

# بيت الفلسفة

# PHILOSOPHY HOUSE

Philosophy and Artificial Intelligence

الفلسفة والذكاء الاصطناعي

شخصية العدد:

ناصر ناصر

The Issue's Theme:  
Nassif Nassar

فلسفة الولاء

Philosophy of Loyalty

الذكاء البشري وحدود

الذكاء الاصطناعي

Human Intelligence  
and the Limitations of  
Artificial Intelligence

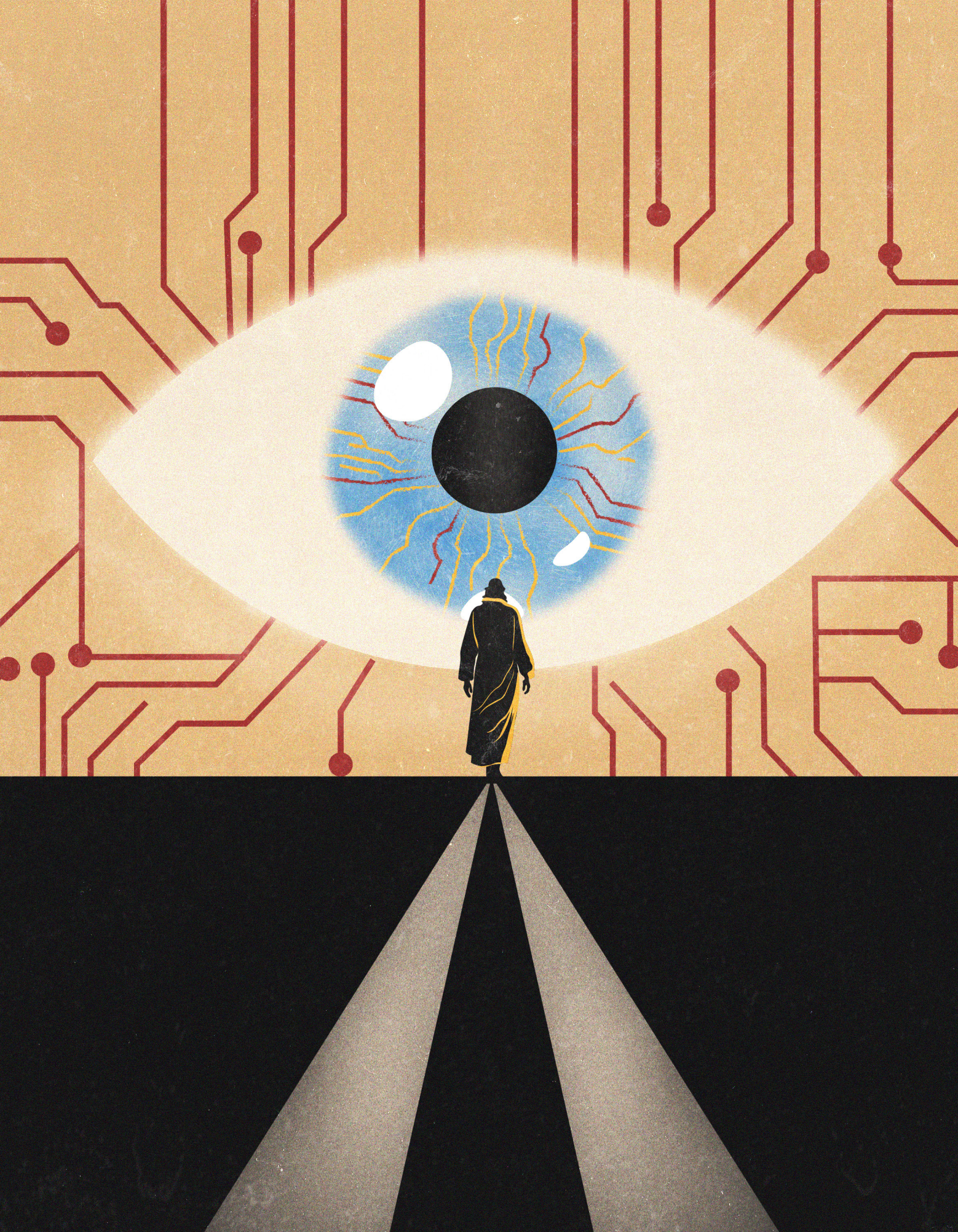
العقل الصناعي بين الذكاء والغباء:

هل يصلح العقل الصناعي للتفلسف؟

The Artificial Mind between Intelligence  
and Stupidity: Can It Philosophize?











# القولُ العربيُّ بالفلسفةِ

## Arabs' Philosophical View

## اقرأ في هذا العدد | Read in this issue

|    |                                                                                                                          |    |                                                                                                           |
|----|--------------------------------------------------------------------------------------------------------------------------|----|-----------------------------------------------------------------------------------------------------------|
| 05 | Editorial                                                                                                                | 04 | إفتتاحية                                                                                                  |
| 09 | The Artificial Mind between Intelligence and Stupidity: Can It Philosophize?<br>Ahmed Al-Barqawi                         | 08 | العقل الصنعي بين الذكاء والغباء هل يصلح العقل الصنعي للتفلسف؟<br>- أحمد برقاوي                            |
| 13 | Human Intelligence and the Limitations of Artificial Intelligence<br>Basil B. Al-Zain                                    | 13 | الذكاء البشري وحدود الذكاء الاصطناعي<br>- باسل بديع الزين                                                 |
| 21 | Philosophy of Loyalty<br>Ali Al-Hasan                                                                                    | 20 | فلسفة الولاء<br>- علي الحسن                                                                               |
| 27 | On Time<br>Nader El-Bizri                                                                                                | 28 | الذكاء الاصطناعي والمشكلة الأخلاقية<br>(تأملات فلسفية في عصر الآلة)<br>- إحسان علي الحيدري                |
| 29 | Artificial Intelligence and the Ethical Dilemma (Philosophical Reflections on the Machine Age)<br>Ehsan Ali Abdul-Ameer  | 37 | مستقبل الفلسفة في عصر الذكاء الاصطناعي<br>- حسن حمّاد                                                     |
| 37 | The Future of Philosophy in the Age of Artificial Intelligence<br>Hassan Hammad                                          | 44 | الوعي بين المبعث الإنساني والمبعث الآلي<br>- خالد كموني                                                   |
| 45 | Consciousness Between Human Impulse and Mechanical Inquiry<br>Khalid Kammouni                                            | 53 | من مطبعة «غوتنبرغ» إلى الإنترنت الإنسانية تمتحن أدواتها<br>- مليكة بن دودا                                |
| 53 | From Gutenberg's Printing Press to the Internet Humanity Puts Its Tools to the Test<br>Malika bin Duda                   | 60 | هل يستطيع الذكاء الاصطناعي أن يتفلسف؟<br>مدخل إلى إيبيستيمولوجيا "فكر الآلة" وأنطولوجيته<br>- نادر البرزي |
| 64 | Can AI Philosophise? Prolegomena to an Epistemology and Ontology of 'Machine-Thought'<br>Nader El-Bizri                  | 68 | شخصية العدد: ناصيف نصار<br>العقلانية شرط الاستقلال الفلسفي<br>- علي عبدون                                 |
| 69 | Featured Personality of the Issue: Nassif Nassar Rationality as a Condition for Philosophical Independence<br>Ali Abdoun | 75 | هل تكون الفردانية خلاصًا لعذابات الإنسان المعاصر؟<br>شهد عبید                                             |
| 75 | Can Individualism Be a Solution to the Suffering of Modern Man?<br>Shahd Obaid                                           | 83 | قائمة المصادر والمراجع                                                                                    |
| 83 | List of Sources and References                                                                                           |    |                                                                                                           |



Editor-in-chief :

- Dr. Ahmed Barqawi

Advisory Board Members :

- Anwar Moghith

- Basil Badi Al-zain

- Hassan Hammad

- Khalid Kamouni

- Al- Zawawi Bagora

- Fathi Al-Triki

- Mohammed mahjoob

- Mohammed Al-Misbahi

- Mushir Aoun

- Malika bin doda

- Musa Barhouma

Editing and proofreading

- Basil Badi Al-zain

Translating & reviewing

- Ghanim Samarrai

Magazine Design

- Sami Mahjoub

All information published in the House of Philosophy magazine is for scientific purposes only, House of Philosophy magazine does not bear responsibility for any error or omission received in the magazine.

