

The need for an AI Bill of Rights through the case of the metaverse

Stefanos Kyprianou

The development of Artificial Intelligence (AI) is a pivotal point in the history of technological development. The efficiency of AI algorithms in completing tasks makes them valuable for businesses and institutions which seek solutions to reduce the cost and time of processes or enhance their monetisation systems and results.¹ However, the power of these algorithms raises concerns regarding their exploitation or employment for nefarious purposes. Considering that AI is being employed as a primary tool in technological advancements such as the virtual reality/metaverse, governing the use of AI is arguably essential to prevent the exacerbation of existing AI-related issues and protect citizens. Thus, governments and institutions are currently drafting regulatory frameworks to govern their use while also balancing their priorities. More particularly, after BREXIT, the UK places innovation and growth at the forefront of their national strategies on AI. This could undermine the protection of human rights. Hence, this essay will support the need for an AI bill of Rights that facilitates safe innovation and place human rights at the forefront. It will illustrate the need by examining potential risks and concerns that arise from the use of AI in the metaverse. Initially, an assessment will be conducted on the National AI strategy, which illustrates the UK's pro-innovation approach for AI regulation. Then, the focus will turn to the human-centred approaches other jurisdictions support, such as the US AI bill of rights. Subsequently, the need for a UK AI bill of rights will be exemplified through the potential effects of the use of AI in the metaverse. The essay will notably examine the effects of using AI in the metaverse for monitoring and manipulating users. Eventually, the need for an AI bill of rights will be evident through the examined effects on citizens.

The UK National AI strategy and the AI bill of rights

¹ Bernd Carsten Stahl, *Artificial Intelligence for a Better Future An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies* (Springer, 18 March 2021)

AI is becoming an integral part of various technological advancements and fields. Due to its importance in modern society and economy, countries seek to promote AI-based innovation through the appropriate frameworks. Most notably, the UK government drafted and published its National AI strategy, which aims to illustrate its vision of creating a trust framework in which innovation and opportunity are at the forefront of any standard and regulation.² They acknowledged that these technologies are disrupting the established norms and rules while also raising concerns for potential risks and harms towards citizens. They highlighted specific concerns regarding the risk such systems might pose to privacy and human agency, along with potential physical or economic harms that might be enabled or exacerbated by AI.³ Therefore, the government emphasised the need for adopting a “secure by design” approach to mitigate risks. Furthermore, they supported the need for the creation of sector-specific legislation as a model of governing the use of AI systems in an efficient way.⁴ It can be argued, though, that such legislation would require careful consideration of other overlapping regulatory frameworks.

Evidently, the UK supports legislation that governs holistically the use of AI to protect citizens while having innovation at its forefront. Meanwhile, the USA’s White House proposed an equivalent regulatory framework through a document opening a discussion for a potential Bill of Rights in an AI-Powered World.⁵ Some discussions about the proposed framework included the right for a meaningful recourse in the case of harm by an algorithm’s recommendation and the right to know when AI systems are affecting one’s civil liberties. The most critical discussions, though, were regarding the rights to freedom from surveillance and freedom from being subjected to AI-based decision-making systems which are made from biased data sets.⁶ Evidently,

² Emre Kazim, Denise Almeida, Nigel Kingsman, Charles Kerrigan, Adriano Koshiyama, Elizabeth Lomas, Airlie Hilliard, *Innovation and opportunity: review of the UK’s national AI strategy* (06 December 2021) *Discover Artificial Intelligence* Vol.1, No.1, pp.1-10, 2

³ Secretary of State for Digital, Culture, Media and Sport, *National AI Strategy* (September 2021)

⁴ *Ibid*,52

⁵ Glenn Glow, *The AI Bill Of Rights: Protecting Americans From The Dangers Of Artificial Intelligence* (Forbes, 9 January 2022)

⁶ *ibid*

such discussions, along with proposals such as the European AI Act and the Chinese New Generation Artificial Intelligence Development Plan, illustrate the intention of governments to regulate AI.⁷ Therefore, further exploring that intention vital to create a unified approach on a regulatory framework such as an AI Bill of Rights.

