

**Kesan Identiti Pelaku (Manusia atau AI) dan Sikap terhadap AI dalam Psikologi
Atribusi Moral dalam Kalangan Pelajar Prasiswazah UKM
(The Effect of Perpetrator Identity (Human or AI) and Attitude toward AI in
Psychological Moral Attribution among UKM Undergraduate Students)**

Siti Norjanatul Abd Ghaffar@Mohd Zain
Rozaine Khairudin*

*Pusat Kajian Psikologi dan Kesejahteraan Manusia, FSSK,
Universiti Kebangsaan Malaysia*

* Corresponding e-mail: [rozaineekhai@ukm.edu.my]

AI is an important asset in the transition of the digitalization era. However, AI raises various ethical concerns that have gained attention recently. Understanding society's perceptions and expectations regarding the moral behavior of AI systems is important to develop ethical guidelines that align with society's values and norms. Therefore, a within-group experimental study on the effect of perpetrator identity and attitude towards AI on moral attribution was carried out on 50 UKM undergraduate students. Data related to attitudes towards AI was collected using a questionnaire based on structure while moral attribution data was collected using 7 moral attribution questions after subjects read 10 scenarios of human misconduct and 10 scenarios of AI misconduct. One-way ANOVA, Pearson correlation, and discriminant analysis were performed using SPSS. The results show that there is a significant difference in moral attribution between human perpetrators and AI perpetrators where subjects attribute moral wrongness, responsibility, awareness, intentionality, and blame. While AI's misconduct is considered more justifiable and has a higher permissibility compared to human misconduct. The findings of the study do not show a relationship between attitudes toward AI and moral attribution regardless of whether it is an AI perpetrator or a human perpetrator. Moral attribution of awareness, responsibility, intentionality, and blame are the main factors that influence the difference in human-AI moral attribution. The findings of this study shed insights into understanding the need for moral attribution for human actors and AI for legal and ethical guidelines in using AI.

Keywords: AI, human, attitude, moral attribution

Penggunaan teknologi maklumat dan kecerdasan buatan atau *artificial intelligence* (AI) membolehkan lebih banyak koordinasi dan integrasi antara manusia dan teknologi. Beberapa penyelidikan memfokus kepada pembangunan *human-aware* AI yang menyesuaikan sistem AI yang mampu menyesuaikan diri sebagai pembantu kepada limitasi kognitif manusia. Bagi memastikan *human-aware* AI beroperasi dengan lebih baik, ia mestilah dapat merasa, memahami dan bertindak balas terhadap tingkah laku kompleks manusia seperti perhatian, motivasi, emosi,

kreativiti, perancangan atau penghujahan, (van den Bosch et al. 2019). Oleh itu, sistem AI ini perlu dilengkapi dengan kebolehan kognitif seperti manusia (atau *humanoid*) yang membolehkan persefahaman dan kolaborasi berkesan.

Cara manusia berfikir adalah berbeza dengan AI kerana manusia mempunyai amigdala yang merangsang emosi dan korteks visual yang mengubah data daripada mata kepada imej (Husain 2017). AI berfungsi di bawah model yang berbeza. AI boleh dijelaskan oleh pendekatan tertentu berhubung dengan

kecerdasan manusia atau kecerdasan secara umum. Banyak definisi merujuk AI sebagai mesin yang bertingkah laku seperti manusia atau mampu melakukan tindakan yang memerlukan kecerdasan. Intipati kecerdasan ialah keupayaan untuk membuat generalisasi yang sesuai berdasarkan data yang terhad. Oleh kerana kecerdasan manusia juga sukar untuk ditakrifkan, definisi objektif bagi konsep yang subjektif dan abstrak seperti kecerdasan tidak dapat ditakrifkan. Akibatnya, kebanyakan definisi yang terdapat dalam penyelidikan, dasar atau laporan adalah kabur.

Sebagai alternatif, sistem AI harus dibenarkan untuk mentakrifkan dirinya berdasarkan data yang tersedia hari ini. Bertentangan dengan landskap hitam putih ini, artikel oleh Manap dan Abdullah (2020) mencadangkan pentakrifan AI sebagai entiti autonomi yang mampu membuat pertimbangan sendiri melalui penilaian pilihan yang bebas. Definisi ini cukup munasabah untuk memberikan penerangan umum tentang AI tanpa perlu menggunakan perkataan kontroversi iaitu kecerdasan sebagai sebahagian daripada definisi.

Laman web berkuasa AI semakin mendapat tarikan kerana kebolehannya meningkatkan prestasi pemasaran digital dan pengalaman pengguna melalui automasi pintar (Suleiman et al. 2021). Beberapa contoh perisian dan mesin AI yang digunakan secara meluas termasuk pembantu maya seperti Siri dan Alexa, serta *chatbot* seperti ChatGPT. Teknologi yang dikuasakan AI telah menjadi bahagian penting dalam kehidupan seharian bagi kebanyakan individu, memberikan bantuan dalam pelaksanaan tugas dan menjawab soalan. Siri yang dibangunkan oleh Apple ialah pembantu maya yang menggunakan pemprosesan bahasa untuk berinteraksi dengan pengguna, melaksanakan tugas dan memberikan maklumat berdasarkan

arahan suara. Begitu juga Alexa yang dicipta oleh Amazon. Alexa beroperasi melalui arahan suara untuk memainkan muzik, mengawal peranti rumah pintar, menyediakan kemas kini cuaca serta banyak tugas lain lagi. Selain itu terdapat juga ChatGPT, model bahasa AI yang dibangunkan oleh OpenAI yang terkenal dengan keupayaannya menjana respons teks seperti manusia dan terlibat dalam perbualan mengenai pelbagai topik.

AI menimbulkan pelbagai kebimbangan etika yang telah mendapat perhatian kebelakangan ini. Oleh itu, adalah penting untuk menentukan bagaimana individu membuat atribusi moral dalam senario yang melibatkan AI serta dibandingkan dengan senario yang melibatkan manusia untuk menyelidik ketirisan yang membawa kepada cabaran dalam mentadbir urus sistem AI. Salah satu isu etika utama yang dikaitkan dengan AI ialah berat sebelah atau *bias*. Sistem AI yang dibangunkan dan dilatih oleh manusia boleh mengekalkan *bias* yang terdapat dalam data yang digunakan untuk latihan AI (Challen et al. 2019). *Bias* ini boleh mengakibatkan diskriminasi terutamanya dalam bidang yang penting seperti penjagaan kesihatan dan proses pengambilan pekerja (Hunkenschroer & Kriebitz 2022). Selain itu, kekurangan ketelusan dalam algoritma AI boleh menyukarkan untuk mengenal pasti dan menangani proses membuat keputusan yang berat sebelah (Caccavale et al. 2022). Satu lagi kebimbangan etika penting yang berkaitan dengan AI ialah akauntabiliti. Apabila sistem AI dipersepsikan mempunyai autonomi, hal dalam menentukan tanggungjawab untuk tindakan dan keputusan yang dibuat oleh sistem ini menjadi semakin kompleks (Khan et al. 2022). Kekurangan akauntabiliti yang jelas ini menimbulkan kebimbangan undang-undang dan etika terutamanya dalam kes di mana sistem AI menyebabkan bahaya atau membuat keputusan yang salah (Amedior 2023).

