

Assessing the Appetite for Trustworthiness and the Regulation of Artificial Intelligence in Europe

Labhaoise NíFhaoláin¹, Andrew Hines¹ and Vivek Nallur¹

¹School of Computer Science, University College Dublin, Ireland
laoibhse.nifhaolain@ucdconnect.ie

Abstract. While Artificial Intelligence (AI) is near ubiquitous, there is no effective control framework within which it is being advanced. Without a control framework, trustworthiness of AI is impacted. This negatively affects adoption of AI and reduces its potential for social benefit. For international trade and technology cooperation, effective regulatory frameworks need to be created. This study presents a thematic analysis of national AI strategies for European countries in order to assess the appetite for an AI regulatory framework. A Declaration of Cooperation on AI was signed by EU members and non-members in 2018. Many of the signatories have adopted national strategies on AI. In general there is a high level of homogeneity in the national strategies. An expectation of regulation, in some form, is expressed in the strategies, though a reference to AI specific legislation is not universal. With the exception of some outliers, international cooperation is supported. The shape of effective AI regulation has not been agreed upon by stakeholders but governments are expecting and seeking regulatory frameworks. This indicates an appetite for regulation. The international focus has been on regulating AI solutions and not on the regulation of individuals. The introduction of a professional regulation system may be a complementary or alternative regulatory strategy. Whether the appetite and priorities seen in Europe are mirrored worldwide will require a broader study of the national AI strategy landscape.

Keywords: Artificial Intelligence, Regulation, Trustworthiness

1 Introduction

At this point, to say that Artificial intelligence (AI) is all pervasive is trite. However despite its inescapable presence, society has yet to identify an effective way to oversee and control this technology. The control of AI is essential to achieve solutions with an overall social benefit which are also safe, trustworthy and legal. Control in this context can take many forms in society, from soft governance through to legislative interventions, from company ethics guidelines to criminal law sanctions [50]. The challenge being faced on a national and global scale is how to ensure that the development and deployment of AI solutions is managed through some control framework. The absence of a control framework for AI impacts trustworthiness, which in turn impacts the adoption and usage of

AI [48,12]. The adoption of technology is influenced by the users' perception of risk [51]. Behavioural decision theory informs how regulation can counterbalance the perceived risk, resulting in increased adoption levels [52].

It has been suggested that placing the focus on global cooperation and safe, responsible development of AI, rather than on technological dominance or leadership is likely to result in greater public confidence in AI solutions [5]. A large number of private, national, and global organisations are examining and documenting how AI is being, can be, and should be developed and deployed. The Council of Europe has identified 348 such documents [7]. In 2018, 29 European countries (the Signatories) signed a Declaration of Cooperation on Artificial Intelligence (The Declaration of Cooperation) [8]. Many of these countries have adopted national strategies on AI (the Strategies) which are examined in this paper. The EU Commission published its Inception Impact Assessment for AI Legislation in July 2020 [12]. This public consultation sought views on options to address ethical and legal issues arising out of AI. The overall policy objective of that consultation process is to ensure the development and uptake of trustworthy and lawful AI.

This study assesses the appetite for a AI regulatory framework based on the published strategies of the Signatories through thematic content analysis. There was an expectation that there would be a high degree of similarity amongst the strategies, given that the Signatories are all European and have the Declaration of Cooperation in common. This review sought to test that hypothesis and to elicit data on outliers in order to identify the anomalous Signatories and seek to understand the reason for their deviation.

2 Background

2.1 International focus on AI

A number of European and global organisations and bodies have emphasised AI over recent years by facilitating collaboration amongst nations and carrying out research on the international AI landscape. These organisations and bodies have differing roles but what they have in common is the legal capacity to require compliance with legal instruments (whether conventions, treaties, laws, decisions or agreements). Their work demonstrates that there is a desire at an international level within a number of organisations to influence how the use of AI develops. Should these organisations choose to create legally binding instruments to regulate AI, then Signatories' receptiveness to supranational regulation will impact on the success of such regulation. It is these organisations and associated governments which enable regulation and governance, rather than institutes and think tanks.