Meanwhile, it is critical to highlight the distinction between the pro-citizen approach adopted by the US in the AI bill of rights proposal and the UK pro-growth, light-touch approach in regulating data-driven technologies.⁸ This is also supported by the UK governments' intention in diluting the right to human review of AI decisions protected by Article 22 of the GDPR (General Data Protection Regulation)⁹. Through their strategy they project important signals to foster innovation. Although, ensuring ethical innovation needs a carefully drafted framework with more attention given to citizen's rights.¹⁰ Therefore, one might argue that a carefully drafted AI bill of rights, would be the appropriate regulatory framework that the UK needs to foster responsible innovation. In order to support the need for an adequate ex-ante framework one could assess potential exploitations of AI such as its use in the metaverse for manipulation.

AI in the Metaverse

At first, illustrating the need for an AI regulation through the effect of such systems in the metaverse, requires an understanding of these uprising immersive digital environments. The metaverse enables people and artificial intelligences to live and interact in a digital space. It was first envisioned as science fiction by Neal Stephenson in his novel *Snow Crash* in 1992¹¹. However, with the help of AI systems, this concept is moving from fiction to reality, with many immersive digital environments already being created and

⁷ Hikari Senju, *Why We Must Create An AI Bill Of Rights* (Forbes, 25 January 2022)

⁸ Anjana Ahuja, *A global AI bill of rights is desperately needed* (Financial Times, 11 October 2021)

⁹ EU General Data Protection Regulation [2016] OJ L 119

¹⁰ Kazim[N2],9

¹¹ Neal Stephenson, *Snow Crash* (Bantam Books 1992)

shared by people through screens or virtual reality goggles. Its development raised various discussions about the potential concerns that might arise, especially regarding the risks posed by using AI in virtual worlds. As Sandra Wachter, an associate professor at the Oxford Internet Institute, argued, “the metaverse will exacerbate problems that we already have”.¹² As it has been stated, one of the thorniest issues of the metaverse revolve around matters of data which are closely correlated with AI systems. The enjoyment of digital environments and worlds will also entail that a massive amount of data could be captured, recorded, processed, and sold.¹³ Therefore, it is important to examine the role of robust data processing systems through AI to understand the risks they pose to citizens. Some of the most notable uses of AI in the metaverse is targeted advertising through AI-generated content or the creation of AI-controlled bots and environments.¹⁴ Even though there are infinite uses of the AI in the metaverse, it is critical for the current examination to consider the uses that showcase the need for an AI Bill of Rights. As Louis Rosenberg, the scientist who developed the first functional augmented reality system stated in his article, the core problematic uses of AI in the metaverse are the “three Ms”.¹⁵ The “three Ms” refer to the ability of AI in the metaverse to Monitor, Manipulate, and eventually monetize us. Consequently, it is of paramount importance to examine the processes of Monitoring and Manipulation, which are closely correlated with AI, to illustrate the need to protect citizens’ rights from the effects of such systems.

Monitor

The first problematic use case of AI in the metaverse is the monitoring of people. Over the last decades, tracking and surveillance of behaviour and preferences have been a common and essential commercial practice of tech companies to enable targeted advertising and monetize data.¹⁶ This pervasive

¹² Sonia Elks, *AI bots to user data: Is there space for rights in the metaverse?* (Thomson Reuters, 12 November 2021)

¹³ *ibid*

¹⁴ John Radoff, *The Metaverse and Artificial Intelligence* (Medium, 15 April 2021)

¹⁵ Louis Rosenberg, *Regulate the Metaverse* (Medium, 21 December 2021)

¹⁶ *ibid*

monitoring was perceived by some as a violation of privacy due to the exploitation of information for polarization and radicalisation of masses in social media.¹⁷ However, the issues created by the attention economy might be exacerbated through AI in the metaverse. Such virtual worlds might allow companies to monitor everything a person does in their everyday life in the metaverse, such as where they go, what they say, what and how long someone is gazing at something.¹⁸ Through the use of AI in movement tracking, companies will also be able to monitor a person's gait, gestures, body language and potentially their surrounding environment.¹⁹ Processing of such information through AI systems can identify and classify users based on how they interact physically (e.g. holding objects) or visually (e.g. gazing in an advertisement) with content and create an extremely accurate consumer profile.

