

---

# Response to the UK IPO's Consultation on Artificial Intelligence and Intellectual Property

This response to the UKIPO's [Call for Views](#) was prepared by academics from the School of Law at the University of Essex. The document addresses different aspects of intellectual property which reflect the contributors' areas of specialisation.

**Authors:** Dr. Alexandros Antoniou  
Prof. Stavroula Karapapa  
Dr. Eden Sarid  
Prof. Lorna Woods  
Prof. Ting Xu

**Corresponding authors:** Dr. Alexandros Antoniou  
Email: [a.antoniou@essex.ac.uk](mailto:a.antoniou@essex.ac.uk)  
Prof. Stavroula Karapapa  
Email: [s.karapapa@essex.ac.uk](mailto:s.karapapa@essex.ac.uk)

**Submission date:** 30 Nov. 2020

## General introduction

It is accepted that intellectual property rights should each be assessed and amended in the light of its primary objectives; in principle, it should not be structured to operate as general regulation of particular technologies by the back door. Of course, the balance of interests is reflected in the respective conditions for the creation of the different types of intellectual property and in the exceptions to protection.

As regards this first aspect, although AI processes can mimic human capacities, AI is not yet conscious of itself. ‘The prospect of a “strong” AI, that is to say one that is conscious of itself, seems after all still to be very futuristic’.<sup>1</sup> An AI system can be ‘instructed’ to generate works, but ‘is unlikely to be motivated’ to generate works.<sup>2</sup> If an AI system can only be instructed to generate works, it does not have ‘intent’ to be creative.

Of particular though not exclusive relevance to the second, we emphasise that the impact on human autonomy and dignity as well as specific human rights of the creation, development and deployment of AI should be borne in mind. This is especially so given the interconnection between intellectual property and other areas of law (e.g., misuse of private information and data protection). Many discussions in intellectual property focus on the relationship between creator/ owner and user/ infringer as these are common across most if not all instances of IP. The position of human subject, when relevant, is not considered save to the extent the subject is either creator/ owner or user. Issues can be seen, for example, as regards the patenting of DNA or, as we discuss below, copyright and deepfakes where this omission is likely to adversely affect minoritised groups, particularly women.

## Patents

### **Question No. 1:**

What role can/ does the patent system play in encouraging the development and use of AI technologies?

An underlying assumption at the basis of British patent law is that patents encourage the development (though not necessarily the use) of new inventions. And while this assertion is widely debated – and empirical data points in different, sometimes contradicting, directions – there is no reason to suggest that this assumption would not apply to the development of AI technologies. Yet, we must also consider the flip side – the cooling effects and deterrent dynamics that patents entail. Patents create legal monopolies that have been argued to stifle innovation, inflate prices for users, and entrench methodologies and path dependencies in research that obstruct further innovation. Against this background, we need to consider the nature of AI technologies – how these are developed, how they are used, and what they create (see discussion below).

---

<sup>1</sup> S Séjourné, ‘Draft Report on Intellectual Property Rights for the Development of Artificial Intelligence Technologies’. European Parliament, Committee on Legal Affairs, 2020/2015(INI), 24 April 2020.

<sup>2</sup> S Chesterman, ‘Artificial Intelligence and the Limits of Legal Personality’, [NUS Law Working Paper 2020/ 025](#), August 2020, p. 23).

Many of the questions laid out in this consultation, and the answers to them, reflect the basic trade-off at the basis of patent law between granting monopolies in exchange for disclosure and incentivising innovation. Keeping in mind that AI systems can be instructed to innovate, but not motivated to do so, any claim that AI systems should gain patents must be backed by very compelling evidence.

**Question No. 4:**

If AI cannot be credited as inventor, will this discourage future inventions being protected by patents? Would this impact on innovation developed using AI? Would there be an impact if inventions were kept confidential rather than made public through the patent system?

While there is very limited data to provide a concrete answer to these questions, it should be noted that there are indications that not crediting AI as inventor, will not discourage innovation.

