



Rialtas na hÉireann
Government of Ireland

 Ref. Ares(2020)3359830 - 26/06/2020

Ireland's National Submission to the Public Consultation on the EU White Paper on Artificial Intelligence

Prepared by Department of Business, Enterprise and
Innovation

June 2020

Ireland's National Submission to the Public Consultation on the EU White Paper on Artificial Intelligence

1. This paper provides the response of the Irish authorities to the Commission's White Paper on AI. The issues raised by the White Paper are of high relevance to Europe's economy and society at this time and Ireland welcomes the opportunity to comment on them. AI is of particular importance to Ireland, Europe and indeed globally, both in providing opportunities to drive productivity but also in benefitting society through the applications based upon it.

2. Ireland agrees that the two issues raised in the White Paper are of critical importance. These issues are how to encourage the adoption of the benefits of AI through the ecosystem of excellence and the need to consider regulation in order to address perceived risks that AI may represent through the ecosystem of trust. Indeed, these ecosystems are mutually interdependent not least because trust in AI is an essential condition for its adoption and use.

Key Messages

3. Ireland is fully committed to the aims of the EU Coordinated Plan on AI and therefore welcomes the contents of the White Paper which set out the means to establish an ecosystem of excellence.

4. Skills are an essential support for the adoption of AI. Without the necessary skills, it will not be possible to develop or deploy AI in the market. Ireland stresses the need to develop skills across all appropriate areas of the ecosystem. These are not just the technical skills required to develop AI but also complimentary skills including regulatory skills and those necessary to interpret AI in its deployment including specific sectoral related skills.

5. AI has the ability to assist SMEs in terms of increasing productivity. In this regard, Ireland is fully supportive of the aims of the creation of a network of European Digital Innovation Hubs to advise and support SMEs in the development and adoption of AI.

6. When considering the necessary technical infrastructural support necessary for the deployment of AI, Ireland advocates continued awareness and leverage of opportunities presented by the latest technological advances in areas such as high performance, edge and quantum computing.

7. As regards the ecosystem of trust, it is felt that the White Paper fails to provide a compelling case for prioritisation for the amount of regulation suggested within it. Many areas in which applications utilising AI are under development have yet to achieve significant traction in the market. Until that happens it is suggested that it is very difficult to stand over any perception of risk because there is insufficient coherent data on which to build an evidence base.

8. It is also considered that there should be a clear understanding of the existing regulatory acquis, not just with regard to liability issues, but also the numerous sectoral pieces of regulation that the Commission have referred to as in need of review. Without this clear understanding it is not possible to determine where new regulation needs to be brought to bear. This should be fully taken into account in any future regulatory proposals so as to ensure consistency of approach and avoidance of unintended consequences.

9. The White Paper lacks explicit mention of the need to comply with international human rights law. Although references are made to fundamental rights, the language of the White Paper appears to restrict these to 'fundamental values' and 'European values'. This raises a concern that AI could fail to receive adequate oversight under international human rights law. In addition, some greater mention of how to cohere the European approach with universal rights at large would be welcome, given that AI development in Europe has the potential to have a global impact.

10. Rather than trying to evaluate all possible consequences in an ex-ante fashion which, of its nature, is difficult and is likely to give rise to considerable uncertainty in implementation, it might be appropriate to adopt an approach that leans more upon incremental learning. This should lead to more emphasis on approaches such as regulatory sandboxes as a way of managing the risks of high-tech developments in an appropriate setting. It might also represent a more agile and collaborative approach than the independent audit envisaged by the White Paper. It is also an approach that is in keeping with the lab to market process set out in the EU Coordinated Plan on AI and could have synergies with the creation of international standards.

11. In addition to the approaches contained in the White paper, is suggested that:

- Suitable design-based approaches to AI regulation should be explored, following the example of and, where appropriate, incorporating data protection by design principles.
- As regards the references to the work ongoing in other multilateral/international fora, it is important that, as far as possible, the various approaches being identified are consistent between themselves.

12. More practically it has been pointed out that, now it has left the EU, the UK is unlikely to emulate the regulatory approach we are considering and will be a (third) natural competitor for AI-related start-ups and businesses in addition to US and China. The UK Government has committed to 'work with businesses to develop an agile approach to regulation that promotes innovation and the growth of new sectors.

