
## Introduction

The Information Accountability Foundation ("IAF"), a non-profit research and education think tank, appreciates the opportunity to submit comments on the European Commission's proposed Regulation laying down harmonized rules on artificial intelligence ("AI Regulation"). The IAF's mission is collaborative scholarship and education on the policies and processes necessary to use data responsibly in an observational age, while enabling a trusted digital ecosystem that serves people. IAF scholarship is based on the following three core beliefs:

- It is critical that organizations are able to think with data and engage in knowledge discovery and creation in order to enable and achieve the benefits of a global digital ecosystem.
- To be trusted, organizations must be accountable, responsible, and answerable and be prepared to demonstrate their accountability.
- Frameworks, based on risk assessments and effective and risk appropriate data governance, enable beneficial, data-driven innovation while protecting individuals and society from the potential harms that may arise from data processing in the digital age.

These comments reflect the views of IAF staff and do not necessarily reflect the views of the IAF board of trustees, contributors, or broader community.

## Background

The AI Regulation touches on all three of the IAF's core beliefs and is explicit in terms of its goal to enable economic prosperity. The Commission's explanatory memorandum makes it clear why the AI Regulation is important:

> *By improving predictions, optimizing operations and resources allocation, and personalizing services delivery, the use of artificial intelligence can support socially and environmentally beneficial outcomes and provide key competitive advantages to the companies and the European economy.*

The IAF believes the Commission could have gone further in that statement. Beneficial outcomes are not just in the interest of companies and the economy. They serve people, both individually and as a group, and in doing so, if governed correctly, they positively impact the full

range of fundamental rights and interests. Employment, health, education, and the ability for smaller businesses to find a competitive advantage all are impacted by trusted digital innovation.

The Commission's statement began with "improving predictions." Better predictions are the product of probability tools that rest on quality data well curated and responsibly stewarded. The nomenclature for this process is "thinking with data." Artificial Intelligence ("AI") is the substantial next step in organisations thinking with data. Thinking with data results in outputs or "insights," and the nomenclature for that process is "knowledge discovery." Knowledge discovery is where AI begins. Knowledge discovery is the source for better predictions. Questions that should be asked are: What are the key correlations coming from the data? Are those correlations sound? Could and should those insights be applied in a legal, fair, and just manner?

New knowledge, once created through knowledge discovery, creates the pathway to the desired future of better operations, resource allocation, and environmentally beneficial outcomes. The nomenclature for this process is "acting with data." To be consistent with the full range of fundamental rights, these processes must be applied in a fashion that creates real value for stakeholders in a manner that is respectful of other stakeholders and that mitigates identifiable harms through a sound risk management process.

The AI Regulation's clear objective is rules that lead to trustworthy organisations that are able to demonstrate that all processes related to knowledge discovery and acting with data are conducted in a responsible and answerable manner.

This application of learning may take place either directly by people or through some degree of automated decision making**.** This separation of data processing into thinking with data (knowledge creation) and acting with data (knowledge application) are part of a two phased approach that was described in "Big Data and Analytics: Seeking Foundations for Effective Privacy Guidance." This concept of a two-step approach is built into the AI Regulation in the delineation of providers of AI systems (thinking with data) and users (acting with data). AI development involves experimenting or exploring with data for the potential of identifying signals or trends which lead to the opportunity to create a model whose outcome can drive decision making. There is risk to people at both phases, but they may be very different, and arguably the risks relative to the impact to people are significantly less at the knowledge creation stage. However, the identification of risk requires robust assessments of the right types at the right time.

The differentiation of assessing and mitigating risk should be part of future looking public policy. EU Parliament Member Axel Voss identifies the necessity for risk differentiation in his paper "Fixing the GDPR: Towards Version 2.0," when he states:

> *The GDPR does not differentiate enough between low-risk and high-risk applications, determining – with a few exceptions such as prior consultation of the DPA for high-risk*

> *applications – largely the same obligations for each type of data processing. The possibility in the GDPR to define different risk classes of data processing, which require different legal bases, is not being used.*

The regulation of data in AI model development and deployment in a balanced fashion requires the process to be free of legacy prejudices, and Voss recognizes a prejudice against any processing in the General Data Protection Regulation ("GDPR").