There is data suggesting that innovators, companies, and investors in advanced technologies, such as AI, tend to rely strongly on market dynamics, such as economics of scale, first in the market advantages, brand recognition, rather than on patents.

The nature of AI subject matter also puts to question the risks of innovations and invention being kept confidential rather than made public, given the extreme speed in which these are being produced and become obsolete.

**Question No. 6:**

If AI was named as sole or joint inventor of a patented invention, who or what should be entitled to own the patent?

This question may hinge on the kind of invention created, and where the information that the AI used was sourced from. For example, if the technology uses data generated through the contributions of the public, there is no doctrinal reason why the public should not be the owner of the patent.

Yet, regardless of the answer – due to the nature of AI technologies, if AI was to be named the sole or joint author – it should be able to grant licenses to use the innovations and inventions it created. While this can be devised into the AI apparatus, for reasons of human rights, public interests such as fair competition and anti-trust, a framework for the granting of fair, reasonable, and non-discriminatory (FRAND) licenses must be implemented. Note, however, the possible consequences of such a licensing approach, discussed in relation to deepfakes below.

## Copyright

**Question No. 1:**

Do you agree with the above description of how AI may use copyright works and databases, when infringement takes place and which exceptions apply? Are there other technical and legal aspects that need to be considered?

It is well-recognised that, given the gap in protection for image rights in the UK, copyright has been used (where the subject owns the relevant rights). Whether or not this is the purpose of copyright, any amendment of the current rules so as to facilitate AI risks disturbing the existing balance, as can be seen if we consider a particular technology: deepfake generation. Much of the

discussion has focussed on the use of others' copyright works when considering the training data for various AI systems, as well as ownership of the end product. Looking purely at this relationship some solutions, e.g., developing a system of remuneration for the owners of any underlying work, could address some of the free rider issues. This does not however deal adequately with the interests of the subject (where the deepfake is to create a doppelganger of a known person), especially given that a significant proportion of deepfakes currently seem to be deployed in the context of pornography and are essentially weaponised against women. Copyright, in some circumstances, might then be a tool to fight against such activity (and a similar argument might be made with regard to performers' rights). While this will not work in all circumstances because copyright ownership lies elsewhere, this in itself is not a justification for removing that protection. Further, considering the possibility of non-subject ownership, institution of a payment system may encourage further use of images – indeed, that is the point – but doing so may legitimise the dehumanisation of the human subject. It might be said that porn deepfakes might not themselves benefit from copyright protection. Even if this is the case this does not address the problem that the deepfakes exist. While this may be a particular issue affecting some sub-groups more than others (consider the position of models in relation to their photographs), it does tie in with a more general concern about the datafication of society and the treating of people as mere sources of data (even in the context of photographs the processing of photographs is also for example in automated facial recognition technology).

What is to be done? One solution would be to tackle the issue of deepfakes directly in a specific regime (see approach in e.g., California); another would be to ensure that the use of photographs and videos (of people) do not benefit from any AI exceptions; a third could look to the subjects of photographs and give them something akin to a moral right to stop demeaning treatment of their image, e.g. an expansion of s. 85 of the Copyright, Designs and Patents Act 1988 (CDPA 1988) (though care would need to be taken to protect the parody exception).

**Question No. 2:**

Is there a need for greater clarity about who is liable when an AI infringes copyright?

The answer to this question links back to the question of the nature of the particular AI technology, its independence, and mostly - the question of ownership.

To the extent that the person who undertook the necessary arrangements for the creation of a computer-generated work is its author by virtue of s. 9(3) of the CDPA 1988, the person who undertook the arrangements for the development of AI that infringes copyright should be expressly held liable in case infringement is established.

Yet, part of the problems that inhere in AI is that it might be impossible to know that it infringed copyright or to be able to locate the creator/ owner of the AI apparatus.

