



YOGYAKARTA PRIMA GUS TEKNIK

TEORI ETIKA DALAM KECERDASAN BUATAN (AI)

Dr. Budi Raharjo, S.Kom., M.Kom., MM.

Dr. Budi Raharjo, S.Kom., M.Kom., MM.

TEORI ETIKA DALAM KECERDASAN BUATAN (AI)



YAYASAN PRIMA AGUS TEKNIK

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

ISBN 978-623-8120-74-1 (PDF)



9 786238 120741

TEORI ETIKA DALAM KECERDASAN BUATAN (AI)

Penulis :

Dr. Budi Raharjo, S.Kom., M.Kom., MM.

ISBN : 9 786238 120741

Editor :

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

Penyunting :

Dr. Joseph Teguh Santoso, M.Kom.

Desain Sampul dan Tata Letak :

Irdha Yuniyanto, S.Ds., M.Kom

Penebit :

Yayasan Prima Agus Teknik

Redaksi :

Jl. Majapahit no 605 Semarang

Telp. (024) 6723456

Fax. 024-6710144

Email : penerbit_ypat@stekom.ac.id

Distributor Tunggal :

Universitas STEKOM

Jl. Majapahit no 605 Semarang

Telp. (024) 6723456

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 kepada Tuhan Yang Maha Esa atas terselesainya buku yang berjudul *“Teori Etika Dalam Kecerdasan Buatan (Ai)”*. Dalam buku ini membahas implikasi etika dan pertanyaan moral yang muncul dari pengembangan dan penerapan teknologi kecerdasan buatan (AI). Buku ini juga meninjau pedoman dan kerangka kerja yang dibuat oleh negara dan wilayah di seluruh dunia untuk mengatasi permasalahan tersebut. Buku ini menyajikan perbandingan antara kerangka kerja utama yang ada saat ini dan isu-isu etika utama, serta menyoroti kesenjangan dalam mekanisme pembagian manfaat yang adil; penugasan tanggung jawab; eksploitasi pekerja; kebutuhan energi dalam konteks perubahan lingkungan dan iklim; dan implikasi AI yang lebih kompleks dan kurang pasti, misalnya mengenai hubungan antarmanusia.

Buku ini membahas implikasi etika dan pertanyaan moral yang muncul dari pengembangan dan penerapan teknologi kecerdasan buatan (AI). Buku ini juga meninjau pedoman dan kerangka kerja yang telah dibuat oleh negara dan wilayah di seluruh dunia untuk mengatasinya. Buku ini menyajikan perbandingan antara kerangka kerja utama saat ini dan isu-isu etika utama, dan menyoroti kesenjangan seputar mekanisme pembagian manfaat yang adil; penugasan tanggung jawab; eksploitasi pekerja; kebutuhan energi dalam konteks perubahan lingkungan dan iklim; dan implikasi AI yang lebih kompleks dan kurang pasti, misalnya mengenai hubungan antarmanusia.

Bab 1 memperkenalkan ruang lingkup laporan dan menjelaskan istilah-istilah utama. Buku ini mengacu pada definisi Komisi Eropa mengenai AI sebagai 'sistem yang menampilkan perilaku cerdas'. Istilah-istilah penting lainnya yang didefinisikan dalam bab ini mencakup kecerdasan dan bagaimana hal ini digunakan dalam konteks AI dan robot cerdas (yaitu robot dengan AI yang tertanam), serta mendefinisikan pembelajaran mesin, jaringan syaraf tiruan, dan pembelajaran mendalam, sebelum melanjutkan ke pertimbangan definisi moralitas dan etika dan bagaimana kaitannya dengan AI.

Pada Bab 2, buku ini memetakan dilema etika utama dan pertanyaan moral terkait penerapan AI. Buku ini dimulai dengan menguraikan sejumlah manfaat potensial yang dapat diperoleh dari AI sebagai konteks untuk menempatkan pertimbangan etika, sosial, dan hukum. Selanjutnya membahas dampak AI terhadap psikologi manusia, sehingga menimbulkan pertanyaan tentang dampak AI terhadap hubungan, seperti dalam kasus robot cerdas yang mengambil peran sosial manusia, seperti keperawatan. Hubungan manusia-robot juga dapat mempengaruhi hubungan manusia-manusia dengan cara yang belum diantisipasi. Bagian ini juga membahas pertanyaan tentang kepribadian, dan apakah sistem AI harus memiliki hak pilihan moral.

Bab 3 mengeksplorasi inisiatif etis di bidang AI. Bab ini pertama-tama menguraikan inisiatif etis yang diidentifikasi dalam buku ini, merangkum fokusnya dan jika memungkinkan mengidentifikasi sumber pendanaan. Kerugian dan kekhawatiran yang diatasi oleh inisiatif-

inisiatif ini kemudian dibahas secara rinci. Isu-isu yang diangkat secara umum dapat diselaraskan dengan isu-isu yang diidentifikasi dalam Bab 2 dan dapat dibagi menjadi beberapa pertanyaan seputar: hak asasi manusia dan kesejahteraan; kerugian emosional; akuntabilitas dan tanggung jawab; keamanan, privasi, aksesibilitas dan transparansi; keamanan dan kepercayaan; kerugian sosial dan keadilan sosial; keabsahan dan keadilan; pengendalian dan penggunaan (atau penyalahgunaan) AI secara etis; kerusakan dan kelestarian lingkungan hidup; penggunaan informasi; risiko eksistensial.

Bab 4 membahas standar dan peraturan etika AI yang muncul. Ada sejumlah standar baru yang mengatasi dampak etika, hukum, dan sosial dari robotika dan AI. Mungkin yang paling awal adalah Panduan BS 8611 untuk Desain Etis dan Penerapan Robot dan Sistem Robot. Hal ini didasarkan pada 20 bahaya dan risiko etika yang berbeda, yang dikelompokkan dalam empat kategori: sosial, penerapan, komersial & keuangan, dan lingkungan. Standar ini mengakui bahaya fisik sebagai bahaya etika dan mengakui bahwa bahaya fisik dan emosional harus seimbang dengan manfaat yang diharapkan bagi pengguna.

Inisiatif kebijakan nasional dan internasional dibahas dalam Bab 5: Strategi Nasional dan Internasional mengenai AI. Strategi UE adalah inisiatif internasional pertama mengenai AI dan mendukung strategi masing-masing Negara Anggota. Namun strateginya berbeda-beda dalam hal sejauh mana strategi tersebut mengatasi permasalahan etika. Di tingkat Eropa, kekhawatiran masyarakat menjadi hal yang menonjol dalam inisiatif AI. Bab 6 merangkum tema-tema yang muncul dari literatur, inisiatif etika, dan strategi nasional dan internasional terkait AI, serta menyoroti kesenjangan yang ada. Bab terakhir dalam buku ini membahas tentang rangkuman yang tersaji dalam buku ini dan penelitian masa depan yang bisa dikembangkan mengenai perancangan robot etis. Akhir kata semoga buku ini berguna bagi para pembaca.

Semarang, November 2023

Penulis

Dr. Budi Raharjo, S.Kom., M.Kom., MM.

DAFTAR ISI

Halaman Judul	i
Kata Pengantar	ii
Daftar Isi	vi
BAB 1 PENGANTAR ARTIFICIAL INTELLIGENCE (AI)	1
1.1. Apa itu Kecerdasan Buatan	1
1.2. Definisi Moralitas Dan Etika, Dan Kaitannya Dengan AI	3
1.3. Struktur Buku	3
BAB 2 ETIKA PENERAPAN DAN PENGEMBANGAN ARTIFICIAL INTELLIGENCE	6
2.1. Dampak Terhadap Masyarakat	7
2.2. Dampak Terhadap Psikologi Manusia	23
2.3. Dampak Terhadap Sistem Keuangan	27
2.4. Dampak Terhadap Sistem Hukum	29
2.5. Dampak Terhadap Lingkungan Dan Planet Ini	36
2.6. Dampak Terhadap Kepercayaan	37
BAB 3 INISIATIF ETIS DI BIDANG KECERDASAN BUATAN	48
3.1. Inisiatif Etika Internasional	48
3.2. Kerugian Dan Kekhawatiran Etis Diatasi Dengan Inisiatif Ini	54
3.3. Studi Kasus	68
BAB 4 STANDAR DAN REGULASI ARTIFICIAL INTELLIGENCE	85
BAB 5 STRATEGI NASIONAL DAN INTERNASIONAL TENTANG AI	89
5.1. Eropa	89
5.2. Amerika Utara	94
5.3. Asia	95
5.4. Afrika	97
5.5. Amerika Selatan	97
5.6. Australia	98
5.7. Inisiatif AI Internasional Selain United Emirat	99
5.8. Kesiapan Pemerintah Terhadap AI	102
BAB 6 TEMA YANG MUNCUL	104
6.1. Mengatasi Masalah Etika Melalui Strategi Nasional Dan Internasional	104
6.2. Mengatasi Tantangan Tata Kelola Yang Ditimbulkan Oleh AI	106
BAB 7 PENUTUP	109
7.1. Membangun Robot Etis	111
Daftar Pustaka	114

BAB 1

PENGANTAR ARTIFICIAL INTELLIGENCE (AI)

Perkembangan pesat dalam kecerdasan buatan (AI) dan pembelajaran mesin membawa potensi manfaat yang sangat besar. Namun kita perlu mengeksplorasi seluruh aspek etika, sosial, dan hukum dari sistem AI jika kita ingin menghindari konsekuensi dan risiko negatif yang tidak diinginkan yang timbul dari penerapan AI di masyarakat.

Bab ini memperkenalkan AI secara luas, termasuk penggunaan dan definisi kecerdasan saat ini. Hal ini juga mendefinisikan robot dan posisinya dalam bidang AI yang lebih luas.

1.1. APA ITU KECERDASAN BUATAN?

Komunikasi Kecerdasan Buatan Komisi Eropa (European Commission, 2018a) mendefinisikan kecerdasan buatan sebagai berikut:

“Kecerdasan Buatan (AI) mengacu pada sistem yang menampilkan perilaku cerdas dengan menganalisis lingkungannya dan mengambil tindakan – dengan tingkat otonomi tertentu – untuk mencapai tujuan tertentu. Sistem berbasis AI dapat murni berbasis perangkat lunak, bertindak di dunia virtual (misalnya asisten suara, perangkat lunak analisis gambar, mesin pencari, sistem pengenalan suara dan wajah) atau AI dapat tertanam dalam perangkat keras (misalnya robot canggih, mobil otonom, drone atau aplikasi Internet of Things).”

Dalam buku ini, kami mempertimbangkan AI berbasis perangkat lunak dan robot cerdas (yaitu robot dengan AI yang tertanam) ketika mengeksplorasi masalah etika. Oleh karena itu, robot cerdas adalah bagian dari AI (baik mereka menggunakan pembelajaran mesin atau tidak).

Bagaimana kita mendefinisikan kecerdasan? Definisi yang jelas adalah bahwa perilaku cerdas adalah 'melakukan hal yang benar pada waktu yang tepat'. Legg dan Hunt (2007) mensurvei berbagai definisi informal tentang kecerdasan, dengan mengidentifikasi tiga ciri umum: bahwa kecerdasan adalah (1) 'suatu sifat yang dimiliki seorang agen ketika ia berinteraksi dengan lingkungan atau lingkungannya', (2) 'terkait pada kemampuan agen untuk berhasil atau memperoleh keuntungan sehubungan dengan beberapa tujuan atau sasaran', dan (3) 'tergantung pada seberapa mampu agen tersebut beradaptasi dengan berbagai tujuan dan lingkungan'. Mereka menunjukkan bahwa kecerdasan melibatkan adaptasi, pembelajaran dan pemahaman. Jadi, secara sederhana, kecerdasan adalah 'kemampuan untuk memperoleh dan menerapkan pengetahuan dan keterampilan serta memanipulasi lingkungannya'.

Dalam menafsirkan definisi kecerdasan ini, kita perlu memahami bahwa untuk robot fisik, lingkungannya adalah dunia nyata, yang dapat berupa lingkungan manusia (untuk robot sosial), jalan kota (untuk kendaraan otonom), rumah perawatan atau rumah sakit. (untuk robot hidup yang dirawat atau dibantu), atau tempat kerja (untuk robot rekan kerja). 'Lingkungan' AI perangkat lunak adalah konteksnya, yang mungkin bersifat klinis (untuk AI

diagnosis medis), atau ruang publik – misalnya untuk pengenalan wajah di bandara, atau virtual untuk pengenalan wajah di media sosial. Namun, seperti robot fisik, AI perangkat lunak hampir selalu berinteraksi dengan manusia, baik melalui antarmuka tanya jawab: melalui teks untuk chatbot, atau melalui ucapan untuk asisten digital di ponsel (yaitu Siri) atau di rumah (yaitu Alexa).

Interaksi dengan manusia inilah yang menimbulkan hampir seluruh permasalahan etika yang disurvei dalam buku ini. Semua AI dan robot saat ini adalah contoh dari apa yang kami sebut sebagai AI: sebuah istilah yang mencerminkan fakta bahwa AI dan robot saat ini biasanya hanya mampu melakukan satu tugas khusus. Tujuan jangka panjang penelitian AI dan robotika adalah apa yang disebut kecerdasan umum buatan (AGI) yang sebanding dengan kecerdasan manusia. Penting untuk dipahami bahwa AI sempit saat ini sering kali lebih baik daripada kebanyakan manusia dalam satu tugas tertentu. Contohnya adalah AI catur atau Go-playing, mesin pencari, atau sistem terjemahan bahasa alami. Namun robot perawatan serba guna yang mampu, misalnya, menyiapkan makanan untuk orang lanjut usia (dan mencuci piring setelahnya), membantu mereka berpakaian atau menanggalkan pakaian, naik dan turun dari tempat tidur atau mandi, dll., masih belum menjadi tujuan penelitian.

Pembelajaran mesin adalah istilah yang digunakan untuk AI yang mampu belajar atau, dalam kasus robot, beradaptasi dengan lingkungannya. Ada berbagai macam pendekatan terhadap pembelajaran mesin, namun pendekatan ini biasanya terbagi dalam dua kategori: pembelajaran yang diawasi dan tanpa pengawasan. Sistem pembelajaran yang diawasi umumnya menggunakan Jaringan Syaraf Tiruan (JST), yang dilatih dengan menampilkan masukan (misalnya, gambar binatang) kepada ANN yang masing-masing ditandai (oleh manusia) dengan keluaran (misalnya jerapah, singa, gorila). Kumpulan masukan dan keluaran yang cocok ini disebut kumpulan data pelatihan. Setelah pelatihan, ANN harus dapat mengidentifikasi hewan mana yang ada dalam gambar yang disajikan (yaitu singa), meskipun gambar singa tersebut tidak ada dalam kumpulan data pelatihan. Sebaliknya, pembelajaran tanpa pengawasan tidak memiliki data pelatihan; sebaliknya, AI (atau robot) harus memikirkan sendiri cara menyelesaikan tugas tertentu (yaitu cara menavigasi keluar dari labirin dengan sukses), umumnya melalui trial and error.

Pembelajaran yang diawasi dan tidak diawasi memiliki keterbatasannya masing-masing. Dengan pembelajaran yang diawasi, kumpulan data pelatihan harus benar-benar mewakili tugas yang diperlukan; jika tidak, AI akan menunjukkan bias. Keterbatasan lainnya adalah ANN belajar dengan memilih fitur-fitur gambar dalam data pelatihan yang tidak diantisipasi oleh perancang manusia. Jadi, misalnya, mereka mungkin salah mengidentifikasi mobil dengan latar belakang bersalju sebagai serigala, karena semua contoh serigala dalam gambar kumpulan data pelatihan memiliki latar belakang bersalju, dan ANN telah belajar mengidentifikasi latar belakang bersalju sebagai serigala, bukan serigala itu sendiri. Pembelajaran tanpa pengawasan umumnya lebih kuat daripada pembelajaran yang diawasi, tetapi memiliki keterbatasan yaitu pembelajaran tersebut umumnya sangat lambat (dibandingkan dengan manusia yang sering kali dapat belajar hanya dari satu percobaan).

Istilah pembelajaran mendalam hanya mengacu pada (biasanya) sistem pembelajaran mesin yang diawasi dengan ANN yang besar (yaitu berlapis-lapis) dan kumpulan data pelatihan yang besar.

Penting untuk diingat bahwa istilah AI dan pembelajaran mesin tidak sama. Banyak AI dan robot berkemampuan tinggi tidak menggunakan pembelajaran mesin.

1.2. DEFINISI MORALITAS DAN ETIKA, DAN KAITANNYA DENGAN AI

Etika adalah prinsip-prinsip moral yang mengatur tingkah laku seseorang atau pelaksanaan suatu kegiatan. Sebagai contoh praktis, salah satu prinsip etika adalah memperlakukan semua orang dengan hormat. Para filsuf telah memperdebatkan etika selama berabad-abad, dan terdapat berbagai prinsip yang terkenal, mungkin salah satu yang paling terkenal adalah imperatif kategoris Kant 'bertindaklah sebagaimana Anda ingin semua orang bertindak terhadap semua orang'.

Etika AI berkaitan dengan pertanyaan penting tentang bagaimana manusia pengembang, produsen, dan operator harus berperilaku untuk meminimalkan kerugian etis yang dapat timbul dari AI di masyarakat, baik yang timbul dari desain yang buruk (tidak etis), penerapan yang tidak tepat, atau penyalahgunaan. Cakupan etika AI mencakup kekhawatiran yang ada saat ini, misalnya, tentang privasi data dan bias dalam sistem AI saat ini; kekhawatiran jangka pendek dan menengah mengenai, misalnya, dampak AI dan robotika terhadap pekerjaan dan tempat kerja; dan kekhawatiran jangka panjang mengenai kemungkinan sistem AI mencapai atau melampaui kemampuan yang setara dengan manusia (yang disebut *superintelligence*).

Dalam 5 tahun terakhir, etika AI telah berubah dari perhatian akademis menjadi perdebatan politik dan publik. Meningkatnya keberadaan ponsel pintar dan aplikasi berbasis AI yang kini diandalkan oleh banyak dari kita setiap hari, fakta bahwa AI semakin berdampak pada semua sektor (termasuk industri, layanan kesehatan, kepolisian & peradilan, transportasi, keuangan, dan rekreasi), serta prospek 'perlombaan senjata' AI, telah mendorong banyak sekali inisiatif nasional dan internasional, mulai dari LSM, kelompok akademis dan industri, badan profesional, dan pemerintah. Inisiatif ini telah menghasilkan publikasi sejumlah besar prinsip etika untuk robotika dan AI (setidaknya 22 prinsip etika yang berbeda telah diterbitkan sejak Januari 2017), standar etika baru bermunculan (terutama dari British Standards Institute dan Asosiasi Standar IEEE), dan semakin banyak negara (dan kelompok negara) yang mengumumkan strategi AI (dengan investasi skala besar) dan membentuk badan penasehat atau kebijakan nasional. Dalam buku ini kami mensurvei inisiatif-inisiatif tersebut untuk mengungkap isu-isu etika utama dalam AI dan robotika.

1.3. STRUKTUR BUKU

Robot dan kecerdasan buatan (AI) hadir dalam berbagai bentuk, sebagaimana diuraikan di atas, yang masing-masing menimbulkan permasalahan etika yang berbeda-beda. Hal ini diuraikan dalam Bab 2: Memetakan dilema etika utama dan pertanyaan moral terkait penerapan AI. Bab ini membahas secara khusus:

Dampak sosial: bagian ini membahas potensi dampak AI terhadap pasar tenaga kerja dan perekonomian serta bagaimana kelompok demografis yang berbeda mungkin terkena dampaknya. Hal ini menjawab pertanyaan mengenai kesenjangan dan risiko bahwa AI akan semakin memusatkan kekuasaan dan kekayaan di tangan segelintir orang. Permasalahan yang berkaitan dengan privasi, hak asasi manusia, dan martabat ditangani begitu pula dengan risiko bahwa AI akan melanggengkan bias, baik disengaja maupun tidak, pada sistem sosial yang ada atau penciptanya. Bagian ini juga menimbulkan pertanyaan tentang dampak teknologi AI terhadap demokrasi, yang menunjukkan bahwa teknologi ini mungkin bermanfaat bagi perekonomian yang dikendalikan oleh negara.

Dampak psikologis: dampak apa saja yang mungkin timbul dari hubungan manusia-robot? Bagaimana kita mengatasi ketergantungan dan penipuan? Haruskah kita mempertimbangkan apakah robot layak diberi status 'kepribadian' dan apa implikasi hukum dan moral dari tindakan tersebut?

Dampak sistem keuangan: potensi dampak AI pada sistem keuangan dipertimbangkan, termasuk risiko manipulasi dan kolusi serta kebutuhan untuk membangun akuntabilitas.

Dampak terhadap sistem hukum: terdapat beberapa cara AI dapat mempengaruhi sistem hukum, termasuk: pertanyaan terkait kejahatan, seperti tanggung jawab jika AI digunakan untuk aktivitas kriminal, dan sejauh mana AI dapat mendukung aktivitas kriminal seperti perdagangan narkoba. Dalam situasi di mana AI terlibat dalam cedera pribadi, seperti tabrakan yang melibatkan kendaraan otonom, maka timbul pertanyaan seputar pendekatan hukum terhadap klaim (apakah itu merupakan kasus kelalaian, yang biasanya menjadi dasar klaim yang melibatkan kecelakaan kendaraan, atau tanggung jawab produk).

Dampak lingkungan: peningkatan penggunaan AI disertai dengan peningkatan penggunaan sumber daya alam, peningkatan kebutuhan energi, dan masalah pembuangan limbah. Namun, AI dapat meningkatkan cara kita mengelola limbah dan sumber daya, sehingga memberikan manfaat bagi lingkungan.

Dampak terhadap kepercayaan: masyarakat bergantung pada kepercayaan. Agar AI dapat melakukan tugas-tugas, seperti pembedahan, masyarakat harus mempercayai teknologinya. Kepercayaan mencakup aspek-aspek seperti keadilan (bahwa AI tidak memihak), transparansi (bahwa kita akan dapat memahami bagaimana AI sampai pada keputusan tertentu), akuntabilitas (seseorang dapat dimintai pertanggungjawaban atas kesalahan yang dibuat oleh AI) dan kontrol (bagaimana kita mungkin 'mematikan' AI yang menjadi terlalu kuat).

Pada Bab 3, Inisiatif etis di bidang kecerdasan buatan, buku ini mengulas berbagai inisiatif etis yang bermunculan sebagai respons terhadap kekhawatiran dan permasalahan etika yang muncul sehubungan dengan AI. Bagian 3.1 membahas isu-isu yang dieksplorasi oleh setiap inisiatif dan mengidentifikasi laporan yang tersedia (per Mei 2019). Bahaya dan permasalahan etika yang diatasi melalui inisiatif yang diuraikan di atas, dibahas di Bagian 3.2. Hal ini secara

garis besar dibagi menjadi 12 kategori: hak asasi manusia dan kesejahteraan; kerugian emosional; akuntabilitas dan tanggung jawab; keamanan, privasi, aksesibilitas, dan transparansi; keamanan dan kepercayaan; kerugian sosial dan keadilan sosial; kerugian finansial; keabsahan dan keadilan; pengendalian dan penggunaan (atau penyalahgunaan) AI secara etis; kerusakan dan kelestarian lingkungan hidup; penggunaan informasi dan risiko eksistensial. Bab ini mengeksplorasi masing-masing topik tersebut dan cara-cara untuk mengatasinya melalui inisiatif-inisiatif tersebut.

Bab 4 menyajikan status standar dan peraturan Etika AI saat ini. Saat ini hanya satu standar (British Standard BS8611, Panduan desain etis robot dan sistem robot) yang secara khusus membahas AI. Namun, IEEE sedang mengembangkan sejumlah standar yang mempengaruhi AI dalam berbagai konteks. Meskipun hal-hal tersebut masih dalam pengembangan, hal-hal tersebut disajikan di sini sebagai indikasi kemajuan standar dan peraturan.

Bab 5 mengeksplorasi strategi nasional dan internasional mengenai AI. Bab ini membahas apa yang diperlukan untuk AI yang dapat dipercaya dan visi masa depan AI yang diartikulasikan dalam strategi nasional dan internasional.

Bab 6 menjelaskan Tinjauan kami terhadap literatur mengenai masalah etika seputar AI dan robot cerdas menyoroti berbagai potensi dampak, termasuk dalam bidang sosial, psikologis, keuangan, hukum, dan lingkungan.

Bab 7 menjadi akhir bab dari buku ini, akan merangkum kajian buku ini serta memberikan pandangan masa depan tentang pengembangan robot etis.

BAB 2

ETIKA PADA PENERAPAN DAN PENGEMBANGAN ARTIFICIAL INTELLIGENCE

Menurut Future of Life Institute (n.d.), AI 'memiliki potensi ekonomi, sosial, medis, keamanan, dan lingkungan yang besar', dengan potensi manfaat yang mencakup:

- Membantu masyarakat memperoleh keterampilan dan pelatihan baru;
- Demokratisasi layanan;
- Merancang dan memberikan waktu produksi yang lebih cepat dan siklus iterasi yang lebih cepat;
- Mengurangi penggunaan energi;
- Menyediakan pemantauan lingkungan secara real-time untuk polusi dan kualitas udara;
- Meningkatkan pertahanan keamanan siber;
- Meningkatkan output nasional;
- Mengurangi inefisiensi layanan kesehatan;
- Menciptakan pengalaman dan interaksi baru yang menyenangkan bagi masyarakat; Dan
- Meningkatkan layanan terjemahan real-time untuk menghubungkan orang-orang di seluruh dunia.



Gambar 2.1: Masalah etika dan moral utama yang terkait dengan pengembangan dan penerapan AI

Dalam jangka panjang, AI dapat menghasilkan 'terobosan' di berbagai bidang, kata Institut tersebut, mulai dari ilmu dasar dan terapan hingga kedokteran dan sistem lanjutan. Namun, selain menjanjikan, sistem cerdas yang semakin mampu menciptakan tantangan etika yang signifikan (Winfield, 2019a). Buku ini merangkum pertimbangan etika, sosial, dan hukum utama dalam penerapan AI, dengan mengambil wawasan dari literatur akademis yang relevan. Permasalahan yang dibahas berkaitan dengan dampak terhadap: masyarakat manusia; psikologi manusia; sistem keuangan; sistem hukum; lingkungan hidup dan planet ini; dan berdampak pada kepercayaan.

2.1 DAMPAK TERHADAP MASYARAKAT

Pasar tenaga kerja

Masyarakat telah mengkhawatirkan perpindahan pekerja karena teknologi selama berabad-abad. Otomasi, kemudian mekanisasi, komputasi, dan yang terbaru adalah AI dan robotika telah diperkirakan akan menghancurkan lapangan kerja dan menciptakan kerusakan permanen pada pasar tenaga kerja. Leontief (1983), mengamati peningkatan dramatis dalam kekuatan pemrosesan chip komputer, khawatir bahwa manusia akan digantikan oleh mesin, seperti halnya kuda menjadi usang karena penemuan mesin pembakaran internal. Namun, di masa lalu, otomatisasi sering kali menggantikan tenaga manusia dalam jangka pendek, namun telah menciptakan lapangan kerja dalam jangka panjang (Autor, 2015).

Namun demikian, terdapat kekhawatiran luas bahwa kecerdasan buatan dan teknologi terkait dapat menciptakan pengangguran massal dalam dua dekade mendatang. Sebuah makalah baru-baru ini menyimpulkan bahwa teknologi informasi baru akan menempatkan 'sebagian besar lapangan kerja, di berbagai jenis pekerjaan, dalam risiko dalam waktu dekat' (Frey dan Osborne, 2013).

AI sudah tersebar luas di bidang keuangan, eksplorasi ruang angkasa, manufaktur maju, transportasi, pengembangan energi, dan perawatan kesehatan. Kendaraan tak berawak dan drone otonom juga menjalankan fungsi yang sebelumnya memerlukan campur tangan manusia. Kita telah melihat dampak otomatisasi terhadap pekerjaan 'kerah biru'; namun, seiring dengan semakin canggihnya komputer, kreatif, dan serba guna, semakin banyak pekerjaan yang akan terpengaruh oleh teknologi dan semakin banyak posisi yang menjadi ketinggalan jaman.

Dampak terhadap pertumbuhan ekonomi dan produktivitas

Para ekonom umumnya antusias dengan prospek AI terhadap pertumbuhan ekonomi. Robotika menambah sekitar 0,4 poin persentase pertumbuhan PDB tahunan dan produktivitas tenaga kerja di 17 negara antara tahun 1993 dan 2007, yang setara dengan dampak pengenalan mesin uap terhadap pertumbuhan di Inggris (Graetz dan Michaels, 2015)

Dampak terhadap tenaga kerja

Sulit untuk mengukur dampak robot, AI, dan sensor terhadap angkatan kerja karena kita berada pada tahap awal revolusi teknologi. Para ekonom juga tidak sepakat mengenai dampak relatif AI dan robotika. Sebuah penelitian menanyakan 1.896 ahli tentang dampak teknologi baru; 48 persen percaya bahwa robot dan agen digital akan menggantikan sejumlah

besar pekerja kerah biru dan putih, dan banyak yang menyatakan kekhawatiran bahwa hal ini akan menyebabkan peningkatan besar dalam kesenjangan pendapatan, sejumlah besar pengangguran, dan kerusakan sistem sosial. ketertiban (Smith dan Anderson, 2014). Namun, separuh ahli lainnya yang menanggapi survei ini (52%) memperkirakan bahwa teknologi tidak akan menggantikan lebih banyak pekerjaan dibandingkan yang diciptakan pada tahun 2025. Para ahli tersebut percaya bahwa meskipun banyak pekerjaan yang saat ini dilakukan oleh manusia akan diambil alih secara substansial oleh robot atau robot. agen digital, mereka yakin bahwa kecerdasan manusia akan menciptakan lapangan kerja, industri, dan cara baru untuk mencari nafkah.

Beberapa orang berpendapat bahwa teknologi telah menghasilkan perubahan besar dalam angkatan kerja:

'Kemajuan teknologi akan meninggalkan sebagian orang, bahkan mungkin banyak orang, seiring dengan pesatnya kemajuan teknologi... Saat ini adalah waktu yang paling tepat untuk menjadi pekerja dengan keterampilan khusus atau pendidikan yang tepat karena orang-orang ini dapat menggunakan teknologi untuk menciptakan dan menangkap nilai. Namun, saat ini adalah saat yang paling buruk untuk menjadi pekerja yang hanya memiliki keterampilan dan kemampuan 'biasa', karena komputer, robot, dan teknologi digital lainnya memperoleh keterampilan dan kemampuan ini dengan kecepatan yang luar biasa' (Brynjolfsson dan McAfee, 2014).

Ford (2009) mengeluarkan peringatan yang sama kuatnya, dan berpendapat bahwa:

'seiring dengan percepatan teknologi, otomasi mesin pada akhirnya dapat menembus perekonomian hingga tingkat upah tidak lagi memberikan pendapatan yang memadai dan kepercayaan diri bagi sebagian besar konsumen di masa depan. Jika masalah ini tidak diatasi, dampaknya akan menyebabkan penurunan perekonomian. Ia memperingatkan bahwa 'pada titik tertentu di masa depan – mungkin bertahun-tahun atau beberapa dekade dari sekarang – mesin akan mampu melakukan pekerjaan yang dilakukan oleh sebagian besar orang-orang 'rata-rata' dalam populasi kita, dan orang-orang ini tidak akan mampu melakukan pekerjaan tersebut. untuk mencari pekerjaan baru'.

Namun, beberapa ekonom membantah klaim ini, dengan mengatakan bahwa meskipun banyak pekerjaan akan hilang karena kemajuan teknologi, lapangan kerja baru akan tercipta. Menurut orang-orang ini, keuntungan dan kerugian pekerjaan akan seimbang dalam jangka panjang.

“Mungkin lebih sedikit orang yang menyortir barang di gudang karena mesin dapat melakukannya lebih baik dibandingkan manusia. Namun pekerjaan menganalisis data besar, menambang informasi, dan mengelola jaringan berbagi data akan tercipta”. (West, 2018).

Jika AI mendorong pertumbuhan ekonomi, hal ini dapat menciptakan permintaan akan lapangan kerja di seluruh perekonomian, termasuk hal-hal yang tidak terkait langsung dengan teknologi. Misalnya, jumlah pekerja di sektor rekreasi dan perhotelan dapat meningkat jika pendapatan rumah tangga meningkat, sehingga memungkinkan masyarakat untuk membeli lebih banyak makanan di luar dan melakukan perjalanan (Furman dan Seamans, 2018).

Meski begitu, jelas bahwa sejumlah sektor akan terkena dampaknya. Frey dan Osborne (2013) menghitung bahwa ada kemungkinan besar bahwa 47 persen pekerja di AS akan melihat pekerjaan mereka terotomatisasi dalam 20 tahun ke depan. Menurut analisis mereka, telemarketer, pemeriksa hak milik, penjahit tangan, teknisi matematika, penjamin emisi asuransi, tukang reparasi jam tangan, agen kargo, ahli pajak, pekerja proses fotografi, pegawai rekening baru, teknisi perpustakaan, dan spesialis entri data memiliki peluang 99 persen untuk terkena penyakit ini. pekerjaan mereka terkomputerisasi. Di sisi lain, terapis rekreasi, supervisor mekanik, direktur manajemen darurat, pekerja sosial kesehatan mental, audiolog, terapis okupasi, pekerja sosial layanan kesehatan, ahli bedah mulut, supervisor pemadam kebakaran, dan ahli gizi memiliki peluang kurang dari satu persen untuk mengalami hal ini.

Dalam studi lebih lanjut, tim tersebut mensurvei 156 pakar akademis dan industri di bidang pembelajaran mesin, robotika, dan sistem cerdas, dan menanyakan tugas apa yang mereka yakini saat ini dapat diotomatisasi (Duckworth dkk., 2019). Mereka menemukan bahwa pekerjaan yang bersifat klerikal, berulang, tepat, dan perseptual semakin dapat diotomatisasi, sedangkan pekerjaan yang lebih kreatif, dinamis, dan berorientasi pada manusia cenderung kurang 'dapat diotomatisasi'.

Yang mengkhawatirkan, delapan kali lebih banyak pekerjaan berada di antara 'sebagian besar' dan 'sepenuhnya' dapat diotomatisasi dibandingkan antara 'sebagian besar tidak' dan 'tidak sama sekali' dapat diotomatisasi, ketika ditimbang berdasarkan jumlah pekerjaan. Aktivitas yang diklasifikasikan sebagai 'penalaran dan pengambilan keputusan' dan 'koordinasi, pengembangan, pengelolaan, dan pemberian nasihat' cenderung lebih kecil kemungkinannya untuk dapat diotomatisasi, sedangkan 'pengelolaan', 'pemrosesan informasi dan data', dan 'melakukan aktivitas yang kompleks dan teknis' kemungkinan besar dapat dilakukan. menjadi lebih dari itu.

Secara keseluruhan model tersebut memperkirakan potensi otomasi yang sangat tinggi pada pekerjaan perkantoran, dukungan administratif, dan penjualan, yang secara keseluruhan mempekerjakan sekitar 38 juta orang di AS. Yang juga berisiko tinggi terhadap otomasi adalah proses fisik seperti produksi, pertanian, perikanan dan kehutanan, serta transportasi dan perpindahan material, yang mempekerjakan sekitar 20 juta orang secara total. Sebaliknya, pekerjaan yang tahan terhadap otomasi mencakup pekerjaan di bidang pendidikan, hukum, pelayanan masyarakat, seni, dan media, dan pada tingkat yang lebih rendah, pekerjaan di bidang manajemen, bisnis, dan keuangan.

Dalam buku ini mengungkapkan bahwa pekerjaan dengan gaji dan tingkat pendidikan tertinggi cenderung paling tidak dapat menerima otomasi. Namun, hal ini pun tidak menjamin bahwa aktivitas suatu pekerjaan tidak dapat diotomatisasi. Seperti yang penulis tunjukkan, pengawas lalu lintas udara memperoleh penghasilan sekitar Rp. 125.000.000 per

tahun, namun diperkirakan sebagian besar tugas mereka dapat diotomatisasi. Sebaliknya, guru prasekolah dan asisten pengajar berpenghasilan di bawah Rp. 30.000.000 per tahun, namun peran mereka dianggap tidak dapat diotomatisasi.

Dampaknya terhadap demografi yang berbeda

Dampak dari perubahan yang cukup besar ini tidak akan dirasakan secara merata oleh seluruh anggota masyarakat. Demografi yang berbeda akan terkena dampak yang berbeda-beda, dan beberapa di antaranya lebih berisiko dibandingkan yang lain akibat teknologi baru. Mereka yang memiliki sedikit keterampilan teknis atau keahlian khusus akan menghadapi kesulitan paling besar. Generasi muda yang memasuki pasar tenaga kerja juga akan terkena dampak yang sangat besar, karena mereka masih dalam tahap awal karir dan akan menjadi generasi pertama yang bekerja bersama AI. Meskipun banyak generasi muda mempunyai waktu untuk memperoleh keahlian yang relevan, hanya sedikit yang mendapatkan pelatihan di bidang sains, teknologi, teknik, dan matematika (STEM), sehingga membatasi kemampuan mereka untuk menghadapi perubahan pekerjaan. Menurut Departemen Pendidikan AS (2014), akan ada peningkatan 14 persen dalam pekerjaan STEM antara tahun 2010 dan 2020 — namun 'hanya 16 persen siswa sekolah menengah atas yang mahir dalam matematika dan tertarik pada karir STEM'.

Perempuan juga mungkin terkena dampak yang tidak proporsional, karena semakin banyak perempuan yang bekerja di posisi pengasuh, salah satu sektor yang mungkin terkena dampak robot. Karena diskriminasi, prasangka dan kurangnya pelatihan, kelompok minoritas dan miskin sudah menderita tingkat pengangguran yang tinggi: tanpa pelatihan keterampilan tinggi, akan lebih sulit bagi mereka untuk beradaptasi dengan perekonomian baru. Banyak dari orang-orang ini juga tidak memiliki akses terhadap internet berkecepatan tinggi, sehingga membatasi kemampuan mereka untuk mengakses pendidikan, pelatihan dan pekerjaan.

Survei khusus Eurobarometer 460 mengidentifikasi bahwa sebagian besar penduduk UE memiliki respons positif terhadap peningkatan penggunaan teknologi digital, mengingat hal itu dapat meningkatkan masyarakat, perekonomian, dan kualitas hidup mereka, dan sebagian besar juga menganggap diri mereka cukup kompeten untuk memanfaatkan teknologi ini. dalam berbagai aspek kehidupan dan pekerjaan mereka. Namun, yang terpenting, sikap ini bervariasi berdasarkan usia, lokasi, dan latar belakang pendidikan — sebuah temuan yang penting dalam isu bagaimana AI akan memengaruhi demografi yang berbeda dan potensi permasalahan yang timbul seputar 'kesenjangan digital'.

Misalnya, laki-laki muda dengan tingkat pendidikan tinggi kemungkinan besar memiliki pandangan positif terhadap digitalisasi dan penggunaan robot — dan juga kemungkinan besar telah mengambil tindakan perlindungan terkait privasi dan keamanan online mereka (sehingga menempatkan mereka berisiko lebih rendah di bidang ini). Pola sosio-demografis seperti ini menyoroti area utama yang menjadi perhatian dalam peningkatan pengembangan dan penerapan AI agar tidak ada orang yang dirugikan atau tertinggal.

Konsekuensi

'Ketika kita berbicara tentang 'AI untuk kebaikan', kita perlu mendefinisikan apa yang dimaksud dengan 'kebaikan'. Saat ini, indikator kinerja utama yang kami gunakan didasarkan

pada PDB. Bukan berarti hal ini buruk, tapi ini tentang mengukur produktivitas dan keuntungan eksponensial.'

Ada kemungkinan bahwa teknologi AI dan robot dapat memperburuk kesenjangan sosial dan ekonomi yang ada, dengan menempatkan kelas pekerjaan saat ini dalam risiko, menghilangkan lapangan kerja, dan menyebabkan pengangguran massal di sektor pekerjaan yang dapat diotomatisasi. Diskriminasi juga bisa menjadi sebuah masalah, dimana generasi muda berpotensi terkena dampak yang tidak proporsional, selain mereka yang tidak memiliki pelatihan keterampilan tinggi.

Ketidaksamaan

Pertanyaan terbesar seputar AI adalah kesenjangan, yang biasanya tidak disertakan dalam perdebatan mengenai etika AI. Ini adalah masalah etika, namun sebagian besar merupakan masalah politik – siapa yang diuntungkan dari AI?'

Teknologi AI dan robotika diharapkan dapat memungkinkan perusahaan untuk mengefektifkan bisnisnya, menjadikannya lebih efisien dan produktif. Namun, ada yang berpendapat bahwa hal ini akan mengorbankan tenaga kerja manusianya. Hal ini tentu berarti bahwa pendapatan akan terbagi ke lebih sedikit orang, sehingga meningkatkan kesenjangan sosial. Akibatnya, individu yang memiliki kepemilikan di perusahaan yang digerakkan oleh AI akan mendapatkan keuntungan yang tidak proporsional.

Ketimpangan: eksploitasi pekerja

Perubahan lapangan kerja yang terkait dengan otomatisasi dan digitalisasi tidak hanya terlihat melalui hilangnya lapangan kerja, karena AI diperkirakan akan menciptakan banyak bentuk lapangan kerja baru, namun juga dalam hal kualitas pekerjaan. Winfield (2019b) menyatakan bahwa pekerjaan baru mungkin membutuhkan pekerja berketerampilan tinggi namun bersifat repetitif dan membosankan, sehingga menciptakan 'pekerjaan kerah putih' yang dipenuhi pekerja yang melakukan tugas-tugas seperti menandai dan memoderasi konten – dengan cara ini, AI dapat menimbulkan biaya tambahan bagi sumber daya manusia yang harus dipertimbangkan ketika mengkarakterisasi manfaat AI bagi masyarakat.

Membangun AI paling sering mengharuskan orang untuk mengelola dan membersihkan data untuk menginstruksikan algoritme pelatihan. AI yang lebih baik (dan lebih aman) memerlukan kumpulan data pelatihan yang besar dan industri outsourcing baru bermunculan di seluruh dunia untuk memenuhi kebutuhan ini. Hal ini telah menciptakan beberapa kategori pekerjaan baru.

Hal ini mencakup: (i) memindai dan mengidentifikasi konten yang menyinggung untuk dihapus, (ii) menandai objek dalam gambar secara manual untuk membuat kumpulan data pelatihan untuk sistem pembelajaran mesin (misalnya, untuk menghasilkan kumpulan data pelatihan untuk AI mobil tanpa pengemudi) dan (iii) menafsirkan pertanyaan (teks atau ucapan) yang tidak dapat dipahami oleh chatbot AI. Secara kolektif pekerjaan-pekerjaan ini kadang-kadang dikenal dengan istilah 'turk mekanis' (dinamai demikian setelah robot pemain catur abad ke-18 yang diketahui dioperasikan oleh seorang master catur manusia yang tersembunyi di dalam kabinet).

Ketika pertama kali diluncurkan, tugas-tugas semacam ini ditawarkan sebagai cara bagi orang-orang untuk mendapatkan uang tambahan di waktu luang mereka, namun Gray dan Suri (2019) menyatakan bahwa 20 juta orang kini dipekerjakan di seluruh dunia, melalui kontraktor pihak ketiga, dalam sebuah pekerjaan on-demand. ekonomi', bekerja di luar perlindungan undang-undang ketenagakerjaan. Pekerjaan biasanya dijadwalkan, diarahkan, disampaikan dan dibayar secara online, melalui antarmuka pemrograman aplikasi (API). Ada beberapa investigasi jurnalistik terhadap para pekerja di bidang pekerjaan ini – yang disebut sebagai 'pekerjaan hantu' oleh peneliti Harvard Mary L. Gray karena sifat 'tersembunyi' dari rantai nilai yang menyediakan kekuatan pemrosesan yang menjadi dasar AI.

Rata-rata konsumen teknologi AI mungkin tidak pernah tahu bahwa seseorang adalah bagian dari proses tersebut – rantai nilainya tidak jelas. Salah satu masalah etika utama adalah – mengingat harga produk akhir – para pekerja sementara ini mendapat imbalan yang tidak adil untuk pekerjaan yang penting agar teknologi AI dapat berfungsi. Hal ini terutama terjadi ketika angkatan kerja berada di negara-negara di luar UE atau AS – misalnya, terdapat industri 'pelabelan data' yang sedang berkembang di Tiongkok dan Kenya. Masalah lainnya adalah pekerja diharuskan menonton dan memeriksa konten ofensif di platform media seperti Facebook dan YouTube. Konten tersebut dapat mencakup perkataan yang mendorong kebencian, pornografi yang mengandung kekerasan, kekejaman, dan terkadang pembunuhan terhadap hewan dan manusia. Sebuah laporan berita menguraikan masalah kesehatan mental (gejala trauma seperti PTSD, serangan panik, dan kelelahan), serta kondisi kerja yang buruk dan konseling yang tidak efektif.

Pasukan tersembunyi yang terdiri dari pekerja sedikit demi sedikit ini melakukan pekerjaan yang sangat membosankan dan dibayar rendah, paling buruk, berbahaya, tidak sehat dan/atau berbahaya secara psikologis. Penelitian Gray menyatakan bahwa para pekerja di bidang ini masih menunjukkan keinginan untuk berinvestasi dalam pekerjaan lebih dari sekadar transaksi pembayaran tunggal, dan menyarankan agar dampak ekonomi, sosial dan psikologis dari 'pekerjaan hantu' harus ditangani secara sistematis.

Menjadikan masukan pekerja lebih transparan pada produk akhir, memastikan rantai nilai meningkatkan distribusi manfaat yang adil, dan memastikan struktur dukungan yang tepat bagi orang-orang yang berurusan dengan konten yang berbahaya secara psikologis merupakan langkah penting untuk mengatasi masalah etika.

Berbagi manfaat

AI mempunyai potensi untuk membawa manfaat yang signifikan dan beragam bagi masyarakat dan memfasilitasi, antara lain, peningkatan efisiensi dan produktivitas dengan biaya lebih rendah (OECD, n.d.). The Future of Life Institute (n.d.) menyatakan bahwa AI mungkin mampu mengatasi sejumlah masalah global yang paling sulit – kemiskinan, penyakit, konflik – dan dengan demikian meningkatkan taraf hidup yang tak terhitung jumlahnya.

Laporan mengenai AI, otomasi, dan ekonomi (2016) menyoroti pentingnya memastikan bahwa potensi manfaat AI tidak terakumulasi secara tidak merata, dan dapat diakses oleh sebanyak mungkin orang. Daripada menggambarkan perkembangan AI dan otomasi sebagai sesuatu yang mengarah pada hasil yang tak terelakkan yang ditentukan oleh

teknologi itu sendiri, laporan tersebut menyatakan bahwa inovasi dan perubahan teknologi 'tidak terjadi dalam ruang hampa': masa depan AI mungkin tidak dibentuk oleh kemampuan teknologi. Namun melalui berbagai insentif non-teknis. Selain itu, penemu atau pengembang AI memiliki potensi besar untuk menentukan penggunaan dan jangkauannya, sehingga menunjukkan perlunya para penemu mempertimbangkan potensi dampak yang lebih luas dari ciptaan mereka.

Otomatisasi lebih dapat diterapkan pada peran tertentu dibandingkan peran lainnya, sehingga menempatkan pekerja tertentu pada posisi yang dirugikan dan berpotensi meningkatkan ketimpangan upah. Dunia usaha mungkin termotivasi oleh profitabilitas. Namun, meskipun hal ini mungkin menguntungkan pemilik bisnis dan pemangku kepentingan, hal ini mungkin tidak menguntungkan pekerja.

Brundage dan Bryson (2016) menyebutkan studi kasus listrik, yang menurut mereka terkadang dianggap analog dengan AI. Meskipun listrik dapat membuat banyak bidang menjadi lebih produktif, menghilangkan hambatan, dan memberikan manfaat serta peluang bagi banyak kehidupan, dibutuhkan waktu puluhan tahun bagi listrik untuk menjangkau beberapa pasar, dan 'memang, lebih dari satu miliar [orang] masih kekurangan akses terhadap listrik'.

Untuk memastikan bahwa manfaat AI didistribusikan secara adil – dan untuk menghindari dinamika siapa pun yang mendesainnya terlebih dahulu – salah satu pilihannya adalah dengan menyatakan terlebih dahulu bahwa AI bukanlah barang pribadi melainkan untuk kepentingan semua orang. Pendekatan seperti ini memerlukan perubahan norma dan kebijakan budaya. Pedoman nasional dan pemerintah yang baru dapat mendasari strategi baru untuk memanfaatkan manfaat AI bagi masyarakat, membantu menavigasi transisi ekonomi yang didorong oleh AI, dan mempertahankan serta memperkuat kepercayaan masyarakat terhadap AI. Brundage dan Bryson (2016) setuju dengan seruan kebijakan dan peraturan ini, dengan menyatakan bahwa 'mendanai penelitian dasar saja tidak cukup dan mengharapkan penelitian tersebut disebarluaskan dan adil di masyarakat oleh pihak swasta'. Namun, skenario masa depan tersebut tidak ditentukan sebelumnya,, dan akan dibentuk oleh kebijakan dan pilihan yang ada saat ini.

Future of Life Institute (n.d.) mencantumkan sejumlah rekomendasi kebijakan untuk mengatasi kemungkinan 'dampak ekonomi, perpindahan tenaga kerja, kesenjangan, pengangguran akibat teknologi', dan ketegangan sosial dan politik yang mungkin menyertai AI. Hilangnya pekerjaan akibat AI akan memerlukan program pelatihan ulang baru serta dukungan sosial dan finansial bagi pekerja yang dipindahkan; permasalahan seperti ini mungkin memerlukan kebijakan ekonomi seperti pendapatan dasar universal dan skema perpajakan robot. Institut ini menyarankan bahwa kebijakan harus fokus pada mereka yang paling berisiko tertinggal – pengasuh, perempuan dan anak perempuan, kelompok masyarakat yang kurang terwakili dan kelompok rentan – dan pada mereka yang membangun sistem AI, untuk menargetkan setiap 'desain produk yang tidak tepat, titik buta, asumsi yang salah [dan] sistem nilai dan tujuan yang dikodekan ke dalam mesin (The Future of Life Institute, n.d.).

Menurut Brundage dan Bryson (2016), mengambil pendekatan proaktif terhadap kebijakan AI bukanlah hal yang 'prematur, salah arah, [atau] berbahaya', mengingat AI 'sudah cukup matang secara teknologi untuk memberikan dampak pada miliaran nyawa triliunan kali sehari'. Mereka menyarankan agar pemerintah berupaya meningkatkan pengetahuan terkait dan lebih mengandalkan para ahli; bahwa penelitian yang relevan mendapat alokasi dana lebih besar; bahwa para pembuat kebijakan merencanakan masa depan, mengupayakan 'kekokohan dan kesiapan dalam menghadapi ketidakpastian'; dan bahwa AI diterapkan secara luas dan secara proaktif dapat diakses (terutama di bidang-bidang yang memiliki nilai sosial yang besar, seperti kemiskinan, penyakit, atau energi bersih).