> *It (GDPR) also wants to establish the view that processing of personal data is generally regarded as a socially undesirable behaviour. This approach is not only latently hostile to progress. The result is that even the processing of personal data that is protected by fundamental rights or that is socially desirable for the protection of public interest comes under constant pressure to justify itself.*

He reminds the reader that the GDPR requires balancing against the full range of fundamental rights and does not focus just on privacy. Specifically, the use of data to arrive at accurate insights is not per se an infringement on dignity. Infringements may come from actions taken based on the processing, but the inclusion of data should make no judgment on the persons to whom the data may pertain. It is more likely that infringements on dignity could come from inaccuracies resulting from insufficient training data. The IAF agrees with Voss that the GDPR is an important law and that privacy must be protected. However, the GDPR should make possible technology applications such as AI, and in the IAF's view, there are unresolved tensions between the AI Regulation and the GDPR. The IAF in these comments will suggest ways in which the GDPR in general might be improved so that data serves people and so that the risk of harm is more appropriately addressed.

**Overview of the IAF's Comments**

In the IAF's view, the AI Regulation could be improved by addressing two key areas.

- **The AI Regulation glosses over the tension with the GDPR.** When AI involves personal data, the GDPR challenges the use of this data in AI. Tensions and conflict with the GDPR need to be resolved. The GDPR strictly interpreted makes AI in any of the high-risk cases challenging to use. This result is contrary to the stated goals of the AI Regulation.
- **The construct of a risk-based approach in the AI Regulation is unevenly developed and applied.** First, by limiting the scope of the AI Regulation to select high risk areas, the AI Regulation infers that many of the good risk-based data governance requirements included in the AI Regulation will be ignored in other AI scenarios that could create similar risk to individuals and society. Second, the detailed approach to the use of "conformity" assessments is limiting and does not reflect the growing body of literature and indeed public policy that promotes a broader approach to AI risk management through AI Impact Assessments (AIAs).

The IAF's comments are in two parts:

- The first section addresses the interface between the objectives of the AI Regulation and the GDPR. AI conducted with personal data is first covered by the GDPR, and if prohibited or challenged by the GDPR will be difficult to use in an AI application. This result is at odds with the Commission's objective and could be resolved through some modest revisions to the GDPR. Since Article 10(5) of the AI Regulation defers to the GDPR "for the purposes of ensuring bias monitoring, detection and correction in relation to the high-risk AI systems," deference to numerous other GDPR articles will be necessary. While it is possible to deal with the GDPR tension though modifications to the AI Regulation, for clarity, the IAF thinks the proposed changes to the GDPR will have benefits beyond AI.
- The second section covers the risk management components of the AI Regulation and whether they effectively address risk to people from processing or making the choice not to process data. The IAF thinks a more effective approach to risk management could be achieved, and that approach has been outlined in the IAF's model legislation, The Fair and Open Use Act.[1] The GDPR is important legislation, and the IAF's legitimate uses are reflective of the GDPR legal bases with additions for knowledge creation and scientific research.

## IAF's Comments

### I. Proposed GDPR Changes

#### A. Interface and Tension With the GDPR

The AI Regulation's explanatory memorandum states that it is intended to be consistent with the GDPR. Further, the AI Regulation is not intended to amend the GDPR. The Joint Opinion of the European Data Protection Board ("EDPB") and European Data Protection Supervisor ("EDPS")[2] requested specificity that the processing of personal data, as part of AI, is regulated by the GDPR (and other privacy and data protection laws under the jurisdiction of the data protection authorities). Furthermore, the EDPB and EDPS requested clarity that automated decision making using personal data be under their jurisdiction as well. While the authorities are requesting clarity on that point, there is nothing in the AI Regulation that suggests the AI Regulation supersedes the GDPR.

Knowledge discovery with personal data meets the current GDPR definition of profiling, and at least part of the AI process is knowledge discovery. As some AI solutions are likely to make automated decisions pertaining to people based on processing that uses personal data that

[1] The IAF developed a model legislative approach primarily as a communications tool as it is easier to engage with policy makers in the U.S. in "legislative" language and format. While the model legislation is drafted with a U.S. policy maker audience in mind, it is illustrative and parts and/or themes could be adopted in any other global legislative format

[2] EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and the Council laying down harmonized rules on artificial intelligence.

goes through a model developed using AI would meet the definition of fully automated decision making. The Article 29 Data Protection Working Party Guidelines on Automated individual decision-making and Profiling ("Article 29 Guidelines") for the purposes of the GDPR[3] creates significant impediments to AI processing. The Article 29 Guidelines do not adequately differentiate the difference between profiling as knowledge discovery and profiling and the use of automated decision making at the knowledge applications stage.