Furthermore, through the monitoring of facial expressions and biosignals such as heart rate, companies might be able to extract the real-time emotional state of people through artificial intelligence algorithms.²⁰ Consequently, the use of AI systems to monitor and profile users of the metaverse in such deep level raises huge concerns about the effect it might have on privacy. Firstly, in the current legal framework, it might be challenging to identify which national law applies in digital spaces.²¹ Subsequently, acquiring consent might become unwieldy, considering the complexity of the needs of each digital world and the number of organisations that might be implicated.²² Also the uninterrupted surveillance of people could render the right to a private life obsolete. Lastly, the outcomes of algorithmic processing such as consumer and emotional profiles of users are considered as inferred data which are yet to be protected by data-protection legislation.²³ Thus, the complexity of the implications which the use of AI in the metaverse introduces, illustrate the need for an efficient AI governance. Through the National AI Strategy, the government pledges to

¹⁷ Louis Rosenberg, Regulation of the Metaverse: A Roadmap (6th ICVARS, March 2022)

¹⁸ *Ibid*, 4

¹⁹ *Ibid*

²⁰ *Ibid*

²¹ Sonia[N12],

²² *Ibid*

²³ *Ibid*

uphold high data protection standards while removing unnecessary barriers to data flow. Therefore, the need to ensure citizens' rights to privacy in such a complex matter and achieve such balance for the use of AI in the metaverse or other technologies, supports the necessity of an AI bill of rights.

Manipulate

The second use case that might pose a risk to human rights and autonomy is the use of AI to manipulate users. For many years since the development of media, advertisers have skilfully attempted to influence public opinions on a macro-scale and personal opinions on a micro-scale.²⁴ This was enhanced through the collection of data from social media and the implementation of targeted advertising technics, nudges and algorithms. While it is argued that microtargeting has been proven harmful in promoting political and ideological content due to its polarising nature, it is widely used for commercial purposes. However, since consumers nowadays are aware of such technics, they are sceptical of advertisements they encounter on platforms.²⁵ However, the metaverse will arguably elevate microtargeting in a different level, with such personalized content that might be hard to resist.²⁶ As it has been established, the metaverse creates the potential of collecting massive amounts of personal data that could be used for highly accurate real-time profiling of user's behaviour, emotions, and preferences. It unlocks new ways of influencing and driving user choices, such as combining information for psychological and emotional states, biological signals, and unconscious cognitive processes. Such integration of information might transform microtargeting into subliminal advertising.²⁷ Subsequently, it can be argued that the influence of such methods cannot be resisted easily without being appropriately informed of their employment.²⁸ Therefore, one can deduce that transparency in the marketing algorithms used by platforms could be critical

²⁴ Rosenberg(2022)[N17]

²⁵ ibid

²⁶ ibid

²⁷ Rune Klingenberg Hansen, Meta's Metaverse With Likely Be Filled With Marketing And Manipulation (Data Ethics, 15 February 2022)

²⁸ ibid

for the protection of human autonomy. Along that, though, the metaverse goes beyond the use of highly accurate profiling and adds another dimension to the potentials of user manipulation through AI-generated content. Product placements and other promotions might be conducted through virtual illustrations created by algorithms designed to micro-target the user.²⁹ Also, platform providers will be able to manipulate the metaverse experience of any user through the interactions they have with objects or bots/virtual salespersons.

Firstly, a simplistic illustration of such manipulations through object projection would be a user walking on a virtual street and being exposed only to certain brands of cars, stores, clothes or other products. In that case, the algorithm would choose the most suitable content based on a users' preferences and emotional state.³⁰ Secondly, there is a possibility of using virtual salespersons with an adoptive personalized pitch for every user. Through the use of data on facial expressions, psychology and vocal inflections, AI systems will be able to generate persuasive and "trustworthy" virtual salespersons.³¹ The interaction with such AI-generated content or bots will feel like a serendipitous encounter and users might think that everyone experiences the same environment without understanding they are being subjected to microtargeting. Thus, one might argue that the capabilities of AI in the metaverse and the potential of manipulation and pervasive influence of people raises various concerns regarding the risks to human autonomy and human rights.