Dengan mempertimbangkan industri energi sebagai contohnya, AI mungkin dapat memodernisasi jaringan energi, meningkatkan keandalannya, dan mencegah pemadaman listrik dengan mengatur pasokan dan permintaan di tingkat lokal dan nasional, kata Wolfe (2017). 'Jaringan pintar' seperti ini akan menghemat uang perusahaan-perusahaan energi namun juga memungkinkan konsumen untuk secara aktif memantau penggunaan energi mereka secara real-time dan melihat penghematan biaya, meneruskan manfaat dari pengembang ke produsen ke konsumen – dan membuka cara-cara baru untuk berhemat, memperoleh penghasilan dan berinteraksi dengan jaringan energi. Jacobs (2017) membahas potensi 'prosumer' (mereka yang memproduksi dan mengonsumsi energi, berinteraksi dengan jaringan listrik dengan cara baru) untuk membantu desentralisasi produksi energi dan menjadi 'kekuatan pengganggu positif' dalam industri ketenagalistrikan – jika strategi energi diatur secara efektif melalui kebijakan dan manajemen yang diperbarui. Memberikan konsumen data yang real-time dan dapat diakses juga akan membantu mereka memilih tarif yang paling hemat biaya, kata Ramchurn dkk. (2013), mengingat memperkirakan secara akurat konsumsi tahunan seseorang dan menguraikan tarif yang rumit merupakan tantangan utama yang dihadapi konsumen energi. Oleh karena itu, hal ini mungkin mempunyai potensi untuk mengentaskan kemiskinan energi, mengingat kenaikan harga energi dan ketergantungan pada jaringan pasokan energi terpusat dapat menyebabkan rumah tangga berada dalam kemiskinan bahan bakar.

Konsentrasi kekuasaan di kalangan elit

'Apakah AI harus meningkatkan kesenjangan? Bisakah Anda merancang sistem yang menargetkan, misalnya, kebutuhan masyarakat termiskin? Jika AI digunakan untuk lebih memberikan manfaat bagi masyarakat kaya dibandingkan memberikan manfaat bagi masyarakat miskin, dan hal ini tampaknya, atau yang lebih meresahkan, memberikan tekanan yang tidak semestinya pada masyarakat yang sudah terpinggirkan, lalu apa yang dapat kita lakukan untuk mengatasinya? Apakah itu penggunaan AI yang tepat?'

Nemitz (2018) menulis bahwa akan menjadi 'naif' untuk mengabaikan bahwa AI akan memusatkan kekuasaan di tangan beberapa raksasa internet digital, karena 'realitas mengenai bagaimana [sebagian besar masyarakat] menggunakan Internet dan apa yang diberikan oleh Internet kepada mereka adalah hal yang sangat penting. dibentuk oleh beberapa perusahaan besar...perkembangan AI didominasi oleh perusahaan-perusahaan besar tersebut dan ekosistem yang bergantung pada mereka.

Akumulasi kekuatan teknologi, ekonomi dan politik di tangan lima pemain teratas – Google, Facebook, Microsoft, Apple dan Amazon – memberi mereka pengaruh yang tidak semestinya di bidang masyarakat yang relevan dengan pembangunan opini di negara-negara demokrasi: pemerintah, legislator, masyarakat sipil, partai politik, sekolah dan pendidikan, jurnalisme dan pendidikan jurnalisme dan sains dan penelitian.

Secara khusus, Nemitz prihatin bahwa penyelidikan terhadap dampak teknologi baru seperti AI terhadap hak asasi manusia, demokrasi, dan supremasi hukum mungkin terhambat oleh kekuatan perusahaan teknologi, yang tidak hanya membentuk pengembangan dan penerapan AI, tetapi juga perdebatan mengenai peraturannya. Nemitz mengidentifikasi beberapa bidang di mana raksasa teknologi menggunakan kekuasaannya:

1. **Keuangan.** Lima pemain teratas tidak hanya mampu berinvestasi besar-besaran dalam pengaruh politik dan sosial, mereka juga mampu membeli ide-ide baru dan start-up di bidang AI, atau bidang lain yang menarik bagi model bisnis mereka – sesuatu yang mereka minati. memang melakukan.
2. **Wacana publik.** Perusahaan teknologi mengendalikan infrastruktur yang menjadi tempat berlangsungnya wacana publik. Situs-situs seperti Facebook dan Google semakin menjadi sumber informasi politik utama, atau bahkan satu-satunya, bagi warga negara, terutama generasi muda, sehingga merugikan pihak keempat. Sebagian besar pendapatan iklan kini juga masuk ke Google dan Facebook, sehingga menghilangkan pendapatan utama dari surat kabar dan membuat jurnalisme investigatif menjadi tidak terjangkau.
3. **Mengumpulkan data pribadi.** Perusahaan-perusahaan ini mengumpulkan data pribadi untuk mendapatkan keuntungan, dan membuat profil orang-orang berdasarkan perilaku mereka (baik online maupun offline). Mereka mengetahui lebih banyak tentang kita dibandingkan diri kita sendiri atau teman-teman kita – dan mereka menggunakan serta menyediakan informasi ini untuk mendapatkan keuntungan, pengawasan, keamanan, dan kampanye pemilu.

Secara keseluruhan, Nemitz menyimpulkan hal itu

'akumulasi kekuasaan di tangan segelintir orang – kekuasaan uang, kekuasaan atas infrastruktur demokrasi dan wacana, kekuasaan atas individu berdasarkan profiling dan dominasi inovasi AI... harus dilihat bersama-sama, dan... harus menjadi masukan bagi masyarakat. perdebatan saat ini tentang etika dan hukum untuk AI'.

Bryson (2019), sementara itu, meyakini pemusatan kekuasaan ini bisa menjadi konsekuensi tak terelakkan dari turunnya biaya teknologi robotik. Biaya yang tinggi dapat menjaga keberagaman dalam sistem perekonomian. Misalnya, ketika biaya transportasi tinggi, seseorang mungkin memilih untuk menggunakan toko lokal daripada mencari penyedia global terbaik untuk suatu barang tertentu. Biaya yang lebih rendah memungkinkan relatif sedikit perusahaan untuk mendominasi, dan jika hanya sedikit penyedia layanan yang menerima seluruh bisnis, maka mereka juga akan menerima seluruh kekayaannya.

Ketidakstabilan politik

Bryson (2019) juga mencatat bahwa kebangkitan AI dapat menyebabkan ketimpangan kekayaan dan pergolakan politik. Ketimpangan sangat berkorelasi dengan polarisasi politik (McCarty et al., 2016), dan salah satu konsekuensi dari polarisasi adalah meningkatnya politik identitas, di mana keyakinan digunakan untuk menandakan status atau afiliasi dalam kelompok (Iyengar et al., 2012; Newman dkk., 2014). Sayangnya hal ini dapat mengakibatkan situasi di mana keyakinan lebih terikat pada afiliasi kelompok seseorang dibandingkan dengan fakta obyektif, dan hilangnya kepercayaan pada para ahli.

'Meskipun kadang-kadang dimotivasi oleh penggunaan yang tidak bertanggung jawab atau bahkan penyalahgunaan posisi oleh beberapa ahli, secara umum kehilangan akses terhadap pandangan para ahli adalah sebuah bencana. Tidak seorang pun, betapapun cerdasnya, dapat menguasai seluruh pengetahuan manusia sepanjang hidupnya. Jika masyarakat mengabaikan simpanan keahlian yang telah mereka bangun – seringkali melalui pendanaan pendidikan tinggi dari pajak – maka masyarakat akan berada pada posisi yang sangat dirugikan' (Bryson, 2019).

Privasi, hak asasi manusia dan martabat

AI akan berdampak besar terhadap privasi dalam dekade mendatang. Privasi dan martabat pengguna AI harus dipertimbangkan dengan cermat saat merancang robot layanan, perawatan, dan pendamping, karena bekerja di rumah penduduk berarti mereka akan mengetahui momen-momen yang sangat pribadi (seperti mandi dan berpakaian). Namun, aspek AI lainnya juga akan memengaruhi privasi. Smith (2018), Presiden Microsoft, baru-baru ini berkomentar:

'Teknologi [Intelligent] mengangkat isu-isu yang menjadi inti perlindungan hak asasi manusia seperti privasi dan kebebasan berekspresi. Masalah-masalah ini meningkatkan tanggung jawab bagi perusahaan teknologi yang menciptakan produk-produk tersebut. Dalam pandangan kami, hal ini juga memerlukan peraturan pemerintah yang bijaksana dan pengembangan norma-norma seputar penggunaan yang dapat diterima.'

Privasi dan hak data

'Manusia tidak akan memiliki hak pilihan dan kendali [atas data mereka] dengan cara apa pun jika mereka tidak diberikan alat untuk mewujudkannya'. (John Havens)

Salah satu cara AI memengaruhi privasi adalah melalui Intelligent Personal Assistants (IPA) seperti Amazon's Echo, Google's Home, dan Apple's Siri. Perangkat yang diaktifkan dengan suara ini mampu mempelajari minat dan perilaku penggunanya, namun ada kekhawatiran mengenai fakta bahwa perangkat tersebut selalu aktif dan mendengarkan di latar belakang.

Sebuah survei terhadap pelanggan IPA menunjukkan bahwa kekhawatiran terbesar terhadap privasi masyarakat adalah peretasan perangkat mereka (68,63%), diikuti oleh

pengumpulan informasi pribadi tentang mereka (16%), mendengarkan percakapan mereka 24/7 (10%), merekam percakapan pribadi (12%), tidak menghormati privasi (6%), menyimpan data (6%) dan sifat perangkat yang 'menyeramkan' (4%) (Manikonda dkk, 2018). Namun terlepas dari kekhawatiran ini, masyarakat sangat positif terhadap perangkat ini, dan merasa nyaman menggunakannya.

Aspek lain dari AI yang memengaruhi privasi adalah Big Data. Teknologi ini berada pada tahap di mana pencatatan jangka panjang dapat disimpan mengenai siapa saja yang menghasilkan data yang dapat disimpan – siapa pun yang memiliki tagihan, kontrak, perangkat digital, atau riwayat kredit, belum lagi tulisan publik dan penggunaan media sosial. Catatan digital dapat dicari menggunakan algoritma pengenalan pola, yang berarti bahwa kita telah kehilangan asumsi default anonimitas karena ketidakjelasan (Selinger dan Hartzog, 2017).

Siapa pun dari kita dapat diidentifikasi melalui perangkat lunak pengenalan wajah atau pengumpulan data dari kebiasaan belanja atau media sosial kita (Pasquale, 2015). Kebiasaan online ini mungkin tidak hanya menunjukkan identitas kita, namun juga kecenderungan politik atau ekonomi kita, dan strategi apa yang mungkin efektif untuk mengubahnya (Cadwalladr, 2017a,b).

Pembelajaran mesin memungkinkan kita mengekstrak informasi dari data dan menemukan pola baru, serta mampu mengubah data yang tampaknya tidak berbahaya menjadi data pribadi yang sensitif. Misalnya, pola penggunaan media sosial dapat memprediksi kategori kepribadian, preferensi politik, dan bahkan hasil kehidupan. Pilihan kata, atau bahkan tekanan tulisan tangan pada stylus digital, dapat mengindikasikan keadaan emosi, termasuk apakah seseorang berbohong. Hal ini berdampak signifikan terhadap privasi dan anonimitas, baik online maupun offline.

Aplikasi AI berdasarkan pembelajaran mesin memerlukan akses ke data dalam jumlah besar, namun subjek data memiliki hak terbatas atas cara data mereka. Baru-baru ini, UE mengadopsi Peraturan Perlindungan Data Umum (GDPR) yang baru untuk melindungi privasi warga negara. Namun, peraturan tersebut hanya berlaku untuk data pribadi, dan bukan kumpulan data 'anonim' yang biasanya digunakan untuk melatih model.

Selain itu, data pribadi, atau informasi tentang siapa saja yang mengikuti rangkaian pelatihan, dalam kasus tertentu dapat direkonstruksi dari sebuah model, yang berpotensi menimbulkan konsekuensi signifikan terhadap regulasi sistem ini. Misalnya, meskipun orang mempunyai hak tentang bagaimana data pribadi mereka digunakan dan disimpan, mereka memiliki hak terbatas atas model terlatih. Sebaliknya, model-model tersebut biasanya dianggap diatur oleh berbagai hak kekayaan intelektual, seperti rahasia dagang. Misalnya, saat ini, tidak ada hak atau kewajiban perlindungan data terkait model setelah model tersebut dibuat, namun sebelum ada keputusan yang diambil mengenai penggunaannya.

Hal ini menimbulkan sejumlah masalah etika. Tingkat kontrol apa yang dimiliki subjek terhadap data yang dikumpulkan tentang mereka? Haruskah individu mempunyai hak untuk menggunakan model tersebut, atau setidaknya mengetahui untuk apa model tersebut digunakan, mengingat kepentingan mereka dalam melatih model tersebut? Mungkinkah sistem pembelajaran mesin yang mencari pola dalam data secara tidak sengaja melanggar

privasi seseorang jika, misalnya, pengurutan genom salah satu anggota keluarga mengungkapkan informasi kesehatan anggota keluarga lainnya?

Masalah etika lainnya adalah bagaimana mencegah identitas, atau informasi pribadi, seseorang yang terlibat dalam pelatihan model agar tidak diketahui (misalnya melalui serangan siber). Veale dkk. (2018) berpendapat bahwa perlindungan ekstra harus diberikan kepada orang-orang yang datanya telah digunakan untuk melatih model, seperti hak untuk mengakses model; mengetahui dari mana barang-barang tersebut berasal, dan kepada siapa barang-barang tersebut diperdagangkan atau disebar; hak untuk menghapus diri mereka sendiri dari model yang terlatih; dan hak untuk menyatakan keinginan agar model tersebut tidak digunakan di masa depan.

Hak asasi Manusia

AI mempunyai dampak penting terhadap demokrasi, dan hak masyarakat atas kehidupan pribadi dan martabat. Misalnya, jika AI dapat digunakan untuk menentukan keyakinan politik masyarakat, maka individu dalam masyarakat kita mungkin rentan terhadap manipulasi. Para ahli strategi politik dapat menggunakan informasi ini untuk mengidentifikasi pemilih mana yang kemungkinan besar akan dibujuk untuk mengubah afiliasi partainya, atau untuk meningkatkan atau menurunkan kemungkinan mereka untuk memilih, dan kemudian menggunakan sumber daya untuk membujuk mereka agar melakukan hal tersebut. Strategi seperti ini diduga berdampak signifikan terhadap hasil pemilu baru-baru ini di Inggris dan Amerika Serikat (Cadwalladr, 2017a; b).

Alternatifnya, jika AI dapat menilai keadaan emosi seseorang dan mengukur kapan mereka berbohong, orang-orang ini dapat menghadapi penganiayaan oleh orang-orang yang tidak menyetujui keyakinan mereka, mulai dari perundungan oleh individu hingga hilangnya peluang karier. Di beberapa masyarakat, hal ini dapat mengakibatkan hukuman penjara atau bahkan kematian di tangan negara.

Pengawasan

'Jaringan kamera yang saling terhubung memberikan pengawasan terus-menerus terhadap banyak kota metropolitan. Dalam waktu dekat, drone, robot, dan kamera yang dapat dipakai dapat memperluas pengawasan ini ke lokasi pedesaan dan rumah seseorang, tempat ibadah, dan bahkan lokasi di mana privasi dianggap sakral, seperti kamar mandi dan ruang ganti. Ketika penerapan robot dan kamera yang dapat dikenakan meluas ke rumah kita dan mulai menangkap dan merekam semua aspek kehidupan sehari-hari, kita mulai mendekati dunia di mana semua orang, bahkan orang yang berada di sekitar, terus-menerus diamati oleh berbagai kamera ke mana pun mereka pergi'.

Ini mungkin terdengar seperti mimpi buruk distopia, namun penggunaan AI untuk memata-matai semakin meningkat. Misalnya, seorang hakim di Ohio baru-baru ini memutuskan bahwa data yang dikumpulkan oleh alat pacu jantung seorang pria dapat digunakan sebagai bukti bahwa dia melakukan pembakaran. Data yang dikumpulkan oleh perangkat Amazon Alexa juga digunakan sebagai bukti. Ratusan perangkat rumah yang terhubung, termasuk peralatan rumah tangga dan televisi, kini secara rutin mengumpulkan data yang dapat digunakan sebagai bukti atau diakses oleh peretas. Video dapat digunakan

untuk berbagai tujuan yang sangat mengganggu, seperti mendeteksi atau mengkarakterisasi emosi seseorang.

AI juga dapat digunakan untuk memantau dan memprediksi potensi pembuat onar. Kapasitas pengenalan wajah diduga digunakan di Tiongkok, tidak hanya untuk mengidentifikasi individu, namun untuk mengidentifikasi suasana hati dan perhatian mereka baik di kamp pendidikan ulang maupun di sekolah biasa (Bryson, 2019). Ada kemungkinan, teknologi tersebut dapat digunakan untuk menghukum siswa yang tidak memperhatikan atau menghukum narapidana yang tidak tampak senang untuk mematuhi pendidikan (ulang) mereka.

Sayangnya, pemerintah tidak selalu mengutamakan kepentingan warganya. Pemerintah Tiongkok telah menggunakan sistem pengawasan untuk menempatkan lebih dari satu juta warganya di kamp pendidikan ulang karena kejahatan dalam mengekspresikan identitas Muslim mereka (Human Rights Watch, 2018). Terdapat risiko bahwa pemerintah yang khawatir akan perbedaan pendapat akan menggunakan AI untuk menekan, memenjarakan, dan merugikan individu.

Lembaga penegak hukum di India sudah menggunakan teknologi AI hibrida yang canggih¹ untuk mendigitalkan catatan kriminal, dan menggunakan pengenalan wajah untuk memprediksi dan mengenali aktivitas kriminal. Ada juga rencana untuk melatih drone untuk mengidentifikasi perilaku kekerasan di ruang publik, dan menguji drone ini di festival musik di India. Sebagian besar program ini bertujuan untuk mengurangi tingkat kejahatan, mengelola ruang publik yang ramai untuk meningkatkan keselamatan, dan memberikan efisiensi pada penegakan hukum. Namun, hal-hal tersebut memiliki implikasi yang jelas terhadap privasi dan hak asasi manusia, karena penampilan dan perilaku publik seseorang dipantau, dikumpulkan, disimpan, dan mungkin dibagikan tanpa izin. AI yang dibahas tidak hanya beroperasi tanpa adanya perlindungan untuk mencegah penyalahgunaan, sehingga rentan terhadap pengawasan dan pelanggaran privasi, namun AI juga beroperasi pada tingkat akurasi yang dipertanyakan. Hal ini dapat menyebabkan penangkapan palsu dan orang-orang dari komunitas yang sangat rentan dan terpinggirkan harus membuktikan bahwa mereka tidak bersalah.

Kebebasan berbicara

Kebebasan berbicara dan berekspresi adalah hak mendasar dalam masyarakat demokratis. Hal ini bisa sangat dipengaruhi oleh AI. AI telah banyak dipuji oleh perusahaan-perusahaan teknologi sebagai solusi terhadap permasalahan seperti ujaran kebencian, ekstremisme kekerasan, dan misinformasi digital (Li dan Williams, 2018). Di India, alat analisis sentimen semakin banyak digunakan untuk mengukur nada dan sifat pembicaraan online, dan sering kali dilatih untuk melakukan penghapusan konten otomatis (Marda, 2018). Pemerintah India juga telah menyatakan minatnya dalam menggunakan AI untuk mengidentifikasi berita palsu dan meningkatkan citra India di media sosial (Seth 2017). Ini adalah tren yang berbahaya, mengingat terbatasnya kompetensi pembelajaran mesin untuk memahami nada dan konteks. Penghapusan konten otomatis berisiko menyensor ucapan yang sah; Risiko ini menjadi lebih besar karena tindakan ini dilakukan oleh perusahaan swasta, yang terkadang

bertindak berdasarkan instruksi pemerintah. Pengawasan ketat berdampak pada kebebasan berekspresi karena mendorong sensor mandiri.

Bias

AI diciptakan oleh manusia, sehingga rentan terhadap bias. Bias sistematis mungkin timbul akibat data yang digunakan untuk melatih sistem, atau akibat nilai-nilai yang dianut oleh pengembang dan pengguna sistem. Hal ini paling sering terjadi ketika aplikasi pembelajaran mesin dilatih pada data yang hanya mencerminkan kelompok demografi tertentu, atau yang mencerminkan bias masyarakat. Sejumlah kasus mendapat perhatian karena mendorong bias sosial yang tidak disengaja, yang kemudian direproduksi atau secara otomatis diperkuat oleh sistem AI.

Contoh bias AI

Organisasi jurnalisme investigasi ProPublica menunjukkan bahwa COMPAS, perangkat lunak berbasis pembelajaran mesin yang digunakan di AS untuk menilai kemungkinan terdakwa melakukan kembali pelanggaran pidana, sangat bias terhadap orang kulit hitam Amerika. Sistem COMPAS lebih cenderung salah memprediksi bahwa terdakwa berkulit hitam akan melakukan pelanggaran kembali, dan pada saat yang sama, secara keliru memprediksi hal sebaliknya dalam kasus terdakwa berkulit putih.

Para peneliti telah menemukan bahwa alat distribusi iklan otomatis lebih cenderung mendistribusikan iklan pekerjaan bergaji tinggi kepada laki-laki dibandingkan perempuan (Datta et al., 2015). Perekrutan berdasarkan informasi AI rentan terhadap bias; alat pembelajaran mandiri Amazon yang digunakan untuk menilai pencari kerja ternyata sangat berpihak pada laki-laki dan memberi peringkat tinggi pada mereka. Sistem telah belajar untuk memprioritaskan lamaran yang menekankan karakteristik laki-laki, dan menurunkan peringkat lamaran dari universitas dengan kehadiran perempuan yang kuat.

Banyak database gambar populer berisi gambar yang dikumpulkan hanya dari beberapa negara (AS, Inggris), sehingga dapat menyebabkan bias dalam hasil penelusuran. Basis data seperti ini secara teratur menggambarkan perempuan melakukan pekerjaan dapur sementara laki-laki sedang berburu, misalnya, dan penelusuran untuk 'gaun pengantin' menghasilkan versi putih standar yang disukai di masyarakat barat, sedangkan gaun pengantin India dikategorikan sebagai 'gaun pertunjukan'. seni' atau 'kostum' (Zhou 2018). Ketika aplikasi diprogram dengan bias seperti ini, hal ini dapat menyebabkan situasi seperti kamera secara otomatis memperingatkan fotografer bahwa subjeknya menutup mata ketika mengambil foto orang Asia, karena kamera telah dilatih tentang stereotip, maskulin, dan maskulin.

ImageNet, yang bertujuan memetakan dunia objek, merupakan kumpulan data besar yang terdiri dari 14,1 juta gambar yang disusun dalam lebih dari 20.000 kategori – yang sebagian besar adalah tumbuhan, batu, hewan. Para pekerja telah mengurutkan 50 gambar dalam satu menit ke dalam ribuan kategori untuk ImageNet – dengan kecepatan seperti itu, terdapat potensi ketidakakuratan yang besar. Pemberian tag yang bermasalah, tidak akurat – dan diskriminatif (lihat Diskriminasi di atas) dapat dipertahankan dalam kumpulan data melalui banyak iterasi

Ada beberapa aktivitas yang menunjukkan bias yang terkandung dalam kumpulan pelatihan data. Salah satunya adalah aplikasi pengenalan wajah (ImageNet Roulette) yang membuat asumsi tentang Anda sepenuhnya berdasarkan foto wajah Anda yang diunggah – mulai dari usia dan jenis kelamin hingga profesi dan bahkan karakteristik pribadi. Model ini telah dikritik karena pelabelannya yang menyinggung, tidak akurat, dan rasis – namun pembuatnya mengatakan bahwa ini adalah antarmuka yang menunjukkan kepada pengguna bagaimana model pembelajaran mesin menafsirkan data dan bagaimana hasilnya bisa sangat mengganggu.

Implikasi

Karena banyak model pembelajaran mesin yang dibuat dari data yang dihasilkan manusia, bias manusia dapat dengan mudah mengakibatkan distribusi data pelatihan yang tidak merata. Jika pengembang tidak berupaya mengenali dan mengatasi bias ini, aplikasi dan produk AI dapat melanggengkan ketidakadilan dan diskriminasi. AI yang bias terhadap kelompok tertentu dalam masyarakat dapat menimbulkan dampak yang luas. Penggunaannya dalam penegakan hukum atau keamanan nasional, misalnya, dapat mengakibatkan sejumlah demografi dipenjarakan atau ditahan secara tidak adil. Penggunaan AI untuk melakukan pemeriksaan kredit dapat mengakibatkan penolakan pinjaman terhadap beberapa individu secara tidak adil, sehingga menyulitkan mereka untuk keluar dari siklus kemiskinan. Jika AI digunakan untuk menyaring orang-orang untuk melamar pekerjaan atau diterima di universitas, hal ini dapat mengakibatkan seluruh lapisan masyarakat dirugikan.

Masalah ini diperburuk oleh kenyataan bahwa aplikasi AI biasanya berupa 'kotak hitam', di mana konsumen tidak mungkin menilai apakah data yang digunakan untuk melatihnya adil atau representatif. Hal ini membuat bias sulit dideteksi dan ditangani. Oleh karena itu, terdapat banyak penelitian terbaru tentang menjadikan pembelajaran mesin adil, akuntabel, dan transparan, dan lebih banyak aktivitas dan demonstrasi yang dapat dilakukan oleh publik akan bermanfaat.

Demokrasi

Seperti telah dibahas, pemusatan kekuatan teknologi, ekonomi dan politik di antara beberapa perusahaan besar dapat memberikan mereka pengaruh yang tidak semestinya terhadap pemerintah – namun penerapan dan penerapan AI juga dapat mengancam demokrasi.

Berita palsu dan media sosial

Sepanjang sejarah, kandidat politik yang berkampanye untuk mendapatkan jabatan hanya mengandalkan bukti dan survei yang bersifat anekdotal dan terbatas untuk memberi mereka gambaran tentang apa yang dipikirkan para pemilih. Kini dengan munculnya Big Data, para politisi memiliki akses terhadap sejumlah besar informasi yang memungkinkan mereka menargetkan kategori pemilih tertentu dan mengembangkan pesan yang paling sesuai dengan mereka.

Hal ini mungkin merupakan hal yang baik bagi para politisi, namun terdapat banyak bukti bahwa teknologi yang didukung AI telah disalahgunakan secara sistematis untuk memanipulasi warga negara dalam pemilu baru-baru ini, sehingga merusak demokrasi.

Misalnya, 'bot' – akun otonom – digunakan untuk menyebarkan berita dan propaganda yang bias melalui Twitter menjelang pemilihan presiden AS tahun 2016 dan pemungutan suara Brexit di Inggris (Pham, Gorodnichenko dan Talavera, 2018). Beberapa dari akun otomatis ini dibuat dan dioperasikan dari Rusia dan, sampai batas tertentu, mampu membiaskan konten yang dilihat di media sosial, sehingga memberikan kesan dukungan yang salah.

Selama pemilu presiden AS tahun 2016, bot pro-Trump ditemukan telah menyusup ke ruang online yang digunakan oleh para aktivis pro-Clinton, tempat mereka menyebarkan konten yang sangat otomatis, sehingga menghasilkan seperempat lalu lintas Twitter tentang pemilu tahun 2016 (Hess, 2016). Bot juga berperan besar dalam mempopulerkan #MacronLeaks di media sosial hanya beberapa hari sebelum pemilihan presiden Prancis tahun 2017 (Polonski, 2017). Mereka membombardir Facebook dan Twitter dengan campuran informasi yang bocor dan laporan palsu, sehingga membangun narasi bahwa Emmanuel Macron adalah seorang penipu dan munafik.

Sebuah laporan baru-baru ini menemukan bahwa setidaknya 28 negara – termasuk negara otoriter dan negara demokrasi – menggunakan 'pasukan siber' untuk memanipulasi opini publik melalui aplikasi jejaring sosial utama (Bradshaw dan Howard, 2017). Pasukan siber ini menggunakan berbagai taktik untuk mempengaruhi opini publik, termasuk melontarkan kata-kata kasar dan melecehkan pengguna media sosial lain yang menyampaikan kritik terhadap pemerintah. Di Rusia, pasukan siber diketahui menysar jurnalis dan pembangkang politik, dan di Meksiko, jurnalis sering kali menjadi sasaran dan dilecehkan melalui media sosial oleh pasukan siber yang disponsori pemerintah (O'Carroll, 2017). Lainnya menggunakan bot otomatis – menurut Bradshaw dan Howard (2017), bot telah digunakan oleh aktor pemerintah di Argentina, Azerbaijan, Iran, Meksiko, Filipina, Rusia, Arab Saudi, Korea Selatan, Suriah, Turki, dan Venezuela. Bot ini sering digunakan untuk membanjiri jaringan media sosial dengan spam dan berita 'palsu' atau bias, dan juga dapat memperkuat suara dan ide marginal dengan meningkatkan jumlah suka, berbagi, dan retweet yang mereka terima, sehingga menciptakan kesan popularitas, momentum, atau artifisial. relevansi. Menurut penulis, rezim otoriter bukanlah satu-satunya atau bahkan yang terbaik dalam manipulasi media sosial yang terorganisir.

Selain membentuk debat online, AI dapat digunakan untuk menargetkan dan memanipulasi pemilih individu. Selama pemilihan presiden AS tahun 2016, perusahaan ilmu data Cambridge Analytica memperoleh akses ke data pribadi lebih dari 50 juta pengguna Facebook, yang mereka gunakan untuk membuat profil psikologis orang-orang guna menargetkan iklan kepada pemilih yang mereka anggap paling reseptif.

Masih ada ketidakpercayaan umum terhadap media sosial di kalangan masyarakat di seluruh Eropa, dan kontennya dipandang dengan hati-hati; survei Eurobarometer pada tahun 2017 menemukan bahwa hanya 7% responden menganggap berita yang dipublikasikan di platform sosial online secara umum dapat dipercaya (Komisi Eropa, 2017). Namun, demokrasi perwakilan bergantung pada pemilu yang bebas dan adil di mana warga negara dapat memilih tanpa manipulasi – dan AI mengancam akan melemahkan proses ini.

Gelembung berita dan ruang gema

Media semakin banyak menggunakan pemberi rekomendasi berita algoritmik (ANR) untuk menargetkan berita yang disesuaikan kepada masyarakat berdasarkan minat mereka (Thurman, 2011; Gillespie, 2014). Namun menyajikan berita kepada pembaca berdasarkan riwayat bacaan mereka sebelumnya akan menurunkan kemungkinan orang menemukan konten, opini, dan sudut pandang yang berbeda dan belum ditemukan (Harambam dkk., 2018). Ada bahaya bahwa hal ini dapat mengakibatkan meningkatnya polarisasi masyarakat, karena masyarakat pada dasarnya hidup dalam 'ruang gema' dan 'gelembung filter' (Pariser, 2011) di mana mereka hanya terpapar pada sudut pandang mereka sendiri. Interaksi berbagai gagasan dan orang-orang dianggap penting untuk berfungsinya demokrasi.

Akhir dari demokrasi

Beberapa komentator mempertanyakan apakah negara demokrasi sangat cocok dengan era AI dan pembelajaran mesin, dan apakah penerapannya akan memungkinkan negara-negara dengan sistem politik lain memperoleh keuntungan. Selama 200 tahun terakhir demokrasi telah berkembang pesat karena kebebasan individu berdampak baik bagi perekonomian. Kebebasan mendorong inovasi, meningkatkan perekonomian dan kekayaan, serta menciptakan masyarakat kaya yang menghargai kebebasan. Namun, bagaimana jika hubungan tersebut melemah? Bagaimana jika pertumbuhan ekonomi di masa depan tidak lagi bergantung pada kebebasan individu dan semangat kewirausahaan?

Perekonomian yang direncanakan secara terpusat dan dikendalikan oleh negara mungkin lebih cocok untuk era baru AI, karena tidak begitu peduli dengan hak-hak individu dan privasi masyarakat. Misalnya saja, besarnya populasi di suatu negara berarti bahwa bisnis di Tiongkok memiliki akses terhadap data dalam jumlah besar, dengan sedikit batasan mengenai bagaimana data tersebut dapat digunakan. Di Tiongkok, tidak ada undang-undang privasi atau perlindungan data, seperti aturan GDPR baru di Eropa. Karena Tiongkok akan segera menjadi pemimpin dunia dalam bidang AI, hal ini berarti Tiongkok dapat menentukan masa depan teknologi dan batasan penggunaannya.

“Beberapa tahun terakhir menunjukkan bahwa teknologi digital berkembang dengan baik di bawah kondisi monopoli: semakin besar suatu perusahaan, semakin banyak data dan daya komputasi yang didapat, dan semakin efisien perusahaan tersebut; semakin efisien, semakin banyak data dan daya komputasi yang didapat, dalam siklus yang terus berlanjut” (Bartlett, 2018). Menurut Bartlett, kecintaan masyarakat terhadap 'kenyamanan' berarti bahwa jika 'machinokrasi' mampu memberikan kekayaan, kemakmuran, dan stabilitas, banyak orang mungkin akan sangat senang dengan hal tersebut.

2.2 DAMPAK TERHADAP PSIKOLOGI MANUSIA

AI menjadi semakin baik dalam memodelkan pemikiran, pengalaman, tindakan, percakapan, dan hubungan manusia. Di zaman di mana kita sering berinteraksi dengan mesin seolah-olah mereka adalah manusia, apa dampaknya terhadap hubungan antarmanusia di dunia nyata?

Hubungan

Hubungan dengan orang lain merupakan inti dari keberadaan manusia. Di masa depan, robot diharapkan dapat melayani manusia dalam berbagai peran sosial: merawat, mengurus rumah, merawat anak-anak dan orang tua, mengajar, dan banyak lagi. Kemungkinan besar robot juga akan dirancang untuk tujuan eksplisit yaitu seks dan persahabatan. Robot-robot ini mungkin dirancang agar terlihat dan berbicara seperti manusia. Manusia mungkin mulai membentuk ikatan emosional dengan robot, bahkan mungkin merasa cinta terhadap robot. Jika ini terjadi, apa pengaruhnya terhadap hubungan manusia dan jiwa manusia?

Hubungan manusia-robot

'Risiko terbesar [AI] yang dihadapi siapa pun adalah hilangnya kemampuan berpikir sendiri. Kita sudah melihat orang-orang lupa cara membaca peta, mereka melupakan keterampilan lainnya. Jika kita kehilangan kemampuan untuk mawas diri, kita kehilangan hak pilihan manusia dan kita berputar-putar.

Salah satu bahayanya adalah penipuan dan manipulasi. Robot sosial yang dicintai dan dipercaya dapat disalahgunakan untuk memanipulasi manusia (Scheutz 2012); misalnya, seorang peretas dapat mengendalikan robot pribadi dan mengeksploitasi hubungan uniknya dengan pemiliknya untuk mengelabui pemiliknya agar membeli produk. Meskipun manusia sebagian besar dicegah melakukan hal ini karena perasaan seperti empati dan rasa bersalah, robot tidak memiliki konsep mengenai hal ini.

Perusahaan dapat merancang robot masa depan dengan cara yang meningkatkan kepercayaan dan daya tariknya. Misalnya saja, jika diketahui bahwa manusia lebih jujur saat menggunakan robot atau AI percakapan (chatbots) dibandingkan dengan manusia lain, maka hanya masalah waktu saja sebelum robot digunakan untuk menginterogasi manusia — dan jika ternyata robot memang lebih jujur, umumnya lebih dapat dipercaya dibandingkan manusia, maka robot kemungkinan besar akan digunakan sebagai perwakilan penjualan.

Ada kemungkinan juga bahwa manusia menjadi tergantung secara psikologis pada robot. Teknologi diketahui memanfaatkan fungsi penghargaan di otak, dan kecanduan ini dapat mengarahkan orang untuk melakukan tindakan yang tidak akan mereka lakukan jika tidak melakukannya.

Mungkin sulit untuk memprediksi dampak psikologis dari menjalin hubungan dengan robot. Misalnya, Borenstein dan Arkin (2019) menanyakan bagaimana hubungan 'bebas risiko' dengan robot dapat memengaruhi perkembangan mental dan sosial pengguna; Agaknya, robot tidak akan diprogram untuk putus dengan manusia, sehingga secara teoritis menghilangkan naik turunnya emosi dari suatu hubungan.

Menikmati persahabatan atau hubungan dengan robot pendamping mungkin melibatkan kesalahan, pada tingkat sadar atau tidak sadar, robot itu adalah orang sungguhan. Untuk mendapatkan manfaat dari hubungan tersebut, seseorang harus 'secara sistematis menipu diri sendiri mengenai sifat sebenarnya dari hubungan mereka dengan [AI]' (Sparrow, 2002). Menurut Sparrow, menuruti 'sentimentalitas yang tercela secara moral' melanggar kewajiban kita sendiri untuk memahami dunia secara akurat. Masyarakat yang rentan akan sangat beresiko menjadi korban penipuan ini (Sparrow dan Sparrow, 2006).

Hubungan manusia-manusia

Robot dapat mempengaruhi stabilitas perkawinan atau hubungan seksual. Misalnya, perasaan cemburu bisa muncul jika pasangan menghabiskan waktu bersama robot, seperti 'pacar virtual' (avatar chatbot). Hilangnya kontak dengan sesama manusia dan mungkin penarikan diri dari hubungan normal sehari-hari juga mungkin terjadi. Misalnya, seseorang yang memiliki robot pendamping mungkin enggan pergi ke acara (misalnya, pernikahan) yang biasanya dihadiri oleh pasangan manusia-manusia. Orang-orang yang menjalin hubungan manusia-robot mungkin mendapat stigma.

Ada beberapa masalah etika yang ditimbulkan oleh manusia yang menjalin hubungan dengan robot:

- Bisakah robot mengubah keyakinan, sikap, dan/atau nilai-nilai yang kita miliki tentang hubungan antarmanusia? Orang-orang mungkin menjadi tidak sabar dan tidak mau berusaha menjalin hubungan antarmanusia ketika mereka dapat menjalin hubungan dengan robot yang 'sempurna' dan menghindari tantangan-tantangan ini.
- Mungkinkah 'robot intim' menyebabkan peningkatan perilaku kekerasan? Beberapa peneliti berpendapat bahwa 'sexbots' akan mengubah persepsi masyarakat tentang nilai-nilai manusia, meningkatkan keinginan atau kesediaan masyarakat untuk menyakiti orang lain. Jika kita bisa memperlakukan robot sebagai alat kepuasan seksual, maka kita akan cenderung memperlakukan orang lain dengan cara yang sama. Misalnya, jika pengguna berulang kali meninju robot pendampingnya, apakah tindakan tersebut tidak etis (Lalji, 2015)? Akankah kekerasan terhadap robot menormalkan pola perilaku yang pada akhirnya akan berdampak pada manusia lain? Namun, beberapa orang berpendapat bahwa robot bisa menjadi pelampiasan hasrat seksual, mengurangi kemungkinan terjadinya kekerasan, atau membantu pemulihan dari penyerangan.

Mesin yang dibuat agar terlihat dan bertindak seperti kita juga dapat memengaruhi 'rangkaiannya sosial' kapasitas yang telah kita kembangkan untuk bekerja sama satu sama lain, termasuk cinta, persahabatan, kerja sama, dan pengajaran (Christakis, 2019). Dengan kata lain, AI dapat mengubah betapa penuh kasih dan baik hati kita — tidak hanya dalam interaksi langsung kita dengan mesin tersebut, namun juga dalam interaksi kita satu sama lain. Misalnya, haruskah kita mengkhawatirkan dampak sikap kasar anak-anak terhadap asisten digital seperti Alexa atau Siri? Apakah hal ini memengaruhi cara mereka memandang atau memperlakukan orang lain?

Penelitian menunjukkan bahwa robot memiliki kapasitas untuk mengubah sikap kooperatif kita. Dalam sebuah eksperimen, sekelompok kecil orang bekerja dengan robot humanoid untuk memasang rel kereta api di dunia virtual. Robot itu diprogram untuk sesekali melakukan kesalahan — dan mengakui kesalahan tersebut serta meminta maaf. Memiliki robot yang kikuk dan meminta maaf sebenarnya membantu kelompok-kelompok ini bekerja lebih baik daripada kelompok kontrol, dengan meningkatkan kolaborasi dan komunikasi di antara anggota kelompok manusia. Hal ini juga berlaku pada percobaan kedua, di mana orang-orang dalam kelompok yang terdiri dari robot-robot yang rawan kesalahan secara konsisten mengungguli orang lain dalam tugas pemecahan masalah.

Kedua penelitian ini menunjukkan bahwa AI dapat meningkatkan cara manusia berhubungan satu sama lain. Namun, AI juga dapat membuat kita berperilaku kurang produktif dan kurang etis. Dalam eksperimen lain, Christakis dan timnya memberikan beberapa ribu subjek uang untuk digunakan dalam beberapa putaran game online. Di setiap putaran, subjek diberitahu bahwa mereka bisa menjadi egois dan menyimpan uang mereka, atau menjadi altruistik dan menyumbangkan sebagian atau seluruh uangnya kepada tetangga mereka. Jika mereka memberikan sumbangan, para peneliti akan mencocokkannya, sehingga menggandakan jumlah uang yang diterima tetangga mereka. Meskipun dua pertiga orang pada awalnya bertindak altruistik, para ilmuwan menemukan bahwa perilaku kelompok tersebut dapat diubah hanya dengan menambahkan beberapa robot (menyamar sebagai pemain manusia) yang berperilaku egois. Akhirnya, para pemain manusia berhenti bekerja sama satu sama lain. Bot kemudian mengubah sekelompok orang yang murah hati menjadi orang yang egois.

Fakta bahwa AI dapat mengurangi kemampuan kita untuk bekerja sama sangatlah memprihatinkan, karena kerja sama adalah ciri utama spesies kita. "Saat AI merasuki kehidupan kita, kita harus menghadapi kemungkinan bahwa AI akan menghambat emosi kita dan menghambat hubungan mendalam antarmanusia, sehingga membuat hubungan kita satu sama lain menjadi kurang timbal balik, atau lebih dangkal, atau lebih narsistik," kata Christakis (2019).

Kepribadian

Ketika mesin semakin banyak mengambil tugas dan keputusan yang biasanya dilakukan oleh manusia, haruskah kita mempertimbangkan untuk memberikan 'kepribadian' dan lembaga moral atau hukum pada sistem AI? Salah satu cara memprogram sistem AI adalah 'pembelajaran penguatan', di mana peningkatan kinerja diperkuat dengan imbalan virtual. Bisakah kita menganggap suatu sistem menderita ketika fungsi imbalannya memberikan masukan negatif? Ketika kita menganggap mesin sebagai entitas yang dapat merasakan, merasakan, dan bertindak, maka kita tidak perlu lagi memikirkan status hukumnya. Haruskah mereka diperlakukan seperti hewan yang kecerdasannya sebanding? Akankah kita mempertimbangkan penderitaan mesin 'perasaan'?

Para sarjana semakin banyak membahas status hukum robot dan sistem AI selama tiga dekade terakhir. Namun, perdebatan tersebut muncul kembali baru-baru ini ketika resolusi parlemen Uni Eropa pada tahun 2017 mengundang Komisi Eropa 'untuk mengeksplorasi, menganalisis dan mempertimbangkan implikasi dari semua solusi hukum yang mungkin, [termasuk]...menciptakan status hukum khusus untuk robot dalam jangka panjang. dijalankan, sehingga setidaknya robot otonom paling canggih dapat ditetapkan memiliki status sebagai orang elektronik yang bertanggung jawab untuk memperbaiki segala kerusakan yang mungkin ditimbulkannya, dan mungkin menerapkan kepribadian elektronik pada kasus di mana robot membuat keputusan otonom atau berinteraksi dengan pihak ketiga secara independen.

Namun resolusi tersebut menimbulkan sejumlah keberatan, termasuk surat terbuka dari beberapa 'Pakar Kecerdasan Buatan dan Robotika' pada bulan April 2018 yang

menyatakan bahwa 'penciptaan Status Hukum 'orang elektronik' untuk 'otonom', 'tidak dapat diprediksi' dan robot 'belajar mandiri' harus dibuang dari sudut pandang teknis, hukum, dan etika. Mengaitkan kepribadian elektronik dengan robot berisiko menghilangkan tanggung jawab moral, akuntabilitas sebab akibat, dan tanggung jawab hukum terkait kesalahan dan penyalahgunaannya, kata surat itu.

Mayoritas penelitian etika mengenai AI tampaknya setuju bahwa mesin AI tidak boleh diberi hak moral, atau dilihat sebagai manusia. Bryson (2018) berargumen bahwa memberi robot hak moral dapat ditafsirkan sebagai tindakan tidak bermoral, karena 'tidak etis jika menempatkan artefak dalam situasi persaingan dengan kita, membuat mereka menderita, atau menjadikannya makhluk fana yang tidak perlu'. Dia melanjutkan dengan mengatakan itu

'Ada biaya yang besar tetapi sedikit atau tidak ada manfaat dari sudut pandang manusia atau robot jika menggunakan dan menerapkan agen atau kesabaran pada artefak cerdas di luar yang biasanya dianggap berasal dari kepemilikan apa pun. Oleh karena itu, tanggung jawab atas tindakan moral apa pun yang diambil oleh suatu artefak harus diserahkan kepada pemilik atau operatornya, atau jika terjadi malfungsi kepada produsennya, seperti halnya artefak konvensional.

2.3 DAMPAK TERHADAP SISTEM KEUANGAN

Salah satu domain pertama yang mulai menerapkan aplikasi otonom adalah pasar keuangan, dengan sebagian besar perkiraan menghubungkan lebih dari setengah volume perdagangan ekuitas AS dengan algoritma (Wellman dan Rajan, 2017).

Pasar sangat cocok untuk otomatisasi, karena sekarang hampir seluruhnya beroperasi secara elektronik, menghasilkan data dalam jumlah besar dengan kecepatan tinggi, yang memerlukan algoritma untuk mencernanya. Dinamisme pasar berarti bahwa respons yang tepat waktu terhadap informasi sangatlah penting, sehingga memberikan insentif yang kuat untuk mengeluarkan orang-orang yang lamban dari pengambilan keputusan. Yang terakhir, dan mungkin yang paling jelas, imbalan yang tersedia untuk keputusan perdagangan yang efektif cukup besar, sehingga menjelaskan mengapa perusahaan berinvestasi pada teknologi ini sebesar yang mereka miliki. Dengan kata lain, perdagangan algoritmik dapat menghasilkan keuntungan dengan kecepatan dan frekuensi yang tidak mungkin dilakukan oleh pedagang manusia.

Meskipun agen otonom saat ini beroperasi dalam lingkup kompetensi dan otonomi yang relatif sempit, mereka tetap mengambil tindakan yang mempunyai konsekuensi bagi masyarakat.

Contoh yang terkenal adalah Knight Capital Group. Selama 45 menit pertama hari perdagangan pada tanggal 1 Agustus 2012, saat memproses 212 pesanan kecil dari pelanggan, agen perdagangan otomatis yang dikembangkan oleh dan beroperasi atas nama Knight Capital secara keliru mengirimkan jutaan pesanan ke pasar ekuitas. Hasilnya, lebih dari empat juta transaksi dilakukan di pasar keuangan, menghasilkan miliaran dolar dalam posisi net long dan short. Perusahaan kehilangan Rp. 460 Miliar karena perdagangan yang tidak disengaja, dan nilai sahamnya sendiri turun hampir 75%.

Meskipun hal ini merupakan contoh kerugian yang tidak disengaja, agen perdagangan otonom juga dapat digunakan secara jahat untuk mengganggu stabilitas pasar, atau merugikan pihak yang tidak bersalah. Meskipun penggunaannya tidak dimaksudkan untuk tujuan jahat, otonomi dan kemampuan beradaptasi dari strategi perdagangan algoritmik, termasuk meningkatnya penggunaan teknik pembelajaran mesin yang canggih membuat sulit untuk memahami bagaimana kinerjanya dalam keadaan yang tidak terduga.

Manipulasi pasar

Raja dkk. (2019) membahas beberapa cara agen keuangan otonom dapat melakukan kejahatan keuangan, termasuk manipulasi pasar, yang didefinisikan sebagai 'tindakan dan/atau perdagangan oleh pelaku pasar yang berupaya mempengaruhi harga pasar secara artifisial' (Spatt, 2014).

Simulasi pasar yang terdiri dari agen perdagangan buatan telah menunjukkan bahwa, melalui pembelajaran penguatan, AI dapat mempelajari teknik spoofing buku pesanan, yang melibatkan penempatan pesanan tanpa niat untuk mengeksekusinya untuk memanipulasi peserta yang jujur di pasar (Lin, 2017).

Bot sosial juga terbukti mengeksploitasi pasar dengan menggelembungkan saham secara artifisial melalui promosi yang curang, sebelum menjual posisinya kepada pihak yang tidak menaruh curiga dengan harga yang melambung (Lin 2017). Misalnya, dalam kasus yang menonjol baru-baru ini, lingkup pengaruh jaringan bot sosial digunakan untuk menyebarkan disinformasi tentang perusahaan publik yang jarang diperdagangkan. Nilai perusahaan naik lebih dari 36.000% ketika saham pennynya melonjak dari kurang dari Rp. 1.000 menjadi di atas Rp. 200.000 per saham dalam waktu beberapa minggu (Ferrara 2015).

Kolusi

Penetapan harga, suatu bentuk kolusi juga dapat muncul dalam sistem otomatis. Karena agen perdagangan algoritmik dapat mempelajari informasi harga hampir secara instan, tindakan apa pun untuk menurunkan harga oleh satu agen kemungkinan besar akan diimbangi dengan agen lain secara instan. Hal ini bukanlah hal yang buruk dan hanya mewakili pasar yang efisien. Namun, kemungkinan penurunan harga akan mengakibatkan pesaing Anda secara bersamaan melakukan hal yang sama dapat menjadi disinsentif. Oleh karena itu, algoritma (jika rasional) akan mempertahankan harga yang lebih tinggi yang disepakati secara artifisial dan diam-diam, dengan tidak menurunkan harga terlebih dahulu (Ezrachi dan Stucke, 2016). Yang terpenting, agar kolusi dapat terjadi, suatu algoritma tidak perlu dirancang secara khusus untuk berkolusi.

Akuntabilitas

Meskipun tanggung jawab atas algoritme perdagangan terletak pada organisasi yang mengembangkan dan menerapkan algoritme tersebut, agen otonom dapat melakukan tindakan — terutama dalam keadaan yang tidak biasa — yang sulit diantisipasi oleh pemrogramnya. Apakah kesulitan tersebut mengurangi tanggung jawab sampai tingkat tertentu?

Misalnya, Wellman dan Rajan (2017) memberikan contoh agen perdagangan otonom yang melakukan operasi arbitrase, yaitu ketika pedagang memanfaatkan perbedaan harga

suatu aset untuk mencapai keuntungan yang hampir pasti. Secara teoritis, agen dapat mencoba untuk memicu peluang arbitrase dengan melakukan tindakan jahat untuk menumbangkan pasar, misalnya dengan menyebarkan informasi yang salah, mendapatkan akses yang tidak tepat terhadap informasi, atau melakukan pelanggaran langsung terhadap aturan pasar.

Jelasnya, akan merugikan bagi agen perdagangan otonom untuk terlibat dalam manipulasi pasar, namun bisakah algoritma otonom memenuhi definisi hukum manipulasi pasar, yang memerlukan 'niat'?

Wellmen dan Rajan (2017) berpendapat bahwa agen perdagangan akan semakin mampu beroperasi pada tingkat yang lebih luas tanpa pengawasan manusia, dan peraturan kini diperlukan untuk mencegah kerugian sosial. Namun, upaya untuk mengatur atau membuat undang-undang mungkin terhambat oleh beberapa masalah.

2.4 DAMPAK TERHADAP SISTEM HUKUM

Penciptaan mesin AI dan penggunaannya di masyarakat dapat berdampak besar pada hukum pidana dan perdata. Seluruh sejarah hukum manusia dibangun berdasarkan asumsi bahwa manusia, dan bukan robot, yang mengambil keputusan. Dalam masyarakat di mana keputusan-keputusan yang semakin rumit dan penting diserahkan kepada algoritma, terdapat risiko bahwa kerangka hukum yang kita miliki untuk bertanggung jawab tidak akan mencukupi.

Bisa dibilang, pertanyaan hukum jangka pendek yang paling penting terkait AI adalah siapa atau apa yang harus bertanggung jawab atas pelanggaran yang merugikan, kriminal, dan kontrak yang melibatkan AI dan dalam kondisi apa.