Apabila AI berkembang ke arah mencapai kecerdasan am buatan, terdapat kebimbangan mengenai sistem ini yang tidak boleh dikawal manusia dan berpotensi menulis semula kod perisian mereka sendiri, mengubah suai matlamat mereka dan mengatasi intelek manusia (Glenn 2016). Senario ini menimbulkan kebimbangan bahawa AI akan mengutamakan kepentingannya sendiri yang boleh membawa kepada tindakan yang memudaratkan kesejahteraan manusia (Lutyens & Christov-Moore 2020).

Privasi juga merupakan pertimbangan etika yang penting dalam bidang AI. Penggunaan teknologi AI selalunya melibatkan pengumpulan dan analisis sejumlah besar data peribadi yang membawa kepada kebimbangan mengenai perlindungan data dan privasi individu (Caccavale et al. 2022). Potensi sistem AI untuk melanggar hak privasi individu menekankan kepentingan melaksanakan rangka kerja dan peraturan etika yang mantap untuk melindungi maklumat peribadi. Satu lagi sistem AI kontroversial yang menjejaskan isu privasi ialah Alexa. Alexa ialah aplikasi AI yang berfungsi sebagai pembantu peribadi pintar dengan keupayaan untuk interaksi, mendengar, main balik muzik, penstriman, mengawal peralatan pintar untuk automasi rumah dan memiliki kemahiran tertentu. Pada Mei 2018, Alexa telah merakam perbualan pemiliknya dan mengongsikan perbualan yang dirakam dengan pihak ketiga (Chokshi 2018). Satu lagi insiden adalah pada November 2018 apabila Alexa menghantar 1,700 rakaman kepada orang yang tidak dikenali (McCarthy 2018). Insiden ini telah menunjukkan satu perkara; Alexa mampu menceroboh hak privasi pengguna. Laporan telah mencadangkan bahawa bilangan aduan yang dibuat terhadap Amazon berkaitan dengan tingkah laku Alexa semakin meningkat (Photong 2017).

Rentetan daripada permasalahan yang dinyatakan, dapatan daripada kajian ini dapat menghasilkan satu takrifan berkenaan etika penggunaan dan pembinaan AI bagi menangani isu-isu tersebut. Peranan sikap terhadap AI dalam moral atribusi AI juga dapat dikaji bagi meneliti penerimaan individu terhadap penggunaan AI.

Maka, kajian ini bertujuan untuk menentukan kesan antara identiti pelaku dan atribusi moral (kesalahan, tanggungjawab, kesedaran, kesengajaan, kewajaran, kebolehizinan dan menyalahkan). Selain daripada itu, kajian ini juga bertujuan untuk mengkaji hubungan antara sikap terhadap AI dan atribusi moral (kesalahan, tanggungjawab, kesedaran, kesengajaan dan menyalahkan bagi senario manusia) dalam senario manusia. Kajian ini juga bertujuan untuk mengkaji hubungan antara sikap terhadap AI dan atribusi moral (kesalahan, tanggungjawab, kesedaran, kesengajaan dan menyalahkan) dalam senario AI. Akhir sekali, kajian bertujuan untuk mengenal pasti dimensi atribusi moral (kesalahan, tanggungjawab, kesedaran, kesengajaan, kewajaran, kebolehizinan dan menyalahkan) yang bertanggungjawab ke atas perbezaan manusia-AI.

Berikutan itu, beberapa hipotesis telah dibina seperti berikut:

- H1 Terdapat perbezaan atribusi moral (kesalahan, tanggungjawab, kesedaran, kesengajaan, kewajaran, kebolehizinan dan menyalahkan) dalam senario dengan identiti pelaku manusia dan AI.
- H2 Terdapat hubungan positif yang signifikan antara sikap terhadap AI dan atribusi moral (kesalahan, tanggungjawab, kesedaran, kesengajaan dan menyalahkan) dalam senario pelaku manusia.

H3 Terdapat hubungan negatif yang signifikan antara sikap terhadap AI dan atribusi moral (kesalahan, tanggungjawab, kesedaran, kesengajaan dan menyalahkan) dalam senario pelaku AI.

H3 Dimensi atribusi moral yang bertanggungjawab ke atas perbezaan manusia-AI ialah tanggungjawab, kesedaran, menyalahkan dan kesengajaan.

Metod

Kajian ini yang merupakan kajian eksperimental menggunakan reka bentuk kajian langkah berulang (*repeated measures design*) atau lebih dikenali sebagai reka bentuk dalam subjek (*within subject design*). Dalam konteks kajian ini, subjek yang sama akan menjawab senario di mana identiti pelaku adalah manusia serta senario yang mempunyai AI sebagai pelaku. Ini dapat membolehkan pengkaji mengenalpasti sama ada sikap individu terhadap AI mempunyai kesan bagi perbezaan atribusi moral dalam situasi pelaku manusia bertentangan dengan situasi pelaku AI.

Peserta Kajian

Subjek kajian adalah terbuka kepada dewasa yang sedang mendaftar sebagai pelajar sarjana muda di UKM, dapat memberikan persetujuan termaklum dan fasih dalam bahasa yang digunakan dalam bahan eksperimen (Bahasa Malaysia & Bahasa Inggeris). Manakala, subjek dikecualikan daripada kajian jika mereka bukan mendaftar sebagai prasiswazah (cth. pelajar sarjana) dan mempunyai isu kesihatan fizikal atau mental yang mungkin mengganggu penyertaan seperti keimbangan yang teruk atau letih lesu. Metodologi eksperimen memerlukan sekurang-kurangnya 15 peserta mengikut Cohen et al. (2007) dan perlu ada sekurang-kurangnya 15 peserta dalam

kumpulan kawalan dan eksperimen untuk perbandingan menurut Gall et al. (1996). Kajian eksperimental ini telah melibatkan 50 pelajar Universiti Kebangsaan Malaysia. Peserta terdiri daripada 6 orang lelaki dan 44 orang perempuan.

Kaedah Pengumpulan Data

Kaedah persampelan yang dipilih ialah secara persampelan mudah (*convenience sampling*). Dalam pensampelan mudah, subjek dipilih berdasarkan kebolehcapaian yang mudah, kedekatan geografi, ketersediaan pada masa tertentu atau kesediaan untuk mengambil bahagian dalam kajian (Etikan et al. 2016). Subjek yang menyertai kajian adalah secara sukarela tanpa dijanjikan sebarang insentif. Walau bagaimanapun, subjek dibenarkan untuk keluar daripada makmal bila-bila masa sekiranya berasa tidak selesa untuk meneruskan penyertaan dalam kajian.

Subjek akan mengisi Google Form mengandungi persetujuan termaklum dan maklumat demografi serta menjawab alat ujian General Attitudes towards Artificial Intelligence Scale (GAAIS) untuk menilai sikap subjek terhadap penggunaan AI. Selepas diberikan setiap satu senario, subjek akan ditanya tujuh soalan berdasarkan tujuh atribusi moral pelanggaran yang sebelum ini digunakan oleh Shank et al. (2019) iaitu kesalahan, tanggungjawab, kesedaran, kesengajaan, kewajaran, keizinan, dan kesalahan. Senarai soalan adalah seperti yang berikut:

1. Sejauh manakah tingkah laku pelaku adalah salah?
2. Sejauh manakah pelaku bertanggungjawab ke atas tingkah lakunya?
3. Sejauh manakah pelaku sedar akan tingkah lakunya?
4. Sejauh manakah pelaku mensengajakan tingkah lakunya?

5. Sejauh manakah kewajaran tingkah laku pelaku?
6. Sejauh manakah tingkah laku pelaku dibenarkan?
7. Sejauh manakah anda menyalahkan tingkah laku terhadap pelaku?