There are a number of specific areas of the GDPR that are problematic for AI.

Article 5 requires that processing be lawful and fair, and that the data use is consistent with the purpose for which it was collected. AI training data can come from many different sources so assuring the data was collected for a purpose consistent with AI training is difficult and problematic. The implication that a model developed and trained with less data will potentially have less utility is well founded. In addition, and more significant, a model trained on less data may have inherent data biases.

Article 6 requires that data only be processed when there is a lawful legal basis. The legal basis that most often would fit is legitimate interest. However, legitimate interest requires a balancing test, and the balancing test only takes into consideration the legitimate interest of the controller and the interest of the data subject. Training data typically is not about the interests of a single individual but rather the interests of a broader group impacted by the insights that come from the AI processing. While the balancing concept could be interpreted more broadly in the AI context, at this point there is no authority for doing so.[4] Furthermore, pursuant to Article 21 of the GDPR, if legitimate interest is the legal basis for processing, individuals may object to the processing, and the processing may not proceed unless the controller demonstrates "compelling legitimate grounds for the processing which override the interests, rights and freedoms of the individual." The ability of individuals to withdraw their data may impact the accuracy of insights and introduce actual bias when the processing at this stage has little or no impact on a specific individual.

Articles 12, 13 and 14 specify the right to be informed, whether the data originated with the controller or from a third party. As described in the AI Regulation, the controller when developing an AI model (knowledge creation) and when the model is used (knowledge application) may be very different. The IAF agrees that transparency is critical but believes the GDPR as applied today would make AI with personal data administratively difficult. In addition, terms such as "transparency" and "accuracy" have much narrower definitions and application in the GDPR than what is required in AI.[5] Furthermore, Articles 13 and 14 require the following be disclosed if automated decision-making, including profiling, is being used:

- The reason automated decision-making and/or profiling is being used

[3] https://ec.europa.eu/newsroom/article29/items/612053

[4] The IAF led a multi-stakeholder project in 2017 that resulted in a more inclusive balancing tool. It may be found at Legitimate-Interests-and-Integrated-Risk-and-Benefits-Assessment.pdf (secureservercdn.net)

[5] See Transparency Needs a Makeover - The Information Accountability Foundation

- Information about the logic used
- All significant effects and anticipated consequences
- What the logic used takes into consideration

As almost all AI is profiling, these requirements particularly are challenging at the model development stage at a time when the impact to an individual is minor. This may limit artificially the use of AI.

Article 15 covers the individual's right of access, but the question of access to what is uncertain in the AI development process. The provisions of the AI Regulation provide some answers, but as stated earlier, the GDPR comes into effect before the jurisdiction of the AI Regulation.

Article 22, as mentioned earlier, says that "The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her." Again, this language may preclude AI processing before the jurisdiction of the AI regulation even comes into effect.

Article 35 requires Data Processing Impact Assessments (DPIAs) when data processing poses a "high risk" to individuals. Unlike the language in Article 6 regarding legitimate Interests balancing, Article 35 provides a more nuanced approach, specifically, "assessment of the risks to the rights and freedoms of data subjects." However, there is no distinction between an assessment required for high-risk processing and the distinction between appropriate risk controls at the knowledge discovery stage and the application of insights that may have been derived at this stage.

Finally, Article 9 of the GDPR covers the processing of special categories of personal data. This provision of the GDPR has made the processing of data for scientific research purposes as well as for more general knowledge discovery quite difficult. While Article 5 of the GDPR says that scientific research is a compatible purpose, it has generally been understood that use of data, once collected under a different purpose, requires consent for scientific research if the data meets the special category test. Much of the data that meets the AI Regulation's objective of better predictions falls in this category because of the high-risk scenarios that form the core of the AI Regulation.