Hukum pidana

Suatu kejahatan terdiri atas dua unsur, yaitu perbuatan pidana yang disengaja atau kelalaian (*actus reus*) dan kesengajaan untuk melakukan kejahatan (*mens rea*). Jika robot terbukti memiliki kesadaran yang cukup, maka mereka dapat bertanggung jawab sebagai pelaku langsung tindak pidana, atau bertanggung jawab atas kejahatan kelalaian. Jika kita mengakui bahwa robot mempunyai pikirannya sendiri, diberkahi dengan keinginan bebas, otonomi, atau kesadaran moral seperti manusia, maka seluruh sistem hukum kita harus diubah secara drastis. Meskipun hal ini mungkin terjadi, namun hal tersebut tidak mungkin terjadi. Namun demikian, robot dapat mempengaruhi hukum pidana dengan cara yang lebih halus.

Beban

Meningkatnya pendelegasian pengambilan keputusan kepada AI juga akan berdampak pada banyak bidang hukum yang memerlukan *mens rea*, atau niat, untuk melakukan kejahatan.

Apa yang akan terjadi, misalnya jika program AI yang dipilih untuk memprediksi kesuksesan investasi dan mengikuti tren pasar melakukan evaluasi yang salah sehingga menyebabkan kurangnya peningkatan modal dan karenanya menyebabkan kebangkrutan perusahaan karena penipuan? Karena persyaratan niat penipuan tidak ada, manusia hanya

dapat bertanggung jawab atas kejahatan kebangkrutan yang lebih kecil yang dipicu oleh evaluasi robot (Pagallo, 2017).

Model pertanggungjawaban yang ada mungkin tidak memadai untuk mengatasi peran AI dalam aktivitas kriminal di masa depan (King et al, 2019). Misalnya, dalam istilah *actus reus*, meskipun agen otonom dapat melakukan tindak pidana atau kelalaian, aspek sukarela dari *actus reus* tidak akan terpenuhi, karena gagasan bahwa agen otonom dapat bertindak secara sukarela masih kontroversial. Hal ini berarti bahwa para agen, baik yang dibuat-buat maupun yang lainnya, berpotensi melakukan tindak pidana atau kelalaian tanpa memenuhi persyaratan pertanggungjawaban atas tindak pidana tersebut.

Ketika pertanggungjawaban pidana didasarkan pada kesalahan, maka hal tersebut juga memerlukan *mens rea* (pikiran yang bersalah). *Mens rea* dapat terdiri dari niat untuk melakukan *actus reus* menggunakan aplikasi berbasis AI, atau pengetahuan bahwa menggunakan agen otonom akan atau dapat menyebabkannya melakukan tindakan kriminal atau kelalaian. Namun, dalam beberapa kasus, kompleksitas pemrograman agen otonom memungkinkan perancang, pengembang, atau penyebar tidak mengetahui atau memprediksi tindakan kriminal atau kelalaian AI. Hal ini memberikan insentif besar bagi agen manusia untuk menghindari mencari tahu apa sebenarnya yang dilakukan sistem pembelajaran mesin, karena semakin sedikit yang diketahui agen manusia, semakin besar kemungkinan mereka menolak tanggung jawab atas kedua alasan tersebut (Williams 2017).

Tindakan robot otonom juga dapat mengarah pada situasi di mana manusia mewujudkan *mens rea*, dan robot melakukan *actus reus*, sehingga memecah komponen kejahatan (McAllister 2017).

Alternatifnya, pembuat undang-undang dapat mendefinisikan pertanggungjawaban pidana tanpa persyaratan kesalahan. Hal ini akan mengakibatkan tanggung jawab dibebankan kepada orang yang menyebarkan AI terlepas dari apakah mereka mengetahuinya, atau dapat memprediksi perilaku ilegal tersebut. Tanggung jawab tanpa kesalahan semakin banyak digunakan untuk tanggung jawab produk dalam hukum perbuatan melawan hukum (misalnya, obat-obatan dan barang konsumsi). Namun, Williams (2017) berpendapat bahwa *mens rea* dengan niat atau pengetahuan itu penting, dan kita tidak bisa mengabaikan persyaratan utama pertanggungjawaban pidana tersebut begitu saja karena sulitnya membuktikannya.

Kingston (2018) merujuk pada definisi yang diberikan oleh Hallevy (2010) tentang bagaimana tindakan AI dapat dipandang berdasarkan hukum pidana. Menurut Hallevy, model hukum ini dapat dibagi menjadi tiga skenario:

1. **Pelaku lewat-yang lain.** Jika suatu pelanggaran dilakukan oleh suatu entitas yang tidak memiliki kapasitas mental *mens rea* – anak-anak, hewan, atau orang yang mengalami gangguan mental – maka mereka dianggap sebagai agen yang tidak bersalah. Namun, jika agen yang tidak bersalah ini diinstruksikan oleh orang lain untuk melakukan kejahatan tersebut, maka instruktur tersebut dianggap bertanggung jawab secara pidana. Dalam model ini, AI dapat dianggap sebagai agen yang tidak bersalah, dengan pemrogram perangkat lunak atau pengguna berperan sebagai pelaku melalui orang lain.

2. **Konsekuensi yang mungkin terjadi secara alami.** Hal ini berkaitan dengan kaki tangan suatu tindak pidana; jika tidak ada konspirasi yang dapat dibuktikan, seorang kaki tangan masih dapat dimintai pertanggungjawaban secara hukum jika tindakan pelaku merupakan konsekuensi alami atau kemungkinan besar dari suatu skema yang didorong atau dibantu oleh seorang kaki tangan. Skenario ini mungkin terjadi ketika AI yang dirancang untuk tujuan 'baik' disalahgunakan dan melakukan kejahatan. Misalnya, robot di pabrik mungkin melukai pekerja di dekatnya yang secara keliru mereka anggap sebagai ancaman terhadap misi terprogram mereka. Dalam kasus ini, pemrogram dapat dimintai pertanggungjawaban sebagai kaki tangan jika mereka mengetahui bahwa pelanggaran pidana merupakan konsekuensi alami atau kemungkinan besar dari rancangan atau penggunaan program mereka. Hal ini tidak berlaku untuk AI yang diprogram untuk melakukan hal yang 'buruk', namun untuk AI yang disalahgunakan. Siapa pun yang mampu dan mungkin memperkirakan AI digunakan untuk tujuan kriminal tertentu dapat dianggap bertanggung jawab berdasarkan skenario ini: pemrogram, vendor, penyedia layanan, atau pengguna (dengan asumsi bahwa batasan sistem dan kemungkinan konsekuensi penyalahgunaan telah disebutkan). dalam instruksi AI – yang kemungkinannya kecil).
3. **Tanggung jawab langsung.** Model ini menghubungkan *actus* dan *mens rea* dengan AI. Namun, meskipun *actus rea* (tindakan atau kelambanan) relatif mudah untuk dikaitkan dengan AI, kata Kingston (2018), menghubungkan *mens rea* (pikiran bersalah) lebih kompleks. Misalnya, program AI yang 'mengendarai' kendaraan otonom yang melebihi batas kecepatan dapat dianggap bertanggung jawab secara pidana karena ngebut – namun untuk skenario tanggung jawab ketat seperti ini, tidak diperlukan niat pidana, dan tidak perlu membuktikan bahwa mobil tersebut melaju dengan sadar. Kingston juga menandai sejumlah kemungkinan masalah yang muncul ketika mempertimbangkan AI untuk bertanggung jawab secara langsung. Misalnya, bisakah AI yang terinfeksi virus mengklaim pertahanan yang mirip dengan pemaksaan atau keracunan, atau AI yang tidak berfungsi mengklaim pertahanan yang mirip dengan kegilaan? Seperti apa hukumannya – dan siapa yang akan dihukum?

Mengidentifikasi siapa sebenarnya yang bertanggung jawab atas tindakan AI memang penting, namun juga berpotensi sulit. Misalnya, 'pemrogram' dapat diterapkan pada banyak kolaborator, atau diperluas hingga mencakup peran seperti perancang program, pakar produk, dan atasan mereka – dan kesalahannya mungkin terletak pada manajer yang menunjuk pakar atau pemrogram yang tidak memadai (Kingston, 2010).

Psikologi

Ada risiko robot AI dapat memanipulasi kondisi mental pengguna untuk melakukan kejahatan. Hal ini ditunjukkan oleh Weizenbaum (1976) yang melakukan eksperimen awal pada interaksi manusia-bot di mana orang mengungkapkan detail pribadi yang tidak terduga tentang kehidupan mereka. Robot juga dapat menormalkan pelanggaran seksual dan kejahatan terhadap manusia, seperti kasus robot seks tertentu (De Angeli, 2009).

Perdagangan, pasar keuangan dan kebangkrutan

Sebagaimana dibahas sebelumnya dalam buku ini, terdapat kekhawatiran bahwa agen otonom di sektor keuangan dapat terlibat dalam manipulasi pasar, penetapan harga, dan kolusi. Kurangnya niat dari agen manusia, dan kemungkinan bahwa agen otonom (AA) dapat bertindak bersama juga menimbulkan masalah serius sehubungan dengan tanggung jawab dan pemantauan. Akan sulit untuk membuktikan bahwa agen manusia bermaksud agar AA memanipulasi pasar, dan juga akan sulit untuk memantau manipulasi tersebut. Kemampuan AA untuk mempelajari dan menyempurnakan kemampuan mereka juga menyiratkan bahwa agen-agen ini mungkin mengembangkan strategi baru, sehingga semakin sulit untuk mendeteksi tindakan mereka (Farmer dan Skouras 2013).

Obat Berbahaya atau Berbahaya

Di masa depan, AI dapat digunakan oleh geng kriminal terorganisir untuk mendukung perdagangan dan penjualan zat terlarang. Penjahat dapat menggunakan kendaraan tak berawak yang dilengkapi AI dan teknologi navigasi otonom untuk menyelundupkan zat terlarang. Karena jaringan penyelundupan terganggu oleh pengawasan dan intersepsi jalur transportasi, penegakan hukum menjadi lebih sulit ketika kendaraan tak berawak digunakan untuk mengangkut barang selundupan. Menurut Europol (2017), drone menghadirkan ancaman nyata dalam bentuk penyelundupan narkoba secara otomatis. Kapal selam penyelundup kokain yang dikendalikan dari jarak jauh telah ditemukan dan disita oleh penegak hukum AS (Sharkey dkk., 2010).

Kendaraan bawah air tak berawak (UUV) juga dapat digunakan untuk kegiatan ilegal, sehingga menimbulkan ancaman signifikan terhadap penegakan larangan narkoba. Karena UUV dapat bertindak secara independen dan tidak bergantung pada operatornya, maka akan lebih sulit untuk menangkap pelaku kejahatan yang terlibat.

Bot sosial juga dapat digunakan untuk mengiklankan dan menjual pornografi atau narkoba kepada jutaan orang secara online, termasuk anak-anak.

Pelanggaran Terhadap Orang tersebut

Bot sosial juga dapat digunakan untuk melecehkan orang. Kini, ketika AI dapat menghasilkan konten palsu yang lebih canggih, bentuk-bentuk pelecehan baru mungkin terjadi. Baru-baru ini, pengembang merilis perangkat lunak yang menghasilkan video sintesis yang dapat secara akurat menggantikan wajah seseorang dengan wajah orang lain. Banyak dari video sintesis ini bersifat pornografi dan kini terdapat risiko bahwa pengguna jahat akan membuat konten palsu untuk melecehkan korbannya.

Robot AI juga dapat digunakan untuk menyiksa dan menginterogasi orang, menggunakan teknik penyiksaan psikologis (misalnya meniru orang yang dikenal sebagai subjek penyiksaan) atau fisik (McAllister 2017). Karena robot tidak dapat memahami rasa sakit atau merasakan empati, mereka tidak akan menunjukkan belas kasihan atau kasih sayang. Oleh karena itu, kehadiran robot interogasi saja dapat menyebabkan subjek berbicara karena takut. Penggunaan robot juga akan menjauhkan pelaku manusia dari tindakan *actus reus*, dan secara emosional menjauhkan diri mereka dari kejahatan yang dilakukan, sehingga penyiksaan lebih mungkin terjadi.

Sebagai mesin yang tidak berpikir, AA tidak dapat memikul tanggung jawab moral atau kewajiban atas tindakan mereka. Namun, salah satu solusinya adalah dengan mengambil pendekatan pertanggungjawaban pidana yang ketat, di mana hukuman atau ganti rugi dapat dijatuhkan tanpa bukti kesalahan, yang akan menurunkan ambang batas niat untuk melakukan kejahatan. Namun bahkan di bawah kerangka pertanggungjawaban yang ketat, pertanyaan mengenai siapa sebenarnya yang harus dijatuhi hukuman penjara karena pelanggaran yang disebabkan oleh AI terhadap seseorang masih sulit dilakukan. Jelas bahwa AA tidak dapat dimintai pertanggungjawaban. Namun, banyaknya aktor yang terlibat menimbulkan masalah dalam memastikan di mana letak tanggung jawabnya—apakah pihak yang menugaskan dan mengoperasikan AA, atau pengembangnya, atau pembuat undang-undang dan pembuat kebijakan yang menyetujui penerapan agen-agen tersebut di dunia nyata.

Pelanggaran Seksual

Ada bahaya bahwa robot yang mengandung AI dapat digunakan untuk mempromosikan obyektifikasi seksual, pelecehan seksual, dan kekerasan. Seperti yang dibahas di bagian 2.1, sexbots memungkinkan orang melakukan simulasi pelanggaran seksual seperti fantasi pemerkosaan. Tindakan-tindakan tersebut bahkan dapat dirancang untuk meniru pelanggaran seksual, seperti pemerkosaan terhadap orang dewasa dan anak-anak. Interaksi dengan bot sosial dan robot seks juga dapat menurunkan kepekaan pelaku terhadap pelanggaran seksual, atau bahkan meningkatkan keinginan mereka untuk melakukan pelanggaran seksual (De Angeli 2009; Danaher 2017).

Siapa yang bertanggung jawab?

Ketika mempertimbangkan kemungkinan konsekuensi dan penyalahgunaan AI, pertanyaan kuncinya adalah: siapa yang bertanggung jawab atas tindakan AI? Apakah pemrogram, produsen, pengguna akhir, AI itu sendiri, atau lainnya? Apakah jawaban atas pertanyaan ini sama untuk semua AI atau mungkin berbeda, misalnya, untuk sistem yang mampu mempelajari dan mengadaptasi perilakunya?

Menurut Resolusi Parlemen Eropa (2017) tentang AI, tanggung jawab hukum atas tindakan (atau kelambanan) AI biasanya dibebankan pada aktor manusia: misalnya pemilik, pengembang, produsen, atau operator AI. Misalnya, mobil self-driving di Jerman saat ini dianggap sebagai tanggung jawab pemiliknya. Namun, masalah muncul ketika mempertimbangkan keterlibatan pihak ketiga, dan sistem canggih seperti jaringan saraf yang belajar mandiri: jika suatu tindakan tidak dapat diprediksi oleh pengembang karena AI telah cukup berubah dari desainnya, dapatkah pengembang bertanggung jawab atas tindakan tersebut? ? Selain itu, infrastruktur legislatif yang ada saat ini dan kurangnya mekanisme peraturan yang efektif menimbulkan tantangan dalam mengatur AI dan menyalahkan orang lain, kata Atabekov dan Yastrebov (2018), dan AI otonom khususnya menimbulkan pertanyaan apakah diperlukan kategori hukum baru untuk mencakup fitur-fiturnya. dan batasannya (Parlemen Eropa, 2017).

Taddeo dan Floridi (2018) menyoroti konsep 'agensis terdistribusi'. Ketika tindakan atau keputusan AI terjadi melalui rantai interaksi yang panjang dan kompleks antara manusia dan

robot – mulai dari pengembang dan perancang hingga produsen, vendor, dan pengguna, masing-masing dengan motivasi, latar belakang, dan pengetahuan yang berbeda – maka dapat dikatakan bahwa hasil AI dapat dikatakan. menjadi hasil dari agensi terdistribusi. Dengan agen terdistribusi datangnya tanggung jawab terdistribusi. Salah satu cara untuk memastikan bahwa AI berfungsi untuk 'mencegah kejahatan dan mendorong kebaikan' di masyarakat adalah dengan menerapkan kerangka moral tanggung jawab terdistribusi yang membuat semua agen bertanggung jawab atas peran mereka dalam hasil dan tindakan AI (Taddeo dan Floridi, 2018).

Penerapan AI yang berbeda mungkin memerlukan kerangka kerja yang berbeda. Misalnya, terkait robot militer, Lokhorst dan van den Hoven (2014) menyatakan bahwa tanggung jawab utama terletak pada perancang dan pelaksana robot, namun robot mungkin dapat memikul tanggung jawab pada tingkat tertentu atas tindakannya.

Mesin pembelajaran dan AI otonom adalah contoh penting lainnya. Penggunaannya dapat menciptakan 'kesenjangan tanggung jawab', kata Matthias (2004), di mana produsen atau operator mesin, pada prinsipnya, tidak dapat memprediksi perilaku AI di masa depan – dan dengan demikian tidak dapat bertanggung jawab atas hal tersebut dalam jangka waktu yang lama. pengertian hukum atau moral. Matthias mengusulkan agar pemrogram jaringan saraf, misalnya, semakin menjadi 'pencipta organisme perangkat lunak', dengan sangat sedikit kendali di luar titik pengkodean. Perilaku AI tersebut menyimpang dari pemrograman awal dan menjadi produk interaksinya dengan lingkungannya – perbedaan yang jelas antara fase pemrograman, pelatihan, dan pengoperasian mungkin hilang, sehingga membuat kesalahan menjadi sangat rumit dan tidak jelas. Kesenjangan tanggung jawab ini memerlukan pengembangan dan klarifikasi praktik moral dan undang-undang yang sesuai bersamaan dengan penerapan automata pembelajaran (Matthias, 2004). Hal ini juga sejalan dengan Scherer (2016), yang menyatakan bahwa AI sejauh ini dikembangkan dalam 'kekosongan peraturan', dengan hanya sedikit undang-undang atau peraturan yang dirancang untuk secara eksplisit mengatasi tantangan unik dan tanggung jawab AI.

Pencurian dan penipuan, serta pemalsuan dan peniruan identitas

AI dapat digunakan untuk mengumpulkan data pribadi dan memalsukan identitas seseorang. Misalnya, bot media sosial yang menambahkan orang sebagai 'teman' akan mendapatkan akses ke informasi pribadi, lokasi, nomor telepon, atau riwayat hubungan mereka (Bilge et al., 2009). AI dapat memanipulasi orang dengan membangun hubungan baik dengan mereka, kemudian mengeksploitasi hubungan tersebut untuk mendapatkan informasi dari atau akses ke komputer mereka (Chantler dan Broadhurst 2006).

AI juga dapat digunakan untuk melakukan penipuan perbankan dengan memalsukan identitas korban, termasuk meniru suara seseorang. Dengan menggunakan kemampuan pembelajaran mesin, perangkat lunak Adobe mampu mempelajari dan mereproduksi pola bicara individu dari rekaman suara orang tersebut selama 20 menit. Menyalin suara pelanggan dapat memungkinkan penjahat untuk berbicara dengan bank orang tersebut dan melakukan transaksi.

Hukum perbuatan melawan hukum

Hukum tort mencakup situasi di mana perilaku seseorang menyebabkan cedera, penderitaan, kerugian yang tidak adil, atau kerugian bagi orang lain. Ini adalah kategori hukum yang luas yang dapat mencakup berbagai jenis klaim cedera pribadi. Undang-undang perbuatan melawan hukum memiliki dua tujuan dasar dan umum:

- 1) Memberikan kompensasi kepada korban atas segala kerugian yang disebabkan oleh pelanggaran yang dilakukan terdakwa; dan
- 2) Untuk mencegah terdakwa mengulangi pelanggaranannya di kemudian hari.

Undang-undang tort kemungkinan akan menjadi fokus utama dalam beberapa tahun ke depan seiring munculnya mobil self-driving di jalan umum. Dalam kasus mobil otonom yang dapat mengemudi sendiri, ketika terjadi kecelakaan, ada dua bidang hukum yang relevan - kelalaian dan tanggung jawab produk.

Saat ini sebagian besar kecelakaan diakibatkan oleh kesalahan pengemudi, yang berarti tanggung jawab atas kecelakaan diatur oleh prinsip kelalaian. Kelalaian adalah doktrin yang membuat orang bertanggung jawab atas tindakan yang tidak wajar dalam situasi tersebut (Anderson dkk, 2009). Untuk membuktikan tuntutan kelalaian, penggugat harus menunjukkan bahwa:

- Kewajiban kehati-hatian merupakan kewajiban tergugat kepada penggugat
- Telah terjadi pelanggaran terhadap kewajiban tersebut oleh terdakwa
- Terdapat hubungan sebab akibat antara pelanggaran kewajiban tergugat dan kerugian yang dialami penggugat, dan;
- Bahwa penggugat menderita kerugian sebagai akibatnya.

Biasanya perusahaan asuransi menentukan pihak yang bersalah, sehingga menghindari tuntutan hukum yang mahal. Namun hal ini menjadi lebih rumit jika ada cacat pada kendaraan yang menyebabkan kecelakaan. Dalam kasus mobil yang dapat mengemudi sendiri, kecelakaan dapat disebabkan oleh kegagalan perangkat keras, kegagalan desain, atau kesalahan perangkat lunak – yaitu cacat pada algoritma komputer.

Saat ini, jika tabrakan disebabkan oleh kesalahan atau cacat pada program komputer, produsen akan bertanggung jawab berdasarkan doktrin Kewajiban Produk, yang mengharuskan produsen, distributor, pemasok, pengecer, dan pihak lain yang menyediakan produk untuk publik bertanggung jawab atas hal tersebut. Cedera yang disebabkan oleh produk tersebut.

Karena sebagian besar tabrakan kendaraan otonom diperkirakan disebabkan oleh kesalahan perangkat lunak, maka cacat tersebut kemungkinan besar harus lulus 'uji utilitas risiko' (Anderson et al., 2010), yaitu suatu produk cacat jika terdapat risiko bahaya yang dapat diperkirakan sebelumnya. Oleh produk dapat dikurangi atau dihindari dengan penerapan desain alternatif yang masuk akal oleh penjual, dan pengabaian desain alternatif menjadikan produk tersebut tidak cukup aman.

Namun, kasus uji utilitas risiko, yang diperlukan untuk membuktikan cacat desain, bersifat kompleks dan memerlukan banyak saksi ahli, sehingga membuat klaim cacat desain menjadi mahal untuk dibuktikan (Gurney dkk, 2013). Sifat bukti, seperti algoritma yang

kompleks dan data sensor juga cenderung membuat litigasi menjadi sangat menantang dan kompleks.

Ini berarti metode yang digunakan untuk memulihkan ganti rugi atas kecelakaan mobil harus beralih dari bidang hukum yang sudah mapan dan langsung menjadi bidang hukum yang rumit dan mahal (tanggung jawab produk). Penggugat memerlukan banyak ahli untuk memulihkan dan menemukan cacat pada algoritme, yang akan berdampak pada kecelakaan kendaraan otonom yang paling sederhana sekalipun. Hal ini kemungkinan besar akan mempengaruhi kemampuan korban untuk mendapatkan kompensasi dan ganti rugi atas cedera yang dideritanya dalam kecelakaan mobil.

2.5 DAMPAK TERHADAP LINGKUNGAN DAN PLANET INI

Teknologi AI dan robotika memerlukan daya komputasi yang besar, yang disertai dengan biaya energi. Bisakah kita mempertahankan pertumbuhan besar-besaran AI dari sudut pandang energik ketika kita dihadapkan pada perubahan iklim yang belum pernah terjadi sebelumnya?

Pemanfaatan sumber daya alam

Ekstraksi nikel, kobalt, dan grafit untuk digunakan dalam baterai lithium ion – yang biasa ditemukan pada mobil listrik dan ponsel pintar – telah merusak lingkungan, dan AI kemungkinan akan meningkatkan permintaan ini. Ketika pasokan yang ada berkurang, operator mungkin terpaksa bekerja di lingkungan yang lebih kompleks dan berbahaya bagi operator manusia – sehingga menyebabkan otomatisasi penambangan dan ekstraksi logam lebih lanjut. Hal ini akan meningkatkan hasil dan laju penipisan logam tanah jarang, sehingga semakin memperburuk lingkungan.

Polusi dan limbah

Pada akhir siklus produknya, barang-barang elektronik biasanya dibuang, sehingga menyebabkan penumpukan logam berat dan bahan beracun di lingkungan. Meningkatkan produksi dan konsumsi perangkat teknologi seperti robot akan memperburuk masalah limbah ini, terutama karena perangkat tersebut kemungkinan besar dirancang dengan 'keusangan bawaan' – sebuah proses di mana produk dirancang untuk menjadi usang 'sebelum waktunya' sehingga pelanggan harus membeli penggantinya. barang – menghasilkan limbah elektronik dalam jumlah besar. Keusangan yang direncanakan menghabiskan sumber daya alam seperti logam tanah jarang, sekaligus meningkatkan jumlah limbah. Sumber menunjukkan bahwa di Amerika Utara, lebih dari 100 juta ponsel dan 300 juta komputer pribadi dibuang setiap tahunnya (Guilina et al., 2009).

Cara untuk mengatasi hal ini adalah dengan 'mendorong konsumen untuk memilih produk dan jasa yang ramah lingkungan dan lebih berkelanjutan' (Dewan Bisnis Dunia untuk Pembangunan Berkelanjutan, 2000). Namun, hal ini terhambat karena konsumen mengharapkan peningkatan yang sering, dan kurangnya kepedulian konsumen terhadap dampak lingkungan ketika mempertimbangkan peningkatan.

Masalah energi

Selain dampak buruk dari peningkatan penambangan dan limbah terhadap lingkungan, penerapan teknologi AI, khususnya pembelajaran mesin, akan membutuhkan lebih banyak data untuk diproses. Dan itu memerlukan energi yang sangat besar. Di Amerika Serikat, pusat data sudah menyumbang sekitar 2 persen dari seluruh listrik yang digunakan. Menurut perkiraan, AlphaGo milik DeepMind – yang mengalahkan Juara Go Lee Sedol pada tahun 2016 – memerlukan kekuatan 50.000 kali lebih besar daripada otak manusia (Mattheij, 2016).

AI juga memerlukan energi dalam jumlah besar untuk produksi dan pelatihan – misalnya, diperlukan waktu berjam-jam untuk melatih model AI berskala besar agar dapat memahami dan mengenali bahasa manusia sehingga dapat digunakan untuk tujuan penerjemahan. Menurut Strubell, Ganesh, dan McCallum (2019), jejak karbon dari pelatihan, penyediaan, dan eksperimen dengan pemrosesan bahasa alami AI adalah lebih dari tujuh kali lipat dari rata-rata manusia dalam satu tahun, dan kira-kira 1,5 kali lipat dari jejak karbon dari sebuah rata-rata mobil, termasuk bahan bakar, sepanjang masa pakainya.

Cara AI dapat membantu planet ini

Alternatifnya, AI sebenarnya bisa membantu kita merawat planet ini dengan lebih baik, dengan membantu kita mengelola limbah dan polusi. Misalnya, penerapan kendaraan otonom dapat mengurangi emisi gas rumah kaca, karena kendaraan otonom dapat diprogram untuk mengikuti prinsip-prinsip mengemudi ramah lingkungan sepanjang perjalanan, mengurangi konsumsi bahan bakar sebanyak 20 persen dan mengurangi emisi gas rumah kaca pada tingkat yang sama. (Iglinski dkk., 2017). Kendaraan otonom juga dapat mengurangi kemacetan lalu lintas dengan merekomendasikan rute alternatif dan rute terpendek, dan dengan berbagi informasi lalu lintas kepada kendaraan lain di jalan raya, sehingga konsumsi bahan bakar lebih sedikit.

Ada juga penerapan AI dalam lingkungan konservasi. Misalnya, teknologi pembelajaran mendalam dapat digunakan untuk menganalisis gambar hewan yang ditangkap oleh kamera sensor gerak di alam liar. Informasi ini kemudian dapat digunakan untuk memberikan informasi yang akurat, terperinci, dan terkini tentang lokasi, jumlah, dan perilaku satwa di alam liar, yang dapat berguna dalam meningkatkan keanekaragaman hayati lokal dan upaya konservasi lokal (Norouzzadeh et al., 2018).

2.6 DAMPAK TERHADAP KEPERCAYAAN

AI diatur untuk mengubah kehidupan kita sehari-hari di berbagai bidang seperti transportasi; industri jasa; kesehatan; pendidikan; keselamatan dan keamanan masyarakat; dan hiburan. Meskipun demikian, sistem ini harus diterapkan dengan cara yang membangun kepercayaan dan pemahaman, serta menghormati hak asasi manusia dan hak sipil (Dignum, 2018). Mereka harus mengikuti prinsip-prinsip dan nilai-nilai dasar kemanusiaan, serta menjaga kesejahteraan manusia dan planet ini.

Konsensus yang sangat besar di kalangan komunitas riset adalah bahwa kepercayaan terhadap AI hanya dapat dicapai melalui keadilan, transparansi, akuntabilitas, dan regulasi. Masalah lain yang berdampak pada kepercayaan adalah seberapa besar kendali yang ingin kita

berikan terhadap mesin AI, dan apakah, misalnya, kita ingin selalu menjaga hubungan dengan manusia, atau memberikan otonomi yang lebih besar pada sistem.

Meskipun robot dan AI sebagian besar dipandang positif oleh masyarakat di seluruh Eropa, robot dan AI juga menimbulkan perasaan campur aduk, menimbulkan kekhawatiran dan kegelisahan (Komisi Eropa 2012; Komisi Eropa 2017). Dua survei Eurobarometer, yang bertujuan untuk mengukur persepsi, penerimaan, dan opini masyarakat mengenai topik tertentu di kalangan warga negara UE di Negara-negara Anggota, telah dilakukan untuk mengkarakterisasi sikap masyarakat terhadap robot dan AI (survei 382), dan terhadap peningkatan digitalisasi dan otomatisasi (survei 460).

Survei-survei ini menunjukkan bahwa masih ada jalan yang harus ditempuh sebelum masyarakat merasa nyaman dengan meluasnya penggunaan robot dan teknologi canggih di masyarakat. Misalnya, meskipun responden menyukai gagasan untuk memprioritaskan penggunaan robot di area yang menimbulkan risiko atau kesulitan bagi manusia – misalnya eksplorasi ruang angkasa, manufaktur, militer, keamanan, serta pencarian dan penyelamatan – mereka merasa sangat tidak nyaman dengan area yang melibatkan area rentan atau berbahaya. wilayah ketergantungan masyarakat. Responden menentang penggunaan robot untuk merawat anak-anak, orang lanjut usia, dan orang cacat; untuk pendidikan; dan untuk layanan kesehatan, meskipun banyak yang memiliki pandangan positif terhadap robot secara umum. Mayoritas dari mereka yang disurvei juga 'sangat tidak nyaman' dengan gagasan membiarkan anjing mereka berjalan-jalan dengan robot, menjalani operasi medis yang dilakukan oleh robot, atau membiarkan anak-anak atau orang tua lanjut usia mereka diawasi oleh robot – sebuah skenario di mana kepercayaan adalah kuncinya.

Mengapa kepercayaan itu penting

'Agar AI dapat mencapai potensi maksimalnya, kita harus membiarkan mesin terkadang bekerja secara mandiri, dan mengambil keputusan sendiri tanpa campur tangan manusia', jelas Taddeo (2017).

Bayangkan sebuah masyarakat dimana tidak ada kepercayaan pada dokter, guru, atau pengemudi. Tanpa kepercayaan, kita harus menghabiskan sebagian besar hidup kita untuk mencurahkan waktu dan sumber daya untuk memastikan orang lain, atau segala sesuatunya, melakukan tugasnya dengan baik. Pengawasan ini akan mengorbankan pekerjaan kita sendiri, dan pada akhirnya akan menciptakan masyarakat yang tidak berfungsi.

'Kami memercayai algoritme pembelajaran mesin untuk menunjukkan keputusan terbaik yang harus diambil saat merekrut calon kolega atau saat memberikan pembebasan bersyarat selama persidangan pidana; untuk mendiagnosis penyakit dan mengidentifikasi kemungkinan penyembuhannya. Kita memercayai robot untuk merawat orang lanjut usia dan balita, berpatroli di perbatasan, dan mengemudikan atau menerbangkan kita keliling dunia. Kami bahkan memercayai teknologi digital untuk mensimulasikan eksperimen dan memberikan hasil yang memajukan pengetahuan ilmiah dan pemahaman kami tentang dunia. Kepercayaan ini tersebar luas dan tangguh. Hal ini hanya dinilai ulang (jarang dilanggar) jika terjadi konsekuensi negatif yang serius.' (Taddeo, 2017)

Faktanya, teknologi digital begitu luas sehingga memercayainya sangatlah penting agar masyarakat kita dapat berfungsi dengan baik. Mengawasi algoritme pembelajaran mesin yang digunakan untuk mengambil keputusan secara terus-menerus akan memerlukan banyak waktu dan sumber daya, sehingga penggunaan teknologi digital menjadi tidak mungkin dilakukan. Namun, pada saat yang sama, tugas-tugas yang kita percayai pada teknologi digital sangatlah relevan sehingga kurangnya pengawasan dapat menimbulkan risiko serius terhadap keselamatan dan keamanan kita, serta terhadap hak-hak dan nilai-nilai yang mendasari masyarakat kita.

Dengan kata lain, penting untuk mengidentifikasi cara efektif untuk memercayai teknologi digital sehingga kita dapat memanfaatkan manfaatnya, sekaligus melindungi hak-hak dasar dan mendorong pengembangan masyarakat informasi yang terbuka, toleran, dan adil. Hal ini sangat penting dalam sistem hibrida yang melibatkan agen manusia dan buatan.

Namun bagaimana kita menemukan tingkat kepercayaan yang tepat? Taddeo berpendapat bahwa dalam jangka pendek, desain dapat memainkan peran penting dalam mengatasi masalah ini. Misalnya, pesan pop-up yang mengingatkan pengguna akan hasil mesin pencari algoritmik yang memperhitungkan profil online pengguna, atau pesan yang menandai bahwa hasil suatu algoritma mungkin tidak objektif. Namun dalam jangka panjang, diperlukan infrastruktur yang menegakkan norma-norma seperti keadilan, transparansi, dan akuntabilitas di semua sektor.

Keadilan

Untuk memercayai AI, AI harus adil dan tidak memihak. Seiring dengan semakin banyaknya keputusan yang didelegasikan kepada AI, kita harus memastikan bahwa keputusan tersebut bebas dari bias dan diskriminasi. Baik itu menyaring CV untuk wawancara kerja, memutuskan penerimaan ke universitas, melakukan pemeringkatan kredit untuk perusahaan pinjaman, atau menilai risiko seseorang melakukan pelanggaran lagi, keputusan yang dibuat oleh AI harus adil dan tidak memperdalam kesenjangan sosial yang sudah mengakar.

Namun bagaimana kita membuat algoritma menjadi adil? Ini tidak semudah kelihatannya. Masalahnya adalah tidak mungkin mengetahui algoritme apa yang sebenarnya dipelajari berdasarkan jaringan saraf saat Anda melatihnya dengan data. Misalnya, algoritme COMPAS, yang menilai seberapa besar kemungkinan seseorang melakukan kejahatan dengan kekerasan ternyata sangat mendiskriminasi orang kulit hitam. Namun algoritme tersebut sebenarnya tidak memberikan masukan pada ras manusia. Sebaliknya, algoritme menyimpulkan data sensitif ini dari informasi lain, misalnya, alamat.

Misalnya, sebuah penelitian menemukan bahwa dua program AI yang secara mandiri belajar mengenali gambar kuda dari perpustakaan yang luas, menggunakan pendekatan yang sangat berbeda (Lapuschkin et al., 2019). Sementara satu AI fokus tepat pada fitur-fitur hewan, AI lainnya mendasarkan keputusannya sepenuhnya pada sekelompok piksel di sudut kiri bawah setiap gambar kuda. Ternyata piksel tersebut berisi label hak cipta untuk gambar kuda tersebut. AI bekerja dengan sempurna untuk alasan yang sepenuhnya salah.

Untuk merancang algoritma yang adil, pertama-tama Anda harus memutuskan seperti apa hasil yang adil itu. Corbett-Davies dkk. (2017) menjelaskan empat definisi keadilan algoritmik yang berbeda untuk suatu algoritma yang menilai risiko seseorang melakukan kejahatan.

1. **Kesetaraan statistik** – dimana proporsi terdakwa yang ditahan di setiap kelompok ras sama besarnya. Misalnya, terdakwa kulit putih dan kulit hitam ditahan dengan tarif yang sama.
2. **Kesetaraan statistik bersyarat** – ketika mengendalikan serangkaian faktor risiko 'sah' yang terbatas, jumlah terdakwa yang ditahan dalam setiap kelompok ras sama besarnya. Misalnya, di antara terdakwa yang mempunyai jumlah hukuman yang sama, terdakwa berkulit hitam dan putih ditahan dengan jumlah yang sama.
3. **Kesetaraan prediktif** – yaitu akurasi keputusan yang sama antar kelompok ras, yang diukur dengan tingkat positif palsu. Ini berarti bahwa di antara para terdakwa yang tidak akan melakukan kejahatan kekerasan jika dibebaskan, tingkat penahanannya sama antar kelompok ras.
4. **Kalibrasi** – di antara terdakwa dengan skor risiko tertentu, proporsi yang melakukan pelanggaran ulang adalah sama di seluruh kelompok ras.

Namun, meskipun dimungkinkan untuk merancang algoritma yang memenuhi beberapa persyaratan ini, banyak gagasan tentang keadilan yang bertentangan satu sama lain, dan tidak mungkin memiliki algoritma yang memenuhi semuanya.

Aspek penting lainnya dari keadilan adalah mengetahui mengapa program otomatis mengambil keputusan tertentu. Misalnya, seseorang berhak mengetahui alasan penolakan pinjaman banknya. Hal ini memerlukan transparansi. Namun seperti yang akan kita ketahui, tidak selalu mudah untuk mengetahui mengapa suatu algoritma mengambil keputusan tertentu – banyak AI menggunakan 'jaringan saraf' yang kompleks sehingga perancangannya pun tidak dapat menjelaskan bagaimana mereka sampai pada jawaban tertentu.

Transparansi

Beberapa tahun yang lalu, sebuah program komputer di Amerika menilai kinerja guru di Houston dengan membandingkan nilai ujian siswanya dengan rata-rata negara bagian (Sample, 2017). Mereka yang memiliki peringkat tinggi mendapat pujian dan bahkan bonus, sedangkan mereka yang memiliki peringkat rendah akan dipecat. Beberapa guru merasa bahwa sistem menandai mereka tanpa alasan yang jelas, namun mereka tidak memiliki cara untuk memeriksa apakah program tersebut adil atau salah karena perusahaan yang membuat perangkat lunak tersebut, SAS Institute, menganggap algoritmanya sebagai rahasia dagang dan tidak akan mengungkapkannya. Para guru membawa kasus mereka ke pengadilan, dan hakim federal memutuskan bahwa program tersebut telah melanggar hak sipil mereka.

Studi kasus ini menyoroti pentingnya transparansi untuk membangun kepercayaan terhadap AI - kita harus selalu dapat mengetahui alasan sistem otonom mengambil keputusan tertentu, terutama jika keputusan tersebut menimbulkan kerugian. Mengingat uji coba autopilot mobil tanpa pengemudi di dunia nyata telah mengakibatkan beberapa kecelakaan fatal, jelas terdapat kebutuhan mendesak akan transparansi untuk mengetahui bagaimana

dan mengapa kecelakaan tersebut terjadi, memperbaiki kesalahan teknis atau operasional, dan membangun akuntabilitas.

Permasalahan ini juga banyak terjadi di kalangan masyarakat, terutama ketika menyangkut layanan kesehatan, yang merupakan permasalahan yang sangat pribadi bagi banyak orang (Komisi Eropa, 2017). Misalnya, di seluruh Eropa, banyak yang menyatakan keprihatinan atas kurangnya kemampuan mereka mengakses catatan kesehatan dan medis; Meskipun sebagian besar dari mereka akan dengan senang hati menyerahkan rekam medis mereka kepada profesional kesehatan, namun jauh lebih sedikit lagi yang akan dengan senang hati memberikan data tersebut kepada perusahaan publik atau swasta untuk tujuan penelitian medis. Sikap-sikap ini mencerminkan kekhawatiran atas kepercayaan, akses data, dan penggunaan data – yang semuanya sangat berkaitan dengan gagasan transparansi dan pemahaman tentang apa yang dikumpulkan AI, alasannya, dan bagaimana seseorang dapat mengakses data yang dikumpulkan mengenai hal tersebut.

Kotak hitam

Transparansi bisa menjadi sangat sulit dengan sistem AI modern, terutama yang berbasis sistem pembelajaran mendalam. Sistem pembelajaran mendalam didasarkan pada jaringan syaraf tiruan (JST), sekelompok node yang saling berhubungan, terinspirasi oleh penyederhanaan cara neuron terhubung di otak. Karakteristik ANN adalah, setelah ANN dilatih dengan kumpulan data, segala upaya untuk memeriksa struktur internal ANN untuk memahami mengapa dan bagaimana ANN membuat keputusan tertentu hampir tidak mungkin dilakukan. Sistem seperti ini disebut sebagai 'kotak hitam'.

Masalah lainnya adalah bagaimana memverifikasi sistem untuk memastikan bahwa sistem tersebut memenuhi persyaratan desain yang ditentukan. Pendekatan verifikasi saat ini biasanya berasumsi bahwa sistem yang diverifikasi tidak akan pernah mengubah perilakunya, namun sistem berdasarkan pembelajaran mesin—menurut definisinya—mengubah perilakunya, sehingga verifikasi apa pun kemungkinan besar akan dianggap tidak valid setelah sistem mempelajarinya (Winfield dan Jirotko, 2018).

AI Now Institute di New York University, yang meneliti dampak sosial AI, baru-baru ini merilis sebuah laporan yang mendesak badan-badan publik yang bertanggung jawab atas peradilan pidana, layanan kesehatan, kesejahteraan dan pendidikan untuk melarang AI kotak hitam karena keputusan mereka tidak dapat dijelaskan. Laporan tersebut juga merekomendasikan agar AI lulus uji coba pra-pelepasliaran dan dipantau 'di alam liar' sehingga bias dan kesalahan lainnya dapat segera diperbaiki (AI Now Report, 2018).

Dalam banyak kasus, dimungkinkan untuk mengetahui bagaimana suatu algoritma mengambil keputusan tertentu tanpa 'membuka kotak hitam AI'. Daripada mengungkap keseluruhan cara kerja AI, para peneliti baru-baru ini mengembangkan cara untuk mengetahui apa yang diperlukan untuk mengubah keputusan AI mereka (Wachter et al., 2018). Metode mereka dapat menjelaskan mengapa AI menolak permohonan hipotek seseorang, misalnya, karena AI mungkin mengungkapkan bahwa pinjaman tersebut ditolak karena pendapatan orang tersebut adalah Rp. 300.000.000, namun akan disetujui jika pendapatannya Rp.

450.000.000. Hal ini akan memungkinkan keputusan tersebut ditentang, dan memberi tahu orang tersebut apa yang perlu mereka tangani untuk mendapatkan pinjaman.

Kroll (2018) berpendapat bahwa, bertentangan dengan kritik yang menyatakan bahwa sistem perangkat lunak kotak hitam tidak dapat dipahami, algoritma pada dasarnya adalah bagian dari teknologi yang dapat dipahami. Ia menegaskan bahwa ketidakjelasan muncul dari dinamika kekuatan yang mengelilingi sistem perangkat lunak, bukan teknologi itu sendiri, yang selalu dibangun untuk tujuan tertentu, dan juga selalu dapat dipahami dalam kaitannya dengan tujuan desain dan operasional, serta masukan, keluaran, dan hasil. Misalnya, meskipun sulit untuk mengetahui mengapa iklan tertentu ditayangkan kepada orang tertentu pada waktu tertentu, hal tersebut mungkin dilakukan, dan tidak dilakukannya hal tersebut hanyalah sebuah pilihan desain, bukan suatu keniscayaan dari kompleksitas masalah besar. sistem – sistem harus dirancang sedemikian rupa sehingga mendukung analisis.

Kroll berargumen bahwa mungkin saja kita terlalu fokus pada pemahaman mekanisme suatu alat, padahal fokus sebenarnya seharusnya adalah pada bagaimana alat tersebut digunakan dan dalam konteks apa.

Permasalahan dan masalah lain terkait transparansi mencakup fakta bahwa perangkat lunak dan data adalah karya milik, yang berarti perusahaan mungkin tidak berkepentingan untuk mengungkapkan cara mereka mengatasi masalah tertentu. Banyak perusahaan memandang perangkat lunak dan algoritme mereka sebagai rahasia dagang berharga yang merupakan kunci mutlak untuk mempertahankan posisi mereka di pasar yang kompetitif.

Transparansi juga bertentangan dengan privasi, karena orang yang terlibat dalam pelatihan model pembelajaran mesin mungkin tidak ingin datanya, atau kesimpulan tentang datanya diungkapkan. Selain itu, masyarakat awam, atau bahkan regulator mungkin tidak memiliki pengetahuan teknologi untuk memahami dan menilai algoritma.

Sistem yang dapat dijelaskan

Beberapa peneliti menuntut sistem menghasilkan penjelasan tentang perilaku mereka. Namun, hal itu memerlukan keputusan tentang apa yang harus dijelaskan, dan kepada siapa. Penjelasan hanya berguna jika mencakup konteks di balik cara alat dioperasikan. Bahayanya adalah penjelasan berfokus pada mekanisme bagaimana alat tersebut beroperasi dan mengorbankan kontekstualisasi pengoperasian tersebut.

Dalam banyak kasus, memahami mekanisme sistem algoritmik secara tepat mungkin tidak diperlukan, sama seperti kita tidak memahami cara manusia mengambil keputusan. Demikian pula, meskipun transparansi sering kali diartikan sebagai pengungkapan kode sumber atau data, kita tidak harus melihat kode sumber komputer agar suatu sistem menjadi transparan, karena hal ini hanya akan memberi tahu kita sedikit tentang perilakunya. Sebaliknya transparansi harus berkaitan dengan perilaku eksternal algoritma. Beginilah cara kita mengatur perilaku manusia — bukan dengan melihat sirkuit saraf otak mereka, namun dengan mengamati perilaku mereka dan menilainya berdasarkan standar perilaku tertentu.

Penjelasan mungkin tidak meningkatkan kepercayaan manusia terhadap sistem komputer, karena jawaban yang salah sekalipun akan menerima penjelasan yang mungkin tampak masuk akal. Bias otomasi, sebuah fenomena di mana manusia menjadi lebih percaya

pada jawaban yang berasal dari mesin (Cummings, 2004), dapat berarti bahwa penjelasan yang menyesatkan tersebut mempunyai bobot yang cukup besar.

Pemahaman yang disengaja

Cara paling sederhana untuk memahami suatu teknologi adalah dengan memahami tujuan teknologi tersebut dirancang, bagaimana teknologi dirancang untuk melakukan hal tersebut, dan mengapa teknologi dirancang dengan cara tertentu dan bukan dengan cara lain (Kroll, 2018). Cara terbaik untuk memastikan bahwa suatu program berjalan sesuai keinginan Anda, dan tidak ada bias, atau konsekuensi yang tidak diinginkan adalah melalui validasi, investigasi, dan evaluasi program secara menyeluruh selama pengembangan. Dengan kata lain, mengukur kinerja suatu sistem selama pengembangan untuk mengungkap bug, bias, dan asumsi yang salah. Bahkan sistem yang dirancang dengan cermat pun dapat kehilangan fakta penting tentang dunia, dan penting untuk memverifikasi bahwa sistem beroperasi sebagaimana mestinya. Hal ini mencakup apakah model secara akurat mengukur apa yang seharusnya – sebuah konsep yang dikenal sebagai validitas konstruk; dan apakah data tersebut secara akurat mencerminkan dunia nyata

Misalnya, model pembelajaran mesin yang bertugas melakukan pemeriksaan kredit dapat secara tidak sengaja mengetahui bahwa kualitas pakaian peminjam berkorelasi dengan pendapatan dan kelayakan kredit mereka. Selama pengembangan perangkat lunak harus diperiksa korelasinya, sehingga dapat ditolak.

Auditor algoritma

Larsson dkk. (2019) menyarankan peran auditor algoritme profesional, yang tugasnya menginterogasi algoritme untuk memastikan algoritme mematuhi standar yang telah ditentukan sebelumnya. Salah satu contohnya adalah auditor algoritme kendaraan otonom, yang dapat memberikan simulasi skenario lalu lintas untuk memastikan bahwa kendaraan tidak meningkatkan risiko secara tidak proporsional terhadap pejalan kaki atau pengendara sepeda dibandingkan dengan penumpang.

Baru-baru ini, para peneliti mengusulkan kelas algoritme baru, yang disebut program pengawasan, yang fungsinya untuk 'memantau, mengaudit, dan menjaga akuntabilitas operasional program AI'. Misalnya, salah satu idenya adalah memiliki algoritma yang melakukan penilaian real-time terhadap jumlah bias yang disebabkan oleh algoritma penyaringan berita, sehingga menimbulkan peringatan jika bias meningkat melampaui ambang batas tertentu.

Akuntabilitas

'Bagaimana para pengambil keputusan memahami keputusan apa yang dibuat oleh teknologi AI dan bagaimana keputusan tersebut berbeda dengan keputusan yang dibuat oleh manusia?... intinya adalah bahwa AI membuat keputusan secara berbeda dari manusia dan terkadang kita tidak memahami perbedaan-perbedaan tersebut. Kami tidak tahu mengapa atau bagaimana mereka mengambil keputusan itu.'

Metode lain untuk memastikan kepercayaan terhadap AI adalah melalui akuntabilitas. Seperti yang telah dibahas, akuntabilitas memastikan bahwa jika AI melakukan kesalahan atau merugikan seseorang, ada seseorang yang dapat dimintai pertanggungjawaban, baik itu

perancang, pengembang, atau perusahaan yang menjual AI tersebut. Jika terjadi kerugian, harus ada mekanisme ganti rugi sehingga korban dapat memperoleh kompensasi yang memadai.

Semakin banyak literatur yang mulai membahas konsep-konsep seperti akuntabilitas algoritmik dan AI yang bertanggung jawab. Akuntabilitas algoritmik, menurut Caplan dkk. (2018), berkaitan dengan pendelegasian tanggung jawab atas kerugian yang timbul sebagai akibat dari keputusan berdasarkan algoritma yang menghasilkan konsekuensi yang diskriminatif atau tidak adil. Salah satu bidang di mana akuntabilitas mungkin menjadi penting adalah pengenalan kendaraan tanpa pengemudi. Jika terjadi kecelakaan, siapa yang harus bertanggung jawab? Sejumlah kecelakaan fatal telah terjadi pada mobil self-driving, misalnya pada tahun 2016, Tesla Model S yang dilengkapi radar dan kamera menentukan bahwa truk di dekatnya ternyata berada di angkasa, sehingga mengakibatkan kecelakaan fatal. Pada bulan Maret 2018, sebuah mobil yang digunakan oleh Uber dalam uji coba kendaraan tanpa pengemudi menabrak dan membunuh seorang wanita di Arizona, AS. Meskipun mobil otonom lebih aman dibandingkan kendaraan yang dikendarai manusia, kecelakaan seperti ini merusak kepercayaan.

Peraturan

Salah satu cara untuk memastikan akuntabilitas adalah regulasi. Winfield dan Jirotko (2018) mengemukakan bahwa teknologi secara umum dipercaya jika membawa manfaat dan aman serta diatur dengan baik. Makalah mereka berpendapat bahwa salah satu elemen kunci dalam membangun kepercayaan terhadap AI adalah tata kelola yang etis – serangkaian proses, prosedur, budaya, dan nilai yang dirancang untuk memastikan standar perilaku tertinggi. Standar perilaku ini perlu diadopsi oleh masing-masing desainer dan organisasi di mana mereka bekerja, sehingga masalah etika ditangani ketika atau sebelum masalah tersebut muncul dengan cara yang berprinsip, daripada menunggu hingga masalah muncul dan mengatasinya secara langsung.