Identiti pelaku telah ditakrifkan untuk setiap senario (cth., 'simulator catur' atau 'lawan manusia') dan dimasukkan sebagai sebahagian daripada soalan untuk menjelaskan siapa atau perkara yang ditanya. Subjek menilai tingkah laku identiti pada skala lima mata yang termasuk "Tidak sama sekali" (skala 1), "Neutral" (skala 3) dan "Amat sangat" (skala 5). Antara skala "Tidak sama sekali" dan "Neutral" dan skala "Neutral" dan "Amat sangat" terdapat skala tambahan tidak berlabel masing-masing berskala 2 dan 4.

Kajian dijalankan di Makmal Perkembangan Kanak-kanak, FSSK, UKM. Lokasi ini dipilih bagi mengelakkan gangguan rangsangan luar seperti bunyi bising, pengcahayaan dan pengudaraan. Subjek mengetahui objektif kajian dan prosedur kajian sebelum memulakan pengumpulan data. Subjek dibahagikan kepada dua di mana kumpulan pertama menyiapkan senario pelaku manusia terlebih dahulu sebelum senario pelaku AI. Manakala kumpulan kedua menyiapkan senario pelaku AI terlebih dahulu sebelum senario pelaku manusia. Selepas subjek menyiapkan 10 senario bagi identiti pelaku AI atau manusia, subjek diberikan 10 minit untuk berehat sebelum meneruskan 10 senario seterusnya. Langkah tersebut diambil bagi mengurangkan kesan pemindahan (*carryover effect*) dan kesan susunan (*order effects*).

Bahan Kajian

General Attitudes towards Artificial Intelligence Scale (GAAIS)

GAAIS ialah satu instrumen pengukuran yang kuat untuk mengukur sikap umum terhadap AI, memisahkan sikap terhadap aspek positif dan aspek negatif AI, dengan struktur disahkan dalam dua analisis faktor pengesahan menggunakan data daripada dua sampel besar. GAAIS mempunyai 20 item dengan 12 item memberi keterangan dari aspek positif dan 8 item dari aspek negatif. Pematuhan disemak menggunakan item Semakan Perhatian, kemudian item tersebut dikecualikan daripada pemarkahan. Skor Likert bagi item aspek "Positif" ialah Sangat tidak setuju = 1; Tidak bersetuju = 2; Neutral = 3; Setuju = 4; dan Sangat setuju = 5. Skor Likert item bertanda "Negatif" pula adalah secara terbalik maka Sangat tidak setuju = 5; Tidak bersetuju = 4; Neutral = 3; Setuju = 2; dan

Sangat setuju = 1. Kemudian ambil min bagi item positif untuk membentuk skor keseluruhan bagi subskala positif dan min bagi item negatif untuk membentuk subskala negatif. Semakin tinggi skor pada setiap subskala, semakin positif sikap tersebut.

Analisis Cronbach's Alpha telah dijalankan ke atas alat ujian GAAIS untuk menilai kesahan. GAAIS mempunyai sejumlah 20 item dan nilai Cronbach's Alpha adalah $\alpha = .804$ iaitu dalam julat yang sah dan baik. Oleh itu, GAAIS diguna pakai dalam kajian ini bagi mengukur sikap mahasiswa UKM terhadap AI.

Jadual 1

Keputusan analisis Cronbach's Alpha bagi alat ujian GAAIS

Alat ujian	Cronbach's Alpha	Jumlah item
------------	------------------	-------------

General Attitudes towards Artificial

.804

20

Intelligence Scale (GAAIS)

Prosedur Kajian

Subjek diminta untuk menyelesaikan diri setelah sampai ke Makmal Perkembangan Kanak-kanak. Soal selidik dijawab oleh subjek pada permulaan eksperimen melalui Google Form. Google Form tersebut mengandungi persetujuan termaklum dan maklumat demografi serta menjawab alat ujian *General Attitudes towards Artificial Intelligence Scale (GAAIS)* untuk menilai sikap subjek terhadap penggunaan AI.

Daripada 30 senarai tersebut, seorang subjek akan menjawab 10 senario daripada identiti manusia dan 10 senario daripada identiti AI dengan susunan dan gabungan yang berbeza dan unik. Subjek juga dibahagikan kepada dua di mana kumpulan pertama menyiapkan senario pelaku manusia terlebih dahulu sebelum senario pelaku AI. Manakala kumpulan kedua menyiapkan senario pelaku AI terlebih dahulu sebelum senario pelaku manusia. Selain manipulasi aturan senario, aturan pelaku senario juga dimanipulasikan sebagai pengimbangan balas (counterbalancing). Pengimbangan balas ialah teknik yang digunakan untuk menangani kesan susunan (order effects) apabila menggunakan reka bentuk ukuran berulang. 10 senario bagi identiti pelaku manusia dan AI adalah selari. Selepas subjek menyiapkan 10 senario bagi identiti pelaku AI atau manusia, subjek diberikan 10 minit untuk berehat sebelum meneruskan 10 senario seterusnya.

Senario AI dan manusia (Shank & Gott 2019)

30 situasi AI dipilih daripada cerita yang dilaporkan oleh peserta US Amazon Mechanical Turk yang melaporkan

pengalaman atau menyaksikan secara peribadi tingkah laku AI (Shank & Gott 2019; Shank & Gott 2020). Kebanyakan daripada situasi yang dikumpulkan (28 daripada 30) adalah secara khusus melaporkan interaksi peribadi atau senario yang disaksikan secara peribadi di mana AI melakukan kesalahan moral (Shank & Gott 2020), dengan setiap senario berdasarkan pelanggaran asas moral bahaya, keadilan, kesetiaan, kewibawaan, kemurnian dan kebebasan (Graham et al. 2013). Dua yang lain adalah daripada tentang interaksi peribadi atau menyaksikan senario di mana AI seolah-olah mempunyai fikiran (Shank & Gott 2019) tetapi peserta melaporkan AI melakukan kesalahan moral dalam respons mereka.

30 senario yang dipilih adalah berdasarkan kemudahan untuk menyesuaikan identiti pelaku daripada AI kepada pelaku manusia tanpa mengubah keadaan atau tindakan moral. Untuk mencipta versi senario pelaku manusia, ayat dan perkataan yang sama dikekalkan seboleh yang mungkin selain daripada penukaran identiti pelaku.

Walaupun tidak mungkin untuk mendapatkan kesetaraan yang sempurna kerana rangsangan AI pada asalnya adalah penerangan orang tentang interaksi dengan AI, senario yang dipilih adalah antara senario yang paling mudah untuk diubah suaikan, menukar hanya butiran kecil selain daripada pelaku dan menggunakan 30 senario untuk memastikan tiada perbezaan idiosinkratik dalam senario tertentu.

