What are the responsibilities of governing or regulatory bodies in determining and regulating the place of intelligent technology? Does the domain of use change anything? How does this translate into responsibilities for the individual programmer?

Ferdinand Krammer

Humans and Intelligent machines CW1

2021

1 Introduction

As we progress in to the 21st century the number of intelligent technologies which we interact with on a daily basis is increasing. These intelligent technologies are based around some form of artificial intelligence, as such will be known as artificial intelligent technologies (AITs) from here on. Governments and regulatory bodies have very little control over the place that AITs are taking in the world, as although they are increasingly investing money in their development, the effect of this is likely to be minimal due to the amounts of money being invested by private industry[1]. Thus the best way in which governments can influence the place of AITs and exercise what they feel to be their responsibilities, is through regulation. The process of regulating AITs is proving to be a contentious issue with prominent people that are involved in their development being both for regulation (Elon Musk called for regulation of AI development in 2017 due to its potential to do harm[2]) and against regulation (intel CEO Brian Krzanich has said that it is too early to regulate as it is in its "infancy"[3]). Current regulation varies vastly around the world with countries taking their own paths, depending on what they see as their responsibilities, while general ethical principles are fairly closely aligned[4]), with regulators generally in consensus that 'something' needs to be done, they are not when it comes to 'what' needs to be done[5]. Because the use and development of AITs is an international endeavour, the effects of regulation across the world has to be considered as in the absence of a unified approach, the regulation in one place can affect what can be developed in another (as has been seen with the uses of date since the introduction of GDPR regulation in the EU)[6]. This leads to the question needing to be asked as to what the responsibility of the governments are, we will attempt to consider the various approaches different governments have taken to AITs and what regulation they are bringing in, in order to gain an understanding of their perceived responsibilities and how these affect the individual programmer.

2 How governments perceive their responsibilities and how this affects their regulation of intelligent technologies

The methods of regulation and the perceived role of government in the process of regulation varies between countries and regions. While these differences in regulation exist for a variety of different reasons, one of the main reasons for their respective approaches is how they see the best way to become a world leader in AI. Some countries, such as China, are seeking to become the world leaders in AI through the establishment of "ethical norms, policies and regulation" in different areas of AI[7] as well as directing the areas of research[8]. The Chinese government want to direct the development of AI through the creation of regulation that establishes a traceable and accountable system[8]; they feel this would enable them to become internationally competitive[7]. However when considering the responsibilities the Chinese government feels it has in determining and regulating the place of AITs, they put the 'good' of the people before the 'good' of the individual[9], meaning that regulations which are put in place to protect people can be broken if "ultimately [it] serve[s] the state[s] ends"[6]), even when its interests go against ethical concerns[7].

At the other end of the spectrum, we find the USA, under both the Obama and Trump administrations the opinion was held that to encourage the development and use of AITs a light handed approach was to be adopted[6, 10], however there have been incidences where they have felt the need to introduce regulation as has been done with self-driving cars[6]. This approach can be termed an 'open market' approach, with the emphasis being placed on self-regulation, with companies such as Microsoft[11] and Google[12] producing lists of ethical principles. However this process, as C. Cath et al.(2017) suggests, ends up favouring "industry over other stakeholders", as well as being heavily reliant on the private sector taking the initiative in the first place. While this general absence of regulation may initially result in an increase in the AI being developed in different spheres, it may also become a hindrance as companies developing these AITs are transnational, and as such will be following regulations in other regions of the world. This has already been seen with the GDPR regulation brought in by the EU back in 2018[6].

As with the China the EU is following a more regulated approach to help it to determine the place of AIT. The EU are trying use regulation to determine the place of AITs based on their risks[13, 14], as set out in their report produced in April 2021[14]). These proposed regulations for the EU would split AITs into different levels, with 'unacceptable risk' AITs being prohibited as they are likely to contravene fundamental rights[13, 14], whereas 'high risk' AITs will be allowed but with increased obligations for monitoring, reporting and transparency, these obligations would be the responsibility of the systems provider([13],7). The final category of 'non-high risk' applications would not require any additional compliance beyond that which already exists[8, 14]. Damian Cyman et al. (2021) suggested that the EU's response could be due to the "fear [of] being left powerless" as well as not wanting to play catch-up with any changes in the field of AITs, but rather have the ability to control the direction of their development and the place AIT's will have in the society of tomorrow. As the EU feels that its responsibilities lie in the protection of its citizens from harm along with protecting their rights[8] from potential abuse by AITs, they already have GDPR which functions to protect citizens' data by laying an emphasis on the transparent use of it and the ability for people to challenge its uses[15]. Although the framework which the EU have laid out might be comprehensive in defining the limits within which AITs can be developed, it risks hindering the development of AITs in the future due to this relatively broad-brush approach which attempts to 'future proof' their legislation[13]. As the AIT industry is international in nature, potential fall-out over regulation will have an international effect, if the AITs in question are to be rolled out in the EU, they will have to be compliant[6]). This comes back

to the question of where to draw the line between the responsibility of government versus industry in the development of AITs.