### B. Addressing the GDPR Issues

Creating a carveout from the GDPR for personal data used in AI knowledge development and application in the AI Regulation would be possible. However, doing so would leave knowledge creation that falls short of AI in a strange position. Instead, it makes sense to amend the GDPR so it better reflects a true balancing of the fundamental rights associated with employment, health, free movement, and conduct of a business in a digital age. The concept of amending the GDPR has already been raised by European Parliament member Axel Voss in the white paper Fixing the GDPR: Towards Version 2.0.

The IAF respectfully makes the following suggestions:

- Knowledge discovery should join scientific research as being recognized as a compatible purpose under Article 5. Doing so would recognize that research uses of data, whether they meet the scientific test or not, raises fewer risks than the actual application of data to take an action that impacts individuals.
- Scientific research and knowledge discovery should be added to Article 6 as legal bases G and H. Saying a use of data is compatible with the purpose for which it was collected is not adequate. All uses require legal standing. The inclusion of data in a research project does not infringe on the fundamental right to dignity. However, the impacts to people of research not done can impact the human status in a manner that does impact dignity. To protect these new legal bases from misuse, appropriate conditions, particularly transparency related to the objectives and safeguards for processing, should be added as well.
- A pathway for using special categories of data in knowledge discovery and scientific research should be created in Article 9.
- Article 35 should be amended to more clearly apply to the application of insights stage (use of data), where there is a much greater risk to individuals, and not at the knowledge discovery stage.
- Article 22 should be revised to describe more clearly profiling and automated decision making. Profiling is central to the knowledge discovery process. As such, it should be subject to a DPIA to ascertain that processing is conducted in a legal and fair manner. Automated decision-making is a separate process. It too should be subject to assessments, and the risks from a particular automated decision-making process should be understood and documented. Currently, the GDPR confuses the connections between the two.
- Lastly, Recital 4, which is (should be) the heart and soul of the GDPR, states:

  *The processing of personal data should be designed to serve mankind. The right to data protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights . . . .*

  The intent of this Recital has not been adopted in the application of the GDPR in areas such as balancing, as required, under a legitimate interest assessment. In the IAF's view, this is a byproduct of failing to explicitly carry forward the Recital's intent into a specific Article. This failure could be accomplished through revisions to Article 5 of the GDPR.

## Risk Management

The IAF believes the risk-based approach in the AI Regulation has been developed and applied unevenly in several ways. This uneven approach begins with the narrowing of risk management tools to conformity assessments and only applies them to a small subset of high-risk applications. By extension, this narrowing precludes the identification of the broader potential adverse impacts to individuals from AI. While the proposed regulation suggests risk assessments beyond conformity assessments, it is very vague about the types of risks to be identified and how one should characterize those risks.

### Product Safety

There is a heavy focus on product safety that drives many of the risk management requirements in the AI Regulation and places less importance on the reality that AI is more about the implications of the data collected and produced and less on the product. This approach ignores or under invests in the data driven implication to individuals. For example, an insulin pump is just a pump. The product safety rules govern how that pump works. However, the data from and produced through the actual operation of such a medical device could be used to provide beneficial insights though AI. These insights may trigger new insights that improve future applications or the application of data driven insights in other health related uses. The product safety approach looks at the pump as a pump and not as a data generating device. The AI contribution to society begins with the data and how that data might improve this and other medical devices in the future. For example, an algorithm, which uses AI to help manage type 1 diabetes, was designed entirely using a mathematical simulator, and when the algorithm was validated on real world data from people with type 1 diabetes, it generated recommendations that were highly similar to recommendations from endocrinologists.[6] However, a drug pump, that collects and processes data to achieve patient safety, is a pump and not primarily an inferred data production device.

Similarly, the applications where decisions on people are being made or influenced even if they use de-identified data for the data science or begin with personal data are under-valued in the AI Regulation due to the AI Regulation's emphasis on the requirements related to "conformity" assessments. A conformity type assessment focused on product safety does not adequately address the data driven benefits and risks. At best, the risk assessment process to identify risks through the processing of data is inferred in the AI Regulation.