Mereka mencontohkan maskapai penerbangan yang dapat dipercaya karena kita tahu bahwa mereka adalah bagian dari industri yang diatur dengan ketat dan memiliki catatan keselamatan yang luar biasa. Alasan mengapa pesawat komersial begitu aman bukan hanya karena desainnya yang bagus, namun juga karena proses sertifikasi keselamatan yang ketat, dan fakta bahwa jika terjadi kesalahan, terdapat proses investigasi kecelakaan udara yang kuat dan dapat dilihat oleh publik.

Winfield dan Jirotko (2018) menyarankan bahwa beberapa jenis robot, misalnya mobil tanpa pengemudi, harus diatur melalui badan yang mirip dengan Otoritas Penerbangan Sipil (CAA), dengan mobil tanpa pengemudi setara dengan Cabang Investigasi Kecelakaan Udara.

Terkait persepsi masyarakat terhadap robot dan teknologi canggih, regulasi dan manajemen menjadi perhatian utama. Dalam dua survei terhadap masyarakat di seluruh UE (Komisi Eropa 2012; Komisi Eropa, 2012), keduanya menunjukkan bahwa secara umum terdapat pandangan positif terhadap robot dan digitalisasi selama hal ini diterapkan dan dikelola dengan hati-hati. Faktanya, antara 88% dan 91% dari mereka yang disurvei menyatakan bahwa robot dan teknologi canggih harus dikelola dengan hati-hati, hal ini

merupakan salah satu hasil terkuat dalam kedua survei tersebut – yang mencerminkan keprihatinan dan area prioritas yang kuat di kalangan warga UE.

Pengendalian

Masalah lain yang mempengaruhi kepercayaan masyarakat terhadap AI adalah pengendalian. Hal ini sebagian besar berkaitan dengan ketakutan terhadap gagasan 'Superintelligence' - bahwa ketika kecerdasan buatan meningkat hingga melampaui kemampuan manusia, kecerdasan buatan mungkin akan mengambil kendali atas sumber daya kita dan mengalahkan spesies kita, sehingga menyebabkan kepunahan manusia. Ketakutan terkait adalah, meskipun agen AI dirancang dengan hati-hati agar memiliki tujuan yang selaras dengan kebutuhan manusia, agen tersebut mungkin akan mengembangkan subtujuan yang tidak diantisipasi. Misalnya, Bryson (2019) mencontohkan robot pemain catur yang diajarkan untuk meningkatkan permainannya. Robot ini secara tidak sengaja belajar menembak orang yang mematikannya di malam hari, sehingga merampas sumber daya vitalnya. Namun, meskipun sebagian besar peneliti sepakat bahwa ancaman ini tidak mungkin terjadi, untuk menjaga kepercayaan terhadap AI, manusia harus memiliki pengawasan penuh terhadap teknologi ini.

Manusia dalam lingkaran

Salah satu gagasan yang dikemukakan oleh para peneliti adalah selalu menjaga human-in-the-loop (HITL). Di sini operator manusia akan menjadi komponen penting dari proses kendali otomatis, mengawasi robot. Bentuk sederhana dari HITL yang sudah ada adalah penggunaan pekerja manusia untuk memberi label pada data untuk melatih algoritma pembelajaran mesin. Misalnya saat Anda menandai email sebagai 'spam', Anda adalah salah satu dari banyak manusia yang berada dalam lingkaran algoritme pembelajaran mesin yang kompleks, yang membantunya dalam upaya berkelanjutan untuk meningkatkan klasifikasi email sebagai spam atau non-spam.

Namun HITL juga dapat menjadi alat yang ampuh untuk mengatur perilaku sistem AI. Misalnya, banyak peneliti berpendapat bahwa operator manusia harus mampu memantau perilaku LAWS, atau 'robot pembunuh', atau algoritma penilaian kredit. Kehadiran manusia memenuhi dua fungsi utama dalam sistem AI HITL (Rahwan, 2018):

1. Manusia dapat mengidentifikasi perilaku buruk melalui sistem yang otonom, dan mengambil tindakan perbaikan. Misalnya, sistem penilaian kredit mungkin salah mengklasifikasikan orang dewasa sebagai tidak memenuhi syarat untuk mendapatkan kredit karena usia mereka salah dimasukkan—sesuatu yang mungkin dilihat manusia dari foto pemohon. Demikian pula, sistem visi komputer pada drone yang dipersenjatai mungkin salah mengidentifikasi warga sipil sebagai kombatan, dan operator manusia—diharapkan—akan mengesampingkan sistem tersebut.
2. Menjaga manusia tetap berada dalam lingkaran juga akan memberikan akuntabilitas - jika sistem otonom menyebabkan kerugian pada manusia, keberadaan manusia dalam lingkaran memberikan kepercayaan bahwa seseorang akan menanggung akibat dari kesalahan tersebut. Menurut Rahwan (2018), sampai kita menemukan cara untuk menghukum algoritma yang merugikan manusia, 'sulit untuk memikirkan alternatif lain'.

Namun, meskipun HITL berguna untuk membangun sistem AI yang harus diawasi, hal ini mungkin tidak cukup. Mesin AI yang mengambil keputusan dengan implikasi sosial yang lebih luas, seperti algoritme yang mengendalikan jutaan mobil tanpa pengemudi atau algoritme penyaringan berita yang memengaruhi keyakinan dan preferensi politik jutaan warga negara, harus diawasi oleh masyarakat secara keseluruhan, sehingga memerlukan pengawasan yang ketat dari masyarakat secara keseluruhan. paradigma 'masyarakat dalam lingkaran'.

Tombol merah besar

Sebagai cara untuk mengatasi beberapa ancaman kecerdasan buatan, para peneliti telah mengusulkan cara untuk menghentikan sistem AI sebelum sistem tersebut lepas dari kendali luar dan menyebabkan kerusakan. Apa yang disebut 'tombol merah besar', atau 'tombol mematikan' akan memungkinkan operator manusia untuk mengganggu atau mengalihkan sistem, sekaligus mencegah sistem mengetahui bahwa intervensi semacam itu merupakan ancaman. Namun, beberapa komentator khawatir bahwa mesin AI yang cukup canggih dapat mengantisipasi langkah ini dan mempertahankan diri dengan belajar menonaktifkan 'tombol mematikan' miliknya sendiri.

Tombol merah memunculkan pertanyaan praktis yang lebih luas tentang mematikan sistem AI agar tetap aman. Apa cara terbaik untuk mencapai hal tersebut, dan untuk jenis sistem AI spesifik apa?

Orseau dan Armstrong (2016) baru-baru ini menerbitkan makalah tentang cara mencegah AI yang diprogram melalui pembelajaran penguatan (RL) agar tidak melihat interupsi sebagai ancaman. Misalnya, suatu algoritme yang mencoba mengoptimalkan kinerja caturinya mungkin belajar untuk menonaktifkan tombol matinya sehingga algoritme tersebut dapat menghabiskan lebih banyak waktu untuk mempelajari cara bermain catur. Atau mungkin belajar untuk menyakiti orang yang mencoba memamatkannya, dll. Apa yang para peneliti usulkan adalah menjauhkan varian pembelajaran penguatan tertentu dari pembelajaran untuk menghindari atau menghalangi interupsi. Dengan cara ini, menurut penulis, suatu sistem dapat menghasilkan kebijakan optimal yang juga dapat diinterupsi. Dengan bersifat 'dapat diinterupsi secara aman', makalah ini menyimpulkan, pembelajaran penguatan tidak akan melemahkan upaya pengawasan dan intervensi yang bertanggung jawab.

Riedl dan Harrison (2017) menyarankan untuk membuat 'tombol merah besar' yang, setelah ditekan, akan mengalihkan AI ke dunia simulasi di mana AI dapat menjalankan fungsi penghargaannya tanpa menimbulkan bahaya apa pun. Alternatifnya, ide lainnya adalah menjaga ketidakpastian sistem mengenai fungsi penghargaan utama, yang akan mencegah AI memberikan nilai pada penonaktifan off-switch.

Namun Arnold dan Schultz (2018) berargumen bahwa pendekatan 'tombol merah' muncul ketika suatu sistem sudah 'menjadi nakal' dan berusaha menghalangi interferensi, dan pendekatan 'tombol merah besar' berfokus pada ancaman jangka panjang, dengan membayangkan sistem yang jauh lebih maju daripada yang ada saat ini dan mengabaikan permasalahan yang ada saat ini dalam menjaga akuntabilitas sistem otomatis. Pendekatan

yang lebih baik, menurut Arnold dan Scheutz, adalah dengan melakukan evaluasi diri dan pengujian yang berkelanjutan sebagai bagian integral dari operasi sistem, untuk mendiagnosis bagaimana kinerja sistem, dan memperbaiki kesalahan apa pun.

Mereka berargumentasi bahwa untuk mencapai hal ini AI harus mengandung inti etika (EC) yang terdiri dari mekanisme pembuatan skenario dan lingkungan simulasi yang digunakan untuk menguji keputusan sistem di dunia simulasi, bukan di dunia nyata. EC ini akan disembunyikan dari sistem itu sendiri, sehingga algoritma sistem akan dicegah untuk mempelajari operasi dan fungsinya, dan pada akhirnya keberadaannya. Melalui pengujian terus-menerus di dunia simulasi, Komisi Eropa akan memantau dan memeriksa perilaku menyimpang - memberikan respons yang jauh lebih efektif dan waspada dibandingkan tombol darurat yang mungkin tidak dapat ditekan tepat waktu.

BAB 3

INISIATIF ETIS DI BIDANG KECERDASAN BUATAN

Sebagaimana dirinci di bagian sebelumnya, ada banyak sekali pertimbangan etis yang menyertai pengembangan, penggunaan, dan dampak kecerdasan buatan (AI). Hal ini berkisar dari potensi dampak AI terhadap hak asasi manusia warga negara dalam suatu masyarakat hingga keamanan dan pemanfaatan data yang dikumpulkan; mulai dari bias dan diskriminasi yang secara tidak sengaja ditanamkan ke dalam AI oleh sekelompok pengembang yang homogen, hingga kurangnya kesadaran dan pemahaman masyarakat tentang konsekuensi dari pilihan dan penggunaan AI tertentu, yang mengarah pada pengambilan keputusan yang kurang tepat dan kerugian selanjutnya.

AI dibangun berdasarkan revolusi TIK dan komputasi sebelumnya, sehingga akan menghadapi sejumlah masalah etika serupa. Meskipun teknologi dapat digunakan untuk kebaikan, namun ada kemungkinan teknologi tersebut disalahgunakan. Kita mungkin melakukan antropomorfisasi dan memanusaiakan AI secara berlebihan, sehingga mengaburkan batas antara manusia dan mesin. Perkembangan AI yang sedang berlangsung akan menimbulkan 'kesenjangan digital' baru, dimana teknologi memberikan manfaat yang lebih besar kepada beberapa kelompok sosio-ekonomi dan geografis dibandingkan kelompok lainnya. Lebih lanjut, AI akan berdampak pada biosfer dan lingkungan kita yang belum memenuhi syarat (Veruggio dan Operto, 2006).

3.1. INISIATIF ETIKA INTERNASIONAL

Meskipun peraturan resmi masih langka, banyak inisiatif independen telah diluncurkan secara internasional untuk mengeksplorasi permasalahan ini – dan permasalahan etika lainnya. Inisiatif-inisiatif yang dieksplorasi pada bagian ini diuraikan pada Tabel 3.1 dan akan dipelajari berdasarkan dampak buruk dan kekhawatiran terkait yang ingin dipahami dan dimitigasi.

Tabel 3.1: Inisiatif etis dan dampak buruknya telah diatasi

Prakarsa	Lokasi	Permasalahan utama telah diatasi	Publikasi	Sumber pendanaan
Institut Etika dalam Kecerdasan Buatan	Jerman	Rekayasa yang berpusat pada manusia dan fokus pada landasan budaya dan sosial dari kemajuan pesat AI, yang mencakup disiplin ilmu termasuk filsafat, etika, sosiologi, dan ilmu politik.		Hibah pendanaan awal (2019) dari Facebook (Rp. 7,5 Miliar selama lima tahun).
Institut AI & Pembelajaran	Britania Raya	Institut ini bertujuan untuk memberdayakan semua orang mulai		tidak dikenal

Mesin yang Etis		dari individu hingga seluruh negara untuk mengembangkan AI, berdasarkan delapan prinsip pembelajaran mesin yang bertanggung jawab: prinsip-prinsip ini berkaitan dengan pemeliharaan kendali manusia, ganti rugi yang tepat untuk dampak AI, evaluasi bias, penjelasan, transparansi, reproduktifitas, mitigasi pengaruh otomatisasi AI pada pekerja, akurasi, biaya, privasi, kepercayaan, dan keamanan.		
Institut Kecerdasan Buatan Etis dalam Pendidikan	Britania Raya	Potensi ancaman terhadap generasi muda dan pendidikan dari pesatnya pertumbuhan teknologi AI baru, dan memastikan pengembangan etis dari EdTech yang dipimpin oleh AI.		tidak dikenal
Institut Kehidupan Masa Depan	Amerika Serikat	Memastikan bahwa pengembangan AI bermanfaat bagi umat manusia, dengan fokus pada keselamatan dan risiko eksistensial: perlombaan senjata otonom, kendali manusia terhadap AI, dan potensi bahaya AI yang bersifat 'umum/kuat' atau super cerdas.	'Prinsip AI Asilomar'	Pribadi. Donor teratas: Elon Musk (SpaceX dan Tesla), Jaan Tallinn (Skype), Matt Wage (pedagang keuangan), Nisan Stiennon (insinyur perangkat lunak), Sam Harris, George Godula (pengusaha teknologi), dan Jacob Trefethen (Harvard).
Asosiasi Mesin Komputasi	Amerika Serikat	Transparansi, kegunaan, keamanan, aksesibilitas, akuntabilitas, dan inklusivitas digital komputer dan jaringan, dalam hal penelitian, pengembangan, dan implementasi.	Pernyataan tentang: transparansi dan akuntabilitas algoritmik (Januari 2017), komputasi dan keamanan jaringan (Mei	tidak dikenal

			2017), Internet of Things (Juni 2017), aksesibilitas, kegunaan, dan inklusivitas digital (September 2017), dan akses wajib terhadap infrastruktur informasi berdasarkan undang-undang penegakan hukum (April 2018).	
Masyarakat Jepang untuk Kecerdasan Buatan (JSAI)	Jepang	Untuk memastikan bahwa penelitian dan pengembangan AI tetap bermanfaat bagi masyarakat, dan bahwa pengembangan serta penelitian dilakukan secara etis dan moral.	'Pedoman Etika'	tidak dikenal
AI4All	Amerika Serikat	Keberagaman dan inklusi dalam AI, untuk memaparkan kelompok-kelompok yang kurang terwakili pada AI demi kebaikan sosial dan manfaat kemanusiaan.		Google
Masyarakat Masa Depan	Amerika Serikat	Dampak dan tata kelola kecerdasan buatan memberikan manfaat luas bagi masyarakat, mencakup penelitian kebijakan, penasehatan dan kecerdasan kolektif, koordinasi tata kelola, hukum, dan pendidikan.	'Draf Prinsip Tata Kelola AI' Diterbitkan Oktober 2017 (kemudian dipublikasikan di situs web mereka pada 7 Februari 2019),	tidak dikenal
Institut AI Sekarang	Amerika Serikat	Implikasi sosial dari AI, khususnya di bidang: Hak dan kebebasan, ketenagakerjaan dan otomatisasi, bias dan inklusi, serta keselamatan dan infrastruktur penting.		Berbagai organisasi, termasuk Luminata, MacArthur Foundation, Microsoft Research,

				Google, Ford Foundation, DeepMind Ethics & Society, dan Ethics & Governance of AI Initiative.
Institut Insinyur Listrik dan Elektronika (IEEE)	Amerika Serikat	Pedoman sosial dan kebijakan untuk menjaga AI dan sistem cerdas tetap berpusat pada manusia, dan melayani nilai-nilai dan prinsip-prinsip kemanusiaan. Berfokus untuk memastikan bahwa semua pemangku kepentingan – baik desain maupun pengembangan – dididik, dilatih, dan diberdayakan untuk memprioritaskan pertimbangan etis dalam hal hak asasi manusia, kesejahteraan, akuntabilitas, transparansi, dan kesadaran akan penyalahgunaan.	'Desain yang Selaras Secara Etis' Edisi Pertama (Maret 2019)	
Kemitraan dalam AI	Amerika Serikat	Praktik terbaik dalam teknologi AI: Keamanan, keadilan, akuntabilitas, transparansi, tenaga kerja dan perekonomian, kolaborasi antara manusia dan sistem, pengaruh sosial dan masyarakat, serta kebaikan sosial.		Kemitraan ini dibentuk oleh sekelompok peneliti AI yang mewakili enam perusahaan teknologi terbesar di dunia: Apple, Amazon, DeepMind dan Google, Facebook, IBM, dan Microsoft.
Yayasan Robotika yang Bertanggung Jawab	Belanda	Robotika yang bertanggung jawab (dalam hal desain, pengembangan, penggunaan, regulasi, dan implementasi). Secara proaktif mempertimbangkan isu-isu yang menyertai inovasi teknologi, dan dampaknya terhadap nilai-nilai sosial seperti keselamatan, keamanan, privasi, dan kesejahteraan.		tidak dikenal
AI4People	Belgium	Dampak sosial AI, dan prinsip-prinsip dasar, kebijakan, dan praktik yang	'Kerangka Etis untuk	Atomium— Institut Sains,

		menjadi landasan untuk membangun 'masyarakat AI yang baik'.	Masyarakat AI yang Baik'	Media, dan Demokrasi Eropa. Sebagian pendanaan diberikan kepada Ketua Komite Ilmiah proyek dari Dewan Penelitian Teknik dan Ilmu Fisika.
Inisiatif Etika dan Tata Kelola Kecerdasan Buatan	Amerika Serikat	Berusaha untuk memastikan bahwa teknologi otomasi dan pembelajaran mesin diteliti, dikembangkan, dan diterapkan dengan cara yang membenarkan nilai-nilai sosial tentang keadilan, otonomi manusia, dan keadilan.		Harvard Berkman Klein Center dan MIT Media Lab. Didukung oleh The Miami Foundation (sponsor fiskal), Knight Foundation, Luminate, Red Hoffman, dan William and Flora Hewlett Foundation.
Saidot: Mewujudkan ekosistem AI yang bertanggung jawab	Finlandia	Membantu perusahaan, pemerintah, dan organisasi mengembangkan dan menerapkan ekosistem AI yang bertanggung jawab, untuk memberikan layanan AI yang transparan, akuntabel, dan dapat dipercaya. Memungkinkan organisasi untuk mengembangkan AI yang berpusat pada manusia, dengan fokus pada peningkatan tingkat kepercayaan dan akuntabilitas dalam ekosistem AI. Platform ini menawarkan perangkat lunak dan sistem algoritmik yang dapat 'memvalidasi keandalan sistem intelijen' (Saidot, 2019)		
euRobotika	Eropa	Mempertahankan dan memperluas bakat dan kemajuan Eropa dalam		Komisi Eropa

		bidang robotika – industrialisasi AI dan dampak ekonomi.		
Pusat Etika dan Inovasi Data	Inggris	Mengidentifikasi dan menutup kesenjangan dalam lanskap peraturan, penggunaan data AI, dan memaksimalkan manfaat AI bagi masyarakat.		Pemerintah Inggris
Kelompok Minat Khusus Kecerdasan Buatan (SIGAI), Asosiasi Mesin Komputasi	Amerika Serikat	Mempromosikan dan mendukung pertumbuhan dan penerapan prinsip dan teknik AI di seluruh komputasi, dan mempromosikan pendidikan dan publikasi AI melalui berbagai forum		Asosiasi Mesin Komputasi
<i>Perkembangan internasional penting lainnya: terkini dan historis</i>				
Deklarasi Montreal	Kanada	Pengembangan AI yang bertanggung jawab secara sosial, mempertemukan 400 peserta dari seluruh sektor masyarakat untuk mengidentifikasi tantangan etika dan moral dalam jangka pendek dan panjang. Nilai-nilai utama: kesejahteraan, otonomi, keadilan, privasi, pengetahuan, demokrasi, dan akuntabilitas.		Université de Montréal dengan dukungan dari Fonds de recherche en santé du Québec dan Palais des congrès de Montréal.
Persatuan Global UNI	Swiss	Gangguan pekerja dan transparansi dalam penerapan AI, robotika, serta pembelajaran data dan mesin di tempat kerja. Melindungi kepentingan pekerja dan menjaga kendali manusia serta keseimbangan kekuasaan yang sehat.	'10 Prinsip Teratas untuk AI yang Etis'	tidak dikenal
Jaringan Penelitian Robotika Eropa (EURON)	Eropa (Koordinator berbasis di Swedia)	Koordinasi penelitian, pendidikan dan pelatihan, penerbitan dan pertemuan, hubungan industri dan hubungan internasional di bidang robotika.	'Peta Jalan Robotetika'	Komisi Eropa (2000-2004)
Platform Robotika Eropa (EUROP)	Eropa	Menyatukan komunitas robotika dan AI Eropa. Didorong oleh industri, fokus pada daya saing dan inovasi.		Komisi Eropa

3.2. KERUGIAN DAN KEKHAWATIRAN ETIS DIATASI DENGAN INISIATIF INI

Semua inisiatif yang disebutkan di atas sepakat bahwa AI harus diteliti, dikembangkan, dirancang, diterapkan, dipantau, dan digunakan dengan cara yang etis – namun masing-masing inisiatif memiliki bidang prioritas yang berbeda. Bagian ini akan mencakup analisis dan pengelompokan inisiatif-inisiatif di atas, berdasarkan jenis permasalahan yang ingin diatasi, dan kemudian menguraikan beberapa pendekatan dan solusi yang diusulkan untuk melindungi dari bahaya.

Sejumlah permasalahan utama muncul dari inisiatif ini, yang secara garis besar dapat dibagi ke dalam kategori berikut:

1. Hak Asasi Manusia dan kesejahteraan
Apakah AI demi kepentingan terbaik umat manusia dan kesejahteraan manusia?
2. Kerugian emosional
Akankah AI menurunkan integritas pengalaman emosional manusia, atau memfasilitasi kerusakan emosional atau mental?
3. Akuntabilitas dan tanggung jawab
Siapa yang bertanggung jawab atas AI, dan siapa yang akan bertanggung jawab atas tindakannya?
4. Keamanan, privasi, aksesibilitas, dan transparansi
Bagaimana kita menyeimbangkan aksesibilitas dan transparansi dengan privasi dan keamanan, terutama terkait data dan personalisasi?
5. Keamanan dan kepercayaan
Bagaimana jika AI dianggap tidak dapat dipercaya oleh publik, atau bertindak dengan cara yang mengancam keselamatan dirinya sendiri atau orang lain?
6. Kerugian sosial dan keadilan sosial
Bagaimana kita memastikan bahwa AI bersifat inklusif, bebas dari bias dan diskriminasi, serta selaras dengan moral dan etika publik?
7. Kerugian finansial
Bagaimana kita mengendalikan AI yang berdampak negatif pada peluang ekonomi dan lapangan kerja, serta mengambil pekerjaan dari pekerja manusia atau menurunkan peluang dan kualitas pekerjaan tersebut?
8. Keabsahan dan keadilan
Bagaimana kita memastikan bahwa AI – dan data yang dikumpulkannya – digunakan, diproses, dan dikelola dengan cara yang adil, merata, dan sah, serta tunduk pada tata kelola dan peraturan yang sesuai? Seperti apa peraturan tersebut? Haruskah AI diberikan 'kepribadian'?
9. Pengendalian dan penggunaan – atau penyalahgunaan – AI secara etis
Bagaimana AI bisa digunakan secara tidak etis - dan bagaimana kita dapat melindunginya dari hal ini? Bagaimana kita memastikan bahwa AI tetap berada di bawah kendali manusia sepenuhnya, bahkan ketika AI berkembang dan 'belajar'?
10. Kerusakan dan kelestarian lingkungan hidup

Bagaimana kita melindungi terhadap potensi kerusakan lingkungan yang terkait dengan pengembangan dan penggunaan AI? Bagaimana kita memproduksinya secara berkelanjutan?

11. Penggunaan yang diinformasikan

Apa yang harus kita lakukan untuk memastikan masyarakat sadar, terdidik, dan mendapat informasi tentang penggunaan dan interaksi mereka dengan AI?

12. Risiko eksistensial

Bagaimana kita menghindari perlombaan senjata AI, melakukan mitigasi dan mengatur potensi bahaya terlebih dahulu, dan memastikan bahwa pembelajaran mesin tingkat lanjut bersifat progresif dan dapat dikelola?



Gambar 3.2: Prinsip-prinsip umum untuk desain, pengembangan, dan implementasi sistem otonom dan cerdas yang beretika dan berbasis nilai (sebagaimana didefinisikan oleh Desain yang Disejajarkan Secara Etis IEEE Edisi Pertama Maret 2019)

Secara keseluruhan, inisiatif-inisiatif ini bertujuan untuk mengidentifikasi dan membentuk kerangka kerja dan sistem etis yang menetapkan manfaat bagi manusia pada tingkat tertinggi, memprioritaskan manfaat bagi masyarakat dan lingkungan (tanpa membuat kedua tujuan ini bertentangan), dan memitigasi risiko dan dampak negatif. terkait dengan AI — dengan fokus untuk memastikan bahwa AI akuntabel dan transparan (IEEE, 2019).

'Desain yang Selaras Secara Etis: Visi untuk Memprioritaskan Kesejahteraan Manusia dengan Sistem Otonom dan Cerdas' (v1; 2019) dari IEEE adalah salah satu dokumen paling penting yang diterbitkan hingga saat ini mengenai masalah etika yang mungkin diangkat oleh AI — dan berbagai cara yang diusulkan untuk memitigasi hal ini.

Bidang-bidang yang terkena dampak utama meliputi pembangunan berkelanjutan; hak data pribadi dan keagenan atas identitas digital; kerangka hukum untuk akuntabilitas; dan kebijakan untuk pendidikan dan kesadaran. Nilai-nilai tersebut termasuk dalam tiga pilar kerangka konseptual Desain yang Selaras Secara Etis: Nilai-nilai kemanusiaan universal; penentuan nasib sendiri secara politik dan badan data; dan ketergantungan teknis.

Kerusakan secara rinci

Dengan mempertimbangkan masing-masing dampak buruk tersebut, bagian ini mengeksplorasi bagaimana dampak buruk tersebut dikonseptualisasikan melalui inisiatif dan beberapa tantangan yang masih ada.

Hak asasi manusia dan kesejahteraan

Semua inisiatif menganut pandangan bahwa AI tidak boleh melanggar hak asasi manusia yang mendasar, seperti martabat manusia, keamanan, privasi, kebebasan berekspresi dan informasi, perlindungan data pribadi, kesetaraan, solidaritas dan keadilan (Parlemen, Dewan dan Komisi Eropa, 2012).

Bagaimana kita memastikan bahwa AI menjunjung tinggi hak asasi manusia dan memprioritaskan kesejahteraan manusia? Atau apakah AI tidak memberikan dampak yang tidak proporsional terhadap kelompok masyarakat yang rentan, seperti anak-anak, penyandang disabilitas, atau orang lanjut usia, atau menurunkan kualitas hidup masyarakat?

Untuk memastikan bahwa hak asasi manusia dilindungi, IEEE merekomendasikan kerangka kerja, standar, dan badan pengatur tata kelola baru yang mengawasi penggunaan AI; menerjemahkan kewajiban hukum yang ada ke dalam kebijakan yang terinformasi, memungkinkan adanya norma budaya dan kerangka hukum; dan selalu mempertahankan kendali penuh manusia atas AI, tanpa memberikan mereka hak atau keistimewaan yang setara dengan manusia (IEEE, 2019). Untuk menjaga kesejahteraan manusia, yang didefinisikan sebagai 'kepuasan manusia terhadap kehidupan dan kondisi kehidupan, serta keseimbangan yang tepat antara pengaruh positif dan negatif' (ibid), IEEE menyarankan untuk memprioritaskan kesejahteraan manusia selama fase desain, dan menggunakan metrik terbaik dan paling diterima secara luas untuk mengukur dengan jelas keberhasilan AI di masyarakat.

Ada persilangan antara akuntabilitas dan transparansi: harus selalu ada cara yang tepat untuk mengidentifikasi dan melacak pelanggaran hak, dan menawarkan ganti rugi dan reformasi yang sesuai. Data pribadi juga merupakan isu utama di sini; AI mengumpulkan segala macam data pribadi, dan pengguna harus tetap memiliki akses dan kendali atas data mereka, untuk memastikan bahwa hak-hak dasar mereka ditegakkan secara hukum (IEEE, 2019).

Menurut Foundation for Responsible Robotics, AI harus dikembangkan secara etis dengan mempertimbangkan hak asasi manusia untuk mencapai tujuan 'robotika yang

bertanggung jawab', yang mengandalkan inovasi proaktif untuk menjunjung tinggi nilai-nilai sosial seperti keselamatan, keamanan, privasi, dan kesejahteraan. Yayasan ini terlibat dengan para pembuat kebijakan, mengatur dan menyelenggarakan acara, menerbitkan dokumen konsultasi untuk mendidik para pembuat kebijakan dan masyarakat, dan menciptakan kolaborasi pemerintah-swasta untuk menjembatani kesenjangan antara industri dan konsumen, untuk menciptakan transparansi yang lebih besar. Hal ini memerlukan pengambilan keputusan yang etis sejak tahap penelitian dan pengembangan, pendidikan konsumen yang lebih baik, serta pembuatan undang-undang dan kebijakan yang bertanggung jawab – yang dilakukan sebelum AI dirilis dan digunakan.

Future of Life Institute mendefinisikan sejumlah prinsip, etika, dan nilai-nilai yang perlu dipertimbangkan dalam pengembangan AI, termasuk kebutuhan untuk merancang dan mengoperasikan AI dengan cara yang sesuai dengan cita-cita martabat manusia, hak, kebebasan, dan budaya. keanekaragaman. Hal ini juga diamini oleh Pedoman Etika Masyarakat Jepang untuk AI, yang menempatkan pentingnya realisasi AI dengan cara yang bermanfaat bagi umat manusia, dan sejalan dengan etika, hati nurani, dan kompetensi peneliti dan masyarakat secara keseluruhan. AI harus berkontribusi terhadap perdamaian, keselamatan, kesejahteraan, dan kepentingan umum masyarakat, kata Society, dan melindungi hak asasi manusia.

Inisiatif Hukum dan Masyarakat Masa Depan menekankan bahwa umat manusia memiliki hak, martabat, dan kebebasan untuk berkembang yang setara, dan berhak atas hak asasi manusia mereka. Dengan mengingat hal ini, sejauh mana kita harus mendelegasikan keputusan yang mempengaruhi manusia kepada mesin? Misalnya, apakah 'hakim' AI dalam profesi hukum bisa lebih efisien, adil, seragam, dan hemat biaya dibandingkan manusia – dan jika memang demikian, apakah ini cara yang tepat untuk menerapkan AI? Deklarasi Montréal bertujuan untuk memperjelas hal ini, dengan menyusun kerangka etika yang mendukung hak asasi manusia yang diakui secara internasional di bidang-bidang yang terkena dampak penerapan AI: 'Prinsip-prinsip deklarasi saat ini bertumpu pada keyakinan umum bahwa umat manusia berupaya untuk tumbuh sebagai makhluk sosial. makhluk yang diberkahi dengan sensasi, pikiran dan perasaan, dan berusaha untuk memenuhi potensi mereka dengan secara bebas menjalankan kapasitas emosional, moral dan intelektual mereka.' Dengan kata lain, AI tidak hanya tidak boleh mengganggu kesejahteraan manusia, namun juga harus secara proaktif mendorong dan mendukungnya untuk berkembang dan berkembang.

Beberapa pihak mendekati AI dari sudut pandang yang lebih spesifik – seperti UNI Global Union, yang berupaya melindungi hak individu untuk bekerja. Lebih dari separuh pekerjaan yang saat ini dilakukan oleh manusia dapat dilakukan lebih cepat dan efisien dengan cara otomatis, kata Union. Hal ini mengidentifikasi kerugian besar yang mungkin disebabkan oleh AI dalam bidang pekerjaan manusia. Persatuan ini menyatakan bahwa kita harus memastikan bahwa AI bermanfaat bagi manusia dan bumi, serta melindungi dan meningkatkan hak asasi manusia, martabat manusia, integritas, kebebasan, privasi, serta keragaman budaya dan gender.

Kerugian emosional

Apa artinya menjadi manusia? AI akan berinteraksi dan berdampak pada pengalaman emosional manusia dengan cara yang belum memenuhi syarat; manusia rentan terhadap pengaruh emosional baik secara positif maupun negatif, dan 'mempengaruhi' – bagaimana emosi dan keinginan mempengaruhi perilaku – adalah bagian inti dari kecerdasan. Pengaruhnya berbeda-beda antar budaya, dan, mengingat sensitivitas budaya dan cara berinteraksi yang berbeda, AI yang afektif dan berpengaruh dapat mulai memengaruhi cara orang memandang masyarakat itu sendiri. IEEE merekomendasikan berbagai cara untuk memitigasi risiko ini, termasuk kemampuan untuk mengadaptasi dan memperbarui norma dan nilai AI sesuai dengan siapa mereka terlibat, dan sensitivitas budaya di mana mereka beroperasi.

Ada berbagai cara yang dapat dilakukan AI untuk menimbulkan kerugian emosional, termasuk keintiman palsu, keterikatan berlebihan, objektifikasi dan komodifikasi tubuh, serta isolasi sosial atau seksual. Hal ini tercakup dalam berbagai inisiatif etis yang disebutkan di atas, termasuk Foundation for Responsible Robotics, Partnership on AI, lembaga AI Now (khususnya mengenai komputasi pengaruh), Deklarasi Montréal, dan Peta Jalan Jaringan Penelitian Robotika Eropa (EURON) (misalnya, bagian mereka tentang risiko humanoids).

Kemungkinan kerugian ini mengemuka ketika mempertimbangkan perkembangan hubungan intim dengan AI, misalnya dalam industri seks. Sistem intim, sebagaimana IEEE menyebutnya, tidak boleh berkontribusi pada seksisme, ketidaksetaraan ras, atau stereotip citra tubuh yang negatif; harus untuk penggunaan positif dan terapeutik; harus menghindari manipulasi seksual atau psikologis terhadap pengguna tanpa persetujuan; tidak boleh dirancang sedemikian rupa sehingga menyebabkan isolasi pengguna dari interaksi manusia; harus dirancang secara transparan mengenai dampaknya terhadap dinamika dan kecemburuan hubungan antarmanusia; tidak boleh mendorong perilaku menyimpang atau kriminal, atau menjadikan praktik seksual ilegal seperti pedofilia atau pemerkosaan sebagai hal yang normal; dan tidak boleh dipasarkan secara komersial sebagai perseorangan (dalam arti hukum atau lainnya).

AI afektif juga terbuka terhadap kemungkinan untuk menipu dan memaksa penggunaannya – para peneliti telah mendefinisikan tindakan AI yang secara halus mengubah perilaku sebagai 'menyenggol', yaitu ketika AI memanipulasi dan memengaruhi penggunaannya secara emosional melalui sistem afektif. Meskipun hal ini mungkin berguna dalam beberapa hal – ketergantungan obat, pola makan sehat – hal ini juga dapat memicu perilaku yang memperburuk kesehatan manusia. Analisis sistematis harus mengkaji etika desain afektif sebelum penerapan; pengguna harus dididik tentang cara mengenali dan membedakan dorongan; pengguna harus memiliki sistem opt-in untuk sistem dorongan otonom; dan kelompok rentan yang tidak dapat memberikan persetujuan, seperti anak-anak, harus mendapatkan perlindungan tambahan. Secara umum, para pemangku kepentingan harus mendiskusikan pertanyaan apakah jalur desain yang mendorong AI, yang cocok untuk penggunaan yang egois atau merugikan, merupakan jalur yang etis untuk dilakukan (IEEE, 2019).

Sebagaimana dikemukakan oleh IEEE (2019), dorongan dapat digunakan oleh pemerintah dan entitas lain untuk mempengaruhi perilaku masyarakat. Apakah pantas secara etis jika robot menggunakan dorongan untuk mendorong, misalnya, perilaku amal atau donasi? Kita harus mengupayakan transparansi penuh mengenai penerima manfaat dari perilaku tersebut, kata IEEE, karena potensi penyalahgunaan.

Masalah lainnya termasuk kecanduan teknologi dan kerugian emosional akibat bias sosial atau gender.

Akuntabilitas dan tanggung jawab

Sebagian besar inisiatif mengamanatkan bahwa AI harus dapat diaudit, untuk memastikan bahwa perancang, produsen, pemilik, dan operator AI bertanggung jawab atas tindakan teknologi atau sistem tersebut, dan dengan demikian dianggap bertanggung jawab atas potensi kerugian yang mungkin ditimbulkannya. Menurut IEEE, hal ini dapat dicapai jika pengadilan mengklarifikasi masalah kesalahan dan tanggung jawab selama fase pengembangan dan penerapan jika memungkinkan, sehingga pihak-pihak yang terlibat memahami kewajiban dan hak mereka; oleh perancang dan pengembang dengan mempertimbangkan keragaman norma budaya yang ada di antara berbagai kelompok pengguna; dengan membangun ekosistem multi-pemangku kepentingan untuk menciptakan norma-norma yang saat ini belum ada, mengingat teknologi berorientasi AI masih terlalu baru; dan dengan menciptakan sistem registrasi dan pencatatan sehingga selalu memungkinkan untuk melacak siapa yang secara hukum bertanggung jawab atas AI tertentu.

Future of Life Institute menangani masalah akuntabilitas melalui Prinsip Asilomar, yang merupakan daftar 23 prinsip panduan yang harus diikuti oleh AI agar menjadi etis dalam jangka pendek dan panjang. Perancang dan pembangun sistem AI yang canggih adalah 'pemangku kepentingan dalam dampak moral dari penggunaan, penyalahgunaan, dan tindakan mereka, dengan tanggung jawab dan peluang untuk membentuk implikasi tersebut' (FLI, 2017); jika AI melakukan kesalahan, AI juga harus dapat memastikan alasannya. Kemitraan AI juga menekankan pentingnya akuntabilitas dalam hal bias. Kita harus peka terhadap fakta bahwa asumsi dan bias ada dalam data dan juga dalam sistem yang dibangun dari data tersebut, dan berusaha untuk tidak meniru asumsi dan bias tersebut – yaitu dengan secara aktif bertanggung jawab dalam membangun AI yang adil dan bebas bias.

Semua inisiatif lainnya menyoroti pentingnya akuntabilitas dan tanggung jawab – baik oleh desainer dan insinyur AI, dan oleh peraturan, hukum, dan masyarakat dalam skala yang lebih besar.

Seks dan Robot

Pada bulan Juli 2017, Foundation for Responsible Robotics menerbitkan laporan tentang 'Our Sexual Future with Robots' (Foundation for Responsible Robotics, 2019). Hal ini bertujuan untuk menyajikan ringkasan obyektif dari berbagai isu dan opini seputar hubungan erat kita dengan teknologi. Banyak negara mengembangkan robot untuk kepuasan seksual; ini sebagian besar cenderung merupakan representasi pornografi dari tubuh manusia – dan sebagian besar adalah perempuan. Representasi ini, jika disertai dengan antropomorfisme manusia, dapat menyebabkan robot dianggap berada di antara makhluk hidup dan benda

mati, terutama jika kepuasan seksual dipadukan dengan unsur keintiman, persahabatan, dan percakapan. Robot juga dapat mempengaruhi persepsi masyarakat tentang gender atau stereotip tubuh, mengikis hubungan dan keintiman antarmanusia, serta menyebabkan isolasi sosial yang lebih besar. Namun, robot juga mempunyai potensi untuk memberikan manfaat emosional dan seksual bagi manusia, misalnya dengan membantu mengurangi kejahatan seksual, dan merehabilitasi korban pemerkosaan atau pelecehan seksual melalui inklusi dalam terapi penyembuhan.

Akses dan transparansi vs. keamanan dan privasi

Kekhawatiran utama terhadap AI adalah transparansi, penjelasan, keamanan, reproduktifitas, dan interpretasi: apakah mungkin untuk mengetahui mengapa dan bagaimana suatu sistem membuat keputusan tertentu, atau mengapa dan bagaimana robot bertindak seperti yang dilakukannya? Hal ini sangat mendesak terutama dalam kasus sistem yang sangat penting bagi keselamatan yang mungkin mempunyai konsekuensi langsung terhadap kerusakan fisik: mobil tanpa pengemudi, misalnya, atau sistem diagnosis medis.

Tanpa transparansi, pengguna mungkin kesulitan memahami sistem yang mereka gunakan – dan konsekuensi yang terkait – dan akan sulit untuk meminta pertanggungjawaban orang-orang terkait. Untuk mengatasi hal ini, IEEE mengusulkan pengembangan standar baru yang merinci tingkat transparansi yang dapat diukur dan diuji, sehingga kepatuhan sistem dapat dinilai secara objektif. Bentuknya mungkin akan berbeda-beda bagi pemangku kepentingan yang berbeda; pengguna robot mungkin memerlukan tombol 'mengapa Anda melakukan itu', sementara lembaga sertifikasi atau penyelidik kecelakaan akan memerlukan akses ke algoritma yang relevan dalam bentuk 'kotak hitam etis' yang memberikan transparansi kegagalan (IEEE, 2019).

AI membutuhkan data untuk terus mempelajari dan mengembangkan pengambilan keputusan otomatisnya. Data ini bersifat pribadi dan dapat digunakan untuk mengidentifikasi identitas fisik, digital, atau virtual individu tertentu (yaitu informasi identitas pribadi, PII). 'Sebagai hasilnya,' tulis IEEE (2017), 'melalui setiap transaksi digital (eksplisit atau teramati) manusia menghasilkan bayangan digital unik dari diri fisik mereka'. Sejauh mana manusia dapat menyadari hak untuk menjaga kerahasiaan informasi tertentu, atau memberikan masukan mengenai bagaimana data tersebut digunakan? Individu mungkin tidak memiliki alat yang tepat untuk mengontrol dan mengembangkan identitas unik mereka serta mengelola implikasi etis yang terkait dengan penggunaan data mereka. Tanpa kejelasan dan edukasi, banyak pengguna AI tidak akan menyadari jejak digital yang mereka ciptakan, dan informasi yang mereka sebarkan ke dunia. Sistem harus diterapkan agar pengguna dapat mengontrol, berinteraksi, dan mengakses data mereka, serta memberi mereka kendali atas kepribadian digital mereka.

PII telah ditetapkan sebagai aset individu (misalnya, berdasarkan Peraturan (UE) 2016/679 di Eropa), dan sistem harus meminta persetujuan eksplisit pada saat data dikumpulkan dan digunakan, untuk melindungi otonomi dan martabat individu. dan hak untuk menyetujui. IEEE menyebutkan kemungkinan 'AI privasi atau agen algoritmik atau wali' yang dipersonalisasi untuk membantu individu mengatur dan mengontrol data pribadi mereka

serta memperkirakan dan mengurangi potensi implikasi etis dari pertukaran data pembelajaran mesin.

Prinsip Asilomar dari Future of Life Institute sepakat dengan IEEE mengenai pentingnya transparansi dan privasi di berbagai aspek: transparansi kegagalan (jika AI gagal, penyebabnya harus diketahui), transparansi peradilan (AI apa pun yang terlibat dalam pengambilan keputusan peradilan) pembuatannya harus memberikan penjelasan yang memuaskan kepada manusia), privasi pribadi (manusia harus memiliki hak untuk mengakses, mengelola, dan mengontrol data yang dikumpulkan dan dibuat oleh AI), dan kebebasan dan privasi (AI tidak boleh secara tidak wajar membatasi kebebasan nyata atau yang dirasakan seseorang) . Saidot mengambil pendekatan yang sedikit lebih luas dan sangat menekankan pentingnya AI yang transparan, akuntabel, dan dapat dipercaya, di mana orang, organisasi, dan sistem cerdas terhubung secara terbuka dan kolaboratif untuk mendorong kerja sama, kemajuan, dan inovasi.

Seluruh inisiatif yang disurvei mengidentifikasi transparansi dan akuntabilitas AI sebagai isu penting. Keseimbangan ini mendasari banyak permasalahan lainnya – seperti keadilan hukum dan peradilan, kompensasi dan hak pekerja, keamanan data dan sistem, kepercayaan publik, dan kerugian sosial.

Otonomi dan agen vs. pasien

Pendekatan AI saat ini tidak dapat disangkal bersifat antroposentris. Hal ini menimbulkan kemungkinan permasalahan seputar perbedaan antara agen moral dan pasien moral, antara yang dibuat-buat dan alami, antara yang mengatur dirinya sendiri dan yang tidak. AI tidak bisa menjadi otonom seperti halnya makhluk hidup dianggap otonom (IEEE, 2019), namun bagaimana kita mendefinisikan otonomi dalam istilah AI? Otonomi mesin menunjukkan bagaimana mesin bertindak dan beroperasi sesuai dengan peraturan, namun setiap upaya untuk menanamkan emosi dan moralitas ke dalam AI 'mengaburkan perbedaan antara agen dan pasien dan dapat mendorong ekspektasi antropomorfik terhadap mesin', tulis IEEE - terutama ketika AI yang diwujudkan mulai terlihat. semakin mirip dengan manusia. Menetapkan perbedaan yang dapat digunakan antara otonomi manusia dan sistem/mesin melibatkan pertanyaan tentang keinginan bebas, keberadaan/penjadian, dan penentuan sebelumnya. Jelas bahwa diskusi lebih lanjut diperlukan untuk memperjelas apa yang dimaksud dengan 'otonomi' dalam kaitannya dengan kecerdasan buatan dan sistem.

Keamanan dan kepercayaan

Jika AI digunakan untuk melengkapi atau menggantikan pengambilan keputusan oleh manusia, terdapat konsensus bahwa AI harus aman, dapat dipercaya, dan dapat diandalkan, serta bertindak dengan integritas.

IEEE mengusulkan untuk menumbuhkan 'pola pikir keselamatan' di kalangan peneliti, untuk 'mengidentifikasi dan mencegah perilaku yang tidak diinginkan dan tidak diantisipasi dalam sistem mereka' dan untuk mengembangkan sistem yang 'dirancang aman'; membentuk dewan peninjau di lembaga-lembaga sebagai sumber daya dan sarana untuk mengevaluasi proyek dan kemajuannya; mendorong komunitas untuk berbagi, untuk menyebarkan berita tentang perkembangan, penelitian, dan alat yang berhubungan dengan keselamatan. Prinsip

Asilomar dari Future of Life Institute menunjukkan bahwa semua yang terlibat dalam pengembangan dan penerapan AI harus dipimpin oleh misi, dengan mengadopsi norma bahwa AI 'hanya boleh dikembangkan untuk memenuhi cita-cita etika yang dianut secara luas, dan untuk kepentingan seluruh umat manusia, bukan untuk kepentingan seluruh umat manusia. satu negara bagian atau organisasi' (Future of Life Institute, 2017). Pendekatan ini akan membangun kepercayaan masyarakat terhadap AI, yang merupakan kunci keberhasilan integrasi AI ke dalam masyarakat.

Perkumpulan AI Jepang mengusulkan agar AI harus bertindak dengan integritas setiap saat, dan AI serta masyarakat harus sungguh-sungguh berupaya untuk belajar dan berkomunikasi satu sama lain. 'Komunikasi yang konsisten dan efektif' akan memperkuat saling pengertian, kata Society, dan '[berkontribusi] pada perdamaian dan kebahagiaan umat manusia secara keseluruhan' (JSAI, 2017). Kemitraan AI sepakat dan berupaya untuk memastikan AI dapat dipercaya dan menciptakan budaya kerja sama, kepercayaan, dan keterbukaan di antara para ilmuwan dan insinyur AI. Institute for Ethical AI & Machine Learning juga menekankan pentingnya dialog; AI menyatukan isu-isu kepercayaan dan privasi dalam delapan prinsip intinya, yang mengamanatkan agar para ahli teknologi AI berkomunikasi dengan para pemangku kepentingan tentang proses dan data yang terlibat untuk membangun kepercayaan dan menyebarkan pemahaman ke seluruh masyarakat.

Kerugian sosial dan keadilan sosial: inklusivitas, bias, dan diskriminasi

Pengembangan AI memerlukan keberagaman sudut pandang. Ada beberapa organisasi yang menetapkan bahwa hal ini harus sejalan dengan sudut pandang masyarakat dan selaras dengan norma, nilai, etika, dan preferensi sosial, bahwa bias dan asumsi tidak boleh dimasukkan ke dalam data atau sistem, dan bahwa AI harus selaras dengan nilai-nilai publik. tujuan, dan perilaku, menghormati keragaman budaya. Inisiatif ini juga berpendapat bahwa semua orang harus memiliki akses terhadap manfaat AI, dan hal ini harus bermanfaat bagi kebaikan bersama. Dengan kata lain, pengembang dan pelaksana AI mempunyai tanggung jawab sosial untuk menanamkan nilai-nilai yang benar ke dalam AI dan memastikan bahwa nilai-nilai tersebut tidak menyebabkan atau memperburuk kerugian yang ada atau di masa depan terhadap bagian masyarakat mana pun.

IEEE menyarankan untuk terlebih dahulu mengidentifikasi norma-norma sosial dan moral dari komunitas tertentu di mana AI akan diterapkan, dan norma-norma di sekitar tugas atau layanan tertentu yang akan ditawarkan; merancang AI dengan gagasan 'pembaruan norma', mengingat norma tidak statis dan AI harus berubah secara dinamis dan transparan seiring dengan budaya; dan mengidentifikasi cara-cara masyarakat menyelesaikan konflik norma, dan melengkapi AI dengan sistem yang dapat digunakan untuk menyelesaikan konflik norma dengan cara yang serupa dan transparan. Hal ini harus dilakukan secara kolaboratif dan lintas upaya penelitian yang beragam, dengan kehati-hatian dalam mengevaluasi dan menilai potensi bias yang merugikan kelompok sosial tertentu.

Sebuah 'kotak hitam etis'

Inisiatif termasuk UNI Global Union dan IEEE menyarankan untuk melengkapi sistem AI dengan 'kotak hitam etis': sebuah perangkat yang dapat mencatat informasi tentang sistem

tersebut untuk memastikan akuntabilitas dan transparansi, namun juga mencakup data yang jelas tentang pertimbangan etis yang dibangun ke dalam sistem. sejak awal (UNI Global Union, n.d.).

Beberapa inisiatif – seperti AI4All dan AI Now Institute – secara eksplisit mengadvokasi inklusi AI yang adil, beragam, merata, dan non-diskriminatif di semua tahap, dengan fokus pada dukungan bagi kelompok yang kurang terwakili. Saat ini, program gelar terkait AI tidak membekali calon pengembang dan desainer dengan pengetahuan etika yang sesuai (IEEE, 2017), dan lingkungan perusahaan serta praktik bisnis tidak memberdayakan secara etis, dengan kurangnya peran ahli etika senior yang dapat mengarahkan dan mendukung inovasi berbasis nilai.

Dalam skala global, kesenjangan kesenjangan antara negara maju dan negara berkembang sangatlah besar. Meskipun AI mempunyai manfaat yang besar dalam hal kemanusiaan, AI tidak boleh memperlebar kesenjangan ini atau memperburuk kemiskinan, buta huruf, ketidaksetaraan gender dan etnis, atau mengganggu lapangan kerja dan tenaga kerja secara tidak proporsional. IEEE menyarankan untuk mengambil tindakan dan berinvestasi untuk mengurangi kesenjangan ketimpangan; mengintegrasikan tanggung jawab sosial perusahaan (CSR) ke dalam pengembangan dan pemasaran; mengembangkan struktur kekuasaan yang transparan; memfasilitasi dan berbagi pengetahuan dan penelitian robotika dan AI; dan secara umum menjaga AI tetap sejalan dengan Tujuan Pembangunan Berkelanjutan AS. Teknologi AI harus tersedia secara merata di seluruh dunia melalui standarisasi global dan perangkat lunak sumber terbuka, dan diskusi lintas disiplin harus diadakan mengenai pendidikan dan pelatihan AI yang efektif (IEEE, 2019).

