REFLECTIONS ON THE EU'S AI ACT AND HOW WE COULD MAKE IT EVEN BETTER

BY
MEERI HAATAJA

&
JOANNA J. BRYSON

CEO & Co-Founder, Saidot Ltd, Espoo, FI. Professor of Ethics and Technology, Centre for Digital Governance, Hertie School, Berlin, DE.

TechREG CHRONICLE
MARCH 2022

Visit **www.competitionpolicyinternational.com** for access to these articles and more!

Reflections on the EU's AI Act and How we Could Make it Even Better

By Meeri Haataja & Joanna J. Bryson

Jurisdictions around the world are preparing regulations for artificial intelligence, as investments in AI technologies continue to increase as a source of efficiency and innovation for companies and governments. One of the most influential regulative proposals for AI is that proposed by the European Commission in April 2021, the "AI Act." The EU's proposed regulation has already inspired some international regulative proposals and is likely to broadly impact AI policies around the world. Yet the Act is still in process, it's strengths could be compromised, or it's weaknesses addressed. In this piece, we analyze the core policy concepts of the AI Act, with focus both on those worth amending and defending. These discussions may provide valuable elements for other regions beyond the EU to consider for their own AI policy. While the AI Act could still be improved to make it even more robust in managing AI-related risks to health, safety, and fundamental rights, and to increase incentives to industry to take actions beneficial to both itself and others, overall we applaud this act.

01 INTRODUCTION

The EU's proposed regulation for artificial intelligence, published in April 2021 and known as the "AI Act,"[2] is probably the most influential AI-focused policy paper published to date. Reflecting an extensive process, and part of an impressive suite of innovative legislation aimed at addressing the challenges of digital governance, the AI Act ("AIA") contains many strong policy ideas well worth proposing, enforcing, and defending. Of course, much has already been said by researchers, policymakers, and industry representatives of various kinds. However, while reading these inputs, we feel that there is still an important gap worth filling, reflecting the expected practical impacts of the proposed AI Act on the providers and deployers[3] of AI technologies. Drawing from this practical perspective, we too do provide suggestions where the proposed regulation could still be improved. At the same time, we also critique some of the previous critiques – amplifying some and providing counterarguments to others. More generally we wish to acknowledge and encourage the positive work of others, and encourage familiarization with the referenced materials for more extensive exploration of our topics. This includes that we want to emphasize and reinforce the good parts of the initial draft of the AIA, to ensure these portions are retained intact or even strengthened through the present process of finalizing the legislation.

Let us nevertheless start by pointing to some areas of the proposal which undeniably require some further iteration. We focus only on critique which we believe has a significant influence on successful implementation, and achieving the targets of the regulation as outlined in the proposal.[4] These therefore should be addressed now, in contrast with the EC's built-in mechanism for continuous improvement of contents referred to in annexes of the proposed regulation. Our first observation is that the impressive suite of digital governance legislation[5] proposed and still to be proposed must of course be carefully monitored to ensure that nothing creates gaps or "wiggle room"; this motivates several of our comments here. While as computer scientists we of course appreciate the EC's attempt to avoid redundancy and therefore potential contradiction between the Acts, we believe the only way to prevent gaps is to add explicit points of contact between them. Explicit connections should be made between the various acts, though of course these should be loosely-coupled "universal joints," allowing maximum flexibility in the other acts, and ensuring that the acts seldom if ever need to be amended in synchrony.

With respect to the AIA itself, we now discuss eight points which, in our opinion, would benefit from some reworking.

Be explicit that all AI, and indeed all software, is a manufactured product and falls under classic product law. This would ensure that product safety, evidence of due diligence – following best practice, avoiding known bad practice, etc. applies to every level of commercially marketed AI and AI development. Something like this is frequently stated in official presentations of the law, but yet it is also often debated on panels. For example, some say the exception for medical devices shows that most AI systems are *not* devices. Note that this specification could also simplify the Digital Services Act ("DSA")[6], and perhaps should be reiterated there, and would presumably link both the DSA and the AIA to the forthcoming liability act.

Define AI in terms of its applications. The definition of AI must focus on use cases rather than specific technologies. This is a minor textual, but substantial and urgent conceptual fix, which unfortunately runs counter to some present member-nation thinking, including the presently proposed

[2] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206.

[3] We're deliberately not calling deployers "users" as the EC has. This is to avoid ambiguation between the terms referring to deployers and end-users. We strongly advise the EC, EP and everyone else to disambiguate the use of this term. The other group potentially labelled "users" we here refer to as "end users," again for clarity.

[4] AI Act, p.4: 1) ensure that AI systems placed on the Union market and used are safe and respect existing law on fundamental rights and Union values; 2) ensure legal certainty to facilitate investment and innovation in AI; 3) enhance governance and effective enforcement of existing law on fundamental rights and safety requirements applicable to AI systems; 4) facilitate the development of a single market for lawful, safe and trustworthy AI applications and prevent market fragmentation.