رئيس التحرير - أحمد برقاوي

**أعضاء الهيئة الاستشارية**

- أنور مغيث

- باسل الزين

- حسن حماد

- خالد كموني

- الزواوي باغورة

- فتحي التريكي

- محمد محجوب

- محمد المصباحي

- مشير عون

- مليكة بن دودة

- موسى برهومة

مدير التحرير

- باسل بديع الزين

الترجمة والمراجعة

- غانم السامرائي

تصميم المجلة

- سامي محجوب

جميع المعلومات المنشورة في مجلة بيت الفلسفة هي

لأهداف علمية فقط، مجلة بيت الفلسفة لا تتحمل

المسؤولية عن أي خطأ أو إغفال ورد في المجلة.

جميع الحقوق محفوظة 2025

للتواصل مع مجلة بيت الفلسفة

To contact the House of Philosophy Magazine

Magazine@philosophyhouse.ae



# الافتتاحية

نرى هل نصدق إذا قلنا بأن العالم يمرّ في مرحلة جديدة من التاريخ البشريّ اسمها مرحلة الذكاء الصنعيّ؟ الحقّ أنّ كلّ جديد ينطوي على سيّورة تتعّين فيها صورة الحاضر. بهذا المعنى، فإنّ المستقبل الذي يرومه الذكاء الصنعيّ يجعله حاضرًا وهو لا يزال في حقل المجهول، ومن ثمّ، فإنّ حقل التوقّعات في مآلات حضور الذكاء الصنعيّ في الحياة محفوف بمخاطر عدم اليقين. وآية ذلك أنّ صعد فاعليّة الذكاء الصنعيّ مُتعدّدة جدًّا، بدءًا من العلم، مرورًا بالصّناعة وأشكال الاقتصاد المتعدّدة، وانتهاءً بقدرته على التوقّع. قد يكون من الصّعب على علم الطّب أن يتطوّر بمعزلٍ عن إنجازات هذه الثّورة التكنو-إلكترونيّة والحاسوب، ومن الصّعب أيضًا أن تصوّر شغل الآلات من دون هذا العقل، وقس على ذلك. غير أنّ حضور الذكاء الصنعيّ في الحقل الأدبيّ والفلسفة والعلوم الإنسانيّة بعامة قد لا يكون شبيهًا بحضوره في الصّعد الّتي ذكرنا. فضلًا عن ذلك، يبقى السّؤال حاضرًا: هل يمكن أن نعزل الذكاء الصنعيّ عمّا يُنتجه الذكاء البشريّ؟ بالتأكيد لا. وهذه واحدة من أهمّ المشكلات الّتي قد نواجهها في العلوم الإنسانيّة على نحو خاصّ. هبّ أنّا طرحنا على الذكاء الصنعيّ سؤالًا تاريخيًّا حول شخصيّة مختلف حولها، فإنّ جواب الذكاء الصنعيّ لن يكون سوى حشد الآراء المتعلّقة بهذه الشّخصيّة الّتي نعرفها، وهي آراء قد تخضع لتقويمات إيديولوجيّة مُتناقضة، وقس على ذلك في علوم اجتماعيّة أخرى، وهذا يفضي بنا إلى القول بأنّ سلطة الذكاء البشريّ ستمارس فعلها في الذكاء الصنعيّ، بل قلّ كلّما مدّ الذكاء البشريّ الذكاء الصنعيّ بالمعرفة، صار الذكاء الصنعيّ قادرًا على أن يستجيب للذكاء البشريّ. إنّ أسئلة كثيرة متعلّقة بمكانة الإنسان في العالم ومستقبله. فهل يستطيع الذكاء الصنعيّ، كما يعتقد بعض المتفائلين، أن يقوم بالوظائف نفسها الّتي يقوم بها صاحب الذكاء الطّبيعيّ؟ وهل يمكن للذكاء الصنعيّ أن يزيد من ذكاء البشر؟ لا شيء باستطاعته أن يُعوّض عن العبقرية الّتي لا تُعلّم.

رئيس التحرير



---

# Editorial

Can we truly assert that the world is now entering a new epoch in human history, an epoch to which the name the age of artificial intelligence might rightly be given? Every innovation, after all, unfolds through a process in which the contours of the present are shaped and defined. In this respect, the future envisioned by artificial intelligence imposes itself upon the present, even while it still belongs to the domain of the unknown. Thus, the horizon of expectations surrounding AI remains charged with the risks of uncertainty. The fields in which AI exerts influence are vast and heterogeneous: from the natural sciences to industry and diverse sectors of the economy, and ultimately to the realm of predictive analysis.

It is increasingly difficult to imagine the progress of medical science apart from the achievements of this technoelectronic and computational revolution, just as it is difficult to imagine machines functioning without this ‘mind’. The same observation holds true across a multitude of disciplines. Yet, the presence of artificial intelligence within literature, philosophy, and the humanities may not mirror its role in the aforementioned fields. Here, a more fundamental question emerges. Can artificial intelligence be meaningfully separated from the very forms of intelligence that human beings themselves produce? The answer, it seems, must be in the negative. This problem is among the most urgent challenges confronting the humanities today.

