PRIVACY INTERNATIONAL

Submission of evidence to the House of Lords Select Committee on Artificial Intelligence

6 September 2017

Submission of evidence to the House of Lords Select Committee on Artificial Intelligence

September 6, 2017

Frederike Kaltheuner, Programme Lead and Policy Officer, Privacy International
frederike@privacyinternational.org

Dana Polatin-Reuben, Technology Officer, Privacy International
dana@privacyinternational.org

Statement of interest

Privacy International welcomes the opportunity to respond to this inquiry by the House of Lords Select Committee on Artificial Intelligence ('AI'). Privacy International is a non-profit, non-governmental organisation based in London, dedicated to defending the right to privacy around the world. Established in 1990, Privacy International undertakes research and investigations into government surveillance and data exploitation in the private sector with a focus on the technologies that enable these practices. To ensure universal respect for the right to privacy, Privacy International advocates for strong national, regional and international laws that protect privacy around the world. It has litigated or intervened in cases implicating the right to privacy in the courts of the United States, the UK, and Europe, including the European Court of Human Rights and the European Court of Justice. It also strengthens the capacity of partner organisations in developing countries to identify and defend against threats to privacy. Privacy International employs technologists, investigators, policy and advocacy experts, and lawyers, who work together to understand the technical underpinnings of emerging technology and to consider how existing legal definitions and frameworks map onto such technology.

The pace of technological change

What is the current state of artificial intelligence and what factors have contributed to this? How is it likely to develop over the next 5, 10 and 20 years? What factors, technical or societal, will accelerate or hinder this development?

1. Artificial intelligence (AI), or intelligent systems which can act without being specifically programmed to follow certain steps or instructions[1], is a term that is often used to refer to a diverse range of applications and use-cases at different levels of complexity and abstraction. The term is employed to encompass everything from machine learning which makes inferences, predictions, and decisions about individuals, and other domain-specific AI algorithms, to fully autonomous and connected objects, as well as the futuristic idea of Singularity. This lack of definitional clarity is a challenge, since different types of AI and different domains of application raise specific ethical and regulatory issues.

2. The most widespread AI methods are collectively known as machine learning, which undergirds everything from text auto correction to drone targeting systems. Machine learning uses algorithms trained with vast amounts of data to improve a system's performance at a task over time. Tasks often involve making decisions or recognising patterns, with many different possible outputs in a range of domains and applications.

3. As an organisation which works on the right to privacy, we are primarily concerned about current and future applications of AI that are designed for the following purposes: (1) to identify and track individuals; (2) to predict or evaluate individuals or groups and their behaviour; (3) to automatically make or feed into consequential decisions about people or their environment; and (4) to generate, collect and share data.

4. **AI applications can be used to identify and thereby track individuals across different devices, in their homes, at work and in public spaces.** For example, while personally identifiable information (PII) is routinely anonymised within datasets, AI can be employed to de-anonymise this data, complicating the distinction between PII and non-PII data on which current data protection regulation is based.

5. **Using machine learning methods, highly sensitive information can also be inferred, or predicted from non-sensitive forms of data.** As a result of such profiling, databases that merely contain data about an individual's behaviour can be used to generate unknown data about their likely identity, attributes, interests, or demographic information. Such predictions may include information about health, political opinions, sexual orientation, or family life.

6. **AI systems can be used to make or inform consequential decisions about people or their environment.** Automated decision-making that relies on AI also plays a role in the personalisation of information and experiences, from news feeds to targeted advertising and recommendation systems. Such personalisation gears information towards individuals' presumed interests or identities, which are derived through profiling.

7. AI-driven consumer products, from smart home appliances to phone applications, are often built for data exploitation. **Consumers are commonly faced with an informational asymmetry as to what kinds of data and how much data their devices, networks, and platforms generate, process, or share**. As we bring ever more smart and connected devices into our homes, workplaces, public spaces and onto our bodies, educating the public about such data exploitation becomes pressing.