[5] At a minimum, this suite consists of the Digital Services Act (DSA), the Digital Markets Act (DMA), the AIA, and the still-forthcoming Liabilities Act. See further below.

[6] Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive. Available at https://eur-lex.europa.eu/legal-content/en/TXT/?qid=1608117147218&uri=COM%3A2020%3A825%3AFIN.

presidential compromise text. The appendix (Annex I)[7] needs to be labelled as indicative, not complete, with all systems producing similar outcomes to the listed technology through automated means being equally covered. The last thing any legislator should want to do is to motivate the use of obscure or novel technology when well-established and transparent techniques are available.[8] We should motivate convergence on technology that easily complies with regulatory requirements.

Clear alignment with the GDPR[9] is a hygiene factor. The AIA applies equally to all systems falling into its scope, whether or not they handle personal data. The EC has specifically avoided overlaps with GDPR and consequently hardly even mentions data protection in the proposal's requirements. We agree with EDPB and EDPS[10] on a need to clarify this relationship and support e.g. the addition of a requirement for compliance with the GDPR in the requirements for high-risk systems (Chapter 2). We believe this is essential also to the establishment of AIA-related processes in provider and deployer organizations as complementary to data protection processes, such as data protection impact assessment ("DPIA"), to encourage governance efficiency.

Lack of public sector enforcement is an elephant in the room. The potential loophole for Member States to leave public authorities without administrative fines is simply unacceptable.[11] Considering that a substantial share of the prohibited and high-risk cases are public sector uses, leaving out enforcement mechanisms from public authorities would undermine the credibility of the proposal in securing both fundamental rights and democracy.[12] This would also give private organizations, who are working as AI providers to public sector organizations, an unfavorable or even unfair position. The regulatory risk in terms of penalties would fall to private sector providers. Yet at the same time, risk of incidents would increase, because public sector clients may not be properly incentivized to comply with the deployer obligations, such as human oversight. Following the same reasons, we call for independence of the market surveillance authorities in Member States.

The EC should make up its mind about the prohibited use cases. As commented by many, considerations on the prohibited use cases appear to us as a compromise which, via the specific exceptional conditions, creates loopholes that allow continued utilization of remote biometric identification in public spaces for law enforcement as usual.[13] For example, kidnapping is unfortunately literally an every-day occurrence, largely driven by custody battles; wanted criminals are similar. Terrorism events may be less common, but only if strictly demarcated as brief, temporary emergencies of extreme violence or danger, as the act indeed presently specifies. Note that as the U.S. demonstrated in 2021, even leading democracies experience the creation of apparent terrorist 'emergencies' around benign political events such as transfer of power. The EC should decide whether or not it is really ready to prohibit such use cases, or whether rather they prefer to carefully regulate them[14] and act accordingly. This aspect has been thoroughly discussed e.g. by the EDPB and EDPS joint position paper. We also acknowledge the challenges of interpreting the scope of prohibition for sublimi-

[7] Annexes to the AI Act. Available at https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_2&format=PDF.

[8] Bryson, Joanna J., Mihailis E. Diamantis & Thomas D. Grant. "Of, for, and by the people: the legal lacuna of synthetic persons." Artificial Intelligence and Law 25, no. 3 (2017): 273-291.

[9] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[10] EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).

[11] AI Act, Article 71 (7-8).

[12] cf "Draft AI Act: EU needs to live up to its own ambitions in terms of governance and enforcement (Submission to the European Commission's Consultation on a Draft Artificial Intelligence Act)" Algorithm Watch, forthcoming.

[13] AI Act, Article 5 (1d, 2-4).

[14] Robbins, Scott. "Facial Recognition for Counter-Terrorism: Neither a Ban Nor a Free-for-All." In Counter-Terrorism, Ethics and Technology, pp. 89-104. Springer, Cham, 2021.

nal techniques,[15] further discussed e.g. by Veale & Borgesius.[16]

The EC should make up its mind about the prohibited use cases

Realistic data governance requirements. Another key requirement, high quality datasets "free of bias," is like the "exceptional" status of the prohibited use cases, completely implausible. Again, in presentations the EC often says they know that even "complete" data must reflect the biases of our imperfect world, yet setting an impossible bar for high-risk AI, like ubiquitous "exceptional" circumstances for prohibition, invites facetious lawsuits and (perhaps worse) ridicule. These problems are serious enough that we would recommend releasing revised text as soon as possible on these two matters. Here we would prefer to see instead indications of the need for documenting due diligence, best practice, and requirements for proportionate effort.