Serangkaian pedoman etika yang diterbitkan oleh Masyarakat Jepang untuk AI menekankan, antara lain, pentingnya a) kontribusi terhadap kemanusiaan, dan b) tanggung jawab sosial. AI harus bertindak demi kepentingan publik, menghormati keragaman budaya, dan selalu digunakan secara adil dan setara.

Foundation for Responsible Robotics mencakup Komitmen terhadap Keberagaman dalam upayanya mewujudkan AI yang bertanggung jawab; Partnership on AI memperingatkan tentang 'titik buta yang serius' dalam mengabaikan adanya bias dan asumsi yang tersembunyi dalam data; Saidot bertujuan untuk memastikan bahwa, meskipun nilai-nilai sosial kita kini 'semakin dimediasi oleh algoritme', AI tetap berpusat pada manusia (Saidot, 2019); Future of Life Institute menyoroti perlunya AI yang dijiwai dengan nilai-nilai kemanusiaan berupa keragaman budaya dan hak asasi manusia; dan Institute for Ethical AI & Machine Learning menyertakan 'evaluasi bias' untuk memantau bias dalam pengembangan dan produksi AI. Bahaya bias dan asumsi manusia merupakan risiko yang sering teridentifikasi dan akan menyertai perkembangan AI yang sedang berlangsung.

Kerugian finansial: Peluang ekonomi dan lapangan kerja

AI dapat mengganggu perekonomian dan menyebabkan hilangnya pekerjaan atau gangguan kerja bagi banyak orang, dan akan berdampak pada hak-hak pekerja dan strategi perpindahan karena banyak jenis pekerjaan menjadi otomatis (dan hilang seiring dengan perubahan bisnis).

Selain itu, dibandingkan hanya berfokus pada jumlah pekerjaan yang hilang atau diperoleh, struktur ketenagakerjaan tradisional perlu diubah untuk memitigasi dampak otomatisasi dan mempertimbangkan kompleksitas ketenagakerjaan.

Perubahan teknologi terjadi terlalu cepat sehingga tenaga kerja tradisional tidak dapat mengimbangnya tanpa pelatihan ulang. Para pekerja harus dilatih untuk dapat beradaptasi, menurut IEEE (2019), dan keahlian baru, dengan menerapkan strategi pengganti bagi mereka yang tidak dapat dilatih kembali, dan program pelatihan dilaksanakan di tingkat sekolah menengah atas atau sebelumnya untuk meningkatkan akses terhadap masa depan. pekerjaan. UNI Global Union menyerukan dibentuknya badan tata kelola AI yang beretika dan multi-pemangku kepentingan di tingkat global dan regional, yang menyatukan para desainer, produsen, pengembang, peneliti, serikat pekerja, pengacara, organisasi masyarakat sipil, pemilik, dan pemberi kerja. AI harus memberikan manfaat dan memberdayakan masyarakat secara luas dan setara, dengan kebijakan yang diterapkan untuk menjembatani kesenjangan ekonomi, teknologi, dan sosial digital, dan memastikan transisi yang adil dengan dukungan terhadap kebebasan dan hak-hak dasar.

AI Now Institute bekerja dengan beragam kelompok pemangku kepentingan untuk lebih memahami implikasi AI terhadap tenaga kerja, termasuk otomatisasi dan integrasi tahap awal AI yang mengubah sifat pekerjaan dan kondisi kerja di berbagai sektor. Future Society secara khusus menanyakan bagaimana AI akan mempengaruhi profesi hukum: 'Jika sistem AI terbukti lebih unggul dibandingkan pengacara manusia dalam aspek-aspek tertentu dalam pekerjaan hukum, apa implikasi etika dan profesionalnya terhadap praktik hukum?' (Masyarakat Masa Depan, 2019)

AI di tempat kerja akan berdampak lebih dari sekedar keuangan pekerja, dan mungkin menawarkan berbagai peluang positif. Sebagaimana ditetapkan oleh IEEE (2019), AI dapat menawarkan solusi potensial terhadap bias di tempat kerja – jika dikembangkan dengan mempertimbangkan hal ini, seperti yang disebutkan di atas – dan mengungkap kekurangan dalam pengembangan produk, sehingga memungkinkan perbaikan proaktif dalam fase desain (dibandingkan dengan perbaikan retroaktif).

'RRI adalah sebuah proses yang transparan dan interaktif di mana para pelaku masyarakat dan inovator menjadi saling responsif satu sama lain dengan tujuan untuk menerima (secara etis), keberlanjutan dan keinginan masyarakat terhadap proses inovasi dan produk-produknya yang dapat dipasarkan (untuk memungkinkan tertanamnya produk-produk tersebut secara tepat). kemajuan ilmu pengetahuan dan teknologi dalam masyarakat kita.' (Von Schomberg, 2013)

Penelitian dan inovasi yang bertanggung jawab (RRI)

RRI adalah bidang yang sedang berkembang, khususnya di UE, yang memanfaatkan etika klasik untuk menyediakan alat yang dapat digunakan untuk mengatasi masalah etika sejak awal proyek. Ketika dimasukkan ke dalam fase desain proyek, RRI meningkatkan kemungkinan desain menjadi relevan dan kuat dalam hal keselarasan etika. Banyak penyandang dana dan organisasi penelitian memasukkan RRI dalam pernyataan misi mereka dan dalam upaya penelitian dan inovasi mereka (IEEE, 2019).

Keabsahan dan keadilan

Beberapa inisiatif menjawab perlunya AI yang sah, adil, adil, dan tunduk pada tata kelola dan peraturan yang tepat dan bersifat pre-emptive. Banyaknya masalah etika kompleks seputar AI diterjemahkan secara langsung dan tidak langsung menjadi tantangan hukum tersendiri. Bagaimana seharusnya AI diberi label: sebagai sebuah produk? Seekor binatang? Seseorang? Sesuatu yang baru?

IEEE menyimpulkan bahwa AI tidak boleh diberikan 'kepribadian' pada tingkat apa pun, dan bahwa, meskipun pengembangan, desain, dan distribusi AI harus sepenuhnya mematuhi semua hukum internasional dan domestik yang berlaku, masih banyak pekerjaan yang harus dilakukan untuk mendefinisikan dan menerapkan AI. perundang-undangan yang relevan. Masalah hukum terbagi dalam beberapa kategori: status hukum, penggunaan oleh pemerintah (transparansi, hak individu), akuntabilitas hukum atas kerugian, dan transparansi, akuntabilitas, dan verifikasi. IEEE menyarankan agar AI tetap tunduk pada peraturan hukum properti yang berlaku; bahwa para pemangku kepentingan harus mengidentifikasi jenis-jenis keputusan yang tidak boleh didelegasikan kepada AI, dan memastikan kontrol manusia yang efektif atas keputusan-keputusan tersebut melalui peraturan dan standar; bahwa undang-undang yang ada harus dicermati dan ditinjau ulang untuk mencari mekanisme yang secara praktis dapat memberikan otonomi hukum bagi AI; dan bahwa produsen dan operator harus diwajibkan mematuhi hukum yang berlaku di semua yurisdiksi tempat AI dapat beroperasi. Mereka juga merekomendasikan agar pemerintah menilai kembali status hukum AI seiring dengan semakin canggihnya AI, dan bekerja sama dengan regulator, pelaku masyarakat dan industri serta pemangku kepentingan lainnya untuk memastikan bahwa kepentingan kemanusiaan – dan bukan pengembangan sistem itu sendiri – tetap menjadi prinsip panduan.

Kontrol dan penggunaan – atau penyalahgunaan – AI secara etis

Dengan AI baru yang lebih canggih dan kompleks, muncul kemungkinan penyalahgunaan yang lebih canggih dan kompleks. Data pribadi dapat digunakan untuk tujuan jahat atau untuk mendapatkan keuntungan, sistem berisiko diretas, dan teknologi dapat digunakan secara eksploitatif. Hal ini terkait dengan penggunaan informasi dan kesadaran masyarakat: saat kita memasuki era baru AI, dengan munculnya sistem dan teknologi baru yang belum pernah diterapkan sebelumnya, masyarakat harus selalu mengetahui risiko yang mungkin timbul baik dari penggunaan atau penyalahgunaan. ini.

IEEE menyarankan cara-cara baru untuk mendidik masyarakat mengenai masalah etika dan keamanan, misalnya peringatan 'privasi data' pada perangkat pintar yang mengumpulkan data pribadi; menyampaikan pendidikan ini dengan cara yang terukur dan efektif; dan mendidik pemerintah, anggota parlemen, dan lembaga penegak hukum seputar masalah ini, sehingga mereka dapat bekerja sama dengan masyarakat – seperti halnya petugas polisi yang memberikan ceramah keselamatan di sekolah – dan menghindari ketakutan dan kebingungan (IEEE, 2019).

Masalah lainnya termasuk manipulasi perilaku dan data. Manusia harus mempertahankan kendali atas AI dan menentang subversi. Sebagian besar inisiatif yang ditinjau menandai hal ini sebagai potensi masalah yang dihadapi AI seiring perkembangannya,

dan menandai bahwa AI harus berperilaku dengan cara yang dapat diprediksi dan dapat diandalkan, dengan cara yang tepat untuk mendapatkan ganti rugi, dan harus melalui validasi dan pengujian. AI juga harus bekerja demi kebaikan umat manusia, tidak boleh mengeksploitasi manusia, dan ditinjau secara berkala oleh pakar manusia.

Kepribadian dan AI

Persoalan apakah AI layak atau tidak memiliki 'kepribadian' menjadi perdebatan seputar akuntabilitas, otonomi, dan tanggung jawab: apakah AI itu sendiri yang bertanggung jawab atas tindakan dan konsekuensinya, atau orang yang menciptakannya?

Konsep ini, alih-alih membiarkan robot dianggap manusia dalam pengertian manusia, akan menempatkan robot pada tingkat hukum yang sama dengan perusahaan. Perlu dicatat bahwa badan hukum suatu perusahaan saat ini dapat melindungi orang-orang yang mendukung mereka dari dampak hukum. Namun, Persatuan Global UNI menegaskan bahwa tanggung jawab hukum ada pada penciptanya, bukan robot itu sendiri, dan menyerukan larangan untuk mengaitkan tanggung jawab pada robot.

Kerusakan dan kelestarian lingkungan

Produksi, pengelolaan, dan penerapan AI harus berkelanjutan dan menghindari kerusakan lingkungan. Hal ini juga terkait dengan konsep kesejahteraan; Aspek utama kesejahteraan yang diakui adalah lingkungan hidup, yang berkaitan dengan udara, keanekaragaman hayati, perubahan iklim, kualitas tanah dan air, dan sebagainya (IEEE, 2019). IEEE (EAD, 2019) menyatakan bahwa AI tidak boleh membahayakan sistem alam bumi atau memperburuk degradasinya, dan berkontribusi dalam mewujudkan pengelolaan, pelestarian, dan/atau pemulihan sistem alam bumi yang berkelanjutan. UNI Global Union menyatakan bahwa AI harus mengutamakan manusia dan bumi, berupaya melindungi dan bahkan meningkatkan keanekaragaman hayati dan ekosistem planet kita (UNI Global Union, n.d.). Foundation for Responsible Robotics mengidentifikasi sejumlah potensi penggunaan AI di tahun-tahun mendatang, mulai dari peran pertanian dan peternakan hingga pemantauan perubahan iklim dan perlindungan spesies yang terancam punah. Hal ini memerlukan kebijakan yang bertanggung jawab dan terinformasi untuk mengatur AI dan robotika, kata Foundation, untuk memitigasi risiko dan mendukung inovasi dan pengembangan yang berkelanjutan.

Penggunaan informasi: pendidikan dan kesadaran publik

Masyarakat harus dididik tentang penggunaan, penyalahgunaan, dan potensi bahaya AI, melalui partisipasi masyarakat, komunikasi, dan dialog dengan masyarakat. Persoalan persetujuan – dan seberapa banyak seseorang dapat memberikan secara wajar dan sadar – adalah inti dari hal ini. Misalnya, IEEE mengemukakan beberapa contoh di mana persetujuan kurang jelas dibandingkan dengan etika: bagaimana jika data pribadi seseorang digunakan untuk membuat kesimpulan yang membuat mereka merasa tidak nyaman atau tidak menyadarinya? Bisakah persetujuan diberikan ketika sistem tidak berinteraksi langsung dengan individu? Edisi terakhir ini diberi nama 'Internet of Other People's Things' (IEEE, 2019). Lingkungan korporasi juga mengangkat isu ketidakseimbangan kekuasaan; banyak karyawan tidak memiliki persetujuan yang jelas tentang bagaimana data pribadi mereka – termasuk data

kesehatan – digunakan oleh perusahaan mereka. Untuk mengatasi hal ini, IEEE (2017) menyarankan penilaian dampak data karyawan untuk menangani nuansa perusahaan ini dan memastikan bahwa tidak ada data yang dikumpulkan tanpa persetujuan karyawan. Data juga hanya boleh dikumpulkan dan digunakan untuk tujuan tertentu, dinyatakan secara eksplisit, sah, selalu diperbarui, diproses secara sah, dan tidak disimpan untuk jangka waktu lebih lama dari yang diperlukan. Dalam kasus di mana subjek tidak memiliki hubungan langsung dengan sistem pengumpulan data, persetujuan harus bersifat dinamis, dan sistem dirancang untuk menafsirkan preferensi data dan batasan pengumpulan dan penggunaan.

Untuk meningkatkan kesadaran dan pemahaman tentang AI, mahasiswa sarjana dan pascasarjana harus dididik tentang AI dan hubungannya dengan pembangunan manusia berkelanjutan, kata IEEE. Secara khusus, kurikulum dan kompetensi inti harus didefinisikan dan disiapkan; program gelar yang berfokus pada bidang teknik dalam pembangunan internasional dan bantuan kemanusiaan harus dihadapkan pada potensi penerapan AI; dan kesadaran harus ditingkatkan mengenai peluang dan risiko yang dihadapi oleh Negara-negara Berpenghasilan Menengah Bawah dalam penerapan AI dalam upaya kemanusiaan di seluruh dunia.

Banyak inisiatif yang berfokus pada hal ini, termasuk Foundation for Responsible Robotics, Partnership on AI, Japanese Society for AI Ethical Guidelines, Future Society, dan AI Now Institute; mereka dan pihak lain berpendapat bahwa dialog yang jelas, terbuka dan transparan antara AI dan masyarakat adalah kunci untuk menciptakan pemahaman, penerimaan, dan kepercayaan.

Risiko eksistensial

Menurut Future of Life Institute, masalah utama yang ada di sekitar AI 'bukanlah kedengkian, melainkan kompetensi' – AI akan terus belajar ketika mereka berinteraksi dengan orang lain dan mengumpulkan data, sehingga mengarahkan mereka untuk memperoleh kecerdasan seiring berjalannya waktu dan berpotensi mengembangkan tujuan-tujuan yang ada pada diri mereka sendiri. bertentangan dengan manusia.

'Anda mungkin bukan pembenci semut jahat yang menginjak semut karena niat jahat,' 'tetapi jika Anda bertanggung jawab atas proyek energi hijau pembangkit listrik tenaga air dan ada sarang semut di wilayah tersebut yang akan terendam banjir, sayang sekali bagi semut. . Tujuan utama penelitian keamanan AI adalah untuk tidak menempatkan manusia pada posisi semut (The Future of Life Institute, 2019).

AI juga memberikan ancaman dalam bentuk sistem senjata otonom (AWS). Karena hal ini dirancang untuk menyebabkan kerusakan fisik, hal ini menimbulkan banyak permasalahan etika. IEEE (2019) memaparkan sejumlah rekomendasi untuk memastikan bahwa AWS tunduk pada kendali manusia yang berarti: rekomendasi tersebut menyarankan jalur audit untuk menjamin akuntabilitas dan kendali; sistem pembelajaran adaptif yang dapat menjelaskan alasannya secara transparan dan mudah dipahami; bahwa manusia yang menjalankan sistem otonom dapat diidentifikasi, bertanggung jawab, dan sadar akan dampak pekerjaan mereka; bahwa perilaku otonom dapat diprediksi; dan bahwa kode etik profesi dikembangkan untuk mengatasi perkembangan sistem otonom – terutama yang dimaksudkan untuk menimbulkan

kerugian. Upaya untuk mencapai AWS dapat mengarah pada perlombaan senjata internasional dan stabilitas geopolitik; oleh karena itu, IEEE merekomendasikan bahwa sistem yang dirancang untuk bertindak di luar batas kendali atau penilaian manusia adalah tidak etis dan melanggar hak asasi manusia serta akuntabilitas hukum atas penggunaan senjata.

Mengingat potensinya yang sangat merugikan masyarakat, kekhawatiran ini harus dikendalikan dan diatur terlebih dahulu, kata Foundation for Responsible Robotics. Inisiatif lain yang mencakup risiko ini secara eksplisit mencakup UNI Global Union dan Future of Life Institute, yang memperingatkan terhadap perlombaan senjata dalam senjata otonom yang mematikan, dan menyerukan upaya perencanaan dan mitigasi terhadap kemungkinan risiko jangka panjang. Kita harus menghindari asumsi yang kuat mengenai batas atas kemampuan AI di masa depan, menegaskan Prinsip Asilomar FLI, dan menyadari bahwa AI yang canggih mewakili perubahan besar dalam sejarah kehidupan di Bumi.

3.3. STUDI KASUS

Studi kasus: robot perawatan kesehatan

Kecerdasan Buatan dan robotika dengan cepat bergerak ke bidang perawatan kesehatan dan akan semakin berperan dalam diagnosis dan perawatan klinis. Misalnya, saat ini atau dalam waktu dekat, robot akan membantu mendiagnosis pasien; pelaksanaan operasi sederhana; dan pemantauan kesehatan pasien dan kesejahteraan mental di fasilitas perawatan jangka pendek dan jangka panjang. Mereka juga dapat memberikan intervensi fisik dasar, bekerja sebagai perawat pendamping, mengingatkan pasien untuk meminum obatnya, atau membantu pasien dengan mobilitasnya. Di beberapa bidang dasar kedokteran, seperti diagnostik citra medis, pembelajaran mesin telah terbukti menyamai atau bahkan melampaui kemampuan kita dalam mendeteksi penyakit.

AI yang diwujudkan, atau robot, sudah terlibat dalam sejumlah fungsi yang memengaruhi keselamatan fisik manusia. Pada bulan Juni 2005, robot bedah di sebuah rumah sakit di Philadelphia tidak berfungsi selama operasi prostat, sehingga melukai pasien. Pada bulan Juni 2015, seorang pekerja di pabrik Volkswagen di Jerman tewas tertabrak robot di jalur produksi. Pada bulan Juni 2016, sebuah mobil Tesla yang beroperasi dalam mode autopilot bertabrakan dengan truk besar, menewaskan penumpang mobil tersebut.

Ketika robot menjadi lebih umum, potensi bahaya di masa depan akan meningkat, terutama dalam kasus mobil tanpa pengemudi, robot bantu, dan drone, yang akan mengambil keputusan yang mempunyai konsekuensi nyata bagi keselamatan dan kesejahteraan manusia. Pertaruhannya jauh lebih tinggi dengan AI yang diwujudkan dibandingkan dengan perangkat lunak belaka, karena robot memiliki bagian-bagian yang bergerak dalam ruang fisik (Lin et al., 2017). Robot apa pun yang bagian fisiknya bergerak menimbulkan risiko, terutama bagi orang-orang yang rentan seperti anak-anak dan orang tua.

Keamanan

Sekali lagi, mungkin isu etika terpenting yang timbul dari pertumbuhan AI dan robotika dalam layanan kesehatan adalah keselamatan dan penghindaran bahaya. Sangat penting bahwa robot tidak membahayakan manusia, dan robot harus aman untuk digunakan. Poin ini

sangat penting khususnya dalam bidang layanan kesehatan yang menangani orang-orang yang rentan, seperti orang sakit, orang lanjut usia, dan anak-anak.

Teknologi perawatan kesehatan digital menawarkan potensi untuk meningkatkan akurasi diagnosis dan pengobatan, namun untuk sepenuhnya membangun keamanan jangka panjang dan investasi kinerja teknologi dalam uji klinis diperlukan. Efek samping yang melemahkan dari implan jaring vagina dan perjuangan hukum yang terus berlanjut melawan produsen (The Washington Post, 2019), merupakan contoh penolakan terhadap pengujian yang lebih cepat, meskipun hal ini menyebabkan penundaan dalam inovasi layanan kesehatan. Investasi dalam uji klinis akan sangat penting untuk menerapkan inovasi layanan kesehatan yang ditawarkan sistem AI dengan aman.

Pemahaman pengguna

Penerapan AI yang benar oleh profesional kesehatan penting untuk memastikan keselamatan pasien. Misalnya, asisten robot bedah yang tepat 'da Vinci' telah terbukti merupakan alat yang berguna dalam meminimalkan pemulihan bedah, namun memerlukan operator yang terlatih (The Conversation, 2018).

Pergeseran dalam keseimbangan keterampilan dalam angkatan kerja medis diperlukan, dan penyedia layanan kesehatan bersiap untuk mengembangkan literasi digital staf mereka selama dua dekade mendatang (NHS' Topol Review, 2009). Dengan semakin tertanamnya genomik dan pembelajaran mesin dalam diagnosis dan pengambilan keputusan medis, para profesional kesehatan perlu melek digital untuk memahami setiap alat teknologi dan menggunakannya dengan tepat. Penting bagi pengguna untuk memercayai AI yang disajikan namun tetap menyadari kekuatan dan kelemahan masing-masing alat, serta mengetahui kapan validasi diperlukan. Misalnya, studi pembelajaran mesin yang secara umum akurat untuk memprediksi risiko komplikasi pada pasien pneumonia secara keliru menganggap penderita asma memiliki risiko rendah. Kesimpulan ini dicapai karena pasien pneumonia asma langsung dibawa ke perawatan intensif, dan perawatan di tingkat yang lebih tinggi dapat menghindari komplikasi. Oleh karena itu, rekomendasi yang tidak akurat dari algoritma tersebut ditolak (Pulmonology Advisor, 2017).

Namun, masih dipertanyakan sejauh mana individu perlu memahami bagaimana sistem AI sampai pada prediksi tertentu agar dapat membuat keputusan yang mandiri dan tepat. Meskipun pemahaman mendalam tentang matematika diwajibkan, kompleksitas dan sifat pembelajaran algoritma pembelajaran mesin sering kali menghalangi kemampuan untuk memahami bagaimana suatu kesimpulan dibuat dari kumpulan data — yang disebut 'kotak hitam' (Schönberger, 2019). Dalam kasus seperti ini, salah satu cara yang mungkin dilakukan untuk memastikan keselamatan adalah dengan memberikan lisensi kepada AI untuk melakukan prosedur medis tertentu, dan 'membubarkan' AI jika terjadi sejumlah kesalahan.

Perlindungan data

Data medis pribadi yang diperlukan untuk algoritme layanan kesehatan mungkin berisiko. Misalnya, terdapat kekhawatiran bahwa data yang dikumpulkan oleh pelacak kebugaran dapat dijual kepada pihak ketiga, seperti perusahaan asuransi, yang dapat menggunakan data tersebut untuk menolak jaminan layanan kesehatan (National Public

Radio, 2018). Peretas juga menjadi perhatian utama karena memberikan keamanan yang memadai untuk sistem yang diakses oleh sejumlah personel medis merupakan suatu masalah.

Pengumpulan data medis pribadi sangat penting bagi algoritme pembelajaran mesin untuk memajukan intervensi layanan kesehatan, namun kesenjangan dalam tata kelola informasi menjadi penghalang terhadap pembagian data yang bertanggung jawab dan etis. Kerangka kerja yang jelas tentang bagaimana staf layanan kesehatan dan peneliti menggunakan data, seperti genomik, dengan cara yang menjaga kerahasiaan pasien diperlukan untuk membangun kepercayaan publik dan memungkinkan kemajuan dalam algoritma layanan kesehatan (NHS' Topol Review, 2009).

Tanggung jawab hukum

Meskipun AI berjanji untuk mengurangi jumlah kecelakaan medis, ketika terjadi masalah, tanggung jawab hukum harus ditetapkan. Jika peralatan dapat dibuktikan rusak maka produsen bertanggung jawab, namun seringkali sulit untuk menentukan apa yang salah selama suatu prosedur dan apakah ada orang, petugas medis atau mesin, yang harus disalahkan. Misalnya, ada tuntutan hukum terhadap asisten bedah da Vinci (Mercury News, 2017), namun robot tersebut tetap diterima secara luas (The Conversation, 2018).

Dalam kasus algoritme 'kotak hitam' yang tidak memungkinkan untuk memastikan bagaimana suatu kesimpulan dicapai, maka akan sulit untuk menetapkan kelalaian di pihak pembuat algoritme (Hart, 2018).

Untuk saat ini, AI digunakan sebagai bantuan dalam pengambilan keputusan oleh para ahli, sehingga para ahli tetap menjadi pihak yang bertanggung jawab dalam banyak kasus. Misalnya, dalam kasus pneumonia yang disebutkan di atas, jika staf medis hanya mengandalkan AI dan memulangkan pasien pneumonia asma tanpa menerapkan pengetahuan khusus mereka, maka hal tersebut merupakan tindakan kelalaian mereka (Pulmonology Advisor, 2017; International Journal of Hukum dan Teknologi Informasi, 2019).

Dalam waktu dekat, kelalaian AI dapat dianggap sebagai kelalaian. Misalnya, di negara-negara kurang berkembang yang kekurangan tenaga profesional medis, tidak adanya AI yang dapat mendeteksi penyakit mata akibat diabetes dan mencegah kebutaan, karena kurangnya dokter mata yang dapat memberikan persetujuan atas diagnosis, dapat dianggap tidak etis (The Guardian, 2019; International Jurnal Hukum dan Teknologi Informasi, 2019).

Bias

Non-diskriminasi adalah salah satu nilai fundamental UE (lihat Pasal 21 Piagam Hak-Hak Fundamental UE), namun algoritme pembelajaran mesin dilatih pada kumpulan data yang sering kali memiliki lebih sedikit data yang tersedia tentang kelompok minoritas, sehingga dapat menjadi bias. Hal ini dapat berarti bahwa algoritme yang dilatih untuk mendiagnosis suatu kondisi cenderung kurang akurat untuk pasien etnis; misalnya, dalam kumpulan data yang digunakan untuk melatih model untuk mendeteksi kanker kulit, kurang dari 5 persen gambar berasal dari individu berkulit gelap, sehingga menimbulkan risiko kesalahan diagnosis bagi orang kulit berwarna (The Atlantic, 2018).

Untuk memastikan diagnosis yang paling akurat diberikan kepada orang-orang dari semua etnis, bias algoritmik harus diidentifikasi dan dipahami. Bahkan dengan pemahaman

yang jelas tentang desain model, ini adalah tugas yang sulit karena sifat pembelajaran mesin yang bersifat 'kotak hitam' yang disebutkan di atas. Namun, berbagai kode etik dan inisiatif telah diperkenalkan untuk menemukan bias sebelumnya.

Partnership on AI, sebuah kelompok industri yang berfokus pada etika diluncurkan oleh Google, Facebook, Amazon, IBM dan Microsoft (The Guardian, 2016) — meskipun, yang mengkhawatirkan, dewan ini tidak terlalu beragam.

Kesetaraan akses

Teknologi kesehatan digital, seperti pelacak kebugaran dan pompa insulin, memberikan pasien peluang untuk berpartisipasi aktif dalam layanan kesehatan mereka sendiri. Beberapa orang berharap bahwa teknologi ini akan membantu mengatasi kesenjangan kesehatan yang disebabkan oleh buruknya pendidikan, pengangguran, dan sebagainya. Namun, terdapat risiko bahwa individu yang tidak mampu memperoleh teknologi yang diperlukan atau tidak memiliki 'literasi digital' yang diperlukan akan tersingkir, sehingga memperkuat kesenjangan kesehatan yang ada (The Guardian, 2019).

Program Pelebaran Partisipasi Digital Layanan Kesehatan Nasional Inggris adalah salah satu contoh bagaimana layanan kesehatan berupaya mengurangi kesenjangan kesehatan, dengan membantu jutaan orang di Inggris yang tidak memiliki keterampilan untuk mengakses layanan kesehatan digital. Program seperti ini sangat penting dalam memastikan kesetaraan akses terhadap layanan kesehatan, dan juga dalam meningkatkan data dari kelompok minoritas yang diperlukan untuk mencegah bias dalam algoritma layanan kesehatan yang dibahas di atas.

Kualitas perawatan

'Terdapat potensi luar biasa dalam teknologi layanan kesehatan digital untuk meningkatkan akurasi diagnosis dan perawatan, efisiensi layanan, dan alur kerja bagi para profesional layanan kesehatan' (NHS' Topol Review, 2019).

Jika diperkenalkan dengan pemikiran dan pedoman yang cermat, robot pendamping dan perawatan, misalnya, dapat meningkatkan kehidupan para lansia, mengurangi ketergantungan mereka, dan menciptakan lebih banyak peluang untuk interaksi sosial. Bayangkan sebuah robot perawatan di rumah yang dapat: mengingatkan Anda untuk meminum obat; ambilkan barang untuk Anda jika Anda terlalu lelah atau sudah di tempat tidur; melakukan tugas pembersihan sederhana; dan membantu Anda tetap berhubungan dengan keluarga, teman, dan penyedia layanan kesehatan melalui tautan video.

Namun, muncul pertanyaan apakah robot yang 'dingin' dan tanpa emosi benar-benar dapat menggantikan sentuhan empati manusia. Hal ini khususnya terjadi dalam perawatan jangka panjang terhadap populasi yang rentan dan sering kali kesepian, yang mendapatkan pendampingan dasar dari para pengasuh. Interaksi manusia sangat penting bagi orang lanjut usia, karena penelitian menunjukkan bahwa jaringan sosial yang luas memberikan perlindungan terhadap demensia. Saat ini, robot masih jauh dari kata sahabat sejati. Meski pun mereka dapat berinteraksi dengan orang lain, dan bahkan menunjukkan simulasi emosi, kemampuan percakapan mereka masih sangat terbatas, dan mereka bukanlah pengganti cinta dan perhatian manusia. Beberapa orang mungkin mengatakan bahwa melarang orang lanjut

usia untuk melakukan kontak dengan manusia adalah tindakan yang tidak etis, dan bahkan merupakan suatu bentuk kekejaman.

Dan apakah menelantarkan lansia kita ke perawatan mesin dingin akan menjadikan mereka obyektif (menurunkan) mereka, atau manusia yang merawat mereka? Sangat penting bahwa robot tidak membuat orang lanjut usia merasa seperti benda, atau bahkan memiliki lebih sedikit kendali atas hidup mereka dibandingkan ketika mereka bergantung pada manusia – jika tidak, mereka mungkin merasa seperti 'gumpalan benda mati: untuk didorong, diangkat, dipompa atau terkuras, tanpa referensi yang tepat pada fakta bahwa mereka adalah makhluk hidup (Kitwood 1997).

Pada prinsipnya, otonomi, martabat dan penentuan nasib sendiri semuanya dapat dihormati sepenuhnya melalui penerapan mesin, namun tidak jelas apakah penerapan peran-peran ini dalam bidang kedokteran yang sensitif akan dianggap dapat diterima. Misalnya, seorang dokter menggunakan perangkat telepresence untuk memberikan prognosis kematian pada pasien California; tidak mengherankan jika keluarga pasien marah dengan pendekatan layanan kesehatan yang tidak bersifat pribadi ini (The Independent, 2019). Di sisi lain, ada pendapat bahwa teknologi baru, seperti aplikasi pemantauan kesehatan, akan memberikan waktu lebih bagi staf untuk berinteraksi lebih langsung dengan pasien, sehingga berpotensi meningkatkan kualitas layanan secara keseluruhan.

Tipu muslihat

Sejumlah 'carebot' dirancang untuk interaksi sosial dan sering disebut-sebut memberikan peran terapi emosional. Misalnya, panti jompo telah menemukan bahwa interaksi seperti binatang yang dilakukan anak anjing laut robotik dengan penghuninya dapat mencerahkan suasana hati mereka, mengurangi kecemasan, dan benar-benar meningkatkan kemampuan bersosialisasi penghuni dengan manusia yang merawat mereka. Namun, batas antara kenyataan dan imajinasi menjadi kabur bagi pasien demensia. Jadi, apakah tidak jujur jika memperkenalkan robot sebagai hewan peliharaan dan mendorong keterlibatan sosial-emosional? (KALW, 2015) Dan jika demikian, apakah secara moral dapat dibenarkan?

Robot pendamping dan robot hewan peliharaan dapat mengurangi kesepian di antara orang lanjut usia, namun hal ini mengharuskan mereka untuk percaya, dalam beberapa hal, bahwa robot adalah makhluk hidup yang peduli terhadap mereka dan memiliki perasaan — sebuah penipuan mendasar. Turkle dkk. (2006) berpendapat bahwa 'fakta bahwa orang tua, kakek-nenek, dan anak-anak kita mungkin mengatakan 'Aku cinta kamu' kepada robot yang akan mengatakan 'Aku cinta kamu' sebagai balasannya, sungguh tidak nyaman; hal ini menimbulkan pertanyaan tentang keaslian yang kami butuhkan dari teknologi kami'. Wallach dan Allen (2009) setuju bahwa robot yang dirancang untuk mendeteksi gerakan sosial manusia dan merespons dengan cara yang sama, semuanya menggunakan teknik yang bisa dibilang merupakan bentuk penipuan. Agar seseorang dapat memperoleh manfaat dari memiliki robot hewan peliharaan, mereka harus terus-menerus menipu diri sendiri tentang sifat sebenarnya dari hubungan mereka dengan hewan tersebut. Terlebih lagi, mendorong orang lanjut usia untuk berinteraksi dengan mainan robot memiliki efek menjadikan mereka kekanak-kanakan.

Otonomi

Penting agar robot layanan kesehatan benar-benar memberikan manfaat bagi pasien itu sendiri, dan tidak hanya dirancang untuk mengurangi beban perawatan bagi seluruh masyarakat — terutama dalam hal perawatan dan AI yang menyertainya. Robot dapat memberdayakan penyandang disabilitas dan lanjut usia serta meningkatkan kemandirian mereka; Faktanya, jika ada pilihan, beberapa orang mungkin lebih memilih robot daripada bantuan manusia untuk tugas-tugas intim tertentu seperti toilet atau mandi. Robot dapat digunakan untuk membantu para lansia agar bisa tinggal di rumah mereka lebih lama, sehingga memberi mereka kebebasan dan otonomi yang lebih besar. Namun, seberapa besar kendali atau otonomi yang harus diberikan kepada seseorang jika kemampuan mentalnya dipertanyakan? Jika seorang pasien meminta robot untuk melemparkannya dari balkon, haruskah robot tersebut melaksanakan perintah tersebut?

Kebebasan dan privasi

Seperti banyak bidang teknologi AI, privasi dan martabat pengguna perlu dipertimbangkan secara cermat saat merancang layanan kesehatan dan robot pendamping. Bekerja di rumah manusia berarti robot akan mengetahui momen-momen pribadi seperti mandi dan berpakaian; jika momen-momen tersebut direkam, siapa yang berhak mengakses informasi tersebut, dan berapa lama rekaman tersebut harus disimpan? Masalah ini menjadi lebih rumit jika kondisi mental orang lanjut usia memburuk dan mereka menjadi bingung - seseorang dengan Alzheimer bisa lupa bahwa ada robot yang mengawasi mereka, dan bisa melakukan tindakan atau mengatakan sesuatu dengan mengira mereka berada dalam privasi di rumah mereka sendiri. Robot perawatan di rumah harus mampu menyeimbangkan privasi penggunaannya dan kebutuhan keperawatan, misalnya dengan mengetuk pintu dan menunggu undangan sebelum memasuki kamar pasien, kecuali dalam keadaan darurat medis.

Untuk memastikan keamanannya, robot terkadang perlu bertindak sebagai pengawas, sehingga membatasi kebebasan mereka. Misalnya, robot dapat dilatih untuk melakukan intervensi jika kompor dibiarkan menyala, atau bak mandi meluap. Robot bahkan mungkin perlu menahan orang lanjut usia untuk melakukan tindakan yang berpotensi membahayakan, seperti memanjat kursi untuk mengambil sesuatu dari lemari. Rumah pintar dengan sensor dapat digunakan untuk mendeteksi seseorang yang berusaha meninggalkan kamarnya, mengunci pintu, atau memanggil staf – namun dengan melakukan hal tersebut, orang lanjut usia akan dipenjara.

Hak pilihan moral

'Ada penelitian yang sangat menarik di mana otak dapat digunakan untuk mengendalikan berbagai hal, seperti mungkin mereka kehilangan fungsi lengannya...di mana menurut saya kekhawatiran sebenarnya terletak pada hal-hal seperti penargetan perilaku: langsung ke hipokampus dan orang-orang menekan 'persetujuan', seperti yang kita lakukan sekarang, untuk akses data'. (John Havens)

Robot tidak memiliki kapasitas untuk melakukan refleksi etis atau dasar moral dalam pengambilan keputusan, dan oleh karena itu manusia saat ini harus memegang kendali penuh atas pengambilan keputusan apa pun. Contoh penalaran etis pada robot dapat ditemukan

dalam film distopia tahun 2004 'I, Robot', di mana karakter Will Smith tidak setuju dengan bagaimana robot di masa fiksi menggunakan logika dingin untuk menyelamatkan nyawanya dibandingkan nyawa anak-anak. Jika layanan kesehatan yang lebih otomatis diupayakan, maka pertanyaan tentang hak pilihan moral akan memerlukan perhatian lebih. Penalaran etis sedang dibangun pada robot, namun tanggung jawab moral lebih dari sekadar penerapan etika – dan masih belum jelas apakah robot di masa depan akan mampu menangani permasalahan moral yang kompleks dalam layanan kesehatan (Goldhill, 2016).

Memercayai

Larosa dan Danks (2018) menulis bahwa AI dapat memengaruhi interaksi dan hubungan antarmanusia dalam bidang layanan kesehatan, khususnya antara pasien dan dokter, dan berpotensi mengganggu kepercayaan yang kita berikan kepada dokter.

'Penelitian psikologi menunjukkan orang-orang tidak memercayai orang-orang yang membuat keputusan moral dengan menghitung biaya dan manfaat — seperti yang dilakukan komputer' (The Guardian, 2017). Ketidakpercayaan kita terhadap robot mungkin juga berasal dari banyaknya robot yang mengamuk dalam fiksi ilmiah distopia. Berita tentang kesalahan komputer — misalnya, tentang algoritme pengidentifikasi gambar yang salah mengira kura-kura sebagai senjata (The Verge, 2017) — serta kekhawatiran akan hal-hal yang tidak diketahui, privasi, dan keamanan merupakan alasan penolakan terhadap penggunaan AI (Global News Kanada, 2016).

Pertama, dokter secara eksplisit tersertifikasi dan memiliki lisensi untuk melakukan praktik kedokteran, dan lisensi mereka menunjukkan bahwa mereka memiliki keterampilan, pengetahuan, dan nilai-nilai khusus seperti 'tidak membahayakan'. Jika robot menggantikan dokter untuk tugas pengobatan atau diagnostik tertentu, hal ini berpotensi mengancam kepercayaan pasien-dokter, karena pasien kini perlu mengetahui apakah sistem tersebut disetujui atau 'dilisensikan' dengan tepat untuk fungsi yang dijalankannya.

Kedua, pasien memercayai dokter karena mereka memandang dokter sebagai teladan keahliannya. Jika dokter hanya dilihat sebagai 'pengguna' AI saja, kita bisa memperkirakan bahwa peran mereka akan diremehkan di mata masyarakat, sehingga mengurangi kepercayaan masyarakat.

Ketiga, pengalaman pasien dengan dokternya merupakan pendorong kepercayaan yang signifikan. Jika pasien memiliki jalur komunikasi yang terbuka dengan dokternya, dan terlibat dalam percakapan tentang perawatan dan pengobatan, maka pasien akan memercayai dokter tersebut. Sebaliknya jika dokter berulang kali mengabaikan keinginan pasien, maka tindakan tersebut akan berdampak buruk pada kepercayaan. Memasukkan AI ke dalam dinamika ini dapat meningkatkan kepercayaan — misalnya jika AI mengurangi kemungkinan kesalahan diagnosis, atau meningkatkan perawatan pasien. Namun, AI juga dapat menurunkan kepercayaan jika dokter mendelegasikan terlalu banyak kewenangan diagnostik atau pengambilan keputusan kepada AI, sehingga melemahkan posisi dokter sebagai otoritas dalam masalah medis.

Seiring dengan semakin banyaknya bukti yang mendukung manfaat terapeutik dari setiap pendekatan teknologi, dan semakin banyaknya sistem interaksi robot yang memasuki pasar, maka kepercayaan terhadap robot kemungkinan akan meningkat. Hal ini telah terjadi pada sistem perawatan kesehatan robotik seperti asisten robot bedah da Vinci (The Guardian, 2014).

Penggantian pekerjaan

Seperti di industri lain, terdapat kekhawatiran bahwa teknologi baru dapat mengancam lapangan kerja (The Guardian, 2017), misalnya, kini terdapat carebot yang dapat melakukan hingga sepertiga pekerjaan perawat (Tech Times, 2018). Meskipun ada kekhawatiran ini, Topol Review NHS (2009) menyimpulkan bahwa 'teknologi ini tidak akan menggantikan tenaga kesehatan profesional namun akan meningkatkan teknologi tersebut ('meningkatkan'), memberi mereka lebih banyak waktu untuk merawat pasien'. Tinjauan tersebut juga menguraikan bagaimana NHS Inggris akan memelihara lingkungan belajar untuk memastikan karyawan memiliki kemampuan digital.

Studi kasus: Kendaraan Otonom

Kendaraan Otonom (AV) adalah kendaraan yang mampu merasakan lingkungannya dan beroperasi dengan sedikit atau tanpa masukan dari pengemudi manusia. Meskipun ide mobil self-driving telah ada setidaknya sejak tahun 1920-an, baru dalam beberapa tahun terakhir teknologi telah berkembang hingga kendaraan AV mulai bermunculan di jalan umum.

Menurut badan standardisasi otomotif SAE Internasional (2018), ada enam tingkat otomatisasi berkendara:

0	Tidak ada otomatisasi	Sistem otomatis mungkin mengeluarkan peringatan dan/atau melakukan intervensi sementara dalam berkendara, namun tidak memiliki kendali kendaraan yang berkelanjutan.
1	Tangan di atas	Pengemudi dan sistem otomatis berbagi kendali kendaraan. Misalnya, sistem otomatis dapat mengontrol tenaga mesin untuk mempertahankan kecepatan yang ditentukan (misalnya Cruise Control), tenaga mesin dan rem untuk mempertahankan dan memvariasikan kecepatan (misalnya Adaptive Cruise Control), atau kemudi (misalnya Bantuan Parkir). Pengemudi harus siap mengambil kembali kendali penuh kapan saja.
2	Lepas tangan	Sistem otomatis mengambil kendali penuh atas kendaraan (termasuk akselerasi, pengereman, dan kemudi). Namun, pengemudi harus memantau jalannya dan bersiap untuk segera melakukan intervensi kapan saja.
3	Tutup mata	Pengemudi dapat dengan aman mengalihkan perhatiannya dari tugas mengemudi (misalnya mengirim pesan teks atau menonton film) karena kendaraan akan menangani situasi apa pun yang memerlukan respons segera. Namun, pengemudi tetap harus bersiap untuk melakukan intervensi, jika diminta oleh AV untuk melakukannya, dalam jangka waktu yang ditentukan oleh produsen AV.

4	Jangan pedulikan	Seperti level 3, namun perhatian pengemudi tidak diperlukan untuk keselamatan, artinya pengemudi dapat tidur atau meninggalkan kursi pengemudi dengan aman.
5	Roda kemudi opsional	Tidak diperlukan campur tangan manusia sama sekali. Contoh AV level 5 adalah robot taksi.

Beberapa otomatisasi tingkat rendah sudah mapan dan ada di pasaran, sementara AV tingkat tinggi sedang menjalani pengembangan dan pengujian. Namun, ketika kita melakukan transisi ke tingkat yang lebih tinggi dan memberikan lebih banyak tanggung jawab pada sistem otomatis dibandingkan pengemudi manusia, sejumlah masalah etika pun muncul.

Dampak Sosial dan Etis dari AV

'Kita tidak bisa membuat alat-alat ini dengan mengatakan, 'kita tahu bahwa manusia bertindak dengan cara tertentu, kita akan membunuh mereka – inilah yang harus dilakukan'.' (John Havens)

Keamanan publik dan etika pengujian di jalan umum

Saat ini, mobil dengan fungsi 'mengemudi dengan bantuan' legal di sebagian besar negara. Khususnya, beberapa model Tesla memiliki fungsi Autopilot, yang menyediakan otomatisasi level 2 (Tesla, nd). Pengemudi secara hukum diperbolehkan menggunakan fungsi mengemudi berbantuan di jalan umum asalkan mereka tetap mengendalikan kendaraan setiap saat. Namun, banyak dari fungsi mengemudi berbantuan ini belum mendapatkan sertifikasi keselamatan independen, sehingga dapat menimbulkan risiko bagi pengemudi dan pengguna jalan lainnya. Di Jerman, sebuah laporan yang diterbitkan oleh Komisi Etika Mengemudi Otomatis menyoroti bahwa sektor publik bertanggung jawab untuk menjamin keselamatan sistem AV yang diperkenalkan dan diberi lisensi di jalan umum, dan merekomendasikan agar semua sistem mengemudi AV harus tunduk pada lisensi dan pemantauan resmi (Komisi Etik, 2017).

Selain itu, ada dugaan bahwa industri AV sedang memasuki fase paling berbahaya, dengan mobil belum sepenuhnya otonom namun operator manusia belum sepenuhnya terlibat (Solon, 2018). Risiko yang ditimbulkan oleh hal ini telah menjadi perhatian luas setelah kematian pejalan kaki pertama yang melibatkan mobil otonom. Tragedi ini terjadi di Arizona, AS, pada Mei 2018, ketika AV level 3 yang sedang diuji oleh Uber bertabrakan dengan Elaine Herzberg yang berusia 49 tahun saat dia sedang mengendarai sepedanya di seberang jalan pada suatu malam. Jaksa menetapkan bahwa Uber 'tidak bertanggung jawab secara pidana' (Shepherdson dan Somerville, 2019), dan laporan awal Dewan Keselamatan Transportasi Nasional AS (NTSB, 2018), yang tidak menarik kesimpulan mengenai penyebabnya, menyatakan bahwa semua elemen dari Uber 'tidak bertanggung jawab secara pidana'. sistem self-driving beroperasi normal pada saat kecelakaan terjadi. Uber mengatakan bahwa pengemudi diandalkan untuk melakukan intervensi dan mengambil tindakan dalam situasi yang memerlukan pengereman darurat – sehingga beberapa komentator menyerukan komunikasi yang menyesatkan kepada konsumen seputar istilah 'mobil self-driving' dan 'autopilot' (Leggett, 2018). Kecelakaan tersebut juga menyebabkan beberapa pihak mengutuk praktik pengujian sistem AV di jalan umum sebagai tindakan yang berbahaya dan tidak etis,

dan menyebabkan Uber menghentikan sementara program mengemudi mandiri (Bradshaw, 2018).

Masalah keselamatan manusia – baik masyarakat maupun penumpang – muncul sebagai isu utama terkait mobil self-driving. Perusahaan-perusahaan besar – Nissan, Toyota, Tesla, Uber, Volkswagen – sedang mengembangkan kendaraan otonom yang mampu beroperasi di lingkungan yang kompleks dan tidak dapat diprediksi tanpa kendali langsung manusia, dan mampu mempelajari, menyimpulkan, merencanakan, dan mengambil keputusan.

Kendaraan tanpa pengemudi dapat memberikan banyak manfaat: statistik menunjukkan bahwa Anda hampir pasti lebih aman berada di dalam mobil yang dikemudikan oleh komputer dibandingkan mobil yang dikemudikan oleh manusia. Hal ini juga dapat mengurangi kemacetan di perkotaan, mengurangi polusi, mengurangi waktu perjalanan dan perjalanan, serta memungkinkan masyarakat menggunakan waktu mereka dengan lebih produktif. Namun hal ini bukan berarti akhir dari kecelakaan lalu lintas. Meskipun mobil tanpa pengemudi memiliki perangkat lunak dan perangkat keras terbaik, risiko tabrakan tetap ada. Sebuah mobil otonom dapat dikejutkan, misalnya oleh seorang anak yang muncul dari belakang kendaraan yang diparkir, dan selalu ada pertanyaan tentang bagaimana caranya: bagaimana seharusnya mobil tersebut diprogram ketika mereka harus memutuskan keselamatan mana yang harus diprioritaskan?

Mobil tanpa pengemudi mungkin juga harus memilih antara keselamatan penumpang dan pengguna jalan lainnya. Katakanlah sebuah mobil melaju di sudut tempat sekelompok anak sekolah sedang bermain; tidak ada cukup waktu untuk berhenti, dan satu-satunya cara agar mobil tidak menabrak anak-anak adalah dengan membelok ke dinding bata — membahayakan penumpang. Keselamatan siapa yang harus diprioritaskan oleh mobil: keselamatan anak-anak atau penumpang?

Proses dan teknologi untuk investigasi kecelakaan

AV adalah sistem kompleks yang sering kali mengandalkan teknologi pembelajaran mesin canggih. Beberapa kecelakaan serius telah terjadi, termasuk sejumlah kematian yang melibatkan AV level 2:

- Pada bulan Januari 2016, Gao Yanning yang berusia 23 tahun meninggal ketika Tesla Model S miliknya menabrak bagian belakang truk penyapu jalan di jalan raya di Hebei, Tiongkok. Keluarga tersebut percaya Autopilot terlibat ketika kecelakaan itu terjadi dan menuduh Tesla melebih-lebihkan kemampuan sistem. Tesla menyatakan bahwa kerusakan pada kendaraan membuat tidak mungkin untuk menentukan apakah Autopilot diaktifkan dan, jika demikian, apakah tidak berfungsi. Kasus perdata atas kecelakaan tersebut sedang berlangsung, dengan penilai pihak ketiga meninjau data dari kendaraan tersebut (Curtis, 2016).
- Pada Mei 2016, Joshua Brown yang berusia 40 tahun meninggal ketika Tesla Model S miliknya bertabrakan dengan truk saat Autopilot sedang digunakan di Florida, AS. Investigasi yang dilakukan oleh Badan Keselamatan Jalan Raya dan Transportasi Nasional menemukan bahwa pengemudinya, dan bukan Tesla, yang bersalah (Gibbs, 2016). Namun,

Administrasi Keselamatan Lalu Lintas Jalan Raya Nasional kemudian menetapkan bahwa autopilot dan ketergantungan berlebihan pengemudi pada alat bantu mengemudi Tesla adalah penyebabnya (Felton, 2017).

- Pada bulan Maret 2018, Wei Huang terbunuh ketika Tesla Model X miliknya menabrak penghalang keselamatan jalan raya di California, AS. Menurut Tesla, tingkat keparahan kecelakaan itu 'belum pernah terjadi sebelumnya'. Dewan Keselamatan Transportasi Nasional kemudian menerbitkan laporan yang menghubungkan kecelakaan itu dengan kesalahan navigasi Autopilot. Tesla kini digugat oleh keluarga korban (O'Kane, 2018).

Sayangnya, upaya untuk menyelidiki kecelakaan ini terhambat oleh kenyataan bahwa standar, proses, dan kerangka peraturan untuk menyelidiki kecelakaan yang melibatkan kendaraan AV belum dikembangkan atau diadopsi. Selain itu, sistem pencatatan data eksklusif yang saat ini dipasang di AV membuat penyelidikan kecelakaan sangat bergantung pada kerja sama produsen untuk menyediakan data penting mengenai kejadian yang mengarah pada kecelakaan (Stilgoe dan Winfield, 2018).

