlindungan data. GDPR mencerminkan pendekatan ini.

Undang-undang nasional juga melestarikan gagasan anonimisasi, misalnya, undang-undang perlindungan data Swiss: Hukum Swiss mewajibkan "data pribadi" ada sebagai awal dari proses anonimisasi. Prosesnya sendiri adalah penghapusan kemungkinan untuk mengidentifikasi seseorang melalui data pribadi. Hal ini dapat dilakukan dengan dua cara: pertama, penghapusan total kemungkinan untuk mengidentifikasi orang di balik data pribadi; kedua, penghapusan kemungkinan sehingga identifikasi orang tersebut menjadi sangat sulit dilakukan. Dengan kata lain, risiko "identifikasi ulang" haruslah cukup rendah. Sebagian besar undang-undang Eropa memperlakukan anonimisasi sebagai "pemrosesan data" data pribadi. Akibat hukumnya adalah, hingga anonimisasi selesai, proses yang digunakan harus mematuhi persyaratan hukum pemrosesan data pribadi.

Persyaratan anonimisasi ini, sebuah fitur yang dimiliki oleh berbagai undang-undang tentang perlindungan data, berpotensi menghambat penggunaan AI secara signifikan. Terdapat dilema bagi mereka yang menggunakan AI dalam lingkungan yang diatur: penggunaan data anonim memerlukan keseimbangan antara risiko: anonimisasi yang terlalu sedikit (yang dalam hal ini data tidak akan benar-benar anonim atau risiko identifikasi ulang akan tinggi); dan terlalu banyak anonimisasi (yang dalam hal ini data yang diolah mungkin menjadi terlalu samar untuk dapat digunakan secara bermakna).

Seberapa efektif anonimisasi di era "Big Data", terutama dalam konteks AI? Kepatuhan terhadap persyaratan hukum untuk menghilangkan atau mengurangi risiko mengidentifikasi orang di balik data pribadi dapat menjadi sangat sulit atau kompleks, terutama dalam kasus di mana teknologi (khususnya AI) beroperasi secara independen dari dorongan manusia dan tidak selalu memenuhi harapan manusia.

Situasi ini dapat menjadi lebih bermasalah dalam kasus di mana dua blok data yang seolah-olah anonim digabungkan. Ketika AI memungkinkan untuk mengidentifikasi orang yang "berada di balik" data melalui kombinasi, tindakan ini dapat melanggar norma perlindungan data bahkan tanpa sepenuhnya dipahami oleh pengguna AI pada umumnya. Dengan kata lain, kombinasi data melalui AI dapat terjadi melalui Pembelajaran Mesin tanpa keterlibatan



manusia secara langsung; hal ini memerlukan pertimbangan yang cermat oleh regulator; tetapi, meskipun dilindungi pada tingkat regulasi, mungkin sulit dideteksi dalam praktik.

2. Anonimisasi: Pembahasan Kasus-Kasus Penerapan

Potensi pertanyaan yang dibahas di atas untuk menjadi masalah hukum yang meluas, dan bukan sekadar isu teoretis, dapat dilihat dalam beberapa konteks dan situasi yang berbeda.

Mobil Swakemudi dan Otonom

Data pribadi sangat penting agar mobil swakemudi dan otonom dapat beroperasi sebagaimana mestinya. Mobil otonom memiliki sensor untuk mengumpulkan data pribadi, misalnya mengenai pergerakan di area tertentu, lokasi, status, dan aktivitas penumpang. Data ini dapat menghubungkan parameter teknologi kendaraan dengan individu yang dapat diidentifikasi. Oleh karena itu, anonimisasi data pribadi sangat penting dalam hal ini.

Perlindungan data diperlukan karena dua alasan utama ketika seseorang mengoperasikan mobil swakemudi dan otonom. Pertama, akumulasi data dalam jumlah besar di dalam kendaraan menimbulkan risiko bagi keselamatan individu. Kendaraan otonom mengumpulkan sejumlah besar data tentang jalan yang digunakan, status kendaraan, status penumpang, lokasi, dan lainnya. Kumpulan data ini membuka wawasan global dan lengkap tentang ranah pribadi individu. Konsentrasi data pribadi yang tinggi dalam satu entitas teknologi ini menjadi sasaran empuk peretasan dan tindakan kriminal terkait pencurian data lainnya. Akibatnya, mobil otonom membutuhkan sistem anonimisasi yang rumit di dalam kendaraan.

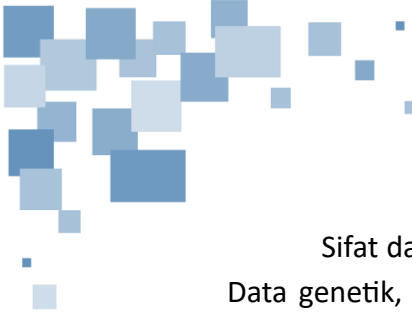
Kedua, kombinasi data dari beberapa entitas AI dapat menjadi pelanggaran privasi. Sebuah kendaraan otonom dapat berkomunikasi dan mentransfer data ke entitas AI lainnya jika sistem cloud terpasang. Kombinasi informasi yang tidak dianonimkan dan yang ditransfer akan sangat memudahkan identifikasi dan analisis individu di balik blok informasi tersebut.

Selain itu, mobil otonom yang beroperasi dalam sistem cloud mungkin memiliki beberapa antarmuka jarak jauh, yang memungkinkan terjadinya serangan terhadap sistem dan data pribadi yang tersimpan; sederhananya, kendaraan otonom tidak dapat mengontrol ke mana data yang ditransmisikannya pergi atau bagaimana data tersebut digunakan. Hal ini mengharuskan data pribadi dianonimkan sebelum ditransmisikan ke entitas AI lainnya.

Mobil self-driving dan mobil otonom membuka pertanyaan lain terkait perlindungan data: Siapa yang memiliki dan berhak mengakses data tersebut? Apakah masih dapat diterima jika produsen mobil memiliki data pribadi saya (seperti lokasi, kecepatan, ritme pribadi, dll.) atau bahkan penyedia telekomunikasi yang menyediakan koneksi internet dengan kartu SIM-nya? Haruskah seseorang dapat meminta agar semua informasi tentang dirinya dihapus dari mobil otonom dalam waktu singkat setelah digunakan? Mengapa tidak?

Data Kesehatan dan, Khususnya, Penyimpanan dan Pengumpulan Materi Genetik/Medis

Beberapa teknologi menggunakan AI untuk analisis dan kategorisasi data genetik. AI menganalisis darah, air liur, dan cairan lain yang diproduksi oleh tubuh manusia, serta potongan kulit dan otot yang diekstraksi. AI memberikan peluang luar biasa untuk meningkatkan kualitas perawatan medis dan penelitian.



Sifat dasar data genetik menghadirkan tantangan besar bagi anonimisasi data pribadi. Data genetik, pada dasarnya, selalu dapat dikaitkan dengan individu tertentu dan (dengan sedikit keraguan) dapat diidentifikasi, karena informasi genetik tidak pernah dibagikan sepenuhnya oleh individu lain.

Selain itu, materi medis dan genetik menyediakan akses ke informasi yang umumnya diklasifikasikan sebagai informasi paling pribadi, kuasi "tempat rahasia" informasi di antara data pribadi. Data genetik, baik secara terpisah maupun ketika digabungkan dengan data lain, memberikan informasi tentang status, kesehatan, gaya hidup, dan banyak aspek lain dari kehidupan pribadi (atau bahkan intim) seseorang.

Oleh karena itu, anonimisasi mungkin bukan pilihan dalam banyak situasi terkait data genetik dan medis. Hal ini karena data genetik, jika dianonimkan dengan benar, akan menjadi terlalu umum untuk bermanfaat sebagaimana mestinya. Hal ini menimbulkan pertanyaan tentang bagaimana mematuhi undang-undang perlindungan data saat menyimpan dan memproses data genetik atau medis, tanpa membuat data yang dikumpulkan terlalu umum untuk bermanfaat. Pseudonimisasi umumnya dipandang sebagai pilihan yang layak.

Anonimisasi mengharuskan bagian-bagian data pribadi yang memungkinkan identifikasi individu di balik data tersebut dihapus sepenuhnya. Sebaliknya, untuk pseudonimisasi, bagian-bagian data pribadi yang memungkinkan identifikasi individu diganti atau dihapus. Misalnya, nama dapat diganti dengan nomor referensi atau pengenal lainnya. Pseudonimisasi memastikan bahwa data tetap bermanfaat, tetapi meminimalkan risiko bagi individu.

Seperti yang diilustrasikan oleh pembahasan di atas, mungkin mustahil untuk memuaskan hasrat AI yang tak terpuaskan terhadap data dengan menganonimkan kumpulan data besar yang saat ini ada. Oleh karena itu, kita perlu mempertimbangkan langkah-langkah lain untuk memastikan bahwa AI mematuhi rezim privasi yang ada, yang akan kita bahas sekarang.

Dasar Hukum dan Persetujuan

Berdasarkan banyak undang-undang privasi, pemrosesan informasi pribadi sah hanya jika alasan organisasi memproses data pribadi tersebut sesuai dengan justifikasi yang diakui dalam undang-undang yang berlaku, dan organisasi mengidentifikasi alasan spesifik untuk pemrosesan informasi pribadi tersebut. Kami menyebut alasan-alasan ini sebagai "dasar hukum" untuk pemrosesan.

Beberapa undang-undang privasi, seperti GDPR, menyebutkan dasar hukum yang dapat diandalkan oleh organisasi, sementara undang-undang privasi lainnya, seperti undang-undang privasi Kanada, menerapkan standar kewajaran untuk persetujuan yang diwajibkan. Apa pun rezim peraturan yang berlaku, operator AI semakin perlu memulai proyek di lingkungan AI hanya setelah melakukan analisis mendalam terhadap data yang terlibat, kebutuhan semua data individual untuk mencapai tujuan, dan dampak negatif apa pun yang dihasilkan dari pemrosesan data yang dimaksud. Dengan melakukan hal tersebut dan menemukan dasar hukum untuk pemrosesan, mereka dapat menghindari pemborosan biaya. Mari kita pertimbangkan masalah yang harus mereka hadapi.

1. Dasar Hukum

Berdasarkan GDPR, operator AI perlu mengidentifikasi dasar hukum untuk pemrosesannya: dasar hukum yang tersedia bagi pengendali data berdasarkan GDPR umumnya bergantung pada kategori informasi pribadi dan tujuan penggunaan data tersebut. Pasal 6 GDPR menyajikan dasar hukum 'default' untuk pemrosesan informasi pribadi, yaitu:

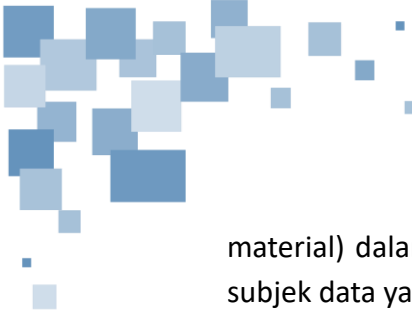
- persetujuan subjek data
- pemrosesan diperlukan untuk pelaksanaan kontrak yang melibatkan subjek data atau untuk mengambil langkah-langkah atas permintaan subjek data sebelum menandatangani kontrak
- pemrosesan diperlukan untuk memenuhi kewajiban hukum pengendali data
- pemrosesan diperlukan untuk melindungi kepentingan vital subjek data atau orang lain
- pemrosesan diperlukan untuk pelaksanaan tugas yang dilakukan untuk kepentingan publik atau dalam pelaksanaan wewenang resmi yang dilimpahkan kepada pengendali
- pemrosesan diperlukan untuk kepentingan sah pengendali data yang tidak dikesampingkan oleh kepentingan atau hak dan kebebasan fundamental subjek data

Persetujuan sering kali dipandang sebagai dasar hukum standar emas untuk pemrosesan informasi pribadi (jika subjek data menyetujui, maka pemrosesan harus sah) dan oleh karena itu cenderung menjadi posisi default bagi banyak pengendali data. Namun, persetujuan sebagai dasar hukum menghadirkan banyak tantangan kepatuhan, terutama bagi operator AI. Di antara tantangan lainnya, validitas persetujuan bergantung pada fakta, sebagaimana dibahas di bawah ini.

Dasar hukum lain dalam Pasal 6 mencakup kriteria kebutuhan. Secara umum, pemrosesan tidak 'diperlukan' untuk tujuan GDPR jika ada cara lain yang efektif dan kurang mengganggu untuk mencapai tujuan yang sama. Persyaratan kebutuhan ini dapat menghadirkan tantangan bagi AI khususnya: AI membutuhkan sejumlah besar data untuk 'belajar' dan, sementara AI mungkin memiliki manfaat individu dan/atau sosial jangka panjang, dalam jangka pendek, pemrosesan informasi pribadi untuk mengajarkan sistem AI bukanlah hal yang diperlukan dan juga bukan merupakan pilihan yang paling tidak mengganggu.

Seperti GDPR, Undang-Undang Privasi Selandia Baru 1993 juga mensyaratkan kebutuhan suatu lembaga dapat mengumpulkan informasi pribadi hanya "untuk tujuan yang sah terkait dengan fungsi atau aktivitas lembaga" dan "pengumpulan informasi diperlukan untuk tujuan itu." Operator AI di Selandia Baru dapat menghindari persyaratan kebutuhan dengan mengandalkan persetujuan; Misalnya, dalam kontrak dengan subjek data untuk aplikasi AI yang memberikan rekomendasi mode bagi penggunanya, operator AI dapat memperoleh persetujuan dari pengguna atas penggunaan data pribadi mereka oleh sistem AI. Namun, hal ini memiliki tantangan tersendiri sebagaimana dibahas di bawah ini.

Mengidentifikasi dasar hukum yang paling tepat dapat menjadi tugas yang sulit. Sekilas, operator AI mungkin berpandangan bahwa apa yang disebut dasar hukum "kepentingan sah" akan berlaku untuk pemrosesan mereka, mengingat, sekilas, cakupannya tampak luas. Namun, dasar ini tidak lugas; operator AI harus mempertimbangkan dan menyeimbangkan kepentingan sah mereka (seperti kepentingan hukum, ekonomi, dan



material) dalam pemrosesan informasi pribadi (melalui sistem AI-nya); dengan kepentingan subjek data yang relevan dan hak-hak fundamental serta kebebasan fundamental.

Semua konsekuensi gangguan harus dipertimbangkan secara komprehensif, meskipun terjadinya konsekuensi tersebut tidak pasti, yaitu hanya risiko yang kemungkinan mengakibatkan "risiko tinggi terhadap hak dan kebebasan subjek data." Saat melakukan uji keseimbangan ini, faktor-faktor yang relevan meliputi sifat dan jumlah data, konsekuensi pemrosesan, dan ekspektasi wajar subjek data berdasarkan hubungan mereka dengan pengendali.

Dampak negatif apa pun yang dihasilkan dari pemrosesan data (misalnya, jika hasil penilaian risiko yang dilakukan oleh sistem AI mencegah subjek data mendapatkan pinjaman) tidak serta merta memicu pelanggaran hukum pemrosesan data; namun, pertimbangan perlu diberikan pada bagaimana risiko tersebut dapat dimitigasi. Hal ini dapat melibatkan pelaksanaan penilaian dampak perlindungan data jika pemrosesan kemungkinan mengakibatkan "risiko tinggi terhadap hak dan kebebasan subjek data."

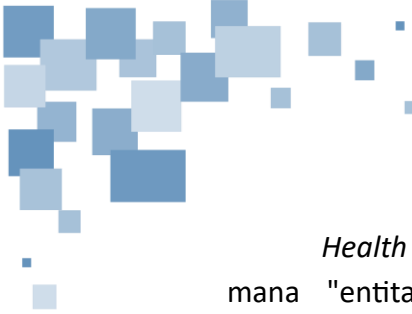
Mengidentifikasi dasar hukum untuk pemrosesan menjadi lebih rumit jika terdapat "kategori khusus" data pribadi yang sedang diproses (misalnya, jika informasi tersebut berisi data kesehatan). Dalam hal ini, landasan hukum lebih lanjut dalam Pasal 9 harus dipenuhi selain landasan hukum Pasal 6 yang disebutkan di atas. Landasan hukum Pasal 9 meliputi:

- ❖ subjek data telah memberikan persetujuan eksplisit untuk pemrosesan informasi pribadi tersebut untuk satu atau lebih tujuan tertentu.
- ❖ pemrosesan diperlukan untuk tujuan melaksanakan kewajiban dan menjalankan hak-hak khusus pengendali atau subjek data di bidang ketenagakerjaan, jaminan sosial, dan hukum perlindungan sosial berdasarkan hukum Uni Eropa.
- ❖ pemrosesan berkaitan dengan informasi pribadi yang secara nyata dipublikasikan oleh subjek data.

Landasan hukum mana yang dapat diandalkan akan bergantung pada sifat data dan tujuan pemrosesan yang tepat.

Sebagaimana dibahas di bawah ini, persetujuan eksplisit merupakan landasan hukum yang ketat dan, sebagaimana dibahas di atas, kebutuhan merupakan standar yang tinggi. Dalam praktiknya, hal ini berarti bahwa operator AI kemungkinan akan dibatasi dalam hal landasan hukum yang tersedia bagi mereka.

Terdapat banyak batasan lain dalam GDPR yang perlu dipatuhi: misalnya, terdapat isu lebih lanjut yang perlu dipertimbangkan jika data pribadi terkait dengan anak atau berisi informasi terkait dengan hukuman pidana atau pelanggaran pidana. Di AS, undang-undang privasi sektor spesifik juga memerlukan dasar hukum. Berdasarkan *Children's Online Privacy Protection Act AS (COPPA)*, pengendali data yang mengoperasikan situs web, aplikasi seluler, atau layanan digital lainnya yang ditujukan kepada anak-anak di bawah usia 13 tahun harus mengunggah kebijakan privasi dan mendapatkan "persetujuan orang tua yang dapat diverifikasi" (VPC) sebelum mengumpulkan informasi pribadi (sebagaimana didefinisikan dalam COPPA) dari anak-anak di bawah usia 13 tahun.



Health Insurance Portability and Accountability Act AS (HIPAA) membatasi keadaan di mana "entitas yang dilindungi" (dan "rekan bisnisnya") dapat menggunakan dan mengungkapkan informasi kesehatan yang dapat diidentifikasi secara individual yang dibuatnya atau diterimanya dari pasien, yang dikenal sebagai "informasi kesehatan yang dilindungi" atau "PHI." Entitas yang dilindungi HIPAA diizinkan untuk menggunakan PHI untuk perawatan, pembayaran, atau operasi perawatan kesehatan dan sebagaimana diizinkan atau diwajibkan oleh hukum, dan dalam beberapa kasus, dengan persetujuan pasien.

Beberapa undang-undang privasi tidak menyebutkan dasar hukum yang spesifik; Sebaliknya, pemrosesan informasi pribadi sah jika diperlukan, sesuai, dan/atau wajar berdasarkan fakta dan keadaan. *Personal Information Protection and Electronic Documents Act* Kanada (PIPEDA) mengizinkan organisasi untuk mengumpulkan, menggunakan, dan mengungkapkan informasi pribadi hanya untuk tujuan yang "dianggap pantas oleh orang yang berakal sehat dalam keadaan tersebut." Berdasarkan Undang-Undang Privasi Federal Australia, suatu organisasi tidak boleh mengumpulkan informasi pribadi kecuali informasi tersebut secara wajar diperlukan untuk atau terkait langsung dengan satu atau lebih fungsi atau aktivitas bisnisnya.

Undang-undang privasi lainnya tidak mewajibkan dasar hukum sebelum memproses informasi pribadi, tetapi mungkin mewajibkan organisasi yang tercakup untuk memberikan pemberitahuan dan pengungkapan. Misalnya, Undang-Undang Privasi Konsumen California secara umum mengizinkan pemrosesan "informasi pribadi" oleh bisnis yang tercakup jika bisnis tersebut memenuhi persyaratan pengungkapan CCPA dan menghormati hak akses, penghapusan, penolakan dan hak layanan yang setara yang diberikan CCPA bagi penduduk California.


Demikian pula, *Gramm-Leach Bliley Act* ("GLBA") mewajibkan lembaga keuangan untuk memberi tahu nasabahnya tentang praktik berbagi informasi dan memberi tahu konsumen tentang hak mereka untuk menolak praktik berbagi tertentu sebelum atau bersamaan dengan pengumpulan dan pemrosesan informasi keuangan pribadi non-publik. Pemberitahuan ini biasanya dilakukan melalui penggunaan formulir pemberitahuan privasi model yang diterbitkan bersama oleh lembaga federal yang mengatur lembaga keuangan.

2. Persetujuan sebagai Dasar Hukum: Tantangan bagi AI

Apa itu Persetujuan?

Sekilas, persetujuan adalah konsep yang lugas. Namun, persetujuan yang sah itu rumit dan menantang. Menurut definisi "persetujuan" dalam Pasal 4 GDPR, persetujuan berdasarkan hukum perlindungan data hanya sah jika merupakan "indikasi yang diberikan secara bebas, spesifik, berdasarkan informasi, dan tidak ambigu tentang keinginan subjek data, yang dengannya ia, melalui pernyataan atau tindakan afirmatif yang jelas, menyatakan persetujuannya terhadap pemrosesan data pribadi yang berkaitan dengannya."

Persetujuan sah jika berdasarkan informasi: "Persetujuan berdasarkan informasi mengharuskan individu diberikan informasi yang cukup untuk membuat keputusan yang tepat..." Demikian pula, cakupan persetujuan harus jelas dan harus mencakup semua aktivitas pemrosesan yang dilakukan untuk tujuan yang sama. Dalam aplikasi AI yang seringkali



mencakup berbagai aktivitas pemrosesan dan transfer data, semua ini harus disajikan sebagai pernyataan terpisah yang dapat disetujui oleh subjek data secara individual.

Saat memperoleh persetujuan, jenis persetujuan (opt-in atau opt-out) bervariasi tergantung pada jenis informasi pribadi, jenis aktivitas pemrosesan, dan yurisdiksi operator AI dan/atau subjek data. Berdasarkan GDPR, pernyataan persetujuan harus eksplisit atau konklusif melalui tindakan proaktif yang hanya dapat dipahami sebagai persetujuan subjek data terhadap pemrosesan datanya.

Mengklik kotak centang, misalnya, akan memenuhi persyaratan persetujuan yang sah dalam banyak kasus, karena setidaknya merupakan pernyataan niat yang aktif. Namun, dalam aplikasi AI yang dicakup oleh Pasal 22 dan memerlukan persetujuan eksplisit, bentuk persetujuan tersebut tidak akan mencukupi. Di sisi lain, kotak yang telah dicentang sebelumnya kemungkinan besar tidak merupakan persetujuan yang sah.

Kantor Komisioner Privasi Kanada merilis pedoman di bawah PIPEDA Kanada dengan tujuh prinsip panduan untuk menentukan persetujuan yang bermakna:

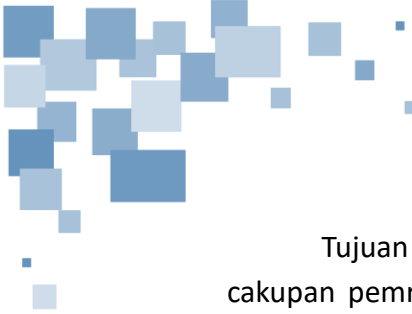
- Menekankan elemen-elemen kunci.
- Memungkinkan individu untuk mengontrol tingkat detail yang mereka dapatkan dan kapan.
- Berikan individu pilihan yang jelas untuk mengatakan 'ya' atau 'tidak'.
- Bersikap inovatif dan kreatif.
- Pertimbangkan perspektif konsumen.
- Jadikan persetujuan sebagai proses yang dinamis dan berkelanjutan.
- Bertanggung jawab: Bersiaplah untuk menunjukkan kepatuhan.