Sandboxes are fine but not enough for SMEs. If you are a startup developing AI for law, public safety, health, or environment – good for you. The intended regulatory sandbox can actually be useful for you by enabling repurposing of personal data within the sandbox to enable the development of public interest AI.[17] For any other SMEs the added value seems low. What really is critical is that the EC clarifies how proportionality works for a startup whose impact grows from 4 to 40M individuals while the intended purpose remains the same. We think this consideration is 'there' in the act, but not yet made clear enough.[18] SMEs will also likely benefit more from access to technological compliance tools than ad hoc consultative support by member state authorities.

Stakeholder engagement remains in the ethics space. Stakeholder participation has become one of the important means for ensuring ethical governance of AI systems. For example, the EU AI HLEG final paper recommends stakeholder participation under its guidance for how to manage diversity, non-discrimination and fairness of AI systems.[19] Maybe surprisingly, the proposed AI regulation ignores this, or at least, leaves it for providers and deployers to consider whether or not such engagement would be meaningful. Based on the EC's proposal, high-risk systems may well be developed also in the future without representation of impacted people. The EC may want to review whether there is enough stakeholder participation of affected communities in the key governance structures of the proposal, e.g. through creation of harmonized standards. Again, this would need to be proportionate, and can be expected to sometimes require significant expansion of effort if a startup finds itself unexpectedly successful and growing rapidly. Resources should be available to help companies deal with such success appropriately.

For the sake of readers' time, we refrain from going into further details that other critics have discussed in detail in position papers referred to throughout this document. For convenience, table 1 summarizes the discussed key critiques along with our next focus: policy concepts and ideas already in the AIA which we believe are fundamentally important for the success of this new legislation, and thus worth defending.

[15] AI Act, Article 5 (1a)

[16] Michael Veale & Frederik Zuiderveen Borgesius, 2021: Demystifying the Draft EU Artificial Intelligence Act, p.7-9. Available at https://arxiv.org/abs/2107.03721.

[17] AI Act, Article 54.

[18] AI Act, Articles 8-9.

[19] Independent High-Level Expert Group on Artificial Intelligence Set Up by the European Commission, 2019: Ethics Guidelines for Trustworthy Artificial Intelligence, available at https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

Table 1: Summary table on the key issues raised

Main element	Key contents	Criticisms	Ideas to defend
Definition of AI	Techniques and mechanisms	I) AI as a manufactured product II) Define in terms of outcomes not processes	i) Breadth of application
Framework on AI risk levels	Unacceptable risk High risk Limited risk Minimal or no risk	V) Scope of prohibited uses	ii) Framework for AI risk levels
Requirements for high-risk systems	Five key requirements Obligations for providers and deployers Notifying authorities and notified bodies Standards, conformity assessment, certificates, registration Post-market monitoring, information sharing, market surveillance Governance	III) Lack of alignment with GDPR and other existing regulations VI) Implausible data governance requirements VII) Missing stakeholder engagement requirements	iii) Proportionality of requirements (though should be refined) iv) Accountability of AI supply chain v) Meaningful documentation requirements
Other	Transparency obligations for certain AI systems Measures in support of innovation Codes of conduct Confidentiality and penalties	IV) Public sector administrative fines VIII) Support for SMEs	vi) Contextual transparency reporting to AI end users vii) EU Database for high-risk systems

Before looking into what is particularly good in the proposal, let us first summarize some of its key aspects, creating a helpful context for our more detailed analysis.[20]

The AIA regulative proposal was announced as part of a broader package, A European Approach to Excellence in AI, targeted to strengthen and foster Europe's potential to compete globally. Therefore, while our focus here is on the proposal itself, it is useful to understand the larger context and the accompanying coordinated plan on AI (2021 review) which details the strategy for fighting for Europe's competitiveness in AI. "Through the Digital Europe and Horizon Europe programmes, the Commission plans to invest €1 billion per year in AI. It will mobilize additional investments from the private sector and the Member States in order to reach an annual investment volume of €20 billion over the course of this decade. And, the newly adopted Recovery and Resilience Facility makes €134 billion available for digital. This will be a game-changer, allowing Europe to amplify its ambitions and become a global leader in developing cutting-edge, trustworthy AI."[21] This corresponds to roughly €65 billion investment volume annually by 2025.[22]

The AIA is part of a continuum of actions that started in 2017 with the European Parliament's Resolution on Civil Law Rules on Robotics and AI[23] and entailed several other key milestones[24] prior to the proposal at hand. It is addressed to AI use cases that pose a high risk to people's health, safety, or fundamental rights. The regulations would apply to all providers and deployers placing on the market or putting into service high-risk AI systems in the European Union, regardless of the origin of the providing entity. In this

[20] Some content from this section has been included in abridged and altered format in Dempsey, M., McBride, K., Haataja, M., & Bryson, J. J. "Transnational digital governance and its impact on artificial intelligence," *The Oxford Handbook of AI Governance*, Oxford University Press, expected 2022.