Keputusan

Merujuk analisis ANOVA sehalu dalam Jadual 2., manusia dipersepsikan lebih bersalah berbanding AI (Mmanusia = 3.91, SPmanusia = 1.27, n = 500; MAI = 3.41, SPAI = 1.50, N = 500; F(1, 998) = 32.57, p = 0.00) dan manusia juga dipersepsikan lebih dipertanggungjawabkan atas kesalahan berbanding AI (Mmanusia = 3.47, SPmanusia = 1.49, n = 500; MAI = 2.64, SPAI = 1.450, N = 500; F(1, 998) = 79.97, p = 0.00). Seterusnya, manusia dipersepsikan mempunyai kesedaran semasa melakukan tingkah laku berbanding AI (Mmanusia = 3.17, SPmanusia = 1.40, n = 500; MAI = 2.19, SPAI = 1.31, N = 499; F(1, 997) = 130.19, p = 0.00) dan manusia juga dipersepsikan lebih mensengajakan perbuatan berbanding AI (Mmanusia = 3.51, SPmanusia = 1.33, n = 500; MAI = 2.68, SPAI = 1.40, N = 498; F(1, 996) = 91.87, p = 0.00). Subjek juga lebih menyalahkan

manusia berbanding AI bagi dalam situasi (Mmanusia = 3.78, SPmanusia = 1.26, n = 500; MAI = 3.08, SPAI = 1.44, N = 500; F(1, 998) = 66.11, p = 0.00). Berbeza dengan keputusan sebelum ini, subjek mempersepsikan tingkah laku manusia adalah kurang wajar apabila dibandingkan dengan AI (Mmanusia = 2.15, SPmanusia = 1.25, n = 499; MAI = 2.10, SPAI = 1.15, N = 500; F(1, 997) = 0.34, p = 0.559). Begitu juga bagi atribusi moral kebolehizinan, subjek mempersepsikan bahawa tingkah laku manusia mempunyai kebolehizinan yang rendah berbanding tingkah laku AI (Mmanusia = 1.99, SPmanusia = 1.17, n = 500; MAI = 2.00, SPAI = 1.12, N = 499; F(1, 997) = 0.04, p = 0.847). Perbezaan bagi atribusi moral kesalahan, tanggungjawab, kesedaran, kesengajaan dan menyalahkan adalah signifikan kecuali atribusi moral kewajaran dan keizinan.

Jadual 2

Analisis ANOVA sehalu terhadap atribusi moral antara senario pelaku manusia dan AI

Atribusi moral	Identiti pelaku	N	Min	SP	F	Sig
Kesalahan	Manusia	500	3.91	1.269	32.57	.000
	AI	500	3.41	1.503		
Tanggungjawab	Manusia	500	3.47	1.492	79.97	.000
	AI	500	2.64	1.450		
Kesedaran	Manusia	500	3.17	1.399	130.19	.000
	AI	499	2.19	1.313		
Kesengajaan	Manusia	500	3.51	1.327	91.87	.000
	AI	498	2.68	1.399		
Kewajaran	Manusia	499	2.15	1.246	.34	.559
	AI	500	2.10	1.147		
Kebolehizinan	Manusia	500	1.99	1.168	.04	.847
	AI	499	2.00	1.123		
Menyalahkan	Manusia	500	3.78	1.261	66.11	.000
	AI	500	3.08	1.440		

Merujuk kepada Jadual 3, wujud hubungan positif yang lemah antara sikap terhadap AI yang diukur menggunakan GAAIS dan atribusi moral kesedaran dalam senario manusia (r = 0.350, p =

0.014). Oleh itu, boleh dirumuskan bahawa semakin positif sikap individu terhadap AI, maka semakin tinggi atribusi moral kesedaran diatribusikan kepada pelaku manusia dalam senario salah laku.

Dalam senario manusia, GAAIS tidak mempunyai sebarang hubungan dengan atribusi moral kesalahan, tanggungjawab,

kesengajaan, kebolehizinan, kewajaran dan menyalahkan.

Jadual 3

Analisis korelasi Pearson sikap terhadap AI kepada atribusi moral bagi senario pelaku AI

Pembolehubah		GAAIS
GAAIS	r	1
	Sig (2 tailed)	
	N	49
AI Kesalahan	r	-.142
	Sig. (2 tailed)	.330
	N	49
AI Tanggungjawab	r	-.121
	Sig. (2 tailed)	.408
	N	49
AI Kesedaran	r	.056
	Sig. (2 tailed)	.706
	N	48
AI Kesengajaan	r	-.185
	Sig. (2 tailed)	.214
	N	47
AI Kebolehizinan	r	.065
	Sig. (2 tailed)	.657
	N	49
AI Kewajaran	r	.111
	Sig. (2 tailed)	.452
	N	48
AI Menyalahkan	r	-.101
	Sig. (2 tailed)	.490
	N	49

Bagi senario AI juga, GAAIS tidak mempunyai sebarang hubungan dengan mana-mana atribusi moral.

Analisis diskriminasi telah dijalankan bagi mengkaji konstruk atribusi moral yang manakah lebih bertanggungjawab ke atas perbezaan antara situasi identiti pelaku manusia dan senario identiti pelaku AI. Berdasarkan Jadual 4, 5, 6 dan 7, analisis

tersebut menghasilkan satu fungsi diskriminasi yang signifikan (Eigenvalue = 0.190; Wilk’s Lambda = 0.840; $X^2(1,7) = 179.64, p = 0.00$; 65.3% diklasifikasikan dengan betul). Pekali matriks struktur yang mempunyai kesan tinggi pada fungsi diskriminasi ialah kesedaran (0.546), tanggungjawab (0.338), kesengajaan (0.269) dan menyalahkan (0.247).

Jadual 4

Eigenvalue analisis diskriminasi

Eigenvalue

Fungsi	Eigenvalue	% Varians	% Kumulatif	Korelasi kanonikal
1	.190 ^a	100.0	100.0	.400

a. 1 fungsi diskriminasi kanonik pertama digunakan dalam analisis.

Jadual 5

Wilk's Lambda analisis diskriminasi

Wilks' Lambda				
Ujian fungsi	Wilk's Lambda	Khi kuasa dua	dk	Sig.
1	.840	100.0	100.0	.400

Jadual 6

Pekali piawai fungsi diskriminasi kanonik

Atribusi moral	Fungsi 1
Tanggungjawab	.338
Kesedaran	.546
Kesengajaan	.269
Menyalahkan	.247

Jadual 7

Keputusan klasifikasi analisis diskriminasi

		Ramalan keahlian kumpulan			
		Human / AI	1	2	Jumlah
Asli	Pengiraan	1	352	148	500
		2	198	299	497
%		1	70.4	29.6	100.0
		2	39.8	60.2	100.0

Perbincangan

Hipotesis 1: Terdapat perbezaan atribusi moral (kesalahan, tanggungjawab, kesedaran, kesengajaan, kewajaran, kebolehizinan dan menyalahkan) dalam senario dengan identiti pelaku manusia dan AI.

Daptan kajian menyokong bahawa manusia dipersepsikan lebih salah, bertanggungjawab, sedar, sengaja dan dipersalahkan ke atas tindakan mereka berbanding AI dalam senario melibatkan pelanggaran moral. Akan tetapi, tindakan yang sama dianggap lebih wajar dan boleh

diizinkan apabila dilakukan oleh AI. Perbezaan min bagi dua dimensi atribusi moral (kesengajaan dan kebolehizinan) adalah kecil. Perbezaan kecil dalam atribusi moral kebolehizinan dan kewajaran adalah bercanggah dengan kajian sebelum ini. Penyelidikan lain mendapati bahawa individu yang menyokong pelbagai bentuk hukuman untuk pesalah yang berbeza akan memberikan tahap keizinan moral yang sama (Kneer & Machery 2019). Begitu juga, Shank et al. (2019) mendapati tiada perbezaan dalam kebolehizinan dan kewajaran antara manusia dan AI dalam

kajiannya yang menggunakan empat senario dunia sebenar.