**Question No. 3:**

Is there a need to clarify existing exceptions, to create new ones, or to promote licensing, in order to support the use of copyright works by AI systems? Please provide any evidence to justify this.

In the light of the exit of the UK from the European Union, there may be scope for the UK to legislatively innovate in this direction and expand the availability of permitted uses of copyright works by AI systems. That should primarily include AI uses in the context of digital humanities' research and other uses carried out for non-commercial purposes.

Supporting the use of copyright works by AI systems would require expanding the scope of the available exception on text mining and data analytics (s. 29, CDPA 1988) and the quotation

exception (s. 30, CDPA 1988). However, licensing may be a more appropriate solution with regard to commercial uses of protected works via AI systems.

By reference to AI-generated deepfakes, please refer to our answer to Question No. 1.

**Question No. 5:**

Should content generated by AI be eligible for protection by copyright or related rights?

The UK and Ireland are some of the few jurisdictions that offer express protection to so-called computer-generated works. Computer-generated works are defined to be works generated by computer in circumstances where the author of the work is not an individual (s. 178, CDPA 1988). S. 9(3) of the UK CDPA and s. 21(f) of the Irish Copyright Act clearly state the author of computer-generated works is the person who undertook the necessary arrangements for the creation of the work. The UK clearly excludes moral rights protection on computer generated works, although Ireland does not seem to include a similarly express provision.

At the time that the provisions were put in place, ownership of copyright in computer generated works was not in question. The reason was that the computer was an instrument in a creative process that was initiated by a human author. To that effect, declaring the person who undertook the arrangements to create a computer-generated work as the author of that work did not create problems. With the evolution of AI technology, AI tools will gradually become autonomous and not be dependent on human input. In cases that the process of creating a work takes place autonomy and without human intervention, s. 9(3) cannot apply as the requirement of a 'person by whom the arrangements necessary for the creation of the work are undertaken' is not met (see also s. 1(a)). With the expansion of the use of AI in the creation of copyright protected works, normative questions are also raised as to the subsistence of copyright in AI operated processes of creating works. One of them concerns the extent to which sufficient incentives for human creativity will exist as soon as machine creativity starts taking over. This is particularly in cases where mass produced works generated via AI could come in conflict with works created by human authors.

A way in which the issue of AI engagement in creative processes can be understood has been discussed at the EU level. The European Parliament in April 2020 issued a draft report arguing that 'certain works generated by AI can be regarded as equivalent to intellectual works and could therefore be protected by copyright'.<sup>3</sup> However, it recommends that 'ownership of rights be assigned to the person who prepares and publishes a work lawfully, provided that the technology designer has not expressly reserved the right to use the work in that way'.<sup>4</sup> The European Parliament's recommendation has raised a question: Should we look at a 'creative output' or 'a creative process' to judge whether works generated by AI are copyrightable?

Whereas works created through AI can in principle be protected by copyright in Ireland and the UK to the extent that human input is involved, compatibility of these provisions with the originality test is dubious. This CJEU-developed test (the 'author's own intellectual creation') looks both at the creative output and the creative process of making a work and, as a result, the extent to which protection can be offered to works generated via AI is doubtful. For these works to be a result of free and creative choices, stamping the *personal touch of the author* and not merely

---

<sup>3</sup> The European Parliament's Committee on Legal Affairs draft [report on intellectual property rights for the development of artificial intelligence technologies](#) (24 April 2020, Motion for a European Parliament Resolution, p. 6).

<sup>4</sup> Ibid.

following rules and instructions, there is need for human input in the creative process.<sup>5</sup> The personhood theories underpinning the ‘author’s own intellectual creation’ test cannot support copyright protection on AI creativity.

We submit that content generated by AI that does not fall under the definition of ss. 9(3) and 178 CDPA 1988 should not be eligible for protection by copyright as this would be in conflict with copyright principles (human authorship and originality test) and theoretical approaches to copyright (incentives, personhood etc).

**Question No. 6:**

If so, what form should this protection take, who should benefit from it, and how long should it last?