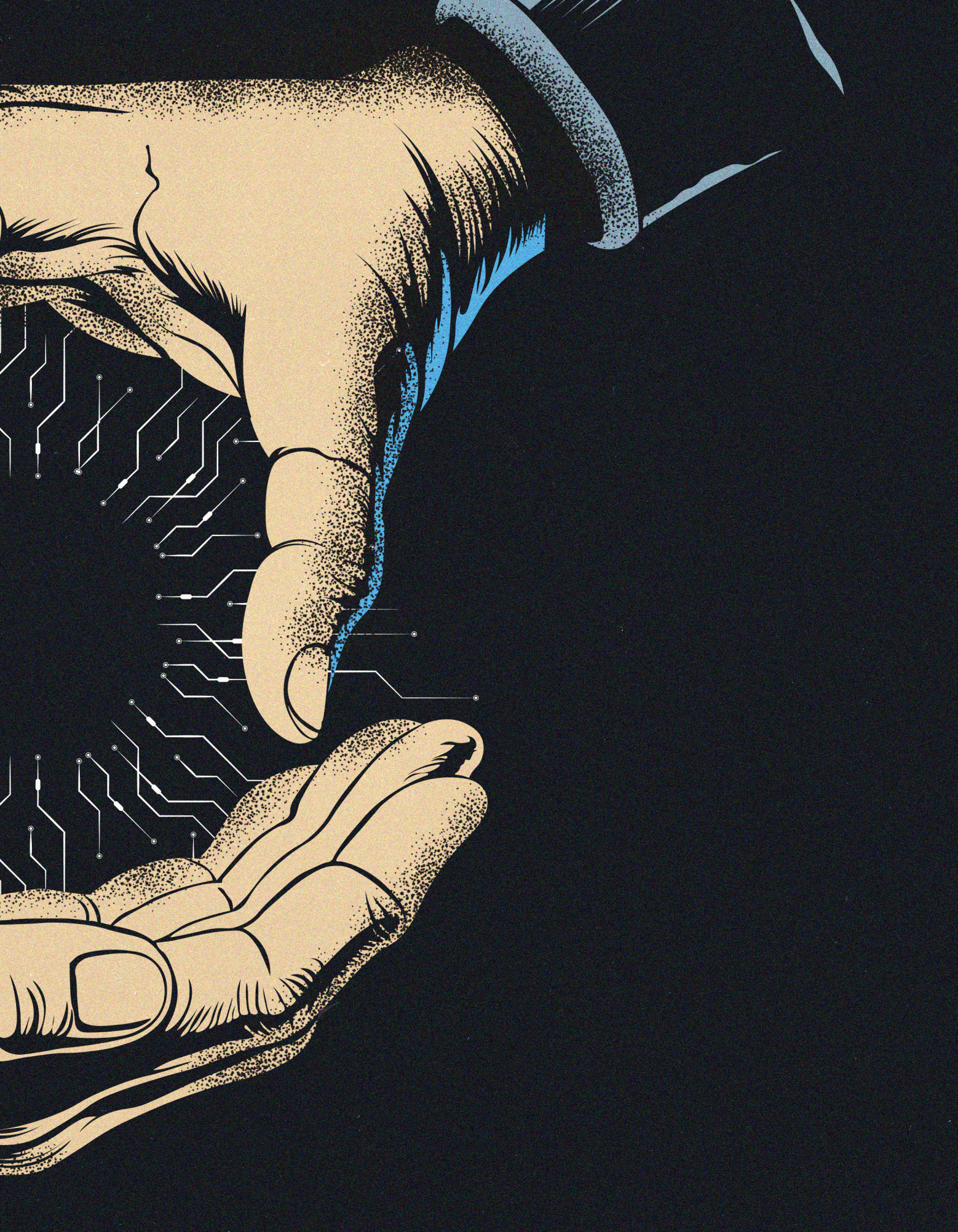
Consider, for example, the case in which we pose to artificial intelligence a historical question concerning a contested figure. The response, at best, would be a compilation of opinions already known to us – opinions that may themselves be subject to conflicting ideological interpretations. The same difficulty arises across the social sciences more broadly. This leads us to a crucial recognition: human intelligence inevitably exerts authority over artificial intelligence. The more human intelligence furnishes AI with knowledge, the more capable AI becomes of responding to it.

Such considerations invite larger philosophical questions concerning the place of the human being in the world and the fate of humanity itself. Can artificial intelligence, as some optimists maintain, truly perform the same functions as natural intelligence? Can it, perhaps, even enhance human intelligence? Yet nothing – no matter how advanced – can serve as a substitute for that genius which resists instruction and remains irreducibly human.

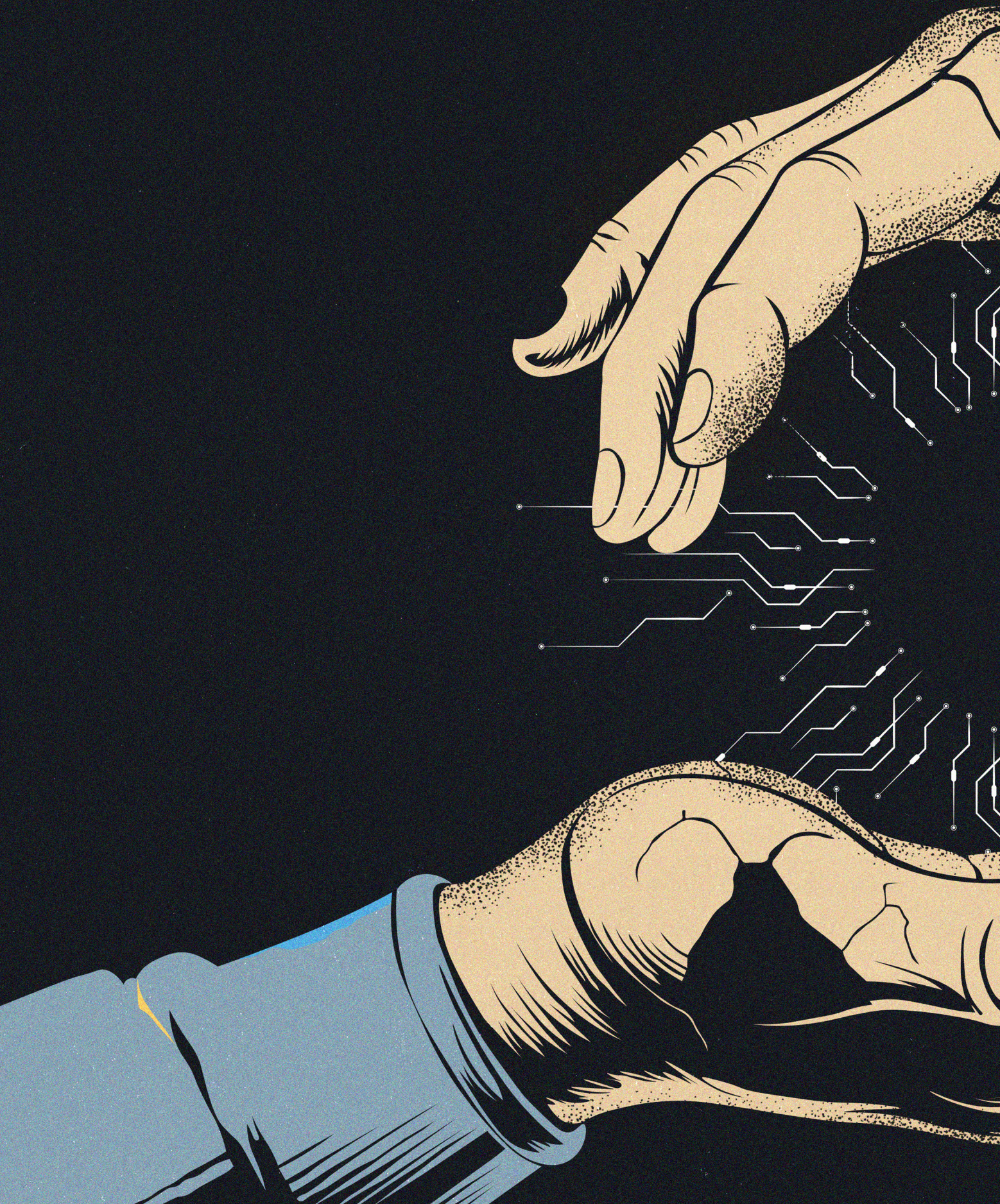
**- Editor in Chief**













# الدّكاء الاصطناعيّ والمشكلة الأخلاقيّة (تأمّلات فلسفيّة في عصر الآلة)

إحسان علي الحيدري

## أولاً: تقاطع الدّكاء الاصطناعيّ والفلسفة

يمثّل ظهور الدّكاء الاصطناعيّ (AI) نقطة تحوّل حاسمة في مسيرة الحضارة البشريّة، متجاوزاً كونه مجرد تطوّر تكنولوجي؛ ليصبح ظاهرة تثير أعمق الأسئلة الفلسفيّة. فالدّكاء الاصطناعيّ الذي صاغ مصطلحه (جون ماكارتي) العام 1955، يُعرّف بأنه مجموع التقنيّات والخوارزميّات التي تسعى إلى محاكاة الدّكاء البشريّ. الحقّ أنّ الثورة المعلوماتيّة لا تقتصر على إيجاد مخترعات وأدوات جديدة، بل تتعدّها إلى إحداث صدمة كبيرة في مفاهيم كانت تُعدّ حتى وقت قريب وقفاً على الإنسان وحده.