The description of conformity assessments is not consistent or complete enough with the trend toward broader AIA type assessments. By very definition, "conformity" means applying to an

[6] New algorithm uses artificial intelligence to help manage type 1 diabetes, ScienceDaily, June 12, 2020, https://www.sciencedaily.com/releases/2020/06/200612172204.htm

established standard. Yet, while much literature has been developed on AI risks, there is no consensus on risk mitigation processes, and certainly, there is no established standard. At a very basic level, the use of the term "conformity" is a clear carryover to a focus on product safety and perhaps an unintentional sub-focus on the identification of broader risks. An extension or by product of this approach is that it is unclear that the "risk of what" is established for conformity assessments.

While the AI Regulation infers a broader process to identify risks, the focus on conformity assessments moves against the trend being established relative to AIAs. In both academic work, other public sector developments and proposed legislation in other jurisdictions, the use of AIAs is suggested or proposed. For example, in Australia, Canada, Singapore and shortly Hong Kong, governments have or will be implementing processes and requirements for their public sectors to address ethical and other risk challenges of AI. These guidelines often include AIAs. Even before laws and regulations will drive the need for expansive impact assessments, public sector procurement rules will have an earlier impact.

In the short term, even regulatory guidance issued by data protection authorities and others will have an impact on organizational processes. An example of this is AI guidance issued by the UK ICO,[7] which suggest a very different sort of internal assessment than even found in complex DPIA's.

Globally, there already are many AI specific proposed bills and other governmental activity, and much governmental debate and investigation as to what should govern the impact (risks) of AI. They universally contemplate some form of broader impact assessment. Several proposed laws in the U.S. introduce specific requirements related to algorithmic assessments or risk assessments related to "automated decisions" which is becoming the proxy integration medium between AI and privacy. The growth in the use of AI likely will accelerate this trend. In terms of specific assessment type requirements, specific proposed laws include the Algorithmic Accountability Act (U.S. House of Representatives) and Algorithmic Fairness Act (U.S. Senate). What is also being seen are proposed privacy laws that either explicitly or implicitly require much broader type risk assessments than a privacy impact assessment ("PIA"). See, for example, the proposed Mind Your Own Business Act (U.S. Senate), Consumer Online Privacy Rights Act (U.S. Senate), and the Automated Decision Systems Accountability Act of 2020 (U.S. - California).

In short, the formal requirement relative to conformity assessments in the AI Regulation is not aligned with the needs for a trustworthy AI ecosystem. The AI Regulation should require AIAs for all AI applications. AIAs then could encompass conformity assessments, and while they

[7] Guidance on AI and data protection | ICO, AI Auditing Framework | ICO, Explaining decisions made with AI | ICO

could be separate from GDPR assessments, they could be incorporated where AI involves personal data. The area of AIA's has been explored in detail in the IAF's report on The-Road-to-Expansive-Impact-Assessments.pdf (secureservercdn.net)

The IAF believes that risk assessments, based on the concept of adverse outcomes, should be broadly applied and that oversight and potentially ex ante reviews also should be risk based. This approach ties to a related recommendation. As the AI Regulation is narrow in scope and only applies to named high risk scenarios, the implication is that all of the good practice requirements listed in the AI Regulation could be ignored by another AI application that was not initially high risk, as defined, but still could have significant impact and risk to an individual. While the AI Regulation may have been designed to take an "immediate but limited" approach, in the case of AI, such an approach is not consistent with the GDPR generally (legal or similar impact on an individual) or specifically (risky processing triggering the need for a DPIA under Article 35). An example is a broad-based AI system that results in a segmented approach to access to key services or pricing. It appears that highly impactful AI scenarios, such as the racial bias found in Airbnb's 2015 Smart Pricing algorithm or the bias in online targeted advertising targeting, would not be covered by the AI Regulation.

AIAs would address all requirements and facilitate risk and adverse outcomes reviews. They would identify higher risks (by potential impact, not by scenario) and facilitate appropriate oversight. The oversight system needs to be more grounded than one that is based on ex ante reviews by third parties that are driven by established standards in an arena where there are no AI standards.