Tantangan dalam Menggunakan Persetujuan Sebagai Dasar Hukum

Sebagaimana akan terlihat dari ikhtisar persyaratan persetujuan, persetujuan memiliki sejumlah tantangan khusus sebagai dasar hukum dalam konteks sistem AI: ini termasuk fakta bahwa berbagai jenis informasi pribadi memerlukan berbagai jenis persetujuan dan bahwa persetujuan dapat dibatalkan.

Secara khusus, terdapat tantangan bahwa subjek data harus memiliki informasi relevan yang cukup tentang penggunaan yang direncanakan dan konsekuensi pemrosesan untuk memastikan bahwa setiap persetujuan yang mereka berikan merupakan pilihan yang terinformasi: membuat persetujuan yang spesifik/eksplisit dalam konteks AI merupakan tantangan karena operator AI tidak dapat mengetahui secara pasti bagaimana informasi pribadi mana yang diproses dan untuk tujuan apa. Hal ini khususnya terjadi pada sistem AI yang belajar mandiri di mana pemrosesan akhir yang dilakukan sistem mungkin berbeda dari yang diinginkan oleh pengembang sistem AI (dan yang telah mendapatkan persetujuan dari pengembang).

Jika terdapat kesulitan dalam mengandalkan landasan hukum lain, persetujuan mungkin merupakan pilihan yang tepat. Namun, ketika mengandalkan persetujuan, operator AI harus berhati-hati untuk memastikan bahwa mereka menjelaskan secara rinci bagaimana data individu akan digunakan dan menggabungkan proses yang memungkinkan individu untuk memberikan indikasi yang jelas bahwa mereka menyetujui penggunaan tersebut.



Tujuan pemrosesan juga perlu ditinjau jika operator AI memutuskan untuk mengubah cakupan pemrosesannya (misalnya, menggunakan data untuk tujuan lain) atau jika sistem belajar mandiri berkinerja berbeda dari yang diantisipasi, maka operator perlu memastikan bahwa individu mengetahui penggunaan baru ini dan, jika perlu, persetujuan mereka diperoleh.

Penting juga untuk diingat bahwa persetujuan dapat ditarik kapan saja, yang akan menempatkan operator AI pada posisi yang sulit jika hal ini terjadi. Oleh karena itu, pertimbangan harus selalu diberikan pada ketersediaan dasar alternatif untuk persetujuan.

3. Kesimpulan

Sebagaimana disebutkan di awal bagian ini, rekomendasi yang bermaksud baik bagi organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI, untuk menghindari pemborosan biaya, adalah memulai proyek di lingkungan AI hanya setelah melakukan analisis mendalam terhadap data yang terlibat, kebutuhan semua data individual untuk mencapai tujuan, dan dampak negatif apa pun yang dihasilkan dari pemrosesan data yang dimaksud. Namun, isu-isu terkait pencarian dasar hukum untuk pemrosesan data dan penghindaran risiko tinggi bagi subjek data, dalam banyak sistem AI, hanya dapat ditangani oleh pengembang.

Tampaknya jelas bahwa kesulitan dalam memperjuangkan dasar hukum yang stabil tidak dapat diselesaikan hanya setelah implementasi oleh pengguna AI. Hal ini mengarah pada istilah populer 'etika berdasarkan desain', yang tampaknya perlu mempertimbangkan masalah privasi sejak awal.

Mengidentifikasi Tujuan

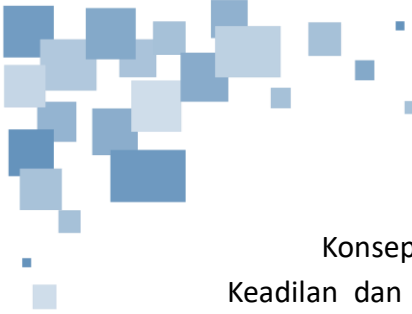
Hal ini bermula dari analisis Dasar Hukum di atas: hukum privasi umumnya mewajibkan perusahaan untuk mengidentifikasi tujuan pengumpulan data pada atau sebelum pengumpulan; tetapi dengan AI, beberapa wawasan baru disadari "setelah kejadian" dan banyak perusahaan menyadari bahwa mereka sekarang dapat memonetisasi "data hukum" dengan cara yang belum pernah mereka pertimbangkan sebelumnya.

Bagaimana prinsip "mengidentifikasi tujuan" dapat dihormati dalam konteks seperti itu? Perusahaan perlu mengidentifikasi tujuan pengumpulan data dan akan digunakan untuk menginformasikan subjek data dan menetapkan dasar hukum mereka untuk pemrosesan. Jika tujuan pemrosesan berubah, hal ini harus dikomunikasikan kepada individu. Oleh karena itu, sangat penting bagi perusahaan untuk mengidentifikasi dengan jelas tujuan pemrosesan apa pun di awal proyek dan juga melakukan tinjauan berkala untuk memastikan keakuratannya.

Keadilan dan Isu-Isu Operasional Terkait

1. Prinsip Keadilan

Kami mengakui bahwa "keadilan" adalah konsep yang jauh melampaui hukum privasi. Bahkan, kami telah memberikan konsep keadilan bab tersendiri dalam buku ini. Konsep ini mencakup tinjauan dampak pemrosesan data terhadap individu untuk memastikan data tidak digunakan dengan cara yang dapat menimbulkan dampak buruk yang tidak dapat dibenarkan bagi individu tersebut.



Konsep ini telah dibahas secara rinci dalam pembahasan kami tentang Prinsip 4 Keadilan dan Non-Diskriminasi. Dalam hukum privasi, konsep ini juga mengartikulasikan ekspektasi mendasar kami seputar pemrosesan data pribadi: Untuk itu, lihat juga pembahasan kami tentang Prinsip 3 Transparansi dan Keterjelasan, yang menekankan pentingnya 'keterjelasan' bagi keadilan AI.

Alih-alih membahas hal ini lagi, di bagian ini kami berfokus pada risiko terhadap keadilan dan AI dari perspektif regulator privasi; kami menggunakan frasa 'Privasi sebagai Keadilan' untuk mencerminkan bagian dari regulasi privasi saat ini dan di masa mendatang yang bertujuan untuk mencerminkan prinsip-prinsip keadilan. Ini diikuti dengan analisis teknologi peningkatan privasi yang dapat membantu memenuhi kepatuhan terhadap prinsip ini.

2. Tantangan dan Risiko Sistem AI terhadap Keadilan

Tantangan dan risiko yang dihadirkan sistem AI terhadap 'privasi-sebagai-keadilan' didasarkan pada fakta sederhana bahwa melatih sistem AI untuk menerapkan privasi-sebagai-keadilan itu sulit: hukum privasi juga tidak absolut, melainkan melibatkan keseimbangan antara hak individu dan kepentingan masyarakat.

Sebagaimana disebutkan dalam pembahasan kami tentang Prinsip 2 Akuntabilitas, perlu tetap ada tingkat akuntabilitas manusia; jadi, 'melatih' sistem AI untuk melakukan keseimbangan antara kepentingan individu dan kolektif mungkin melewati 'garis merah' tetapi bagaimanapun juga akan menjadi tugas yang sangat sulit, terutama karena mungkin tidak ada jawaban yang hitam dan putih.

Data Pelatihan

Data yang diberikan kepada sistem AI mungkin tidak diatur, tidak difilter, tidak bersih, tidak representatif, dan terkadang, bahkan sengaja dimiringkan, sehingga sistem AI apa pun yang dilatih menggunakan data tersebut akan mencerminkan bias dan ketidakakuratan yang melekat dalam data pelatihan. Pelatihan dibahas lebih rinci dalam Bab Keadilan. Kami menambahkan satu contoh di sini untuk mengilustrasikan tantangan terhadap keadilan yang dihadirkan oleh data pelatihan:

Sebuah perusahaan global berupaya mengembangkan sebuah program, sebagai bagian dari proses identifikasi talenta terbaik, yang akan memberikan skor kepada pelamar, mulai dari satu hingga lima bintang, berdasarkan resume atau lamaran mereka. Namun, pada tahun 2015, perusahaan tersebut menyadari bahwa sistem tersebut tidak menilai kandidat untuk posisi teknis secara netral gender. Sistem tersebut dilatih berdasarkan resume yang telah diserahkan selama periode sepuluh tahun, yang sebagian besar, mencerminkan dominasi laki-laki di sektor teknologi, adalah laki-laki. Pada dasarnya, sistem AI melatih dirinya sendiri untuk lebih menyukai kandidat laki-laki, dengan memberikan skor yang lebih rendah pada aplikasi atau resume yang mengandung kata "perempuan" atau menunjukkan bahwa kandidat tersebut pernah kuliah di perguruan tinggi khusus perempuan.

Kumpulan data pelatihan harus dipantau untuk menentukan apakah data tersebut bias. Hal ini dapat menjadi bagian dari kepatuhan terhadap standar 'privasi-sebagai-keadilan'.

Area Terkait yang Perlu Diperhatikan

Meskipun pelatihan merupakan perhatian utama di sini, terdapat area terkait yang dipertimbangkan dalam pembahasan Prinsip 4 Keadilan dan Non-Diskriminasi yang relevan dengan isu bagaimana regulasi privasi dapat mendukung keadilan yang lebih besar: (i) Sistem AI sebagai "Kotak Hitam" (yaitu, bagaimana keputusan diambil tidak transparan); dan (ii) Tidak Adanya Tinjauan atau Pengujian Independen (yaitu, hasil sistem AI seringkali tidak dipantau atau tidak dinilai secara ketat atau independen untuk mendeteksi bias). Selain itu, dan sebagaimana dibahas di bawah ini, konsep pengumpulan terbatas/minimalisasi data dalam undang-undang Privasi Data bertentangan dengan kebiasaan dasar sistem AI untuk mengonsumsi data dalam jumlah besar yang sekali lagi menimbulkan risiko terhadap keadilan.

Jika kita mengkaji isu-isu ini pada tingkat paling dasar, isu-isu ini menggambarkan tema umum: kurangnya kendali manusia dan masyarakat atas serta pemantauan pengembangan, pengujian, implementasi, dan penggunaan sistem AI. Lebih lanjut, risiko yang timbul dari kurangnya kendali ini berdampak langsung pada konsep keadilan kita:

- (i) informasi pribadi dapat digunakan dengan cara yang tidak diantisipasi oleh individu dan di luar ekspektasi wajar individu;
- (ii) hasil yang dihasilkan oleh sistem AI dapat bias atau diskriminatif, sehingga menimbulkan dampak buruk yang tidak dapat dibenarkan bagi individu; dan
- (iii) dampak risiko-risiko ini diperkirakan akan meningkat, seiring kemampuan sistem AI yang belajar mandiri untuk beroperasi dengan cara yang melampaui apa yang dibayangkan oleh para pengembang, membuka pintu bagi sistem yang, setelah beroperasi, akan membuat penentuan privasi mereka sendiri, tentang apakah akan mengidentifikasi ulang data anonim, tentang bagaimana informasi pribadi dapat digunakan, atau tentang pengungkapan informasi pribadi apa yang dapat diterima.

3. Pendekatan untuk Meningkatkan Keadilan dalam Sistem AI

Terlepas dari tantangan signifikan terhadap privasi yang dihadirkan oleh AI, prinsip keadilan tidak akan dikorbankan demi kemanfaatan AI. Terdapat praktik manajemen, teknologi peningkatan privasi, serta alat dan teknik deteksi bias yang, jika dipelihara dan dikembangkan dengan tepat, dapat menyediakan sarana untuk menyegarkan kembali perlindungan privasi dan prinsip keadilan dalam dunia AI. Pada bagian ini, kita akan membahas masing-masing secara berurutan.

Manajemen Sistem AI

Pentingnya pendekatan manajemen dalam menangani ancaman privasi yang dijelaskan di bagian ini adalah bahwa pendekatan tersebut tidak mewakili proses baru yang dirancang khusus untuk sistem AI, melainkan merupakan penerapan pendekatan manajemen yang ada terhadap tantangan AI saat ini. Dalam menghadapi risiko terhadap privasi dan keadilan yang ditimbulkan oleh hilangnya kendali atas pengembangan, pengujian, penerapan, dan penggunaan AI, pendekatan tersebut merupakan respons manajemen yang bijaksana untuk menegaskan kembali kendali.

Privasi Berdasarkan Desain

Metodologi Privasi Berdasarkan Desain menyediakan templat tentang cara yang tepat untuk mendapatkan kembali kendali atas sistem AI. Fokus metodologi Privasi Berdasarkan Desain adalah pada langkah-langkah pencegahan untuk memitigasi risiko privasi sebelum terjadi. Privasi dimasukkan sebagai komponen penting dan inti dari sistem yang dirancang, sementara fungsionalitas penuh dan perlindungan keamanan dipertahankan di sepanjang siklus hidup informasi. Kerangka kerja Privasi Berdasarkan Desain tercermin dalam GDPR, di mana Pasal 25 GDPR mewajibkan pengendali data untuk mengintegrasikan perlindungan privasi ke dalam sistem dan memastikan bahwa perlindungan data dijaga dalam pengaturan standar sistem, dengan mematuhi kerangka kerja Privasi Berdasarkan Desain.

Kepatuhan terhadap prinsip ini mengharuskan perancang sistem AI untuk mengintegrasikan kontrol privasi ke dalam sistem data dan teknologi informasi sejak awal. Konferensi Internasional Komisioner Perlindungan Data dan Privasi 2018 memang mengadopsi *Declaration on Ethics and Data Protection in Artificial Intelligence* yang mengesahkan prinsip panduan berikut sebagai salah satu pendekatannya untuk mengatasi tantangan terkait AI:

4. Sebagai bagian dari pendekatan "etika berdasarkan desain" secara keseluruhan, sistem kecerdasan buatan harus dirancang dan dikembangkan secara bertanggung jawab, dengan menerapkan prinsip privasi berdasarkan bawaan dan privasi berdasarkan desain...

Dewan Etika

Dewan etika semakin banyak diadopsi oleh organisasi sektor publik dan swasta untuk menangani berbagai isu kontroversial atau sensitif dan dapat digunakan untuk mendukung prinsip keadilan.

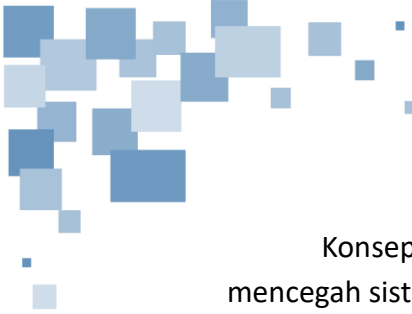
Penilaian Dampak Perlindungan Data (DPIA)

Penilaian dampak perlindungan data mencerminkan prinsip bahwa pihak yang memproses data pribadi memiliki kewajiban untuk menilai risiko yang terlibat dan merupakan perluasan dari rekomendasi untuk menerapkan mekanisme penilaian dampak yang lebih umum, sebagaimana dibahas dalam Bab Akuntabilitas. Otoritas Perlindungan Data Norwegia telah memperluas gagasan ini secara khusus ke sistem AI. Dalam makalah mereka, *Artificial Intelligence and Privacy*, Otoritas menyarankan agar kriteria khusus diterapkan sehubungan dengan penilaian dampak sistem AI.

Teknologi AI yang Meningkatkan Privasi

Berbeda dengan ancaman yang mungkin ditimbulkan oleh sistem AI terhadap privasi individu, terdapat fokus yang semakin meningkat pada penggunaan AI dan metodologinya untuk memitigasi dampak negatif AI atau untuk mengatasi masalah privasinya. Faktanya, para ahli teknologi telah memprediksi penerapan AI ini selama beberapa waktu, menunjukkan bahwa penggunaan AI dengan cara ini mungkin merupakan satu-satunya cara realistis untuk memantau model AI yang sangat kompleks dan terus berubah secara memadai dan efektif.

Dalam satu contoh teknologi peningkatan privasi tersebut, para peneliti sedang menyelidiki penggunaan agen berbasis AI, atau auditor AI, untuk melindungi privasi. Proposalnya adalah untuk mengembangkan AI yang dirancang untuk melindungi dari AI lain dalam pengelolaan data pribadi. Sistem semacam itu dapat digunakan untuk mendeteksi ancaman siber, bertindak sebagai sistem peringatan cerdas untuk jaringan dan basis data.



Konsep auditor AI ini juga dapat diperluas sehingga auditor AI akan bertindak untuk mencegah sistem AI lain mengidentifikasi ulang data yang sebelumnya dianonimkan dan akan digunakan untuk mengidentifikasi hasil algoritmik yang tidak adil dan diskriminatif. Auditor AI ini mungkin di masa mendatang akan digunakan oleh pemerintah atau lembaga regulator untuk menganalisis sistem AI lain untuk memastikan kewajaran, akurasi, dan kepatuhan.

Deteksi Bias

Contoh-contoh yang dikutip di atas menunjukkan bahwa sistem AI pada dasarnya tidak netral, tetapi justru rentan terhadap bias dan masalah diskriminasi. Tantangannya, dari perspektif kewajaran, adalah mendeteksi, mengukur, dan memitigasi dampak diskriminatif ini. Seiring dengan meningkatnya penelitian tentang teknologi yang meningkatkan privasi, fokus pada bias dalam algoritma dan pengembangan algoritma serta teknik statistik baru untuk mendeteksi dan memitigasi bias ini juga meningkat, misalnya strategi kewajaran observasional, strategi anti-klasifikasi, dan strategi kalibrasi.

Pembuatan Profil dan Pengambilan Keputusan Otomatis

1. Pendahuluan

'Pembuatan Profil' dan 'Pengambilan Keputusan Otomatis' adalah aktivitas yang umumnya bergantung pada AI dan melibatkan data pribadi oleh karena itu, terdapat kaitan erat dengan regulasi privasi data: di bagian ini, kami berfokus pada bagaimana GDPR64 saat ini menangani kedua isu praktis ini; karena GDPR64 menggambarkan isu-isu tersebut dengan baik dan merupakan model yang dapat membantu mengelola risiko terkait AI di area ini. Namun, bab ini sebaiknya dibaca bersama dengan pembahasan kami tentang Prinsip 3 Transparansi dan Keterjelasan, untuk mendeteksi isu-isu tersebut secara komprehensif.


Pasal 4(4) GDPR mendefinisikan pembuatan profil sebagai

segala bentuk pemrosesan data pribadi secara otomatis yang terdiri dari penggunaan data pribadi yang berkaitan dengan seseorang untuk mengevaluasi aspek-aspek pribadi tertentu yang berkaitan dengan seseorang, khususnya untuk menganalisis atau memprediksi aspek-aspek kinerja orang tersebut di tempat kerja, situasi ekonomi, kesehatan, preferensi pribadi, minat, keandalan, perilaku, lokasi, atau pergerakannya. (Pasal 4(4))

Keputusan otomatis didefinisikan dan diatur oleh Pasal 22 GDPR: "keputusan otomatis individual" adalah keputusan yang berkaitan dengan individu yang didasarkan pada pemrosesan mesin. Contohnya adalah pengenaan denda berdasarkan gambar yang direkam oleh kamera kecepatan otomatis.

Dalam praktiknya, penting untuk membedakan antara: (i) pembuatan profil yang merupakan pemrosesan otomatis, yang harus mematuhi prinsip-prinsip perlindungan data dalam Pasal 5 dan 6; dan (ii) pembuatan profil yang merupakan bagian dari pengambilan keputusan otomatis sebagaimana dimaksud dalam Pasal 22, yang tunduk pada aturan tambahan.

Dari definisi istilah 'pembuatan profil' dan 'keputusan otomatis' individual ini, jelas bahwa, pada prinsipnya, mereka yang mengembangkan, menerapkan, atau menggunakan sistem ini tunduk pada aturan yang berlaku untuk pembuatan profil dalam GDPR (dengan asumsi, tentu saja, bahwa data pribadi, sebagaimana didefinisikan dalam Pasal 4(1),



digunakan). Namun, sejauh mana dan dengan konsekuensi apa GDPR berlaku akan bergantung pada cakupan pembuatan profil yang terlibat.

Ada tiga kemungkinan cara pembuatan profil:

- (i) pembuatan profil umum;
- (ii) pengambilan keputusan berdasarkan pembuatan profil; dan
- (iii) pengambilan keputusan otomatis semata, termasuk pembuatan profil, yang menghasilkan akibat hukum atau secara signifikan memengaruhi subjek data.

Batasan Tujuan Masalah Tujuan Ganda

Salah satu masalah yang sering muncul dalam aplikasi AI (terutama analisis Big Data dan berpotensi dalam pembuatan profil maupun keputusan otomatis) adalah pembuatan profil dilakukan untuk lebih dari satu tujuan. Seringkali, kumpulan data yang dikumpulkan digunakan untuk menganalisis pertanyaan spesifik yang muncul selama pengembangan algoritma.

Akibatnya, tujuan awal tidak didefinisikan dan ditetapkan secara jelas ketika data dikumpulkan untuk pertama kalinya, atau data tersebut digunakan kembali untuk tujuan lain. Penggunaan kembali data tersebut untuk tujuan selain tujuan awal dapat diizinkan, jika subjek data telah menyetujui pemrosesan atau jika pemrosesan lebih lanjut dianggap sesuai dengan tujuan awal.

Menurut pertimbangan GDPR, pemrosesan lebih lanjut untuk tujuan pengarsipan demi kepentingan publik, tujuan penelitian ilmiah atau historis, atau tujuan statistik harus dianggap sebagai operasi pemrosesan yang sah dan sesuai. Namun, masih menjadi pertanyaan mengenai aktivitas pemrosesan mana yang dapat digolongkan dalam istilah 'tujuan ilmiah dan statistik'.

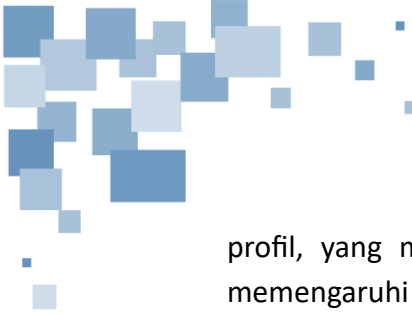
GDPR menginterpretasikan pemrosesan data pribadi untuk tujuan penelitian ilmiah secara luas, termasuk misalnya pengembangan dan demonstrasi teknologi, penelitian fundamental, penelitian terapan, dan penelitian yang didanai swasta. Berdasarkan definisi ini, pengembangan AI dapat merupakan penelitian ilmiah, meskipun dalam banyak kasus 'pengembangan' AI tidak dapat dibedakan secara jelas dari penerapan AI semata. Ketika sebuah model terus-menerus dikembangkan, sulit untuk menentukan di mana penelitian berhenti dan di mana penggunaan dimulai. Proyek harus dievaluasi kasus per kasus untuk mengidentifikasi tujuan awal, dan tujuan sekunder apa pun yang dianggap kompatibel.