The draft report of the European Parliament in April 2020 recommends that ‘ownership of rights be assigned to the person who prepares and publishes a work lawfully, provided that the technology designer has not expressly reserved the right to use the work in that way’.<sup>6</sup> The European Parliament’s recommendation has raised a question: Should we look at a ‘creative output’ or ‘a creative process’ to judge whether works generated by AI are copyrightable? In responding to this question, it may be helpful to look at ‘a creative process’ which involves the human intention to generate and publish a creative output. So certain works generated by AI randomly and repeatedly should be excluded from copyright protection. The same ought to apply to AI created works where the creative process did not require human input.

The distinction between ‘creative output’ and ‘creative process’ is meaningful with regards to the question of authorship and ownership. Indeed, the identification of the ‘person who undertook the necessary arrangements for the creation of the work’ — at least in the context of computer-generated works — can result in controversy: is it the software developer (otherwise put the person who initiated the creative process), or the person who used or supervised the AI tool in order to produce the AI created output?

In cases of autonomous AI algorithms, the creative outcome may be the produce of the mere push of a button and in such cases when it comes to artificial intelligence algorithms that are capable of generating independent work often when the user’s contribution is the mere push of a button. In such cases the distinction between content and signal copyright could prima facie be relevant.<sup>7</sup>

We recommend the following:

- There is already legal protection for various facets of AI and copyright, and it is advisable to retain the status quo before introducing new legislation, expanding rights or launching sui generis rights.
- The development of AI technology is already subject to protection under copyright for computer programmes. It may also be subject to patent protection in some jurisdictions and can also receive protection — at least in part — as a trade secret. This protection is deemed appropriate as it can offer a holistic spectrum of proprietary entitlements. Some authors contemplate a sui generis protection for the protection of AI algorithms but there

<sup>5</sup> S Karapapa, *Defences to Copyright Infringement: Creativity, Innovation and Growth on the Internet* (OUP 2020) 56-57.

<sup>6</sup> The European Parliament’s Committee on Legal Affairs draft [report on intellectual property rights for the development of artificial intelligence technologies](#) (24 April 2020, Motion for a European Parliament Resolution, p. 6).

<sup>7</sup> R Arnold, ‘Content copyrights and signal copyrights: the case for a rational scheme of protection’ (2011) 1(3) QMJIP 272–279.

is no compelling argument why protection for software which implements AI should be stronger or different than the protection applicable to other computer programmes.

- The outputs of the creative process that have been produced as a result of human intervention can be subject to copyright protection, with a very limited term of protection, starting from the creation of the work.
- Although automated processes of producing works have been found on occasion to attract copyright protection (see e.g., *Cummins v Bond* [1927] 1 Ch. 167), AI outputs created *without human intervention* should not be subject to legal protection under copyright.

**Question No. 7:**

Do other issues need to be considered in relation to content produced by AI systems?

An important aspect to consider is human creativity and how that can remain incentivised if AI-related creativity becomes the more and more prominent. Relevant in this consideration is whether AI created works come in direct competition with the works that have been used, if at all, for the development of the AI algorithm. Another aspect concerns deepfakes, discussed under Question No. 1 (copyright).

**Question No. 8:**

Does copyright provide adequate protection for software which implements AI?

Please refer to our answer to Question No. 6 (copyright) above.

## Trade marks

**Question No. 1:**

If AI technology becomes a primary purchaser of products, what impact could this have on trade mark law?