Salah satu solusinya adalah menyesuaikan semua AV masa depan dengan perekam data peristiwa standar industri – yang disebut 'kotak hitam etis' – yang dapat diakses oleh penyelidik kecelakaan independen. Hal ini mencerminkan model yang sudah ada dalam investigasi kecelakaan udara.

Kecelakaan nyaris celaka

Saat ini, belum ada sistem yang mampu mencatat kecelakaan nyaris celaka secara sistematis. Meskipun ada kemungkinan bahwa produsen sudah mengumpulkan data ini, mereka tidak berkewajiban untuk melakukannya – atau untuk membagikan data tersebut. Satu-satunya pengecualian saat ini adalah negara bagian California, AS, yang mewajibkan semua perusahaan yang secara aktif menguji AV di jalan umum untuk mengungkapkan frekuensi di mana pengemudi manusia dipaksa untuk mengambil kendali kendaraan demi alasan keselamatan (dikenal sebagai 'disengagement').

Pada tahun 2018, jumlah pelepasan yang dilakukan oleh produsen AV bervariasi secara signifikan, dari satu pelepasan untuk setiap 11.017 mil yang dilakukan oleh Waymo AV hingga satu pelepasan untuk setiap 1,15 mil yang digerakkan oleh Apple AV (Hawkins, 2019). Data mengenai pelepasan diri ini memperkuat pentingnya memastikan bahwa faktor-faktor yang mendorong keselamatan manusia tetap terlibat. Namun, proses pengumpulan data di California telah dikritik, dengan beberapa orang mengklaim bahwa kata-kata yang ambigu dan kurangnya pedoman yang ketat memungkinkan perusahaan untuk menghindari pelaporan kejadian tertentu yang dapat disebut nyaris celaka.

Tanpa akses terhadap jenis data ini, pembuat kebijakan tidak dapat memperhitungkan frekuensi dan signifikansi kecelakaan nyaris celaka, atau menilai langkah-langkah yang diambil oleh produsen sebagai akibat dari kecelakaan nyaris celaka tersebut. Sekali lagi, pembelajaran dapat diambil dari model yang diikuti dalam investigasi kecelakaan udara, di mana semua kejadian nyaris celaka dicatat secara menyeluruh dan diselidiki secara independen. Para pembuat kebijakan memerlukan statistik komprehensif mengenai semua kecelakaan dan kejadian nyaris celaka agar dapat menjadi masukan bagi peraturan.

Privasi data

Menjadi jelas bahwa produsen mengumpulkan sejumlah besar data dari AV. Ketika kendaraan-kendaraan ini menjadi semakin umum di jalan raya, pertanyaan yang muncul adalah: sejauh mana data-data ini membahayakan privasi dan hak perlindungan data pengemudi dan penumpang?

Masalah pengelolaan data dan privasi telah muncul, dan beberapa diantaranya menimbulkan kekhawatiran tentang potensi penyalahgunaan data AV untuk tujuan periklanan (Lin, 2014). Tesla juga mendapat kecaman karena penggunaan log data AV yang tidak etis. Dalam investigasi yang dilakukan oleh The Guardian, surat kabar tersebut menemukan beberapa contoh di mana perusahaan membagikan data pribadi pengemudi kepada media setelah terjadi kecelakaan, tanpa izin mereka, untuk membuktikan bahwa teknologinya tidak bertanggung jawab (Thielman, 2017). Pada saat yang sama, Tesla tidak mengizinkan pelanggan melihat log data mereka sendiri.

Salah satu solusi yang diusulkan oleh Komisi Etika Jerman untuk Mengemudi Otomatis adalah memastikan bahwa semua pengemudi AV diberikan kedaulatan data penuh (Komisi Etika, 2017). Hal ini akan memungkinkan mereka mengontrol bagaimana data mereka digunakan.

Pekerjaan

Pertumbuhan kendaraan AV kemungkinan akan menempatkan pekerjaan tertentu – terutama supir bus, taksi, dan truk – dalam risiko. Dalam jangka menengah, pengemudi truk menghadapi risiko terbesar karena truk jarak jauh berada di garis depan teknologi AV (Viscelli, 2018). Pada tahun 2016, pengiriman bir komersial pertama dilakukan menggunakan truk yang dapat mengemudi sendiri, dalam perjalanan sejauh 120 mil dan tidak melibatkan manusia (Isaac, 2016). Tahun lalu terjadi perjalanan tanpa pengemudi pertama dengan truk yang dapat mengemudi sendiri, dengan AV menempuh perjalanan sejauh tujuh mil tanpa satu orang pun di dalamnya (Cannon, 2018).

Ke depan, pengemudi bus juga kemungkinan besar akan kehilangan pekerjaan karena semakin banyak bus yang tidak memiliki pengemudi. Sejumlah kota di dunia telah mengumumkan rencana untuk memperkenalkan layanan antar-jemput tanpa pengemudi di masa depan, termasuk Edinburgh (Calder, 2018), New York (BBC, 2019a) dan Singapura (BBC 2017). Di beberapa tempat, visi ini telah menjadi kenyataan; pesawat ulang-alik Las Vegas terkenal mengalami awal yang sulit ketika mengalami tabrakan pada hari pertama pengoperasiannya (Park, 2017), dan wisatawan di kota kecil Neuhausen Rheinfall di Swiss kini dapat naik bus tanpa pengemudi ke kunjungi air terjun terdekat (CNN, 2018). Dalam jangka menengah, bus tanpa pengemudi kemungkinan besar akan dibatasi pada rute yang 100% menggunakan jalur bus khusus. Meskipun demikian, kemajuan layanan angkutan mandiri telah menciptakan ketegangan dengan organisasi buruh dan pejabat kota di AS (Weinberg, 2019). Tahun lalu, Serikat Pekerja Transportasi Amerika membentuk koalisi dalam upaya menghentikan bus otonom memasuki jalan-jalan di Ohio (Pfleger, 2018).

Taksi yang sepenuhnya otonom kemungkinan besar hanya akan terwujud dalam jangka panjang, setelah teknologi AV telah teruji sepenuhnya dan terbukti pada level 4 dan 5.

Meskipun demikian, dengan adanya rencana untuk memperkenalkan taksi tanpa pengemudi di London pada tahun 2021 (BBC, 2018), dan layanan taksi otomatis sudah tersedia di Arizona, AS (Sage, 2019), mudah untuk melihat mengapa pengemudi taksi merasa tidak nyaman.

Kualitas lingkungan perkotaan

Dalam jangka panjang, AV mempunyai potensi untuk mengubah lingkungan perkotaan kita. Beberapa dari perubahan ini mungkin mempunyai dampak negatif bagi pejalan kaki, pengendara sepeda, dan penduduk setempat. Ketika mengemudi menjadi lebih otomatis, kemungkinan akan ada kebutuhan akan infrastruktur tambahan (misalnya jalur khusus AV). Mungkin juga ada dampak yang lebih luas terhadap perencanaan kota, dengan otomatisasi yang mengatur perencanaan segala sesuatu mulai dari kemacetan lalu lintas dan parkir hingga ruang hijau dan lobi (Marshall dan Davies, 2018). Peluncuran AV juga memerlukan perluasan jangkauan jaringan 5G secara signifikan — sekali lagi, hal ini berdampak pada perencanaan kota (Khosravi, 2018).

Dampak lingkungan dari mobil self-driving juga harus dipertimbangkan. Meskipun mobil self-driving berpotensi mengurangi penggunaan bahan bakar dan emisi terkait secara signifikan, penghematan ini dapat diatasi dengan fakta bahwa mobil self-driving membuat berkendara jarak jauh menjadi lebih mudah dan menarik (Worland, 2016). Oleh karena itu, dampak otomatisasi terhadap perilaku mengemudi tidak boleh dianggap remeh.

Tanggung jawab hukum dan etika

Dari sudut pandang hukum, siapa yang bertanggung jawab atas kecelakaan yang disebabkan oleh robot, dan bagaimana seharusnya korban mendapat kompensasi (jika memang ada) ketika kendaraan yang dikendalikan oleh suatu algoritma menyebabkan cedera? Jika pengadilan tidak dapat menyelesaikan masalah ini, produsen robot mungkin akan mengeluarkan biaya tak terduga yang akan menghambat investasi. Namun, jika korban tidak mendapatkan kompensasi yang layak maka kendaraan otonom tidak akan dipercaya atau diterima oleh masyarakat.

Robot perlu membuat penilaian panggilan dalam kondisi ketidakpastian, atau situasi 'tidak menang'. Namun, pendekatan atau teori etis manakah yang harus diikuti oleh robot jika tidak ada panduan hukum? Seperti Lin dkk. jelaskan, pendekatan yang berbeda dapat menghasilkan hasil yang berbeda, termasuk jumlah korban jiwa akibat kecelakaan.

Selain itu, siapa yang harus memilih etika kendaraan otonom — pengemudi, konsumen, penumpang, produsen, politisi? Loh dan Loh (2017) berpendapat bahwa tanggung jawab harus dibagi antara para insinyur, pengemudi, dan sistem mengemudi otonom itu sendiri.

Namun, Millar (2016) menyarankan bahwa pengguna teknologi, dalam hal ini penumpang mobil yang dapat mengemudi sendiri, harus dapat memutuskan prinsip etika atau perilaku apa yang harus diikuti oleh robot. Dengan menggunakan contoh dokter, yang tidak memiliki otoritas moral untuk membuat keputusan penting mengenai perawatan di akhir hayat tanpa persetujuan pasiennya, ia berpendapat bahwa akan ada protes moral jika para insinyur merancang mobil tanpa meminta pengemudinya. langsung untuk masukan mereka, atau memberi tahu pengguna sebelumnya bagaimana mobil diprogram untuk berperilaku dalam situasi tertentu.

Studi kasus: Peperangan dan persenjataan

Meskipun sebagian sistem otonom dan cerdas telah digunakan dalam teknologi militer setidaknya sejak Perang Dunia Kedua, kemajuan dalam pembelajaran mesin dan AI menandakan titik balik dalam penggunaan otomatisasi dalam peperangan.

AI sudah cukup maju dan canggih untuk digunakan di berbagai bidang seperti analisis citra satelit dan pertahanan dunia maya, namun cakupan penerapannya yang sebenarnya masih belum sepenuhnya terwujud. Sebuah laporan baru-baru ini menyimpulkan bahwa teknologi AI memiliki potensi untuk mengubah peperangan ke tingkat yang sama, atau bahkan lebih besar, dibandingkan munculnya senjata nuklir, pesawat terbang, komputer, dan bioteknologi (Allen dan Chan, 2017). Beberapa dampak utama AI terhadap militer diuraikan di bawah ini.

Dilema etika dalam pembangunan

Pada tahun 2014, inisiatif Open Roboethics (ORi 2014a, 2014b) melakukan jajak pendapat yang menanyakan pendapat masyarakat tentang apa yang harus dilakukan oleh mobil otonom yang mereka penumpangi jika ada anak yang melangkah di depan kendaraan dalam terowongan. Mobil tidak punya waktu untuk mengerem dan menyelamatkan anak tersebut, namun bisa berbelok ke dinding terowongan, menewaskan penumpangnya. Ini merupakan putaran dari 'dilema troli' klasik, di mana seseorang mempunyai pilihan untuk mengalihkan troli yang melarikan diri dari jalur yang dapat melukai beberapa orang ke jalur yang hanya akan melukai satu orang.

36% peserta mengatakan bahwa mereka lebih memilih mobilnya membelok ke tembok, sehingga menyelamatkan anak tersebut; Namun, mayoritas (64%) mengatakan mereka ingin menyelamatkan diri mereka sendiri, sehingga mengorbankan anak tersebut. 44% peserta berpendapat bahwa penumpang harus dapat memilih tindakan yang akan diambil dalam mobilnya, sementara 33% mengatakan bahwa anggota parlemen harus memilih. Hanya 12% yang mengatakan bahwa produsen mobil harus mengambil keputusan. Hasil ini menunjukkan bahwa masyarakat tidak menyukai gagasan bahwa para insinyur mengambil keputusan moral atas nama mereka.

Meminta masukan penumpang dalam setiap situasi adalah hal yang tidak praktis. Namun, Millar (2016) menyarankan prosedur 'penyiapan' di mana orang dapat memilih pengaturan etika mereka setelah membeli mobil baru. Meskipun demikian, memilih cara mobil bereaksi terlebih dahulu dapat dianggap sebagai tindakan yang merugikan, jika, misalnya, pengguna memprogram kendaraannya untuk selalu menghindari tabrakan dengan cara membelok ke arah pengendara sepeda. Hal ini akan meningkatkan akuntabilitas dan tanggung jawab pengguna, sekaligus mengalihkan tanggung jawab dari produsen.

Senjata otonom yang mematikan

Ketika sistem otomatis dan otonom menjadi lebih mampu, militer menjadi lebih bersedia untuk mendelegasikan wewenang kepada mereka. Hal ini kemungkinan akan berlanjut dengan meluasnya adopsi AI, yang mengarah pada perlombaan senjata yang terinspirasi oleh AI. Komite Industri Militer Rusia telah menyetujui rencana agresif dimana 30% kekuatan tempur Rusia akan sepenuhnya terdiri dari platform robotik yang dikendalikan

dari jarak jauh dan otonom pada tahun 2030. Negara-negara lain kemungkinan besar akan menetapkan tujuan serupa. Meskipun Departemen Pertahanan Amerika Serikat telah memberlakukan pembatasan penggunaan sistem otonom dan semi-otonom yang menggunakan kekuatan mematikan, negara-negara lain dan aktor non-negara mungkin tidak melakukan pengendalian diri seperti itu.

Teknologi drone

Harga pesawat militer standar bisa lebih dari Rp. 10 Triliyun per unit; Namun, Kendaraan Udara Tak Berawak quadcopter berkualitas tinggi saat ini berharga sekitar Rp. 10.000.000, yang berarti bahwa dengan harga satu pesawat kelas atas, militer dapat memperoleh satu juta drone. Meskipun drone komersial saat ini memiliki jangkauan yang terbatas, di masa depan mereka mungkin memiliki jangkauan yang sama dengan rudal balistik, sehingga membuat platform yang ada menjadi usang.

Pembunuhan robot

Ketersediaan robot yang berbiaya rendah, berkemampuan tinggi, mematikan, dan otonom dapat membuat pembunuhan yang ditargetkan menjadi lebih luas dan lebih sulit untuk diidentifikasi. Robot penembak otomatis dapat membunuh sasaran dari jauh.

Alat Peledak yang Ditingkatkan dengan Robot Seluler

Ketika teknologi robot komersial dan kendaraan otonom semakin meluas, beberapa kelompok akan memanfaatkan teknologi ini untuk membuat Alat Peledak Improvisasi (IED) yang lebih canggih. Saat ini, kemampuan teknologi untuk mengirimkan bahan peledak dengan cepat ke sasaran yang tepat dari jarak beberapa mil jauhnya terbatas pada negara-negara yang kuat. Namun, jika pengiriman paket jarak jauh dengan drone menjadi kenyataan, biaya pengiriman bahan peledak dari jarak jauh akan turun dari jutaan dolar menjadi ribuan atau bahkan ratusan. Demikian pula, mobil tanpa pengemudi dapat membuat bom mobil bunuh diri lebih sering terjadi dan menimbulkan dampak buruk karena tidak lagi memerlukan pengemudi yang ingin bunuh diri.

Hallaq dkk. (2017) juga menyoroti bidang-bidang utama di mana pembelajaran mesin kemungkinan besar akan memengaruhi peperangan. Mereka menggambarkan contoh di mana seorang Komandan (CO) dapat menggunakan Intelligent Virtual Assistant (IVA) dalam lingkungan medan perang yang dinamis yang secara otomatis memindai citra satelit untuk mendeteksi jenis kendaraan tertentu, sehingga membantu mengidentifikasi ancaman terlebih dahulu. Ia juga dapat memprediksi niat musuh, dan membandingkan data situasional dengan database tersimpan dari ratusan latihan wargame sebelumnya dan pertempuran langsung, memberikan CO akses ke tingkat akumulasi pengetahuan yang tidak mungkin diperoleh jika tidak dilakukan.

Penggunaan AI dalam peperangan menimbulkan beberapa pertanyaan hukum dan etika. Salah satu kekhawatirannya adalah sistem senjata otomatis yang mengecualikan penilaian manusia dapat melanggar Hukum Humaniter Internasional, dan mengancam hak dasar kita untuk hidup dan prinsip martabat manusia. AI juga dapat menurunkan ambang batas perang, sehingga mempengaruhi stabilitas global.

Hukum Humaniter Internasional menetapkan bahwa setiap serangan harus membedakan antara kombatan dan non-kombatan, proporsional dan tidak boleh menargetkan warga sipil atau objek sipil. Selain itu, serangan tidak boleh memperburuk penderitaan para kombatan. AI mungkin tidak dapat memenuhi prinsip-prinsip ini tanpa keterlibatan penilaian manusia. Secara khusus, banyak peneliti khawatir bahwa Lethal Autonomous Weapon Systems (LAWS) – sejenis robot militer otonom yang dapat secara mandiri mencari dan 'melibatkan' target menggunakan kekuatan mematikan – mungkin tidak memenuhi standar yang ditetapkan oleh Hukum Humaniter Internasional, sebagaimana adanya. tidak dapat membedakan warga sipil dari kombatan, dan tidak dapat menilai apakah kekuatan serangan tersebut proporsional mengingat kerugian yang akan ditimbulkan oleh warga sipil.

Amoroso dan Tamburrini (2016, p. 6) berpendapat bahwa: '[HUKUM harus] mampu menghormati prinsip-prinsip pembedaan dan proporsionalitas setidaknya serta menjadi prajurit manusia yang kompeten dan teliti'. Namun, Lim (2019) menyatakan bahwa meskipun LAWS yang gagal memenuhi persyaratan ini tidak boleh diterapkan, suatu hari nanti LAWS akan cukup canggih untuk memenuhi persyaratan pembedaan dan proporsionalitas. Sementara itu, Asaro (2012) berpendapat bahwa tidak peduli seberapa bagus HUKUM yang didapat; sudah merupakan persyaratan moral bahwa hanya manusia yang boleh melakukan kekuatan mematikan, dan secara moral salah jika mendelegasikan keputusan hidup atau mati kepada mesin.

Beberapa orang berpendapat bahwa mendelegasikan keputusan untuk membunuh manusia ke mesin adalah pelanggaran terhadap martabat dasar manusia, karena robot tidak merasakan emosi, dan tidak memiliki gagasan tentang pengorbanan dan apa artinya mengambil nyawa. Seperti yang dijelaskan oleh Lim dkk (2019), 'sebuah mesin, tanpa darah dan tanpa moralitas atau kematian, tidak dapat memahami pentingnya penggunaan kekerasan terhadap manusia dan tidak dapat memberikan keadilan terhadap beratnya suatu keputusan'.

Robot juga tidak punya konsep apa artinya membunuh orang yang 'salah'. 'Hanya karena manusia dapat merasakan kemarahan dan penderitaan yang menyertai pembunuhan manusia maka mereka dapat memahami pengorbanan dan penggunaan kekerasan terhadap manusia. Hanya dengan cara itulah mereka dapat menyadari 'beratnya keputusan' untuk membunuh'.

Namun, ada pula yang berpendapat bahwa tidak ada alasan khusus mengapa terbunuh oleh mesin akan menjadi pengalaman yang lebih buruk, atau kurang bermartabat, dibandingkan terbunuh oleh serangan rudal jelajah. 'Yang penting adalah apakah korban mengalami rasa terhina dalam proses pembunuhan. Korban yang diancam dengan potensi bom tidak akan peduli apakah bom tersebut dijatuhkan oleh manusia atau robot' (Lim dkk, 2019). Selain itu, tidak semua manusia memiliki kapasitas emosional untuk mengkonseptualisasikan pengorbanan atau emosi relevan yang menyertai risiko. Di tengah panasnya pertempuran, tentara jarang mempunyai waktu untuk memikirkan konsep

pengorbanan, atau membangkitkan emosi yang relevan untuk membuat keputusan setiap kali mereka mengerahkan kekuatan mematikan.

Selain itu, siapa yang harus bertanggung jawab atas tindakan sistem otonom — komandan, pemrogram, atau operator sistem? Schmit (2013) berpendapat bahwa tanggung jawab untuk melakukan kejahatan perang harus berada pada individu yang memprogram AI, dan komandan atau pengawas (dengan asumsi bahwa mereka mengetahui, atau seharusnya mengetahui, sistem senjata otonom telah diprogram dan digunakan dalam suatu negara). kejahatan perang, dan mereka tidak melakukan apa pun untuk menghentikannya).

BAB 4

STANDAR DAN REGULASI ARTIFICIAL INTELLIGENCE

Generasi kecil standar etika baru bermunculan seiring dengan semakin dipahaminya dampak etika, hukum, dan sosial dari kecerdasan buatan dan robotika. Baik suatu standar dengan jelas mengartikulasikan permasalahan etika yang tersurat maupun tersirat, semua standar mewujudkan semacam prinsip etika (Winfield, 2019a). Standar-standar yang ada masih dalam tahap pengembangan dan informasi yang tersedia untuk umum mengenai standar-standar tersebut masih terbatas.

Mungkin standar etika eksplisit paling awal dalam robotika adalah Panduan BS 8611 untuk Desain Etis dan Penerapan Robot dan Sistem Robot (British Standard BS 8611, 2016). BS8611 bukanlah kode praktik, namun panduan tentang bagaimana desainer dapat mengidentifikasi potensi bahaya etika, melakukan penilaian risiko etika terhadap robot atau AI mereka, dan memitigasi risiko etika apa pun yang teridentifikasi. Hal ini didasarkan pada 20 bahaya dan risiko etika yang berbeda, yang dikelompokkan dalam empat kategori: sosial, penerapan, komersial & keuangan, dan lingkungan.

Nasihat mengenai langkah-langkah untuk memitigasi dampak setiap risiko diberikan, bersama dengan saran tentang bagaimana langkah-langkah tersebut dapat diverifikasi atau divalidasi. Bahaya sosial mencakup, misalnya, hilangnya kepercayaan, penipuan, pelanggaran privasi dan kerahasiaan, kecanduan, dan kehilangan pekerjaan. Penilaian Risiko Etis juga harus mempertimbangkan penyalahgunaan yang dapat diperkirakan, risiko yang mengarah pada stres dan ketakutan (dan minimalisasinya), kegagalan kontrol (dan efek psikologis terkait), konfigurasi ulang dan perubahan terkait tanggung jawab, bahaya yang terkait dengan aplikasi robotika tertentu. Perhatian khusus diberikan pada robot yang dapat belajar dan implikasi peningkatan robot yang timbul, dan standar tersebut berpendapat bahwa risiko etika yang terkait dengan penggunaan robot tidak boleh melebihi risiko aktivitas yang sama jika dilakukan oleh manusia.

British Standard BS 8611 berasumsi bahwa bahaya fisik berarti bahaya etis, dan mendefinisikan bahaya etis sebagai dampak terhadap 'kesejahteraan psikologis dan/atau masyarakat dan lingkungan.' Hal ini juga mengakui bahwa bahaya fisik dan emosional harus seimbang dengan manfaat yang diharapkan bagi pengguna.

Standar ini menyoroti perlunya melibatkan masyarakat dan pemangku kepentingan dalam pengembangan robot dan memberikan daftar pertimbangan desain utama termasuk:

- Robot tidak boleh dirancang terutama untuk membunuh manusia;
- Manusia tetap menjadi agen yang bertanggung jawab;
- Harus mungkin untuk mengetahui siapa yang bertanggung jawab atas robot apa pun;
- Robot harus aman dan sesuai tujuannya;
- Robot tidak boleh dirancang untuk menipu;
- Prinsip kehati-hatian harus diikuti;

- Privasi harus dimasukkan dalam desain;
- Pengguna tidak boleh didiskriminasi, atau dipaksa menggunakan robot.

Pedoman khusus disediakan bagi para ahli robotika, khususnya mereka yang melakukan penelitian. Hal ini mencakup kebutuhan untuk melibatkan masyarakat, mempertimbangkan kekhawatiran masyarakat, bekerja sama dengan para ahli dari disiplin ilmu lain, memperbaiki informasi yang salah dan memberikan instruksi yang jelas. Metode khusus untuk memastikan penggunaan robot secara etis meliputi: validasi pengguna (untuk memastikan robot dapat/dioperasikan sesuai harapan), verifikasi perangkat lunak (untuk memastikan perangkat lunak berfungsi sesuai antisipasi), keterlibatan pakar lain dalam penilaian etika, penilaian ekonomi dan sosial terhadap hasil yang diharapkan. Penilaian terhadap segala implikasi hukum, pengujian kepatuhan terhadap standar yang relevan. Jika diperlukan, pedoman dan kode etik lain harus dipertimbangkan dalam desain dan pengoperasian robot (misalnya kode medis atau hukum yang relevan dalam konteks tertentu). Standar ini juga menyatakan bahwa penerapan robot untuk keperluan militer tidak menghilangkan tanggung jawab dan akuntabilitas manusia.

Asosiasi Standar IEEE juga telah meluncurkan standar melalui inisiatif globalnya mengenai Etika Sistem Otonom dan Cerdas. Dengan memposisikan 'kesejahteraan manusia' sebagai prinsip utama, inisiatif IEEE secara eksplisit berupaya untuk memposisikan kembali robotika dan AI sebagai teknologi untuk meningkatkan kondisi manusia, bukan sekadar kendaraan untuk pertumbuhan ekonomi (Winfield, 2019a). Tujuannya adalah untuk mendidik, melatih dan memberdayakan pemangku kepentingan AI/robot untuk 'memprioritaskan pertimbangan etis sehingga teknologi ini maju demi kepentingan umat manusia.'

Saat ini terdapat 14 kelompok kerja standar IEEE yang berupaya menyusun apa yang disebut standar 'manusia' yang mempunyai implikasi terhadap kecerdasan buatan (Tabel 4.1).

Tabel 4.1: 'Standar manusia' IEEE dengan implikasinya terhadap AI

	Standar	Maksud/Tujuan
P7000	Proses Model untuk Mengatasi Masalah Etis Selama Desain Sistem	Untuk menetapkan proses desain etis Sistem Otonom dan Cerdas.
P7001	Transparansi Sistem Otonom	Untuk memastikan transparansi sistem otonom kepada berbagai pemangku kepentingan. Ini secara khusus akan membahas: <ul style="list-style-type: none"> • Pengguna: memastikan pengguna memahami apa yang dilakukan sistem dan alasannya, dengan tujuan membangun kepercayaan; • Validasi dan sertifikasi: memastikan sistem tersebut diawasi dengan cermat; • Kecelakaan: memungkinkan penyelidikan kecelakaan untuk melakukan penyelidikan; • Pengacara dan saksi ahli: memastikan bahwa, setelah terjadi kecelakaan, kelompok-kelompok ini mampu memberikan bukti;

		<ul style="list-style-type: none"> • Teknologi yang disruptif (misalnya mobil tanpa pengemudi): memungkinkan masyarakat menilai teknologi (dan, jika perlu, membangun kepercayaan diri).
P7002	Proses Privasi Data	Untuk menetapkan standar penggunaan etis data pribadi dalam proses rekayasa perangkat lunak. Buku ini akan mengembangkan dan menjelaskan penilaian dampak privasi (PIA) yang dapat digunakan untuk mengidentifikasi kebutuhan dan efektivitas tindakan pengendalian privasi. Ini juga akan memberikan daftar periksa bagi mereka yang mengembangkan perangkat lunak yang menggunakan informasi pribadi.
P7003	Pertimbangan Bias Algoritmik	Untuk membantu pengembang algoritme memperjelas cara mereka berupaya menghilangkan atau meminimalkan risiko bias dalam produk mereka. Hal ini akan mengatasi penggunaan informasi yang terlalu subyektif dan membantu pengembang memastikan bahwa mereka mematuhi undang-undang mengenai karakteristik yang dilindungi (misalnya ras, gender). Ini mungkin mencakup: <ul style="list-style-type: none"> • Proses benchmarking untuk pemilihan kumpulan data; • Pedoman untuk mengkomunikasikan batas-batas algoritma yang telah dirancang dan divalidasi (menjaga terhadap konsekuensi yang tidak diinginkan dari penggunaan yang tidak terduga); • Strategi untuk menghindari kesalahan penafsiran keluaran sistem oleh pengguna.
P7004	Standar Tata Kelola Data Anak dan Siswa	Khusus ditujukan pada institusi pendidikan, panduan ini akan memberikan panduan dalam mengakses, mengumpulkan, menyimpan, menggunakan, membagikan, dan memusnahkan data anak/siswa.
P7005	Standar Tata Kelola Data Perusahaan yang Transparan	Mirip dengan P7004, namun ditujukan untuk pengusaha.
P7006	Standar untuk Agen Kecerdasan Buatan (AI) Data Pribadi	Menjelaskan elemen teknis yang diperlukan untuk membuat dan memberikan akses ke AI yang dipersonalisasi. Hal ini akan memungkinkan individu untuk mengatur dan membagikan informasi pribadi mereka dengan aman pada tingkat yang dapat dibaca mesin, dan memungkinkan AI yang dipersonalisasi untuk bertindak sebagai proxy untuk keputusan mesin-ke-mesin.
P7007	Standar Ontologis untuk Sistem Robotika dan Otomasi yang Didorong Secara Etis	Standar ini menyatukan teknik dan filosofi untuk memastikan bahwa kesejahteraan pengguna dipertimbangkan sepanjang siklus hidup produk. Hal ini bertujuan untuk mengidentifikasi cara-cara memaksimalkan manfaat dan meminimalkan dampak negatif, dan juga akan mempertimbangkan cara-cara agar komunikasi dapat dilakukan dengan jelas antara berbagai komunitas.
P7008	Standar untuk Dorongan yang Didorong Secara Etis	Berdasarkan 'teori dorongan', standar ini berupaya untuk menggambarkan dorongan saat ini atau potensi yang mungkin dilakukan oleh robot atau sistem otonom. Hal ini mengakui bahwa

	untuk Sistem Robotik, Cerdas, dan Otonom	dorongan dapat digunakan untuk berbagai alasan, namun dorongan tersebut bertujuan untuk mempengaruhi penerima secara emosional, mengubah perilaku dan dapat bersifat manipulatif, serta berupaya untuk menguraikan metodologi untuk desain etis AI dengan menggunakan dorongan.
P7009	Standar untuk Desain Sistem Otonom dan Semi-Otonom yang Aman dari Kegagalan	Untuk menciptakan metodologi yang efektif untuk pengembangan dan implementasi mekanisme fail-safe yang kuat, transparan dan akuntabel. Ini akan membahas metode untuk mengukur dan menguji kemampuan sistem untuk gagal dengan aman.
P7010	Standar Metrik Kesejahteraan untuk Kecerdasan Buatan yang Etis dan Sistem Otonom	Untuk menetapkan dasar bagi metrik yang digunakan untuk menilai faktor-faktor kesejahteraan yang dapat dipengaruhi oleh sistem otonom, dan bagaimana kesejahteraan manusia dapat ditingkatkan secara proaktif.
P7011	Standar Proses Identifikasi dan Penilaian Keterpercayaan Sumber Berita	Berfokus pada informasi berita, standar ini bertujuan untuk menstandarisasi proses penilaian keakuratan faktual sebuah berita. Ini akan digunakan untuk menghasilkan skor 'kepercayaan'. Standar ini berupaya mengatasi dampak negatif dari berita 'palsu' yang tidak terkendali, dan dirancang untuk memulihkan kepercayaan terhadap penyedia berita.
P7012	Standar untuk Ketentuan Privasi Pribadi yang Dapat Dibaca Mesin	Untuk menentukan bagaimana ketentuan privasi disajikan dan bagaimana ketentuan tersebut dapat dibaca dan diterima oleh mesin.
P7013	Standar Inklusi dan Penerapan Teknologi Analisis Wajah Otomatis	Untuk memberikan pedoman mengenai data yang digunakan dalam pengenalan wajah, persyaratan keragaman, dan tolok ukur aplikasi serta situasi di mana pengenalan wajah tidak boleh digunakan.

BAB 5

STRATEGI NASIONAL DAN INTERNASIONAL TENTANG AI

Seiring dengan kemajuan teknologi di balik AI yang melampaui ekspektasi, inisiatif kebijakan bermunculan di seluruh dunia untuk mengimbangi perkembangan ini. Strategi nasional pertama mengenai AI diluncurkan oleh Kanada pada bulan Maret 2017, diikuti oleh pemimpin teknologi Jepang dan Tiongkok. Di Eropa, Komisi Eropa mengedepankan komunikasi mengenai AI, memprakarsai pengembangan strategi independen oleh Negara-negara Anggota. Inisiatif AI Amerika diharapkan segera dilakukan, bersamaan dengan upaya intensif di Rusia untuk meresmikan 10 poin rencana AI mereka.

Inisiatif-inisiatif ini sangat berbeda dalam hal tujuan, besaran investasi, dan komitmen mereka terhadap pengembangan kerangka etika, yang diulas di sini pada bulan Mei 2019.



Gambar 5.1: Strategi Nasional dan Internasional mengenai AI yang dipublikasikan pada Mei 2019

5.1. EROPA

Komunikasi Komisi Eropa mengenai Kecerdasan Buatan (Komisi Eropa, 2018a), yang dirilis pada bulan April 2018, membuka jalan bagi strategi internasional pertama mengenai AI.

Dokumen tersebut menguraikan pendekatan terkoordinasi untuk memaksimalkan manfaat dan mengatasi tantangan yang ditimbulkan oleh AI.

Komunikasi mengenai AI diresmikan sembilan bulan kemudian dengan presentasi rencana terkoordinasi mengenai AI (Komisi Eropa, 2018b). Rencana tersebut merinci tujuh tujuan, yang mencakup pembiayaan start-up, investasi Rp. 6 Triliyun di beberapa 'pusat keunggulan penelitian', mendukung master dan PhD di bidang AI dan menciptakan ruang data umum Eropa.

Tujuan rencana ini adalah untuk mengembangkan 'pedoman etika dengan perspektif global'. Komisi menunjuk kelompok ahli tingkat tinggi yang independen untuk mengembangkan pedoman etika mereka, yang – setelah konsultasi – diterbitkan dalam bentuk finalnya pada bulan April 2019 (Kelompok Pakar Tingkat Tinggi Komisi Eropa untuk Kecerdasan Buatan, 2019). Pedoman ini mencantumkan persyaratan utama yang harus dipenuhi oleh sistem AI agar dapat dipercaya.

Tujuh persyaratan UE untuk AI yang dapat dipercaya:

1. Keagenan dan pengawasan manusia
2. Ketahanan dan keamanan teknis
3. Privasi dan tata kelola data
4. Transparansi
5. Keberagaman, non-diskriminasi dan keadilan
6. Kesejahteraan masyarakat dan lingkungan
7. Akuntabilitas

Kelompok Pakar Tingkat Tinggi AI Uni Eropa segera setelah itu merilis serangkaian pedoman kebijakan dan investasi lebih lanjut untuk AI yang dapat dipercaya (European Commission High-Level Expert Group on AI, 2019b), yang mencakup sejumlah rekomendasi penting seputar perlindungan manusia, peningkatan penyerapan AI di sektor swasta, memperluas kapasitas penelitian Eropa di bidang AI dan mengembangkan praktik pengelolaan data yang etis.

Dewan Eropa juga memiliki berbagai proyek yang sedang berjalan terkait penerapan AI dan pada bulan September 2019 membentuk Komite Ad Hoc untuk Kecerdasan Buatan (CAHAI). Komite ini akan menilai elemen-elemen potensial dari kerangka hukum untuk pengembangan dan penerapan AI, berdasarkan prinsip-prinsip dasar Dewan mengenai hak asasi manusia, demokrasi, dan supremasi hukum (Council of Europe, 2019a).

Ke depan, Presiden Komisi Eropa berikutnya, Ursula von der Leyen, telah mengumumkan AI sebagai prioritas Komisi berikutnya, termasuk undang-undang untuk pendekatan terkoordinasi mengenai 'implikasi AI pada manusia dan etika' (Kayali, 2019; von der Leyen, 2019).

Komisi Eropa memberikan kerangka pemersatu untuk pengembangan AI di UE, namun Negara-negara Anggota juga diwajibkan untuk mengembangkan strategi nasional mereka sendiri. Finlandia adalah Negara Anggota pertama yang mengembangkan program nasional AI (Kementerian Urusan Ekonomi dan Ketenagakerjaan Finlandia, 2018a). Program ini didasarkan pada dua laporan, Era Kecerdasan Buatan Finlandia dan Pekerjaan di Era

Kecerdasan Buatan (Kementerian Urusan Ekonomi dan Ketenagakerjaan Finlandia, 2017, 2018b). Tujuan kebijakan fokus pada investasi untuk daya saing dunia usaha dan pelayanan publik. Meskipun rekomendasi-rekomendasi telah dimasukkan ke dalam kebijakan, kelompok pengarah AI di Finlandia akan berjalan hingga akhir masa jabatan Pemerintah saat ini, dan laporan akhir diharapkan segera dikeluarkan.

Sejauh ini, Denmark, Perancis, Jerman, Swedia dan Inggris juga telah mengumumkan inisiatif nasional mengenai AI. Strategi Nasional Kecerdasan Buatan Denmark (Pemerintah Denmark, 2019) dirilis pada Maret 2019 dan mengikuti 'Strategi Pertumbuhan Digital' (Pemerintah Denmark, 2018). Kerangka kerja komprehensif ini mencantumkan tujuan-tujuan termasuk membangun landasan yang bertanggung jawab untuk AI, menyediakan data berkualitas tinggi, dan meningkatkan investasi AI secara keseluruhan (khususnya di sektor pertanian, energi, layanan kesehatan, dan transportasi). Terdapat fokus yang kuat pada etika data, termasuk tanggung jawab, keamanan dan transparansi, serta pengakuan akan perlunya kerangka etika. Pemerintah Denmark menguraikan enam prinsip AI yang etis – penentuan nasib sendiri, martabat, tanggung jawab, penjelasan, kesetaraan dan keadilan, dan pembangunan (solusi yang mendukung pengembangan dan penggunaan AI yang bertanggung jawab secara etis untuk mencapai kemajuan masyarakat) – dan akan membentuk sebuah Data Dewan Etik untuk memantau perkembangan teknologi di dalam negeri.

Di Perancis, 'AI for Humanity' diluncurkan pada bulan Maret 2018 dan berkomitmen untuk mendukung talenta Perancis, memanfaatkan data dengan lebih baik dan juga membangun kerangka etika mengenai AI (AI For Humanity, 2018). Presiden Macron telah berkomitmen untuk memastikan transparansi dan penggunaan AI yang adil, yang akan tertanam dalam sistem pendidikan. Strategi ini terutama didasarkan pada karya Cédric Villani, ahli matematika dan politisi Perancis, yang laporannya pada tahun 2018 tentang AI memberikan rekomendasi di bidang kebijakan ekonomi, infrastruktur penelitian, ketenagakerjaan, dan etika (Villani, 2018).

Strategi AI Jerman segera diadopsi pada bulan November 2018 (Die Bundesregierung, 2018) dan berisi tiga janji utama: menjadikan Jerman pemimpin global dalam pengembangan dan penggunaan AI, menjaga pengembangan dan penggunaan AI yang bertanggung jawab, dan mengintegrasikan AI dalam masyarakat dalam hal etika, hukum, budaya dan kelembagaan. Sasaran individu mencakup pengembangan Pusat Keunggulan untuk penelitian, penciptaan 100 jabatan profesor tambahan untuk AI, pendirian observatorium AI di Jerman, pendanaan 50 aplikasi unggulan AI untuk memberikan manfaat bagi lingkungan, pengembangan pedoman AI yang sesuai dengan undang-undang perlindungan data, dan pembentukan sebuah 'Dana Masa Depan Pekerjaan dan Masyarakat Digital' (De.digital, 2018).

Pendekatan Swedia terhadap AI (Government Offices of Sweden, 2018) memiliki istilah yang kurang spesifik, namun memberikan panduan umum mengenai pendidikan, penelitian, inovasi, dan infrastruktur untuk AI. Rekomendasinya mencakup membangun basis penelitian yang kuat, kolaborasi antar sektor dan dengan negara lain, mengembangkan upaya untuk mencegah dan mengelola risiko, serta mengembangkan standar untuk memandu penggunaan AI secara etis. Dewan AI Swedia, yang terdiri dari para ahli dari industri dan akademisi, juga

telah dibentuk untuk mengembangkan 'model Swedia' untuk AI, yang menurut mereka akan berkelanjutan, bermanfaat bagi masyarakat dan mendorong pertumbuhan ekonomi jangka panjang (Dewan AI Swedia, 2019).

Pemerintah Inggris mengeluarkan 'Kesepakatan Sektor AI' yang komprehensif pada bulan April 2018 (GOV.UK, 2018), yang merupakan bagian dari 'Strategi Industri' yang lebih besar, yang bertujuan untuk meningkatkan produktivitas melalui investasi dalam bisnis, keterampilan, dan infrastruktur (GOV.UK, 2019). Mereka menjanjikan hampir Rp. 1 Triliyun untuk mempromosikan AI di Inggris, dengan lima tema utama: ide, sumber daya manusia, infrastruktur, lingkungan bisnis, dan tempat.

Kebijakan utama mencakup peningkatan investasi penelitian dan pengembangan hingga total 2,4% PDB pada tahun 2027; menginvestasikan lebih dari Rp. 400 miliar dalam matematika, pendidikan digital dan teknik; mengembangkan skema pelatihan ulang nasional untuk menutup kesenjangan keterampilan dan berinvestasi dalam infrastruktur digital seperti kendaraan listrik dan jaringan serat optik. Selain komitmen investasi ini, kesepakatan ini juga mencakup pembentukan 'Pusat Etika dan Inovasi Data' (CDEI) untuk memastikan penggunaan AI yang aman dan etis. Pertama kali diumumkan pada anggaran tahun 2017, CDEI akan menilai risiko AI, meninjau kerangka peraturan dan tata kelola, serta memberikan saran kepada pemerintah dan pencipta teknologi mengenai praktik terbaik (Departemen Digital, Budaya, Media & Olahraga Pemerintah Inggris, 2019).

Beberapa negara Eropa lainnya sedang bersiap untuk merilis strategi nasional. Austria telah membentuk 'Dewan Robot' untuk membantu Pemerintah mengembangkan Strategi AI nasional (Austrian Council on Robotics and Artificial Intelligence, 2019). Buku putih yang disiapkan oleh Dewan menjadi landasan bagi strategi ini. Dokumen yang berfokus pada sosial ini mencakup tujuan untuk mempromosikan penggunaan AI yang bertanggung jawab, mengembangkan langkah-langkah untuk mengenali dan memitigasi bahaya, menciptakan kerangka hukum untuk melindungi keamanan data, dan melahirkan dialog publik seputar penggunaan AI (Austrian Council on Robotics and Artificial Intelligence, 2018).

Estonia biasanya cepat memanfaatkan teknologi baru, termasuk AI. Pada tahun 2017, Penasihat Inovasi Digital Estonia Marten Kaevats menggambarkan AI sebagai langkah selanjutnya dalam 'e-governance' di Estonia (Plantera, 2017). Faktanya, AI sudah banyak digunakan oleh pemerintah, yang saat ini sedang merancang strategi AI nasional (Castellanos, 2018). Rencana tersebut dilaporkan akan mempertimbangkan implikasi etis dari AI, serta menawarkan insentif ekonomi praktis dan program percontohan.

Satuan tugas AI telah dibentuk oleh Italia (Agency for Digital Italy, 2019) untuk mengidentifikasi peluang yang ditawarkan AI dan meningkatkan kualitas layanan publik. Buku putih mereka (Satuan Tugas Kecerdasan Buatan Badan Digital Italia, 2018), yang diterbitkan pada bulan Maret 2018, menggambarkan etika sebagai tantangan pertama bagi keberhasilan penerapan AI, yang menyatakan perlunya menjunjung tinggi prinsip bahwa AI harus menjadi prioritas utama. pelayanan kepada warga negara dan untuk memastikan kesetaraan dengan menggunakan teknologi untuk memenuhi kebutuhan universal. Gugus tugas ini menguraikan lebih lanjut tantangan-tantangan yang berkaitan dengan pengembangan teknologi,

kesenjangan keterampilan, aksesibilitas dan kualitas data, serta kerangka hukum. Buku ini memberikan total 10 rekomendasi kepada pemerintah, namun belum direalisasikan melalui kebijakan.

Malta, sebuah negara yang sebelumnya sangat fokus pada teknologi blockchain, kini telah mengumumkan rencananya untuk mengembangkan strategi AI nasional, menempatkan Malta 'di antara 10 negara teratas yang memiliki strategi nasional untuk AI' (Malta AI, 2019). Sebuah gugus tugas telah dibentuk yang terdiri dari perwakilan industri, akademisi, dan pakar lainnya untuk membantu merancang kebijakan bagi Malta yang akan fokus pada AI yang beretika, transparan, dan bertanggung jawab secara sosial sambil mengembangkan langkah-langkah yang mengumpulkan investasi asing, yang mencakup pengembangan keahlian dan infrastruktur yang dibutuhkan untuk mendukung AI di Malta.

Polandia juga sedang mengerjakan strategi AI nasionalnya. Sebuah laporan yang baru-baru ini dirilis oleh Digital Poland Foundation (2019) berfokus pada ekosistem AI di Polandia, sebagai cikal bakal strategi AI nasional. Meskipun buku ini memberikan gambaran komprehensif mengenai teknologi terkini di Polandia, buku ini tidak memberikan rekomendasi khusus kepada pemerintah, dan tidak menyebutkan masalah etika seputar AI.

Meskipun media melaporkan perkembangan AI yang berfokus pada militer di Rusia (Apps, 2019; Bershidski, 2017; Le Miere, 2017; O'Connor, 2017), negara tersebut saat ini tidak memiliki strategi nasional mengenai AI. Setelah konferensi 'Kecerdasan Buatan: Masalah dan Solusi' pada tahun 2018, Kementerian Pertahanan Rusia merilis daftar rekomendasi kebijakan, yang mencakup pembentukan sistem negara untuk pendidikan AI dan pusat nasional untuk AI. Laporan terbaru menunjukkan bahwa Presiden Putin telah menetapkan batas waktu hingga 15 Juni 2019 bagi pemerintahannya untuk menyelesaikan strategi nasional mengenai AI.

Di seluruh UE: Sikap masyarakat terhadap robot dan digitalisasi

Secara keseluruhan, survei mengenai perspektif Eropa terhadap AI, robotika, dan teknologi canggih (Komisi Eropa 2012; Komisi Eropa 2017) menunjukkan bahwa masyarakat secara umum mempunyai pandangan positif terhadap perkembangan ini, dan memandangnya sebagai dampak positif bagi masyarakat, perekonomian, dan warga negara. 'hidup. Namun, sikap ini berbeda-beda berdasarkan usia, jenis kelamin, tingkat pendidikan, dan lokasi dan sangat bergantung pada paparan seseorang terhadap robot dan informasi yang relevan — misalnya, hanya sejumlah kecil dari mereka yang disurvei yang benar-benar memiliki pengalaman menggunakan robot (dulu atau sekarang), dan mereka yang berpengalaman lebih cenderung memandangnya secara positif dibandingkan mereka yang tidak berpengalaman.

Tren umum persepsi masyarakat dari survei-survei tersebut menunjukkan bahwa responden adalah:

- Mendukung penggunaan robot dan digitalisasi dalam pekerjaan yang menimbulkan risiko atau kesulitan bagi manusia (seperti eksplorasi ruang angkasa, manufaktur, dan militer);
- Prihatin bahwa teknologi tersebut memerlukan pengelolaan yang efektif dan hati-hati;

- Khawatir bahwa otomatisasi dan digitalisasi akan mengakibatkan hilangnya lapangan kerja, dan tidak yakin apakah hal ini akan menstimulasi dan meningkatkan peluang kerja di seluruh UE;
- Tidak mendukung penggunaan robot untuk merawat anggota masyarakat yang rentan (orang lanjut usia, orang sakit, hewan peliharaan yang menjadi tanggungan mereka, atau mereka yang menjalani prosedur medis);
- Khawatir dalam mengakses dan melindungi data dan informasi online mereka, dan kemungkinan besar telah mengambil tindakan perlindungan dalam hal ini (perangkat lunak antivirus, perubahan perilaku penelusuran);
- Tidak bersedia mengendarai mobil tanpa pengemudi (hanya 22% yang bersedia melakukannya);
- Ketidakpercayaan terhadap media sosial, dengan hanya 7% yang menganggap cerita yang dipublikasikan di media sosial 'secara umum dapat dipercaya'; Dan
- Kecil kemungkinannya untuk menganggap meluasnya penggunaan robot sebagai hal yang bersifat jangka pendek, melainkan menganggapnya sebagai skenario yang akan terjadi setidaknya 20 tahun ke depan.

Oleh karena itu, kekhawatiran ini menonjol dalam inisiatif AI di Eropa, dan mencerminkan opini umum mengenai penerapan robot, AI, otomasi, dan digitalisasi di seluruh bidang kehidupan, pekerjaan, kesehatan, dan banyak lagi.

5.2. AMERIKA UTARA

Kanada adalah negara pertama di dunia yang meluncurkan strategi AI nasional pada bulan Maret 2017. Strategi Kecerdasan Buatan Pan-Kanada (Canadian Institute For Advanced Research, 2017) didirikan dengan empat tujuan utama, untuk: meningkatkan jumlah AI peneliti dan lulusan di Kanada; mendirikan pusat keunggulan ilmiah (di Edmonton, Montreal dan Toronto); mengembangkan kepemimpinan pemikiran global dalam implikasi AI terhadap ekonomi, etika, kebijakan dan hukum; dan mendukung komunitas riset nasional di bidang AI.

Sebuah program terpisah untuk AI dan masyarakat didedikasikan untuk membahas implikasi sosial dari AI, yang dipimpin oleh kelompok kerja terkait kebijakan yang mempublikasikan temuan mereka untuk pemerintah dan masyarakat. Bekerja sama dengan Pusat Penelitian Ilmiah Nasional Perancis (CNRS) dan Penelitian dan Inovasi Inggris (UKRI), program AI dan masyarakat baru-baru ini mengumumkan serangkaian lokakarya interdisipliner untuk mengeksplorasi isu-isu termasuk kepercayaan pada AI, dampak AI dalam layanan kesehatan. sektor ini dan bagaimana AI memengaruhi keragaman dan ekspresi budaya (Canadian Institute For Advanced Research, 2019).

Di AS, Presiden Trump mengeluarkan Perintah Eksekutif yang meluncurkan 'Inisiatif AI Amerika' pada bulan Februari 2019 (Gedung Putih, 2019a), yang segera diikuti dengan peluncuran situs web yang menyatukan semua inisiatif AI lainnya (Gedung Putih, 2019b), termasuk AI untuk Inovasi Amerika, AI untuk Industri Amerika, AI untuk Pekerja Amerika, dan AI untuk Nilai-Nilai Amerika. Inisiatif AI Amerika memiliki lima bidang utama: berinvestasi dalam penelitian dan pengembangan, memanfaatkan sumber daya AI (yaitu data dan daya

komputasi), menetapkan standar tata kelola, membangun tenaga kerja AI, dan keterlibatan internasional. Departemen Pertahanan juga telah menerbitkan strategi AI-nya sendiri (Departemen Pertahanan AS, 2018), dengan fokus pada kemampuan militer AI.

Pada bulan Mei, AS memajukan hal ini dengan AI Initiative Act, yang akan menginvestasikan Rp. 2,2 Miliar untuk mengembangkan strategi AI nasional, serta mendanai penelitian dan pengembangan federal. Undang-undang tersebut, yang berupaya untuk 'membentuk inisiatif Federal yang terkoordinasi untuk mempercepat penelitian dan pengembangan kecerdasan buatan untuk ekonomi dan keamanan nasional Amerika Serikat' berkomitmen untuk membentuk Kantor Koordinasi AI Nasional, menciptakan standar evaluasi AI, dan mendanai 5 penelitian AI nasional. pusat. Program ini juga akan mendanai National Science Foundation untuk meneliti dampak AI terhadap masyarakat, termasuk peran bias data, privasi dan akuntabilitas, serta memperluas upaya penelitian berbasis AI yang dipimpin oleh Departemen Energi (Kongres AS, 2019).