Minimalisasi dan Pembatasan Penyimpanan Data

Area lain yang relevan dengan pembuatan profil dan keputusan otomatis adalah prinsip minimisasi dan pembatasan penyimpanan data. Data harus disimpan dalam bentuk yang memungkinkan identifikasi subjek data tidak lebih lama dan tidak lebih luas dari yang diperlukan untuk tujuan pemrosesan data pribadi. Aplikasi AI cenderung melibatkan pengumpulan dan analisis data sebanyak mungkin. Hal ini harus ditinjau dan data tidak boleh disimpan tanpa batas waktu.

2. Cakupan larangan keputusan individual berdasarkan AI contoh GDPR

Pasal 22 GDPR menegaskan bahwa subjek data berhak untuk tidak menjadi subjek keputusan yang semata-mata didasarkan pada pemrosesan otomatis, termasuk pembuatan



profil, yang menghasilkan suatu hasil (keputusan atau tindakan) yang secara signifikan memengaruhi seseorang. Kriteria lebih lanjut dari Pasal 22 adalah bahwa evaluasi yang mengarah pada keputusan otomatis harus memiliki kompleksitas tertentu. Tidak semua keputusan 'jika-maka' dapat dicakup oleh larangan pengambilan keputusan individu.

Satu-satunya pengecualian dari larangan ini adalah: (i) jika pembuatan profil diperlukan untuk membuat, atau melaksanakan, kontrak antara subjek data dan pengendali data; (ii) jika diizinkan oleh hukum setempat yang menetapkan langkah-langkah yang sesuai untuk melindungi hak-hak subjek data; atau (iii) jika didasarkan pada persetujuan eksplisit dari subjek data. Oleh karena itu, pengecualian diterapkan tergantung pada dasar hukum mana yang diandalkan, oleh karena itu penting untuk mengidentifikasi dasar hukum tersebut sejak awal.

Tidak Ada Intervensi Manusia?

Agar Pasal 22 berlaku, keputusan otomatis harus didasarkan semata-mata pada pemrosesan otomatis, yang berarti tidak boleh ada bentuk intervensi manusia apa pun dalam proses pengambilan keputusan. “Intervensi manusia” berarti bahwa seseorang telah melakukan penilaian independen terhadap data pribadi yang mendasarinya, dan berwenang untuk memeriksa kembali rekomendasi yang telah dihasilkan oleh model tersebut.

Oleh karena itu, dalam beberapa aplikasi kecerdasan tertambah, Pasal 22 tidak akan berlaku, mengingat Pasal 22 hanya membantu pengambilan keputusan dengan, misalnya, menganalisis kumpulan data yang besar. Lebih lanjut, dapat dikatakan bahwa Pembelajaran Mesin (sebagai ekspresi spesifik AI) juga dapat mencakup tindakan manusia; karena selama fase pelatihan, interaksi antara mesin dan insinyur merupakan komponen penting dari peningkatan sistem yang inheren.

3. Hak Tambahan dalam Proses Pengambilan Keputusan Otomatis

Perlindungan

Pasal 22(3) menetapkan bahwa pengendali wajib melindungi hak-hak subjek data setidaknya hak untuk mendapatkan intervensi manusia dari pihak pengendali untuk menyatakan sudut pandanganya dan untuk menentang keputusan tersebut (prinsip yang konsisten dengan prinsip akuntabilitas manusia yang diuraikan dalam Bab Akuntabilitas). Menentang suatu keputusan dapat bertujuan untuk membalikkan atau membatalkan keputusan tersebut dan kembali ke status di mana belum ada keputusan yang dibuat, atau untuk mengubah hasil dan menerima keputusan alternatif.

Jika Pasal 22(3) dibaca sebagai satu kesatuan, subjek data harus menggunakan hak-haknya berdasarkan paragraf ini secara bersamaan; dan dengan demikian, intervensi manusia menjadi diperlukan untuk membentuk keputusan baru. Namun, Pasal 22(3) juga dapat ditafsirkan sedemikian rupa sehingga hak-hak tersebut dapat digunakan secara terpisah (misalnya, data masukan harus dipantau, dengan keputusan baru yang sepenuhnya dibuat oleh algoritma). Penafsiran ketentuan ini oleh pengadilan dan otoritas perlindungan data masih harus dilihat. Namun demikian, bagaimanapun juga, selama pengembangan proses AI, langkah-langkah harus diterapkan yang memungkinkan keputusan untuk digugat dan/atau diintervensi manusia.

Hak Atas Penjelasan

Lebih lanjut, ketentuan ini dapat mengarah pada kesimpulan bahwa terdapat hak terpisah di luar hak subjek data dalam Pasal 12, 13, dan 15 yang mengharuskan penjelasan tentang bagaimana keputusan otomatis dibuat, agar subjek data dapat menggugat keputusan tersebut. Hak penjelasan tersebut, jika ada, hanya dapat berlaku untuk pengambilan keputusan otomatis sebagaimana dimaksud dalam Pasal 22.

Pertimbangan 71 GDPR mengatakan bahwa sehubungan dengan pemrosesan data untuk pengambilan keputusan otomatis, hal itu harus tunduk pada perlindungan yang sesuai, yang harus mencakup penyediaan informasi spesifik kepada subjek data dan hak untuk memperoleh intervensi manusia, untuk mengekspresikan sudut pandangnya, untuk memperoleh penjelasan tentang keputusan yang dicapai setelah penilaian tersebut dan untuk menantang keputusan tersebut. Namun, GDPR mengharuskan pengontrol untuk memberikan informasi yang berarti tentang logika yang terlibat, tidak harus penjelasan yang rumit tentang algoritma yang digunakan atau pengungkapan algoritma lengkap. Ini juga sesuai dengan kata-kata Pertimbangan 71, yang menggunakan istilah yang tidak memiliki ruang lingkup yang jelas dan dapat menjadi subjek interpretasi; khususnya tujuan hak atau informasi apa yang harus diungkapkan, tidak ditentukan; dan selanjutnya, logika internal algoritma tidak perlu dijelaskan.

Informasi yang diberikan harus cukup komprehensif agar subjek data dapat memahami alasan keputusan tersebut. Subjek data harus diinformasikan tentang kriteria penilaian yang menjadi dasar keputusan dan aspek serta temuan mana yang menentukan dalam kasus tertentu, sehingga subjek data dapat mempertanyakan keputusan tersebut, mengajukan keberatan, dan memicu peninjauan yang melibatkan manusia.

Hak Subjek Data (misalnya, Akses dan Penghapusan Subjek)


1. Hak atas Informasi dan Hak untuk Mengakses

Perundang-undangan Perlindungan Data didasarkan pada pengakuan hak-hak individu: "Hak Subjek Data" ini dapat menjadi tantangan tersendiri dalam pengembangan dan penggunaan kecerdasan buatan, karena prinsip transparansi.

Sebagaimana disebutkan di atas (dalam Keputusan Otomatis) berdasarkan GDPR, pengendali harus menginformasikan subjek data tidak hanya mengenai keberadaan pengambilan keputusan dan pembuatan profil otomatis, tetapi juga memberikan informasi yang bermakna tentang logika yang terlibat, serta signifikansi dan konsekuensi yang diperkirakan dari pemrosesan tersebut bagi subjek data. Dalam konteks AI, hal ini harus mencakup penyediaan deskripsi logika algoritma (yang dapat mencakup daftar sumber data atau variabel), tetapi bukan pengungkapan pengambilan keputusan dan algoritma itu sendiri.

2. Hak untuk Penghapusan

Berdasarkan GDPR (Pasal 17), subjek data berhak untuk memperoleh penghapusan data pribadi yang berkaitan dengan dirinya dari pengendali tanpa penundaan yang tidak semestinya. Namun, dalam beberapa aplikasi AI, secara teknis mustahil untuk menghapus data dan oleh karena itu, perlu untuk memahami istilah penghapusan dalam pengertian yang netral terhadap teknologi: Penghapusan referensi apa pun terhadap individu dapat menjadi



alternatif untuk penghapusan. Perlu juga dicatat bahwa hak ini hanya dipicu dalam keadaan tertentu, seperti ketika persetujuan diandalkan sebagai dasar hukum, oleh karena itu perlu untuk mengidentifikasi dasar relevan yang diandalkan sejak awal.

Tidak ada definisi istilah "Penghapusan data terkait pribadi," baik dalam bab-bab yang mengikat dari GDPR maupun dalam pertimbangan yang tidak mengikat. Menurut Pasal 4(2), penghapusan dan penghapusan digunakan sebagai dua istilah yang berbeda, yang tidak boleh dipahami secara identik, mengingat bahwa istilah-istilah tersebut disebutkan sebagai bentuk pemrosesan data yang berbeda. Ini berarti bahwa penghapusan tidak selalu menyiratkan penghancuran akhir. Yang relevan untuk penghapusan hanyalah hasilnya, yaitu ketidakmungkinan faktual untuk memahami informasi yang tercakup dalam data yang akan dihapus.

Dengan latar belakang ini, pengendali data memiliki ruang lingkup diskresi tertentu terkait jenis dan cara penghapusan. Penghapusan referensi apa pun kepada seseorang dapat menjadi alat yang tepat untuk menghapus data sesuai dengan Pasal 4(2) sehubungan dengan Pasal 17(1). Namun, harus dipastikan bahwa baik pengendali data itu sendiri maupun pihak ketiga tidak dapat memulihkan referensi pribadi tanpa upaya yang tidak proporsional. Selain itu, CNIL telah menerima bahwa dalam aplikasi tertentu (misalnya, blockchain), pengendali data dapat membuat data praktis tidak dapat diakses, dan oleh karena itu semakin mendekati efek penghapusan data.

Keamanan

Konsep keamanan melekat dalam privasi. Jika informasi pribadi tidak dilindungi oleh langkah-langkah keamanan yang wajar, informasi tersebut pada dasarnya dapat diakses oleh siapa pun dan seluruh gagasan privasi menjadi ilusi, baik dalam konteks AI maupun di tempat lain. Namun, masalah keamanan dalam sistem AI melampaui masalah privasi, dan menyentuh aspek AI yang lebih mendasar.

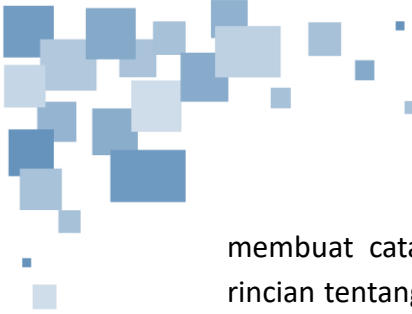
8.3 MENGGUNAKAN AI UNTUK MEMENUHI PERSYARATAN PRIVASI

Setelah meninjau berbagai isu dan tantangan yang ditimbulkan oleh penggunaan AI di bidang privasi, penting untuk mempertimbangkan bagaimana teknologi AI dapat mendukung kepatuhan terhadap peraturan privasi.

Mengelola Data

Organisasi besar yang kaya data semakin memiliki kumpulan data yang besar, yang sepenuhnya tidak terstruktur dan tanpa alokasi untuk tugas-tugas tertentu yang membutuhkan data tersebut. Oleh karena itu, salah satu langkah pertama untuk mematuhi peraturan privasi adalah menganalisis data mana yang diproses di dalam perusahaan: untuk melakukan ini, karyawan yang bertanggung jawab atas berbagai departemen ditanyai kategori data pribadi mana yang diproses di departemen masing-masing di perusahaan. Baru pada langkah kedua dilakukan penilaian tentang tujuan penggunaan dan/atau kebutuhan data serta aktivitas pemrosesan yang relevan.

Salah satu tugas utama terkait akuntabilitas berdasarkan Pasal 24 GDPR adalah menyimpan catatan aktivitas pemrosesan (Pasal 30). Oleh karena itu, pengawas wajib



membuat catatan aktivitas pemrosesan dalam bentuk tertulis atau elektronik, termasuk rincian tentang tujuan pemrosesan, deskripsi kategori subjek data dan kategori data pribadi, serta kategori penerima, dan terakhir pemeriksaan kebolehan pemrosesan data (dasar hukum). Dokumen ini memungkinkan pengawas untuk menunjukkan kepatuhan terhadap GDPR.

Dengan perangkat AI, pemindaian, pengambilan, dan pengkategorian data dapat dilakukan secara otomatis. Misalnya, MinerEye (<https://minereye.com>) mengembangkan perangkat lunak AI berbasis visi komputer dan pembelajaran mesin untuk merayapi, mengidentifikasi, dan mengklasifikasikan semua data perusahaan.

Penilaian Risiko

GDPR mengikuti pendekatan berbasis risiko. Oleh karena itu, hanya jika suatu jenis pemrosesan data kemungkinan besar mengakibatkan risiko tinggi terhadap hak dan kebebasan orang perseorangan, pengawas, sebelum melakukan pemrosesan, diwajibkan untuk melakukan penilaian dampak operasi pemrosesan yang direncanakan terhadap perlindungan data pribadi. Penilaian dampak perlindungan data (DPIA) adalah proses yang dirancang untuk mendeskripsikan pemrosesan, menilai kebutuhan dan proporsionalitasnya, serta membantu mengelola risiko terhadap hak dan kebebasan individu yang diakibatkan oleh pemrosesan data pribadi, dengan menilai dan menentukan langkah-langkah untuk mengatasinya. Kewajiban pengendali untuk melakukan DPIA dalam keadaan tertentu harus dipahami dengan latar belakang kewajiban umum mereka untuk mengelola risiko yang ditimbulkan oleh pemrosesan data pribadi secara tepat.


Oleh karena itu, kemungkinan besar akan segera ada permintaan besar akan keahlian dalam risiko pemrosesan data. Penilaian risiko merupakan tugas yang seringkali menggunakan solusi AI. Penerapan perangkat lunak semacam itu menimbulkan pertanyaan tentang tanggung jawab jika terjadi kesalahan klasifikasi risiko.

Bahkan dengan klasifikasi risiko yang disepakati secara luas, perlu dipahami bahwa setiap organisasi memiliki fleksibilitas dan kewajiban untuk mempertimbangkan elemen risiko tambahan apa pun yang spesifik terhadap konteks dan pemrosesannya sendiri dan apa yang akan memicu perlunya penyesuaian perangkat lunak tersebut dengan keadaan spesifik perusahaan.

Memastikan Hak-hak Subjek Data

AI juga dapat membantu pengendali dalam memenuhi hak-hak subjek data. Pertimbangan 59 GDPR menyatakan bahwa modalitas harus disediakan untuk memfasilitasi pelaksanaan hak-hak subjek data; termasuk mekanisme untuk meminta dan (jika berlaku) memperoleh akses gratis dan perbaikan atau penghapusan data pribadi serta pelaksanaan hak untuk menolak.

Pengendali juga harus menyediakan sarana agar permintaan dapat diajukan secara elektronik, terutama jika data pribadi diproses secara elektronik. Hal ini khususnya berarti dalam praktiknya bahwa permintaan juga dapat dijawab melalui surel, jika permintaan itu sendiri juga dikirim secara elektronik.



GDPR mensyaratkan cara yang lebih canggih untuk menjawab subjek data terkait hak akses. Jika memungkinkan, pengendali harus dapat menyediakan akses jarak jauh ke sistem aman yang akan memberikan subjek data akses langsung ke data pribadinya. Tanpa menciptakan kewajiban hukum, GDPR memungkinkan penerapan antarmuka web yang memfasilitasi subjek data untuk menjalankan hak aksesnya.

Selain itu, Pasal 21(5) secara eksplisit menyatakan bahwa subjek data dapat menjalankan hak keberatannya melalui cara otomatis menggunakan spesifikasi teknis. Legislator ingin memfasilitasi pelaksanaan hak keberatan melalui penerapan solusi teknis cerdas. Dalam konteks ini, "Teknologi Jangan Lacak" yang memungkinkan subjek data untuk secara otomatis menyatakan keberatan melalui pengaturan peramban terhadap pelacakan di situs web akan menjadi relevan.

Kerangka Kerja Saat Ini: Pendekatan Internasional

Sebagaimana dicatat oleh banyak penelitian pemerintah dan sektoral, sebagian besar ekonomi global telah lama menyadari pergeseran paradigma yang ditimbulkan oleh AI dan privasi data. Oleh karena itu, sebagian besar pemerintah di negara-negara maju semakin mengadopsi pendekatan untuk mengatur AI dan data terkait yang mencerminkan sistem politik, ekonomi, budaya, dan sosial mereka sendiri.

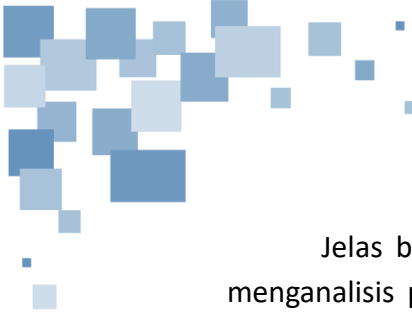
Perusahaan yang menggunakan pembelajaran mesin diwajibkan untuk menerapkan peraturan, seperti *General Data Protection Regulation (GDPR)* dan peraturan privasi data lainnya, pada algoritma mereka. Seiring dengan para profesional teknologi dan privasi yang terus berfokus pada transparansi, kepercayaan, dan ketertelusuran, perusahaan dan konsumen harus memiliki prinsip-prinsip yang jelas dan terfokus yang memandu penggunaan data pribadi.

Prediksi ke Depan dan Solusi yang Diusulkan

Seiring AI terus menjadi arus utama, arahan dan panduan privasi global mungkin terpaksa melihat kerangka kerja mereka saat ini di bawah lensa privasi yang ketat. Saat ini, AI hanya memiliki sedikit undang-undang yang secara khusus menargetkan regulasinya. Sebaliknya, AI perlu menyesuaikan diri, meskipun agak kurang nyaman, dengan rezim privasi saat ini. Misalnya, kami menyarankan agar perusahaan yang menggunakan AI membutuhkan Petugas Etika AI yang terpisah, dalam peran yang serupa dengan Petugas Perlindungan Data berdasarkan GDPR, tetapi dengan mandat khusus untuk mempertimbangkan etika AI.

Terdapat pertentangan antara kumpulan data besar yang dibutuhkan untuk hasil yang akurat dan konflik dengan kerangka kerja terkait privasi saat ini. Dengan meluasnya penggunaan pembelajaran mesin sebagai layanan (MLaaS), bisnis didorong untuk beroperasi di ambang inovasi yang ramah konsumen, sementara legalitas terkait privasi terkadang diabaikan.

Lebih sulit lagi, perusahaan yang mencoba mengikuti aturan dipaksa untuk membuat prediksi terkait konsekuensi dari teknologi yang sedang dikembangkan. Tidaklah masuk akal bagi para pemangku kepentingan bisnis untuk menanggung seluruh beban mencari solusi hukum yang pro-bisnis sekaligus memprediksi persimpangan teknologi masa depan dengan ekonomi dan kerangka hukum dalam matriks global.



Jelas bahwa untuk mematuhi kewajiban yang dijelaskan di atas, operator AI perlu menganalisis proses mereka saat ini untuk mengidentifikasi apakah proses tersebut perlu diperbarui atau diubah dengan cara apa pun untuk memastikan bahwa privasi merupakan inti dari pekerjaan AI mereka.

Misalnya, apakah organisasi telah mengidentifikasi dasar hukum yang relevan untuk pemrosesan? Apakah individu telah diberikan informasi yang cukup untuk memastikan bahwa persetujuan yang diberikan sah? Apakah ada proses yang diterapkan untuk memastikan bahwa pemrosesan ditinjau dan risikonya diidentifikasi? Ini bukanlah tugas yang mudah, namun, seperti yang telah disoroti di atas, AI itu sendiri dapat dimanfaatkan untuk mengurangi dampak terhadap sumber daya.

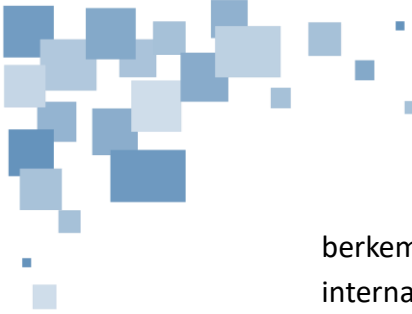
Kesimpulan Keseluruhan

1. Menghormati Privasi Sembari Mengizinkan AI

- ✚ Terdapat konflik yang inheren dan terus berkembang antara: meningkatnya penggunaan sistem AI untuk mengelola data pribadi, terutama data pribadi; dan meningkatnya perlindungan regulasi yang diberikan secara internasional terhadap data pribadi dan data pribadi lainnya.
- ✚ Organisasi yang menggunakan sistem AI memiliki kewajiban hukum dan etika untuk menghormati privasi pribadi; tetapi menghadapi tantangan untuk mematuhi peraturan privasi internasional yang semakin ketat, sambil tetap mengembangkan dan menggunakan teknologi AI.
- ✚ Jika kita ingin AI dapat berkembang (untuk tujuan yang positif dan bermanfaat), tantangan praktis yang dihadapi melalui 'benturan' dengan 'norma' dan peraturan privasi data ini perlu diatasi dan peraturan yang relevan perlu dikembangkan dan disempurnakan untuk mencoba mengakomodasi penggunaan AI yang sebenarnya namun terkendali.
- ✚ Pada gilirannya, organisasi yang menggunakan sistem AI untuk mengelola dan memproses data pribadi harus mengalokasikan sumber daya ke area kepatuhan peraturan, termasuk khususnya privasi data.

2. Peraturan Privasi dalam AI

- ❖ Saat ini, AI hanya memiliki sedikit undang-undang yang secara khusus menargetkan peraturannya: sebaliknya, AI harus menyesuaikan diri, dengan sedikit kurang nyaman, dengan rezim privasi saat ini.
- ❖ Jelas bahwa untuk mendukung norma-norma global peraturan privasi, diperlukan batasan yang dapat ditegakkan terhadap penggunaan AI.
- ❖ Ketika mempertimbangkan perlunya regulasi privasi untuk AI, kami menemukan 'patokan' atau titik awal yang sesuai, yaitu undang-undang privasi data Eropa, GDPR; undang-undang ini merupakan seperangkat prinsip perlindungan data terkini yang ada saat ini dan telah menyentuh aspek-aspek spesifik AI.
- ❖ Namun, tantangan AI terkait privasi data tidak diselesaikan oleh GDPR; kami masih perlu mencari kerangka hukum yang wajar untuk lingkungan AI, yang dapat



berkembang seiring tren menuju 'standar emas' untuk tingkat harmonisasi internasional regulasi Privasi AI.

3. Tantangan Operasional yang akan Dihadapi Pengguna AI

- ✓ Untuk memenuhi kewajiban yang dijelaskan di atas, operator AI perlu menganalisis proses mereka saat ini untuk mengidentifikasi apakah proses tersebut perlu diperbarui atau diubah dengan cara apa pun guna memastikan privasi menjadi inti dari pekerjaan AI mereka.
- ✓ Operator AI dalam banyak kasus tidak dapat menghindari undang-undang privasi melalui anonimisasi data; tetapi mematuhi peraturan privasi menghadirkan tantangan; karena memenuhi persyaratan persetujuan seringkali tidak praktis, yang mengarah pada kebutuhan akan justifikasi proses yang lebih besar dan perlindungan operasional tambahan perlindungan ini dirancang ke dalam sistem AI sejak awal.
- ✓ Kami menyarankan agar perusahaan yang menggunakan AI memerlukan Petugas Etika AI yang terpisah, dengan peran yang serupa dengan Petugas Perlindungan Data berdasarkan GDPR, tetapi dengan mandat khusus untuk mempertimbangkan etika dan kepatuhan peraturan atas penggunaan AI mereka.
- ✓ Meskipun demikian, tidaklah masuk akal bagi para pemangku kepentingan bisnis untuk menanggung seluruh beban mencari solusi hukum yang pro-bisnis sekaligus memprediksi persimpangan teknologi masa depan dengan kerangka hukum dalam matriks global.
- ✓ Perlu dilakukan upaya bersama para pemangku kepentingan bisnis ini untuk menemukan kerangka regulasi yang lebih baik untuk AI dan privasi, dengan mengatasi permasalahan praktis yang muncul.