EU Commission: The EU Commission's High Level Expert Group on Artificial Intelligence has carried out work on Trustworthy AI [21,20] and the EU Commission has issued reports on the Safety and Liability Aspects of AI [10] and on the Ethics of Automated Vehicles [11]. In 2020 the EU Commission sought views on a proposal for AI specific legislation [12] and that process is ongoing.

In April and May of 2018, the Signatories (which comprise both EU and non EU states) signed the Declaration of Cooperation pursuant to which they agreed to work towards an integrated European approach to AI, though a single documented strategy is not referred to. The Signatories are Austria, Belgium, Bulgaria, Cyprus, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Norway, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden and the UK. The Signatories agreed, among other things, that they would "ensure an adequate legal and ethical framework, building on EU fundamental rights and values, including privacy and protections of personal data, as well as principles such as transparency and accountability." The Declaration of Cooperation is not enforceable and, as the title suggests, is based on cooperation. However, the EU Commission is a powerful body as it can both initiate legislation and bring actions against states and businesses for breaches of EU Law.

Council of Europe (CoE): The CoE is an international organisation which promotes human rights democracy and the rule of law and is comprised of 47 member states (including all 29 Signatories). The CoE has created a common legal space in which the members have a legal obligation to guarantee rights as set out in the European Convention on Human Rights. Specifically in relation to AI "The Council of Europe's aim is to identify intersecting areas between AI and our standards on human rights, democracy and rule of law, and to develop relevant standard setting or capacity-building solutions" [6]. The large number of documents identified by the CoE [7] include guidelines, charters, papers, reports and strategies. The authoring bodies are not confined to one sector of society and include organisations, companies, bodies and nation states.

United National Educational, Scientific and Cultural Organisation (UNESCO): Formed in 1945, the United National Educational, Scientific and Cultural Organisation (UNESCO) is global organisation whose mission is to "contribute to the building of a culture of peace, the eradication of poverty, sustainable development and intercultural dialogue through education, the sciences, culture, communication and information". At UNESCO's Scientific 40th session in November 2019 the organisation commenced a two year process to achieve a "global standard-setting instrument on ethics of artificial intelligence". In pursuit of this goal, UNESCO forums and conferences on AI have taken place to gather stakeholder views. The most recent draft text of a recommendation on the ethics of AI of the UNESCO Ad Hoc Expert Group was issued in September 2020 [57] and includes a call for legislative gaps to be filled. While UNESCO Recommendations are not legally binding, its Conventions are.

The Organisation for Economic Co-operation and Development (OECD): The OECD is an intergovernmental economic organisation, formed in 1961 to stimulate economic progress and world trade. It comprises 37, mostly European, member countries. A recommendation of the OECD Council on Artificial Intelligence was adopted on 21 May 2019 by the OECD countries and various non-member adherents [39]. This Recommendation promotes AI that is innovative, trustworthy and upholds democratic values and human rights and

was a culmination of a work carried out by a diverse expert group. While OECD Recommendations are not legally binding, it has the capacity to create legally binding instruments.

2.2 Concept of Trustworthiness in AI

"Trustworthiness in AI" is a term used liberally though it is a troubled concept. There is no single definition of trustworthiness in AI, or indeed in of trustworthiness in software systems. The following quality attributes have been identified as necessary as a foundation for a trustworthy software system [19,2]:

1. **Correctness:** absence of faults
2. **Safety:** absence of catastrophic consequences on the environment
3. **Quality of service:** encompassing (i) availability (probability of readiness for correct service) (ii) reliability (probability of correct service for a given duration of time) and (iii) performance (response time, throughput)
4. **Security:** absence of unauthorised access to a system
5. **Privacy:** absence of unauthorised disclosure of information

Ensuring trust spans the AI lifecycle from creation, through validation and interaction with AI. Becker et al. [2] outlined the need for the above quality attributes to exist to allow for certification which in turn results in a trustworthy software system. Lewis et al. [29] noted that it is necessary to specify both who trusts whom and in what aspects of AI creation, operation and use. When we talk about a "trustworthy product" the trust exists between us and the supplier – we trust that the product attributes have been delivered by the supplier, not that the product has the free will and volition to behave in a trustworthy manner. The need to regulate for human responsibility is addressed by the principles of robotics [4] rather than Azimov's robot centric laws [1].