Pada bulan Juni 2019, Rencana Strategis Penelitian dan Pengembangan Kecerdasan Buatan Nasional dirilis, yang didasarkan pada rencana sebelumnya yang dikeluarkan oleh pemerintahan Obama dan mengidentifikasi delapan prioritas strategis, termasuk melakukan investasi jangka panjang dalam penelitian AI, mengembangkan metode efektif untuk AI manusia. kolaborasi, mengembangkan kumpulan data publik bersama, mengevaluasi teknologi AI melalui standar dan tolok ukur, serta memahami dan mengatasi implikasi AI terhadap etika, hukum, dan sosial. Dokumen tersebut memberikan strategi terkoordinasi untuk penelitian dan pengembangan AI di AS (National Science & Technology Council, 2019).

5.3. ASIA

Asia dalam banyak hal telah memimpin strategi AI, dan Jepang menjadi negara kedua yang meluncurkan inisiatif nasional mengenai AI. Dirilis pada bulan Maret 2017, Strategi Teknologi AI Jepang (Dewan Strategis Jepang untuk Teknologi AI, 2017) memberikan peta jalan industrialisasi, termasuk bidang prioritas di bidang kesehatan dan mobilitas, yang penting dengan mempertimbangkan populasi lansia di Jepang. Jepang membayangkan rencana pengembangan AI dalam tiga tahap, yang berpuncak pada ekosistem AI yang sepenuhnya terhubung dan bekerja di semua bidang masyarakat.

Singapura juga tidak ketinggalan. Pada bulan Mei 2017, AI Singapura diluncurkan, sebuah program lima tahun untuk meningkatkan kemampuan negara tersebut dalam bidang AI, dengan empat tema utama: industri dan perdagangan, kerangka kerja dan pengujian AI, bakat dan praktisi AI, serta Penelitian dan Pengembangan (AI Singapura, 2017). Pada tahun berikutnya, Pemerintah Singapura mengumumkan inisiatif tambahan yang berfokus pada tata kelola dan etika AI, termasuk pembentukan Dewan Penasihat untuk Penggunaan AI dan Data yang Etis, yang diresmikan dalam 'Model AI Governance Framework' (Personal Data Protection Commission Singapore, 2019). Kerangka kerja ini memberikan serangkaian prinsip-prinsip etika, yang diterjemahkan ke dalam langkah-langkah praktis yang dapat diterapkan oleh dunia usaha, termasuk cara mengelola risiko, cara memasukkan pengambilan keputusan manusia ke dalam AI, dan cara meminimalkan bias dalam kumpulan data.

Perekonomian Tiongkok telah mengalami pertumbuhan besar dalam beberapa dekade terakhir, menjadikannya negara dengan perekonomian terbesar kedua di dunia (World Economic Forum, 2018). Untuk menjadikan Tiongkok sebagai pemimpin dunia dalam bidang AI, Pemerintah Tiongkok merilis 'Rencana Pengembangan AI Generasi Berikutnya' pada bulan Juli 2017. Rencana terperinci tersebut menguraikan tujuan industrialisasi, penelitian dan pengembangan, pendidikan, standar etika, dan keamanan (Foundation for Law and International Affairs, 2017). Sejalan dengan Jepang, ini adalah strategi tiga langkah untuk pengembangan AI, yang berpuncak pada tahun 2030 dengan menjadi pusat inovasi AI terkemuka di dunia.

Terdapat fokus besar pada tata kelola, dengan maksud untuk mengembangkan peraturan dan norma etika untuk AI dan 'berpartisipasi aktif' dalam tata kelola global teknologi ini. Diformalkan dalam 'Rencana Aksi Tiga Tahun untuk Mempromosikan Pengembangan Industri Kecerdasan Buatan Generasi Baru', strategi ini mengulangi empat tujuan utama, yaitu: meningkatkan pengembangan produk-produk utama AI (dengan fokus pada kendaraan cerdas, robot layanan, diagnosis medis dan sistem identifikasi gambar video); meningkatkan kompetensi inti di bidang AI secara signifikan; memperdalam pengembangan manufaktur cerdas; dan membangun landasan bagi sistem pendukung industri AI (New America, 2018).

Di India, AI berpotensi menambah 1 triliun INR pada perekonomian pada tahun 2035 (NITI Aayog, 2018). Strategi AI di India, yang diberi nama AI untuk Semua, bertujuan untuk memanfaatkan manfaat AI untuk pertumbuhan ekonomi, serta pembangunan sosial dan 'pertumbuhan inklusif', dengan fokus signifikan pada pemberdayaan masyarakat untuk mendapatkan pekerjaan yang lebih berkualitas. Laporan tersebut memberikan 30 rekomendasi bagi pemerintah, yang mencakup pembentukan Pusat Penelitian Unggulan untuk AI (CORE, masing-masing memiliki Dewan Etiknya sendiri), mendorong peningkatan keterampilan karyawan, membuka kumpulan data pemerintah, dan mendirikan 'Pusat Studi Keberlanjutan Teknologi'. Hal ini juga menetapkan konsep India sebagai 'Garasi AI', yang memungkinkan solusi yang dikembangkan di India dapat diterapkan ke negara-negara berkembang di seluruh dunia.

Bersamaan dengan hal tersebut, Taiwan merilis 'Rencana Aksi AI' pada bulan Januari 2018 (AI Taiwan, 2018), yang sangat berfokus pada inovasi industri, dan Korea Selatan mengumumkan 'Strategi Pengembangan Industri Informasi AI' pada bulan Mei 2018 (H. Sarmah, 2019). Laporan yang mendasari hal ini (Pemerintah Republik Korea, 2016) memberikan rekomendasi yang cukup luas bagi pemerintah, dalam hal pengelolaan data, metode penelitian, AI dalam pemerintahan dan layanan publik, pendidikan, serta reformasi hukum dan etika.

Perdana Menteri Malaysia mengumumkan rencana untuk memperkenalkan kerangka AI nasional pada tahun 2017 (Abas, 2017), yang merupakan perpanjangan dari 'Kerangka Analisis Big Data' yang sudah ada dan dipimpin oleh Malaysia Digital Economy Corporation (MDEC). Belum ada kabar terbaru dari pemerintah sejak tahun 2017. Baru-baru ini, pengusaha terkaya di Sri Lanka, Dhammika Perera, menyerukan strategi AI nasional di negara tersebut,

pada sebuah acara yang diadakan bekerja sama dengan Computer Society of Sri Lanka (Cassim, 2019), Namun belum ada janji resmi dari pemerintah.

Di Timur Tengah, Uni Emirat Arab adalah negara pertama yang mengembangkan strategi AI, yang dirilis pada bulan Oktober 2017 dan dengan penekanan pada peningkatan kinerja pemerintah dan ketahanan keuangan (Pemerintah UEA, 2018). Investasi akan difokuskan pada pendidikan, transportasi, energi, teknologi, dan ruang angkasa. Etika yang mendasari kerangka ini cukup komprehensif; Pedoman Etika AI Dubai menentukan prinsip-prinsip utama yang menjadikan sistem AI adil, akuntabel, transparan, dan dapat dijelaskan (Smart Dubai, 2019a). Bahkan terdapat alat penilaian mandiri yang tersedia untuk membantu pengembang teknologi AI mengevaluasi etika sistem mereka (Smart Dubai, 2019b).

Pemimpin dunia di bidang teknologi, Israel, belum mengumumkan strategi AI nasional. Mengakui perlombaan global untuk kepemimpinan AI, laporan terbaru dari Otoritas Inovasi Israel (Otoritas Inovasi Israel, 2019) merekomendasikan agar Israel mengembangkan strategi AI nasional yang 'dibagikan oleh pemerintah, akademisi, dan industri'.

5.4. AFRIKA

Afrika menaruh perhatian besar pada AI; sebuah buku putih baru-baru ini menunjukkan bahwa teknologi ini dapat memecahkan beberapa masalah paling mendesak di Afrika Sub-Sahara, mulai dari hasil pertanian hingga penyediaan layanan keuangan yang aman (Access Partnership, 2018). Dokumen tersebut memberikan unsur-unsur penting bagi strategi pan-Afrika mengenai AI, yang menunjukkan bahwa kurangnya keterlibatan pemerintah hingga saat ini telah menjadi hambatan dan mendorong pemerintah Afrika untuk mengambil pendekatan proaktif terhadap kebijakan AI. Kebijakan ini mencantumkan undang-undang mengenai privasi dan keamanan data, inisiatif untuk mendorong adopsi cloud secara luas, peraturan yang memungkinkan penggunaan AI untuk penyediaan layanan publik, dan penerapan standar data internasional sebagai elemen kunci dari kebijakan tersebut, meskipun belum ada kebijakan yang mengatur hal tersebut. muncul.

Namun Kenya telah mengumumkan gugus tugas AI (dan blockchain) yang diketuai oleh mantan Sekretaris Kementerian Informasi dan Komunikasi, yang akan menawarkan rekomendasi kepada pemerintah tentang cara terbaik untuk memanfaatkan teknologi ini (Kenyan Wallstreet, 2018). Tunisia juga telah membentuk satuan tugas untuk menyusun strategi nasional mengenai AI dan mengadakan lokakarya pada tahun 2018 bertajuk 'Strategi AI Nasional: Membuka potensi kemampuan Tunisia' (ANPR, 2018).

5.5. AMERIKA SELATAN

Meksiko sejauh ini merupakan satu-satunya negara di Amerika Selatan yang meluncurkan strategi AI. Hal ini mencakup lima tindakan utama, yaitu: mengembangkan kerangka tata kelola yang memadai untuk mendorong dialog multi-sektoral; memetakan kebutuhan industri; mempromosikan kepemimpinan internasional Meksiko di bidang AI; mempublikasikan rekomendasi untuk konsultasi publik; dan bekerja sama dengan para ahli dan masyarakat untuk mencapai keberlangsungan upaya ini (México Digital, 2018).

Strateginya adalah dengan meresmikan Buku Putih (Martinho-Truswell dkk., 2018) yang disusun oleh Kedutaan Besar Inggris di Meksiko, perusahaan konsultan Oxford Insights, dan lembaga pemikir C Minds, bekerja sama dengan Pemerintah Meksiko.

Strategi ini menekankan peran masyarakat dalam pengembangan AI di Meksiko dan potensi penerapan AI di bidang sosial, seperti peningkatan layanan kesehatan dan pendidikan. Buku ini juga menjawab fakta bahwa 18% dari seluruh pekerjaan di Meksiko (total 9,8 juta) akan terkena dampak otomatisasi dalam 20 tahun mendatang dan memberikan sejumlah rekomendasi untuk meningkatkan pendidikan dalam pendekatan komputasi.

Negara-negara Amerika Selatan lainnya kemungkinan akan mengikuti langkah serupa jika mereka ingin mengimbangi pasar negara berkembang di Asia. Laporan terbaru menunjukkan bahwa AI dapat melipatgandakan perekonomian di Argentina, Brasil, Chili, Kolombia, dan Peru (Ovanessoff dan Plastino, 2017).

5.6. AUSTRALIA

Australia belum memiliki strategi nasional mengenai AI. Namun Australia memiliki 'Strategi Ekonomi Digital' (Pemerintah Australia, 2017) yang membahas pemberdayaan masyarakat Australia melalui 'keterampilan dan inklusi digital', yang mencantumkan AI sebagai teknologi utama yang sedang berkembang. Sebuah laporan mengenai 'Masa Depan Teknologi Australia' merinci lebih lanjut rencana AI, termasuk penggunaan AI untuk meningkatkan layanan publik, meningkatkan efisiensi administratif, dan meningkatkan pengembangan kebijakan (Pemerintah Australia, 2018).

Buku ini juga merinci rencana untuk mengembangkan kerangka etika dengan industri dan akademisi, di samping reformasi legislatif untuk menyederhanakan pembagian dan pelepasan data sektor publik. Rancangan kerangka etika (Dawson et al., 2019) didasarkan pada studi kasus dari seluruh dunia tentang AI yang 'salah' dan menawarkan delapan prinsip inti untuk mencegah hal ini, termasuk keadilan, akuntabilitas, dan perlindungan privasi. Ini adalah salah satu kerangka etika komprehensif yang diterbitkan sejauh ini, meskipun belum diterapkan.

Upaya juga sedang dilakukan untuk meluncurkan strategi nasional di Selandia Baru, di mana AI berpotensi meningkatkan PDB hingga Rp. 54 Triliyun (AI Forum New Zealand, 2018). Forum AI Selandia Baru dibentuk untuk meningkatkan kesadaran dan kemampuan AI di negara tersebut, yang mempertemukan masyarakat, industri, akademisi, dan pemerintah.

Laporan mereka yang berjudul 'Kecerdasan Buatan: Membentuk Masa Depan Selandia Baru' (AI Forum New Zealand, 2018) memaparkan sejumlah rekomendasi bagi pemerintah untuk mengoordinasikan pengembangan strategi (yaitu mengoordinasikan investasi penelitian dan penggunaan AI dalam layanan pemerintah); meningkatkan kesadaran akan AI (termasuk melakukan penelitian mengenai dampak AI terhadap perekonomian dan masyarakat); membantu adopsi AI (dengan mengembangkan sumber daya praktik terbaik untuk industri); meningkatkan aksesibilitas data tepercaya; menumbuhkan kumpulan bakat AI (mengembangkan kursus AI, memasukkan AI ke dalam daftar keterampilan berharga bagi para imigran); dan terakhir beradaptasi dengan dampak AI terhadap hukum, etika, dan

masyarakat. Hal ini termasuk rekomendasi untuk membentuk kelompok kerja etika AI dan masyarakat untuk menyelidiki masalah moral dan mengembangkan pedoman praktik terbaik dalam AI, yang selaras dengan badan-badan internasional.

Tantangan terhadap adopsi AI oleh pemerintah

Forum Ekonomi Dunia, melalui konsultasi dengan para pemangku kepentingan, telah mengidentifikasi lima hambatan utama dalam penerapan AI oleh pemerintah:

1. *Penggunaan data yang efektif* - Kurangnya pemahaman tentang infrastruktur data, tidak menerapkan proses tata kelola data (misalnya mempekerjakan petugas dan alat data untuk mengakses data secara efisien).
2. *Keterampilan data dan AI* - Sulit bagi pemerintah, yang memiliki anggaran perekrutan yang lebih kecil dibandingkan banyak perusahaan besar, untuk menarik kandidat dengan keterampilan yang diperlukan untuk mengembangkan solusi AI terbaik.
3. *Ekosistem AI* - Ada banyak perusahaan berbeda yang beroperasi di pasar AI dan pasar ini berubah dengan cepat. Banyak perusahaan rintisan yang memelopori solusi AI memiliki pengalaman terbatas dalam bekerja sama dengan pemerintah dan mengembangkan proyek-proyek besar.
4. *Budaya warisan* – Mengadopsi teknologi transformatif di pemerintahan bisa jadi sulit, karena terdapat praktik dan proses yang sudah mapan dan mungkin kurang adanya dorongan bagi pegawai untuk mengambil risiko dan berinovasi dibandingkan di sektor swasta.
5. *Mekanisme pengadaan* - Sektor swasta memperlakukan algoritme sebagai kekayaan intelektual, sehingga mungkin menyulitkan pemerintah untuk menyesuaikannya sesuai kebutuhan. Mekanisme pengadaan publik juga bisa berjalan lambat dan rumit (misalnya syarat dan ketentuan yang panjang, waktu tunggu yang lama sejak penyerahan tanggapan tender hingga keputusan akhir).

5.7. INISIATIF AI INTERNASIONAL, SELAIN UE

Selain UE, terdapat semakin banyak strategi internasional mengenai AI, yang bertujuan untuk memberikan kerangka kerja pemersatu bagi pemerintah di seluruh dunia dalam pengelolaan teknologi baru dan canggih ini.

Visi Bersama G7 untuk Masa Depan AI

Pada pertemuan G7 tahun 2018 di Charlevoix, Kanada, para pemimpin G7 (Kanada, Prancis, Jerman, Italia, Jepang, Inggris, dan Amerika Serikat) berkomitmen terhadap 12 prinsip AI, yang dirangkum di bawah ini:

1. Mempromosikan AI yang berpusat pada manusia dan adopsi AI secara komersial, dan terus memajukan pendekatan teknis, etika, dan netral teknologi yang sesuai.
2. Mendorong investasi dalam penelitian dan pengembangan AI yang menghasilkan pengujian publik terhadap teknologi baru dan mendukung pertumbuhan ekonomi.
3. Mendukung pendidikan, pelatihan dan pelatihan ulang keterampilan bagi angkatan kerja.

4. Mendukung dan melibatkan kelompok yang kurang terwakili, termasuk perempuan dan individu yang terpinggirkan, dalam pengembangan dan penerapan AI.
5. Memfasilitasi dialog multi-pemangku kepentingan tentang cara memajukan inovasi AI untuk meningkatkan kepercayaan dan adopsi.
6. Mendukung upaya untuk meningkatkan kepercayaan terhadap AI, dengan perhatian khusus untuk melawan stereotip yang merugikan dan mendorong kesetaraan gender. Menumbuhkan inisiatif yang mempromosikan keselamatan dan transparansi.
7. Mempromosikan penggunaan AI oleh usaha kecil dan menengah.
8. Mempromosikan kebijakan pasar tenaga kerja yang aktif, pengembangan tenaga kerja dan program pelatihan untuk mengembangkan keterampilan yang diperlukan untuk pekerjaan baru.
9. Mendorong investasi pada AI.
10. Mendorong inisiatif untuk meningkatkan keamanan digital dan mengembangkan kode etik.
11. Memastikan pengembangan kerangka privasi dan perlindungan data.
12. Mendukung lingkungan pasar terbuka untuk aliran data yang bebas, dengan tetap menghormati privasi dan perlindungan data.

Deklarasi Wilayah Nordik-Baltik tentang AI

Deklarasi yang ditandatangani oleh Wilayah Nordik-Baltik (terdiri dari Denmark, Estonia, Finlandia, Kepulauan Faroe, Islandia, Latvia, Lituania, Norwegia, Swedia, dan Kepulauan Åland) bertujuan untuk mempromosikan penggunaan AI di wilayah tersebut, termasuk meningkatkan peluang untuk pengembangan keterampilan, peningkatan akses terhadap data dan tujuan kebijakan khusus untuk mengembangkan 'pedoman, standar, prinsip dan nilai yang etis dan transparan' mengenai kapan dan bagaimana AI harus digunakan (Nordic Co-operation, 2018).

Prinsip OECD tentang AI

Pada tanggal 22 Mei 2019, Organisasi untuk Kerja Sama dan Pembangunan Ekonomi mengeluarkan prinsip-prinsip AI, yang merupakan standar internasional pertama yang disetujui oleh pemerintah untuk pengembangan AI yang bertanggung jawab. Hal ini mencakup rekomendasi kebijakan praktis serta prinsip-prinsip berbasis nilai untuk 'pengelolaan AI yang bertanggung jawab', yang dirangkum di bawah ini:

- AI harus memberi manfaat bagi manusia dan bumi dengan mendorong pertumbuhan inklusif, pembangunan berkelanjutan, dan kesejahteraan.
- Sistem AI harus menghormati supremasi hukum, hak asasi manusia, nilai-nilai demokrasi dan keberagaman, dan harus mencakup perlindungan yang tepat untuk memastikan masyarakat yang adil.
- Harus ada transparansi seputar AI untuk memastikan bahwa masyarakat memahami hasil dan dapat menentangnya.
- Sistem AI harus berfungsi dengan kuat, aman, dan terlindungi sepanjang siklus hidupnya dan risikonya harus terus dinilai.

- Organisasi dan individu yang mengembangkan, menerapkan, atau mengoperasikan sistem AI harus bertanggung jawab.

Prinsip-prinsip ini telah disetujui oleh pemerintah dari 36 Negara Anggota OECD serta Argentina, Brasil, Kolombia, Kosta Rika, Peru, dan Rumania (OECD, 2019a). Prinsip-prinsip AI yang berpusat pada manusia di G20 dirilis pada bulan Juni 2019 dan diambil dari Prinsip-prinsip OECD (G20, 2019).

Persatuan negara-negara

PBB memiliki beberapa inisiatif terkait AI, antara lain:

- KTT Global AI untuk Kebaikan- KTT yang diadakan sejak tahun 2017 berfokus pada strategi untuk memastikan pengembangan AI yang aman dan inklusif (International Telecommunication Union, 2018a,b). Acara ini diselenggarakan oleh International Telecommunication Union, yang bertujuan untuk 'menyediakan platform netral bagi pemerintah, industri, dan akademisi untuk membangun pemahaman bersama tentang kemampuan teknologi AI yang sedang berkembang dan kebutuhan akan standarisasi teknis dan panduan kebijakan.'
- Pusat AI dan Robotika UNICRI - Lembaga Penelitian Kejahatan dan Keadilan Antarwilayah PBB (UNICRI) meluncurkan program AI dan Robotika pada tahun 2015 dan akan membuka pusat yang didedikasikan untuk topik-topik ini di Den Haag (UNICRI, 2019).
- Laporan UNESCO tentang Etika Robotika - Komisi Dunia UNESCO untuk Etika Pengetahuan Ilmiah dan Teknologi (COMEST) telah menulis laporan tentang 'Etika Robotika', yang membahas tantangan etika robot dalam masyarakat dan memberikan prinsip-prinsip dan nilai-nilai etika, dan kerangka etika berbasis teknologi (COMEST, 2017).

Forum Ekonomi Dunia

Forum Ekonomi Dunia (WEF) membentuk Dewan AI Global pada bulan Mei 2019, yang diketuai bersama oleh pengembang pengenalan suara Kai-Fu Lee, sebelumnya dari Apple, Microsoft dan Google, dan Presiden Microsoft saat ini Bradford Smith. Salah satu dari enam dewan 'Revolusi Industri Keempat', Dewan AI Global akan mengembangkan panduan kebijakan dan mengatasi kesenjangan tata kelola, guna mengembangkan pemahaman bersama di antara negara-negara yang memiliki praktik terbaik dalam kebijakan AI (World Economic Forum, 2019a).

Pada bulan Oktober 2019, mereka merilis kerangka kerja untuk mengembangkan strategi nasional AI untuk memandu pemerintah yang belum atau sedang mengembangkan strategi nasional AI. WEF menggambarkannya sebagai cara untuk menciptakan strategi AI yang 'minimum layak' dan mencakup empat tahap utama:

- 1) Menilai prioritas strategis jangka panjang
- 2) Menetapkan tujuan dan target nasional
- 3) Membuat rencana untuk elemen-elemen strategis yang penting
- 4) Mengembangkan rencana implementasi

WEF juga telah mengumumkan rencana untuk mengembangkan 'AI toolkit' untuk membantu dunia usaha menerapkan AI dengan sebaik-baiknya dan membentuk dewan etika mereka sendiri, yang akan dirilis pada konferensi Davos tahun 2020 (Vanian, 2019).

5.8. KESIAPAN PEMERINTAH TERHADAP AI

Sebuah laporan yang dibuat oleh Pusat Penelitian Pembangunan Internasional Kanada (Oxford Insights, 2019) mengevaluasi 'kesiapan AI' pemerintah di seluruh dunia pada tahun 2019, dengan menggunakan serangkaian data termasuk tidak hanya keberadaan strategi AI nasional, namun juga undang-undang perlindungan data, statistik tentang startup AI dan keterampilan teknologi.

Singapura berada di peringkat nomor 1 dalam estimasi mereka, dan Jepang sebagai satu-satunya negara Asia lainnya yang berada di peringkat 10 besar. Enam puluh persen negara yang masuk dalam 10 besar adalah negara-negara Eropa, dan sisanya berasal dari Amerika Utara.

Keterwakilan Eropa yang kuat dalam analisis ini mencerminkan nilai kerangka UE yang menyatukan, serta kekuatan ekonomi Eropa. Analisis ini juga memuji strategi kebijakan masing-masing negara Eropa, yang, yang penting, telah dikembangkan dalam budaya kolaborasi. Contoh pendekatan kolaboratif ini mencakup Deklarasi Kerja Sama UE mengenai AI (Komisi Eropa, 2018d), yang menyatakan negara-negara anggota sepakat untuk bekerja sama dalam meningkatkan kapasitas Eropa dalam bidang AI, dan kemitraan individual antara negara-negara anggota, seperti Finlandia, Estonia, dan Swedia. , bekerja sama untuk menguji coba aplikasi AI baru.

Tabel 5.1 Peringkat 10 teratas untuk Kesiapan AI Pemerintah tahun 2018/19.

Pangkat	Negara	Skor
1	Singapura	9.19
2	Britania Raya	9.07
3	Jerman	8.81
4	Amerika Serikat	8.80
5	Finlandia	8.77
6	Swedia	8.67
6	Kanada	8.67
8	Perancis	8.61
9	Denmark	8.60
10	Jepang	8.58

Singapura menempati peringkat tertinggi di antara semua negara, sementara Jepang, negara kedua di dunia yang merilis strategi nasional mengenai AI, berada di peringkat ke-10. Posisi Tiongkok di peringkat ke-21 dunia diperkirakan akan meningkat pada tahun depan seiring dengan mulai membuahkan hasil atas investasi Tiongkok di bidang AI. Kemajuan di Asia secara keseluruhan tidak seimbang, dengan dua negara di kawasan ini juga berada di peringkat sepuluh terbawah di seluruh dunia, yang mencerminkan ketimpangan pendapatan di kawasan ini.

Meskipun perkembangan strategi nasional mereka relatif lambat, Amerika Serikat berada di peringkat ke-4, diikuti Kanada. Kedua negara ini didukung oleh perekonomian yang kuat, tenaga kerja yang berketerampilan tinggi, inovasi sektor swasta, dan data yang

melimpah, sehingga wilayah-wilayah yang tidak termasuk dalam 10 besar negara – Afrika, Amerika Selatan, dan Australasia – tidak mampu bersaing.

Kerangka kerja ini memberikan metrik yang sangat berguna untuk menilai kemampuan pemerintah dalam memanfaatkan potensi AI di tahun-tahun mendatang. Namun yang tidak dipertimbangkan oleh analisis ini adalah seberapa kuat setiap negara mempertimbangkan isu moral dan etika seputar penggunaan AI, yang akan kita bahas di bawah ini.

BAB 6

TEMA YANG MUNCUL

. Hal ini berkaitan dengan masalah kepercayaan dan ditangani dengan berbagai cara melalui inisiatif etika yang muncul. Standar dan peraturan juga mulai dikembangkan untuk mengatasi permasalahan ini. Namun, fokus dari banyak strategi AI yang ada saat ini adalah untuk memungkinkan pengembangan teknologi dan, meskipun permasalahan etika telah diatasi, kesenjangan yang signifikan dapat diidentifikasi.

6.1. MENGATASI MASALAH ETIKA MELALUI STRATEGI NASIONAL DAN INTERNASIONAL

Ada beberapa tema yang dimiliki bersama oleh berbagai strategi nasional mengenai AI, di antaranya industrialisasi dan produktivitas mungkin menduduki peringkat tertinggi. Semua negara mempunyai strategi industri untuk AI, dan hal ini terutama menonjol di negara-negara berkembang di Asia Tenggara. Sebagian besar strategi tersebut mengacu pada pentingnya AI bagi daya saing bisnis dan beberapa di antaranya, termasuk Jerman, Korea Selatan, Taiwan, dan Inggris, mengumumkan pendanaan tambahan dan inkubator khusus untuk perusahaan rintisan yang berfokus pada AI.

Baik di sektor swasta maupun publik, pentingnya penelitian dan pengembangan juga diakui secara universal, dengan hampir semua strategi menjanjikan peningkatan pendanaan untuk penelitian dan banyak di antaranya yang mendirikan 'pusat keunggulan' yang sepenuhnya didedikasikan untuk penelitian AI, termasuk strategi dari Kanada, Jerman dan India.

Hal yang penting dalam mengembangkan upaya penelitian yang kuat adalah bakat, sehingga investasi pada sumber daya manusia dan pendidikan juga merupakan hal yang penting dalam sebagian besar strategi. Inggris telah mengumumkan 'Turing Fellowships' untuk mendanai akademisi baru yang mengeksplorasi pendekatan komputasi, sementara Jerman telah menyediakan setidaknya 100 profesor tambahan yang bekerja di bidang AI – keduanya di bawah payung komitmen UE untuk melatih, menarik, dan mempertahankan talenta. Di Asia, Korea Selatan telah berkomitmen untuk mengembangkan enam program pascasarjana baru untuk melatih total 5.000 spesialis AI, sementara Taiwan telah berkomitmen untuk melatih dua kali lipat jumlah tersebut pada tahun 2021.

Sebagian besar strategi juga mempertimbangkan dampak revolusi AI terhadap angkatan kerja yang tidak melek teknologi, yang mungkin merupakan kelompok pertama yang kehilangan pekerjaan karena otomatisasi. Meskipun hal ini juga mencakup pertimbangan etis, beberapa strategi membuat komitmen praktis untuk program pelatihan ulang guna membantu mereka yang terkena dampak mendapatkan pekerjaan baru. Hal ini merupakan tujuan utama dalam rencana UE, dan juga merupakan rencana Negara-negara Anggotanya. Inggris misalnya akan memulai skema pelatihan ulang senilai > Rp. 700 Miliar untuk membantu masyarakat memperoleh keterampilan digital dan Jerman telah meluncurkan 'Strategi Pelatihan Lanjutan Nasional' yang serupa. Tentu saja, negara-negara yang paling

membutuhkan pelatihan ulang mempunyai dana yang paling sedikit tersedia untuk itu. Namun strategi Meksiko menekankan pentingnya pemikiran komputasi dan matematika dalam pengajaran seumur hidup, termasuk untuk membantu warganya melakukan pelatihan ulang, sementara India berjanji untuk mempromosikan lembaga pelatihan informal dan menciptakan insentif finansial untuk meningkatkan keterampilan karyawan. Namun strategi lain menyarankan bahwa pelatihan ulang merupakan tanggung jawab masing-masing perusahaan dan tidak mengalokasikan dana terpisah untuk pelatihan tersebut.

Kolaborasi antar sektor dan negara merupakan benang merah lainnya, namun dimaknai berbeda oleh negara yang berbeda. Pendekatan yang dilakukan India misalnya adalah berbagi; Konsep 'AI Garage' yang disebutkan dalam strategi mereka berarti solusi berbasis AI yang dikembangkan di India akan diterapkan ke negara-negara berkembang yang menghadapi masalah serupa. Sebaliknya, Pemerintah Eksekutif AS mengenai AI bertujuan untuk 'mempromosikan lingkungan internasional yang mendukung AI Amerika' sekaligus melindungi keunggulan teknologi negara tersebut dari 'musuh asing'. Tentu saja, strategi negara-negara anggota UE menunjukkan kecenderungan untuk melakukan kolaborasi lintas batas. Swedia misalnya menyatakan perlunya mengembangkan kemitraan dan kolaborasi dengan negara-negara lain 'terutama di dalam UE', sementara strategi Denmark juga menekankan kerja sama yang erat dengan negara-negara Eropa lainnya.

Demokratisasi teknologi berpotensi mengurangi kesenjangan dalam masyarakat, dan inklusi serta pembangunan sosial merupakan tujuan penting dari banyak inisiatif AI nasional, khususnya di negara-negara berkembang. Strategi India membahas AI untuk 'kebaikan yang lebih besar', dengan fokus pada kemungkinan akses yang lebih baik terhadap layanan kesehatan, pertumbuhan ekonomi bagi kelompok yang sebelumnya tidak termasuk dalam produk keuangan formal, dan penggunaan data untuk membantu petani skala kecil. Strategi Meksiko mencantumkan inklusi sebagai salah satu dari lima tujuan utamanya, yang mencakup tujuan untuk mendemokratisasi produktivitas dan mendorong kesetaraan gender. Perancis juga menginginkan AI yang 'mendukung inklusivitas', mengupayakan kebijakan yang mengurangi kesenjangan sosial dan ekonomi.

Menentukan siapa yang bertanggung jawab atas tindakan dan perilaku AI sangatlah penting dan menantang baik dari segi moral maupun hukum. Saat ini, AI kemungkinan besar dianggap sebagai tanggung jawab hukum dari aktor manusia yang relevan – sebuah alat yang ada di tangan pengembang, pengguna, vendor, dan sebagainya. Namun, kerangka kerja ini tidak memperhitungkan tantangan unik yang ditimbulkan oleh AI, dan terdapat banyak area abu-abu. Sebagai satu contoh saja, ketika sebuah mesin belajar dan berevolusi menjadi berbeda dengan pemrograman awalnya melalui banyak iterasi, mungkin akan menjadi lebih sulit untuk memberikan tanggung jawab atas perilakunya kepada pemrogram. Demikian pula, jika pengguna atau vendor tidak diberi pengarahan yang memadai tentang keterbatasan agen AI, mereka mungkin tidak dapat dimintai pertanggungjawaban. Tanpa membuktikan bahwa agen AI bermaksud melakukan kejahatan (*mens rea*) dan dapat bertindak secara sukarela, yang keduanya merupakan konsep kontroversial, maka agen AI tidak dapat dianggap bertanggung jawab dan berkewajiban atas tindakannya sendiri.

6.2. MENGATASI TANTANGAN TATA KELOLA YANG DITIMBULKAN OLEH AI

Saat ini terdapat dua kerangka kerja internasional utama untuk tata kelola AI: kerangka UE (lihat Bagian 5.1) dan Organisasi untuk Kerja Sama dan Pembangunan Ekonomi (OECD). OECD meluncurkan serangkaian prinsip AI pada Mei 2019 (OECD, 2019a) yang saat itu diadopsi oleh 42 negara. Kerangka kerja OECD menawarkan lima prinsip dasar pengoperasian AI (lihat bagian 5.1.1) serta rekomendasi praktis yang menyertainya bagi pemerintah untuk mencapainya. G20 segera mengadopsi prinsip-prinsip AI mereka sendiri yang berpusat pada manusia, yang diambil dari (dan pada dasarnya merupakan versi ringkasan) prinsip-prinsip OECD (G20, 2019).

Prinsip-prinsip OECD juga didukung oleh Komisi Eropa, yang memiliki strateginya sendiri mengenai AI sejak April 2018 (Komisi Eropa, 2018b). Kerangka kerja UE mencakup rencana investasi yang komprehensif, namun juga melakukan persiapan menghadapi perubahan sosio-ekonomi yang kompleks dan dilengkapi dengan serangkaian pedoman etika terpisah (European Commission High-Level Expert Group on AI, 2019a).

Kesenjangan dalam kerangka AI

Kerangka kerja ini mengatasi dilema moral dan etika yang diidentifikasi dalam buku ini pada tingkat yang berbeda-beda, dengan beberapa kesenjangan yang mencolok. Mengenai masalah lingkungan hidup, meskipun OECD merujuk pada pengembangan AI yang memberikan hasil positif bagi planet ini, termasuk melindungi lingkungan alam, dokumen tersebut tidak menyarankan cara untuk mencapai hal ini, juga tidak menyebutkan tantangan lingkungan spesifik apa pun yang perlu dipertimbangkan.

Komunikasi UE mengenai AI tidak membahas lingkungan hidup. Namun pedoman etika yang menyertainya didasarkan pada prinsip pencegahan bahaya, yang mencakup kerusakan terhadap lingkungan alam dan seluruh makhluk hidup. Kesejahteraan masyarakat dan lingkungan (termasuk keberlanjutan dan 'keramahan lingkungan') adalah salah satu persyaratan UE untuk AI yang dapat dipercaya dan daftar penilaiannya mencakup pertimbangan eksplisit atas risiko terhadap lingkungan atau hewan. Contoh-contoh khusus juga diberikan mengenai cara mencapai hal ini (misalnya penilaian kritis terhadap penggunaan sumber daya dan konsumsi energi di seluruh rantai pasokan).

Dampak terhadap psikologi manusia, termasuk cara manusia berinteraksi dengan AI dan dampak selanjutnya terhadap cara manusia berinteraksi satu sama lain, dapat dibahas lebih lanjut dalam kerangka kerja ini. Dampak psikososial AI tidak dipertimbangkan dalam Prinsip OECD atau Komunikasi UE. Namun, persyaratan UE untuk mempertimbangkan kesejahteraan masyarakat memang membahas 'dampak sosial', yang mencakup kemungkinan perubahan pada hubungan sosial dan hilangnya keterampilan sosial. Pedoman tersebut menyatakan bahwa dampak tersebut harus 'dipantau dan dipertimbangkan secara hati-hati' dan bahwa interaksi AI dengan manusia harus memberikan sinyal yang jelas bahwa interaksi sosialnya merupakan simulasi. Namun, pertimbangan yang lebih spesifik dapat diberikan pada hubungan manusia-robot atau dampak yang lebih kompleks pada jiwa manusia, seperti yang diuraikan di atas.

Prinsip pertama OECD mengenai pertumbuhan inklusif, pembangunan berkelanjutan dan kesejahteraan menyatakan bahwa AI harus dikembangkan dengan cara yang mengurangi 'kesenjangan ekonomi, sosial, gender dan lainnya'. Hal ini juga tercakup dalam prinsip kedua OECD, yang menyatakan bahwa sistem AI harus menghormati keberagaman dan mencakup upaya perlindungan untuk memastikan masyarakat yang adil, namun rincian mengenai bagaimana hal ini dapat dicapai masih kurang.

Pedoman etika UE lebih komprehensif dalam hal ini dan memasukkan keberagaman, non-diskriminasi dan keadilan sebagai persyaratan terpisah. Pedoman tersebut menguraikan bahwa kesetaraan adalah landasan mendasar bagi AI yang dapat dipercaya dan menyatakan bahwa AI harus dilatih berdasarkan data yang mewakili kelompok berbeda untuk mencegah keluaran yang bias. Pedoman tersebut mencakup rekomendasi tambahan untuk menghindari bias yang tidak adil.

Kedua kerangka tersebut memasukkan hak asasi manusia dan nilai-nilai demokrasi sebagai prinsip utama. Hal ini mencakup privasi, yang merupakan salah satu nilai-nilai OECD yang berpusat pada manusia dan persyaratan utama pedoman etika UE, yang menguraikan pentingnya tata kelola data dan aturan akses data. Masalah privasi juga tercakup dalam pedoman perlindungan data OECD yang ada (OECD, 2013).

Implikasi AI terhadap demokrasi hanya disebutkan secara singkat oleh OECD, dan tidak membahas isu-isu tertentu yang dihadapi pemerintah saat ini, seperti Deepfake atau manipulasi opini melalui berita yang ditargetkan. Ancaman terhadap demokrasi tidak disebutkan sama sekali dalam Komunikasi UE, meskipun masyarakat dan demokrasi adalah tema utama dalam pedoman etika terkait, yang menyatakan bahwa sistem AI harus berfungsi untuk menjaga demokrasi dan tidak melemahkan 'proses demokrasi, musyawarah manusia, atau sistem pemungutan suara yang demokratis.'

Permasalahan ini merupakan bagian dari pertanyaan yang lebih besar seputar perubahan sistem hukum yang mungkin diperlukan di era AI, termasuk pertanyaan penting seputar tanggung jawab atas pelanggaran yang melibatkan AI. Masalah tanggung jawab secara eksplisit dibahas oleh UE baik dalam pedoman Komunikasi maupun etika. Memastikan kerangka hukum yang tepat merupakan persyaratan utama Komunikasi UE tentang AI, yang mencakup panduan mengenai tanggung jawab produk dan eksplorasi masalah keselamatan dan keamanan (termasuk penggunaan kriminal). Pedoman etika yang menyertainya juga dapat menangani masalah ini dengan tepat, termasuk memberikan panduan bagi pengembang tentang cara memastikan kepatuhan hukum. Perubahan terkait regulasi dibahas lebih lanjut dalam Kebijakan AI dan Rekomendasi Investasi terbaru (European Commission High-Level Expert Group on AI, 2019b), yang mengeksplorasi potensi perubahan pada undang-undang UE saat ini dan perlunya kewenangan regulasi baru.

Prinsip-prinsip OECD lebih terbatas dalam hal ini. Meskipun peraturan ini memberikan panduan bagi pemerintah untuk menciptakan 'lingkungan kebijakan yang mendukung' AI, termasuk rekomendasi untuk meninjau dan mengadaptasi kerangka peraturan, hal ini dinyatakan bertujuan untuk mendorong 'inovasi dan persaingan' dan tidak mengatasi masalah tanggung jawab. untuk kejahatan yang dibantu AI.

Namun pertanyaan-pertanyaan ini juga dapat muncul dalam isu akuntabilitas, yang dapat diatasi dengan baik oleh kedua kerangka tersebut. OECD mencantumkan akuntabilitas sebagai prinsip utama dan menyatakan bahwa 'organisasi dan individu yang mengembangkan, menerapkan, atau mengoperasikan sistem AI harus bertanggung jawab atas berfungsinya sistem tersebut dengan baik' (OECD, 2019a). Hal ini juga merupakan prinsip inti dari pedoman etika UE, yang menyediakan lebih dari 10 syarat akuntabilitas dalam daftar penilaiannya untuk AI yang dapat dipercaya.

Banyak permasalahan di atas yang pada akhirnya penting untuk membangun kepercayaan terhadap AI, yang juga mengharuskan AI bersikap adil dan transparan. Masalah-masalah ini merupakan dasar dari pedoman etika UE yang membahasnya dengan sangat rinci. OECD juga menyatakan bahwa sistem AI harus menjamin 'masyarakat yang adil dan adil'. Transparansi dan kemampuan menjelaskan adalah prinsip inti OECD, dengan penekanan kuat pada fakta bahwa masyarakat harus mampu memahami dan menantang sistem AI. Prinsip-prinsip OECD kurang memberikan konteks mengenai isu-isu ini dan tidak mempertimbangkan cara-cara praktis untuk memastikan hal ini (misalnya audit algoritma), yang dipertimbangkan oleh pedoman etika UE. Pedoman etika juga mempertimbangkan perlunya pengawasan manusia (termasuk diskusi tentang pendekatan human-in-the-loop dan perlunya 'tombol berhenti', yang keduanya tidak disebutkan dalam prinsip-prinsip OECD).

Yang terakhir, meskipun keduanya mengakui manfaat penggunaan AI di bidang keuangan (Bagian 2.3), tidak ada kerangka kerja yang mampu mengatasi potensi dampak negatif pada sistem keuangan, baik melalui kerugian yang tidak disengaja atau aktivitas jahat. Potensi kejahatan keuangan yang dibantu oleh AI merupakan salah satu potensi yang penting dan saat ini belum ditangani oleh kerangka kerja internasional mana pun. Namun, G7 baru-baru ini menyuarakan kekhawatirannya mengenai mata uang digital dan berbagai produk keuangan baru lainnya yang sedang dikembangkan (Reuters, 2019), yang menunjukkan bahwa perubahan peraturan dalam hal ini sedang terjadi.

BAB 7

PENUTUP

Apa yang dijelaskan dalam buku ini adalah keragaman dan kompleksitas permasalahan etika yang timbul dari pengembangan kecerdasan buatan; mulai dari masalah berskala besar seperti hilangnya pekerjaan akibat otomatisasi, degradasi lingkungan, dan kesenjangan yang semakin parah, hingga permasalahan moral yang lebih bersifat pribadi seperti bagaimana AI dapat memengaruhi privasi kita, kemampuan kita untuk menilai apa yang nyata, dan hubungan pribadi kita.

Yang juga jelas adalah bahwa ada berbagai pendekatan terhadap etika. Prinsip-prinsip etika yang kuat sangat penting di masa depan teknologi yang berkembang pesat ini, namun tidak semua negara memahami etika dengan cara yang sama. Ada sejumlah inisiatif etika independen untuk AI, seperti Institut Etika AI Jerman, yang didanai oleh Facebook, dan Future of Life Institute yang didanai oleh donor swasta di AS. Semakin banyak negara yang juga mengembangkan strategi AI nasional, dengan komponen etikanya masing-masing. Sejumlah negara telah berkomitmen untuk membentuk dewan etika AI, termasuk Jerman, Inggris, India, Singapura, dan Meksiko. UEA juga memprioritaskan etika dalam strategi nasionalnya, dengan mengembangkan 'Perangkat AI yang Etis' dan alat penilaian mandiri untuk pengembang, sementara beberapa negara lain hanya memberikan referensi sekilas; etika hampir sepenuhnya ditinggalkan oleh Jepang, Korea Selatan dan Taiwan.

Penilaian kami menunjukkan bahwa sebagian besar permasalahan etika yang diidentifikasi di sini juga ditangani dalam beberapa bentuk oleh setidaknya salah satu kerangka kerja internasional yang ada saat ini; Komunikasi UE (dilengkapi dengan pedoman etika terpisah) dan Prinsip OECD tentang AI.

Kerangka kerja yang ada saat ini menjawab permasalahan etika utama dan membuat rekomendasi bagi pemerintah untuk mengelolanya, namun terdapat kesenjangan yang nyata. Hal ini mencakup dampak lingkungan, termasuk peningkatan konsumsi energi yang terkait dengan pemrosesan dan pembuatan data AI, dan kesenjangan yang timbul dari distribusi manfaat yang tidak merata dan potensi eksploitasi pekerja. Pilihan kebijakan yang berkaitan dengan dampak lingkungan termasuk memberikan mandat yang lebih kuat untuk keberlanjutan dan tanggung jawab ekologi; mewajibkan pemantauan penggunaan energi, dan publikasi jejak karbon; dan kemungkinan kebijakan yang mengarahkan inovasi teknologi ke arah prioritas lingkungan yang mendesak.

Dalam kasus ketidaksetaraan, pilihan yang bisa diambil adalah dengan menyatakan AI sebagai barang milik publik, bukan milik pribadi. Hal ini memerlukan perubahan norma budaya dan strategi baru untuk membantu menavigasi transisi menuju ekonomi berbasis AI. Menetapkan standar minimum untuk pelaporan tanggung jawab sosial perusahaan akan mendorong perusahaan-perusahaan transnasional yang lebih besar untuk menunjukkan dengan jelas bagaimana mereka berbagi manfaat AI. Kebijakan ekonomi mungkin diperlukan untuk mendukung pekerja yang kehilangan pekerjaan akibat AI; kebijakan-kebijakan tersebut

harus berfokus pada mereka yang paling berisiko tertinggal dan mungkin mencakup kebijakan-kebijakan yang dirancang untuk menciptakan struktur pendukung bagi pekerja tidak tetap. Penting bagi pengulangan kerangka kerja ini di masa depan untuk mengatasi kesenjangan ini dan kesenjangan lainnya agar dapat mempersiapkan diri secara memadai terhadap implikasi penuh dari masa depan AI. Selain itu, untuk memperjelas masalah tanggung jawab yang berkaitan dengan perilaku AI, kerangka moral dan legislatif perlu diperbarui seiring dengan perkembangan teknologi itu sendiri.

Pemerintah juga perlu mengembangkan bentuk-bentuk penilaian teknologi yang baru dan terkini – yang memungkinkan mereka untuk memahami teknologi tersebut secara mendalam sementara teknologi tersebut masih dapat dibentuk, seperti Unit Penilaian Teknologi Kantor Akuntabilitas di AS atau platform European Foresight (<http://www.foresight-platform.eu/>). Bentuk-bentuk baru dari penilaian teknologi TA harus mencakup proses Penilaian Risiko Etis, seperti yang ditetapkan dalam BS8611, dan bentuk-bentuk evaluasi etika lainnya yang saat ini sedang dirancang dalam seri standar etika IEEE Standards Association P7000; P7001 misalnya menetapkan metode untuk mengukur transparansi AI.

Terdapat kebutuhan yang jelas untuk mengembangkan undang-undang dan kebijakan yang layak dan dapat diterapkan untuk menghadapi berbagai tantangan terkait AI, termasuk potensi pelanggaran prinsip-prinsip etika mendasar. Para pembuat kebijakan berada dalam posisi yang berharga karena mampu mengembangkan kebijakan yang secara aktif membentuk pengembangan AI dan ketika pendekatan berbasis data dan pembelajaran mesin mulai mengambil peran yang semakin besar dalam masyarakat, strategi yang bijaksana dan terperinci tentang cara berbagi manfaat dan mencapai tujuan hasil terbaik, sekaligus mengelola risiko secara efektif, sangatlah penting.

Selain kemajuan kebijakan yang sangat menggembirakan sejauh ini, buku ini juga mengungkapkan adanya kesenjangan yang memprihatinkan antar wilayah. Keberhasilan pengembangan AI memerlukan investasi yang besar, dan ketika otomatisasi dan mesin cerdas mulai mendorong proses pemerintahan, terdapat risiko nyata bahwa negara-negara berpendapatan rendah – yaitu negara-negara Selatan – akan tertinggal.

Untuk kembali ke tema utama kita, pertimbangan etis juga harus menjadi komponen penting dalam setiap kebijakan mengenai AI. Hal ini menunjukkan bahwa negara dengan peringkat tertinggi dalam Indeks Kesiapan AI Pemerintah tahun 2019 ini sangat memprioritaskan etika dalam Strategi AI nasional mereka. Singapura adalah salah satu dari sedikit negara yang membentuk Dewan Etika AI dan telah memasukkan berbagai pertimbangan etis ke dalam kebijakannya. Mengatasi permasalahan etika juga merupakan poin penting pertama dalam kerangka Forum Ekonomi Dunia untuk mengembangkan strategi AI nasional. Jadi, terlepas dari potensi kewajiban moral, tampaknya tidak mungkin pemerintah yang tidak menganggap serius etika akan berhasil dalam forum global yang kompetitif.

7.1 MEMBANGUN ROBOT ETIS

Di masa depan, kemungkinan besar mesin cerdas harus mengambil keputusan yang berdampak pada keselamatan manusia, psikologi, dan masyarakat. Misalnya, robot SAR harus bisa 'memilih' korban yang akan ditolong terlebih dahulu setelah gempa bumi; mobil otonom harus bisa 'memilih' apa atau siapa yang akan ditabrak ketika kecelakaan tidak dapat dihindari; robot perawatan di rumah harus mampu menyeimbangkan privasi penggunanya dan kebutuhan perawatan mereka. Namun bagaimana kita mengintegrasikan nilai-nilai sosial, hukum, dan moral ke dalam perkembangan teknologi AI? Bagaimana kita bisa memprogram mesin untuk membuat keputusan etis - sejauh mana pertimbangan etis bisa ditulis dalam bahasa yang dimengerti komputer?

Merancang metode untuk mengintegrasikan etika ke dalam desain AI telah menjadi fokus utama penelitian selama beberapa tahun terakhir. Pendekatan terhadap pengambilan keputusan moral umumnya terbagi dalam dua kubu, pendekatan 'top-down' dan 'bottom-up' (Allen et al., 2005). Pendekatan top-down melibatkan pemrograman aturan dan keputusan moral secara eksplisit ke dalam agen buatan, seperti 'jangan membunuh'. Sebaliknya, pendekatan bottom-up melibatkan pengembangan sistem yang secara implisit dapat belajar membedakan antara perilaku bermoral dan tidak bermoral.

Pendekatan dari bawah ke atas

Pendekatan bottom-up memungkinkan robot mempelajari etika secara mandiri, misalnya dengan menggunakan pembelajaran mesin. Santos-Lang (2002) menunjukkan bahwa ini adalah pendekatan yang lebih baik, karena manusia sendiri terus belajar untuk bersikap etis. Keuntungannya adalah sebagian besar pekerjaan dilakukan oleh mesin itu sendiri, sehingga robot tidak terpengaruh oleh bias desainer. Namun kelemahannya adalah mesin dapat menunjukkan perilaku yang tidak diinginkan dan menyimpang dari tujuan yang diinginkan. Misalnya, jika sebuah robot diprogram untuk 'memilih perilaku yang paling mengarah pada kebahagiaan', mesin tersebut mungkin menemukan bahwa robot tersebut dapat lebih cepat mencapai tujuannya untuk memaksimalkan kebahagiaan dengan terlebih dahulu meningkatkan efisiensi pembelajarannya sendiri, dan 'sementara' beralih dari perilaku tersebut. tujuan awal. Karena perubahan tersebut, mesin bahkan mungkin memilih perilaku yang mengurangi kebahagiaan untuk sementara, jika perilaku tersebut pada akhirnya membantunya mencapai tujuannya. Misalnya sebuah mesin dapat mencoba merampok, berbohong, dan membunuh, agar nantinya menjadi teladan etis.