4. AI Sebagai Alat untuk Mendukung Privasi


- Meskipun terdapat tantangan dari perspektif privasi terkait penggunaan AI, pada gilirannya, munculnya teknologi AI juga dapat digunakan untuk membantu organisasi mematuhi kewajiban privasi.

Prinsip

Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI dan undang-undang nasional yang mengatur penggunaan tersebut wajib berupaya memastikan bahwa sistem AI mematuhi norma dan peraturan privasi, dengan mempertimbangkan karakteristik unik sistem AI, dan perkembangan standar privasi.

Menemukan Keseimbangan

- ❖ Terdapat konflik yang inheren dan terus berkembang antara meningkatnya penggunaan sistem AI untuk mengelola data pribadi, khususnya data pribadi; dan meningkatnya perlindungan regulasi yang diberikan secara internasional terhadap data pribadi dan data pribadi lainnya.
- ❖ Pemerintah yang mengatur implikasi privasi sistem AI harus melakukannya dengan cara yang mengakui karakteristik spesifik AI dan tidak terlalu menghambat inovasi AI.
- ❖ Organisasi yang mengembangkan, menerapkan, dan menggunakan sistem AI harus menganalisis proses mereka saat ini untuk mengidentifikasi apakah proses tersebut



perlu diperbarui atau diubah dengan cara apa pun untuk memastikan bahwa penghormatan terhadap privasi menjadi pertimbangan utama.

Tantangan Operasional ke Depan bagi Pengguna AI

- Sistem AI menimbulkan tantangan khususnya terkait kepraktisan pemenuhan persyaratan di bawah sejumlah rezim legislatif nasional, seperti terkait persetujuan dan anonimisasi data. Oleh karena itu, organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI dan undang-undang nasional apa pun yang mengatur penggunaan tersebut harus menyediakan dasar hukum alternatif untuk pengumpulan dan pemrosesan data pribadi oleh sistem AI.
- Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus mempertimbangkan penerapan perlindungan operasional untuk melindungi privasi seperti prinsip privasi berdasarkan desain yang secara khusus disesuaikan dengan fitur spesifik sistem AI yang diterapkan.
- Organisasi yang mengembangkan, menerapkan, dan menggunakan sistem AI harus menunjuk seorang Petugas Etika AI, dengan peran yang serupa dengan Petugas Perlindungan Data berdasarkan GDPR, tetapi dengan kewenangan khusus untuk mempertimbangkan etika dan kepatuhan regulasi atas penggunaan AI mereka.

AI Sebagai Alat untuk Mendukung Privasi

- Meskipun terdapat tantangan dari perspektif privasi terkait penggunaan AI, pada gilirannya, munculnya teknologi AI juga dapat digunakan untuk membantu organisasi mematuhi kewajiban privasi.



BAB 9

AI DAN INTELEKTUAL PROPERTY

Implikasi AI terhadap hukum kekayaan intelektual (HKI) kemungkinan besar akan sangat signifikan dan luas, terutama seiring berkembangnya AI dan berbagai penerapannya. Dalam konteks ini, penting untuk mempertimbangkan sejauh mana kita perlu memperhatikan penetapan hak HKI dalam AI itu sendiri serta aset atau karya yang diciptakan oleh AI dan apakah hak HKI yang ada memberikan perlindungan yang memadai atau apakah hak baru perlu diciptakan. Dalam hal ini, apakah konsep hukum HKI tradisional, misalnya, "orisinalitas" dan "pencipta" masih tepat ketika melihat karya yang diciptakan oleh AI?

AI sendiri merupakan teknologi yang sangat inovatif atau serangkaian teknologi yang sangat inovatif. Lebih lanjut, aset atau karya yang diciptakan oleh AI dapat mencakup teknologi inovatif, perangkat lunak, karya seni, informasi rahasia, dan aset berwujud maupun tidak berwujud lainnya, yang semuanya akan dilindungi oleh hak HKI jika diciptakan oleh manusia. Namun, sebagaimana akan ditunjukkan di bawah ini, undang-undang HKI yang berlaku di banyak negara mungkin kurang tepat untuk menangani masalah kepemilikan karya yang diciptakan secara otonom oleh teknologi AI.

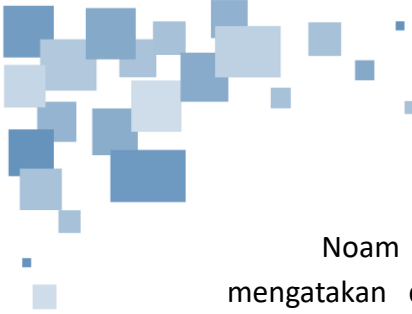
Khususnya dalam konteks HKI, pertanyaan tentang status hukum AI, dan apakah sistem AI dapat menjadi pemilik hak HKI atas ciptaannya, telah mengemuka. Bab ini membahas dua bidang utama, yaitu Hak Cipta dan Paten, yang bisa dibilang menjadi tempat munculnya isu-isu paling krusial. Namun, kami juga mempertimbangkan implikasinya terhadap hak HKI lainnya, khususnya rahasia dagang, hak basis data, perlindungan merek dagang dan merek, serta perlindungan desain. Pertanyaan umum dan penting tentang AI dan Status Hukum dibahas dalam Prinsip Akuntabilitas.

Konsekuensi Praktis

Ada sejumlah skenario unik yang perlu diingat ketika mempertimbangkan komplikasi antara AI dan HKI, ketika AI menciptakan, bertanggung jawab atas penciptaan, atau digunakan untuk menciptakan produk, karya, dan/atau proses baru.

Meskipun tampaknya sulit membayangkan sistem AI mengembangkan gaya seni baru (dapatkah sistem AI apa pun menciptakan lukisan Picasso atau lukisan Mark Rothko?), Prinsip Akuntabilitas mengacu pada contoh terkenal tentang bagaimana AI telah digunakan dalam dunia seni, yaitu proyek "Next Rembrandt", yang melibatkan grup perbankan Belanda ING, J Walter Thomson, Museum Rumah Rembrandt, Microsoft, dan lainnya, yang dipublikasikan pada tahun 2016. Proyek ini "dimaksudkan untuk memicu percakapan tentang hubungan antara seni dan algoritma, antara data dan desain manusia, dan antara teknologi dan emosi."

Selain itu, pada Oktober 2018, Christie's New York mencetak sejarah ketika menjadi balai lelang besar pertama yang menjual karya seni yang dibuat dengan kecerdasan buatan (AI). Karya tersebut terjual seharga Rp 4.325.000.000, lebih dari 4.320% dari estimasi tertingginya, yaitu Rp 100 juta.



Noam Brown, mahasiswa PhD Ilmu Komputer di Carnegie Mellon University, mengatakan dalam sebuah wawancara pada Desember 2018 bahwa ia tidak dapat membayangkan (setidaknya dalam waktu dekat) AI menciptakan novel pemenang penghargaan. Namun, sebuah program AI Jepang dilaporkan pada tahun 2016 telah ikut menulis novel pendek yang lolos seleksi tahap pertama untuk penghargaan sastra nasional, Penghargaan Sastra Hoshi Shinichi Jepang. Memang, pada tahun itu, 11 dari 1.450 pengajuan untuk Penghargaan (yang secara tegas terbuka untuk pelamar non-manusia, khususnya, "program AI dan lainnya") ditulis setidaknya sebagian oleh non-manusia. Skenario-skenario ini, di mana bahkan mungkin terdapat nilai finansial yang signifikan yang dikaitkan dengan sebuah karya yang dihasilkan oleh AI, menunjukkan betapa pentingnya menentukan siapa yang memiliki hak reproduksi atas karya tersebut.

Di bidang ilmu hayati dan pengurutan DNA, Blueprint Genetics menerapkan AI dalam interpretasi klinis untuk mengotomatiskan proses interpretasi yang melelahkan secara manual dan untuk memberdayakan ahli genetika dan klinisi dalam menginterpretasikan data uji pasien secara akurat dan konsisten.

Contoh lain adalah paten AS yang diberikan untuk sistem kecerdasan buatan yang lengkap untuk akuisisi dan analisis informasi hibridisasi susunan asam nukleat. Khususnya, pemilik paten tersebut adalah sebuah perusahaan, Iris BioTechnologies Inc, tetapi muncul pertanyaan tentang siapa yang mungkin memiliki penemuan yang diidentifikasi oleh sistem AI yang dipatenkan tersebut.

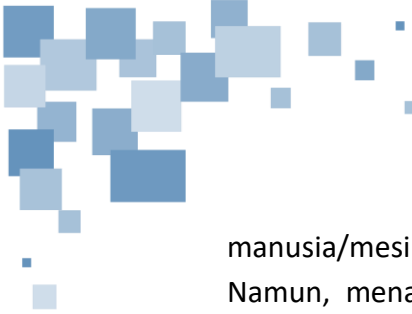
Dalam hal ini, sistem AI telah menunjukkan kemampuan untuk "berinovasi" dalam berbagai cara. Agen AI yang dikembangkan oleh Facebook menciptakan bahasa baru dalam upaya mengoptimalkan proses negosiasi. Bot AI yang dibuat oleh Google untuk menyelidiki metode navigasi rintangan menciptakan desain dan metode penggerak baru untuk robot pejalan kaki yang dapat menaklukkan medan yang sulit.

Sebagaimana akan ditunjukkan di bawah ini, undang-undang dan sistem kekayaan intelektual saat ini mungkin tidak menawarkan jawaban yang jelas untuk situasi di mana hak kekayaan intelektual tidak dapat melindungi aset yang dihasilkan oleh AI. Situasi ini umumnya sama di banyak negara di seluruh dunia. Apakah masuk akal atau praktis untuk melanjutkan pendekatan di mana tidak jelas siapa pemilik aset tak berwujud potensial yang diciptakan?

Apa yang Kita Bicarakan Ketika Kita Merujuk pada AI?

Salah satu dari banyak definisi AI adalah "pemahaman ilmiah tentang mekanisme yang mendasari pemikiran dan perilaku cerdas serta perwujudannya dalam mesin." Lebih lanjut, ketika mempertimbangkan perlindungan yang diberikan kepada AI oleh hak kekayaan intelektual, penting untuk mempertimbangkan berbagai jenis AI yang mungkin terlibat, seperti pembelajaran mesin, jaringan saraf tiruan, dan pembelajaran mendalam, di antara yang lainnya.

Lebih lanjut, mungkin terdapat isu dan/atau solusi yang lebih disukai, tergantung pada sejauh mana kita melihat AI yang sepenuhnya otonom dibandingkan dengan AI semi-otonom. Untuk keperluan bab ini, kami menganggap AI otonom melibatkan mesin yang bertindak di luar kendali manusia. Sementara itu, AI semi-otonom tunduk pada kendali bersama



manusia/mesin, tetapi tetap dapat memberikan hasil dan melakukan tindakan tanpa arahan. Namun, menarik juga untuk mempertimbangkan klasifikasi otonomi, dan perlunya skala otonomi ketika membahas AI (yaitu, klasifikasi dari 0 hingga 5, 0 merupakan AI non-otonom), yang dibahas dalam laporan tentang AI yang diterbitkan pada tahun 2018 atas permintaan Pemerintah Prancis dan di bawah koordinasi Cédric Villani. Semakin tinggi derajat otonomi AI, semakin signifikan implikasinya terhadap hukum kekayaan intelektual.

Apa Tujuan Perlindungan Kekayaan Intelektual?

Ketika mempertimbangkan opsi perlindungan kekayaan intelektual untuk produk, karya, dan proses yang dikembangkan oleh AI, dan, tentu saja, apakah perlindungan tersebut dibenarkan, penting untuk mengingat tujuan perlindungan kekayaan intelektual. Hal ini akan berdampak pada sejauh mana perubahan hukum kekayaan intelektual mungkin diperlukan untuk memberikan perlindungan kepada produk, karya, dan proses yang dikembangkan oleh AI.

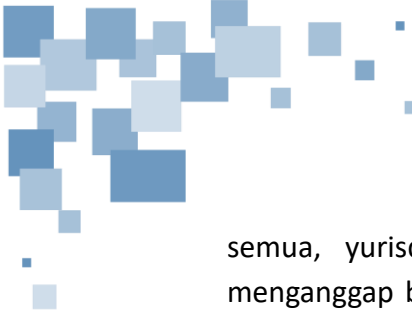
Berbagai hak kekayaan intelektual telah dikembangkan untuk melindungi produk, karya, dan proses, dengan tujuan untuk mendorong penciptaan karya intelektual dan oleh karena itu juga untuk mendorong inovasi di semua bidang. Hak Kekayaan Intelektual (HKI) umumnya memberikan imbalan atas hasil karya yang tidak berwujud.

Ada baiknya merenungkan bagaimana penciptaan karya intelektual dapat didorong oleh hak-hak HKI. Dalam beberapa kasus, pengakuan kepemilikan saja sudah cukup sebagai imbalan dan mungkin dapat diterima oleh pencipta perorangan, tetapi dalam banyak kasus, potensi monetisasi karya intelektual merupakan dorongan untuk penciptaannya.

Ketika seorang pencipta dipekerjakan dan dibayar untuk menciptakan karya intelektual, pemberi kerja akan ingin mendapatkan keuntungan finansial dari karya-karya tersebut, sehingga setidaknya diperlukan beberapa hak HKI, atau hak atas karya yang cukup untuk mendapatkan keuntungan finansial. Hal-hal tersebut perlu diingat ketika mempertimbangkan karya AI atau karya yang diciptakan oleh mesin ketika AI atau mesin tersebut didukung secara finansial oleh suatu entitas bisnis. Hingga mesin mulai menciptakan dan membuat mesin lain, pada suatu saat setidaknya satu manusia akan terlibat dalam proses yang mengarah pada sebuah karya yang diciptakan oleh AI/mesin.

Hukum HKI menetapkan aturan tentang siapa yang memiliki hak tidak berwujud yang mana. WIPO menjelaskan mengapa mempromosikan dan melindungi kekayaan intelektual bermanfaat. Dinyatakan bahwa kemajuan dan kesejahteraan umat manusia bergantung pada penciptaan dan invensi karya-karya baru di bidang teknologi dan budaya. Perlindungan hukum atas karya-karya baru ini mendorong komitmen sumber daya tambahan untuk inovasi lebih lanjut, yang pada gilirannya memacu pertumbuhan ekonomi, menciptakan lapangan kerja dan industri baru, serta meningkatkan kualitas dan kenikmatan hidup.

Yurisdiksi yang berbeda mungkin memiliki alasan yang berbeda untuk melindungi HKI. Di satu sisi, seseorang dapat mengklaim bahwa perlindungan HKI merupakan kompensasi atas upaya pencipta/penemu. Di sisi lain, dengan merujuk secara khusus pada perlindungan hak cipta, perlindungan yang diberikan kepada karya seni mencakup ekspresi kepribadian pencipta. Namun, alasan komersial kemungkinan besar penting di sebagian besar, jika tidak



semua, yurisdiksi. Hal ini tercermin dalam pendekatan pemerintah Tiongkok, yang menganggap bahwa perlindungan HKI merupakan kunci bagi perkembangan ekonomi yang sehat di Tiongkok.

Lingkungan perlindungan HKI yang baik dipandang penting dalam menarik investasi baik berupa uang maupun sumber daya manusia sehubungan dengan kegiatan penelitian dan pengembangan yang menyumbangkan teknologi baru dan yang akan menguntungkan kegiatan berbagai industri. Lebih lanjut, berdasarkan perjanjian internasional tertentu, Tiongkok memiliki kewajiban untuk membangun rezim perlindungan HKI guna membantu memastikan bahwa entitas asing dapat menjalankan bisnis di Tiongkok dengan aman.

Oleh karena itu, pihak yang mendukung peningkatan hak HKI menekankan manfaat perlindungan inovasi dan kreativitas, serta dorongan investasi melalui perlindungan. Hak IP juga memfasilitasi perdagangan dan jual beli karena para pihak dapat membeli, menjual, dan melisensikan hak dengan cara yang umumnya jelas bagi mereka. Ada juga cara yang mapan untuk menyelesaikan perselisihan tentang hak IP. Argumen lain yang mendukung hak IP adalah bahwa mereka merangsang penelitian dan inovasi. Inovasi hanya bernilai jika dapat bermanfaat bagi masyarakat.

Pendanaan diperlukan untuk mengembangkan inovasi ke tingkat di mana ia dapat disebarluaskan dan dimanfaatkan oleh masyarakat. Mereka yang darinya pendanaan dicari memerlukan pengembalian investasi mereka. Akibatnya, harus ada insentif dan perlindungan untuk inovasi jika ingin menarik investasi dan dibawa ke kebaikan masyarakat yang lebih besar.

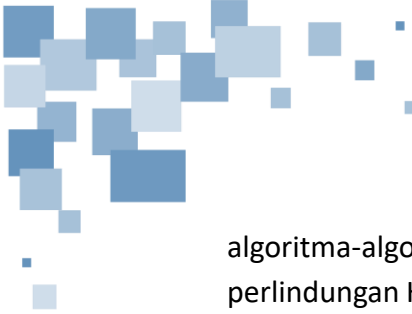
Biaya penelitian tinggi dan investor tidak akan membayar biaya itu tanpa peluang yang wajar untuk mendapatkan pengembalian investasi mereka. Yang juga penting adalah penyebaran informasi teknis melalui publikasi aplikasi paten. Hukum dan sistem yang memungkinkan hak IP ada terkadang rumit dan mahal. Namun, prospek tidak memiliki hak kekayaan intelektual (HKI) meningkatkan risiko kekacauan dan ketidakpastian, yang kemungkinan besar berdampak buruk bagi bisnis dan perekonomian. Oleh karena itu, penting bagi bisnis yang terlibat dalam inovasi semacam itu untuk mampu melindungi hak atas karya yang dihasilkan dari penggunaan AI.

Tentu saja, ada juga pendukung kuat kebebasan penuh terkait hak kekayaan intelektual, yang berpandangan bahwa pemberian hak monopoli justru menghambat dan menghambat pembangunan, dan bahwa tanpa hak kekayaan intelektual akan ada kesetaraan ekonomi dan politik yang lebih besar.

Oleh karena itu, dalam konteks AI, penting untuk mempertimbangkan apakah pemberian perlindungan HKI akan mendorong atau mematikan inovasi. Jika perubahan hukum HKI tidak diterapkan, telah dikemukakan bahwa kegagalan untuk mengakui AI sebagai sesuatu yang tidak dapat dibedakan dari manusia dalam hal kemampuan untuk memiliki hak HKI, hingga pada kesimpulan logisnya, berpotensi untuk membatalkan seluruh sistem HKI.

9.1 MELINDUNGI HAK DALAM AI

Kepemilikan dan perlindungan AI dan karya yang dihasilkan AI merupakan isu yang sulit ditentukan, karena modifikasi otonom dalam sistem AI terhadap algoritma pembelajaran atau



algoritma-algoritma yang responsif terhadap masukan ke sistem AI dan kurangnya perlindungan HKI khusus untuk algoritma. Bagian ini mengulas perlindungan hak dalam AI dan pentingnya mencapai keseimbangan yang tepat antara pemilik atau pengembang teknologi AI dan pihak ketiga yang menggunakan teknologi tersebut, serta penyedia set data.

Meningkatnya pentingnya AI dan algoritma menimbulkan pertanyaan terkait perlindungan. Salah satunya adalah apakah terdapat kesulitan dalam melindungi algoritma penyedia AI, terutama mengingat perbedaan rezim hukum di seluruh dunia (HKI, rahasia dagang). Haruskah penyedia AI meminta perlindungan HKI yang lebih ketat dalam berbagai undang-undang nasional tentang algoritma? Atau, haruskah pengguna AI yang menyediakan set data penting yang memungkinkan penyedia AI, atau bahkan sistem AI itu sendiri, untuk mempelajari dan meningkatkan algoritmanya, mendapatkan bagian HKI? Atau, alternatifnya, apakah undang-undang yang ada yang memungkinkan kepemilikan bersama, baik untuk hak cipta atau paten, sudah memadai?

Hak Cipta

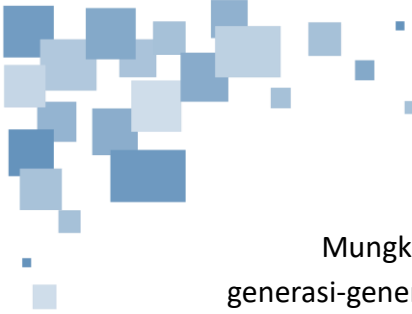
AI dan algoritma pengendalinya biasanya diimplementasikan dalam perangkat lunak. Meskipun perangkat lunak tersebut dilindungi hak cipta, hal ini mungkin tidak berlaku untuk algoritma yang diwujudkan. Di sebagian besar negara Eropa (misalnya, Swiss, Jerman, Prancis, Italia, Austria, Inggris, dll.) dan di AS, algoritma, sebagai ide, prinsip, atau metode, hanya dilindungi secara tidak langsung atau tidak sempurna berdasarkan hukum hak cipta. Algoritma hanya dapat dilindungi hak cipta setelah menjadi sebuah ekspresi, yaitu setelah diintegrasikan ke dalam kode sumber perangkat lunak yang dilindungi hak cipta.

Lebih lanjut, hukum hak cipta tidak menghalangi seseorang untuk menggunakan salinan perangkat lunak guna mengamati, mempelajari, atau menguji fungsi perangkat lunak tersebut guna menentukan ide dan prinsip yang mendasari setiap elemen perangkat lunak, saat menjalankan perangkat lunak tersebut. Pengguna tersebut kemudian dapat menciptakan perangkat lunak yang menerapkan ide dan prinsip yang sama secara independen tanpa menjadi pelanggar hak cipta.

Dengan demikian, meskipun implementasi algoritma dalam perangkat lunak dapat memperoleh manfaat dari perlindungan hak cipta, algoritma itu sendiri mungkin tidak terlindungi. Hal ini menunjukkan bahwa mengandalkan perlindungan algoritma kognitif hanya pada hak cipta tidak akan memberikan perlindungan yang memadai.

Paten

Selain memperlakukan AI dan algoritma sebagai rahasia dagang, instrumen perlindungan yang tersisa untuk dipertimbangkan bagi AI adalah perlindungan paten. Sifat teknologi AI yang terus berkembang juga kemungkinan akan menimbulkan masalah jika mempertimbangkan perlindungan paten. Untuk paten, tidak hanya akan ada masalah tentang siapa yang berhak mengajukan perlindungan paten, tetapi data dan algoritma akan menimbulkan sejumlah masalah mendasar terkait HKI. Misalnya, bagaimana Anda menciptakan hak milik dalam algoritma yang terus berubah, sehingga suatu penemuan tidak akan sama bahkan satu tahun setelah permohonan paten diajukan? Inti dari "daya cipta" AI terletak pada algoritma yang mendasarinya.