From a legal perspective, the concept of a trustworthy product is problematic at a European and national level. Taking the Irish jurisdiction as an example, liability may arise as a result of trust arising between people but a "trustworthy product" does not exist as a concept. Instead, laws relating to product liability and safety are relevant. In contract law, goods supplied must be of merchantable quality (that is to say that they are fit for the purpose for which goods are usually purchased, taking all relevant circumstances into account). Failing to meet this level of quality may be a breach of an implied term of the contract [42,43]. Goods delivered under contract must comply with that contract [44]. Under statute, a defective product is one which does not reach a level of safety which a person is entitled to expect [41]. Outside of statute and contract law, under the law of tort (a civil wrong), a supplier may be liable in negligence if a product is found to be defective. Criminal law may also apply if an unsafe product has been supplied [45]. Whether European Law on product liability is sufficient to address liability in AI has been examined recently in a report released by the European Parliament [3]. The author of that report found that the regime as it stands is not sufficient to address the liability arising in the context of AI.

The Standards community has defined trust as the "degree to which a user...or other stakeholder...has confidence that a product or system...will behave as intended" while trustworthiness is seen as an "ability to meet stakeholders'...expectations in a verifiable way" in a technical report issued by ISO and the International Electrotechnical Commission (IEC) [24].

The EU Commission's High-Level Expert Group on AI [20] considered that in order for AI to be trustworthy it must be lawful, ethical and robust (from a technical and social perspective).

From a computer science perspective, outlined above, the focus is on the software as a service/product and the process driven approach which dictates that once the requisite technological bar has been reached, certification could be granted. The EU Commission's High-Level Expert Group's approach to AI is significantly broader than the foundational quality attributes referred to by Becker et al. and the definition produced by the ISO and IEC.

2.3 Documented Approaches to AI

Research into whether there is global agreement on ethics questions has been published in articles which include reviews of principles and guidelines issued by private companies, public sector organisations and research institutions [25,29]. Jobin et al. revealed five unifying ethical principles of transparency, justice and fairness, non-maleficence, responsibility and privacy [25]. In a similar vein, Fjeld et al. [13] in their review of 36 principles and documents, which included some national strategies, identified seven common themes of privacy, accountability, safety and security, transparency and explainability, fairness and non-discrimination, human control of technology, professional responsibility, promotion of human value. The EU Commission's Science for Policy Report includes an analysis of 13 national strategies on AI from a public services perspective [36]. As the above demonstrates, there is considerable amount of policy work being carried out worldwide on AI. This survey has focussed on Europe and in particular on the governmental strategies of the Signatories.

3 Comparing National AI Strategies

The National Strategies were identified through a review of the relevant countries' governmental websites, internet searches (Google), EU Commission's AI Watch [9], CoE's AI Initiatives [7] and Canadian Institute for Advanced Research (CIFAR) [27]. Strategies of Cyprus, Italy, Latvia, Poland and the Netherlands were not available in English. These were translated using Microsoft Office translation tool and then reviewed. Given the thematic nature of the review, this translation method was not a significant limitation. The final search for published national strategies on AI for the remaining Signatories was carried out on 14 November 2020.

A content analysis methodology was adopted as this method reduces data, is systematic and is flexible [14]. An initial review of the material and application

of emergent coding resulted in the umbrella themes under which the strategies were analysed, with a view to delivering trustworthy AI. In order to minimise subjectivity, binary values were attributed rather than weighted values. Any absence of evidence under the chosen headings was used to identify outliers. The themes were categorised as follows:

1. Intention to adopt ethical, trustworthy or legal AI (that is to say whether there appeared to be an intention to highlight the importance of ethical/trustworthy AI on an ongoing basis through government):
 - Ethics, trustworthy AI, law or regulation identified as a focus area – where this is highlighted through a standalone section
 - International cooperation – where desire is expressed to cooperate and collaborate with other countries (aside from cooperation to merely attract inward investment)
 - National body for AI (for ethics or otherwise) – in the form of a standing committee, council or body, whether proposed or in existence
2. Tools for implementation of ethical/ responsible AI:
 - Certification or standards of AI solutions proposed or in existence – this includes references to both certification and auditing processes and industry standards
 - AI specific legislation referred to – where intention is expressed to promulgate legislation targeted at AI and associated technology
 - Regulation – whether through use of regulatory sandboxes or otherwise, whether proposed or in existence

Twenty-one Signatories have published National Strategies on AI, 3 of which had published a strategy prior to signing the Declaration of Cooperation (Finland, UK and France). The remaining 8 Signatories, which have yet to adopt a strategies, have instituted a process to do so (being Austria, Belgium, Bulgaria, Croatia, Greece, Ireland, Romania and Slovenia).

The methodologies used by governments to produce strategies varied and included delegation of policy formulation to experts in academia and/or industry, public consultation and production of strategies by government departments or a combination of methodologies.

Signatories' appetite for the regulation of AI is not demonstrated by the mere existence of a strategy on AI which may, for example focus on how to attract inward investment, or the need for education in the field. Indeed, while in general high level national strategies do not necessarily translate into action and in many cases are purely aspirational, at the very least, they demonstrate an acknowledgement that AI is an area which warrants specific attention from national governments.

An intention to comply with the law appears in the majority of strategies and this reflects the commitment made by the Signatories in the Declaration of Cooperation. However, a statement that a Nation State will ensure that the relevant laws are complied with is not significant in itself, given that it is simply a statement of commitment to the rule of law.

Signatory	Certification/ Auditing / Standards	National Council/ Body	AI Specific Legislation referred to	Regulation (incl. sandbox) proposed	Ethics/Trust/Legal Regulation as focus	International Cooperation emphasised
Cyprus [56]	●	●	●	●	●	●
Czech Republic [34]	●	●	●	●	●	●
Denmark [53]	●	●	●	●	●	●
Estonia[17]	○	○	●	●	●	○
Finland [32]	●	●	○	●	●	●
France [15]	●	●	●	●	●	●
Germany [54]	●	●	●	●	●	●
Hungary [31]	●	○	●	●	●	●
Italy [30]	●	●	●	●	●	●
Latvia [28]	●	●	●	●	●	●
Lithuania [33]	●	●	●	●	●	●
Luxembourg [55]	●	●	○	●	●	●
Malta [46]	●	●	●	●	●	●
Netherlands [47]	●	●	●	●	●	●
Norway [38]	●	●	○	●	●	●
Poland* [49]	○	●	○	●	●	●
Portugal [16]	○	●	○	●	●	●
Slovakia [40]	○	●	○	●	●	●
Spain [35]	●	●	○	○	●	●
Sweden [18]	●	●	○	●	●	●
UK [23]	○	●	○	●	○	○

Table 1: **Thematic analysis of strategies adopted by Signatories of the Declaration of Cooperation.** Black circles: subject appears in the National Strategy; white circles: subject not emphasised in the National Strategy. *Ratification expected December 2020.

As Table 1 demonstrates, there is a high degree of homogeneity amongst the Signatories in many categories. The areas worthy of comment relate to the countries which have not included the categories in their strategies.

Estonia and the UK are the only two countries which do not reference international cooperation. There are a number of possible explanations for this. It may be that those countries identify international cooperation as more appropriately located in other documents and strategies, or may consider that it is addressed in other arenas. For example, in September 2020 the UK and US have signed a Declaration of Cooperation on Research and Development [22]. Both Estonia and UK devote part of their strategies to the potential for the commercialisation of AI and, in relation to an international dynamic, the emphasis is on attracting inward investment. This could suggest that those countries place a higher value

on developing AI solutions within their borders than necessarily collaborating at an international level on ethics and regulation.

With the exception of the UK, all countries devote a section of their strategies to ethics, trustworthy AI, legal considerations or regulation. In most cases, the sections appear at the end of the strategies. This could be interpreted as a reflection of the esteem in which the topic is held or it might be a tactic to avoid being viewed by the private sector as being driven by regulation. Nonetheless, their inclusion indicates an acknowledgement of the importance of these areas.