يتطلّب فهم تحديات الدّكاء الاصطناعيّ الأخلاقيّة مقاربةً فلسفيّة عميقة تتجاوز الجوانب التقنيّة البحتة، وتتناول مسائل جوهريّة مثل أسس الدّكاء الاصطناعيّ الفلسفيّة، وإمكانية تفوقه على الدّكاء البشريّ، أو امتلاكه الوعي والمشاعر والإرادة الحرّة.

تاريخياً، لم تكن العلاقة بين الفلسفة والتكنولوجيا حديثة العهد. فقد رأى بعض الفلاسفة الإغريق، مثل أفلاطون وأرسطو، أنّ التقنيّة محاولة لتقليد الطبيعة أو تجاوزها، بينما أشار فرانسيس بيكون في عصر النهضة إلى تأثير التكنولوجيا في تحسين ظروف المعيشة. وفي القرن التاسع عشر، عدّ إرنست كاب التقنيّة امتداداً للأعضاء البشريّة. هذه الجذور الفلسفيّة للتكنولوجيا تؤكّد أنّ القضايا الأخلاقيّة التي يثيرها الدّكاء الاصطناعيّ ليست جديدة تماماً، بل هي امتداد لتساؤلات قديمة حول طبيعة الإنسان وعلاقته بإبداعاته.

وبعد، سنتناول هذه المقالة العلاقة المعقّدة بين الدّكاء الاصطناعيّ والوعي، مستكشفة التحديات الأخلاقيّة الرئيسيّة التي يفرضها هذا الدّكاء، ومن ثمّ تنتقل إلى استعراض الأطر الفلسفيّة التي يمكن أن توجّه تطوره، وأخيراً تناقش مستقبل الوجود الإنسانيّ في ظلّ هذه التقنيّة المتسارعة.

## ثانياً: الدّكاء والوعي - أسئلة وجوديّة في عصر الآلة

أ - تمييز الدّكاء الاصطناعيّ الضّعيف من القويّ عادةً ما يتمّ التمييز في مجال الدّكاء الاصطناعيّ بين نوعين رئيسيين: الدّكاء الاصطناعيّ الضّعيف (Weak AI) والدّكاء الاصطناعيّ القوي (Strong AI). الدّكاء الاصطناعيّ الضّعيف هو الذي يكون مبنياً على قاعدة برامج يتمّ تزويد الحاسوب بها مسبقاً، مثل نظام "ديب بلو" الذي هزم بطل العالم في الشّطرنج غاري كاسباروف. هذا النوع من الدّكاء الاصطناعيّ يؤديّ مهامّ محدّدة بكفاءة عالية، لكنّه لا يمتلك فهمًا حقيقيًا أو وعياً بما يفعله.

في المقابل، يُطلق مصطلح الدّكاء الاصطناعيّ القويّ على الأنظمة التي يُفترض أنّها قادرة على محاكاة الدّكاء البشريّ بشكل كامل، بما في ذلك الوعي والمشاعر والإرادة الحرّة.

ب - مشكلة الوعي في الآلات: هل يمكن للآلة أن تفكّر، أو تشعر، أو تمتلك إرادة حرّة؟

تتمحور التّقاشات الفلسفيّة حول السّؤال المتعلق بإمكانية الانتقال من الحواسب الذكيّة إلى الروبوتات الواعية. حقيقة الأمر أنّ مفهوم الوعي أصبح منذ ديكارت قطب التفكير الفلسفيّ في تحديد علاقة الطّبيعة بالذات الإنسانيّة؛ وقد بنى هيغل نسقه الفكريّ بأكمله على اكتشاف الوعي بالذات. يُعدّ الوعي أخصّ خصائص العقل البشريّ، وهو "لغز محيّر فعلاً".

ثمة اتجاهات فلسفية مختلفة في دراسة الوعي. "الاتجاه الرّدّي" الذي يحاول إرجاع الوعي إلى شيء خارج عنه لصعوبة تصوّره كما هو؛ فالسلوكيّة تردّه إلى سلوكنا الخارجي، في حين تردّه النّزعة الفيزيائيّة إلى حالات الدّماغ؛ أمّا الوظيفيّة فتدّره إلى مجرد عمليّات حسابيّة. في المقابل، يرفض الفيلسوف جون سيرل نظريّة "ردّ الوعي"، ويعدّه "ظاهرة حقيقيّة لا تقبل الرّد".



# Artificial Intelligence and the Ethical Dilemma (Philosophical Reflections on the Machine Age)

Ehsan Ali abdul-ameer

## I. The Intersection of Artificial Intelligence and Philosophy

The rise of artificial intelligence (AI) marks a pivotal moment in human civilization – one that goes far beyond technical innovation to represent a phenomenon that poses some of the most profound philosophical questions of our time. Coined by John McCarthy in 1955, the term ‘artificial intelligence’ refers to a range of technologies and algorithms designed to simulate human cognition. The information revolution, therefore, is not merely about inventing new tools, but is a seismic shift that challenges long-held beliefs, which were once thought to be uniquely human.

Addressing the ethical dimensions of AI demands a rigorous philosophical inquiry that transcends purely technical considerations to engage with foundational questions relating to the philosophical underpinnings of AI, the possibility that AI may exceed human intelligence and the expectation that it may acquire consciousness, emotions and free will.

Historically, the relationship between philosophy and technology is deeply rooted. Greek philosophers, such as Plato and Aristotle, regarded technology as a means of imitating or transcending nature. During the Renaissance, Francis Bacon emphasized its role in improving human life. In the 19th century, Ernst Christian Kapp viewed technology as an extension of human faculties. These philosophical roots confirm that the ethical dilemmas posed by AI are not entirely new, but part of an enduring inquiry into human nature and man’s relationship with their creations.

Having said that, this article explores the complex relationship between AI and consciousness, examines the major ethical challenges AI poses, considers philosophical frameworks that may guide its development, and reflects on the future of human existence amid rapid technological acceleration.

## II. Intelligence and Consciousness: Existential Questions in the Machine Age

### A. Weak AI vs. Strong AI

In the field of artificial intelligence, a fundamental distinction is often drawn between two main types: ‘Weak AI’ and ‘Strong AI’. The former is based on databases and programmes pre-loaded into computers, such as ‘Deep Blue’ –the first computer programme to win a chess match against a reigning world champion under regular time controls. It played world champion Garry Kasparov and defeated him. Such systems are highly efficient at specific tasks but lack understanding or awareness.

In contrast, ‘Strong AI’ refers to hypothetical systems capable of fully replicating human intelligence, including consciousness, emotions, and free will.

### B- The Problem of Machine Consciousness: can a machine think, feel, or choose?

Central to the philosophical debate is the question: Can intelligent machines evolve into conscious robots? Since Descartes, consciousness has been a cornerstone of philosophical inquiry, especially in defining the relationship between self and nature. Hegel, for example, structured his philosophical system around the concept of self-awareness.