<https://www.parliament.uk/documents/lords-committees/Artificial-Intelligence/AI-Government-Response2.pdf>. If the EU adopts an unduly restrictive approach, Europe may be at a serious competitive disadvantage in relation to sourcing AI-related investment opportunities and developing partnership approaches between Member States.

Response to White Paper

13. The White Paper address two ecosystems. The **ecosystem of excellence** addresses issues around how to encourage the adoption of the benefits of AI while the **ecosystem of trust** is intended to consider the need to regulate in order to address perceived risks that AI may represent. This document will therefore follow that structure.

The ecosystem of excellence

14. Ireland is fully supportive of the need to engage **Member State's cooperation** through the working of the EU Coordinated Plan on AI. We believe that this provides the opportunity to maximise the impact of investments while leveraging the **research and innovation community**. It will also have the effect of creating synergies between and increasing the excellence of R&I facilities, which in turn will both attract the best talent and produce the best technology.

15. The issue of AI definitions is dealt with in more depth under the ecosystem of trust but in the context of the ecosystem of excellence, clarity should be provided as to what is meant by AI. It is essential to define AI as accurately and appropriately as possible to ensure that the right funding is going to the right projects and that these are given the necessary time and space in the research, innovation and deployment facilities being developed as recommended in the EU Coordinated Plan on AI.

16. On the matter of ensuring that we have the appropriate **skills**, which is vital in supporting an ecosystem of excellence, there are several areas where further consideration is warranted. There is a need to make more explicit the requirement for **non-technical** skillsets to effectively deploy, manage and regulate AI for economic and public benefit.

17. Most of the AI training programmes worldwide are based on the fundamentals of data science, machine/deep learning, and ethical aspects using generic use-cases and examples. While they build the fundamental skills of the AI workforce, a significant gap continues to exist in the immediate employability of the trainees and their ability to apply their AI skills directly on real-world research and business problems in various domains such as environmental sciences, financial services, biotechnology, material sciences and autonomous systems. Thus, there is a need to customise the curricula of training programmes to address the specific AI skills required for such specific domains, particularly focusing on relevant datasets, problems, AI algorithms, and implications on privacy, ethics and security aspects.

18. However, a similar concern arises with regard to the need for specialist curricula in technical fields. For instance, the methods and tools for dataset preparation, AI algorithms, AI solution deployment and ethical/privacy implications for areas such as environmental sciences, healthcare, material sciences are radically different. Thus, developing AI skills for

each of these domains needs customisation of the curriculum of the training programmes to address the requirements and specificities of the problems in each domain.

19. In addition, the skills needs of the public and private sector stakeholders, including strategists and managers of companies implementing AI, as well as Government regulators, need to be taken into account, as they will be critical to building the 'ecosystem of trust'.

20. The following sector specific skills needs have also been highlighted:

- As regards autonomous driving, there are needs for both vehicle drivers and those working in public authorities to develop AI related skills or at least a level of AI-literacy.
- Determining and illustrating the benefits of the use of AI for education; for the learner; the teacher and the educational institution and for the wider society. There is a need to ensure that teachers/education professionals are engaged in determining the relevance of AI in their work and pedagogical practices.

21. With regard to the **focus on SMEs** Ireland is keen to play its part in the network of future European Digital Innovation Hubs which will provide support to SMEs to understand and adopt AI which is of critical importance to improving SME productivity levels. This work is progressing towards designating hubs in time to allow the selected hubs to commence their work in 2021.

22. Ireland encourages the development of a **Public-Private Partnership** in AI, data and robotics **that** will harness the efforts of the private sector through R&I and co-investment.

23. As regards **Public Sector adoption of AI** we consider that in matters of procurement, in order to earn public trust in the roll out of AI systems in the public sector, consideration should be given to approaches integrating principles of trustworthy AI design in the procurement process.

24. More specifically, in relation to the creation of an ecosystem of excellence strong consideration should be given to the need to build capacity and expertise across the **Justice sector**. In this respect actions 1, 2 & 4 of the White Paper should take account of discussions held by the European Council on setting up an EU Innovation Hub for Internal Security. It will also be necessary to develop and deploy an AI/data science workforce for the Justice Sector to support the development and uptake of innovative technologies by law enforcement, the judiciary, court administrators and legal practitioners.

25. On **access to data and computing infrastructures**, it is suggested that an area that could be addressed going forward is the optimisation of data management and AI software for edge computing devices, which have performance and energy limitations, and seamlessly integrating the workflows of edge processing with cluster-/cloud-based HPC systems while preserving data privacy and governance policies.