Despite not including such a section, the UK has acknowledged the importance of ethics in AI and established the Centre for Data Ethics and Innovation in 2018 with the stated purpose "to advise the government on the measures which are needed to ensure safe and ethical innovation in data and AI". However, this body only provides advice to the government, without a parallel positive commitment in the National Strategy to act on the advice.

The concept of trustworthy AI arises in many strategies and several strategies refer to the report issued by the EU Commission's High Level Expert Group on Artificial Intelligence in relation to trustworthy AI. However there is no consensus in the strategies on what constitutes trustworthy AI and this aligns with the lack of consensus on what constitutes trustworthy AI generally.

There is a reasonably even spread of those countries that do and do not propose AI specific legislation. Given that the EU Commission is due to complete an impact assessment on the proposal for AI specific legislation [12], it seems that further engagement with Nation States will be required in order to gain support, should AI specific legislation be pursued.

A smaller number of countries propose a certification or auditing process. It seems likely that those countries which have proposed AI specific legislation and certification/auditing have not yet carried out detailed studies on how to draft and frame this type of legislation or process, given its complex nature. Had they already done so it is likely that reference would have been made in the strategies. The certification/ auditing process is a way to encourage the development of AI solutions by creating a new market while enacting legislation tends to operate as a way to protect consumers and society in the event of damage occurring. Identifying these options in a high level strategy is the first step which indicates a desire to adopt a legislative or certification framework.

Many Signatories whose populations were on the lower end of the scale stated that as a small country they would be obliged to follow the lead of other countries. An outlier was Malta which took a bold step in relation to certification. It stated an intention to expand the list of technologies that Malta Digital Innovation Authority to include AI technologies with a view to being the first country with a national AI certification programme. This innovative approach could stem from a variety of sources. For example, English is an official language of the country and therefore they may not be focussed on language matters which other Signatories with minority languages were concerned about. Further, its legal system is a mix of civil code and common law, which may also allow for agile legal responses.

4 Conclusion/Future Directions

This study establishes that in general, through their National Strategies and as borne out by the Table (a) the Signatories highlight AI ethics and regulation as an area requiring attention and (b) there is an expectation that Signatories themselves will attempt to regulate AI or that they will join other countries to regulate. This willingness to seek solutions in conjunction with other European countries suggests an openness to multi-jurisdiction regulation. The Estonian and UK strategies were exceptions to this and in those strategies, there was an emphasis on the business opportunities in AI.

Identifying options such as certification, auditing, standards and AI specific legislation in a high level strategy indicates a desire to subject AI solutions to a control framework. There is no clear consensus on the type of control framework that is desired. Even if there were consensus, the question of enforcement would remain. Many of the published national strategies refer to specific regulatory measures for AI and focus on product and corporate regulation. Professional regulation for AI, in which individuals are responsible for ensuring compliance, could be a complementary or alternative regulatory strategy [37] but it has not been proposed in any of the national strategies.

Bridging the gap between the definitions and concepts used by policymakers and those of AI researchers' will be important [26] when designing the regulatory framework for AI from a technical perspective. Whether the appetite and priorities seen in Europe are mirrored worldwide will require a broader study of the national AI strategy landscape. The next stage of this research, which is underway, is to expand the study to include non-signatories of the Declaration of Cooperation to evaluate, based on national strategies, the international appetite for regulation.

Acknowledgements This work was funded by Science Foundation Ireland through the SFI Centre for Research Training in Machine Learning (18/CRT/6183).