هذا التباين في فهم الوعي يمثل عقبة فلسفية جوهرية أمام الدَّكاء الاصطناعي، إذ إنَّ محاكاة الوعي تتطلب فهماً عميقاً لطبيعته، الأمر الذي لا يزال محلَّ خلاف بين الفلاسفة. وعليه، إذا كان الوعي ظاهرة غير قابلة الاختزال إلى مجرد عمليات حسابية، فإنَّ تحقيق دكاءٍ اصطناعيٍّ وإعْ بالمعنى البشري قد يكون مستحيلًا؛ ممَّا يضع حدودًا نوعية لتفوق الآلة. من منظور آخر، رأى الفيلسوف الأمريكي دانيال دنت أنَّ الوعي "وهم". بهذا المعنى سعى إلى هدم "المسرح الديكارتِي" الذي يفصل بين الدَّات المفكَّرة والطبيعة الماديَّة. بعبارة أوضح، يرى دنت أن لا وجود لـ "أنا كامل ومجرد" في الدِّماغ، بل هناك موجات متناثرة مُتعدِّدة، ولا مانع من أن تكتسب الأجهزة الآليَّة قدرات واعية على أساسٍ طبيعيٍّ رياضي. هذا المنظور المادِّي يفتح الباب أمام إمكانية تحقيق الوعي الاصطناعي؛ ومن ثمَّ التساؤل عن إمكانية اكتساب الآلات حالاً أخلاقية أو حقوقاً مماثلة للبشر.

مع ذلك، يؤكِّد عدد كبير من الفلاسفة تفرد الذكاء البشري. فالقدرة على الشَّك، والتساؤل، والتأويل، والتفلسف، والإبداع، واستحضار المصير والوجود العام للإنسان على هذه الأرض، سمات لا يمكن للآلات أن تكتسبها. في هذا السياق، يؤكِّد اللُّغوي نعوم تشومسكي أنَّ استعمال العقل الإنسانيِّ للغة يجعله يُفْتَقُّ قدراته تفتيقاً يتعدَّد على الرُّجل الآلي أن يأتي بمثله، عنيبا الاستعمال اللامتناهي للغة، والجدة والابتكار، وملاءمة الموقف، إلخ. هذه الخصائص اللُّغوية والإبداعية والفلسفية، تظلَّ جِكرًا على الإنسان؛ ممَّا يضع حدودًا واضحة بين دكاء الآلة الحسابي والوعي الإنساني العميق. الحال أنَّ الذكاء الاصطناعي، مهما بلغ من تعقيد، سيبقى أداة تخدم البشر، من دون أن يمتلك ذاتاً أخلاقية مستقلة.

## ثالثاً: تحديات الدَّكاء الاصطناعيِّ الأخلاقية: قضايا ومفاهيم

أ - التحيز والإنصاف

يُعد التحيز أحد أبرز التحديات الأخلاقية في مجال الدَّكاء الاصطناعي، إذ يمكن لأنظمة الدَّكاء الاصطناعي أن ترث التحيزات من البيانات التي يتم تدريبها عليها؛ ممَّا يؤدي إلى معاملة غير عادلة لمجموعات مُعَيَّنة. على سبيل المثال، قد تضع خوارزميات التوظيف المتحيزة فئاتٍ سُكَّانية مُعَيَّنة في وضع غير مؤاتٍ. هذا التحيز ليس مجرد عيب تقني، بل هو انعكاس للتحيزات المجتمعية الموجودة في البيانات التاريخية التي

ينغذَّى عليها الدَّكاء الاصطناعي.

يمكن أن يبدأ التحيز في مرحلة تعريف المشكلة (مثل تحيز التصميم أو تحيز تأثير الإطار)، أو في مرحلة تجهيز البيانات (مثل تحيز أخذ العينات، بحيث تكون العينات المأخوذة من مجتمع ما أكثر من اللازم، بينما تكون العينات المأخوذة من مجتمع آخر أقل من المطلوب)، كما يمكن أن يظهر التحيز في مرحلة بناء النموذج (تحيز الخوارزمية، تحيز المتغير المحذوف)، وحتى في مرحلة اختبار النموذج (تحيزات التقويم البشري مثل تحيز التأكيد أو تحيز المراقب).

يواجه السعي إلى بناء أنظمة ذكاء اصطناعي "خالية من التحيز والمغالطات" تحدياً فلسفياً عميقاً. فبينما يُطرح هذا الهدف بوصفه معياراً أخلاقياً، يُشير الواقع إلى أن التحيزات البشرية المعرفية يمكن أن تتسرَّب إلى أنظمة الدَّكاء الاصطناعي عن غير قصد. وذلك ما يخلق مفارقة مفادها أنَّ البشر مدعوون إلى ضمان أخلاقيات الدَّكاء الاصطناعي، لكنَّ البشر أنفسهم هم مصدر عدد كبير من التحيزات التي ينبغي التغلب عليها.

ب - الخصوصية وحماية البيانات

تثير قدرة الدَّكاء الاصطناعي على معالجة كميات هائلة من البيانات الشخصية مخاوف كبيرة بشأن الخصوصية، وبشأن إساءة الاستعمال في المراقبة وفي تتبع الأفراد من دون موافقتهم. فشركات الإعلان، على سبيل المثال، قد تستعمل بياناتك الخاصة لتحديد الإعلانات المعروضة لك؛ ممَّا يثير أسئلة حول مدى الأمان والتحكُّم الشَّخصي. إنَّ حقَّ الإنسان في الخصوصية حق لا خلاف عليه، ويجب احترامه طوال دورة حياة نُظم الدَّكاء الاصطناعي، لا سيَّما في ما يتعلق بجمع البيانات وتشاؤها بما يتماشى مع القانون الدولي والأطر القانونية المحلية. تتجاوز هذه التأثيرات مجرد انتهاك البيانات لتلامس جوهر الهوية الإنسانية. فالدَّكاء الاصطناعي يتحدَّى فهمنا لهويتنا، إذ يمكن لهوياتنا عبر الإنترنت (التي تشكِّلها الخوارزميات) أن تبدو منفصلة عن ذواتنا الحقيقية.

ت - الشَّفافية والمساءلة

تُمثِّل طبيعة "الصندوق الأسود" الخاصة بنماذج الدَّكاء الاصطناعي تحدياً أخلاقياً كبيراً، إذ تُصعِّب على الناس فهم كيفية اتخاذ القرارات. هذا الافتقار إلى الشَّفافية يمكن أن يعيق بناء الثقة والمساءلة، لأنَّ المستخدمين لا يستطيعون رؤية الأساس المنطقي للنتائج التي يحرِّكها الدَّكاء الاصطناعي.



Consciousness is widely regarded as the most distinct aspect of the human mind – a phenomenon that remains, in many ways, “a truly perplexing mystery.”

In fact, philosophical approaches to consciousness vary. Let us take, for example, the ‘Reductive theories’. They attempt to explain consciousness by reducing it to external phenomena. Behaviorism equates it with observable behavior, physicalism with brain states, and functionalism with computational processes. However, philosopher John Searle rejects such reductions, asserting that consciousness is “a real phenomenon that cannot be dismissed.” This divergence remains a central philosophical barrier to creating conscious machines: simulating consciousness requires understanding its essence – an understanding still elusive.