The concepts of author, designer or inventor do not have an equivalent in trade mark law and therefore AI does not affect the trade mark system in the same way as it affects copyright, design rights and patents. That being stated, there are areas of trade mark law that are likely to be impacted by AI, which is increasingly used across a very wide range of industry sectors. For the purposes of determining trade mark registrability and infringement, current trade mark principles are largely based on twentieth-century concepts of how products are purchased. However, the emergence of e-commerce platforms and the increasing use of AI are likely to reduce, and potentially completely remove, the element of human interaction from the process of purchasing goods and services. If AI technology becomes a primary purchaser of products the decisive factors in purchasing products are likely to be product qualities (e.g., its specification, functionality etc), price, speed of delivery and user ratings. This means that established principles of trade mark law such as distinctiveness, average consumer, imperfect recollection and the multi-factorial test of likelihood of confusion as to the origin of the goods or services need to evolve and adapt to the new ways such goods and services are bought. What is more, there may be new types of trade mark infringement and unfair competition, e.g., in instances where similar (but not identical) products are placed under the same listing and ranked by order of product

price; see for instance *Jadebay Ltd v Clarke-Coles Ltd (t/a Feel Good UK)* [2017] EWHC 1400 (IPEC).

**Question No. 2:**

Are there, or could there be, any difficulties with applying the existing legal concepts in trade mark law to AI technology?

**Question No. 3:**

Does AI affect the concept of the “average consumer” in measuring likelihood of confusion?

Human engagement with messages conveyed by (or associated with) signs used in the purchasing process constitutes one of the foundations of trade mark law. However, consumers’ decision-making processes in online marketplaces are increasingly influenced by AI assistants, search engines and automated customer service chatbots, which can identify customer preferences and suggest products based on criteria like price, speed of delivery and availability. Product purchasing processes in online retail shopping are thus likely to transition from a reactive to a predictive mode, the radical extreme of which can see the consumer’s input being completely removed from the conventional purchasing cycle: products can be automatically ordered and delivered on a buyer’s behalf. An implication of such a development is that the “average consumer”, who is assumed to have an imperfect recollection of the relevant marks (such that the contesting signs are not compared side by side; C-342/ 97, *Lloyd Schuhfabrik*) and whose level of attention is likely to vary according to the category of goods or services in question, may ultimately be taken over by an intelligence agent which does not display the inherent flaws of human nature, currently reflected into trade mark law principles. It is not therefore unrealistic to suppose that this long-held concept of the average consumer may become less relevant in the age of AI or even abandoned (in online retailing) in favour of a more sophisticated AI-based tool with enhanced analytical capabilities that help provide more granularity and heightened accuracy in the selection between products or services. It is expected that these enhanced analytical capabilities will help combating unfair competitive practices in marketing approaches or setting prices.

Moreover, the functions, law and practice around trade marks need to be reconsidered in light of the proliferation of AI used by consumers in the context of Internet of Things (IoT) applications. As technology is rapidly moving towards voice commerce, the rising use of voice assistants (e.g., Siri, Alexa and Cortana)<sup>8</sup> is likely to disrupt the analytical framework for comparing trade marks (sight, sound and meaning analysis; see C-251/ 95, *Sabèl*). The phonetic comparison of marks is expected to gain dominance, while the visual and conceptual aspects could be relegated in the process (though not completely eliminated, as retail sales will continue offline too). So, when looking at the type of purchase, for example, a comparison of trade marks must integrate (among other established considerations) an assessment not only of the way a particular public pronounces a mark, but also how an AI assistant interprets this (and perhaps what steps it is programmed to take in order to minimise the likelihood of confusion between similarly named or look-alike products e.g., by seeking confirmation or clarification from the consumer). In the longer run, AI visual comparisons may be feasible via optical recognition technologies used in IoT applications and machine learning may enable conceptual comparisons between products on the basis of product characteristics, user feedback and relevant parameters.

---

<sup>8</sup> A study from Juniper Research estimated that digital voice assistants will triple in use by 2023 surging to over 8 billion device; see <<https://www.juniperresearch.com/press/press-releases/digital-voice-assistants-in-use-to-8-million-2023>> accessed 13 November 2020.