References

1. Asimov, I.: Runaround. *Astounding Science Fiction* **29**(1), 94–103 (1942)
2. Becker, S., Hasselbring, W., Paul, A., Boskovic, M., Koziolok, H., Ploski, J., Dhama, A., Lipskoch, H., Rohr, M., Winteler, D., Giesecke, S., Meyer, R., Swaminathan, M., Happe, J., Muhle, M., Warns, T.: Trustworthy Software Systems. *ACM SIGSOFT Software Engineering Notes* **31**(6), 1–18 (11 2006)
3. Bertolini, A.: Artificial Intelligence and Civil Liability. (2020), [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU\(2020\)621926_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/621926/IPOL_STU(2020)621926_EN.pdf)
4. Boden, M., Bryson, J., Caldwell, D., Dautenhahn, K., Edwards, L., Kember, S., Newman, P., Parry, V., Pegman, G., Rodden, T., Sorrell, T., Wallis, M., Whitby, B., Winfield, A.: Principles of robotics: regulating robots in the real world. *Connection Science* **29**(2), 124–129 (4 2017). <https://doi.org/10.1080/09540091.2016.1271400>
5. Cave, S., ÓhÉigeartaigh, S.S.: An AI Race for Strategic Advantage: Rhetoric and Risks. In: *AIES 2018 - Proceedings of the 2018 AAAI/ACM Conference on AI,*

- Ethics, and Society. pp. 36–40. Association for Computing Machinery, Inc (2018). <https://doi.org/10.1145/3278721.3278780>
6. Council of Europe: Information Documents - Work on Artificial Intelligence (2019), <https://www.coe.int/AI>
 7. Council of Europe: AI Initiatives. (2020), https://docs.google.com/spreadsheets/d/1mU2brATV_fg5MRGfT2ASOFepAI1pivwhGm0VCT22_U/edit#gid=0
 8. EU Commission: Declaration of Cooperation on Artificial Intelligence (2018), <https://ec.europa.eu/jrc/communities/sites/jrccties/files/2018aideclarationatdigitaldaydocxpdf.pdf>
 9. EU Commission: AI WATCH. (2020), https://ec.europa.eu/knowledge4policy/ai-watch_en
 10. EU Commission: Report on the Safety and Liability Implications of Artificial Intelligence, the Internet of Things and Robotics. (2020), https://ec.europa.eu/info/sites/info/files/report-safety-liability-artificial-intelligence-feb2020_en_1.pdf
 11. EU Commission Independent Expert Report: Ethics of Connected and Automated Vehicles. (2020), https://ec.europa.eu/info/sites/info/files/research_and_innovation/ethics_of_connected_and_automated_vehicles_report.pdf
 12. EU Commission Lead DG (Responsible Unit) CNECT: Inception Impact Assessment: Proposal for a legal act of the European Parliament and the Council laying down requirements for Artificial Intelligence (2020), <https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12527-Requirements-for-Artificial-Intelligence>
 13. Fjeld, J., Achten, N., Hillgoss, H., Nagy, A.C., Srikumar, M.: Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI. (2020), <https://ssrn.com/abstract=3518482>
 14. Flick, U.: The SAGE Handbook of Qualitative Data Analysis. SAGE Publications Inc, London (2014). <https://doi.org/10.4135/9781446282243.n12>
 15. Gouvernement de la République française: AI for humanity: French Strategy for Artificial Intelligence (2018), <https://www.aiforhumanity.fr/en/>
 16. Government of Portugal: AI Portugal - National Initiative on Digital Skills. (2019), https://www.incode2030.gov.pt/sites/default/files/julho_incode_brochura.pdf
 17. Government of the Republic of Estonia: Estonia’s National AI Strategy. (2019), <https://www.kratid.ee/in-english>
 18. Government Offices of Sweden: National Approach to Artificial Intelligence. (2018), <https://www.government.se/4a7451/contentassets/fe2ba005fb49433587574c513a837fac/national-approach-to-artificial-intelligence.pdf>
 19. Hasselbring, W., Reussner, R.: Toward Trustworthy Software Systems. *Computer* **39**(4), 91–92 (4 2006). <https://doi.org/10.1109/mc.2006.142>
 20. High-Level Expert Group on Artificial Intelligence set up by the EU Commission: Ethics Guidelines for Trustworthy AI. (2019), https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419
 21. High-Level Expert Group on Artificial Intelligence set up by the EU Commission: Policy and Investment Recommendations for Trustworthy AI. (2019), https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60343
 22. HM Government: Declaration of the United States of America and the United Kingdom of Great Britain and Northern Ireland on Cooperation in Artificial Intelligence Research and Development 25 September 2020, <https://www.gov.uk/government/publications/declaration-of-the-united-states-of-america-and-the-united-kingdom-of-great-britain-and-northern-ireland-on-cooperation-in-ai-research-and-development>