26. Looking further ahead, quantum computing is foreseen to be the next genuine disruption. There are major initiatives worldwide to develop practical quantum computing platforms. To this end, it is important to (1) integrate quantum computing platforms with conventional HPC infrastructure, (2) provide national access with user support, and (3) develop capabilities to leverage quantum computing for AI challenges.

27. As regards **international aspects** thought might be given to leveraging the EU's ethical approach in order to obtain international support from bodies such as the OECD.

The ecosystem of trust

28. The White Paper sets out that the ecosystem of trust will be created by key elements of a future European regulatory framework for AI and is intended to give citizens confidence to accept AI. These rules are intended to address "high risk" AI systems and will adopt the human-centric approach advocated in the High-Level Expert Group's Ethical Guidelines.

Problem Definition

29. The issues that should be taken into account are set out in the Paper as risks to fundamental rights, with issues regarding, for example, discrimination by race and gender, privacy and data protection, as well as safety and liability related risks.

30. In order to make this message more accessible it is suggested that there should be a recognition of a 'do no harm' principle in the use of AI; that is, that the application of AI should, in the achievement of positive goals, avoid exacerbating disparities, avoid discrimination among persons and eschew creating or exacerbating environmental degradation.

31. Fundamental rights issues traverse sectoral boundaries and require a cross governmental and cross-sectoral approach both domestically and at EU level. Appended to this submission is a document highlighting particular use cases relating to different implications of the consideration of fundamental rights. Firstly, the implications may be of a primarily positive nature, but the delivery of the potential benefits can be called into question by other negative connotations brought about by competing rights issues. Secondly, the wide-ranging nature of fundamental rights means that there are a plethora of concerns that can potentially lead to the classification of a use as high risk and the possibility that rights issues therefore become all encompassing. This would clearly be at odds with the stated intentions of the White Paper about not stifling innovation. In order to avoid this, it would be reasonable to have expected some suggested criteria as to the consideration of fundamental rights in arriving at decisions as to the identification of high-risk applications might have been set out in the White Paper.

32. It has been noted that there is a lack of explicit mention of the need to comply with international human rights law. Two issues that arise are:

- While acknowledging that human rights are referred to as fundamental rights the language of the White Paper appears to restrict these to 'fundamental values' and 'European values'. This raises a concern that AI would fall outside the realm of scrutiny within the international human rights framework, and, by doing so, fail to receive adequate oversight under international human rights law;
- Development of AI within Europe has the potential to have a global impact, e.g. users of AI systems developed in Europe will not be limited to European citizens who have their rights guaranteed under the Charter of Fundamental Rights, European Convention of Human Rights. Therefore, some larger mention to how to tie the European approach into universal rights at large would be welcome;

Suitability of current legislative regime

33. Europe already has a **significant body of regulation** which applies, inter alia, to AI. The Commission has undertaken to collate this and Ireland would contend that further regulation should not be undertaken until this work has been undertaken and published. Some of this *acquis* is sectorally specific while other parts are more generally applicable in areas such as consumer and data protection. This *acquis* should be reviewed to determine whether it needs to be updated to take into account aspects of the digital transformation and should be fully taken into account in future regulation to ensure consistency of approach and avoidance of unintended consequences due to duality of approach between specific pieces of regulation.

34. In addition to the body of regulation many Member States have a body of case law which will be of relevance. It is of interest to note that currently tort rules are largely non-harmonised and remain a national competence. The thrust of the white paper suggests that harmonisation, perhaps by way of a new legislative instrument, would be desirable in the case of AI systems, not least to support the internal market. Support, or otherwise, for such an instrument would be critically dependent on the scope and content of the rules contained therein. Also in this respect a report issued by the Expert Group on Liability and New Technologies – New Technologies Formation¹ has concluded that, inter alia, operators of AI products (such as robots, cars) should be strictly liable for damage resulting from the operation of these products and that a producer of an AI product should still be liable for damage caused by defects in the product even if the defect was caused by changes made to the product after being placed on the market, as long as the producer was still in control of updates/upgrades to the technology. Any implementation of this report should be carried out with great sensitivity to ensure that innovation is not stifled by such an absolutist approach.