8. **These applications of AI have the potential to undermine fundamental rights and liberties, including the rights of privacy, freedom of expression and assembly, and raise very serious concerns surrounding discrimination.**

[1] Negnevitsky, M., 2005. *Artificial intelligence: a guide to intelligent systems*. Pearson Education.

9. **They also have the potential to transform society as we know it.** Today, AI CCTV security systems can classify people, follow them through a crowd and detect 'suspicious behaviour'[2]; tomorrow, CCTV cameras and drones may be able to transcribe conversations through lip reading.[3] Today, insurance companies analyse how many exclamation points we use in social media posts to determine whether we are a safe driver[4]; tomorrow, marketers could assess our credit worthiness from objects and facial expressions in the pictures we share on social media platforms.

Is the current level of excitement which surrounds artificial intelligence warranted?

10. AI, if implemented responsibly, can have many exciting impacts on society. AI systems could improve crop yield in large-scale farming by tracking potential issues such as pests[5], and interactive robots are already improving the social skills of people on the autism spectrum[6].

11. Privacy International is not against the use of artificial intelligence; however, as is the case with most emerging technologies, there is a very real risk that commercial and government uses of AI fall into the trap of technological solutionism – the urge to fix problems that don't exist, or for which there is no technological solution, or for which a technological solution will exacerbate existing problems and fail to address underlying issues.

Impact on society

How can the general public best be prepared for more widespread use of artificial intelligence?

12. **Novel applications and recent advances in artificial intelligence could negatively affect the right to privacy. This is significant since privacy is the lynchpin of indispensable individual values** such as human dignity, personal autonomy, freedom of expression, freedom of association, and freedom of choice,[7] as well as broader societal norms.[8]

13. The privacy implications of AI stem from its ability to recognise patterns and increasingly "derive the intimate from the available"[9]. AI methods are being used to identify people who wish to remain anonymous; infer and generate sensitive information about people from their non-sensitive data; profile people based upon population-scale data; and make consequential decisions using this data which profoundly affect people's lives.

14. For instance, machine learning systems have been able to identify about 69% of protesters who are wearing caps and scarves to cover their faces.[10] FindFace, a Russian face recognition application launched in early 2016, allows users to photograph people in a crowd and compares their picture to profile pictures on the popular social network VKontakte, identifying their online

[2] Toomey, M., 2017, Hitachi built an AI security system that follows you through a crowd. *Quartz*. Available from: https://qz.com/958467/hitachi-built-an-ai-security-system-that-follows-you-through-a-crowd/. [Accessed 1st August 2017]
[3] Morgan, T., 2016, Lip-reading technology breakthrough to be used on CCTV. *The Telegraph*. Available from: http://www.telegraph.co.uk/news/2016/03/25/lip-reading-technology-breakthrough-to-be-used-on-cctv/. [Accessed 1st August 2017]
[4] Ruddick, G., 2016, Admiral to price car insurance based on Facebook posts. *The Guardian*. Available from: https://www.theguardian.com/technology/2016/nov/02/admiral-to-price-car-insurance-based-on-facebook-posts. [Accessed 1st August 2017]
[5] McFarland, M., 2017, Farmers turn to artificial intelligence to grow better crops. *CNN*. Available from: http://money.cnn.com/2017/07/26/technology/future/farming-ai-tomatoes/index.html. [Accessed 1st August 2017]
[6] Available from: https://robots4autism.com. [Accessed 1st August 2017]
[7] Payton, T. and Claypoole, T., 2014. *Privacy in the age of Big data: Recognizing threats, defending your rights, and protecting your family*. Rowman & Littlefield.
[8] Post, R.C., 1989. The social foundations of privacy: Community and self in the common law tort. *California Law Review*, pp.957-1010. Summarizing Post *see* Doyle, T., 2012. Daniel J. Solove, Nothing to Hide: The False Tradeoff between Privacy and Security. ("As the legal theorist Robert Post has argued, privacy is not merely a set of restraints on society's rules and norms. Instead, privacy constitutes a society's attempt to promote civility. Society protects privacy as a means of enforcing order in the community. Privacy isn't the trumpeting of the individual against society's interests but the protection of the individual based on society's own norms and values").
[9] Calo, R., 2017. Artificial Intelligence Policy: A Roadmap.https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3015350
[10] Singh, A., Patil, D., Reddy, G.M. and Omkar, S.N., 2017. Disguised Face Identification (DFI) with Facial KeyPoints using Spatial Fusion Convolutional Network. *arXiv preprint arXiv:1708.09317*. ACM. https://arxiv.org/pdf/1708.09317.pdf

profile with 70% reliability.[11] The technology has also been used to identify the real names of sex workers in adult films.[12]