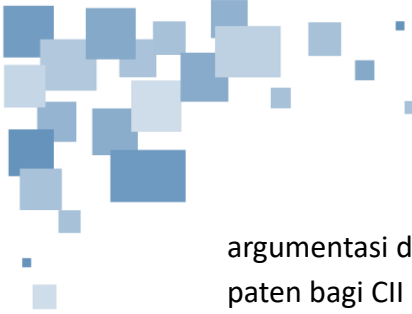
Mungkinkah algoritma generasi pertama yang asli dapat dilindungi, sementara generasi-generasi berikutnya merupakan produk dari algoritma awal dan dapat dilindungi melalui "produk per proses" (kemungkinan hanya dapat dilindungi hingga generasi kedua berdasarkan hukum paten Inggris)? Kuncinya adalah sejauh mana fungsi algoritma yang mendasarinya dimodifikasi.

Jika arsitektur fungsional tetap tidak berubah dan modifikasinya terbatas pada data parametrik, seperti bobot filter, kemungkinan tidak ada penemuan lebih lanjut dan algoritma generasi pertama bisa dibilang masih ada dan akan terus mendapatkan manfaat dari perlindungan paten yang diperoleh untuknya. Jika arsitektur fungsional dimodifikasi, mungkin ada pengaturan teknologi baru, yang mungkin disebut "sebuah penemuan," dan yang mungkin tidak termasuk dalam klaim paten atas penemuan generasi pertama sehingga perlindungan paten lebih lanjut harus diupayakan.

Melindungi penemuan perangkat lunak atau, sebagaimana terkadang disebut, penemuan yang diimplementasikan komputer (CII) telah menjadi semakin sulit dalam beberapa tahun terakhir di Uni Eropa dan khususnya di AS juga. Banyak orang di Eropa akan senang dengan posisi seperti di AS. Namun, keputusan Alice Corp. v. CLS Bank International telah menciptakan ketidakpastian yang cukup besar untuk perlindungan paten di AS. Keputusan Mahkamah Agung AS dalam "Alice" menetapkan bahwa metode berbisnis yang hanya diimplementasikan oleh komputer adalah "ide abstrak" yang tidak dapat dipatenkan.

Sejak keputusan dalam "Alice," pengadilan telah menunjukkan kecenderungan kuat untuk menilai kategori besar CII sebagai "ide abstrak" belaka yang dianggap tidak memenuhi syarat untuk perlindungan paten, terlepas dari apakah mereka diarahkan ke metode bisnis. Beberapa perubahan legislatif diantisipasi untuk memperjelas posisi tersebut. Kantor Paten dan Merek Dagang AS ("USPTO") masih bergulat dengan masalah ini dan pada Januari 2019 mengeluarkan pedoman revisi paten terkait dengan penemuan yang diimplementasikan komputer. Selain itu, Kongres AS mungkin perlu melengkapi Undang-Undang Paten untuk membahas patentabilitas. Bahkan ada beberapa pakar paten yang menggambarkan situasi terkini di AS setelah keputusan "Alice" sebagai "di luar kendali."

Oleh karena itu, meskipun sistem paten di Eropa kontinental pada awalnya dianggap kurang liberal dibandingkan dengan sistem di AS, ada beberapa komentator yang berpendapat bahwa kini sebenarnya lebih mudah untuk memperoleh perlindungan paten di Eropa untuk CII daripada di AS. Dalam hal ini, orang dapat mengajukan pertanyaan yang agak provokatif, apakah situasi ini bahkan dapat menjadi peluang bagi lanskap paten Eropa untuk meningkatkan perlindungan untuk CII dan meningkatkan posisi kompetitifnya dibandingkan dengan AS. Namun, setidaknya untuk saat ini, peluang perlindungan CII melalui paten juga tidak jauh lebih baik di Eropa. Sebagai contoh, Kantor Paten Eropa ("EPO") tidak terlalu sering menggunakan argumen subjek yang tidak memenuhi syarat paten. Sebaliknya, EPO cenderung mengabaikan fitur-fitur invensi yang diyakini EPO bersifat non-teknis, misalnya, tidak memecahkan "masalah teknis", dalam penilaian langkah inventif, yang mengakibatkan penolakan berdasarkan kurangnya daya inventif untuk fitur-fitur yang tersisa. Meskipun jalur



argumentasi di AS dan Eropa mungkin berbeda, kecenderungan untuk menolak perlindungan paten bagi CII saat ini tampaknya cukup serupa.

Beberapa pakar paten percaya bahwa alasan umum mengapa perlindungan CII melalui paten menjadi begitu sulit adalah kenyataan bahwa teknologi berkembang begitu pesat sehingga menghasilkan pemahaman yang sangat dinamis tentang apa yang mungkin merupakan "masalah dan solusi teknologi" dan, dengan demikian, kantor paten dan pengadilan tidak cukup cepat untuk mengikuti perkembangan yang mencengangkan ini. Akibatnya, beberapa teknologi yang sangat maju mungkin tidak mendapatkan manfaat dari ketersediaan perlindungan paten yang diperlukan untuk melindungi investasi besar dalam pengembangannya.

Terlepas dari tantangan-tantangan ini dalam memperoleh perlindungan paten untuk sistem berbasis AI dan CII, terdapat banyak paten untuk teknologi-teknologi ini yang telah diberikan oleh USPTO dan EPO dalam beberapa tahun terakhir. Meskipun sistem paten mungkin lambat beradaptasi dengan teknologi yang sedang berkembang, dan meskipun masih terdapat masalah yang cukup signifikan dalam melindungi AI secara memadai, tampaknya lembaga-lembaga ini sedang meningkatkan kemampuan mereka untuk memberikan perlindungan paten bagi kemajuan di bidang ini. Misalnya, EPO telah memperbarui praktiknya dalam mengkaji metode matematika. Berdasarkan praktik baru EPO, suatu metode matematika dapat berkontribusi pada karakter teknis suatu invensi, sehingga berkontribusi pada langkah inventif, jika klaim tersebut dibatasi pada "aplikasi teknis" spesifik dari metode matematika tersebut.

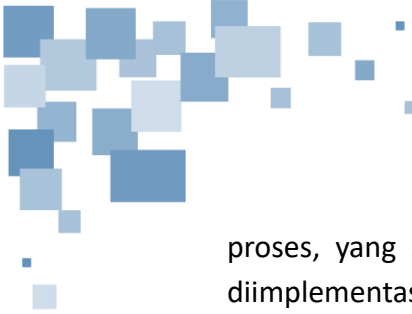
Hak Sui Generis untuk Melindungi Algoritma?

Pilihan lain yang perlu dipertimbangkan adalah hak sui generis untuk algoritma. Namun, sejauh ini, opsi untuk perlindungan hukum algoritma kognitif ini tampaknya belum ada dalam undang-undang apa pun, dan tampaknya tidak ada penyedia AI utama yang meminta amandemen legislatif tersebut.

Bagaimanapun, kita dapat mempertanyakan secara provokatif apakah sekitar 25 tahun yang lalu merupakan ide yang tepat untuk memperkenalkan hak sui generis untuk semua upaya inovatif yang berkaitan dengan perangkat lunak (program komputer, algoritma, hak topografi sirkuit, dll.) alih-alih memaksakannya ke dalam hukum hak cipta dan hukum kekayaan intelektual lainnya, yang didasarkan pada konsep yang bisa dibilang kurang tepat untuk perlindungan perangkat lunak.

Hak sui generis juga tidak memperhitungkan bahwa diperlukan upaya yang sangat besar untuk menyiapkan set data yang tepat agar sistem AI dapat belajar. Menerapkan hak sui generis khusus untuk perlindungan algoritma tampaknya belum menjadi keharusan saat ini, dan bagaimanapun juga, hal itu memerlukan pertimbangan dan kehati-hatian yang cermat agar tidak menggeser keseimbangan yang ada antara pemilik hak dan pengguna ke salah satu pihak.

Berbagai undang-undang kekayaan intelektual yang ada tampaknya masing-masing melindungi beberapa aspek kepemilikan hak dalam AI, meskipun tidak satu pun dari undang-undang tersebut memberikan perlindungan yang seragam. Karena algoritma adalah sebuah



proses, yang seringkali dinyatakan secara matematis dan/atau logis, dan mungkin dapat diimplementasikan dengan melibatkan cara-cara teknis, dengan menggunakan analisis ini sebagai titik awal, dimungkinkan untuk mengidentifikasi undang-undang kekayaan intelektual yang tepat (atau modifikasi apa pun darinya) yang dapat digunakan untuk memberikan perlindungan.

Dengan demikian, untuk sementara waktu, pengembang sistem AI harus mencari perlindungan AI mereka dengan menggabungkan fragmen-fragmen perlindungan yang ditawarkan oleh berbagai undang-undang kekayaan intelektual, dengan undang-undang hak cipta yang memberikan perlindungan untuk implementasi perangkat lunak spesifik suatu algoritma, undang-undang paten yang memberikan perlindungan terbatas untuk CII, dan undang-undang rahasia dagang yang memberikan perlindungan terbatas untuk pengetahuan rahasia. Namun, hal ini tidak berbeda dengan cara banyak produk yang sudah dilindungi, yaitu melalui kumpulan hak kekayaan intelektual (HKI). Meskipun terkadang, dan tergantung pada keadaan spesifiknya, hal ini terasa tidak nyaman, undang-undang HKI telah terbukti fleksibel dan dapat disesuaikan untuk menangani teknologi baru. Kemungkinan besar, hal ini juga akan berlaku untuk perlindungan AI.

Kecuali jika muncul kebutuhan yang terbukti dan mendesak untuk memperkenalkan perlindungan sui generis baru bagi AI yang kami yakini belum muncul saat ini tampaknya disarankan untuk membiarkan para pembuat undang-undang secara bertahap mengklarifikasi, mengembangkan, dan menyesuaikan undang-undang HKI yang ada jika diperlukan untuk mengatasi tantangan spesifik yang ditimbulkan oleh AI. Setiap perubahan signifikan secara keseluruhan kemungkinan besar perlu ditangani dalam skala internasional, dan badan-badan terkait harus mulai mempertimbangkan solusi semacam itu.

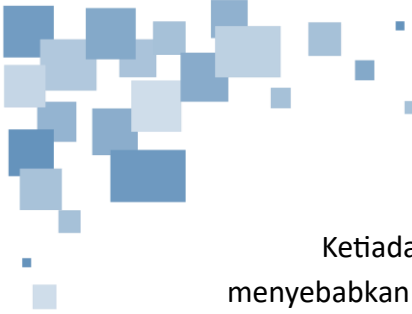
9.2 SIAPA PEMILIK HAK ATAS KARYA YANG DIHASILKAN OLEH AI?

Tidak adanya kreativitas manusia dalam tindakan modifikasi otonom sistem AI, tetapi kreativitas manusia dalam merancang penemuan, menyebabkan ketidaksesuaian terkait kepemilikan karya yang dihasilkan menggunakan AI: karya yang diciptakan oleh algoritma mungkin tidak dapat dilindungi tanpa perubahan peraturan.

Pada bagian ini, kami mempertimbangkan apakah mesin atau proses AI dapat menjadi pencipta suatu karya atau penemu suatu invensi, atau apakah pengertian penemu atau pencipta harus secara eksplisit hanya merujuk pada orang perseorangan. Jika AI otonom itu sendiri dianggap memiliki aset tak berwujud, kami mempertimbangkan langkah-langkah apa yang mungkin perlu diambil oleh legislatif untuk memungkinkan kepemilikan HKI.

Kami juga mempertimbangkan posisi di mana terdapat kolaborasi antara AI dan manusia dalam arti bahwa sistem AI otonom dan manusia bersama-sama menciptakan HKI baru. Dapatkah AI memiliki hak tersebut dalam karya yang dihasilkan oleh AI, jika tidak, siapa yang seharusnya memilikinya, jika ada? Apakah HKI tersebut akan dimiliki oleh pemilik atau pengguna sistem AI, penyedia kumpulan data yang relevan, atau apakah HKI tersebut berada dalam domain publik?

Hak Cipta



Ketiadaan kreativitas oleh atau hubungan dengan orang perseorangan dapat menyebabkan suatu karya tidak dapat dilindungi. Baik melalui pengertian hubungan seperti pemberi kerja/karyawan, kerja kolektif, atau status independen AI mungkin masih sulit untuk melindungi karya yang diciptakan oleh sistem AI.

1. Aturan Umum

Di sejumlah yurisdiksi, termasuk AS, Spanyol, Italia, Swiss, Prancis, Austria, dan Jerman, hanya karya yang diciptakan oleh manusia yang dapat dilindungi oleh hak cipta. Sebagai contoh, berdasarkan hukum AS, sebuah karya harus merupakan karya cipta asli, dituangkan dalam media ekspresi yang nyata, dan juga memiliki setidaknya tingkat kreativitas minimal dan Kantor Hak Cipta AS telah menyatakan bahwa mereka akan "mendaftarkan sebuah karya cipta asli, dengan syarat karya tersebut diciptakan oleh manusia."

Hal ini mengikuti yurisprudensi AS yang telah menetapkan bahwa hukum hak cipta hanya melindungi "hasil kerja intelektual" yang "didasarkan pada kekuatan kreatif pikiran." Di Australia, Pengadilan Federal menyatakan bahwa kode sumber, yang seharusnya dapat dilindungi hak cipta, tidak dapat dilindungi oleh hak cipta karena kode tersebut dihasilkan dengan intervensi komputer dan bukan semata-mata dibuat oleh manusia.


2. Pendekatan Hukum Perdata

Hal ini tampaknya konsisten dengan hukum hak cipta di sebagian besar negara yang menganut hukum perdata. Posisi terkait hak moral, yang umumnya diakui dalam yurisdiksi sipil dan tercantum dalam hukum Uni Eropa, juga relevan. Hak moral mengakui kepenulisan suatu karya, berbeda dengan hak kekayaan intelektual (HKI) material, termasuk hak untuk mengeksploitasi dan memonetisasi suatu karya. Sebagai contoh, meskipun hukum hak cipta Italia tidak secara tegas menyatakan bahwa suatu proses kreatif wajib dilakukan oleh seseorang, berdasarkan hukum Italia, ketentuan mengenai hak moral (yang tidak dapat dialihkan dan dipindahtangankan) memperjelas bahwa hanya manusia yang dapat didefinisikan sebagai pencipta.

Lebih lanjut, yurisprudensi Italia, sejalan dengan interpretasi yang diberikan oleh Mahkamah Kehakiman Uni Eropa, menafsirkan persyaratan kreativitas sebagai "refleksi kepribadian pencipta dan mengekspresikan pilihan bebas dan kreatifnya dalam produksi" karya yang dilindungi. Interpretasi ini tidak dapat secara masuk akal diadaptasi untuk tugas-tugas kreatif yang dilakukan oleh mesin dan AI.

Berdasarkan interpretasi tersebut, beberapa komentator Italia berpendapat bahwa karya yang diciptakan oleh sistem AI mungkin setara dengan karya yatim piatu sehingga menyamakan karya yang diciptakan AI dengan karya lain yang tidak ditetapkan sebagai domain publik, tetapi pengarangnya tidak diketahui. Di sejumlah yurisdiksi, di mana pengarangnya tidak diketahui, pengarang akan dianggap sebagai orang yang pertama kali menerbitkan karya tersebut. Pencantuman karya yang diciptakan AI dalam daftar karya yatim piatu, atau dalam daftar yang serupa, dapat mencegah karya tersebut jatuh ke dalam domain publik, sambil menunggu pembuat undang-undang memberikan perlindungan khusus.

Menarik juga untuk mempertimbangkan posisi di bawah hukum Spanyol. Menurut Undang-Undang Kekayaan Intelektual Spanyol, dianggap bahwa pengarang adalah orang



perseorangan yang menciptakan karya seni, sastra, atau ilmiah, sehingga tidak mencakup karya apa pun yang diciptakan oleh AI. Namun, Undang-Undang Kekayaan Intelektual Spanyol mengatur beberapa konsep hukum seperti karya kolektif, di mana suatu badan hukum dapat menjadi pemilik hak cipta atas karya kolektif tersebut, ketika badan hukum tersebut adalah orang yang mengungkapkannya, atau karya kolaboratif tersebut, dan menerbitkannya, atau karya-karya tersebut, atas nama mereka.

Dalam hal ini, apakah mesin tersebut merupakan agen kreatif? Dapatkah dianggap sebagai rekan penulis? Dapatkah dianggap sebagai karya kolaborasi antara sistem AI dan pemilik dan/atau pengguna sistem? Dalam hal ini, berapa persentase partisipasi yang sesuai untuk masing-masing? Kriteria apa yang akan diterapkan untuk menentukan persentase ini?

Perlu diperhitungkan bahwa sistem AI memiliki pembelajaran progresif, baik dengan dilatih maupun belajar mandiri, sehingga hal ini kemungkinan akan membuat posisi semakin rumit dari perspektif hukum, khususnya sistem AI yang lebih otonom dalam belajar. Selain itu, dapatkah kreasi sistem AI dianggap sebagai karya komposit? Ini berarti bahwa aspek-aspek yang berbeda dari karya tersebut kemudian akan membentuk karya hak cipta yang terpisah, yang mungkin memiliki penulis yang berbeda dan dilindungi untuk jangka waktu yang berbeda.


3. Komisi Aturan Hukum Perdata tentang Robotika

Pada 16 Februari 2017, Parlemen Eropa mengadopsi sebuah laporan berisi rekomendasi kepada Komisi Aturan Hukum Perdata tentang Robotika. Laporan ini dirujuk dalam Prinsip 2 Akuntabilitas dan, khususnya, kemungkinan menciptakan status hukum khusus untuk robot dalam jangka panjang. Laporan tersebut meminta Komisi untuk mengajukan proposal arahan tentang aturan hukum perdata tentang robotika, mengikuti rekomendasi terperinci yang tercantum dalam lampiran laporan.

Laporan ini membahas sejumlah isu yang secara tegas berkaitan dengan HKI. Laporan ini menyerukan Komisi untuk mendukung "pendekatan horizontal dan tekno-logika netral terhadap properti intelektual" terkait kecerdasan buatan. Namun, laporan tersebut menyatakan bahwa, terkait aliran data, meskipun tidak ada ketentuan hukum yang secara khusus berlaku untuk data, "rezim dan doktrin hukum yang ada dapat dengan mudah diterapkan pada robotika." Situasinya pada kenyataannya kemungkinan besar tidak sesederhana itu.

Memang, inilah kesimpulan yang dicapai oleh para komentator di Universitas Maastricht, seraya juga mengakui bahwa "...akhirnya, muncul seruan untuk mengesahkan teknologi-teknologi ini. Sebaiknya... tanpa menghambat inovasi dan dengan mempertimbangkan implikasi serta konsekuensi etis dan hukum." Hal ini mencerminkan seruan dari Komite Industri, Riset, dan Energi kepada Komite Urusan Hukum untuk sebuah resolusi bahwa "rekomendasi mengenai lisensi harus menghormati kebebasan kontraktual dan memberikan ruang bagi rezim lisensi inovatif" seraya memperingatkan terhadap "pengenalan hak kekayaan intelektual baru di bidang robotika dan kecerdasan buatan yang dapat menghambat inovasi dan pertukaran keahlian."

Dalam Pernyataan Penjelasan laporan tersebut, Komite meminta Komisi "untuk mengemukakan pendekatan yang seimbang terhadap hak kekayaan intelektual ketika



diterapkan pada standar dan kode perangkat keras dan perangkat lunak yang melindungi inovasi dan sekaligus mendorong inovasi,” dan bahwa “perumusan kriteria untuk ‘karya intelektual sendiri’ untuk karya berhak cipta yang dihasilkan oleh komputer atau robot sangat dibutuhkan.”

4. Perlindungan Hukum yang Ada untuk Karya Cipta yang Dihasilkan Komputer

Pilihan untuk memberikan kepengarangan hak cipta kepada pemrogram atau pengendali mesin telah diadopsi di negara-negara seperti Inggris, Hong Kong, India, Irlandia, dan Selandia Baru. Sebagai contoh, hukum Inggris secara tegas mengatur kepemilikan hak cipta jika terdapat karya cipta yang dihasilkan komputer. Umumnya, untuk memenuhi syarat perlindungan hak cipta berdasarkan undang-undang, karya tersebut harus "asli" dan, berdasarkan yurisprudensi, merupakan "karya intelektual sendiri" dari penciptanya. Tanpa intervensi legislatif apa pun, kemungkinan besar akan sulit untuk berargumen bahwa suatu karya yang diciptakan oleh AI dapat dianggap "asli" berdasarkan uji ini.


Di AS, terdapat ambang batas yang rendah untuk orisinalitas, dengan perlindungan diberikan kepada "elemen-elemen penyusunnya yang memiliki lebih dari kuantum kreativitas de minimis." Ambang batas yang rendah ini dapat berarti bahwa, berdasarkan hukum AS, karena sebagian besar karya akan "asli" dan karenanya dapat dilindungi hak cipta, doktrin orisinalitas tidak, dengan sendirinya, mencegah kepengarangan algoritmik.

Penulis suatu karya secara otomatis merupakan pemilik pertama hak cipta, dan istilah "penulis", tanpa pengecualian, didefinisikan berdasarkan hukum Inggris sebagai orang yang menciptakan karya tersebut, dengan klarifikasi tambahan yang diberikan untuk jenis karya tertentu. Namun, undang-undang tersebut secara tegas juga menetapkan bahwa untuk karya yang dihasilkan komputer, penulis akan dianggap sebagai "orang yang olehnya pengaturan yang diperlukan untuk penciptaan karya tersebut dilakukan."

Hal ini mungkin terkait dengan doktrin hak cipta AS "karya yang dibuat untuk disewa" berdasarkan Undang-Undang Hak Cipta 1976, yang menyatakan bahwa pemberi kerja atau entitas pemberi tugas pada awalnya akan memiliki hak cipta kecuali para pihak sepakat sebaliknya secara tertulis. Hal ini menunjukkan bahwa terdapat keadaan di mana hak materiil (berbeda dengan hak moral) tidak selalu menjadi milik penulis atau pencipta asli. Meskipun doktrin karya yang dibuat untuk disewa tidak mengizinkan kepengarangan AI, telah dikemukakan bahwa amandemen dapat dilakukan terhadap Undang-Undang Hak Cipta AS untuk mencakup karya yang dihasilkan komputer.

Dampak ketentuan hak cipta di Inggris terkait karya yang dihasilkan komputer belum dinilai oleh pengadilan. Meskipun belum ditafsirkan oleh yurisprudensi, pendekatan ini dapat berarti bahwa orang yang mengoperasikan komputer dianggap sebagai pencipta karya yang dihasilkan komputer.

Dalam sejumlah situasi, ini dapat menjadi cara yang tepat untuk mengatasi masalah ini dan memastikan bahwa karya yang dihasilkan komputer mendapatkan perlindungan hak cipta. Namun, tanpa adanya yurisprudensi substantif mengenai aspek hukum Inggris ini, masih terdapat sejumlah pertanyaan terbuka. Khususnya, mungkin tidak jelas siapa manusia "yang mengatur penciptaan karya" dan apakah, misalnya, pengguna manusia perangkat AI



merupakan pencipta karya tersebut, atau apakah ia merupakan pemrogram asli perangkat AI tersebut, atau bahkan hanya pemilikinya.