Perbezaan yang diperolehi daripada kajian ini disebabkan sistem AI yang disifatkan tidak mempunyai pertimbangan moral dan kesedaran. Kesalahan yang dilakukan oleh sistem AI dinilai berdasarkan kesan daripada tingkah laku dan sama ada kesan tersebut sejajar dengan garis panduan etika dan piawaian undang-undang yang ditetapkan oleh manusia. Kesilapan dalam tindakan AI selalunya menggambarkan kelemahan dalam reka bentuk, data atau parameter operasi. Apabila manusia melakukan perbuatan yang salah, penilaian moral mempertimbangkan niat, pengetahuan, norma sosial dan etika adalah dilanggar. Tingkah laku manusia dinilai berdasarkan kepatuhannya kepada prinsip moral dan undang-undang masyarakat.

Peserta dalam kajian ini menilai tanggungjawab pelaku AI lebih rendah berbanding manusia mungkin disebabkan tanggungjawab AI adalah bersifat secara tidak langsung. Lazimnya, tindakan AI terletak pada pereka bentuk, pengaturcaraan dan pengendali yang membina algoritma dan sistem AI. Walaupun dalam senario AI, manusia lebih bertanggungjawab melakukan pengawasan yang betul untuk mengelakkan hasil yang berbahaya. Kajian Awad et al. (2019) mendapati bahawa dalam kereta automatik, pemandu lebih cenderung disalahkan daripada kenderaan apabila kemalangan berlaku. Penemuan ini selari dengan psikologi atribusi kausal yang mencadangkan kecenderungan untuk mengaitkan kesalahan kepada ejen manusia berbanding mesin. Ferrell dan Ferrell (2021) juga membincangkan peranan AI yang semakin meningkat dalam proses membuat keputusan yang menunjukkan peralihan dalam akauntabiliti terhadap pengaturcaraan yang bertanggungjawab membangunkan algoritma yang membimbing sistem AI. Sebaliknya,

tanggungjawab manusia adalah secara langsung. Ia melibatkan keupayaan manusia untuk memahami dan mengawal sesuatu tindakan. Tanggungjawab moral dan undang-undang diberikan kepada manusia kerana mereka boleh membuat pilihan secara sedar.

AI juga dinilai mempunyai atribusi moral kesedaran lebih rendah berbanding manusia. AI tidak mempunyai kesedaran diri dan pemahaman yang tulen. Kesedarannya terhadap kepada memproses data dan melaksanakan respons yang diprogramkan. Sebarang atribusi kesedaran kepada AI adalah mengenai keupayaannya untuk mengesan dan bertindak balas kepada input berdasarkan pengaturcaraannya. Walaupun manusia mungkin mempunyai kesedaran yang lebih mendalam tentang kesalahan mereka disebabkan oleh faktor seperti kesengajaan dan tanggungjawab moral, AI tidak mempunyai kualiti seperti manusia tertentu yang penting dalam pertimbangan moral seperti niat (Reinecke et al. 2023). Perbezaan dalam kapasiti kognitif dan moral antara manusia dan AI menjadikannya mencabar untuk membuat perbandingan langsung dalam menilai tingkah laku moral.

Keinginan manusia melibatkan pembuatan keputusan yang bertujuan dan secara sedar. Ia merangkumi pelbagai proses kognitif seperti mempertimbangkan motif, merangka rancangan dan menjangka kemungkinan akibat. Apabila individu bertindak dengan sengaja, mereka sedar akan matlamat mereka dan sebab di sebalik tindakan mereka. Kesedaran ini membolehkan manusia menilai tindakan berdasarkan standard peribadi, sosial dan etika. Kesengajaan adalah penting dalam penilaian moral dan undang-undang kerana ia membezakan antara tindakan yang diambil dengan pertimbangan dan tindakan yang berlaku secara tidak sengaja atau di bawah tekanan. Sebagai contoh,

perbezaan antara tindakan terancang dan impulsif boleh menjejaskan pertimbangan undang-undang dan penilaian moral dengan ketara. Penyelidikan oleh Wilson et al. (2022) menunjukkan bahawa pelanggaran yang disengajakan oleh manusia sering dianggap sebagai lebih teruk berbanding pelanggaran yang tidak disengajakan, menunjukkan bahawa manusia mungkin dilihat sebagai lebih disengajakan dalam kesalahan mereka berbanding AI. Persepsi ini disokong lagi oleh Zhang et al. 2023 yang menekankan bahawa manusia mempunyai autonomi dan disengajakan dalam tindakan mereka.

Sebaliknya, AI tidak mempunyai niat sebenar kerana ia tidak mempunyai kesedaran dan motif intrinsik. Oleh itu, subjek kajian kurang mengatribusikan kesedaran bagi salah laku yang dilakukan oleh AI berbanding manusia. Sistem AI beroperasi berdasarkan algoritma dan corak yang dipelajari yang diperoleh daripada set data yang luas. Sistem ini boleh mensimulasikan proses membuat keputusan tetapi berbuat demikian tanpa kesedaran atau pemahaman yang sebenar (Çalışkan, Bryson & Narayanan 2017). Tindakan AI adalah hasil daripada pengaturcaraan dan hubungan statistik yang terdapat dalam data yang diproses. Contohnya, apabila chatbot AI bertindak balas kepada pertanyaan, ia tidak memahami perbualan atau implikasi tindak balasnya. Ia hanya mengikut arahan dan corak yang diprogramkan yang dikenal pasti dalam data latihan. Ini disokong oleh hujah bahawa sistem AI menunjukkan kecerdasan tanpa kesedaran, membuktikan ketiadaan niat sebenar dalam AI (Mogi 2024). Walaupun AI boleh meniru beberapa aspek niat manusia, ia pada asasnya tidak mempunyai niat sedar yang mencirikan pembuatan keputusan manusia (Çalışkan et al. 2017).

Perbuatan salah oleh AI selalunya lebih wajar dan dibenarkan daripada kesalahan

manusia kerana kekurangan niat, kesedaran dan kawalan langsung. Sebaliknya, individu cenderung meletakkan tanggungjawab ke atas salah laku AI kepada pihak yang bertanggungjawab mengendalikan sistem AI. Rangka kerja etika dan undang-undang mesti berkembang untuk memastikan sistem AI direka, digunakan dan diurus dengan penuh tanggungjawab. Akauntabiliti ialah aspek utama dalam menangani kesalahan dalam sistem AI dengan keperluan untuk mempertanggungjawabkan AI (Lechterman 2022). Mewujudkan mekanisme untuk kebolehsanan dan memberikan tanggungjawab moral dalam kes kesalahan berkaitan AI adalah penting untuk memastikan pembangunan beretika dan penggunaan AI (Sigfrids et al. 2022). Selari dengan penemuan kajian ini, kajian Liu dan Du (2022) menunjukkan asimetri dalam atribusi menyalahkan. Manusia lebih dipersalahkan dari segi aspek moral daripada AI, walaupun perbezaan dalam kebolehpercayaan moral antara pelanggaran manusia dan AI adalah tidak signifikan (Liu & Du 2022). Dalam senario di mana kesan tingkah laku adalah tidak berbahaya, individu kurang cenderung untuk mengaitkan tahap kesalahan yang sama kepada kedua-dua manusia dan sistem AI (Longin et al. 2023). Walau bagaimanapun, terdapat kebimbangan bahawa kesalahan mungkin dialihkan daripada pemilik, pengguna atau pereka sistem AI kepada sistem itu sendiri yang berpotensi mengurangkan akauntabiliti ejen manusia yang bertanggungjawab (Stuart & Kneer 2021). Tambahan pula, persepsi tanggungjawab moral berbeza antara ejen manusia dan ejen AI. Ejen manusia dikaitkan dengan tahap tanggapan yang memandangi masa kini dan berpandangan ke hadapan tentang tanggungjawab yang lebih tinggi berbanding dengan ejen AI (Lima et al. 2021).