35. The White Paper describes certain areas in which the current *acquis* is capable of being improved. We support addressing the issue that, within the scope of product safety

¹<https://ec.europa.eu/transparency/regexpert/index.cfm?do=groupDetail.groupMeetingDoc&docid=36608>

legislation, software is typically not considered a product. This is particularly so with regard to online products such as applications. If such a situation is allowed to pertain there will be a significant impact on the concept of safety in the area of AI and in the creation of new risks, for example risks arising from cyber-attacks on AI products.

Scope of a future EU regulatory framework

36. The White Paper states that in order to develop a suitable framework there is a need for a **definition** of AI that is flexible enough to allow technical progress while also being precise enough to provide legal certainty. Defining AI is a key issue which affects all the proposals of the White Paper not just those that relate to regulation. Furthermore, definitions should be agreed by all the stakeholders to facilitate communication and to assist identification of inherent biases. On the face of this it is necessary to classify the various types and uses of AI. One clear candidate for this approach would be machine learning which is differentiated from mere programming in that the machine is trained on data as opposed to programmed to carry out a task in a specific manner.

37. Various sectoral issues have been raised with regard to the definition of AI:

- Education. The definition of AI will affect what we mean by competencies/skills for the effective use of digital technologies in teaching and learning and our understanding of 21st century skills.
- Consumer/Product safety. While acknowledging that agreeing on a definition of AI would be difficult in the extreme there is still a need to be as precise as possible when discussing possible changes to legislation or the introduction of new EU harmonised laws in the area of AI.² In order to meet their obligations, Market Surveillance Authorities (MSAs) need to have well thought out and comprehensive legislation in order to enforce safety provisions in the EU and that begins with careful definitions in the legal text.

38. The framework proposed for the ecosystem of trust is expected to follow a risks-based approach, focussing on “high-risk applications”. Differentiation according to risk helps assure proportionality.

39. The White Paper sets out a position that the definition of high-risk should rely on the combination of the sector (“where, given the characteristics of the activities typically undertaken, significant risks can be expected to occur) and use (“where the use is of such a manner that significant risks are likely to occur”). Mandatory requirements would only apply to a system which met both factors. The existing acquis, it states, would continue to apply to low-risk applications.

² See page 1 of the European Strategy for AI published April 2018 for a possible definition: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN>

40. It would appear to be essential to develop a range of criteria both for the identification of high-risk AI and the identification of the relevant sectors. None has been supplied. Without such criteria, the 'legal certainty' to which the White Paper aspires will not be attained.

41. In any event, the concentration on sectors, and certainly the requirement that the AI should not only be a high risk use but should also be developed in a high risk sector, is likely to cause problems to be missed where the application is capable of being used in a number of sectors but a high-risk use, possibly for purposes for which the AI was not originally intended, arises in a non-high-risk sector.

42. Furthermore, it is suggested that given the range of potential high-risk sectors how is it meaningfully possible to prioritise between these without ending up with an overly prescriptive set of proposals. It is suggested that these issues justify concentrating upon high-risk uses per se. Concentration on use allows potential regulation to focus on two factors; the purpose for which the AI is designed and other uses to which it can be put.

43. Dealing firstly with the intended purpose, this can be broken down into two further considerations. Firstly, is the purpose consistent with the body of fundamental rights? The second issue relates to the robustness of the AI in carrying out the purposed task. In this respect it will be important to fully define what is required to show that the AI in a product is robust. It is suggested that this should go beyond mere fitness for purpose, accuracy or consistency but should contain some form of undertaking from the developer to provide the necessary connectivity and updates to software and data to support the AI for its stated purpose.

44. Where an AI is put to use for other than its intended purpose there needs to be a careful allocation of responsibilities/liabilities between producer and user based on concepts such as foreseeability/duty of care. In certain circumstances, where the AI is put to a use other than its intended one, without reference to the producer, there is an argument that strict liability may be appropriately fixed to the actor responsible for the adoption of that use.

45. It is suggested that it would be preferable to develop a horizontal **framework** from which would hang a number of specific vertical interventions. These verticals would concentrate on behaviours which should be prevented and be designed to tackle certain high risk uses, such as, perhaps the deployment of artificial intelligence using identification through biometric features in a public space. The horizontal approach should deal with matters such as product safety, consumer protection, liability issues and guidance to developers as set out in the assessment list developed in the HLEG Ethical Guidelines as updated from time to time. It is accepted also that AI systems may, require repeated assessments over the lifetime of their application, including to mitigate the potential for possible emerging risks. This is particularly true with machine learning systems; as target variables and the specific of the environments in which they operate change, the systems predictions will become less accurate unless appropriate monitoring and readjustments take place.