15. A 2015 study by researchers at the French Institute for Research in Computer Science showed that 75% of mobile phone users can be re-identified within a dataset using machine learning methods and just two smartphone apps, with the probability rising to 95% if four apps are used.[13]

16. Emotional states, such as confidence, nervousness, sadness, and tiredness, for instance, can be predicted from typing patterns on a computer keyboard.[14] The Big-Five personality traits (extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience) can be predicted from standard mobile phone logs.[15] In 2012, Cambridge researchers used predictive modelling to analyse a dataset of Facebook Likes, demographic profiles, and psychometric tests from 58,000 Americans. From the Likes data, the model could discriminate between heterosexual and homosexual men in 88% of cases; African Americans and Caucasian Americans in 95% of cases; and Democrats and Republicans in 85% of cases.[16]

17. **While such profiling using machine learning can be highly privacy-invasive, there is also no guarantee that the profile that is created in the process is even accurate, given that machine learning methods are inherently probabilistic.** Poor quality data, or systematically biased data are common concerns. Yet, even if profiling was based on perfect data, individuals could still be misclassified, misidentified or misjudged, and such errors may disproportionately affect certain groups of people (see our response to the next question).

18. Profiling, whether it relies on complex machine learning or more straightforward methods, merely determines that an individual is *highly likely* to be female, *likely* to be unworthy of credit, or *unlikely* to be married, homosexual or an introvert. **Since individuals are often unaware about the fact that they are being profiled, it can be difficult to challenge or correct inaccurately inferred or predicted information.** Do we want to rely on probabilistic knowledge to make decisions about life or death? And do we feel comfortable using uncertain and possibly discriminatory inferences to limit an individual's freedom?

Who in society is gaining the most from the development and use of artificial intelligence and data? Who is gaining the least? How can potential disparities be mitigated?

19. **AI's benefits and harms are currently distributed unequally.** Industry gains most from AI, with large tech companies (and selected government agencies) having unprecedented access to vast troves of data on billions of people around the world. Consumers and citizens are frequently unaware about the scope, granularity, and sensitivity of data that third parties hold about them, or that their data is being used to train and develop AI systems.

20. **Risky applications of AI often disproportionately affect those that are already most vulnerable in society.** A good example is AI-driven automated decision-making in hiring. Highly skilled job seekers have the ability to demonstrate their skills and character in a personal interview, while the low-wage sector with high turnover increasingly relies on automated and often proprietary and opaque hiring software that may rely on poor quality or inaccurate data and produce biased, inaccurate, discriminatory or unfair decisions. Such selective reliance on AI-driven decision-making is also evident in in policing. While predictive policing is becoming

[11] Available from: https://www.theguardian.com/technology/2016/may/17/findface-face-recognition-app-end-public-anonymity-vkontakte . [Accessed 1st August 2017]
[12] Available from: http://www.newsweek.com/porn-actress-facial-recognition-findface-sex-worker-453357. [Accessed 1st August 2017]
[13] Achara, J.P., Acs, G. and Castelluccia, C., 2015, October. On the unicity of smartphone applications. In *Proceedings of the 14th ACM Workshop on Privacy in the Electronic Society* (pp. 27-36). ACM. https://arxiv.org/pdf/1507.07851v2.pdf
[14] Epp, C., Lippold, M. and Mandryk, R.L., 2011, May. Identifying emotional states using keystroke dynamics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 715-724). ACM. http://hci.usask.ca/uploads/203-p715-epp.pdf
[15] de Montjoye, Y.A., Quoidbach, J., Robic, F. and Pentland, A., 2013, April. Predicting Personality Using Novel Mobile Phone-Based Metrics. In *SBP* (pp. 48-55). https://link.springer.com/content/pdf/10.1007/978-3-642-37210-0.pdf#page=63
[16] Kosinski, M., Stillwell, D. and Graepel, T., 2013. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), pp.5802-5805. http://www.pnas.org/content/110/15/5802.full#F1

increasingly common in UK law enforcement, it is predominantly used to fight street-level crime, rather than white collar crime such as tax evasion or fraud.