Subsequently, there is a need for a framework that will place restrictions on platforms and protect citizens, especially the vulnerable members of society, from such manipulation.³² Psychometric targeting and the exploitation of such personal inferred data will need to be restricted to protect human autonomy and prevent companies from taking advantage of people that might be in vulnerable physical or psychological states.³³ It is essential also to highlight

²⁹ Rosenberg(2022)[N17]

³⁰ ibid

³¹ ibid

³² Klingenberg[N27]

³³ ibid

the aforementioned complexity on acquiring consent for the use of such methods and the lack of transparency on the way the algorithms functions. The National AI strategy explicitly stated their intention to protect the choices, safety and rights of UK citizens while promoting innovation.³⁴ Thus, one might deduce that the capabilities of AI in the metaverse for manipulation and the risks posed towards human autonomy and rights, support the need for a regulatory framework, such as the AI bill of rights, which prioritizes the protection of citizens' choices and rights. Such legislation will need to impose transparency standards on the use of AI and restrictions on the methods that platforms might use for advertising.

Conclusion

Conclusively, the implementation of AI algorithms in various aspects of people's commercial and personal lives have raised severe concerns about the effect of such systems. Since AI algorithms have already surpassed human intelligence powers in several tasks, their employment by companies and institutions is increasing rapidly. However, organisations' pro-innovation and profitable approach might sometimes overshadow concerns about the effect such systems might have on users. Subsequently, governments and bodies such as the EU are projecting their intention to regulate this disruptive technology to set innovation in line with their priorities. The UK National AI Strategy takes a pro-innovation approach illustrating their priority in placing UK as a leading country in the AI industry. However, it is evident from those above that the employment of AI in the current and future digital world could be exploited in a harmful way. The introduction of the metaverse risks exacerbating the current issues raised by AI exploitation. While data protection regulation has been an ex-post measure to protect the data of users in the digital world, the introduction of a new virtual reality arguably threatens privacy on a new level. The monitoring, computational and influencing capabilities that AI-enabled systems might have in the metaverse are a flagrant indication of the need for adequate regulation of AI. Through the

³⁴ NAS[3],7

National AI Strategy, the government mentions the need for an appropriate regulatory framework. They support that this framework will need to maximize growth and protect the safety, security, choices, and rights of UK citizens.³⁵ Therefore, to achieve a balance of protecting people's rights and maximizing growth a regulation such as the AI Bill of Rights proposed by the US government would be essential. While innovation might be a priority for the government, such a framework would require a human-centred approach to protect adequately human rights and notions such as human autonomy and dignity.

³⁵ *ibid*

Bibliography

Legislation

- EU General Data Protection Regulation [2016] OJ L 119

Books

- Neal Stephenson, *Snow Crash* (Bantam Books 1992)

Articles

- Anjana Ahuja, *A global AI bill of rights is desperately needed* (Financial Times, 11 October 2021)
- Bernd Carsten Stahl, *Artificial Intelligence for a Better Future An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies* (Springer, 18 March 2021)
- Emre Kazim, Denise Almeida, Nigel Kingsman, Charles Kerrigan, Adriano Koshiyama, Elizabeth Lomas, Airlie Hilliard, *Innovation and opportunity: review of the UK's national AI strategy* (06 December 2021) Discover Artificial Intelligence Vol.1, No.1, pp.1-10
- Glenn Glow, *The AI Bill Of Rights: Protecting Americans From The Dangers Of Artificial Intelligence* (Forbes, 9 January 2022)
- Hikari Senju, *Why We Must Create An AI Bill Of Rights* (Forbes, 25 January 2022)
- John Radoff, *The Metaverse and Artificial Intelligence* (Medium, 15 April 2021)

- Louis Rosenberg, Regulation of the Metaverse: A Roadmap (6th ICVARS, March 2022)
- Louis Rosenberg, *Regulate the Metaverse* (Medium, 21 December 2021)
- Rune Klingenberg Hansen, Meta's Metaverse With Likely Be Filled With Marketing And Manipulation (Data Ethics, 15 February 2022)
- Sonia Elks, *AI bots to user data: Is there space for rights in the metaverse?* (Thomson Reuters, 12 November 2021)
-