46. By adopting such a horizontal approach with the associated verticals referred to above it is less likely that such a loophole would be created. In fact, the attempt by the White Paper to deal with these issues in a catch-all approach is possibly misplaced. Such an approach in a fast-evolving space may be insufficiently flexible to deal with new developments.

47. Ireland supports the call made by the European Commission's Advisory Committee on Equal Opportunities for Women and Men for 1) measures to ensure that the data used to train algorithms is unbiased and of a good quality, and 2) for research into 'non-discrimination by design' algorithms that "can incorporate equal treatment laws into their decision-making processes to ensure that no direct discrimination is possible and that no decision is based directly on gender/sex", and notes the recommendation that "National equality bodies and other actors could explore ways to apply existing legislation in a way that could help create suitable jurisprudence on AI."

Sectoral Issues

48. Attention has been drawn to other sectoral specific concerns raised in designing a regulatory framework.

- **48.1 In the Justice sector**, when regulatory intervention is to be targeted at areas where risks are more likely to occur, it would seem to be the case that the use of AI in relation to aspects of courts business could potentially be of high risk and might well justify regulatory intervention. However, caution is needed until such time as it is clear what such intervention might entail.
 - Also, the development of mechanisms by which to judge whether AI is leading to discriminatory practices is considered prerequisite to the effective development and deployment of high-risk AI applications in the Justice sector. The extent to which current statistical modelling techniques could achieve this, or whether a bespoke formal process is required needs to be determined.
- **48.2 The following issues have been raised with regards to Competition and Consumer Protection.**
 - A key factor in the ability to apply the existing competition rules will be transparency regarding automated decision making and algorithms.
 - To give a prominent example of the application of AI to competition law, it is useful to look at the use of AI-powered pricing. The possibility of inadvertent collusion between firms driven by the use of the same AI enabled products also exists within this consideration. If the use of algorithms leads to a coordination of prices between different undertakings, Art. 101 of the Treaty on the Functioning of the European Union (TFEU) might apply. This provision does not require that an undertaking possesses market power but that more than one undertaking is involved in the coordination. Article 101 TFEU prohibits coordination between undertakings and spells out three alternative modes of coordination as being forbidden, namely: "agreements",

“decisions by associations of undertakings” and “concerted practices”. The latter is most likely to involve the use of AI.

- By contrast, Art. 102 TFEU when applied to the use of AI, requires market dominance on the part of the undertaking in question and can, for example, be applied if an undertaking which possesses market power uses price algorithms for implementing a price discrimination strategy vis-à-vis its competitors and/or customers. The White paper points to the availability of data and ability to access relevant datasets for AI use as a strong focus in the ecosystem of excellence. This must be considered as part of a broader initiative at EU level reflected in the European Data Strategy as well as the debate surrounding ex-ante regulation of platforms.
- Risks to consumers have been identified from services using AI technology where they could be targeted in a malign way and enticed to purchase certain products or services. These risks are already apparent and should also be considered a priority for action at EU level.

- **48.3 Concerns in the Transport sector** are highlighted as follows:

- A European regulatory framework for AI could be very useful in the area of automated driving. The issue of enforcement authorities or affected persons lacking the means to verify how a given decision involving the use of AI was made is relevant in the case of automated driving.
- While there is support for the need for a common legislative framework and EU approach to AI in the case of automated driving, this may even need to be a global approach and work in bodies such as the United Nations Economic Commission for Europe is very relevant.
- Prior conformity assessment is very relevant to automated/autonomous vehicles. However, at present it is a concern as to whether there is sufficient expertise or capacity to provide this. A European governance structure on AI and the development of standards could help in this regard.

- **48.4 In the Education Sector**, in the context of devising a regulatory framework it is important to identify and eliminate ethical concerns that are specific to education and to identify how data can help to inform policy direction and actions.

Types of requirements

49. The High-Level Expert Group's (HLEG) Ethical Guidelines and the results of the piloting process on these guidelines suggest a number of requirements. In this regard, particular support is expressed for ensuring human validation and the need for human intervention in the ongoing process of applications. The revisiting of the HLEG work on ethics is welcomed with an expectation that its contents will be more concise and accessible. A concern has also been expressed as to whether Government will have to open up its decision-making processes to scrutiny, especially where they are being used to prevent a crime, for instance

in such disparate areas as fishing quotas or monitoring agricultural activities that impact climate change.