Ketidakpastian ini ditunjukkan dengan jelas oleh posisi di Tiongkok, di mana opini paling populer adalah bahwa karya yang diciptakan secara independen oleh AI tidak mungkin dianggap sebagai karya berdasarkan hukum hak cipta. Namun, beberapa akademisi berpendapat bahwa komisioner manusia dapat dianggap sebagai pemberi kerja AI, sehingga, jika karya AI dapat dilindungi oleh hukum hak cipta, komisioner manusia akan menjadi pemilik karya tersebut.

Di Jepang, dengan perubahan baru yang diperkenalkan ke dalam Undang-Undang Hak Cipta, yang dirujuk lebih lanjut dalam bagian IV di bawah ini tentang Pelanggaran dan Penambahan Teks dan Data, kini terdapat perlindungan hak cipta yang diberikan kepada karya yang dihasilkan komputer di Jepang.

solusi yang masuk akal dan pragmatis untuk kepemilikan hak cipta atas karya yang dihasilkan oleh AI mungkin mengikuti pendekatan ini, dan mengubah undang-undang agar komisioner manusia dari AI otonom menjadi pemilik aset tak berwujud. Namun, solusi ini juga menyisakan pertanyaan tentang siapa yang akan bertanggung jawab atas tindakan dan keputusan AI otonom, di mana AI menjadi lebih maju, dan tugas-tugas yang dialokasikan untuk AI memungkinkan sistem AI lebih bebas untuk membuat keputusannya sendiri.

Dalam keadaan ini, mungkin sulit untuk mengatakan siapa yang membuat pengaturan yang diperlukan untuk penciptaan suatu karya atau, bahkan, apakah ada orang yang membuat pengaturan yang diperlukan. Dalam skenario di mana AI sepenuhnya otonom, jika tidak ada orang yang membuat pengaturan yang diperlukan untuk penciptaan suatu karya yang membutuhkan orisinalitas agar hak cipta dapat berlaku, maka, tanpa solusi yang teridentifikasi dengan jelas, tidak akan ada hak cipta atas karya tersebut karena tidak akan ada penciptanya.

Apakah instruksi tingkat tinggi dari operator cukup untuk kepengarangan, apakah peran pemrogram sistem AI perlu diperhitungkan, atau apakah kita harus mengakui bahwa karya tersebut diciptakan tanpa campur tangan manusia sama sekali, yang mengakibatkan karya tersebut tidak dilindungi hak cipta? Hal ini dapat berdampak dramatis bagi pihak yang ingin memonetisasi karya yang dihasilkan menggunakan perangkat AI.

Semakin otonom suatu mesin, semakin besar tanggung jawab dan akuntabilitas yang dapat dibebankan kepadanya atas tindakannya, yang dapat mencakup penciptaan atau kepengarangan karya secara independen. Sebuah mesin yang secara wajar diberi tugas penciptaan atau kepengarangan karya dapat dianggap memiliki unsur kepribadian tersebut dan oleh karena itu dianggap sebagai seseorang untuk aktivitas tersebut, setidaknya untuk aktivitas tersebut.

Oleh karena itu, seseorang dapat memiliki kepengarangan, tetapi konsep kepribadian, meskipun parsial, dapat terbukti bermasalah dalam menentukan kepemilikan, karena kepemilikan HKI melalui kepemilikan mesin dapat dipandang sebagai bentuk perbudakan. Namun, gagasan tentang mesin yang menjadi karyawan juga penuh dengan kesulitan; misalnya jika undang-undang ketenagakerjaan berlaku. Hal ini mengarah pada kesimpulan bahwa konsep kepengarangan perlu dimodifikasi jika hal tersebut akan menentukan kepemilikan.

Patent

Mematenkan sistem dan solusi yang diimplementasikan melalui perangkat lunak dapat terus menjadi tantangan yang sangat besar dan oleh karena itu dapat menjadi penghalang atau penghambat dalam mencari perlindungan paten untuk sistem dan algoritma AI. Lebih lanjut, persyaratan umum dalam undang-undang paten untuk mengidentifikasi seorang penemu dan bahkan memperoleh pengalihan hak paten dari seorang penemu agar suatu badan hukum dapat menjadi pemohon paten dapat mengakibatkan suatu invensi yang diciptakan oleh AI tidak dapat memperoleh perlindungan paten hanya karena persyaratan untuk mengidentifikasi dan/atau memperoleh pengalihan hak paten dari seorang penemu tidak dapat dipenuhi.


Situasi kepemilikan invensi yang diciptakan dengan bantuan AI serupa dalam banyak hal dengan hukum hak cipta. Hukum paten di sebagian besar yurisdiksi menganggap seorang penemu sebagai pemilik pertama invensi tersebut. Hukum paten yang sama ini umumnya menganggap seorang penemu sebagai siapa pun yang benar-benar membayangkan, atau merancang, invensi tersebut dengan cara tertentu. Di AS, "pertanyaan ambang batas dalam menentukan status penemu adalah siapa yang membayangkan invensi tersebut. Kecuali seseorang berkontribusi pada konsepsi invensi tersebut, ia bukanlah seorang penemu." Di Inggris, pendekatan dua langkah diterapkan untuk menentukan status penemu; pertama, identifikasi konsep inventif dan kemudian tentukan siapa yang merancang konsep tersebut.

Dengan definisi istilah yang begitu luas, AI dapat dipandang sebagai "penemu" dalam berbagai keadaan. Baik AI digunakan untuk mengidentifikasi senyawa farmasi baru, maupun untuk menentukan solusi optimal untuk desain pesawat terbang, argumen yang dapat diajukan adalah bahwa AI berperan dalam "menggagas" atau "merancang" solusi tersebut. Namun, di Eropa, AS, Tiongkok, dan di tempat lain, dipahami bahwa AI tidak dapat menjadi penemu, atau pemilik, suatu invensi karena ia bukan "orang" dalam pengertian hukum apa pun dan tidak memiliki hak hukum yang terkait dengan kepribadian.

Ada banyak aktor yang terlibat dalam memproduksi atau menciptakan invensi yang diciptakan oleh AI, seperti pemrogram perangkat lunak, pencipta algoritma, penyedia data untuk digunakan dalam sistem AI, pelatih dan pemeriksa sistem AI yang memasukkan dan memeriksa input dan output data, pemilik sistem AI, seperti kepemilikan IBM atas Watson, dan operator sistem. Sejauh mana masing-masing menjadi pemilik AI atau apakah AI satu-satunya pemilik?

Untuk saat ini, tampaknya tidak ada yurisprudensi khusus yang menilai dalam kondisi apa manusia yang berinteraksi dengan perangkat atau sistem AI dapat dianggap sebagai penemu ajaran teknis yang diciptakan dengan bantuan AI. Kondisi diskusi di yurisdiksi tertentu dapat diartikan bahwa, untuk dianggap sebagai penemu, manusia tidak diharuskan "menciptakan" suatu ajaran teknis.

Sebaliknya, pandangan AI yang diakui dan masuk akal dalam situasi ini adalah sebagai alat, seperti alat lainnya, yang digunakan oleh satu atau lebih individu untuk mencapai solusi teknis. Untuk dianggap sebagai penemu, cukup dengan "mengenali" suatu ajaran teknis dalam beberapa data yang dihasilkan mesin yang memecahkan masalah teknis. Dalam situasi ini,



manusia yang mengenali solusi teknis dalam keluaran tersebut dianggap sebagai penemu untuk penemuan tersebut. Di AS, konsepsi merupakan tolok ukur penemu dan setiap penemu harus berkontribusi pada konsepsi tersebut. Selain itu, kepemilikan mengikuti penemu. Jadi siapa pemiliknya? Namun, dalam kasus AI otonom yang menghasilkan suatu penemuan, paradigma pencipta-sebagai-penemu dan penemu-sebagai-pemilik yang ditemukan di sebagian besar yurisdiksi dapat menyebabkan konflik yang inheren. Jika AI diketahui telah menggagas atau merancang penemuan tersebut tanpa intervensi yang memadai oleh seorang individu, mungkin tidak ada pemilik sah.

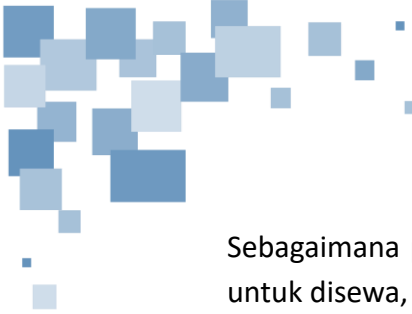
AI yang merupakan satu-satunya "penemu" dalam skenario ini tidak dapat memiliki penemuan tersebut, karena tidak memiliki status badan hukum dan hak kepemilikan yang sebenarnya. Meskipun konflik ini paling jelas terlihat dalam kasus sistem AI yang sepenuhnya otonom (yang masih jauh dari kenyataan), bahkan penggunaan AI saat ini pun dapat menimbulkan masalah. AI sering digunakan saat ini untuk mengidentifikasi cara optimal untuk mendapatkan hasil yang diinginkan dengan kata lain, hasilnya diketahui dan disediakan oleh pengguna, dan AI mengembangkan cara untuk mendapatkan hasil tersebut.

Yurisprudensi AS menunjukkan bahwa sekadar memberikan hasil yang diinginkan tidaklah cukup untuk menjadikan seseorang sebagai "penemu". Untuk memperhitungkan potensi konflik ini, yurisprudensi dan rezim hukum harus beradaptasi seperlunya untuk mengakui peran AI sebagai alat bagi manusia, yang hanya menyediakan bantuan, meskipun dalam bentuk yang agak maju, dan untuk mengakui individu yang mengarahkan dan menggunakan AI sebagai penemu untuk setiap pengembangan yang dirancang atau dirancang dengan bantuan AI.

Perlu diperhatikan bahwa di era digital, upaya inovatif biasanya tidak hanya memengaruhi satu pihak, melainkan merupakan upaya pengembangan bersama, dalam hal ini, antara penyedia AI dan pengguna AI. Dengan demikian, perlindungan HKI untuk masukan kreatif adalah sesuatu yang semakin banyak diminta oleh pengguna AI.

Meskipun undang-undang HKI yang ada saat ini tidak memberikan jawaban yang mudah atas pertanyaan kepemilikan terkait hak HKI, undang-undang tersebut telah beradaptasi di masa lalu dan, dengan pendekatan yang dibahas di bagian ini, hanya diperlukan sedikit perubahan, jika memang diperlukan, saat ini untuk mengklarifikasi atau mengubah undang-undang HKI yang ada, juga untuk melindungi kreasi yang dikembangkan dengan perangkat yang dibantu AI. Salah satu area di mana pengadilan dan legislator dapat mengeksplorasi klarifikasi untuk mencakup kreasi dan invensi AI, adalah apakah menyimpulkan bahwa komisaris manusia dari AI otonom harus menjadi pemilik hak HKI.

Namun, pendekatan semacam itu pasti akan menimbulkan kesulitan dalam menentukan siapa yang membuat pengaturan yang diperlukan untuk penciptaan suatu karya atau, seiring perkembangan AI, apakah ada yang membuat pengaturan yang diperlukan. Kebutuhan untuk membahas dimensi internasional juga tetap penting di sini. Namun, ketika mempertimbangkan siapa yang memiliki, atau siapa yang seharusnya memiliki, karya yang dibuat menggunakan AI, tetap penting untuk membedakan kepemilikan dari kepengarangan.



Sebagaimana pemilik belum tentu merupakan pengarang, misalnya pada karya yang dibuat untuk disewa, pendekatan yang sama dapat diterapkan pada karya yang dibuat oleh AI.

9.3 PELANGGARAN

Apa yang merupakan, atau seharusnya merupakan, pelanggaran juga menimbulkan isu yang kompleks, termasuk di mana tanggung jawab atas pelanggaran seharusnya berada. Namun, hal ini harus konsisten dengan penentuan siapa pengarangnya. Jika AI dapat menjadi pengarang, AI juga harus menjadi pelanggar.

Hak Cipta

1. Pelanggaran Hak Cipta yang Diakibatkan oleh Pelatihan Sistem AI

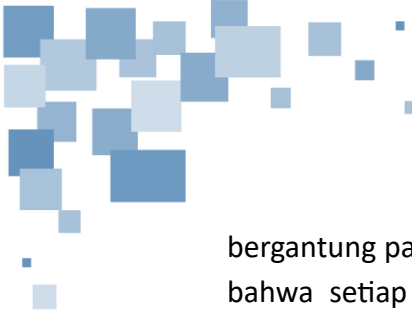
Pelatihan dan pengembangan AI seringkali bergantung pada pembelajaran mesin yang, pada gilirannya, membutuhkan kumpulan data yang sangat besar. Misalnya, pemrosesan bahasa alami bergantung pada analisis dan pemrosesan teks dalam jumlah besar untuk tujuan pelatihan, dan penciptaan novel tidak akan mungkin terjadi tanpa analisis tersebut.

Kecuali dalam kasus di mana set data pelatihan spesifik yang disediakan dan dilisensikan untuk melatih AI digunakan, data pelatihan yang digunakan untuk melatih AI biasanya mencakup sejumlah besar materi, seperti novel, artikel surat kabar, gambar, dll., yang mungkin memiliki hak cipta pihak ketiga dan materi tersebut harus disalin dan dimodifikasi ketika digunakan untuk melatih AI. Hal ini menimbulkan pertanyaan apakah penggunaan materi pelatihan berhak cipta untuk melatih AI tersebut merupakan pelanggaran hak cipta penulis, dan jika demikian, apakah pengecualian harus diperkenalkan dalam undang-undang hak cipta untuk mengecualikan penggunaan tersebut.

Untuk menyelesaikan masalah ini, beberapa pihak menyerukan pembedaan antara penggunaan ekspresif dan non-ekspresif. Penggunaan non-ekspresif, menurut argumen tersebut, akan jauh lebih mirip dengan penggunaan karya yang diizinkan dan tidak melanggar hak cipta (seperti pembacaan buku oleh orang yang membelinya di toko) daripada pembuatan salinan karya untuk eksploitasi komersial. Terlebih lagi, hal itu tidak akan menyampaikan karya asli penulis kepada publik dan juga tidak berfokus pada elemen ekspresif dari karya tersebut. Sebaliknya, tujuan penyalinan adalah untuk mendapatkan ide dan metodologi yang mendasarinya (seperti sintaksis, tata bahasa, frekuensi kata, dll.), yang karenanya tidak dilindungi oleh hak cipta.

Argumen ini sebagian besar didasarkan pada doktrin penggunaan wajar, dengan mempertimbangkan khususnya fakta bahwa pemrosesan untuk tujuan pelatihan AI tidak menggantikan karya asli dan tidak berdampak pada potensi pasar dan nilai karya berhak cipta. Pihak lain mendukung pandangan bahwa penambangan teks dan data (TDM) yang merupakan bentuk teknologi AI yang digunakan untuk melatih AI tidak boleh berada di bawah kendali pemilik hak karena perbedaan antara ide dan ekspresi dan karena hal tersebut bukan merupakan eksploitasi atas karya berhak cipta.

Hal di atas tidak diakui secara luas, khususnya di Eropa, dan "penggunaan wajar" tetap menjadi konsep di AS. Yurisdiksi Eropa umumnya mengambil pendekatan yang jauh lebih formalistis mengenai apa yang dimaksud dengan reproduksi yang memerlukan izin dan



bergantung pada daftar panjang pengecualian spesifik dari prinsip yang luas dan menyeluruh bahwa setiap penyalinan apa pun tujuannya melanggar hak cipta penulis. Pengecualian tersebut mencakup, misalnya, hak untuk menyalin karya untuk tujuan pengarsipan, reproduksi gratis untuk tujuan pembelajaran, untuk kutipan, serta hak untuk membuat salinan sementara atau insidental sebagai bagian dari transmisi jaringan atau penggunaan yang sah.

Kebutuhan untuk memasukkan pengecualian semacam itu termasuk khususnya yang berkaitan dengan salinan sementara atau insidental dalam hukum hak cipta Eropa menggarisbawahi bahwa pembelaan penggunaan non-ekspresif, sebagaimana diuraikan di atas, kemungkinan akan sulit jika diadili di pengadilan di Eropa. Hal ini semakin dipertegas oleh diskusi terkini tentang revisi hukum hak cipta Uni Eropa, yang akan memperkenalkan pengecualian TDM yang sempit, yang telah menimbulkan beberapa kontroversi dalam diskusi mengenai topik tersebut.


2. Pengecualian Penambangan Teks dan Data

Sejak 2014, undang-undang hak cipta Inggris telah memasukkan pengecualian TDM tambahan. Pengecualian ini secara khusus mengizinkan salinan karya untuk dibuat guna analisis teks dan data untuk penelitian nonkomersial. Ini berarti bahwa jika seseorang memiliki akses yang sah ke suatu karya, mereka dapat membuat salinannya untuk tujuan melakukan analisis komputasional atas apa pun yang tercatat dalam karya tersebut, dengan ketentuan bahwa hal ini semata-mata untuk tujuan penelitian nonkomersial dan salinan tersebut disertai dengan pengakuan yang memadai (kecuali karena alasan kepraktisan hal ini tidak memungkinkan).

Ketentuan ini selanjutnya menetapkan bahwa hak cipta dilanggar jika salinan yang dibuat dialihkan kepada orang lain, atau digunakan untuk tujuan selain yang diizinkan oleh pengecualian tersebut, kecuali diizinkan oleh pemilik hak cipta. Terdapat ketentuan tegas bahwa salinan yang dibuat untuk analisis teks dan data tidak boleh dijual atau disewakan. Menariknya, ketentuan tersebut menyatakan bahwa setiap ketentuan kontraktual yang bertujuan untuk mencegah atau membatasi pembuatan salinan yang diizinkan berdasarkan pengecualian tersebut tidak akan dapat diberlakukan.

Pada tahun 2016, Undang-Undang Republik Digital diadopsi di Prancis, memperkenalkan ketentuan mengenai pengecualian TDM. Ketentuan ini menyatakan bahwa penulis tidak dapat lagi menolak "salinan atau reproduksi digital" yang dibuat "dari sumber yang sah" dari karya mereka "untuk tujuan mengeksplorasi teks dan data yang disertakan atau terkait dengan tulisan ilmiah untuk keperluan penelitian publik." Oleh karena itu, dan dengan penekanan pada fakta bahwa karya tersebut bukan untuk tujuan komersial, peneliti secara teoritis berhak menggunakan algoritma yang menyalin dan menganalisis data tertentu, seperti artikel atau laporan, untuk analisis massal.

Ketentuan serupa juga diperkenalkan terkait hak basis data sui generis, yang memberikan pengecualian terhadap hak produsen basis data terkait reproduksi yang dibuat untuk tujuan penggalian teks atau data, dengan ketentuan bahwa data yang dimaksud telah diperoleh secara sah, TDM hanya dilakukan untuk "kebutuhan penelitian publik," dan bukan untuk "tujuan komersial apa pun."



Sebuah dekrit diharapkan akan merinci lebih lanjut terkait penerapan ketentuan baru ini, tetapi belum ada dekrit yang disahkan oleh pemerintah Prancis sehingga ketentuan TDM dalam CPI tidak dapat diberlakukan. Ketentuan terkait hak sui generis hanya sebagian, tetapi tidak sepenuhnya, dapat diberlakukan karena belum ada otoritas yang ditunjuk untuk mengontrol penyimpanan dan komunikasi berkas yang dihasilkan oleh TDM. Draf dekrit untuk mengatur ketentuan TDM yang baru ditolak pada Mei 2017, dan tidak ada kemajuan lebih lanjut yang dicapai sejak saat itu, bahkan mengingat usulan undang-undang Uni Eropa tentang masalah ini.


Di Eropa, Direktif baru mencakup pengecualian TDM terhadap perlindungan hak cipta untuk tujuan penelitian ilmiah. Meskipun hal ini tidak secara tegas mengatur AI, hal ini berpotensi memiliki implikasi yang signifikan bagi AI. Pengecualian TDM hanya berlaku untuk "organisasi penelitian" dan hanya sejauh TDM dipraktikkan secara nirlaba. Akibatnya, penyedia AI komersial mungkin memerlukan lisensi lebih lanjut untuk tujuan penerapan TDM pada data yang telah mereka akses secara sah, yang memaksa mereka untuk menghubungi setiap pemilik data sebelum menerapkan AI pada data tersebut.

Secara khusus, terdapat banyak perdebatan mengenai apakah pengecualian tersebut harus bersifat wajib atau hanya opsional. Pandangan yang dikemukakan adalah bahwa menjadikan pengecualian tersebut wajib akan "membuka jalan bagi pengecualian TDM yang luas dan efektif di seluruh Negara Anggota, yang memungkinkan peneliti dan perusahaan Eropa untuk tetap kompetitif dalam persaingan AI global." Tak perlu dikatakan lagi bahwa pembatasan apa pun atas penerapan pengecualian TDM untuk tujuan penelitian non-komersial akan secara signifikan membatasi dampak yang diinginkan dari pengecualian TDM tersebut untuk kegiatan komersial.

Perusahaan dari semua ukuran di Eropa terus berinvestasi dalam analitik data dan analitik prediktif mereka menggunakan TDM untuk memproses konten dalam jumlah besar dan mengekstrak (atau menemukan) wawasan baru dari data. Oleh karena itu, banyak penyedia AI menginginkan pengecualian TDM yang luas agar perusahaan dapat memanfaatkan pengecualian tersebut untuk melindungi investasi mereka dan memungkinkan perusahaan di Eropa bersaing dengan perusahaan di wilayah lain di dunia. Perusahaan di negara-negara di luar Eropa mendapatkan manfaat dari pengecualian yang luas dan pengecualian penggunaan wajar untuk TDM, dan yang terpenting, karena akses yang sah terhadap karya tersebut merupakan prasyarat, TDM tidak memengaruhi pasar karya asli.

Akibatnya, dapat disimpulkan bahwa pengecualian TDM yang sempit akan menimbulkan inkonsistensi yang parah antara Uni Eropa dan negara-negara penting lainnya di dunia, di mana perusahaan-perusahaan diuntungkan oleh pengecualian TDM yang luas berdasarkan prinsip-prinsip penggunaan wajar. Untuk mencegah inkonsistensi tersebut dan mencegah perusahaan-perusahaan di Eropa menghadapi kerugian kompetitif tambahan, beberapa pelaku industri AI percaya bahwa entitas mana pun yang memiliki akses sah ke karya yang dilindungi hak cipta seharusnya mendapatkan manfaat dari pengecualian TDM.

Mereka mungkin akan berargumen bahwa pengecualian TDM semacam itu tidak akan berdampak negatif terhadap pasar karya asli, dan bahkan mungkin memicu peningkatan minat



terhadap karya tersebut. Masyarakat Hak Cipta Eropa (*European Copyright Society*) sangat mendukung pengecualian yang lebih luas, dengan alasan bahwa hal ini akan mendorong lebih banyak inovasi dan "memastikan kesetaraan dengan Amerika Serikat, di mana perusahaan yang terlibat dalam aktivitas TDM kemungkinan besar akan mendapatkan manfaat dari pengecualian penggunaan wajar."

Dengan demikian, undang-undang hak cipta Eropa yang ada cenderung mengkualifikasikan penggunaan karya berhak cipta untuk tujuan pelatihan AI sebagai pelanggaran hak pencipta atau pemilik hak cipta, kecuali undang-undang hak cipta yang berlaku memberikan pengecualian yang mengizinkan penggunaan khusus tersebut.