21. **Finally, AI systems can contribute to the perpetuation of existing injustices and inequalities in society through inbuilt bias and discrimination**. In the United States, risk assessment software purporting to predict the likelihood of reoffending has been used to aid sentencing decisions since the early 2000s. A 2016 study by the non-profit news organisation ProPublica revealed this software's bias against African-Americans, who are more likely to be given a higher risk score compared with white offenders charged with similar crimes.[17] Another important case is facial recognition software. The US House Committee on Oversight and Government Reform found that the FBI facial recognition database contains photos of half of US adults without consent, and the algorithm is not only wrong nearly 15% of time, but is also more likely to misidentify black people.[18]

22. Machine learning can unintentionally, indirectly, and often unknowingly recreate discrimination from past data. Since profiling using machine learning can create uncannily personal insights, there is a risk of it being used against those who are already marginalised. Even if data controllers take measures to avoid sensitive attributes in automated processing, trivial information can correlate with sensitive information, potentially leading to illegal but indirect discrimination. In racially segregated cities, for instance, postcodes may be a proxy for race. Therefore, without explicitly identifying a data subject's race, profiling may nonetheless identify attributes, or other information that would lead to discriminatory outcomes, if they were to be used to inform or make a decision.

23. Machine learning can also lead to "rational discrimination" – when data analysis finds an accurate correlation, that society nonetheless would consider discriminatory. An example would be if an algorithm found that men are less reliable in paying back loans, and hence their interest rate should be higher. Would we want to discriminate based on gender? And finally, there is simply unfairness, which might not be illegal, but could nonetheless be seen as unfair. If, for instance, a hiring software based on machine learning concludes that users of Internet Explorer are less qualified candidates[19], we could consider this unfair.

Public Perception

Should efforts be made to improve the public's understanding of, and engagement with, artificial intelligence? If so, how?

24. **In the public imagination, AI is always something that isn't quite there yet, that is embodied, futuristic, but not yet widespread. This need to change.** This misconception risks steering the focus of regulatory discussions on speculative technologies that have yet to be implemented on a mass scale, if at all.

25. In particular, we find that **the public's understanding of AI to identify individuals across devices and in public space, and to gain highly sensitive insights from everyday traces of data, is low.**[20] Since *informed* consent is one legal ground for the processing of personal data, this lack of understanding raises concerns.

26. **Similar challenges apply to the privacy and security risks of AI-driven consumer products**. A good example is iRobot, the Roomba robotic vacuum. The product's chief executive suggested that the company might begin selling floor plans of customers' homes,

[17] Angwin, J., Larson, J., Mattu, S and Kirchner, L., 2016, Machine Bias. *ProPublica.* Available from: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing. [Accessed 1st August 2017]
[18] See https://oversight.house.gov/newsarticle/facial-recognition-database-used-fbi-control-house-committee-hears/
[19] 2013. Robot Recruiters. *The Economist.* Available from: https://www.economist.com/news/business/21575820-how-software-helps-firms-hire-workers-more-efficiently-robot-recruiters. [Accessed 1st August 2017]
[20] The Royal Society, 2017, Machine learning: the power and promise of computers that learn by example. *Royal Society*. Available from https://royalsociety.org/~/media/policy/projects/machine-learning/publications/machine-learning-report.pdf. [Accessed 1st August 2017]

derived from the movement of their autonomous cleaner, to Amazon, Apple, and Google Alphabet. [21]

27. Relevant actors, including the government, the EU Commission [the European Data Protection Board], supervisory authorities and civil society must design and develop a plan to educate data subjects and consumers about the various ways in which their data is being used by data controllers. They must also be made aware of how to gain information about processing of their data, how to exercise their rights in relation to such processing, and how to obtain redress, which requires effective implementation and enforcement of the rights of data subjects as set out in the upcoming General Data Protection Regulation (GDPR) and any related UK legislation.