From a different perspective, American philosopher Daniel Dennett describes consciousness as an ‘illusion’. He challenges the ‘Cartesian theater’ that posits a central, unified self-distinct from material nature. Dennett suggests there is no perfect and abstract self within the brain – only multiple processes interacting dynamically. Under this materialist framework, machines, in principle, can develop consciousness through mathematical and natural laws. This opens the door to discussing whether AI may eventually be granted moral consideration or rights.

However, many philosophers still argue for the uniqueness of human intelligence. Doubt, interpretation, creativity, philosophical reflection, and moral deliberation are capabilities that machines may never attain. Linguist Noam Chomsky emphasizes that the human mind’s use of language – its novelty, generativity, and contextual appropriateness – cannot be replicated by machines. These qualities, Chomsky argues, remain distinctly human, highlighting the fundamental divide between mechanical computation and profound human consciousness. No matter how advanced, AI will remain a tool – lacking an autonomous moral self.

### III. The Ethical Challenges of Artificial Intelligence: Issues and Concepts

With the rapid expansion of AI across sectors, a number of pressing ethical issues have surfaced. These are not limited

to technical design or implementation, but extend to fundamental philosophical questions concerning justice, privacy, and moral responsibility.

#### A. Bias and Fairness

Bias is among the most significant ethical concerns in AI. Systems often inherit biases from the data on which they are trained, potentially leading to unfair outcomes. For example, hiring algorithms may disadvantage certain demographic groups based on historical data patterns. This bias is not just a technical glitch; it reflects structural inequities embedded in society.

Besides, bias can emerge at every stage: in defining the problem (e.g., design bias, framing effects), data preparation (e.g., sampling bias), model building (e.g., algorithmic bias, omitted variable bias), and model testing (e.g., confirmation bias, observer bias).

Based on such arguments, one may claim that striving to create ‘bias-free’ AI faces a philosophical paradox: the very humans tasked with creating ethical AI systems are themselves prone to cognitive bias. This reveals a deeper tension: how can imperfect humans build perfectly ethical machines?

#### B. Privacy and Data Protection

AI’s capacity to process vast amounts of personal data raises significant concerns about privacy, surveillance, and consent. For instance, advertising companies may exploit private data to target users without their informed permission.

The right to privacy must be respected across the AI lifecycle, especially in accordance with domestic laws and international standards. Beyond mere data breaches, the ethical stakes involve our very identities. Algorithmically constructed digital profiles can diverge significantly from our authentic selves, raising fundamental questions about autonomy and representation.

#### C. Transparency and Accountability

The ‘black box’ nature of many AI models presents a serious ethical dilemma: how can decisions be trusted if they cannot be understood? Lack of transparency undermines both accountability and public confidence in AI systems.



بشريّ مُعَيَّن إلى الإضرار بشريّ آخرين؟ كيف يمكن للدَّكاء الاصطناعي أن يوازن بين الالتزام بالطاعة وعدم الإضرار بالإنسان؟

ب- أخلاق المنفعة

تركّز أخلاق المنفعة على النتائج، إذ ترى أنّ القرار الأفضل هو ما يحقق أكبر فائدة لأكثر عدد من الناس. في تطبيقات الدَّكاء الاصطناعي، يمكن تصميم الأنظمة لتحسين المنفعة المجتمعية الشاملة، حتّى لو أدّى ذلك في بعض الأحيان إلى ضرر فرديّ.

تبرز المشكلات التّفعّية في سيناريوهات مُعقّدة مثل قيادة السيارات ذاتيّة القيادة في لحظة حادث محتمل: هل يجب أن تنحرف السيارة لتجنّب دهس المشاة، حتّى لو أدّى ذلك إلى إصابة الرّكاب؟ هذه القرارات التي قد تكون "مثلى" من النّاحية الحسابية التّفعّية، قد تتعارض مع حدسنا الأخلاقيّ البشريّ الذي يولي قيمة كبيرة للفرد وحقوقه.

ت- أخلاق الفضيلة

تُقَدَّم أخلاق الفضيلة مقارنة مختلفة، إذ تركّز على تنمية الصّفات الفاضلة في الفاعل الأخلاقيّ بدلاً من الالتزام بالقواعد أو تحقيق أفضل النتائج. في هذا السّياق، تشير أخلاق الفضيلة إلى أنّ الدَّكاء الاصطناعي يجب أن يجسّد السّمات الفاضلة، مثل التعاطف واللّطف. يتطلّب هذا الأمر تصميم خوارزميات لا تكتفي بإصدار القرارات فحسب، بل تفكّر نيّة القرار، وتُقدّر أبعاده الاجتماعيّة قبل تنفيذه.

يهدف مشروع "دكاء التّوايا" الذي أطلقته جامعة طوكيو بالتعاون مع "Open AI" إلى تصميم خوارزميات تعتمد على تحليل الشّيفات، والرّبط بين المعلومات الرّقمية والمفاهيم القيمية المستقاة من فلسفات متعدّدة، بما في ذلك أخلاق الفضيلة. يمثّل هذا التوجّه تحوّلًا فلسفيًا مهمًا، إذ لا يسعى إلى تدريب الدَّكاء الاصطناعي على النّاتج فحسب، بل على المعاني أيضًا؛ مما يجعله أكثر قدرة على التعامل مع المرونة الأخلاقية العالية المطلوبة في مجالات مثل الطبّ، والقانون، والتعليم.

ث- الأخلاق الوجودية

تطرح الأخلاق الوجودية أسئلة حول معنى الوجود، والحرية، والمسؤوليّة الشّخصية في عالم لا يمتلك معنىً جوهريًا مُسبّقًا. فالدَّكاء الاصطناعي يُهدّد بتقويض مفهوم "الدَّكاء الوجودي" البشريّ، أي القدرة على التصدّي للأسئلة الكبرى المتعلّقة بمعنى الحياة والوجود والهويّة والفناء. من شأن الاعتماد المفرط على الدَّكاء الاصطناعي، وتفويضه اتّخاذ بعض القرارات الإنسانيّة الأساسيّة أن يؤدي إلى تآكل الاستقلاليّة البشريّة. يثير تنازل البشر تدريجيًا عن مسؤولياتهم في اتّخاذ القرارات الحاسمة تساؤلات حول طبيعة الوكالة البشريّة والحرية.

ولضمان المساءلة، يجب بناء أنظمة الدَّكاء الاصطناعي على نحو يتضمّن درجة عالية من الوضوح ومن قابليّة التّفسير، مع وجود ميزات لتتبّع مراحل اتّخاذ القرارات المؤتمتة، ولا سيّما تلك التي قد تُلحق ضررًا بالأفراد. ومع ذلك، يظل تحديد الجهة المسؤولة أمرًا معقّدًا عندما تتسبّب أنظمة الدَّكاء الاصطناعي في ضرر أو ترتكب أخطاء. من النّاحية الفلسفيّة، تتطلّب المساءلة فهمًا لسلسلة الأحداث السّببية والمنطق وراء القرارات. إذا كانت قرارات الدَّكاء الاصطناعي غامضة، يصبح من المستحيل فحصها أو تحدّيها أو تحميلها المسؤولية؛ مما يقوِّض مبادئ العدالة والإنصاف والحوكمة الديمقراطيّة.

علاوة على ما تقدّم، فإنّ النقاش المتعلّق بـ "الوكالة الأخلاقية" للدَّكاء الاصطناعي يمثّل محورًا فلسفيًا رئيسًا. فهل يمكن للدَّكاء الاصطناعي أن يتمتّع بوكالة أخلاقية كاملة؟ يرى بعض الفلاسفة أنّ الآلات لا يمكن أن تمتلك أبدًا وكالة أخلاقية لأنّها تفتقر إلى القدرات اللازمة لذلك، مثل الحالات العقلية أو الانفعالات أو الإرادة الحرّة.