23. HM Government: Industrial Strategy Artificial Intelligence Sector Deal. (2018), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/702810/180425_BEIS_AI_Sector_Deal__4_.pdf
24. International Standards Organisation, International Electrotechnical Commission: ISO/IEC TR 24028:2020(en) Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence. (2020), <https://www.iso.org/obp/ui/#iso:std:iso-iec:tr:24028:ed-1:v1:en>
25. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. *Nature Machine Intelligence* **1**(9), 389–399 (2019). <https://doi.org/10.1038/s42256-019-0088-2>
26. Krafft, P.M., Young, M., Katell, M., Huang, K., Bugingo, G.: Defining AI in policy versus practice. In: AIES 2020 - Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society. pp. 72–78. Association for Computing Machinery, Inc (2020). <https://doi.org/10.1145/3375627.3375835>
27. Kung, J., Boskovic, G., Stix, C.: Building an AI World: Report on National and Regional AI Strategies, 2nd Ed. (2020), <https://www.cifar.ca/docs/default-source/ai-reports/building-an-ai-world-second-edition-f.pdf>
28. Latvijas Republikas Ministru kabineta: Informatīvais ziņojums "Par mākslīgā intelekta risinājumu attīstību". (2020), [http://tap.mk.gov.lv/doc/2020_02/IZ_MI\[1\].2.docx](http://tap.mk.gov.lv/doc/2020_02/IZ_MI[1].2.docx)
29. Lewis, D., Hogan, L., Filip, D., Wall, P.: Global Challenges in the Standardization of Ethics for Trustworthy AI. *Journal of ICT* **8**(2), 123–150 (2020), <https://journals.riverpublishers.com/index.php/JICTS/article/view/2645/1741>
30. Ministero dello Sviluppo Economico: Proposte per una Strategia Italiana per l'intelligenza artificiale. (2020), https://www.mise.gov.it/images/stories/documenti/Proposte_per_una_Strategia_italiana_AI.pdf
31. Ministry for Innovation and Technology: Hungary's Artificial Intelligence Strategy 2020-2030 (2020), <https://ai-hungary.com/files/de/9b/de9b302fdf9b5a57901c1fce977273b4.pdf>
32. Ministry of Economic Affairs and Employment: Finland's Age of Artificial Intelligence (2017), https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161688/41_19_Leading%20the%20way%20into%20the%20age%20of%20artificial%20intelligence.pdf?sequence=4&isAllowed=y
33. Ministry of Economics and Innovation: Lithuanian Artificial Intelligence Strategy: A Vision of the Future. (2019), <http://kurkl.lt/wp-content/uploads/2018/09/StrategyIndesignpdf.pdf>
34. Ministry of Industry and Trade of the Czech Republic: National Artificial Intelligence Strategy of the Czech Republic. (2019), https://www.mpo.cz/assets/en/guidepost/for-the-media/press-releases/2019/5/NAIS_eng_web.pdf
35. Ministry of Science Innovation and Universities: Spanish RDI Strategy in Artificial Intelligence. (2019), https://www.ciencia.gob.es/stfls/MICINN/Ciencia/Ficheros/Estrategia_Inteligencia_Artificial_EN.PDF
36. Misuraca, G., van Noordt, C.: Overview of the use and impact of AI in public services in the EU. , Publications Office of the European Union,, Luxembourg (2020). <https://doi.org/10.2760/039619>
37. NíFhaoláin, L., Hines, A.: Could regulating the creators deliver trustworthy AI? pp. 1–5 (2020), <http://arxiv.org/abs/2006.14750>
38. Norwegian Ministry of Local Government and Modernisation: National Strategy for Artificial Intelligence. (2020), https://www.regjeringen.no/contentassets/1febbb2c4fd4b7d92c67ddd353b6ae8/en-gb/pdfs/ki-strategi_en.pdf