Based on the above, there seems to be a need to diversify the current legal principles for trade mark comparison to take account of who/ what is involved in the purchasing decision-making process. Where AI takes control (or at least materially influences this process) and a human being plays a passive role in assessing a product, then some of the habitually relevant factors concerning the likelihood of confusion (e.g., the degree of attention paid by the relevant public to those goods and services, the degree of similarity between the signs, and whether the impression produced by any one of the levels of comparison is more important, and the distinctiveness of the earlier mark etc.) could lose their significance (or even become redundant). When, however, a human being takes an active role in the purchasing decision by instructing the AI assistant which product or service to order, the element of confusion is likely to remain relevant. In cases involving a combination of human interaction and machine intelligence, the aforementioned concepts would not lose their salience and a degree of responsibility for accurate product searching would rest with the consumer.

**Question No. 5:**

Can the actions of AI infringe a trade mark?

**Question No. 6:**

If AI can cause trade mark infringement, does this change who could be liable? Should it be the owner, the operator, the programmer, the trainer, the provider of training data, or some other party?

Being a statistical engine that produces probabilistic results, the proposition that AI cannot get confused does not withstand close scrutiny. It is not unfathomable that it can make incorrect predictions, depending on the nature of analysis in algorithmic decision making. When recommendations by AI result in trade mark infringement, with whom does the responsibility for AI's actions ultimately lie? The parameters of the purchasing decision, as they are provided by the algorithm in question, will need to take centre stage and, as a consequence, the courts will need to engage in a thorough review of the performance and robustness of algorithms supporting the operation of AI applications. Previous CJEU rulings offer some guidance which can be transposed in the field of trade marks and AI, and particularly the context of liability of AI applications. Decisions such as C-236/ 08, Google France and C-324/ 09, L'Oréal suggest that the provider of an AI application may be held liable for infringement on their platform(s) if they have played an active role in the promotion/ sale of the trademarked goods (i.e., when they provide assistance which entails optimising the presentation of the online offers for sale or employ sophisticated algorithms that perform market monitoring activities picking up infringements) or gained knowledge of circumstances that should have put them on notice that the offers for sale were unlawful and, in the event of it being so aware, failed to promptly remove the infringing information from their data sets (or otherwise disable access to it).

That being stated, it is anticipated that AI will also be used to monitor online product listings and detect infringement that can take place, for instance, via the sale of counterfeit goods.

# Trade secrets

## **Question No. 3:**

What are the advantages and disadvantages of using trade secrets in the AI sector? Could information that is not shared inhibit AI development?

Trade secrets offer optimal protection in technologies which are characterised by rapid development and are difficult to detect, develop independently or reverse engineer. Because of the sophistication, depth and complexity of AI systems, AI-enabled technologies lend themselves well for protection by confidentiality/ trade secrets. Trade secrets generally protect broader subject matter than other formal IP rights, covering technical information (e.g., software, manufacturing processes), valuable know-how and even facts (e.g., client lists, supplier or distribution lists). The indefinite protection of confidentiality and trade secrets is in contrast to other forms of IP such as patents (20 years) and copyright (variably 50, 70 or 100 years after the death of the work's creator). There is potentially no geographic limitation so long as the information remains a secret in every relevant territory. This is another advantage over registrable forms of IP, for which there is a cost attached for each country where protection is sought. Trade secrets can also be protected without any procedural formalities or need for governmental/ regulatory approval.

That being stated, it cannot be said that protecting trade secrets is free, as there will usually be costs associated with the implementation of safety or control measures and information protection policies, including physical means of protecting information. It is likely that there will also be a substantial litigation cost if the confidentiality of the information at issue needs to be tested in court. Moreover, lack of registration means that trade secrets are not always well-defined. Claimants may struggle to identify what aspects of a large amount of technical information is secret/ confidential. Importantly, trade secrets do not give their owners a monopoly over the subject of the trade secret. Thus, ownership of a trade secret is only protected against improper acquisition, use or disclosure and does not afford a right to prevent someone else from discovering that information by their own independent effort. As soon as a trade secret is leaked (accidentally or by malice), its secret status is lost and so a trade secret or know-how may become worthless. A patent may also be applied for by a third party precisely because the information is not in the public domain. Another disadvantage of using trade secrets in the AI sector concerns decision making processes and the extent to which certain issues can remain subject to strong trade secret protection, e.g., the parameters used to develop an AI algorithm and/ or use of (anonymised) personal data. Finally, a claim cannot be brought simply because a piece of confidential information is no longer confidential. There needs to be evidence that the disclosure has caused, or is certain to cause, actual damage to a claimant's economic interests. This may not be always straightforward.