Pendekatan dari atas ke bawah

Pendekatan top-down melibatkan agen pemrograman dengan aturan ketat yang harus mereka ikuti dalam situasi tertentu. Misalnya, pada mobil self-driving, kendaraan dapat diprogram dengan perintah 'jangan mengemudi lebih cepat dari 130 km/jam di jalan raya'. Masalah dengan pendekatan top-down adalah bahwa pendekatan tersebut memerlukan penentuan teori moral mana yang sebaiknya diterapkan. Contoh teori moral yang bersaing termasuk etika utilitarian, etika deontologis dan pandangan komensal serta Doktrin Efek Ganda.

Utilitarianisme didasarkan pada gagasan bahwa moralitas suatu tindakan harus dinilai berdasarkan konsekuensinya. Dengan kata lain, suatu tindakan dinilai benar secara moral jika konsekuensinya membawa manfaat yang lebih besar. Berbagai teori utilitarian berbeda-beda dalam hal definisi 'kebaikan' yang ingin dimaksimalkan. Misalnya, Bentham (1789) mengusulkan bahwa agen moral harus bertujuan untuk memaksimalkan kebahagiaan total suatu populasi.

Sebaliknya, etika deontologis (berbasis tugas) berpendapat bahwa tindakan harus dinilai bukan berdasarkan hasil yang diharapkan, tetapi berdasarkan apa yang dilakukan orang. Etika berbasis kewajiban mengajarkan bahwa tindakan itu benar atau salah, terlepas dari akibat baik atau buruk yang mungkin dihasilkan. Dalam bentuk etika ini Anda tidak dapat membenarkan suatu tindakan dengan menunjukkan bahwa tindakan tersebut menghasilkan konsekuensi yang baik.

Terkadang teori moral yang berbeda bisa saling bertentangan. Misalnya saja dalam kasus mobil self-driving yang harus memutuskan apakah akan berbelok untuk menghindari hewan yang dilewatinya. Dalam pandangan komensal, nyawa hewan diperlakukan seolah-olah hanya bernilai sebagian kecil dari nilai nyawa manusia, sehingga mobil akan berbelok jika kecil kemungkinannya menyebabkan kerugian pada manusia (Bogosian, 2017). Namun, pandangan incommensal tidak akan membiarkan manusia ditempatkan pada risiko kematian tambahan demi menyelamatkan hewan. Karena pandangan ini pada dasarnya menolak asumsi-asumsi pihak lain, dan berpendapat bahwa trade-off tidak diperbolehkan, maka tidak ada 'titik tengah' yang jelas di mana prinsip-prinsip yang bersaing dapat bertemu.

Bonnemains dkk. (2018) menggambarkan sebuah dilema ketika sebuah drone yang diprogram untuk menghancurkan sebuah rudal yang mengancam pabrik amunisi sekutu tiba-tiba diperingatkan akan adanya ancaman kedua - sebuah rudal yang mengarah ke sejumlah warga sipil. Drone harus memutuskan apakah akan melanjutkan misi aslinya, atau mengeluarkan rudal baru untuk menyelamatkan warga sipil. Hasil keputusannya berbeda-beda tergantung pada apakah Anda menggunakan utilitarianisme, etika deontologis, dan Doktrin Efek Ganda - sebuah teori yang menyatakan bahwa jika melakukan sesuatu yang baik secara moral mempunyai efek samping yang buruk secara moral, maka secara etis boleh saja melakukannya asalkan ada sisi buruknya.

Beberapa teori tidak mampu menyelesaikan masalah. Misalnya, dari sudut pandang deontologis, kedua keputusan tersebut valid, karena keduanya muncul dari niat baik. Dalam kasus etika utilitarian, tanpa informasi mengenai jumlah warga sipil yang berada dalam bahaya, atau nilai dari pabrik strategis tersebut, akan sulit bagi drone untuk mengambil keputusan. Untuk mengikuti doktrin utilitarian dan membuat keputusan yang memaksimalkan 'hasil yang baik', agen artifisial perlu mengidentifikasi semua kemungkinan konsekuensi dari sebuah keputusan, dari sudut pandang semua pihak, sebelum membuat keputusan tentang konsekuensi mana yang lebih disukai. Hal ini tidak mungkin terjadi di lapangan. Persoalan lainnya adalah bagaimana seharusnya drone memutuskan hasil mana yang diinginkan padahal penilaiannya subjektif? Apa yang Baik? Memberikan jawaban terhadap persoalan

filosofis yang luas ini hampir tidak mungkin dilakukan oleh agen otonom, atau orang yang memprogramnya.

Berdasarkan Doktrin Efek Ganda, drone tidak diperbolehkan untuk mencegah rudal dan menyelamatkan warga sipil, karena efek samping yang buruk (hancurnya drone itu sendiri) akan menjadi sarana untuk memastikan efek yang baik (menyelamatkan manusia). Oleh karena itu, mereka akan terus mengejar tujuan awalnya dan menghancurkan peluncurnya, sehingga membiarkan warga sipil mati.

Jika para filsuf tidak dapat menyepakati manfaat dari berbagai teori, perusahaan, pemerintah, dan peneliti akan semakin sulit memutuskan sistem mana yang akan digunakan untuk agen buatan (Bogosion, 2017). Penilaian moral pribadi seseorang juga bisa sangat berbeda ketika dihadapkan pada dilema moral (Greene et al., 2001), khususnya ketika mereka mempertimbangkan isu-isu yang dipolitisasi seperti keadilan rasial dan kesenjangan ekonomi. Bogosian (2017) berpendapat bahwa kita seharusnya merancang mesin yang secara fundamental tidak memiliki ketidakpastian mengenai moralitas.

DAFTAR PUSTAKA

- Abas, A. (2017). Najib unveils Malaysia's digital 'to-do list' to propel digital initiatives implementation. [online] Nst.com.my. Available from: <https://www.nst.com.my/news/nation/2017/10/292784/najib-unveils-malaysias-digital-do-list-propel-digital-initiatives> [Accessed 8 May 2019].
- Access Partnership and the University of Pretoria (2018). Artificial Intelligence for Africa: An Opportunity for Growth, Development and Democratisation.
- Acemoglu, D. and Restrepo, P. (2018) Low-skill and high-skill automation. *Journal of Human Capital*, 2018, vol. 12, no. 2.
- Agency for Digital Italy (2019). Artificial Intelligence task force. [online] IA-Gov. Available from: <https://ia.italia.it/en/> [Accessed 10 May 2019].
- AI For Humanity (2018). AI for humanity: French Strategy for Artificial Intelligence [online] Available from: <https://www.aiforhumanity.fr/en/> [Accessed 10 May 2019].
- AI Forum New Zealand (2018). Artificial Intelligence: Shaping a Future New Zealand.
- AI Now Insitute, (2018). AI Now Report. AI Now Institute, New York University.
- AI Singapore. (2018). AI Singapore. [online] Available from: <https://www.aisingapore.org> [Accessed 26 Apr. 2019].
- AI Taiwan. (2019). AI Taiwan. [online] Available from: <https://ai.taiwan.gov.tw> [Accessed 28 Apr. 2019].
- AI4All (2019). What we do [online] Available from: <http://ai-4-all.org> [Accessed 11/03/2019].
- Allen, C., Smit, I., and Wallach, W. (2005). Artificial morality: Top-down, bottom-up, and hybrid approaches. *Ethics and Information Technology*. doi:10.1007/s10676-006-0004-4.
- Allen, G., and Chan, T., (2017). Artificial Intelligence and National Security. Available from: <https://www.belfercenter.org/sites/default/files/files/publication/AI%20NatSec%20-%20final.pdf>
- Amoroso, D., and Tamburrini, G. (2018). The Ethical and Legal Case Against Autonomy in Weapons Systems. *Global Jurist* 18 (1), DOI: 10.1515/gj-2017-0012.
- Anderson, J. M., Heaton, P. and = Carroll, S. J. (2010). The U.S. Experience with No-Fault Automobile Insurance: A Retrospective. Santa Monica, CA: RAND Corporation. Available from: <https://www.rand.org/pubs/monographs/MG860.html>.

- ANPR (2018). National AI Strategy: Unlocking Tunisia's capabilities potential [online] Available from: <http://www.anpr.tn/national-ai-strategy-unlocking-tunisia-capabilities-potential/>. [Accessed 6 May 2019].
- Apps, P. (2019). Commentary: Are China, Russia winning the AI arms race? [online] U.S. Available from: <https://www.reuters.com/article/us-apps-ai-commentary/commentary-are-china-russia-winning-the-ai-arms-race-idUSKCN1P91NM>.
- Arnold, T., and Scheutz, M. (2018). The 'big red button' is too late: an alternative model for the ethical evaluation of AI systems. *Ethics and Information Technology*. 20 (1), 59–69.
- Asaro, P. (2012). On Banning Autonomous Weapon Systems: Human Rights, Automation, and the Dehumanization of Lethal Decision-Making. *International Review of the Red Cross*. 94 (886), 687-703.
- Atabekov, A. and Yastrebov, O. (2018) Legal status of Artificial Intelligence: Legislation on the move. *European Research Studies Journal* Volume XXI, Issue 4, 2018 pp. 773 - 782
- Australian Government (2017). The Digital Economy: Opening Up The Conversation. Department of Industry, Innovation and Science. Available from: <https://www.archive.industry.gov.au/innovation/Digital-Economy/Documents/Digital-Economy-Strategy-Consultation-Paper.pdf>
- Australian Government (2018). Australia's Tech Future. Department of Industry, Innovation and Science. Available from: <https://www.industry.gov.au/sites/default/files/2018-12/australias-tech-future.pdf>
- Austrian Council on Robotics and Artificial Intelligence (2019). Österreichischer Rat für Robotik und Künstliche Intelligenz. [online] Available from: <https://www.acrai.at/>
- Austrian Council on Robotics and Artificial Intelligence (2018). Die Zukunft Österreichs mit Robotik und Künstlicher Intelligenz positiv gestalten. White Paper des Österreichischen Rats für Robotik und Künstliche Intelligenz.
- Autor, D. H. (2015). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives*. 29(3), 3–30.
- Bandyopadhyay, A., and Hazra, A. (2017). A comparative study of classifier performance on spatial and temporal features of handwritten behavioural data.
- Baron, E. (2017). Robot surgery firm from Sunnyvale facing lawsuits, reports of death and injury. *Mercury News*. Available from:

<https://www.mercurynews.com/2017/10/22/robot-surgery-firm-from-sunnyvale-facing-lawsuits-reports-of-death-and-injury/>

- Bartlett, J. (2018) How AI could kill off democracy. *New Statesman*. Available from: <https://www.newstatesman.com/science-tech/technology/2018/08/how-ai-could-kill-democracy-0>
- Basu, S. Das, P. Horain, and S. Bhattacharya (eds.). (2016) *Intelligent Human Computer Interaction: 8th International Conference, IHCI 2016, Pilani, IndiaCham: Springer International Publishing*, 111–121.
- BBC News (2017). Singapore to use driverless buses 'from 2022'. BBC. Available from: <https://www.bbc.co.uk/news/business-42090987>
- BBC News. (2018). Addison Lee plans self-driving taxis by 2021. BBC. Available from: <https://www.bbc.co.uk/news/business-45935000>
- BBC News. (2019a). Autonomous shuttle to be tested in New York City. BBC. Available from: <https://www.bbc.co.uk/news/technology-47668886>
- BBC News. (2019b). Uber 'not criminally liable for self-driving death. BBC. Available from: <https://www.bbc.co.uk/news/technology-47468391>
- Beane, M. (2018). Young doctors struggle to learn robotic surgery – so they are practicing in the shadows. *The Conversation*. Available from: <https://theconversation.com/young-doctors-struggle-to-learn-robotic-surgery-so-they-are-practicing-in-the-shadows-89646>
- Bentham, J. (1789). *A Fragment of Government and an Introduction to the Principles of Morals and Legislation*, London.
- Berger, S. (2019). Vaginal mesh has caused health problems in many women, even as some surgeons vouch for its safety and efficacy. *The Washington Post*.
- Bershidsky, L (2017). Elon Musk warns battle for AI supremacy will spark Third World War. *The Independent*. [online] Available from: <https://www.independent.co.uk/life-style/gadgets-and-tech/news/elon-musk-ai-artificial-intelligence-world-war-three-russia-china-robots-cyber-warfare-replicants-a7931981.html>
- Biavaschi, C., Eichhorst, W., Giuliotti, C., Kendzia, M., Muravyev, A., Pieters, J., Rodriguez-Planas, N., Schmidl, R., and Zimmermann, K. (2013). *Youth Unemployment and Vocational Training*. World Development Report. World Bank.
- Bilge, L., Strufe, T., Balzarotti, D., Kirda, K., and Antipolis, S. (2009). All your contacts are belong to us: Automated identity theft attacks on social networks, In *WWW '09: Proceedings of*

the 18th international conference on World Wide Web, WWW '09, April 20-24, 2009, Madrid, Spain. New York, NY, USA. pp. 551– 560.

- Bogosian, K. (2017) Implementation of Moral Uncertainty in Intelligent Machines. *Minds & Machines* 27 (591).
- Bonnemains, V., Saurel, C. & Tessier, C. (2018) Embedded ethics: some technical and ethical challenges. *Ethics and Information Technology*. 20 (41).
- Borenstein, J. and Arkin, R.C. (2019) *Robots, Ethics, and Intimacy: The Need for Scientific Research*.
- Bradshaw, S., and Howard, P. (2017) *Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation*. In Woolley, S. and Howard, P. N. (Eds.)
- Bradshaw, T. (2018) Uber halts self-driving car tests after pedestrian is killed. *Financial Times*. 19 March, 2018.
- Brynjolfsson, E., and McAfee, A (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York, W. W. Norton & Company..
- Bryson, J.. (2018) *Patience is not a virtue: the design of intelligent systems and systems of ethics*. *Ethics and Information Technology*, 20 (1). 15–26
- Bryson, J. J. (2019). *The Past Decade and Future of AI's Impact on Society*. In Baddeley, M., Castells, M., Guiora, A., Chau, N., Eichengreen, B., López, R., Kanbur, R. and Burkett, V. (2019) *Towards a New Enlightenment? A Transcendent Decade*. Madrid, Turner.
- Burgmann, T. (2016). *There's a cure for that: Canadian doctor pushes for more wearable technology*. *Global News Canada*.
- Cadwalladr, C. (2017a). *Revealed: How US billionaire helped to back Brexit*. *The Guardian*.
- Cadwalladr, C. (2017b). *Robert Mercer: The big data billionaire waging war on mainstream media*. *The Guardian*.
- Calder, S. (2018). *Driverless buses and taxis to be launched in Britain by 2021*. *The Independent*. Available from: <https://www.independent.co.uk/travel/news-and-advice/self-driving-buses-driverless-cars-edinburgh-fife-forth-bridge-london-greenwich-a8647926.html>
- Canadian Institute For Advanced Research (2017) *Pan-Canadian Artificial Intelligence Strategy*.
- Canadian Institute For Advanced Research (2019). *AI & Society Workshops: Call Two*.

- Cannon, J. (2018). Starsky Robotics completes first known fully autonomous run without a driver in cab. Commercial Carrier Journal.
- Caplan, R., Donovan, J., Hanson, L. and Matthews, J. (2018). Algorithmic Accountability: A Primer. New York, Data & Society.
- Cassim, N. (2019). Dhammika makes strong case for national strategy for AI. [online] Financial Times.
- Castellanos, S. (2018). Estonia's CIO Tackles AI Strategy For Government. [online] WSJ.
- CDEI (2019). 'The Centre for Data Ethics and Innovation (CDEI) 2019/ 20 Work Programme.
- Chantler, A., & Broadhurst, R. (2006). Social engineering and crime prevention in cyberspace. Technical report., Justice, Queensland University of Technology.
- Chen, A. (2017) 'The Human Toll of Protecting the Internet from the Worst of Humanity'. The New Yorker.
- Chesney, R., & Citron, D. (2018). Deep fakes: A looming crisis for national security, democracy and privacy? Lawfare.
- Christakis, N.A & Shirado, H. (2017) Locally Noisy Autonomous Agents Improve Global Human Coordination in Network Experiments. Nature. 545(7654), 370–374.
- Christakis, N.A (2019) How AI Will Rewire Us. The Atlantic Magazine.
- Citron, D. K., & Pasquale, F. A. (2014). The scored society: due process for automated predictions. Washington Law Review, 89, 1–33.
- CNN. (2018). Self-driving electric bus propels Swiss town into the future. CNN.
- COMEST (2017). Report of COMEST on Robotics Ethics. UNESCO.
- Conn, A. (2018) AI Should Provide a Shared Benefit for as Many People as Possible, Future of Life Institute.
- Consultative Committee of the Convention for the Protection of Individuals with regard to the Processing of Personal Data (2019) Guidelines on Artificial Intelligence and Data Protection.
- Corbe-Davies, S., Pierson, S., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic decision making and the cost of fairness. In Proceedings of KDD '17, Halifax, NS, Canada, August 13-17, 2017, 10 pages. DOI: 10.1145/3097983.3098095
- Council of Europe (2019a). Ad Hoc Committee on Artificial Intelligence – CAHAI.

- Cummings M. (2004). Automation bias in intelligent time critical decision support systems. In AIAA: 1st Intelligent Systems Technical Conference. AIAA 2004, 20-22 September 2004, Chicago, Illinois. pp. 6313.
- Curtis, J. (2016). Shocking dashcam footage shows Tesla 'Autopilot' crash which killed Chinese driver when futuristic electric car smashed into parked lorry. Daily Mail.
- Danaher, J. (2017). Robotic rape and robotic child sexual abuse: Should they be criminalised? *Criminal Law and Philosophy*, 11(1), 71–95.
- Dastin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters.
- Datta, A., Tschantz and M.C., Datta, A. (2015). Automated Experiments on Ad Privacy Settings – A Tale of Opacity, Choice, and Discrimination. *Proceedings on Privacy Enhancing Technologies*. 1, 92–112, DOI: 10.1515/popets-2015-0007
- Dawson, D., Schleiger, E., Horton, J., McLaughlin, J., Robinson, C., Quezada, G., Scowcroft, J., and Hajkovicz, S. (2019). Artificial Intelligence: Australia's Ethics Framework.
- De Angeli, A. (2009). Ethical implications of verbal disinhibition with conversational agents. *Psychology Journal*, 7(1), 49–57.
- De Angeli, A., & Brahnem, S. (2008). I hate you! Disinhibition with virtual partners. *Interacting with Computers*. 20(3), 302–310
- Delvaux, M. (2017). 'With recommendations to the Commission on Civil Law Rules on Robotics'
- Die Bundesregierung (2018) Strategie Künstliche Intelligenz der Bundesregierung.
- Digital Poland Foundation (2019). Map of the Polish AI. Digital Poland Foundation..
- Dignum, V. (2018). Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20: 1.
- Duckworth, P., Graham, L., Osborne and M. AI (2019). Inferring Work Task Automatability from AI Expert Evidence. *AAAI / ACM Conference on Artificial Intelligence, Ethics and Society*. University of Oxford.
- Ethics Commission (2017). Ethics's Commission's complete report on automated and connected driving. Federal Ministry of Transport and Infrastructure.
- Etzioni, A. and Etzioni, O. (2016). AI assisted ethics. *Ethics and Information Technology*, 18(2), 149–156 European Commission (2012) Special Eurobarometer 382: Public Attitudes towards Robots.

- European Commission (2018a). Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions on Artificial Intelligence for Europe
- European Commission (2018b). Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions - Coordinated Plan on Artificial Intelligence (COM(2018) 795 final).
- European Commission (2018c). High-level expert group on artificial intelligence: Draft ethics guidelines for trustworthy AI. Brussels.
- European Commission 2015/2103(INL).
- European Parliament, 2017. EP Resolution with Recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)).
- European Parliament, Council and Commission, (2012). Charter of Fundamental Rights of the European Union. Official Journal of the European Union
- Ezrachi, A., & Stucke, M. E. (2016). Two artificial neural networks meet in an online hub and change the future (of competition, market dynamics and society). Oxford Legal Studies Research Paper, No. 24/2017; University of Tennessee Legal Studies Research Paper, No. 323.
- Farmer, J. D., & Skouras, S. (2013). An ecological perspective on the future of computer trading. *Quantitative Finance*. 13(3), 325–346
- Felton, R. (2017). Limits of Tesla's Autopilot and driver error cited in fatal Model S crash. *Jalopnik*.
- Felton, R. (2018). Two years on, a father is still fighting Tesla over autopilot and his son's fatal crash. *Jalopnik*.
- Floridi, L. (2016). Tolerant paternalism: Pro-ethical design as a resolution of the dilemma of toleration. *Science and Engineering Ethics*. 22(6), 1669–1688.
- Floridi, L., & Taddeo, M. (2016). What is data ethics? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083).
- Foot, P. (1967). The problem of abortion and the doctrine of the double effect. *Oxford review*. 5. Oxford, Oxford University Press.
- Ford, M. (2009) *The Lights in the Tunnel: Automation, Accelerating Technology, and the Economy of the Future*.

- Foundation for Law & International Affairs (2017) China's New Generation of Artificial Intelligence Development Plan. FLIA.
- Frey, C. B. and Osborne, M. A. (2013). The Future of Employment: How Susceptible Are Jobs to Computerisation? Oxford Martin Programme on the Impacts of Future Technology.
- Furman, J & Seamans, R. (2018). AI and the Economy. NBER working paper no.24689
- Future of Life Institute (2019). National and International AI Strategies. Future of Life Institute.
- G20 (2019) G20 Ministerial Statement on Trade and Digital Economy: Annex.
- G7 Canadian Presidency (2018). Charlevoix Common Vision for the Future of Artificial Intelligence.
- Gagan, O. (2018) Here's how AI fits into the future of energy, World Economic Forum, 25 May 2018.
- Garfinkel, S. (2017). Hackers are the real obstacle for self-driving vehicles. MIT Technology Review.
- Gibbs, S. (2017). Tesla Model S cleared by safety regulator after fatal Autopilot crash. The Guardian.
- Gillespie T. (2014). The relevance of algorithms. In Gillespie, T., Boczkowski, P. J., Foot, K. A. (eds.)
- Gogarty, B., & Hagger, M. (2008). The laws of man over vehicles unmanned: The legal response to robotic revolution on sea, land and air. *Journal of Law, Information and Science*, 19, 73–145.
- Goldhill, O. (2016). Can we trust robots to make moral decisions? Quartz.
- GOV.UK (2019). The UK's Industrial Strategy. GOV.UK.
- GOV.UK. (2018b). Centre for Data Ethics and Innovation (CDEI).
- Government Offices of Sweden (2018). National approach to artificial intelligence. Ministry of Enterprise and Innovation.
- Graetz, G. and Michaels, G. (2015). Robots at Work. Centre for Economic Performance Discussion Paper No. 1335.
- Gray, M. L. and Suri, S. (2019). *Ghost Work*, Houghton Mifflin Harcourt.

- Greene, J. D., Sommerville, R. B., Nystrom, L., Darley, J., and Cohen, J. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105–2108. doi:10.1126/science.1062872.
- Guiltinan, J. (2009). Creative destruction and destructive creations: Environmental ethics and planned obsolescence. *Journal of Business Ethics*. 89 (1). pp.1928.
- Gurney, J. K., (2013). Sue My Car, Not Me: Products Liability and Accidents Involving Autonomous Vehicles. unpublished manuscript
- Hadfield-Menell, D., Dragan, A., Abbeel, P., and Russell, S. (2016). The off-switch game. In: IJCAI-ECAI- 2018: International Joint Conference on Artificial Intelligence. IJCAI-ECAI-2018, 13-19 July 2018, Stockholm, Sweden.
- Hallaq, B., Somer, T., Osula, A., Ngo, K., & Mitchener-Nissen, T. (2017). Artificial intelligence within the military domain and cyber warfare. In: 16th European Conference on Cyber Warfare and Security (ECCWS 2017), 29-30 June 2017, Dublin, Ireland. Published in: Proceedings of 16th European Conference on Cyber Warfare and Security.
- Hancock, J. T., Curry, L. E., Goorha, S., and Woodworth, M. (2007). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes*. 45(1): 1–23.
- Harambam, J., Helberger, N., and Van Hoboken, J. (2018). Democratizing algorithmic news recommenders: how to materialize voice in a technologically saturated media ecosystem. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376 (2133).
- Hardt, M. (2014). How Big Data is Unfair. Medium. [online] Available from [accessed 9 Apr. 2019]
- Hart, R, D. (2018). Who's to blame when a machine botches your surgery? Quartz.
- Hawkins, A. J. (2019). California's self-driving car reports are imperfect, but they're better than nothing. The Verge.
- Hawksworth, J. and Fertig, Y. (2018) What will be the net impact of AI and related technologies on jobs in the UK? PwC UK Economic Outlook, July 2018.
- Hern, A. (2016). 'Partnership on AI' formed by Google, Facebook, Amazon, IBM and Microsoft. The Guardian.
- Hess, A,. (2016). On Twitter, a Battle Among Political Bots. The New York Times. Available from: <https://www.nytimes.com/2016/12/14/arts/on-twitter-a-battle-among-political-bots.html>

https://ai4d.ai/wp-content/uploads/2019/05/ai-gov-readiness-report_v08.pdf

Human Rights Watch. (2018). 'Eradicating ideological viruses': China's campaign of repression against Xinjiang's Muslims. Technical report, Human Rights Watch.

IEEE (2019). Homepage [online] Available from: <https://www.ieee.org>

Iglinski, H., Babiak, M. (2017). Analysis of the Potential of Autonomous Vehicles in Reducing the Emissions of Greenhouse Gases in Road Transport. *Procedia Eng.*192, 353–358.

International Telecommunication Union (2018). AI for Good Global Summit 2018 [online] Available from: <https://www.itu.int/en/ITU-T/AI/2018/Pages/default.aspx> [Accessed 14 May 2019].

Isaac, M. (2016). Self-driving truck's first mission: a 120-mile beer run. *New York Times*.

Israel Innovation Authority (2019). Israel Innovation Authority 2018-19 Report.

Jacobs, S. B. (2017) *The Energy Prosumer*, 43*Ecology L. Q.*519.

Japanese Strategic Council for AI Technology (2017). Artificial Intelligence Technology Strategy. Available from: <https://www.nedo.go.jp/content/100865202.pdf>

Johnson, A., and Axinn, S. (2013). The Morality of Autonomous Robots. *Journal of Military Ethics.* 12 (2), 129-141

Johnston, A. K. (2015). Robotic seals comfort dementia patients but raise ethical concerns. *KALW*.

JSAI (2017). Ethical Guidelines. [online] Available from: <http://ai-elsi.org/wp-content/uploads/2017/05/JSAI-Ethical-Guidelines-1.pdf>

JSAI (2019). Overview: Inaugural Address of President Naohiko Uramoto, Artificial Intelligence expanding its scope and impact in our society. [online] Available from: <https://www.ai-gakkai.or.jp/en/about/about-us/>

Kayali, L. (2019). Next European Commission takes aim at AI. *POLITICO*.

Kenyan Wall Street (2018). Kenya Govt unveils 11 Member Blockchain & AI Taskforce headed by Bitange Ndemo. *Kenyan Wallstreet*.

Khakurel, J., Penzenstadler, B., Porras, J., Knutas, A., and Zhang, W. (2018). The Rise of Artificial Intelligence under the Lens of Sustainability. *Technologies.* 6(4), 100.

Khosravi, B. (2018). Autonomous cars won't work – until we have 5G. *Forbes*

- King, T.C., Aggarwal, N., Taddeo, M. et al. (2019). Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions. *Sci Eng Ethics*. pp.1-32
- Kingston, J. K. C. (2018) Artificial Intelligence and Legal Liability. Available at: <https://arxiv.org/ftp/arxiv/papers/1802/1802.07782.pdf> [Accessed 17/08/19].
- Kitwood, T. (1997). *Dementia Reconsidered: The Person Comes First*. Buckingham, Open University Press.
- Knight, W. (2019). The World Economic Forum wants to develop global rules for AI. [online] MIT Technology Review.
- Kroll, J.A. (2018). The fallacy of inscrutability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 376 (2133).
- Lalji, N. (2015). Can we learn about empathy from torturing robots? This MIT researcher is giving it a try. *YES! Magazine*.
- Lapuschkin, S., Wäldchen, S., Binder, A., Montavon, G., Samek, W. and Müller, K.R. (2019). Unmasking Clever Hans predictors and assessing what machines really learn. *Nature Communications*. 10, (1096)
- LaRosa, E., & Danks, D. (2018). Impacts on Trust of Healthcare AI. In: *AAAI / ACM Conference on Artificial Intelligence, Ethics and Society*. AEIS: 2018, 1-3 February, 2018, New Orleans, USA.
- Larsson, S., Anneroth, M., Felländer, A., Felländer-Tsai, L., Heintz, F., Cedering Ångström, R. (2019). Sustainable AI report. AI Sustainability Centre.
- Lashbrook, A. (2018). AI-driven dermatology could leave dark-skinned patients behind. *The Atlantic*.
- Le Miere, J. (2017). Russia is developing autonomous 'swarms of drones' it calls an inevitable part of future warfare. *Newsweek*.
- Leggett, T. (2018) Who is to blame for 'self-driving car' deaths? *BBC Business News*.
- Leontief, Wassily,. (1983). *National Perspective: The Definition of Problems and Opportunities.. The Long-Term Impact of Technology on Employment and Unemployment*. Washington, DC: The National Academies Press. doi: 10.17226/19470.
- Lerner, S. (2018). NHS might replace nurses with robot medics such as carebots: could this be the future of medicine? *Tech Times*.

- Levin, S. (2018). Video released of Uber self-driving crash that killed woman in Arizona. *The Guardian*.
- Li, H., Milani, S., Krishnamoorthy, V., Lewis, M., & Sycara, K. (2019). Perceptions of Domestic Robots' Normative Behavior Across Cultures. In: *AAAI / ACM Conference on Artificial Intelligence, Ethics and Society*. AEIS: 2019, 27-28 January, 2019, Honolulu, Hawaii, USA.
- Li, S., Williams, J. (2018). Despite what Zuckerberg's testimony may imply, AI Cannot Save Us. *Electronic Frontier Foundation*.
- Lim, D., (2019). Killer Robots and Human Dignity. In: *AAAI / ACM Conference on Artificial Intelligence, Ethics and Society*. AEIS: 2019, 27-28 January, 2019, Honolulu, Hawaii, USA.
- Lin, P. (2014). What if your autonomous car keeps routing you past Krispy Kreme? *The Atlantic*.
- Lin, P., Jenkins, R., & Abney, K. (2017). *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*. Oxford University Press.
- Lin, T. C. W. (2017). The new market manipulation. *Emory Law Journal*, 66, 1253.
- Loh, W. & Loh, J. (2017). Autonomy and responsibility in hybrid systems. In P. Lin, et al. (Eds.), *Robot ethics 2.0*. New York, NY: Oxford University Press: 35–50.
- Lokhorst, G.-J. and van den Hoven, J. (2014) Chapter 9: Responsibility for Military Robots. In *Robot Ethics: The Ethical and Social Implications of Robotics* edited by Lin, Abney and Bekey (10 Jan. 2014, MIT Press).
- Malta AI (2019). *Malta AI: Towards a National AI Strategy* [online] Available at: <https://malta.ai>
- Manikonda, L., Deotale, A., & Kambhampati, S., (2018). What's up with Privacy? User Preferences and Privacy Concerns in Intelligent Personal Assistants. In: *AAAI / ACM Conference on Artificial Intelligence, Ethics and Society*. AEIS: 2018, 1-3 February, 2018, New Orleans, USA.
- Marda, V., (2018). Artificial intelligence policy in India: a framework for engaging the limits of data- driven decision-making. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 376 (2133).
- Marshall, A. and Davies, A. (2018). Lots of lobbies and zero zombies: how self-driving cars will reshape cities. *Wired*.
- Martinho-Truswell, E., Miller, H., Nti Asare, I., Petheram, A., Stirling, R., Gómez Mont, G. and Martinez, C. (2018). *Towards an AI Strategy in Mexico: Harnessing the AI Revolution*.

- Mattheij, J. (2016) 'Another Way Of Looking At Lee Sedol vs AlphaGo'. Jacques Mattheij: Technology, Coding and Business. Blog. 17th March 2016.
- Matthias, A. (2004) The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, Sept 2004, Vol. 6, Issue 3, pp.175-183.
- Mazzucato, M. (2018) *Mission-Oriented Research & Innovation in the European Union*. European Commission: Luxembourg.
- Mbadiwe, T. (2017). The potential pitfalls of machine learning algorithms in medicine. *Pulmonology Advisor*.
- McAllister, A. (2017). Stranger than science fiction: The rise of AI interrogation in the dawn of autonomous robots and the need for an additional protocol to the UN convention against torture. *Minnesota Law Review*. 101, 2527–2573.
- McCarty, N. M., Poole, K. T., and Rosenthal, H. (2016). *Polarized America: The Dance Of Ideology And Unequal Riches*. Cambridge, MA: MIT Press, 2nd edition.
- Meisner, E. M. (2009). *Learning controllers for human–robot interaction*. PhD thesis. Rensselaer Polytechnic Institute.
- México Digital (2018). *Estrategia de Inteligencia Artificial MX 2018*. gob.mx.
- Millar, J. (2016). An Ethics Evaluation Tool for Automating Ethical Decision-Making in Robots and Self- Driving Cars. *30(8)*, 787-809.
- Min, W. (2018) *Smart Policies for Harnessing AI*, OECD-Forum, 17 Sept 2018
- Ministry of Economic Affairs and Employment of Finland (2017). *Finland's Age of Artificial Intelligence*.
- Ministry of Economic Affairs and Employment of Finland (2018a). *Artificial intelligence programme*. [online] Available from: <https://tem.fi/en/artificial-intelligence-programme>
- Ministry of Economic Affairs and Employment of Finland (2018b). *Work in the Age of Artificial Intelligence*.
- Mizoguchi, R. (2004). The JSAI and AI activity in Japan. *IEEE Intelligent Systems* 19 (2).
- Moon, M., (2017). Judge allows pacemaker data to be used in arson trial. *Engadget*.
- National Science & Technology Council (2019) *The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update*.

- Nemitz, P,. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 376 (2133).
- Nevejans, N. et al. (2018). Open letter to the European Commission on Artificial Intelligence and Robotics.
- Newman, E. J., Sanson, M., Miller, E. K., Quigley-McBride, A., Foster, J. L., Bernstein, D. M., and Garry, M. (2014). People with easier to pronounce names promote truthiness of claims. *PLoS ONE*.9(2).
- NHS Digital. (2019). Widening Digital Participation. NHS Digital.
- NHS' Topol Review. (2019). Preparing the healthcare workforce to deliver the digital future. Available from: <https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf>
- NITI Aayog (2018). National Strategy for Artificial Intelligence #AIFORALL.
- Nordic cooperation (2018). AI in the Nordic-Baltic region. [online] Available from: <https://www.norden.org/en/declaration/ai-nordic-baltic-region>
- Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C. and Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences of the United States of America*, 115, E5716– E5725.
- NTSB (2018) Preliminary Report Released for Crash Involving Pedestrian, Uber Technologies, Inc., Test Vehicle. National Transport Safety Board News Release. May 24, 2018. Available at: <https://www.nts.gov/news/press-releases/Pages/NR20180524.aspx>
- O'Carroll, T. (2017). Mexico's misinformation wars. Medium.
- O'Connor, T. (2017). Russia is building a missile that can makes its own decisions. Newsweek.
- O'Donoghue, J. (2010). E-waste is a growing issue for states. Deseret News.
- OECD (2013) Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data [OECD/LEGAL/0188]
- OECD (n.d.) OECD initiatives on AI [online] Available at: <http://www.oecd.org/going-digital/ai/>
- O'Kane, S (2018). Tesla defends Autopilot after fatal Model S crash. The Verge.
- O'Neil, C,. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. USA: Crown Publishers.

- O'Neill, S. (2018). As insurers offer discounts for fitness trackers, wearers should step with caution. National Public Radio.
- Ori. (2014b). My (autonomous) car, my safety: Results from our reader poll. Robohub.org.
- Ori.(2014a). If Death by Autonomous Car is Unavoidable, Who Should Die? Reader Poll Results. Robohub.org.
- Orseau, L. & Armstrong, S. (2016). Safely interruptible agents. In: Uncertainty in artificial intelligence: 32nd Conference (UAI). UAI: 2016, June 25-29, 2016, New York City, NY, USA. AUAI Press 2016
- Ovanessoff, A. and Plastino, E. (2017). How Artificial Intelligence Can Drive South America's Growth. Accenture.
- Oxford Insights (2019) Government Artificial Intelligence Readiness Index. Available from:
- P. and Rajan, U. (2017). Ethical Issues for Autonomous Trading Agents. *Minds & Machines* 27
- Pagallo, U. (2017). Apples, oranges, robots: four misunderstandings in today's debate on the legal status of AI systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 376 (2133).
- Pariser E. (2011). *The filter bubble: what the Internet is hiding from you*. London, UK, Penguin.
- Park, M. (2017). Self-driving bus involved in accident on its first day. CNN Business.
- Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge, MA, Harvard University Press.
- Personal Data Protection Commission Singapore (2019). *A Proposed Model Artificial Intelligence Governance Framework*.
- Pfleger, P. (2018). Transportation workers form coalition to stop driverless buses in Ohio. WOSU Radio.
- Pham, T., Gorodnichenko, Y. and Talavera, O. (2018). Social Media, Sentiment and Public Opinions: Evidence from #Brexit and #USElection. NBER Working Papers w24631. The National Bureau of Economic Research; Cambridge, MA.
- Piesing, M. (2014). Medical robotics: Would you trust a robot with a scalpel? *The Guardian*.
- Plantera, F. (2017). Artificial Intelligence is the next step for e-governance in Estonia, State adviser reveals. *e-Estonia*.

- Polonski, V. (2017). #MacronLeaks changed political campaigning. Why Macron succeeded and Clinton failed. World Economic Forum.
- Press Association (2019). Robots and AI to give doctors more time with patients, says report. The Guardian.
- ProPublica (2016). Machine Bias. ProPublica.
- Rahwan, I. (2018). Society-in-the-loop: programming the algorithmic social contract. Ethics and Information Technology. 20: 5. <https://doi.org/10.1007/s10676-017-9430-8>
- Ramchurn, S. D. et al. (2013) AgentSwitch: Towards Smart Energy Tariff Selection. Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013), Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA.
- Reuters (2019). G7 urges tight regulations for digital currencies, agrees to tax digital giants locally. VentureBeat.
- Riedl, M.O., and Harrison, B. (2017). Enter the matrix: A virtual world approach to safely interruptible autonomous systems. arXiv. preprint arXiv:1703.10284
- Roberts, S. (2016) 'Digital Refuse: Canadian Garbage, Commercial Content Moderation and the Global Circulation of Social Media's Waste'. Media Studies Publications.
- Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., Schultz, J., Hale, T. M., and Stern M.J. (2015) Digital Inequalities and Why They Matter. Information, Communication & Society. 18 (5), 569- 592.
- SAE International. (2018). SAE International releases updated visual chart for its 'levels of driving automation' standard for self-driving vehicles. SAE International.
- Sage, A. (2018). Waymo unveils self-driving taxi service in Arizona for paying customers. Reuters.
- Sample, I. (2017). Computer says no: why making AIs fair, accountable and transparent is crucial. The Guardian.
- Sample, I. (2017). Give robots an 'ethical black box' to track and explain decisions, say scientists. The Guardian.
- Santos-Lang, C. (2002). Ethics for Artificial Intelligences. In Wisconsin State-Wide technology Symposium 'Promise or Peril?'. Reflecting on computer technology: Educational, psychological, and ethical implications. Wisconsin, USA.

- Sarmah, H. (2019). Looking East: How South Korea Is Making A Strategic Move In AI. *Analytics India Magazine*.
- Sathe G. (2018). Cops in India are using artificial intelligence that can identify you in a crowd. *Huffington Post*.
- Sauer, G. (2017). A Murder Case test's Alexa's Devotion to your Privacy. *Wired*.
- Scherer, M. U. (2016) *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 *Harv. J. L. & Tech.* 353 (2015-2016)
- Scheutz, M. (2012). The inherent dangers of unidirectional emotional bonds between humans and social robots. In: Lin, P., Abney, K. and Bekey, G. (eds.). *Robot Ethics: The Ethical and Social Implications of Robotics*, MIT Press, pp.205-221.
- Schmitt, M.N., (2013). *Tallinn manual on the international law applicable to cyber warfare*. Cambridge University Press.
- Schönberger, D. (2019). Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. *International Journal of Law and Information Technology* 27 (2), 171–203. <https://doi.org/10.1093/ijlit/eaz004>
- Selbst, A. D. and Barocas. S. (2018). The intuitive appeal of explainable machines. 87 *Fordham Law Review* 1085 Preprint, available from: <https://ssrn.com/abstract=3126971>
- Selbst, A. D. and Powles, J. (2017) Meaningful information and the right to explanation. *Int. Data Privacy Law* 7, 233–242. (doi:10.1093/idpl/ix022)
- Selinger, E. and Hartzog, W. (2017). Obscurity and privacy. In: Pitt, J. and Shew, A. (eds.). *Spaces for the Future: A Companion to Philosophy of Technology*, New York: Routledge.
- Servoz, M. (2019) *The Future of Work? Work of the Future! On How Artificial Intelligence, Robotics and Automation Are Transforming Jobs and the Economy in Europe*, 10 May 2019.
- Seth, S. (2017). Machine Learning and Artificial Intelligence Interactions with the Right to Privacy. *Economic and Political Weekly*, 52(51), 66–70
- Sharkey, A., and Sharkey, N. (2012). Granny and the robots: ethical issues in robot care for the elderly. *Ethics and Information Technology*. 14 (1): 27-40.
- Sharkey, N., Goodman, M., & Ross, N. (2010). The coming robot crime wave. *IEEE Computer Magazine*. 43(8), 6–8.

- Shepherdson, D. and Somerville, H. (2019) Uber not criminally liable in fatal 2018 Arizona self-driving crash – prosecutors. Reuters News. March 5, 2019.
- Shewan, D. (2017). Robots will destroy our jobs – and we're not ready for it. The Guardian.
- Smith, A., & Anderson, J. (2014). AI, Robotics, and the Future of Jobs. Pew Research Center
- Smith, B. (2018). Facial recognition technology: The need for public regulation and corporate responsibility. Microsoft on the Issues.
- Snaith, E. (2019). Robot rolls into hospital ward and tells 97-year-old man he is dying. The Independent.
- Solon, O. (2018). Who's driving? Autonomous cars may be entering the most dangerous phase. The Guardian.
- Sparrow, R., (2002). The march of the robot dogs. Ethics and Information Technology. 4 (4), 305–318.
- Sparrow, R., and Sparrow, L. (2006). In the hands of machines? The future of aged care. Minds and Machines. 16, 141-161.
- Spatt, C. (2014). Security market manipulation. Annual Review of Financial Economics, 6(1), 405–418.
- Stahl, B.C., & Coeckelbergh, M. (2016). Ethics of healthcare robotics: Towards responsible research and innovation. Robotics and Autonomous Systems. 86, 152-161.
- Stilgoe, J. and Winfield, A. (2018). Self-driving car companies should not be allowed to investigate their own crashes. The Guardian.
- Strubell, E., Ganesh, A. and McCallum, A. (2019) Energy and Policy Considerations for Deep Learning in NLP, arXiv:1906.02243
- Swedish AI Council. (2019). Swedish AI Council. [online] Available from: <https://swedishaicouncil.com> [Accessed 10 May 2019].
- Taddeo, M. (2017). Trusting Digital Technologies Correctly. Minds & Machines. 27 (4), 565.
- Taddeo, M. and Floridi, L. (2018) How AI can be a force for good. Science vol. 361, issue 6404, pp.751- 752. DOI: 10.1126/science.aat5991
- Task Force on Artificial Intelligence of the Agency for Digital Italy (2018). White Paper on Artificial Intelligence at the service of citizens.

- Tesla. (nd). Support: autopilot. Tesla. Available from: <https://www.tesla.com/support/autopilot>
- The Danish Government (2018). Strategy for Denmark's Digital Growth. Ministry of Industry, Business and Financial Affairs.
- The Danish Government (2019). National Strategy for Artificial Intelligence. Ministry of Finance and Ministry of Industry, Business and Financial Affairs.
- The Foundation for Responsible Robotics (2019). About us: Our mission [online] Available from: <http://responsiblerobotics.org/about-us/mission/>
- The Future of Life Institute (n.d.) AI Policy Challenges and Recommendations. Available at: <https://futureoflife.org/ai-policy-challenges-and-recommendations/#top>
- The Institute for Ethical AI & Machine Learning (2019). Homepage [online] Available from: <https://ethical.institute/index.html> [Accessed 11 Mar.2019].
- The Institute of Electrical and Electronics Engineers (IEEE) (2017). Ethically Aligned Design: First Edition. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems. (EADv2).
- The Institute of Electrical and Electronics Engineers (IEEE) (2019). Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems (EAD1e)
- The Partnership on AI (2019). About us [online] Available from: <https://www.partnershiponai.org/about/> [Accessed 11 Mar.2019].
- The White House (2019a). Accelerating America's Leadership in Artificial Intelligence. [online] Available from: <https://www.whitehouse.gov/articles/accelerating-americas-leadership-in-artificial-intelligence/>
- The White House (2019b). Artificial Intelligence for the American People [online] Available from: <https://www.whitehouse.gov/ai/>
- Thiagarajan, K. (2019). The AI program that can tell whether you may go blind. The Guardian.
- Thielman, S. (2017). The customer is always wrong: Tesla lets out self-driving car data – when it suits. The Guardian.
- Thomson, J. (1976). Killing, letting die, and the trolley problem. The Monist. 59, 204–217.
- Thurman N. (2011). Making 'The Daily Me': technology, economics and habit in the mainstream assimilation of personalized news. Journalism. 12, 395–415.

- Tindera, M. (2018). Government data says millions of health records are breached every year. Forbes.
- Torres Santeli, J. and Gerdon, S. (2019). 5 challenges for government adoption of AI. [online] World Economic Forum.
- Turkle, S., Taggart, W., Kidd, C.D. and Dasté, O.,(2006). Relational Artifacts with Children and Elders: The Complexities of Cyber companionship. *Connection Science*, 18 (4) pp 347-362.
- U.S. Department of Education, (2014). Science, Technology, Engineering and Math.
- UAE Government (2018). UAE Artificial Intelligence Strategy 2031. [online] Available from: <http://www.uaesai.ae/en/> [Accessed 28 Apr. 2019].
- UK Government Department for Digital, Culture, Media & Sport (2019). Centre for Data Ethics and Innovation: 2-year strategy.
- UK Government Office for Science (2015) Artificial intelligence: opportunities and implications for the future of decision making.
- UNI Global Union (n.d.) Top 10 principles for Ethical Artificial Intelligence [online]. Available from: http://www.thefutureworldofwork.org/media/35420/uni_ethical_ai.pdf
- United Kingdom Commission for Employment and Skills, (2014). The Future of Work: Jobs and Skills in 2030.
- US Department of Defence (2018). Summary of the 2018 Department of Defence Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity
- Vanian, J. (2019). World Economic Forum Wants to Help Companies Avoid the Pitfalls of Artificial Intelligence [online] Fortune. Available at: <https://fortune.com/2019/08/06/world-economic-forum-artificial-intelligence/>
- Veale, M., Binns, R & Edwards, L. (2018). Algorithms that remember: model inversion attacks and data protection law. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 376 (2133).
- Vincent J. (2018). Drones taught to spot violent behavior in crowds using AI. The Verge..
- Vincent, J. (2017). Google's AI thinks this turtle looks like a gun, which is a problem. The Verge.
- Viscelli, S. (2018). Driverless? Autonomous trucks and the future of the American trucker. Center for Labor Research and Education, University of California, Berkeley, and

Working Partnerships USA. Available from:
<http://driverlessreport.org/files/driverless.pdf>

von der Leyen, U. (2019) Political guidelines for the next European Commission: 2019 – 2024.

Wachter S., Mittelstadt B. & Floridi L. (2017). Why a right to explanation of automated decision making does not exist in the general data protection regulation. *Int. Data Privacy Law* 7, 76–99. (doi:10.1093/idpl/ix005).

Wachter, S., Mittelstadt, B. & Russell, C. (2018). Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. *Harvard Journal of Law & Technology*. 31 (2).

Wagner, A.R. (2018). An Autonomous Architecture that Protects the Right to Privacy. In: AAAI / ACM Conference on Artificial Intelligence, Ethics and Society. AIES: 2018, 1-3 February, 2018, New Orleans, USA.

Wallach, W. and Allen, C.,(2009). *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press, New York.

Weinburg, C. (2019). Self-driving shuttles advance in cities, raising jobs concerns. *The Information*.

Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. Oxford, W. H. Freeman & Co.

West, D. M. (2018). *The Future of Work: Robots, AI, and Automation*. Brookings Institution Press Washington DC.

Winfield, A. F. (2019a). Ethical standards in Robotics and AI. *Nature Electronics*, 2(2), 46-48.

Winfield, A.F.T., & Jirotko, M. (2018). Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 376 (2133).

Worland, J. (2016). Self-driving cars could help save the environment – or ruin it. It depends on us. *Time*. Available from: <http://time.com/4476614/self-driving-cars-environment/>

World Business Council for Sustainable Development (WBCSD). (2000). *Eco-Efficiency: Creating more Value with less Impact*. WBCSD: Geneva, Switzerland.

Youyou, W., Kosinski, M., and Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences*. 112(4), 1036– 1040.

Zhao, J., Wang, T., Yatskar, M., Ordonez, V., & Chang, K. W. (2017). Men also like shopping: Reducing gender bias amplification using corpus-level constraints. arXiv. preprint arXiv:1707.09457

Zou, J. & Schiebinger, L. (2018). 'AI can be sexist and racist — it's time to make it fair', Nature Available from: <https://www.nature.com/articles/d41586-018-05707-8>

TEORI ETIKA DALAM KECERDASAN BUATAN (AI)

Dr. Budi Raharjo, S.Kom., M.Kom., MM.

BIODATA PENULIS



Dr. Budi Raharjo, S.Kom., M.Kom., MM lahir di Semarang, tanggal 22 Februari 1985. Beliau adalah Alumni dari Universitas Bina Nusantara (BINUS University) Jakarta dan juga alumni Universitas Kristen Satya wacana (UKSW) Salatiga. Dr. Budi Raharjo telah menjadi Dosen pada Universitas STEKOM pada mata kuliah Kepemimpinan (Leadership), mata kuliah Pengantar Akuntansi, Manajemen Proses, Manajemen Akuntansi dan Manajemen Resiko Bisnis. Selain sebagai dosen Universitas STEKOM, Dr. Budi Raharjo, M.Kom, MM juga mempunyai bisnis sendiri dalam bidang perhotelan dan juga sebagai wirausaha dalam bidang pemasok unggas (ayam) beku, ke berbagai kota besar, khususnya Jakarta dan sekitarnya.

Pengalaman beliau berwirausaha menjadi bekal utama dalam penulisan buku ajar yang diterbitkan oleh Yayasan Prima Agus Teknik (YPAT) Semarang. Oleh sebab itu bukunya berisi langkah langkah praktis yang mudah diikuti oleh para mahasiswa, saat mahasiswa mengikuti proses perkuliahan pada Universitas Sains dan Teknologi Komputer (Universitas STEKOM). Jabatan struktural yang di embannya saat ini adalah Wakil Rektor 1 (Akademik) Universitas STEKOM Semarang.



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

ISBN 978-623-8120-74-1 (PDF)



9 786238 120741