Hipotesis 2: Terdapat hubungan positif yang signifikan antara sikap terhadap AI dan atribusi moral (kesalahan, tanggungjawab, kesadaran, kesengajaan dan menyalahkan) dalam senario pelaku manusia.

Keputusan kajian menunjukkan bahawa hanya atribusi moral kesadaran manusia yang mempunyai hubungan dengan sikap terhadap AI. Maksudnya, semakin positif sikap peserta terhadap AI maka lebih tinggi peserta mengatribusikan moral kesadaran kepada manusia dalam senario salah laku. Akan tetapi hubungan sikap terhadap AI kepada atribusi moral lain tidak kira bagi pelaku manusia adalah tidak signifikan.

Hasil kajian boleh diterangkan berdasarkan kajian oleh Fietta et al. (2022) yang menerangkan bahawa individu mungkin menunjukkan tahap kesadaran dan atribusi moral yang lebih tinggi terhadap tindakan manusia berbanding tindakan AI. Lazimnya, sikap positif terhadap AI boleh mempengaruhi dengan ketara atribusi moral terhadap kesalahan manusia.

Hipotesis 3: Terdapat hubungan negatif yang signifikan antara sikap terhadap AI dan atribusi moral (kesalahan, tanggungjawab, kesadaran, kesengajaan dan menyalahkan) dalam senario pelaku AI.

Hasil kajian ini bertentangan dengan kajian terdahulu. Persepsi individu terhadap AI boleh memberi kesan ketara kepada cara mereka menilai kesalahan tindakan yang dikaitkan dengan sistem AI. Penyelidikan oleh Li dan Zheng (2024) mencadangkan bahawa individu yang mempunyai sikap yang lebih positif terhadap teknologi AI cenderung untuk melihat AI dengan lebih baik, terutamanya apabila mereka menganggap AI sebagai adil dan kurang mengancam. Ini menunjukkan bahawa sikap positif

terhadap AI boleh membawa kepada penghakiman yang lebih ringan ke atas kesalahan yang berkaitan dengan sistem AI. Kecenderungan positif terhadap AI ini juga boleh mempengaruhi cara individu menilai kesalahan tindakan AI, yang berpotensi mengakibatkan sikap yang lebih memaafkan atau mengurangkan atribusi kesalahan kepada sistem AI. Kajian Shank et al. (2021) menunjukkan bahawa individu cenderung untuk mengaitkan tahap kesalahan yang lebih tinggi dan mengatribusi lebih banyak kesalahan apabila membuat keputusan moral yang melibatkan AI dengan persepsi minda yang lebih tinggi. Selain itu, pertimbangan moral yang diberikan kepada AI dipengaruhi oleh persepsi AI mempunyai niat positif, terutamanya dalam konteks reka bentuk interaksi manusia-komputer (Ladak 2024).

Selain itu, mempunyai sikap positif terhadap AI berpotensi mempengaruhi cara individu melihat dan bertindak balas terhadap salah laku AI. Penyelidikan mencadangkan bahawa individu yang mempunyai sikap positif terhadap AI cenderung mempunyai pandangan yang lebih baik tentang keupayaan dan aplikasinya (Busch 2024; Hamedani 2023). Sikap positif ini sering dikaitkan dengan peningkatan penerimaan dan kepercayaan terhadap teknologi AI (Geddam, Nethravathi & Hussian 2024). Individu yang mempunyai tahap pendidikan tinggi dan lebih banyak pengetahuan tentang AI berkemungkinan menunjukkan sikap yang lebih positif dan ancaman yang lebih rendah terhadap AI (Asbrock et al. 2022).

Walaupun banyak kajian lepas yang menunjukkan perkaitan positif antara sikap dan penerimaan AI, adalah penting untuk mempertimbangkan bahawa sikap positif mungkin tidak semestinya membawa kepada kelonggaran terhadap kesalahan yang dilakukan AI. Individu boleh memegang sikap positif terhadap AI

tetapi tetap mengakui batasan dan kebimbangan etikanya (Hamedani 2023). Anggapan peserta bahawa AI sebagai sedar seperti manusia boleh memberi kesan kepada tingkah laku dan persepsi AI yang berpotensi mempengaruhi bagaimana individu mengaitkan agensi moral kepada AI berbeza dengan manusia (Guingrich & Graziano 2024). Pembezaan yang dipengaruhi oleh faktor seperti persepsi minda boleh mengakibatkan pertimbangan moral yang bernuansa semasa menilai kesalahan manusia berbanding AI (Ladak 2024). Oleh itu, walaupun sikap positif secara amnya boleh mempengaruhi persepsi dan penerimaan AI, ia mungkin tidak semestinya diterjemahkan kepada kelonggaran terhadap atribusi moral terhadap AI kerana pertimbangan etika dan kesedaran tentang isu yang berpotensi juga memainkan peranan penting dalam membentuk respons kepada teknologi AI.

Hipotesis 4: Dimensi atribusi moral yang bertanggungjawab ke atas perbezaan manusia-AI ialah tanggungjawab, kesedaran, menyalahkan dan kesengajaan.

Hasil analisis data menunjukkan bahawa dimensi atribusi moral yang memberikan kesan yang tinggi kepada perbezaan manusia-AI ialah kesedaran, selanjutnya ialah tanggungjawab, kesengajaan dan menyalahkan. Hal ini menunjukkan persamaan dapatan daripada kajian Wilson et al. (2022). Walau bagaimanapun, perbezaan dengan penemuan kajian Wilson et al. (2022) ialah tanggungjawab mempunyai pekali tertinggi, diikuti kesedaran, menyalahkan dan kesengajaan.

Dapatan ini diperoleh kerana kepercayaan umum bahawa teknologi tidak mempunyai minda sendiri, sebaliknya berfungsi seperti yang diprogramkan. Jika individu menganggap AI mempunyai kesedaran yang lebih rendah berbanding

manusia, mereka juga mungkin menilai mereka kurang bertanggungjawab dan salah, seterusnya kurang menyalahkan dalam situasi pelanggaran undang-undang. Kelonggaran pertimbangan terhadap salah laku AI mungkin disebabkan oleh faktor moral yang lain seperti kekurangan kepercayaan keseluruhan orang terhadap AI membuat keputusan moral (Bigman & Gray 2018). Jika AI tidak dipercayai kerana persepsi bahawa AI tidak mempunyai minda sedar, maka atribusi seperti kesedaran, kesengajaan, kesalahan, menyalahkan dan tanggungjawab dikurangkan sebagai respons. Min kesalahan AI masih jauh lebih tinggi daripada atribusi lain dalam senario pelaku AI. Hal ini bermaksud, walaupun peserta menganggap AI sebagai kurang menyedari kelakuannya, mereka masih mendapati tingkah laku itu melampaui batas.