Meskipun diskusi di Eropa masih berlangsung, yurisdiksi lain mengambil pendekatan yang berbeda. Misalnya, Parlemen Jepang mengesahkan undang-undang yang memperbarui Undang-Undang Hak Cipta pada 12 September 2018, yang pada dasarnya memungkinkan penambahan data teks oleh semua pengguna dan untuk semua tujuan, baik komersial maupun non-komersial.

Khususnya, Pasal 30-4 undang-undang amandemen tersebut mengizinkan eksploitasi karya berhak cipta apa pun untuk tujuan "analisis informasi", termasuk "ekstraksi, perbandingan, klasifikasi, atau analisis statistik lainnya atas bahasa, suara, gambar, atau elemen lain yang membentuk sejumlah besar karya atau sejumlah besar informasi." Ketentuan-ketentuan ini kemungkinan akan menjadikan Jepang "surga" bagi AI dan pembelajaran mesin, menempatkan negara ini di garda terdepan dalam persaingan ini.


Pelanggaran Hak Cipta atau Paten Pihak Ketiga oleh AI

1. Hak Cipta

Ketika digunakan dalam praktik, AI dapat menyebabkan pelanggaran hak cipta pihak ketiga. Misalnya, AI otonom yang digunakan oleh penerbit untuk menulis novel atau artikel surat kabar dapat menyimpulkan dari riwayat membaca penggunanya bahwa semakin sering AI menyalin konten dari buku atau surat kabar lain, semakin baik umpan balik pembaca, sehingga mulai hanya menyalin konten dari sumber pihak ketiga, sehingga melanggar hak cipta pihak ketiga tersebut. Hal ini menimbulkan pertanyaan apakah hal ini merupakan pelanggaran dan, jika demikian, apakah AI itu sendiri atau, sebagai tambahan atau alternatifnya, "pemilik" atau orang lain yang mengendalikan AI dapat dimintai pertanggungjawaban.

Dalam kerangka hukum Eropa saat ini, sebagaimana dibahas lebih lanjut dalam Prinsip 2 Akuntabilitas, AI tidak memiliki kepribadian hukum; ia tidak dianggap sebagai "orang". Pada saat yang sama, undang-undang hak cipta biasanya menganggap bahwa "seseorang" adalah pencetus pelanggaran. Dengan demikian, kerangka hukum saat ini tidak memperkirakan bahwa sistem otonom dapat dikualifikasikan sebagai pencetus pelanggaran hak cipta yang dapat dimintai pertanggungjawaban atas tindakan atau kelalaiannya.

Oleh karena itu, pihak yang dirugikan akan berupaya meminta pertanggungjawaban seseorang yang memiliki badan hukum yang dapat dikaitkan dengan AI. Misalnya, ini bisa berupa pengguna yang menyebarkan dan merilis AI, pengembang yang memprogram algoritma, atau produsen yang menyediakan AI untuk digunakan oleh pihak ketiga. Apakah



sebagian atau semua orang ini dapat dimintai pertanggungjawaban akan bergantung pada rezim pertanggungjawaban yang berlaku.

Seringkali, rezim semacam itu akan mengatur pertanggungjawaban atas kerugian berdasarkan hukum perdata, yang biasanya mengharuskan orang yang dirugikan untuk menunjukkan bahwa ia menderita kerugian, bahwa kerugian tersebut disebabkan oleh orang yang dianggap bertanggung jawab, dan bahwa orang tersebut bertindak dengan kesalahan.

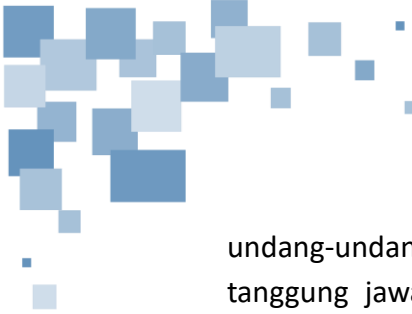
Dalam praktiknya, menunjukkan hubungan sebab akibat dan kesalahan dapat terbukti sangat sulit karena kurangnya pemahaman yang memadai tentang fungsi AI dan tindakan serta kelalaian terduga pelanggar terkait pembuatan atau penggunaan AI yang memicu kerugian. Mengingat kesulitan-kesulitan tersebut, dapat dipertimbangkan apakah akan menerapkan tanggung jawab mutlak (*strict liability*) untuk AI yang hanya membutuhkan pembuktian kerugian dan sebab akibat atau pendekatan manajemen risiko yang tidak berfokus pada orang yang bertindak bersalah, melainkan pada orang yang memiliki kemampuan untuk meminimalkan risiko dan menangani dampak negatif. Namun, perubahan semacam itu tidak boleh diterapkan tanpa mempertimbangkan dampaknya secara saksama. Khususnya, penerapan tanggung jawab mutlak yang luas dapat berdampak buruk pada inovasi, menghambat adopsi AI sehingga merugikan kepentingan publik, dan mengalihkan tanggung jawab secara keliru kepada orang-orang yang tidak memiliki pengaruh atas AI dan tidak menyadari kekurangannya.

Lebih lanjut, perlu dipertimbangkan bahwa kesulitan dalam membuktikan sebab akibat dan kesalahan tidak hanya terjadi pada pelanggaran yang disebabkan oleh AI, tetapi juga harus sering diatasi dalam gugatan pelanggaran apa pun yang melibatkan teknologi lain. Untuk saat ini, tampaknya masih diragukan apakah AI telah mencapai tahap yang memerlukan aturan khusus. Oleh karena itu, perubahan pada aturan tanggung jawab yang ada hanya boleh diterapkan jika dan di mana terdapat kegagalan hukum yang berlaku yang terbukti.

2. Paten

Sehubungan dengan paten, apa yang merupakan pelanggaran juga akan memiliki kompleksitas yang signifikan. Mungkin terdapat berbagai entitas yang terlibat pada berbagai tahap proses, seperti yang telah disebutkan sebelumnya, mengembangkan teknologi AI, menyediakan kumpulan data, yang masing-masing dapat menimbulkan kerumitan dalam hal tanggung jawab dan liabilitas terkait paten. Namun, secara umum, hal ini tidak berbeda dengan banyak skenario pelanggaran paten, terutama yang melibatkan pelanggaran kontribusi atau tidak langsung atau perbuatan melawan hukum bersama. Setelah tindakan "pelanggaran" diidentifikasi, di mana tindakan tersebut terjadi dan siapa yang melakukan tindakan tersebut merupakan tahap penyelidikan selanjutnya. "Siapa" mungkin menjadi kesulitan jika pelanggaran melibatkan mesin, khususnya mesin otonom. Pada tahap penyelidikan inilah pertimbangan serupa terkait kepribadian mesin kemungkinan akan muncul.

Jadi, siapa yang dapat dimintai pertanggungjawaban atas pelanggaran dalam skenario yang melibatkan tindakan mesin otonom, dan apakah perlu mempertimbangkan kemungkinan untuk meminta pertanggungjawaban AI itu sendiri sebagai pelanggar? Bagaimana jika AI atau mesin otonom mengembangkan produk baru yang melanggar paten orang lain? Secara umum,



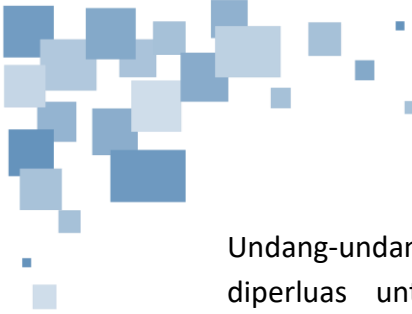
undang-undang paten di sebagian besar yurisdiksi dirancang secara luas untuk mengenakan tanggung jawab atas berbagai aktivitas yang berkaitan dengan produk yang melanggar. Undang-undang paten Inggris mengenakan tanggung jawab untuk membuat, membuang, menawarkan untuk membuang, menggunakan, mengimpor, atau menyimpan produk yang dipatenkan tanpa persetujuan.

Demikian pula, undang-undang paten AS mengenakan tanggung jawab untuk setiap pembuatan, penggunaan, penjualan, penawaran untuk menjual, atau mengimpor produk yang dipatenkan tanpa izin. Akibatnya, untuk setiap produk yang melanggar biasanya ada beberapa entitas yang dapat dikaitkan dengan tanggung jawab dari produsen, distributor, penjual, dan bahkan pengguna akhir. Jadi, bahkan jika AI atau mesin otonom mungkin terlibat dengan produk yang melanggar, hampir pasti akan ada entitas individu atau korporat yang sebenarnya yang dapat dikenakan tanggung jawab.

Memang, dalam hampir setiap skenario yang melibatkan produk yang melanggar hak paten, pemilik paten tidak perlu atau ingin membebankan tanggung jawab kepada AI atau mesin otonom yang tidak dapat menagih ganti rugi. Sebaliknya, pemilik paten hanya akan menargetkan individu atau badan usaha yang menghasilkan pendapatan dari pelanggaran hak paten tersebut untuk memaksimalkan ganti rugi sama seperti pemilik paten yang memilih untuk menargetkan perusahaan bernilai jutaan dolar yang memproduksi produk yang melanggar hak paten, alih-alih individu yang relatif insolven yang menggunakannya.

Isu pelanggaran dan pertanggungjawaban yang lebih kompleks yang melibatkan AI kemungkinan besar melibatkan proses atau metode yang dipatenkan, alih-alih produk. Jika suatu metode atau proses yang dipatenkan memerlukan serangkaian langkah, dan salah satu langkah tersebut dilakukan oleh AI otonom, apakah pelanggaran tersebut ada dan siapa yang dapat dimintai pertanggungjawaban? Dengan kata lain, dapatkah calon pelanggar terhindar dari pertanggungjawaban dengan membiarkan AI otonom melakukan satu atau lebih langkah dari suatu metode yang dipatenkan, atau bahkan semua langkah dari metode tersebut? Jawaban atas pertanyaan ini tidaklah sesederhana itu.

Misalnya, meskipun undang-undang paten AS mengenakan tanggung jawab untuk "siapa pun yang secara aktif mendorong pelanggaran paten", Mahkamah Agung AS telah memutuskan bahwa pelanggaran yang didorong tersebut memerlukan tindakan pelanggaran "langsung" di mana semua langkah dapat dikaitkan dengan satu aktor. Jika satu aktor tidak melakukan semua langkah metode tersebut, suatu entitas tidak dapat bertanggung jawab kecuali jika "mengarahkan atau mengendalikan" kinerja langkah-langkah lain oleh entitas atau entitas lain. Dengan cara ini, undang-undang paten AS menggabungkan prinsip-prinsip tanggung jawab keagenan, juga ditemukan di Inggris dan yurisdiksi lain, di mana prinsipal dapat bertanggung jawab atas tindakan agennya. Namun, dalam kasus AI otonom yang mungkin dapat mengembangkan atau memilih sendiri cara melakukan tugas tertentu, dapatkah dikatakan bahwa entitas lain mengarahkan atau mengendalikan perilakunya? Kita dapat membayangkan skenario di mana perancang sistem AI berubah atau berevolusi begitu jauh dari desain awalnya sehingga perancang tidak lagi bertanggung jawab atas tindakan sistem AI, dan AI bukan lagi agen karena, pada dasarnya, telah "melampaui wewenangnya."



Undang-undang pertanggungjawaban keagenan di berbagai yurisdiksi mungkin perlu diperluas untuk memperhitungkan skenario semacam itu dan untuk mengaitkan pertanggungjawaban kepada perancang awal. Landasan untuk perluasan pertanggungjawaban tersebut dapat ditemukan dalam konsep "wewenang aktual tersirat" dan "wewenang biasa".

Dalam mengeksplorasi isu-isu pelanggaran, klarifikasi atau pengembangan area "penggunaan wajar" atau beberapa pengecualian lain akan menjadi penting. Secara khusus, pertimbangan harus diberikan apakah penggunaan materi berhak cipta untuk melatih AI memenuhi syarat sebagai bentuk penggunaan yang diizinkan atau tidak. Namun, di masa mendatang, dalam hal menangani isu-isu pelanggaran, AI harus dianggap sebagai alat yang tidak memiliki kepribadian hukum. Akibatnya, kemungkinan besar AI tidak dapat bertanggung jawab atas perilaku pelanggaran apa pun dan, bergantung pada skenarionya, perancang, penyedia data, atau pemilik AI harus bertanggung jawab atas tindakan tersebut berdasarkan prinsip hukum yang berlaku.

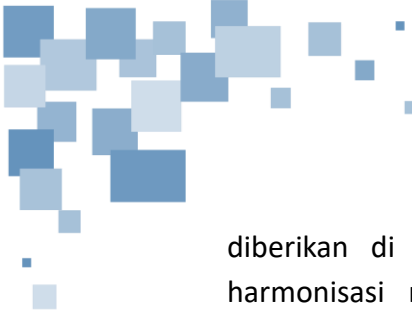
9.4 ISU-ISU KEKAYAAN INTELEKTUAL LAINNYA

Rahasia Dagang

Pilihan yang sering digunakan adalah memperlakukan AI dan algoritma sebagai rahasia dagang dan melindunginya berdasarkan hukum kontrak, dengan klausul dan perjanjian kerahasiaan dan non-pengungkapan yang sesuai. Data mentah, misalnya data pengukuran, berpotensi menjadi rahasia dagang/informasi rahasia. Rahasia dagang dapat menjadi bentuk perlindungan Kekayaan Intelektual yang bermanfaat untuk AI, seperti kode sumber, algoritma, dan set data pelatihan AI. Memang, para komentator telah mempertanyakan apakah perlindungan rahasia dagang mungkin lebih tepat untuk invensi AI, dan untuk set data yang digunakan dalam AI, misalnya, daripada mengandalkan perlindungan paten, terutama mengingat teknologi yang berubah dengan cepat, kesulitan dalam mendeteksi pelanggaran paten, dan menemukan kebaruan yang diperlukan terkait dengan invensi AI. Rahasia dagang mungkin efektif untuk memberikan perlindungan kepada invensi atau set data AI, sejauh rekayasa balik tidak memungkinkan, atau bahwa invensi atau set data yang dimaksud tidak dapat dengan mudah dikembangkan secara independen.

Jelas, perlu dipastikan bahwa persyaratan yang relevan berlaku untuk invensi atau set data tersebut. Uni Eropa baru-baru ini menerapkan arahan baru tentang Rahasia Dagang. Arahan ini memberikan definisi baru yang telah diselaraskan tentang rahasia dagang, yang mensyaratkan bahwa informasi tersebut bersifat rahasia (dalam arti tidak diketahui secara umum, atau tidak mudah diakses oleh, orang-orang yang biasanya menangani informasi ini), memiliki nilai komersial karena dirahasiakan, dan langkah-langkah yang wajar telah diambil untuk merahasiakannya "oleh setiap orang atau badan hukum yang secara sah mengendalikan informasi tersebut."

Definisi ini secara eksplisit mensyaratkan pemegang rahasia dagang adalah orang atau badan hukum, sehingga, tentu saja, akan mengecualikan "rahasia dagang" apa pun yang diciptakan oleh AI otonom. Namun demikian, mungkin ada perlindungan tambahan yang



diberikan di yurisdiksi yang berbeda, karena Arahan ini merupakan undang-undang harmonisasi minimal sehingga Negara Anggota diizinkan untuk memberikan tingkat perlindungan yang lebih tinggi.

Rahasia dagang dilindungi di AS, baik di tingkat federal maupun di tingkat negara bagian, oleh undang-undang rahasia dagang negara bagian. Berdasarkan hukum federal dan negara bagian, perlindungan rahasia dagang berlaku secara luas untuk informasi bisnis, keuangan, dan teknis (seperti kode sumber dan data) dengan ketentuan bahwa informasi tersebut tidak diketahui secara umum atau dapat ditentukan di luar organisasi dan kendali pemilikinya, dan bahwa pemilik informasi tersebut memperoleh nilai ekonomi atau keuntungan bisnis independen dari informasi yang tidak diketahui secara umum tersebut dan telah melakukan upaya yang wajar untuk menjaga kerahasiaannya. Namun, jika rahasia dagang diandalkan, dan kemudian terjadi rekayasa balik oleh AI, kecil kemungkinannya untuk menetapkan tanggung jawab secara signifikan.

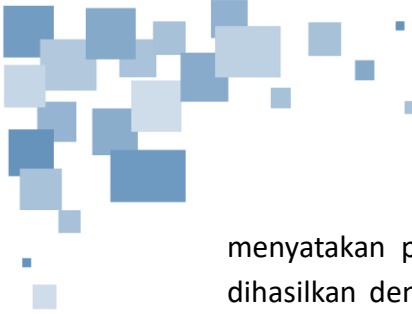
Bagaimanapun, upaya untuk mengandalkan rahasia dagang untuk perlindungan penemuan atau kumpulan data AI dapat bertentangan dengan persyaratan atau keinginan untuk transparansi dan keadilan terkait AI. Hal ini khususnya relevan ketika AI digunakan dalam lingkungan industri di mana algoritma yang tidak berfungsi dengan baik tidak hanya menghasilkan keluaran yang tidak diinginkan atau salah, tetapi juga dapat mengakibatkan insiden keselamatan yang dapat menyebabkan bahaya serius bagi manusia atau lingkungan. Satu kegagalan saja bahkan dapat menyebabkan kerugian. Dalam kasus-kasus penggunaan yang sangat penting bagi keselamatan, baik probabilitas kegagalan maupun keterjelasan keputusan AI sangatlah penting. Jika probabilitas kegagalan tidak dapat dinilai karena kurangnya transparansi atau keterjelasan, AI tersebut sebaiknya tidak digunakan untuk kasus-kasus penggunaan yang relevan dengan keselamatan atau sebaiknya didukung oleh sistem lain yang dapat mengoreksi dan mengguguli keputusan AI.

Hak Basis Data

Ilmuwan data menghabiskan sekitar 80% waktu mereka untuk mengumpulkan, membersihkan, dan mengelola data. Ini berarti bahwa sebagian besar upaya untuk menciptakan nilai dengan AI berkaitan dengan pengumpulan, penyiapan, dan pengelolaan data untuk analisis. Dengan demikian, AI dan basis data berjalan beriringan, dan AI pada dasarnya penting dalam kaitannya dengan lanskap teknologi basis data yang terus berubah.

Uni Eropa mengambil langkah pada Februari 1996 untuk memberikan perlindungan khusus kepada basis data. Arahan Basis Data Uni Eropa (96/9/EC) memberikan hak sui generis yang melindungi data yang terstruktur dalam sebuah "basis data" terlepas dari orisinalitasnya, tetapi dengan syarat telah ada investasi substansial dalam memperoleh, memverifikasi, atau menyajikan isinya. Hak ini merupakan tambahan atas perlindungan yang diberikan oleh hak cipta terhadap struktur basis data dan pemilihan atau pengaturan isinya.

Pada bulan April 2018, Komisi Eropa meninjau dampak dari Arahan Basis Data, khususnya dalam konteks pengembangan teknologi baru. Apakah Arahan Basis Data melindungi data yang dihasilkan mesin dan AI merupakan salah satu pertanyaan yang secara tegas dipertimbangkan oleh Komisi Eropa dalam laporan ini. Namun, laporan tersebut



menyatakan pandangan bahwa Arahan Basis Data tidak berlaku untuk "basis data yang dihasilkan dengan mesin, sensor, dan teknologi baru lainnya," seperti kecerdasan buatan. Laporan tersebut berpandangan bahwa tidak jelas bagaimana teknologi ini diatur dan apakah definisi "basis data" saat ini dalam Arahan Basis Data mencakupnya. Namun, laporan tersebut juga memunculkan pertanyaan lebih lanjut mengenai apakah teknologi ini harus mendapatkan perlindungan berdasarkan hak basis data sui generis. Dalam konteks ini, Laporan mengidentifikasi bahwa pandangan mengenai apakah perlindungan harus diberikan dalam hal ini sangat terpolarisasi: mungkin tidak mengherankan dengan pembuat basis data umumnya mendukung pemberian perlindungan, sementara pengguna menginginkan hak akses ke data.

Dalam mencapai kesimpulannya, laporan ini meninjau dampak sejumlah keputusan CJEU terkait hak sui generis. Dampak dari keputusan-keputusan ini tampaknya adalah bahwa hak sui generis tidak mencakup banyak layanan ekonomi data, termasuk layanan berbasis kecerdasan buatan. Tanpa klarifikasi lebih lanjut tentang cakupan hak oleh CJEU, ketidakpastian hukum akan tetap ada. Setiap intervensi kebijakan yang signifikan terhadap hak sui generis harus substansial dan mempertimbangkan perdebatan kebijakan seputar ekonomi data secara lebih umum. Oleh karena itu, reformasi hak sui generis saat ini dianggap sangat tidak proporsional dengan potensi kebijakan keseluruhan dan nilai yang dikaitkan dengannya oleh para pemangku kepentingan.

Laporan ini mempertimbangkan pandangan responden konsultasi publiknya yang telah mempertimbangkan cara terbaik untuk mengatasi ketidakpastian yang muncul terkait interaksi antara AI dan teknologi baru lainnya serta hak basis data sui generis. Secara khusus, laporan ini membahas pandangan Profesor Mathias Leistner, seorang ahli hukum dan pakar Jerman di bidangnya, yaitu bahwa "karena sifat khusus data yang dihasilkan oleh sensor/mesin, M. Leistner berpendapat bahwa hak sui generis harus diubah untuk data tersebut. Hal ini disebabkan oleh ambang batas investasi substansial yang relatif rendah, ketidakpastian di balik situasi spin-off, dan fakta bahwa pembuat basis data lain akan membutuhkan kumpulan data lengkap lainnya untuk membuat kumpulan data baru, sehingga secara otomatis akan melanggar hak sui generis."

Laporan tersebut menyatakan bahwa: "salah satu kesimpulan mungkin adalah berupaya mempertahankan hak sui generis dalam beberapa bentuk (agar tetap memiliki nilai dalam hukum yang ada bagi mereka yang menghargainya), tetapi pada saat yang sama mencoba meminimalkan dampak negatif (biaya) terhadap ekonomi pengetahuan dan ekologi data." Dengan demikian, laporan tersebut menyarankan bahwa, meskipun tetap mempertahankan hak basis data di seluruh Uni Eropa, Komisi Eropa dapat mempertimbangkan untuk menyediakan hak tersebut melalui pendaftaran dan mengatasi perbedaan implementasi di antara negara-negara anggota dengan memperkenalkan hak tersebut melalui Peraturan Uni Eropa, alih-alih melalui Direktif.

Kesimpulan ini dapat menimbulkan kesulitan bagi bisnis yang bergerak di bidang AI, yang mungkin menganggap bahwa mereka memiliki peluang terbatas untuk melindungi investasi mereka. Namun demikian, Komisi Eropa telah menegaskan bahwa penerapan hak sui generis dalam konteks ekonomi data harus dipantau secara ketat.

Merek Dagang dan Perlindungan Merek

AI tentu akan memberikan manfaat bagi pemilik merek dalam hal waktu dan biaya pencarian merek dagang serta persiapan pemberitahuan penghapusan. Namun, mungkin ada konsekuensi lain bagi pemilik merek ketika teknologi AI digunakan, khususnya ketika keputusan pembelian dibuat oleh AI, alih-alih oleh manusia. Hal ini dapat berdampak pada konsep konsumen rata-rata dan apakah konsep ini perlu diubah untuk memperhitungkan pembelian dengan AI. Selain itu, sejauh mana kebutuhan untuk menunjukkan kebingungan atau kemungkinan kebingungan mungkin perlu diadaptasi. Gartner telah memprediksi bahwa 85% interaksi pelanggan akan dikelola oleh kecerdasan buatan pada tahun 2020, dan hal ini kemungkinan akan berdampak signifikan pada merek, meskipun para komentator berpendapat bahwa merek dapat menjadi usang sebagai akibatnya.