Ethics

What are the ethical implications of the development and use of artificial intelligence? How can any negative implications be resolved?

28. **AI-driven applications sort, score, categorise, assess, and rank people, often without their knowledge or consent.** We have already mentioned the privacy implications of this, but it is important to stress that other human rights are affected as well. This view is echoed by the United Nations Human Rights Council, which on 22 March 2017 noted with concern "that automatic processing of personal data for individual profiling may lead to discrimination or decisions that otherwise have the potential to affect the enjoyment of human rights, including economic, social and cultural rights".[22]

29. When all data, from how we fill out a form, to our location data can be used to gain even more intimate details about our lives and make consequential decisions, from access to credit and insurance to policing, this might result in widespread chilling effects. Individuals might pre-emptively self-censor their speech and behaviour, if the data it generates might be used against them.

30. AI also plays a role in personalisation of information, products, and experiences. By excluding content deemed irrelevant or contradictory to the user's beliefs or presumed interests, such forms of personalisation may reduce the diversity of information users encounter.[23] Personalisation of not just information but also our perception of the world around us will become increasingly important as we move towards connected spaces, like smart cities, as well as augmented and virtual reality. An environment that knows your preferences and adapts itself according to these presumed interests would be highly personalised, but would also raise important questions around autonomy and the ethics of subtle manipulation.

In what situations is a relative lack of transparency in artificial intelligence systems (so-called 'black boxing') acceptable? When should it not be permissible?

31. We would like to draw the Committee's attention to the work of Jenna Burrell[24], who distinguishes between three forms of opacity: (1) opacity as intentional corporate or state secrecy; (2) opacity as technical illiteracy; and (3) an opacity that arises from the characteristics of machine learning algorithms and the scale required to apply them usefully. Only the latter implies that the system's outcomes are not predictable by its designer, whereas users, regulators or the general public will find all the instances opaque. We are most concerned about highly complex AI systems that have the potential to produce harmful or dangerous outcomes that are *neither predictable by their designer nor easily discoverable by the public*.

32. **Black boxing should not be permissible wherever AI systems are used to make or inform consequential decisions about individuals or their environment**; in such instances, a lack

[21] Hern, A., 2017, Roomba maker may share maps of users' homes with Google, Amazon or Apple. *The Guardian*. Available from: https://www.theguardian.com/technology/2017/jul/25/roomba-maker-could-share-maps-users-homes-google-amazon-apple-irobot-robot-vacuum. [Accessed 1st August 2017]
[22] U.N. Human Rights Council Resolution on the Right to Privacy in the Digital Age, U.N. Doc. A/HRC/34/7, 23 Mar. 2017, para.2
[23] Pariser, E., 2011. *The filter bubble: What the Internet is hiding from you*. Penguin UK.
[24] Jenna Burrell, *supra note* 7.

of transparency is highly problematic. Consequential decisions are decisions that produce irreversible effects, or effects that can significantly affect an individual's life or infringe on their fundamental and human rights.

33. We would like to stress that we **are equally concerned about opaque AI systems that *automatically make* and those that *inform* decisions**, that is decisions that are formally attributed to humans but are *de facto* determined by an opaque AI system. A good example is the use of automated risk scores in the criminal justice system. Proprietary software, such as the COMPAS risk assessment that was sanctioned by the Wisconsin Supreme Court in 2016,[25] calculates a score predicting the likelihood of committing a future crime. Even if final decisions are made by a judge, the software's automated decisions can be decisive, especially if judges rely on them exclusively or haven't been warned about their risks, including that the software may produce inaccurate, illegal, discriminatory, or unfair decisions.