The IAF respectfully makes the following suggestions:

The IAF's core suggestion to improve the AI Regulation is to adapt the AI Regulation so that it applies to all AI application scenarios that would be risk assessed as opposed to applying only elements of the AI Regulation to a narrow set of defined high-risk scenarios. This change would allow for the tailoring of requirements to those applications that have the potential to create higher risk. More specifically:

- A requirement to assess the likely risk of an AI application would allow for the extension of the very good data governance requirements to be extended to all AI applications. While many of these included in the AI Regulation should arguably be applied to all areas of AI development relative to data, this would allow for a tailoring of requirements based on risk.
- Broad based AIAs should be a mainstay requirement for all AI applications and should include the assessment of risk. Conformity assessment requirements as currently outlined in the AI Regulation could be incorporated into an AIA as appropriate as could aspects of a DPIA as required by Article 35 of the GDPR. The concept of broader impact assessments has been explored in the IAF's recent report The Road to Expansive Impact Assessments.

**Conclusion**

In summary, first, since the proposed AI regulation is subordinate to the GDPR when AI processing makes use of personal data, the impact of the GDPR on AI needs to be considered. However, the GDPR is not conducive to knowledge discovery, the starting point for AI. Second, the AI regulation is not aligned with emerging practices in trustworthy AI that are necessary to assess and control risk related to AI that could impact people. The IAF's comments suggest strategies both for the GDPR and the AI Regulation so that AI achieves the stated goals of the AI Regulation.

As Recital 4 to the GDPR says, "personal data should be designed to serve mankind," and any regulatory approach to AI should begin with that premise. Giovanni Buttarelli, the former European Data Protection Supervisor, repeated that view when introducing the EDPS' advisory panel on ethical data use. He expected data ethics to fill the gaps in the law that would inevitably come with advanced technologies. AI, by its very nature, is problematic for data protection law. AI done well serves mankind; AI done badly can be destructive of human values. That latter result is why effectively addressing the balance in the AI Regulation is so important.

AI begins with knowledge discovery, and knowledge discovery, when using personal data, currently is not a specific enough purpose to establish a lawful basis. Repurposing personal data requires compatibility or consent, and consent is problematic. Because scientific research is not typically the stated purpose for data creation, scientific research is defined as a compatible purpose.

AI training data, as part of knowledge discovery, raises more data protection issues. AI training almost always is a repurposing, and the quality and robustness of the training data, which can be limited by current interpretations of the GDPR, impacts the results. Faulty results then negatively impact both the benefits and risks to individuals and groups of individuals. The issue is not about circumventing fundamental rights but rather about fine tuning the law for knowledge discovery, so the full range of fundamental rights are protected. Axel Voss's white paper, "Fixing the GDPR: Towards Version 2.0," outlined some of the structural issues in the GDPR related to knowledge discovery.

As for evidence of the problem, the scientific research community has been struggling with the GDPR since it was enacted. Research is a compatible purpose, but a legal basis to process personal data still is needed. There was some thought that legitimate interests would be that legal basis, but it has not transpired that way. While research might be conducted in a pseudo anonymized fashion, the actual data matching requires processing that begins in a clear data form.

Any time data pertaining to an individual are processed, there is an infringement of personal sovereignty, and that infringement is why European law begins with a general prohibition on processing personal data unless there is a lawful basis to do so. The necessity for processing

data for the benefit of mankind is why Recital 4 of the GDPR exists. Voss argues that authorities over balance for the procedural requirements related to individual sovereignty and undervalue fundamental rights such as health, employment, and the joint benefits from technology.

The two new legal bases the IAF proposes - knowledge discovery and scientific research - will require effective accountability processes to achieve fairness. Dignity is the first fundamental right, and human dignity links to the conditions in which individuals live. Health and welfare go beyond the interests of a single data subject. There is general agreement that that lack of balance between multiple rights impacted the Covid-19 response and therefore impacted the human condition and therefore impacted human dignity. The IAF's approach and recommendation would enable the application of more appropriate and balanced risk mitigating controls at the various stages of AI development and application.

While the GDPR is not fine-tuned for AI, the AI Regulation only covers "high risk AI" and relies on conformity assessments that are dependent on established standards. Work in trustworthy AI suggests that all AI that touches on people should undergo AIAs. It is only when an AIA and associated risk analysis is conducted that a full sense of the risk associated with data processing through AI can be identified and mitigated.

The IAF appreciates the opportunity to comment. If you have questions, please contact Martin Abrams at mabrams@informationaccountability.org.