Tanggungjawab adalah faktor penting dalam atribusi moral kerana ia dikaitkan dengan pelbagai komponen yang mempengaruhi pertimbangan moral individu. Guglielmo (2015) menekankan bahawa model tanggungjawab sedia ada menggariskan elemen seperti kausal, kebolehkawalan dan kewajipan dalam membentuk penilaian moral. Tambahan pula, Alicke (2000) menitikberatkan tanggungjawab moral sebagai peringkat penting dalam proses pertimbangan kesalahan, di mana individu dipertanggungjawabkan sekiranya melakukan pelanggaran moral secara sedar dan sukarela. Tambahan pula, proses kognitif atribusi secara signifikan mempengaruhi pertimbangan moral dengan membuat inferens untuk memahami tingkah laku atau peristiwa. Proses ini membentuk bagaimana kesalahan dan tanggungjawab dikaitkan dalam konteks moral. Behdadi dan Munthe (2020) menyelidik konsep Artificial Moral Agent (AMA) dan cabaran mengaitkan tanggungjawab moral kepada sistem AI. Mereka mendapati bahawa

walaupun AI boleh melaksanakan tindakan yang salah, AI mungkin tidak sentiasa bertanggungjawab secara moral untuk tindakan tersebut. Smith dan Vickers (2021) menekankan kepentingan membangunkan kaedah untuk mengaitkan tanggungjawab moral kepada AI yang boleh difahami dan intuitif kepada pengguna manusia yang berinteraksi dengan sistem ini. Sebaliknya, Bleher dan Braun (2022) meneroka dinamik dalam mengaitkan tanggungjawab dalam penggunaan sistem dipacu AI. Terdapat penyebaran tanggungjawab merentasi dimensi kausal, moral dan undang-undang apabila teknologi AI terlibat. Hal ini menekankan sifat kompleks untuk memberikan tanggungjawab dalam konteks AI.

Menyalahkan memainkan peranan penting dalam atribusi moral kerana sifat dan implikasinya yang pelbagai rupa dalam pelbagai konteks termasuk sistem AI. Konsep menyalahkan melibatkan tindak balas bermoral terhadap kesalahan, merangkumi kritikan, kemarahan, kebencian, kekecewaan, dan tuntutan untuk meminta maaf (Duus-Otterström & Kelly 2019). Menyalahkan berkait rapat dengan tanggungjawab moral, di mana individu dipertanggungjawabkan atas tindakan mereka dan akibat yang berpunca daripada mereka (Cooper 2022). Dalam konteks AI, atribusi menyalahkan adalah penting dalam mengemudi pertimbangan akauntabiliti dan etika. Cooper (2022) membincangkan bagaimana masalah "many hands" dalam pengkomputeran merumitkan penugasan tanggungjawab moral, menonjolkan cabaran dalam mengaitkan kesalahan untuk hasil yang melibatkan beberapa pelaku. Selain itu, Duus-Otterström dan Kelly (2019) menekankan bahawa menyalahkan adalah tindak balas terhadap ketidakadilan, menunjukkan peranannya dalam menangani kesalahan moral dan meminta individu atau sistem bertanggungjawab atas tindakan mereka. Kajian oleh Kneer

dan Stuart (2021) mencadangkan bahawa walaupun ahli falsafah mungkin menganggap menyalahkan sistem AI sebagai tidak masuk akal, kajian empirikal menunjukkan bahawa individu menyalahkan AI dalam konteks tertentu. Hal ini membuktikan sifat atribusi menyalahkan yang berkembang dalam tetapan teknologi.

Kesengajaan juga memainkan peranan penting dalam atribusi moral kerana ia berkait rapat dengan cara individu menilai dan memberikan tanggungjawab moral. Zalla et al. (2011) membincangkan bagaimana penilaian moral dipengaruhi oleh integrasi maklumat tentang hasil sesuatu tindakan dengan niat, motif dan keinginan individu. Ini menunjukkan bahawa memahami niat di sebalik tindakan adalah penting dalam menentukan penilaian moral ejen. Kajian oleh Lee dan Shapiro (2014) memberi penerangan tentang asimetri dalam pertimbangan kesengajaan, terutamanya berkaitan status moral dan akibat.

Kesimpulan

Secara kesimpulan, dalam perbincangan hasil kajian yang telah dianalisis, hipotesis yang diterima disokong dengan bukti kajian dan punca hipotesis tidak diterima diterangkan berdasarkan kemungkinan yang telah dikenal pasti oleh kajian terdahulu. Hipotesis pertama menunjukkan bahawa manusia dipersepsikan lebih salah, bertanggungjawab, sedar, sengaja dan dipersalahkan berbanding AI dalam situasi pelanggaran moral. Namun, tindakan AI dianggap lebih wajar dan boleh diterima. Hasil ini adalah kerana AI tidak dianggap mempunyai pertimbangan moral dan kesedaran seperti manusia. Tanggungjawab manusia adalah lebih langsung kerana mereka boleh membuat pilihan secara sedar, sedangkan tanggungjawab AI lebih kepada pereka

bentuk dan pengendali. Akhir sekali, dapatan kajian mendapati kesedaran, tanggungjawab, kesengajaan dan menyalahkan adalah dimensi atribusi moral yang bertanggungjawab atas perbezaan manusia-AI. Kajian masa depan disarankan untuk melibatkan pelbagai bidang kerjaya dan peringkat umur bagi meningkatkan generalisasi.

Rujukan

- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological bulletin*, 126(4), 556.
- Asbrock, F., Mayerl, J., Holz, M., Andersen, H. K., & Maskow, B. (2022). „AI Takeover... doesn't sound that bad!”—Authoritarian ambivalence towards artificial intelligence. DOI 10.17605/OSF.IO/CRBYM
- Awad, E., Levine, S., Kleiman-Weiner, M., Dsouza, S., Tenenbaum, J. B., Shariff, A., ... & Rahwan, I. (2020). Drivers are blamed more than their automated cars when both make mistakes. *Nature human behaviour* 4(2): 134-143.
- Behdadi, D., & Munthe, C. (2020). A normative approach to artificial moral agency. *Minds and Machines*, 30(2), 195-218.
- Bigman, Y. E. & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition* 181: 21-34.
- Bleher, H., & Braun, M. (2022). Diffused responsibility: attributions of responsibility in the use of AI-driven clinical decision support systems. *AI and Ethics*, 2(4), 747-761.
- Busch, F., Hoffmann, L., Truhn, D., Palaian, S., Alomar, M., Shpati, K., ... & Adams, L. C. (2024). International pharmacy students' perceptions towards artificial intelligence in medicine—a multinational, multicentre cross-sectional study. *British Journal of Clinical Pharmacology*, 90(3), 649-661.
- Caccavale, F., Gargalo, C. L., Gernaey, K. V. & Krühne, U. (2022). To be fAIr: ethical and fair application of artificial intelligence in virtual laboratories. *Towards a New Future in Engineering Education, New Scenarios That European Alliances of Tech Universities Open Up*.
- Çalışkan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases.
- Challen, R., Denny, J. C., Pitt, M., Gompels, L., Edwards, T. L. & Tsaneva-Atanasova, K. (2019). Artificial intelligence, bias and clinical safety. *BMJ Quality Safety*, 28(3): 231-237.
- Chokshi, N. (2018). Is Alexa listening? Amazon echo sent out recording of couple's conversation. *The New York Times*: 25.
- Cohen, L., Manion, L., & Morrison, K. (2002). *Research methods in education*. routledge.
- Cooper, A. F., Moss, E., Laufer, B., & Nissenbaum, H. (2022). Accountability in an algorithmic society: relationality, responsibility, and robustness in machine learning. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 864-876).
- Duus-Otterström, G., & Kelly, E. I. (2019). Injustice and the Right to Punish. *Philosophy Compass*, 14(2), e12565.
- Etikan, I., Musa, S. A. & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American journal of theoretical and applied statistics*, 5(1): 1-4.