34. **It is also crucial to define what kind of remedies different stakeholders require.** Individuals should be provided with sufficient information to enable them to fully comprehend the scope, nature, and application of AI, in particular with regards to what kinds of data these systems generate, collect, process, and share. When AI algorithms are used to generate insights or make decisions about individuals, users as well as regulators should be able to determine how a decision has been made, and whether the regular use of these systems violates existing laws, particularly regarding discrimination, privacy, and data protection. Governments and corporations who rely on AI should publish, at a very minimum, aggregate information of the kind of systems being developed and deployed.[26]

The role of the Government

What role should the Government take in the development and use of artificial intelligence in the United Kingdom? Should artificial intelligence be regulated? If so, how?

35. The question of whether artificial intelligence *can* or *should* be regulated is complicated by the fact that artificial intelligence lacks a stable, consensus definition or instantiation.[27] Furthermore, an identical AI application can raise different regulatory and ethical concerns, depending on the domains in which it is employed.

36. Take for instance "SKYNET", a programme by the United States National Security Agency (NSA) which reportedly collects in bulk the metadata communication of the entire Pakistani mobile phone network, and then uses a random forest machine learning algorithm to rate "each person's likelihood of being a terrorist".[28] The insights and classifications that machine learning generates are inherently probabilistic – there are always false positive and false negatives. But the implications of this are vastly different, depending on where exactly machine learning is being applied. An exceptionally low false positive rate is remarkable in business applications, such as targeted advertising. In the case of government surveillance, however, even an error rate as low as "0.008 percent of the Pakistani population" still corresponds to 15,000 people potentially being misclassified as "terrorists".[29]

37. **What clearly should be regulated is the following: the data that feeds into AI systems; the data (and insights) that AI systems generate; as well as how and whether AI systems should be used to make or inform consequential decisions about individuals and groups, especially if these systems are highly complex and opaque.**

[25] Citron, D., 2016, (Un)Fairness of Risk Scores in Criminal Sentencing. *Forbes*. Available from: https://www.forbes.com/sites/daniellecitron/2016/07/13/unfairness-of-risk-scores-in-criminal-sentencing/#6074794b4ad2. [Accessed 1st August 2017]
[26] *Cf.* Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, U.N. Doc. A/HRC/23/40, paras. 91-92 (17 April 2013).
[27] Calo, R., 2017. Artificial Intelligence Policy: A Roadmap. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3015350
[28] For more information, *see* Cole, D., 2014. We kill people based on metadata. *The New York Review of Books*, *10*, p .2014.; Grothoff, C. and Porup, J., 2016. The NSA's SKYNET program may be killing thousands of innocent people. *Ars Technica.*, *available at* https://arstechnica.co.uk/security/2016/02/the-nsas-skynet-program-may-be-killing-thousands-of-innocent-people/.
[29] ibid.

38. While data is central to the development of AI, in particular machine learning, governments and regulators have a responsibility to ensure that the current excitement about AI does not become a pretext for exploiting people's data without their knowledge or unambiguous and informed consent, for processing purposes that are often unexpected and may result in tangible harm.[30] We would like to draw the Committee's attention to principles such as "data minimisation", "privacy and security by design", as well as "purpose limitation" that are designed to mitigate the power imbalance between data controllers and data subjects.

39. The upcoming GDPR contains provisions that specifically address profiling and automated individual decision-making. These are necessary but not sufficient to address all privacy concerns of AI. However, a number of viable expressions in the GDPR are unclear or ambiguous, which may lead to confusion, enforcement gaps or asymmetries. We encourage the government to support additional guidance that clarifies ambiguous terms in a way that guarantees the strongest protections for data subjects.

40. We strongly believe that civil society organisations should be able to investigate and lodge complaints independently or on behalf of data subjects if processing is unlawful. There is urgent need for clear EU-wide guidelines on how to claim redress in front of national supervisory authorities or national courts for violations of their rights in relation to profiling, AI, and the use of machine-learning algorithms.

[30] See for instance Hill, K., 2016, This sex toy tells the manufacturer every time you use it. *Fusion*. Available from: http://fusion.kinja.com/this-sex-toy-tells-the-manufacturer-every-time-you-use-1793861000. [Accessed 1st August 2017]