The proposition that trade secrets may hinder innovation is not new, but this concern is perhaps intensified in the context of AI technologies. Whilst copyright and patents reward disclosure of creativity and innovation to society, trade secrets are conditioned on secrecy and thus can be seen as breaking this virtuous circle by indirectly propelling 'hidden' or invisible innovation development. This could mean that valuable ideas with the capacity to result in significant AI advances may not be shared. However, trade secrets can also be seen as another form of incentive to invention. Because of the advantages associated with trade secrets (notably their wider definition compared to the definition of patentable subject matter) they arguably reach into a number of corners patent law cannot. In addition, trade secret law serves as a substitute for the physical and contractual restrictions that companies would otherwise have to impose for fear of

losing the competitive edge their secrets provide.<sup>9</sup> Companies may be more willing to contract with third parties for manufacturing or developing AI-related products and solutions, if they feel confident that adequate levels of protection and effective means of redress are in place, in case their trade secrets are misused. One proposition is to concentrate efforts on developing stronger legal protection and more robust enforcement of trade secrets. This would facilitate the commercialisation (e.g., acquiring, sharing or licensing) of AI technology and improvement of secret AI-related inventions, further encouraging the flow of information and collaboration between potential business partners, researchers and innovators across different countries.

**Question No. 4:**

Do trade secrets cause problems for the ethical oversight of AI inventions?

AI is increasingly becoming more and more embedded in our everyday lives. It is being deployed in healthcare; it scrutinises CVs and judges our creditworthiness. However, AI is not ready to replace human cognitive capabilities for many tasks and is not necessarily immune from error. Evidence suggests that AI can be demanding, fragile and it may also ‘bake in and scale human and societal biases’.<sup>10</sup> The fallibility of human cognition can intentionally or unintentionally influence AI through algorithms, interaction and underlying data, the inclusion or exclusion of which may disadvantage certain groups of people. As AI algorithms grow more advanced and are deployed into critical domains, it is important that humans are able to understand how algorithmic decisions have been reached.<sup>11</sup> In order to build trust in the use of AI technologies, potential trade-offs might need to be carefully weighed between individuals’ right to know that AI algorithms are fairly applied and companies’ legitimate interests in protecting their proprietary information through trade secrets. One way of tackling this challenge would be imposing proportionate obligations on developers of AI applications in the interests of legitimate public policy considerations. For instance, suitable and accessible disclosures of the extent to which AI shapes an organisation’s decision-making process, the rationale for deploying an AI system and its design parameters should be made available whenever an AI system has a significant impact on people’s lives or where accidents might have occurred. Another way in which ethical oversight of AI inventions can be ensured would be through record-keeping obligations and transparency.

---

<sup>9</sup> M Lemley, ‘The Surprising Virtues of Treating Trade Secrets as IP Rights’ (2008) 61(2) Stanford Law Review 311.

<sup>10</sup> J Silberg and J Manyika, ‘[Notes from the AI frontier: Tackling bias in AI \(and in humans\)](#)’ (McKinsey Global Institute, June 2019); see also T Panch, H Mattie and R Atun, ‘Artificial intelligence and algorithmic bias: implications for health systems’ (2019) 9(2) Journal of Global Health DOI:10.7189/jogh.09.020318

<sup>11</sup> See further the key requirements proposed by the High-Level Expert Group in the Ethics guidelines for trustworthy AI: European Commission, [Ethics Guidelines for Trustworthy AI](#).