50. As regards the development of transparency requirements relating to the operation of algorithms, due regard should be paid to developing intellectual property legislation in the area of Artificial Intelligence. The proper protection of intellectual property rights to create legal certainty and to promote innovation and creativity in the EU economy should be ensured.

Addressees

51. The White Paper states that, “each obligation should be addressed to the actor(s) who is (are) best placed to address any potential risk.” While this is a laudable statement it remains to be seen what it will actually look like in practice and this issue will need careful monitoring when it comes to actual discussion of the legislation at Council level. It may not always be easily apparent what economic operator will be in the best place to address risks due to the blurring of boundaries between the various actors involved in producing, adapting and adopting AI technology.

52. A definition of the specific circumstances where a consumer may be contributing towards product failure (for example by deliberately failing to install updates they had been notified about for a smart product) would also increase clarity on the respective responsibilities of producers and consumers.

Compliance and Enforcement

53. The White Paper sets out a rationale for prior conformity assessment for certain high-risk AI applications. The conformity assessments would include procedures for testing, inspection and calibration that already exist for other products placed on the market. The White Paper sees this conformity assessment being carried out by existing notified bodies. If this approach were to be pursued, we believe that it would be important to ensure that national accreditation bodies acquire the necessary expertise in order that they are able to audit AI systems in accordance with the relevant requirements identified.

Voluntary Labelling for non-high-risk AI applications

54. Such a voluntary scheme would involve interested parties subjecting their product/service to assessment similar to the mandatory requirements or as otherwise considered appropriate. This would allow them to show that their AI is trustworthy and give users confidence and promote uptake of the technology. Once the interested party entered the voluntary scheme, the requirements would be binding upon them

55. In respect of the Voluntary labelling scheme and with relevance to the conformity assessment recommendations, for any such scheme to be meaningful and reliable, prerequisites need to be addressed. These are; a body of implementable safe harbour standards for trustworthy AI, and robust market-based conformity assessment schemes to

demonstrate compliance with those standards. Without these it is considered that it will not be possible to introduce a consumer-facing ethics labelling scheme.

56. The international and European standardisation community needs to develop implementable requirements and process guidance standards for trustworthy AI. This work is already underway in at an International Standards level in ISO/IEC JTC 1/SC 42 *Artificial Intelligence*³ and will be available for CEN/CLC⁴ European adoptions gradually in the future but not sooner than in 2 years. Ethical and human rights principles as promulgated, for example, in the International Standard ISO 26000:2010 *Guidance on social responsibility*⁵, need to be mapped onto implementable and verifiable requirements and process guidance standards in the areas of AI development, risk management, business use and organizational governance. Only preliminary informative work exists in this area at the moment.

57. When a reliable body of knowledge and safe harbour standards for ethical principles preserving AI exist, a conformity assessment scheme can be developed to enable reliable market surveillance, as well as verifiable and auditable (second and) third party conformity assessment and certification.

58. Without the above described prerequisites, there is a considerable risk that, any potential ethics labelling scheme will be semantically empty and effectively undermine end user and consumer trust in AI as used by industry and governments. Notwithstanding the criticism, government enforced labelling and market surveillance will need to be preceded by a voluntary labelling period based on the same body of safe harbour standards that will iron out any interoperability and auditability issues before the scheme can start to be normatively referenced and enforced by regulators.

59. The question arises as to whether a labelling scheme is entirely unsuitable for the AI market? Although it may be a scheme that is easily recognisable for the consumer, it might be simply too basic for attaching to an AI product, even low-risk ones. It is suggested that the use of a CE marking-type scheme would be more suitable as the regulatory framework and market surveillance structures are already in place.

Governance

60. The White Paper envisages **governance being provided by Member States national authorities supported by specially designated test centres to conduct the conformity assessments** including possibly those licensed to provide assessment outside the Community. A European governance structure, incorporating these bodies and involving

³ ISO/IEC JTC 1/SC 42 Information Technology – Artificial Intelligence
<https://www.iso.org/committee/6794475.html>

⁴ CEN and CENELEC are the European Standards Organisations
<https://www.cenelec.eu/Pages/default.aspx>

⁵ <https://www.iso.org/standard/42546.html>

stakeholder participation, could provide support by sharing information and best practice as well as sourcing a panel of experts. Existing sectoral bodies should be incorporated into this structure.