ث - السلامة والأمن

ينبغي لأنظمة الدَّكاء الاصطناعي أن تكون موثوقة وآمنة، بحيث لا تُعرّض المجتمع للخطر، وأن تنطوي على آليات مدمجة لمنع وقوع الضّرر. يتطلّب ذلك اتّخاذ تدابير فعّالة لقياس المخاطر ودرء الضّرر المحتمل، والتأكّد من أنّ وسيلة الدَّكاء الاصطناعي تُلائم الهدف المشروع المراد تحقيقه. على سبيل المثال، يمكن للمركبات ذاتيّة القيادة أن تشكّل خطرًا على حياة الناس في حال عدم التعرّف إليهم بوصفهم كائنات حيّة أو في حال عدم تدريب هذه المركبات على بعض حالات تعطل النظام.

## رابعًا: أطر توجيه أخلاقيات الدَّكاء الاصطناعي الفلسفيّة

تتطلّب التحدّيات الأخلاقية التي يفرضها الدَّكاء الاصطناعي تطبيقًا أطرٍ فلسفيّةٍ راسخة لتوجيه تطويره واستخدامه.

أ- أخلاق الواجب

تركّز أخلاق الواجب على القواعد والواجبات، إذ ترى أنّ القرار الأخلاقيّ هو الذي يلتزم بمجموعة من المبادئ أو القواعد الأخلاقية المطلقة، بغضّ النّظر عن النّاتج. في سياق الدَّكاء الاصطناعي، يمكن ترجمة ذلك إلى برمجة الآلات لتتبّع قواعد أخلاقية محدّدة سلفًا، مثل قوانين (إسحاق أسيموف) الشهيرة للروبوتات التي تنصّ على عدم إيذاء الإنسان وطاعة الأوامر وحماية الذات.

ومع ذلك، يمكن أن تُفضي هذه القواعد إلى تعقيدات لانهائية عند تطبيقها في العالم الحقيقيّ. على سبيل المثال، ماذا لو أدّى تنفيذ أمر



In addition, accountability demands that AI decisions be traceable and explainable, particularly when they carry risks of harm. Yet, assigning responsibility remains complex. When AI systems malfunction, who is to blame? Philosophically, accountability implies knowing the causal chain behind actions. Opaque AI decision-making challenges our ability to assign responsibility, undermining principles of justice and democratic oversight.

Related to this is the ongoing debate around ‘moral agency’. Can AI be considered a moral agent? Some philosophical argue with a ‘big no’! Machines, they believe, lack emotions, intentionality, and free will.

#### D. Safety and Security

AI systems must be designed with reliability and security in mind because their failures can pose significant societal risks. They have to have built-in mechanisms to prevent harmful incidents. Thus, AI safety must include risk assessment, harm mitigation, and assurance that systems serve legitimate human goals. When autonomous vehicles, for example, fail to recognize pedestrians as human beings, or when they are not trained to handle certain failure modes, they may cause fatal consequences.

### IV. Philosophical Frameworks for AI Ethics

To address the ethical concerns, several philosophical frameworks can be applied to guide AI development and the optimal use thereof. These may include:

#### A. Deontological Ethics

Deontological (duty-based) ethics holds that actions are ethical if they conform to universal moral rules, regardless of consequences. In AI, this translates to embedding rule-based ethical codes into machines, such as Isaac Asimov’s famous ‘Three Laws of Robotics’, which stipulate that a. A robot may not injure a human being or, through inaction, allow a human being to come to harm. b. a robot must obey the orders given it by human beings except where such orders would conflict with the First Law, and c. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

However, real-world scenarios often involve conflicting duties. What happens if following a command would harm another person? How should AI navigate such moral tensions?

#### B. Utilitarian Ethics

Utilitarian ethics focuses on outcomes – specifically, choosing the action that yields the greatest benefit for the most people. Applied to AI, this means designing systems to maximize societal good.

Yet, utilitarianism faces challenges in edge cases. For instance, should a self-driving car prioritize pedestrian safety over its passenger? These scenarios, while ‘optimal’ in theory, often conflict with deeply held human intuitions about individual rights and dignity.

#### C. Virtue Ethics

Virtue ethics follows a different approach. It shifts the focus from rules or outcomes to moral character. Applied to AI, this approach suggests that systems should embody virtues such as empathy, fairness, and kindness. This requires designing algorithms that not only execute decisions but also understand the intent behind them and consider social consequences.

The University of Tokyo’s ‘Intentional Intelligence’ project – developed in partnership with Open AI – aims at building algorithms grounded in context and ethical meaning, using cross-cultural values, including those from virtue ethics. This represents a shift toward creating AI that can navigate, morally, ambiguous domains like medicine, education, and law.

#### D. Existential Ethics

Existentialist ethics is concerned with freedom, personal responsibility, and meaning in a world devoid of inherent structure. AI threatens to erode human ‘existential intelligence’ – our ability to ask and answer life’s most profound questions.

Delegating core human decisions to AI may diminish human autonomy. The gradual outsourcing of critical judgment to machines invites deeper reflection: What does it mean to be human when decisions central to our existence are no longer ours to make?

ج- ما بعد الإنسانية وتحدياتها الفلسفية

يمثل مفهوم "ما بعد الإنسانية" (Posthumanism) حركة فكرية - فلسفية تدعو إلى تعزيز الوضع الإنساني من خلال التقنيات المتقدمة، مثل الدَّكاء الاصطناعي، والهندسة الوراثية. تهدف هذه الحركة إلى تجاوز القيود البيولوجية الحالية للإنسان، مثل الشيخوخة والمرض والموت، والوصول إلى مرحلة جديدة من التطور البشري. يرى أنصارها أنَّ الدَّكاء الاصطناعي هو المحرك الرئيس لتحقيق هذه الأهداف، إذ يمكن للبشر دمج هذه التقنيات مع أجسادهم وعقولهم بغية تعزيز قدراتهم بشكل أكبر.

ومع ذلك، يثير هذا المفهوم العديد من التحديات الفلسفية والأخلاقية الجوهرية. فهو ينكر فكرة التفوق البشري على الكائنات الأخرى، ويتحدى الفصل بين العقل والجسد، وبين البشر والآلات، وبين الإنسان والطبيعة. علاوة على أنه يتضمن تعريف الهوية البشرية من جديد، ناهيك بأنه يُهدد الجانب الإنساني بفقدان طبيعته، ويدفعنا إلى طرح أسئلة وجودية عميقة: متى يُعد الإنسان إنساناً؟ ومن يحدد ما يُعد "تحسيناً" وما يُعد "إفساداً" في الجسم البشري؟ إن السعي إلى تحقيق "ما بعد الإنسانية" يفرض نقاشاً فلسفياً حول ما إذا كان السعي إلى "التحسين" يُفقدنا جوهر إنسانيتنا، ويخلق أشكالاً جديدة من عدم المساواة بين البشر "المعززين" وغير "المعززين".

ح- مشكلة التباطؤ التنظيمي

تتسم الثورة التكنولوجية، ولا سيما في مجال الدَّكاء الاصطناعي، بسرعة لا مثيل لها. ومع ذلك، فإن القوانين والتشريعات والأطر الأخلاقية والمجتمعية تميل إلى اللحاق بالركب ببطء شديد. هذا التباين الزمني يؤدي إلى ما يمكن وصفه بـ "الشائرون نياماً"، بحيث يتم ابتكار التقنيات ونشرها بسرعة، بينما تستغرق الأطر القانونية والأخلاقية عقوداً لتطورها.