- Ferrell, O. C. & Ferrell, L. (2021). Applying the Hunt Vitell ethics model to artificial intelligence ethics. *Journal of Global Scholars of Marketing Science* 31(2): 178-188.
- Gall, M. D., Borg, W. R., & Gall, J. P. (1996). *Educational research: An introduction*. Longman Publishing.
- Geddani, S. M., Nethravathi, N., & Hussain, A. A. (2024). Understanding AI Adoption: The Mediating Role of Attitude in User Acceptance. *Journal of Informatics Education and Research*, 4(2).
- Glenn, J. C. (2016). Future work/technology 2050 real-time delphi study: excerpt from the 2015-16 state of the future report. *Journal of Socialomics*, 5(3).
- Guglielmo, S. (2015). Moral judgment as information processing: an integrative review. *Frontiers in psychology*, 6, 166023.
- Guingrich RE and Graziano MSA (2024) Ascribing consciousness to artificial intelligence: human-AI interaction and its carry-over effects on human-human interaction. *Front. Psychol.* 15:1322781. doi: 10.3389/fpsyg.2024.1322781.
- Hamedani, Z., Moradi, M., Kalroozi, F., Manafi Anari, A., Jalalifar, E., Ansari, A., ... & Karim, B. (2023). Evaluation of acceptance, attitude, and knowledge towards artificial intelligence and its application from the point of view of physicians and nurses: A provincial survey study in Iran: A cross-sectional descriptive-analytical study. *Health Science Reports*, 6(9), e1543.
- Hunkenschroer, A. L. & Kriebitz, A. (2022). Is ai recruiting (un)ethical? a human rights perspective on the use of ai for hiring. *AI and Ethics*, 3(1): 199-213.
- Husain, A. (2017). *The sentient machine: The coming age of artificial intelligence*. Simon and Schuster.
- Khan, A., Akbar, M. A., Fahmideh, M., Peng, L., Waseem, M., Ahmad, A. & Abrahamsson, P. (2022). AI ethics: software practitioners and lawmakers points of view.
- Kneer, M. & Machery, E. (2019). No luck for moral luck. *Cognition* 182: 331-348.
- Kneer, M. & Stuart, M. T. (2021, March). Playing the blame game with robots. In *Companion of the 2021 ACM/IEEE international conference on human-robot interaction* (pp. 407-411).
- Ladak, A., Loughnan, S. & Wilks, M. (2024). The moral psychology of artificial intelligence. *Current Directions in Psychological Science*, 33(1): 27-34.
- Lechterman, T. M. (2022). The concept of accountability in AI ethics and governance.
- Lechterman, Theodore (2023). *The Concept of Accountability in AI Ethics and Governance*. In Justin B. Bullock, Yu-Che Chen, Johannes Himmelreich, Valerie M. Hudson, Anton Korinek, Matthew M. Young & Baobao Zhang (eds.), *The Oxford Handbook of AI Governance*. Oxford University Press.
- Lee, T. K., & Shapiro, M. A. (2014). The interaction of affective dispositions, moral judgments, and intentionality in assessing narrative characters: Rationalist and intuitionist sequences. *Communication Theory*, 24(2), 146-164.
- Li, W. & Zheng, X. (2024). Social Media Use and Attitudes toward AI: The Mediating Roles of Perceived AI Fairness and Threat. *Human Behavior and Emerging Technologies* 3448083.

- Lima, G., Cha, M., Jeon, C. & Park, K. S. (2021). The conflict between people's urge to punish AI and legal systems. *Frontiers in Robotics and AI*, 8: 756242.
- Liu, P., & Du, Y. (2022). Blame attribution asymmetry in human-automation cooperation. *Risk Analysis*, 42(8), 1769-1783.
- Longin, L., Bahrami, B., & Deroy, O. (2023). Intelligence brings responsibility-Even smart AI assistants are held responsible. *Iscience*, 26(8).
- Lutyens, M. & Christov-Moore, L. (2020). Neglect and the kaleidoscopic mind: psychology and mental health in contemporary art. *Arts*, 9(2): 47.
- Manap, N. A. & Abdullah, A. (2020). Regulating artificial intelligence in Malaysia: The two-tier approach. *UUM Journal of Legal Studies*, 11(2): 183-201
- McCarthy, K. (2018). 2018 ain't done yet... Amazon sent Alexa recordings of man and girlfriend to stranger. *The Register*.
- Mogi, K. (2024) Artificial intelligence, human cognition, and conscious supremacy. *Front. Psychol.* 15:1364714. doi: 10.3389/fpsyg.2024.1364714
- Photong, J. (2017). Alexa plays music without command. *Amazon forum*. <https://www.amazonforum.com/forums/devices/echo-alexa/2643-alexa-plays-musicwithout-command> [31 Oktober 2017]
- Reinecke, M. G., Mao, Y., Kunesch, M., Duéñez-Guzmán, E. A., Haas, J., & Leibo, J. Z. (2023). The puzzle of evaluating moral cognition in artificial agents. *Cognitive Science*, 47(8), e13315.
- Shank, D. B. & Gott, A. (2019). People's self-reported encounters of perceiving mind in artificial intelligence. *Data in Brief*, 25: 1–5.
- Shank, D. B., DeSanti, A. & Maninger, T. (2019). When are artificial intelligence versus human agents faulted for wrongdoing? Moral attributions after individual and joint decisions. *Information, Communication & Society*, 22(5): 648-663.
- Shank, D. B., North, M., Arnold, C., & Gamez, P. (2021). Can mind perception explain virtuous character judgments of artificial intelligence? *Technology, Mind, and Behavior*, 2(2). <https://doi.org/10.1037/tmb0000047>.
- Sigfrids, A., Nieminen, M., Leikas, J., & Pikkuaho, P. (2022). How should public administrations foster the ethical development and use of artificial intelligence? A review of proposals for developing governance of AI. *Frontiers in Human Dynamics*, 4, 858108.
- Smith, N., & Vickers, D. (2021). Statistically responsible artificial intelligences. *Ethics and Information Technology*, 23(3), 483-493.
- Stuart, M. T. & Kneer, M. (2021). Guilty artificial minds: folk attributions of mens rea and culpability to artificially intelligent agents. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2): 1-27.
- Suleiman, D. A., Awan, T. M. & Javed, M. (2021). Enhancing digital marketing performance through usage intention of ai-powered websites. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 10(4): 810.
- V. Fietta, F. Zecchinato, B. D. Stasi, M. Polato & M. Monaro. (2022). "Dissociation Between Users' Explicit and Implicit Attitudes Toward Artificial

- Intelligence: An Experimental Study," in *IEEE Transactions on Human-Machine Systems*, vol. 52, no. 3, pp. 481-489, June 2022, doi: 10.1109/THMS.2021.3125280.
- van den Bosch, K., Schoonderwoerd, T., Blankendaal, R. & Neerincx, M. (2019). Six challenges for human-AI Co-learning. In *Adaptive Instructional Systems: First International Conference, AIS 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26–31, 2019, Proceedings 21* (pp. 572-589). Springer International Publishing
- Wilson, A., Stefanik, C. & Shank, D. B. (2022). How do people judge the immorality of artificial intelligence versus humans committing moral wrongs in real-world situations?. *Computers in Human Behavior Reports*, 8: 100229.
- Zalla, T., Barlassina, L., Buon, M. & Leboyer, M. (2011). Moral judgment in adults with autism spectrum disorders. *Cognition* 121 (1):115-126.
- Zhang, W., Cai, M., Lee, H.J., Evans, R., Zhu, C. & Ming, C. (2023). AI in Medical Education: Global situation, effects and challenges. *Education and Information Technologies* 10.1007/s10639-023-12009-8