61. Market Surveillance Authorities (MSAs) could be hampered by a lack of technical knowledge and skill when it comes to inspecting and testing AI products. There is also a risk that EU harmonised standards will not be capable of assessing whether AI products comply with the safety aspects of existing legislation.

62. Product Safety is based on MSAs having access to identified economic operators in the EU (currently: manufacturer, authorised representative, distributor, importer with the addition of a fulfilment service provider from June 2021). The characteristics of AI products have the potential to muddy the waters when it comes to identifying a responsible economic operator for a product that has the capability of adapting and modifying itself. This lack of clear traceability and accountability gives rise to a serious risk from the point of view of MSAs. The ability to hold an economic operator accountable for the safety of a particular product is essential in ensuring that the product safety legislative regime can function in an efficient way across the EU.

63. It could be beneficial in devising a governance system for any regulatory framework to include some reference to wider society through, perhaps, national citizens assemblies. Government's definition of what is trustworthy can be biased by its particular knowledge and perspective. It would also be advisable to ensure that all citizens are represented, particularly when AI is being deployed in the provision of public services.

Synergies with other Commission initiatives

64. Digital issues pervade the majority of areas of society and the economy. It is, therefore, not surprising that some of the issues covered in the White Paper on AI also arise in other European initiatives in other digital spaces, specifically the proposals for a Digital Services Act and a Data Strategy. This section seeks to capture some of the instances where this occurs.

The Digital Services Act Package

65. Automation, involving the use of AI applications, allows online services to operate complex systems for the delivery and moderation of content at scale often in conjunction with periodic review of the parameters of the systems and in certain instances targeted human input and oversight. Online services are increasingly viewed by their users as quasi-public spaces. The discourse within those spaces can significantly impact our societies. Often this requires online services to respond in a swift and effective manner to detect and remove illegal content. It is suggested that some consideration of the implications of how they carry out such content moderation through the use of AI is likely to arise in the consideration of potential regulation of AI. Consistency with the approach taken to responsibilities of information society service providers within the DSA will be important.

66. Like with many areas of activity, inherent risk in the case of online services arises from the nature of the activity in question rather than the specific use of AI systems to carry out that activity. However, the use of AI systems, which tend to prioritise speed and operate with less capacity to comprehend context than a person, may amplify these inherent risks, increasing the risk profile of the activity. It is not unlikely that this is the case in respect of content delivery and content moderation. Accordingly, content delivery and content moderation should be identified as a high-risk application of AI systems.

67. Similarly, risks to consumers have been identified from services using AI technology where they could be targeted in a malign way and enticed to purchase certain products or services. These risks are already apparent and should also be considered a priority for action at EU level.

68. There is also the risk posed by data capture and the danger of SMEs losing out from a competitive point of view to larger entities due to an inability to access or hold a sufficient amount of data. This could lead to a growing gap in productivity improvements. The development of high-quality AI technology is dependent on the ability to harvest relevant datasets in sufficient amounts and that gives an advantage to the larger platforms. This should be considered as a priority area to be addressed by any future AI strategy and the policy identified in the White Paper through the ecosystem of excellence to use EU Digital Innovation Hubs to provide support to SMEs in this context is welcome.

A European Data Strategy

69. There is concern that one of the main reasons that European data is underutilised may be nervousness around GDPR. This may mean we do not fully exploit our healthcare data, for example, to improve public health. The Commission needs to provide a clear guidance framework which helps users through the privacy versus value dilemma. There is considerable narrative about GDPR but a lack of clear examples of instances where the needs of public interest can be met, and proportionality respected and what type of safeguards should be used. This would help create consistency of understanding and application across Europe especially where some consider that the provisions of GDPR provide sufficient regulation on AI given that AI relies on data to work.

APPENDIX: Artificial Intelligence (AI) Use cases and potentials impact on human rights

Healthcare diagnostics

One of the main areas of application of AI is in the area of healthcare is diagnostics. Image recognition technology can now assist healthcare professionals in quickly and accurately diagnosing serious illnesses leading to better outcomes for patients. In the Irish context, in 2008, the Health Service Executive initiated a programme called [NIMIS: National Integrated Medical Imaging System](#), to capture and store Radiology, Cardiology and other diagnostic images electronically. The HSE's vision was to modernise the delivery of diagnostic Imaging services throughout the public healthcare system.