39. OECD: Recommendation of the Council on Artificial Intelligence (2019), <https://legalinstruments.oecd.org/api/print?ids=648&lang=en>
40. Office of the Deputy Prime Minister of the Slovak Republic: Strategy of the Digital Transformation of Slovakia 2030. (2019), <https://www.mirri.gov.sk/wp-content/uploads/2019/11/Brochure-SMALL.pdf>
41. Oireachtas Éireann: Liability For Defective Products Act, 1991, <http://www.irishstatutebook.ie/eli/1991/act/28/enacted/en/html>
42. Oireachtas Éireann: Sale of Goods Act, 1893, <http://www.irishstatutebook.ie/eli/1893/act/71/enacted/en/print>
43. Oireachtas Éireann: Sale of Goods and Supply of Services Act, 1980, <http://www.irishstatutebook.ie/eli/1980/act/16/enacted/en/html>
44. Oireachtas Éireann: S.I. No. 11/2003 - European Communities (Certain Aspects of the Sale of Consumer Goods and Associated Guarantees) Regulations 2003, <http://www.irishstatutebook.ie/eli/2003/si/11/made/en/print>
45. Oireachtas Éireann: S.I. No. 199/2004 - European Communities (General Product Safety) Regulations 2004, <http://www.irishstatutebook.ie/eli/2004/si/199/made/en/print>
46. Parliamentary Secretariat for Financial Services Digital Economy and Innovation: Malta the Ultimate Launchpad - a Strategy and Vision for Artificial Intelligence in Malta 2030. (2019), https://malta.ai/wp-content/uploads/2019/11/Malta_The_Ultimate_AI_Launchpad_vFinal.pdf
47. Rijksoverheid: Strategisch Actieplan voor Artificiële Intelligentie. (2019), <https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/beleidsnotas/2019/10/08/strategisch-actieplan-voor-artificiele-intelligentie/Rapport+SAPAI.pdf>
48. Rossi, F.: Building Trust in Artificial Intelligence. *Journal of International Affairs* **72**(1), 127–134 (2019)
49. Rząd Polski: Polityka Rozwoju Sztucznej Inteligencji w Polsce. (2020), <https://www.gov.pl/attachment/0aa51cd5-b934-4bcb-8660-bfecb20ea2a9>
50. Shaffer, G.C., Pollack, M.A.: Hard vs. Soft Law: Alternatives, Complements, and Antagonists in International Governance. *Minnesota Law Review* **94**(3), 706–799 (2010)
51. Slovic, P.: Risk Perception. In: *Carcinogen Risk Assessment*, pp. 171–181. Springer US (1988). https://doi.org/10.1007/978-1-4684-5484-0_13
52. Slovic, P., Fischhoff, B., Lichtenstein, S.: Behavioral decision theory perspectives on risk and safety. *Acta Psychologica* **56**(1-3), 183–203 (8 1984). [https://doi.org/10.1016/0001-6918\(84\)90018-0](https://doi.org/10.1016/0001-6918(84)90018-0)
53. The Danish Government, Ministry of Finance and Ministry of Industry Business and Financial Affairs: National Strategy for Artificial Intelligence. (2019), https://en.digst.dk/media/19337/305755_gb_version_final-a.pdf
54. The Federal Government of Germany: Artificial Intelligence Strategy. (2018), www.ki-strategie-deutschland.de
55. The Government of the Grand Duchy of Luxembourg: Artificial Intelligence: a strategic vision for Luxembourg. (2019), <https://gouvernement.lu/dam-assets/fr/publications/rapport-etude-analyse/minist-digitalisation/Artificial-Intelligence-a-strategic-vision-for-Luxembourg.pdf>
56. Τμήμα Ηλεκτρονικών Επικοινωνιών: Εθνική Στρατηγική Τεχνητής Νοημοσύνης (TN): Δράσεις για την Αξιοποίηση και Ανάπτυξη της TN στην Κύπρο. (2020), <https://dec.dmid.gov.cy/dmid/dec/dec.nsf/All/21122CD12D52BD70C2258505002E43D1>
57. UNESCO: Recommendation on the Ethics of Artificial Intelligence. (2020), <https://unesdoc.unesco.org/ark:/48223/pf0000373434/>