Until recently the UK has seen its role in determining the place of AITs in a similar way to the US where the desired effect is to drive economic growth and a light touch approach is considered the best way of achieving this[16]), yet with the willingness to introduce regulation if it was felt to be needed in an international setting, as indicated by Theresa May in a speech at Davos in 2018[16]. However, with the release of the UK National AI strategy[17] in September 2021, there has been a marked change in the UK governments understanding of its role in determining the role of AIT through the use of regulation. While it still aims to drive economic growth and AIT development, there is a recognition that AITs can be harmful and that the current regular framework is not up to standard[17]. To balance the potential good with the bad, the UK is following a path between the US and EU regulatory approaches, this reflects the UK perception that their responsibility is the protection of their citizens as well as encouraging the development of AITs to enable them to play a greater role[17]. The UK government intends to achieve this through a decentralised approach as they recognise that each sector where AITs can be used has its own specific requirements, with regulations intended to set the boundaries in which innovation can be undertaken. This would not be satisfied by a broad-brush approach, as employed by the EU, however the UK realizes that there may be instances where they need to produce overarching regulation in a similar fashion to the GDPR regulation[17]. This approach is at risk of becoming overly fragmented, with the potential confusion as to who is responsible for regulation of every step of the development process being "hugely detrimental in terms of efficiency and efficacy"[10]. However we will have more clarity on this when the white paper on governing and regulating AI is released by the office for artificial intelligence in early 2022[17].

3 How the domain of use affects responsibilities

Although many AITs are in their infancy, the role of a government is to create constructive regulation that directs their development, even though there is much uncertainty about what their abilities will be[18]. With these emerging AITs, upsides and downsides are also emerging, and the question is what regulatory approaches need to be taken in the various domains of use.

Medicine is one of the few domains which has its own pre-existing regulatory body in most countries, in the UK this is the Medicines & Healthcare products Regulatory Agency (MHRA). In medicine, as expressed by the MHRA's recent guidance released September 2021[19], the protection of patients is paramount. Unlike in most domains, in medicine, a lot of the potential uses of AITs have potentially life changing consequences if things go wrong; this presents challenges in increasing the adoption of these systems which become more accurate as they are further developed. As AITs are increasingly used in support roles to help with detection and diagnosis[5]), regulatory bodies have the responsibility to ensure these systems and processes transparent and explainable, as otherwise why should doctors and patients trust such technologies[5]. This was evidenced in a study proposing the use of a robot with a high success rate, or a surgeon with a lower success rate; all bar one of those asked picked the surgeon[20]. Regulators also have the responsibility to of working out where accountability would lie if things go wrong with an AIT[21]; without this the desired acceptance of such systems would not be possible, as just one case would damage any trust that was had in them.

The use of AITs in synthetic media to create Deepfakes came to prominence in 2017 with the swapping of celebrity faces on to phonographic videos[22]. The use of Deepfakes is an example of an area where if we were to consider the government's responsibility to protect

people from manipulation, then as these have the potential to manipulate people (demonstrated when fraudsters mimicked a CEO's voice to get a manager to transfer $243,000)[23], the logical solution would be an outright ban. However, deepfakes also have potential uses in legitimate media as shown by the Princess Leia in Rogue One[24], although this produces ethical problems if the person dies without giving their consent for them to be recreated. As shown by these two examples, governments need to fine-tune their responses to domains as their possible uses aren't black and white; the same AITs can be used for both legitimate and illegitimate uses. Possible regulation should require transparency along with regulation of the uses of data sets as proposed by the EU regulation[14], however the data is so widespread that this can be done without being truthful removing any accountability.