These AI applications can streamline clinical processes and allow for human resources to be used more efficiently. However, despite the best of intentions with such technologies, there can be drawbacks in terms of their impact on human rights. The main rights that could be impacted negatively are the right to privacy and data protection. This is because of the highly sensitive nature of the information handled by such systems. AI-based diagnostic systems require the collection of vast quantities of sensitive data relating to an individual's often immutable health characteristics, raising serious privacy concerns. Other rights that may be affected positively are:

- Right to Life, Liberty, and Security of Person - AI-based diagnostic systems enhance the enjoyment of the right to life by making accurate, high-quality diagnostic services more widely available.
- Right to Desirable Work - The improved health outcomes that AI-based diagnostic systems are likely to produce will reduce the number of people who are excluded from the dignity of work for medical reasons.
- Right to an Adequate Standard of Living - By detecting diseases earlier and more accurately, AI-based diagnostic systems will improve living standards and quality of life.
- Right to Education - Should AI-based diagnostic systems deliver on their promise, fewer people will be excluded from the enjoyment of the right to education due to ill-health.

Criminal justice

Criminal justice systems around the world are utilising AI to improve and efficiency and efficacy of their operations. There are obvious benefits in terms of crime detection and prevention but the impact on civil liberties and human rights could be far-reaching. An

example of such a technology is the Harm Assessment Reduction Tool (HART)⁶, developed in-house by Durham Constabulary in the UK in collaboration with the University of Cambridge in 2015/16 and deployed across the force at the point of custody decision. The aim of HART is to identify a middle stratum of risk where individuals do not need to be charged, and to reduce the number of people entering the justice system, and by doing so, hopefully reducing the number of people re-entering it.

The growing use of AI in the criminal justice system risks interfering with rights to be free from interferences with personal liberty. Risk-scoring systems are not prescribed by law and use inputs that may be arbitrary, hence decisions informed by these systems may be unlawful or arbitrary. By rating a defendant as high or low risk of (re)offending, they attribute a level of future guilt, which may interfere with the presumption of innocence required in a fair trial. When individuals are denied bail or given a certain sentence for reasons they will never know and that cannot be articulated by the government authority charged with making that decision, trials may not be fair, and this right may be violated. Some of the human rights positively and negatively impacted on include:

- Freedom from Discrimination / Right to Equality Before the Law - These systems may reproduce and perpetuate biases in the training data, but other evidence suggests they may reduce racial and other disparities in bail and sentencing, nonetheless.
- Right to Life, Liberty, Personal Security - Low-risk individuals may benefit from greater pre-trial release and shorter sentences, and the community might benefit from a lower crime rate.
- Freedom from Arbitrary Arrest, Detention, and Exile - Complex algorithms may erroneously classify certain individuals as “high-risk,” raising the possibility of arbitrary pre-trial or post-conviction detention.
- Right to Fair Public Hearing / Right to be Considered Innocent Until Proven Guilty - The proprietary nature of these tools, their inherent complexity, and the inscrutability of the results they produce makes it hard to challenge them in court.
- Right to Privacy - Automated risk assessment systems are premised on the collection, storage, and analysis of vast amounts of personal data, which raises significant privacy concerns.

Facial recognition technology (FRT)

This area of AI is one where there has been significant technological progress. FRT technologies can now out-perform human operators in this domain. It is a form of artificial intelligence that has direct application in law enforcement, border control and in the detection

⁶ See pages 45 – 47 in <https://www.lawsociety.org.uk/support-services/research-trends/algorithm-use-in-the-criminal-justice-system-report/>

of social welfare fraud⁷. Its main advantage is that it can streamline the process of identifying possible offenders or those on a 'watch list'. It can also be operated in a live scenario, comparing the faces in a crowd to a database, something that would not have been possible heretofore.

However, there are several drawbacks to FRT mainly because it is driven by biometric data. Indiscriminately capturing such data in public spaces could potentially lead to serious rights violations. Some of the rights that may be compromised by this technology are:

- Respect for private life and protection of personal data
- The right to non-discrimination
- The rights of children and the vulnerable
- Freedom of expression, assembly and association
- Right to good administration
- Right to effective remedy

⁷ <https://www.independent.ie/irish-news/politics/facial-recognition-software-catches-serial-fraudsters-31262647.html>