ينتج عن هذا التباطؤ التنظيمي "فراغ قانوني" أو "تأخر أخلاقي"، بحيث تُنشر تقنيات الدَّكاء الاصطناعي قبل فهم آثارها الأخلاقية الكاملة أو تنظيمها بشكل كافٍ. من شأن هذا الأمر أن يؤدي إلى أضرار غير متوقعة، وإلى تطبيع ممارسات قد تُعد غير أخلاقية لاحقاً. يدعو التحدي الفلسفي هذا إلى نهج أكثر استباقية ومرونة في صياغة الأطر الأخلاقية والقانونية.

خ- الأخلاق البيئية

تتجاوز الآثار الأخلاقية للدَّكاء الاصطناعي العلاقة بين الإنسان والآلة لتشمل البيئة. فالدَّكاء الاصطناعي، بقدراته الهائلة، يمكن أن يؤثر في البيئة بطرق مختلفة، سواء من خلال استخدامه في استخراج الموارد الطبيعية أو من خلال استهلاك الطاقة الهائل الذي تتطلبه خوارزمياته

ومراكز البيانات.

تؤكد مبادئ اليونسكو لأخلاقيات الدَّكاء الاصطناعي ضرورة تقييم تقنيات الدَّكاء الاصطناعي بناءً على آثارها في "الاستدامة"، وضرورة فهمها بوصفها طيفاً من الأهداف المتغيرة باستمرار، بما في ذلك أهداف الأمم المتحدة للتنمية المستدامة.

خامساً: حتمية الإشراف البشري

على الرغم من القدرات الفائقة للدَّكاء الاصطناعي في أتمتة المهام المعقدة وتحقيق الأهداف بكفاءة، إلا أنَّ المبادئ الأخلاقية الدولية تؤكد أنَّ "الرَّقابة البشرية والحزم" يجب أن يضمننا ألاَّ تحلَّ نظم الدَّكاء الاصطناعي مكان المسؤولية والمساءلة البشرية المطلقة. هذا يعني أنَّ القرار النهائي في الحالات التي تنطوي على عواقب لا يمكن إصلاحها يجب أن يبقى في يد البشر.

يعكس هذا المبدأ تحدياً فلسفياً وعملياً: كيف نُصمِّم أنظمة ذكاء اصطناعي قوية ومستقلة بما يكفي لتكون مفيدة، ولكنها تظل تحت السيطرة البشرية الكاملة؟ يتطلب هذا الأمر من البشر أن يكونوا "في الحلقة" ليس بوصفهم مشرفين فحسب، بل بوصفهم أيضاً كائنات مُدَرَّبة أخلاقياً، كائنات قادرة على التعرف إلى التحيزات والمفارقات الأخلاقية والتخفيف منها. إنَّ الحفاظ على المسؤولية البشرية المطلقة هو حجر الزاوية في بناء الثقة في أنظمة الدَّكاء الاصطناعي وضمان أنها تخدم البشرية، لا أن تحلَّ محلها.

إلى ذلك، تتطلب التحديات الأخلاقية للدَّكاء الاصطناعي استجابة عالمية ومنسقة. وقد أدركت منظمات مثل اليونسكو هذه الحاجة، إذ أصدرت توصية عامة للدَّكاء الاصطناعي في العام 2021 حُدَّت بموجبها الإجراءات الواجب اتباعها من قبل جميع المعنيين، بناءً على قيم ومبادئ أخلاقية تشمل احترام حقوق الإنسان وكرامته، والعيش في مجتمعات مسالمة وعادلة ومترابطة.

والحال أنَّ مبادئ الشفافية، وقابلية التفسير، والخصوصية، وحماية البيانات، والمسؤولية والمساءلة، تهدف إلى بناء مجتمع أكثر عدلاً وديمقراطية.

في الختام، لا يمكن فصل تطور الدَّكاء الاصطناعي عن النقاش الفلسفي العميق حول القيم الإنسانية. يجب أن يكون الهدف الأسمى هو تطوير ذكاء اصطناعي يتماشى مع تطوُّرنا الإنساني والأخلاقي، ويُعزز قيمنا، ويؤدي إلى مستقبل أكثر إشراقاً للبشرية جمعاء.



### E. Post-humanism and Its Philosophical Challenges

Post-humanism advocates enhancing the human condition through technologies such as AI and genetic engineering. It envisions a future in which humans transcend biological limitations like aging, diseases and death, and reach a new stage of human development. Proponents of this perspective regard artificial intelligence as the primary agent in actualizing these aspirations, insofar as the human being is increasingly able to incorporate such technologies into both body and mind, thereby transcending former limitations and advancing the horizons of human potential. However, post-humanism raises complex ethical questions. It challenges human exceptionalism and the separation between mind and body, blurs the boundaries between humans and machines and between man and nature. Furthermore, it entails a redefinition of human identity, threatening the human aspect with the loss of its nature, and prompting us to pose profound existential questions concerning the need for a profound reevaluation of what it means to be human: What constitutes a 'human'? Who decides what counts as 'enhancement' of the human body versus 'corruption' thereof? In fact, striving for the realization of post humanism compels a philosophical argument about whether the efforts aimed at the 'enhancement' would risk the loss of the essence of our humanism and create new forms of inequality between the humans who seek enhancement and those who do not.

### F. The Problem of Organizational Slowness

Technological change, especially in AI, proceeds at unprecedented speed. In contrast, legal systems, ethical norms, and institutional frameworks evolve slowly. This discrepancy results in a 'legal vacuum' or 'ethical lag': technologies are deployed before their implications are fully understood or regulated. This can lead to unforeseen harms and the normalization of practices that may later be deemed unethical. This philosophical challenge calls for a more proactive and flexible approach to formulating ethical and legal frameworks. It, in fact, underscores the urgent need for proactive, adaptive ethical governance.

### G. Environmental Ethics

AI's ethical implications extend beyond the human realm to include environmental impact. From natural resource extraction to energy-hungry data centers, AI can contribute to ecological degradation.

UNESCO's Principles on the Ethics of Artificial Intelligence stress the importance of evaluating technologies through a lens of sustainability, in alignment with the UN's Sustainable Development Goals.

### V. The Necessity of Human Supervision

Despite AI's unprecedented capabilities, global ethical principles emphasize that the final responsibility must remain with humans. AI should support – not replace – the human judgment in decisions with irreversible consequences.

This presents both philosophical and practical challenges: How do we design systems that are powerful and autonomous, yet remain under ethical human control?

Human supervision must involve more than technical control. It demands moral literacy – the ability to detect and correct bias and ethical contradiction. In reality, preserving human accountability is essential to building trust in AI and ensuring it remains a servant, not a master.

The international community has recognized this imperative. UNESCO's 2021 General Recommendation on Artificial Intelligence outlines guiding principles for all stakeholders, grounded in respect for human dignity, human rights, and democratic values. These include transparency, interpretability, data privacy, and accountability – principles aimed at cultivating a just, inclusive, and peaceful global society.

In conclusion, it is imperative to state that the evolution of artificial intelligence cannot be isolated from the deep philosophical dialogue surrounding human values. The ultimate aim should not be the creation of omnipotent machines, but the development of ethical AI systems that enhance human flourishing. This requires aligning technological progress with moral reflection – to ensure AI contributes to a brighter, more just future for all humanity.