In the military and defence domain there are many potential uses of AITs, these can be split into 3 main areas: sustainment and support, adversarial and non-kinetic, adversarial and kinetic[25]. Whilst some countries see their role as protector of their people, thus are seeking to limiting the development of AITs to these areas, others see AITs as a way of gaining advantage over potential opponents, such as the development of lethal autonomous weapons systems (LAWS)[26], some countries, such as the UK, have already banned the development of such AITs[27]. One of the biggest concerns with such systems is the lack of human input or intervention, with it being suggested by Strous et al. 2020 that as whilst a human can react to changes in the environment, a machine might not be able to (such as with a target walking in to hospital), as this removes accountability from the equation. In view of the fact that different countries will approach these in different ways for different reasons, finding a solution that determines the place we want AITs to take is probably the most pressing area where governments need to work together.

4 Responsibilities of the individual Programmer

The individual programmer has currently very little to worry about regarding the regulation of AI, as it is currently virtually non-existent. However as the regulatory landscape changes, as outlined in section 2, the responsibilities of the individual will increase, with the onus being placed on transparency and explainability. This responsibility will lie with the individual, as they have an intimate understanding of the program, however Buiten(2019) suggests that this idea of transparency will break down as programs become more complex especially ones which adapt over time, as even the code being used only provides a snapshot in time[21], this however could be overcome through backtracing. In the absence of any meaningful input from governments or regulatory bodies, the individual programmer has both an ethical and moral responsibly to ensure the reliability as well as to limit the ability for the programs developed, as laid out by ethical codes of conduct such as that produced by IEEE[28]. While the self-regulatory approach to regulation, as taken by the US, doesn't impose any hard boundaries in what can be done in the development and place of AITs, it can at least act to impose preventative boundaries, as suggested by Wischmeyer et al.(2020), that could be "socially sanctionable".

5 Conclusion

As has been seen different governments understand their responsibilities towards regulating and determining the place of AITs in society differently. In the EU and UK, governments understand their responsibility to be the protection of their citizens, as has been shown in section 2, this has led to the need the impose regulation to define the boundaries within which AIT has a place in society. However it has been seen that not all countries agree over where

there responsibilities lie, as demonstrated by the US which appears to see its responsibilities lying more with the 'open market', and China which seems to see its responsibility more to the state. These differences in perceived responsibilities has led to differences in regulatory approaches, and we are still in a unique situation, as although countries are starting to create regulation determining the place of AITs, there is currently no robust regulation in place. This lack of robust regulation alows the potential for international regulation[16]) and this is particularly needed as so many of the companies are international. In determining and regulating the place of AITs, governments and regulatory authorities need to be careful that any boundaries they create do not result in the limiting of development, and with different domains the required intervention of regulatory authorities will also change. If it was decided that it was the responsibility of the government to ensure that AITs are transparent and explainable, this transparency would also go a long way in fulfilling governments' responsibilities in determining the place of AITs, although this will not necessarily resolve all of the potential issues which could arise in different domains. The responsibly to try to ensure transparency also lies with the individual programmer due to their intimate knowledge of the respective systems. Both governments and regulatory bodies, as well as the individual programmer have the responsibility to react quickly to prevent unintended consequences of the use of AITs. As although an AITs might have been designed for one use, as they are malleable systems the intended use might not be the end use or even a foreseen use of such systems.

References

1. Why The Race For AI Dominance Is More Global Than You Think. Available from: https://www.forbes.com/sites/cognitiveworld/2020/02/09/why-the-race-for-ai-dominance-is-more-global-than-you-think/ [Accessed on: 2021 Oct 20]
2. Elon Musk Warns Governors: Artificial Intelligence Poses 'Existential Risk'. 2017 Jul. Available from: https://www.npr.org/sections/thetwo-way/2017/07/17/537686649/elon-musk-warns-governors-artificial-intelligence-poses-existential-risk?t=1635092715358 [Accessed on: 2021 Oct 20]
3. Kharpal A. Intel CEO Brian Krzanich: It's too early to regulate AI. 2017 Nov. Available from: https://www.cnbc.com/2017/11/07/ai-infancy-and-too-early-to-regulate-intel-ceo-brian-krzanich-says.html [Accessed on: 2021 Oct 20]
4. Hagendorff T. The ethics of AI ethics: An evaluation of guidelines. Minds and Machines 2020; 30:99–120
5. Wischmeyer T and Rademacher T. Regulating artificial intelligence. Springer International Publishing, 2019 :1–388. DOI: 10.1007/978-3-030-32361-5
6. Thierer A. U.S. Artificial Intelligence Governance in the Obama-Trump Years. IEEE Transactions on Technology and Society 2021 :1–1. DOI: 10.1109/tts.2021.3113616
7. Roberts H, Cowls J, Morley J, Taddeo M, Wang V, and Floridi L. The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. AI and Society 2021; 36(1):59–77. ISSN: 14355655. DOI: 10.1007/s00146-020-00992-2
8. Cyman D, Gromova E, and Juchnevicius E. Regulation of Artificial Intelligence in BRICS and the European Union. BRICS Law Journal 2021; 8(1):86–115. ISSN: 2412-2343. DOI: 10.21684/2412-2343-2021-8-1-86-115
9. Steele LG and Lynch SM. The pursuit of happiness in China: Individualism, collectivism, and subjective well-being during China's economic and social transformation. Social indicators research 2013; 114:441–51