Ketika AI bertanggung jawab atas keputusan pembelian, ada kemungkinan di masa depan AI tidak akan memperhitungkan merek setidaknya untuk pembelian tertentu, misalnya harga, ketersediaan, umpan balik pelanggan yang semuanya kemungkinan akan berdampak pada nilai merek dan merek dagang.


Konsep kunci dalam hukum merek dagang bagi konsumen rata-rata, dalam hukum saat ini, dianggap cukup terinformasi, cukup cermat, dan jeli, tetapi jarang ada kesempatan untuk membuat perbandingan langsung antar merek dan sebaliknya harus bergantung pada ingatan yang tidak sempurna terhadap merek yang relevan. Lebih lanjut, tingkat perhatian konsumen rata-rata bervariasi sesuai dengan kategori barang atau jasa yang bersangkutan. Tidak jelas sejauh mana parameter ini masih berlaku jika AI bertanggung jawab atas keputusan pembelian, khususnya karena, sebagai program komputer, AI mampu mengingat dengan sempurna.

Demikian pula, mempertimbangkan konsep kebingungan atau kemungkinan kebingungan sepertinya tidak relevan jika AI terlibat, sekali lagi, karena, sebagai program komputer, AI mampu mengingat dengan sempurna. Mempertimbangkan isu pelanggaran, riwayat kasus hingga saat ini yang melibatkan interaksi antara AI dan hukum merek dagang, pengadilan telah menyimpulkan bahwa perusahaan seperti Amazon, Google, dan eBay tidak bertanggung jawab atas pelanggaran. Namun, jika dampak AI meningkat, seperti yang tampaknya mungkin, haruskah ada konsep tanggung jawab baru yang diperkenalkan untuk memberikan perlindungan?

Perlindungan Desain

Bagian ini membahas isu-isu yang timbul dari penggunaan AI dalam kaitannya dengan desain. Secara khusus, kemungkinan besar penggunaan AI dalam industri fesyen, konsumen, dan produk industri akan melibatkan isu-isu yang berkaitan dengan desain.

Perlindungan hak desain di berbagai yurisdiksi seringkali kompleks, dengan kombinasi hak desain terdaftar dan tidak terdaftar, terutama di Uni Eropa. (Tingkat perlindungan yang sama tidak ada di Amerika Serikat, maupun di sejumlah yurisdiksi lain, kecuali beberapa yurisdiksi yang mengatur perlindungan hak desain untuk desain industri berdasarkan Perjanjian Den Haag Mengenai Deposit Internasional Desain Industri.) Di Uni Eropa, hak desain telah terbukti sangat berharga dalam industri fesyen dan produk konsumen.



Secara umum, pemegang pertama hak desain umumnya adalah pencipta desain; atau, jika desain tersebut diciptakan selama masa kerja, pemberi kerja. Para komentator telah menekankan pentingnya perlindungan hak desain untuk produk AI dan robotika, tetapi hanya sedikit, jika ada, komentar mengenai konsekuensi AI dan perlindungan hak desain, dalam situasi di mana AI bertanggung jawab untuk menghasilkan desain yang relevan, atau untuk melanggar desain tertentu. Namun, kemungkinan besar setiap isu tentang kepemilikan atau pelanggaran akan membahas isu yang serupa dengan yang dibahas terkait paten.


AI akan memiliki dampak yang signifikan terhadap semua aspek HKI, terutama dalam dua bidang utama, yaitu hak cipta dan paten. Untuk mendukung dan mendorong inovasi dalam pengembangan penggunaan AI yang akan bermanfaat bagi masyarakat, penting untuk memungkinkan bisnis menarik investasi dan memperoleh imbal hasil investasi. Dalam konteks tersebut, perlindungan HKI yang tepat akan menjadi kunci. Bisnis harus memiliki undang-undang yang memadai agar mereka dapat mengambil langkah-langkah yang diperlukan untuk melindungi hak-hak mereka, jika sesuai, termasuk memperoleh hak cipta dan paten, jika berlaku, dan menerapkan ketentuan kontraktual untuk memungkinkan perlindungan sebagai rahasia dagang.

Saat ini, undang-undang HKI belum sepenuhnya memadai untuk menangani penciptaan karya baru sebagai hasil dari AI. Sebagian besar yurisdiksi mewajibkan keterlibatan manusia dalam kaitannya dengan penciptaan hak HKI. Legislatur perlu memutuskan apakah akan mempertahankan persyaratan ini untuk beberapa elemen manusia dan, jika demikian, bagaimana menerapkannya pada karya yang diciptakan oleh AI otonom dan karya yang didukung AI.

Dianggap bahwa perlindungan tertentu harus diberikan kepada karya yang diciptakan menggunakan AI, khususnya dengan mengklarifikasi dan mengembangkan hak-hak yang ada. Legislatur pasti akan terus mempertimbangkan apakah kepemilikan hak harus diberikan kepada pengembang asli AI, atau apakah hak sui generis baru harus diberikan untuk karya yang diciptakan oleh AI, bahkan hak yang mewajibkan pendaftaran oleh orang yang mengklaim hak tersebut.

Pada saat yang sama, meskipun undang-undang selalu perlu berkembang untuk mengakomodasi perkembangan baru, tampaknya tepat untuk mengambil pendekatan yang hati-hati ketika menilai dan memutuskan apakah ada kebutuhan untuk menerapkan aturan baru dan khusus AI dalam undang-undang HKI. Memperkenalkan hak baru apa pun kemungkinan besar hanya akan memperumit kumpulan hak HKI yang sudah kompleks. Banyak sistem hukum telah terbukti cukup fleksibel dan netral terhadap teknologi, sehingga dapat menangani tantangan yang muncul akibat perkembangan teknologi baru tanpa perlu menerapkan perubahan signifikan.

Hanya jika undang-undang yang ada terbukti gagal menangani isu-isu terkait AI dengan tepat, revisi atau perluasan undang-undang yang ada perlu dilakukan. Jika diperlukan, pemerintah harus memperkenalkan undang-undang yang sesuai dan/atau interpretasi yang didorong oleh AI terhadap undang-undang yang ada untuk memperjelas perlindungan HKI atas



karya yang didukung AI maupun yang diciptakan oleh AI, tanpa berupaya menciptakan hak HKI baru pada tahap ini.

Oleh karena itu, para legislator harus memantau perkembangan yang relevan dengan saksama, dan bertindak aktif setelah terbukti gagal nya undang-undang HKI yang ada dalam menangani tantangan terkait AI. Namun, ketika memperluas atau mengubah undang-undang HKI yang ada, atau menerapkan undang-undang HKI yang baru, penting untuk selalu berupaya menyeimbangkan kepentingan semua pemangku kepentingan terkait sejauh mungkin. Selain itu, perlu diperhatikan untuk tidak mengambil langkah-langkah yang mengarah pada perlindungan berlebihan, karena hal ini dapat berdampak buruk pada tujuan akhir perlindungan HKI.

Pada saat yang sama, mereka harus mengingat bahwa perkembangan di bidang ini dari perspektif teknologi bergerak sangat cepat, sehingga kemungkinan besar, pada saat legislator mengambil langkah apa pun untuk mengatasi masalah tersebut, penggunaan AI sudah berkembang sedemikian rupa sehingga langkah-langkah yang telah diambil telah digantikan. Hingga saat itu, penting untuk mencapai konsensus internasional, idealnya dengan perlunya perjanjian internasional untuk mengatasi masalah kepemilikan HKI, untuk mencapai konsistensi di berbagai yurisdiksi dan di berbagai hak HKI.

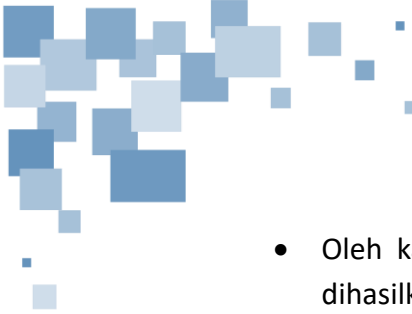
Akan lebih baik untuk menemukan solusi yang menjamin bahwa semua negara dapat memperoleh manfaat dari penyebaran teknologi baru yang cepat. WIPO berpendapat bahwa, alih-alih "menetapkan hak kekayaan intelektual baru, untuk melepaskan potensi ekonomi digital yang sangat besar demi pertumbuhan yang lebih inklusif di seluruh dunia, para pembuat kebijakan harus merangkul lingkungan hukum dan peraturan yang memungkinkan arus data lintas batas tanpa hambatan dalam skala global. Lingkungan semacam itu mengharuskan tidak adanya persyaratan lokasi data yang tidak beralasan serta penerapan aturan yang jelas dan dapat ditegakkan." Hal ini menggarisbawahi perlunya kerja sama global dalam mencapai solusi.

Prinsip

Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus mengambil langkah-langkah yang diperlukan untuk melindungi hak atas karya yang dihasilkan melalui penerapan hukum hak kekayaan intelektual yang berlaku secara tepat dan terarah. Pemerintah harus menyelidiki bagaimana karya yang diciptakan oleh AI dapat dilindungi lebih lanjut, tanpa berupaya menciptakan hak kekayaan intelektual baru pada tahap ini.

Mendukung Insentif dan Perlindungan untuk Inovasi

- Inovasi hanya bernilai jika dapat bermanfaat bagi masyarakat. Pendanaan diperlukan untuk mengembangkan inovasi hingga ke tingkat yang dapat disebarluaskan dan dimanfaatkan oleh masyarakat. Pihak yang dimintai pendanaan membutuhkan pengembalian atas investasi mereka. Akibatnya, harus ada insentif dan perlindungan untuk inovasi jika ingin menarik investasi dan membawa manfaat yang lebih besar bagi masyarakat.

- 
- Oleh karena itu, organisasi harus diizinkan untuk melindungi hak atas karya yang dihasilkan dari penggunaan AI, baik karya yang diciptakan oleh AI maupun karya yang difasilitasi oleh AI.
 - Namun, perlu diperhatikan agar tidak mengambil langkah-langkah yang dapat dianggap sebagai perlindungan berlebihan, karena hal ini dapat merugikan tujuan akhir perlindungan HKI.

Perlindungan Hak HKI

- ❖ Saat ini, undang-undang HKI belum memadai untuk menangani penciptaan karya oleh AI otonom.
- ❖ Organisasi yang mengembangkan, menerapkan, atau menggunakan sistem AI harus mengambil langkah-langkah yang diperlukan untuk melindungi hak-hak atas karya yang dihasilkan. Jika sesuai, langkah-langkah ini harus mencakup penegasan atau perolehan hak cipta, perolehan paten, jika berlaku, dan pencarian ketentuan kontraktual untuk memungkinkan perlindungan sebagai rahasia dagang dan/atau untuk mengalokasikan hak-hak tersebut secara tepat di antara para pihak.

Pengembangan Undang-Undang HKI Baru

- Pemerintah harus berhati-hati dalam merevisi undang-undang HKI yang ada.
- Pemerintah harus menjajaki pengenalan undang-undang yang sesuai (atau interpretasi undang-undang yang ada) untuk memperjelas perlindungan HKI atas karya yang didukung AI maupun yang diciptakan AI, tanpa berupaya menciptakan hak HKI baru pada tahap ini.
- Ketika mengubah undang-undang kekayaan intelektual yang ada atau menerapkan undang-undang kekayaan intelektual yang baru, pemerintah harus berupaya menyeimbangkan kepentingan semua pemangku kepentingan terkait secara memadai.
- Pemerintah juga harus menjajaki konsensus terkait hak AI dan kekayaan intelektual untuk memungkinkan arus data lintas batas yang lancar dan penyebaran teknologi baru yang cepat, serta berupaya mengatasi masalah ini melalui perjanjian internasional.

DAFTAR PUSTAKA

- Abrardi, L., Cambini, C., & Rondi, L. (2022). Artificial intelligence, firms and consumer behavior: A survey. *Journal of Economic Surveys*, 36(4), 969-991.
- Adams, C., Pente, P., Lemermeyer, G., & Rockwell, G. (2023). Ethical principles for artificial intelligence in K-12 education. *Comput. Educ. Artif. Intell.*, 4, 100131.
- Almeida, D., Shmarko, K., & Lomas, E. (2022). The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of US, EU, and UK regulatory frameworks. *AI and Ethics*, 2(3), 377-387.
- Askin, S., Burkhalter, D., Calado, G., & El Dakrouni, S. (2023). Artificial intelligence applied to clinical trials: opportunities and challenges. *Health and technology*, 13(2), 203-213
- Atabekov, A. (2023). Artificial intelligence in contemporary societies: Legal status and definition, implementation in public sector across various countries. *Social Sciences*, 12(3)
- Aziz, L. A. R., & Andriansyah, Y. (2023). The role artificial intelligence in modern banking: an exploration of AI-driven approaches for enhanced fraud prevention, risk management, and regulatory compliance. *Reviews of Contemporary Business Analytics*, 6(1), 110-132.
- Barrett, A., & Pack, A. (2023). Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 20(1), 59.
- Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: themes, knowledge gaps and future agendas. *Internet Research*, 33(7), 133-167.
- Čartolovni, A., Tomičić, A., & Mosler, E. L. (2022). Ethical, legal, and social considerations of AI-based medical decision-support tools: a scoping review. *International Journal of Medical Informatics*, 161, 104738.
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International journal of educational technology in higher education*, 20(1), 38.
- Chen, B., Wu, Z., & Zhao, R. (2023). From fiction to fact: the growing role of generative AI in business and finance. *Journal of Chinese Economic and Business Studies*, 21(4), 471-496.
- Chiappetta, A. (2023). Navigating the AI frontier: European parliamentary insights on bias and regulation, preceding the AI Act. *Internet Policy Review*, 12(4), 1-26.
- Chiu, T. K. (2021). A holistic approach to the design of artificial intelligence (AI) education for K-12 schools. *TechTrends*, 65(5), 796-807.
- Cupać, J., Schopmans, H., & Tuncer-Ebetürk, İ. (2024). Democratization in the age of artificial intelligence: introduction to the special issue. *Democratization*, 31(5), 899-921.

- De Almeida, P. G. R., Dos Santos, C. D., & Farias, J. S. (2021). Artificial intelligence regulation: a framework for governance. *Ethics and Information Technology*, 23(3), 505-525.
- Demaidi, M. N. (2025). Artificial intelligence national strategy in a developing country. *Ai & Society*, 40(2), 423-435.
- Eiras, F., Petrov, A., Vidgen, B., Schroeder, C., Pizzati, F., Elkins, K., ... & Foerster, J. (2024). Risks and opportunities of open-source generative AI. *arXiv preprint arXiv:2405.08597*.
- Erkkilä, T. (2023). Global indicators and AI policy: Metrics, policy scripts, and narratives. *Review of Policy Research*, 40(5), 811-839.
- Falco, G., Shneiderman, B., Badger, J., Carrier, R., Dahbura, A., Danks, D., ... & Yeong, Z. K. (2021). Governing AI safety through independent audits. *Nature Machine Intelligence*, 3(7), 566-571.
- Flores-Viva, J. M., & García-Peñalvo, F. J. (2023). Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4). *Comunicar: Media Education Research Journal*, 31(74), 35-44.
- Floridi, L. (2021). The European legislation on AI: a brief analysis of its philosophical approach. *Philosophy & Technology*, 34(2), 215-222.
- Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. *Machine learning and the city: Applications in architecture and urban design*, 535-545.
- Folorunso, A., Olanipekun, K., Adewumi, T., & Samuel, B. (2024). A policy framework on AI usage in developing countries and its impact. *Global Journal of Engineering and Technology Advances*, 21(01), 154-166.
- Holmes, W., & Miao, F. (2023). *Guidance for generative AI in education and research*. Unesco Publishing.
- Jackson, I., Ivanov, D., Dolgui, A., & Namdar, J. (2024). Generative artificial intelligence in supply chain and operations management: a capability-based framework for analysis and implementation. *International Journal of Production Research*, 62(17), 6120-6145.
- Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2022). Aligning artificial intelligence with climate change mitigation. *Nature Climate Change*, 12(6), 518-527.
- Kamila, M. K., & Jasrotia, S. S. (2025). Ethical issues in the development of artificial intelligence: recognizing the risks. *International Journal of Ethics and Systems*, 41(1), 45-63.
- Kulkov, I., Kulkova, J., Rohrbeck, R., Menvielle, L., Kaartemo, V., & Makkonen, H. (2024). Artificial intelligence-driven sustainable development: Examining organizational, technical, and processing approaches to achieving global goals. *Sustainable Development*, 32(3), 2253-2267.
- Laux, J., Wachter, S., & Mittelstadt, B. (2024). Trustworthy artificial intelligence and the European Union AI act: On the conflation of trustworthiness and acceptability of risk. *Regulation & Governance*, 18(1), 3-32.
- Madiega, T. (2021, April). *Artificial intelligence act*.

- Malik, A., Budhwar, P., & Kazmi, B. A. (2023). Artificial intelligence (AI)-assisted HRM: Towards an extended strategic framework. *Human Resource Management Review*, 33(1), 100940.
- McLarney, E., Gawdiak, Y., Oza, N., Mattmann, C., Garcia, M., Maskey, M., ... & Little, C. (2021). NASA framework for the ethical use of artificial intelligence (AI).
- Medaglia, R., Gil-Garcia, J. R., & Pardo, T. A. (2023). Artificial intelligence in government: Taking stock and moving forward. *Social Science Computer Review*, 41(1), 123-140.
- Mennella, C., Maniscalco, U., De Pietro, G., & Esposito, M. (2024). Ethical and regulatory challenges of AI technologies in healthcare: A narrative review. *Heliyon*, 10(4).
- Miazi, M. A. N. (2023). Interplay of legal frameworks and artificial intelligence (AI): A global perspective. *Law and Policy Review*, 2(2), 01-25.
- Murdoch, B. (2021). Privacy and artificial intelligence: challenges for protecting health information in a new era. *BMC medical ethics*, 22(1), 122.
- Nasim, S. F., Ali, M. R., & Kulsoom, U. (2022). Artificial intelligence incidents & ethics a narrative review. *International Journal of Technology Innovation and Management (IJTIM)*, 2(2), 52-64.
- Nemorin, S., Vlachidis, A., Ayerakwa, H. M., & Andriotis, P. (2023). AI hyped? A horizon scan of discourse on artificial intelligence in education (AIED) and development. *Learning, Media and Technology*, 48(1), 38-51.
- Neumann, O., Guirguis, K., & Steiner, R. (2024). Exploring artificial intelligence adoption in public organizations: a comparative case study. *Public Management Review*, 26(1), 114-141.
- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and information technologies*, 28(4), 4221-4241.
- Novelli, C., Hacker, P., Morley, J., Trondal, J., & Floridi, L. (2025). A robust governance for the AI act: AI office, AI Board, scientific panel, and national authorities. *European Journal of Risk Regulation*, 16(2), 566-590.
- Okunlaya, R. O., Syed Abdullah, N., & Alias, R. A. (2022). Artificial intelligence (AI) library services innovative conceptual framework for the digital transformation of university education. *Library Hi Tech*, 40(6), 1869-1892.
- Plata, S., De Guzman, M. A., & Quesada, A. (2023). Emerging research and policy themes on academic integrity in the age of chat GPT and generative AI. *Asian Journal of University Education (AJUE)*, 19(4), 743-758.
- Radu, R. (2021). Steering the governance of artificial intelligence: national strategies in perspective. *Policy and society*, 40(2), 178-193.
- Reddy, S., Rogers, W., Makinen, V. P., Coiera, E., Brown, P., Wenzel, M., ... & Kelly, B. (2021). Evaluation framework to guide implementation of AI systems into healthcare settings. *BMJ health & care informatics*, 28(1), e100444., 178..
- Slimi, Z., & Carballido, B. V. (2023). Navigating the Ethical Challenges of Artificial Intelligence in Higher Education: An Analysis of Seven Global AI Ethics Policies. *TEM journal*, 12(2).

- Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and society*, 40(2), 137-157.
- Van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. *AI and Ethics*, 1(3), 213-218.
- Xia, B., Lu, Q., Zhu, L., Lee, S. U., Liu, Y., & Xing, Z. (2024, April). Towards a responsible ai metrics catalogue: A collection of metrics for ai accountability. In *Proceedings of the IEEE/ACM 3rd International Conference on AI Engineering-Software Engineering for AI* (pp. 100-111).
- Xia, Q., Chiu, T. K., Lee, M., Sanusi, I. T., Dai, Y., & Chai, C. S. (2022). A self-determination theory (SDT) design approach for inclusive and diverse artificial intelligence (AI) education. *Computers & education*, 189, 104582.
- Yigitcanlar, T., Corchado, J. M., Mehmood, R., Li, R. Y. M., Mossberger, K., & Desouza, K. (2021). Responsible urban innovation with local government artificial intelligence (AI): A conceptual framework and research agenda. *Journal of Open Innovation: Technology, Market, and Complexity*, 7(1), 71.
- Yigitcanlar, T., David, A., Li, W., Fookes, C., Bibri, S. E., & Ye, X. (2024). Unlocking artificial intelligence adoption in local governments: Best practice lessons from real-world implementations. *Smart Cities*, 7(4), 1576-1625.

KERANGKA KEBIJAKAN GLOBAL PEMAKAIAN AI (ARTIFICIAL INTELLIGENCE)

Dr. Ir. Agus Wibowo, M.Kom, M.Si, MM

BIO DATA PENULIS



Penulis memiliki berbagai disiplin ilmu yang diperoleh dari Universitas Diponegoro (UNDIP) Semarang. dan dari Universitas Kristen Satya Wacana (UKSW) Salatiga. Disiplin ilmu itu antara lain teknik elektro, komputer, manajemen dan ilmu sosiologi. Penulis memiliki pengalaman kerja pada industri elektronik dan sertifikasi keahlian dalam bidang Jaringan Internet, Telekomunikasi, Artificial Intelligence, Internet Of Things (IoT), Augmented Reality (AR), Technopreneurship, Internet Marketing dan bidang pengolahan dan analisa data (komputer statistik).

Penulis adalah pendiri dari Universitas Sains dan Teknologi Komputer (Universitas STEKOM) dan juga seorang dosen yang memiliki Jabatan Fungsional Akademik Lektor Kepala (Associate Professor) yang telah menghasilkan puluhan Buku Ajar ber ISBN, HAKI dari beberapa karya cipta dan Hak Paten pada produk IPTEK. Sejak tahun 2023 penulis tercatat sebagai Dosen luar biasa di Fakultas Ekonomi & Bisnis (FEB) Universitas Diponegoro Semarang. Penulis juga terlibat dalam berbagai organisasi profesi dan industri yang terkait dengan dunia usaha dan industri, khususnya dalam pengembangan sumber daya manusia yang unggul untuk memenuhi kebutuhan dunia kerja secara nyata.



YAYASAN PRIMA AGUS TEKNIK

PENERBIT :

YAYASAN PRIMA AGUS TEKNIK
Jl. Majapahit No. 605 Semarang
Telp. (024) 6723456. Fax. 024-6710144
Email : penerbit_ypat@stekom.ac.id

ISBN 978-634-7227-36-2 (PDF)



9

786347

227362