10. Cath C, Wachter S, Mittelstadt B, Taddeo M, and Floridi L. Artificial Intelligence and the 'Good Society': the US, EU, and UK approach. Science and Engineering Ethics 2018; 24(2):505–28. ISSN: 14715546. DOI: 10.1007/s11948-017-9901-7

11. Responsible AI principles from Microsoft. Available from: https://www.microsoft.com/en-us/ai/responsible-ai?activetab=pivot1:primaryr6 [Accessed on: 2021 Oct 20]

12. Perspectives on Issues in AI Governance. Available from: https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf [Accessed on: 2021 Oct 20]

13. Beverly AT. Decoding the Proposed European Union Artificial Intelligence Act. American Society of Intonational Law 2021 Sep; 25(20)

14. European Comission. Regulation of the European Parliament and of the Council - Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and amending Certain Union Legislative Acts. 2021; COM(2021) 206 final

15. Malgieri G. Automated decision-making in the EU Member States: The right to explanation and other "suitable safeguards" in the national legislations. Computer law & security review 2019; 35:105327

16. Turner J. Robot Rules. Springer International Publishing, 2019. ISBN: 978-3-319-96234-4. DOI: 10.1007/978-3-319-96235-1. Available from: http://link.springer.com/10.1007/978-3-319-96235-1

17. HM Goverment. National AI Strategy Presented to Parliament by the Secretary of State for Digital, Culture, Media and Sport. HM Goverment, 2021 Sep. Available from: https://www.gov.uk/government/publications/national-ai-strategy [Accessed on: 2021 Oct 20]

18. Wischmeyer T and Rademacher T. Regulating Artificial Intelligence. Springer, 2020

19. Medicines Healthcare products Regulatory Agency. Software and AI as a Medical Device Change Programme. 2021 Sep. Available from: https://www.gov.uk/government/publications/software-and-ai-as-a-medical-device-change-programme/software-and-ai-as-a-medical-device-change-programme [Accessed on: 2021 Oct 20]

20. Kundu S. AI in medicine must be explainable. Nature Medicine 2021; 27(8):1328. ISSN: 1546170X. DOI: 10.1038/s41591-021-01461-z

21. Buiten MC. Towards intelligent regulation of artificial intelligence. European Journal of Risk Regulation 2019; 10(1):41–59. ISSN: 21908249. DOI: 10.1017/err.2019.8

22. Pavis M. Rebalancing our regulatory response to Deepfakes with performers' rights. Convergence 2021; 27(4):974–98. ISSN: 17487382. DOI: 10.1177/13548565211033418

23. Kietzmann J, Lee LW, McCarthy IP, and Kietzmann TC. Deepfakes: Trick or treat? Business Horizons 2020; 63:135–46

24. Star Wars: Millie Bobby Brown Is Princess Leia In New Video. Available from: https://screenrant.com/star-wars-millie-bobby-brown-princess-leia-deepfake-video/ [Accessed on: 2021 Oct 20]

25. Taddeo M, McNeish D, Blanchard A, and Edgar E. Ethical Principles for Artificial Intelligence in National Defence. Philosophy & Technology 2021 :1–23

26. Strous L, Johnson R, Grier D, and Swade D. Unimagined Futures–ICT Opportunities and Challenges. Springer, 2020

27. Article 36. The United Kingdom and lethal autonomous weapons systems. 2016 Apr. Available from: https://article36.org/wp-content/uploads/2016/04/UK-and-LAWS.pdf [Accessed on: 2021 Oct 20]

28. IEEE-CS/ACM Joint Task Force on Software Engineering Ethics and Professional Practices. Code of Ethics. Available from: https://www.computer.org/education/code-of-ethics [Accessed on: 2021 Oct 20]