[21] European Commission, A European approach to Artificial intelligence, available at https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence.

[22] European Commission, Impact assessment accompanying the AI Act, p.70.

[23] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)).

[24] E.g. A report "Ethics Guidelines for Trustworthy Artificial Intelligence" by EU AI HLEG and European Parliament resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies (2020/2012(INL)).

way, the proposal seeks to level the playing field for EU and non-EU players and has mechanisms to influence far beyond its immediate scope ("regulatory export").[25]

We now turn to discuss concepts of the AIA which, based on our examination to date, are solid and actionable concepts forming the core of the regulative proposal. These concepts may well also be the most important elements for other regions beyond the EU to consider for their own AI policy.

Clear and actionable framework for AI risk levels. The proposal suggests a risk-based approach with different rules tailored to four levels of risk: unacceptable, high, limited, and minimal risk. At the highest level of risk, the unacceptable systems are systems that conflict with European values and are thus prohibited. Such a ban is a victory to all digital human rights advocates and delivers a strong message: First, do no harm. In the next level, the high-risk systems cover a variety of applications where foreseeable risks to health, safety, or fundamental rights demand specific care and scrutiny. According to the EC's impact assessment, roughly 5-15 percent of all AI systems would fall into this high-risk category.[26] Limited-risk systems are those that interact with natural persons and therefore require specific transparency measures to maintain continued human agency and to avoid deceptive uses. All other AI systems – the great majority – belong to the minimal risk category for which the AIA introduces no new rules.

> *We now turn to discuss concepts of the AIA which, based on our examination to date, are solid and actionable concepts forming the core of the regulative proposal*

We find this model both simple and actionable. The EC's list of high-risk use cases cover domains from product safety components to biometric identification, management of critical infrastructure, education, employment and workers' management, essential private and public services, law enforcement, and migration to justice and democratic processes. The list is a synthesis of EC's screening of a large pool of high-risk use cases suggested in reports by European Parliament, ISO, AI Watch, AI HLEG as well as public and targeted stakeholder consultations. It would be hard to challenge this list. Having discussed with organizations deploying AI in these high-risk domains, and based on our experience, such organizations rarely challenge these categorizations either.

Worth noting is the way the detailed list of high-risk systems is provided in the Annexes (II-III). There's a reason for this, other than the convenience of reading. By adding the definitions of all key concepts in the annexes, the EC has secured a smooth mechanism for updating such key concepts that may evolve as the industry, research, and standards around AI mature, by the delegated acts.[27]

Proportionality. An aspect largely neglected by previous critics is the principle of proportionality. By proportionality, we mean an attempt to have the requirements rightly sized in relation to the potential risks, and regulate only what is necessary. We believe proportionality is fundamentally important especially in such a domain, where both technology, as well as use cases, are under fast-paced development and the current exposure to the risks and impacts in many domains is still limited. The EC has elsewhere done a good job in introducing several vehicles while seeking to minimize the added regulatory burden and minimize the costs of compliance, for example in the DSA.[28]

In the AIA, the EC presently claims to address proportionality primarily via the previously-discussed risk-based approach and varied requirements depending on the system risk level. The majority of AI systems in the market would face only transparency requirements as mandatory if any. Unfortunately, all standards–including regulatory levels–are subject to regulatory capture and may be used as barriers to market entry. *We would like to ensure that proportionality goes beyond the strict levels and into finer-grained concerns*. More generally, we advise proportionality with respect to standards. For example, we recommend specifying that compliance with certification should be taken as evidence of due diligence rather than be mandated. We

[25] Peukert, Christian, Stefan Bechtold, Michail Batikas & Tobias Kretschmer, Regulatory export and spillovers: How GDPR affects global markets for data, https://voxeu.org/article/how-gdpr-affects-global-markets-data September 30, 2020.

[26] AIA Impact Assessment, p. 71.

[27] "Delegated acts are non-legislative acts adopted by the European Commission to amend or supplement legislation. Delegated acts are used, for example, when acts have to be adapted to take account of technical and scientific progress."

[28] This care is widely seen as addressing one error in the GDPR, which was that the non-differentiated costs were more excluding for smaller businesses.

would also prefer to see proportionate transparency requirements deployed for all software systems, regardless of the use of techniques presently labelled as AI. Proportional transparency and liability assurance could largely be self-assessed as is suggested in the DSA. The existing AIA levels could then be used to dictate lower bounds e.g. on the extent of transparency by application area, though these still should perhaps be ameliorated by the scale of the system's impact. But where companies self assess potential risks of impacts, they could engage with a proportionate amount of the requirements specified for products in the next-higher level of risk. Should they indeed come to be recategorized as higher risk perhaps after a public incident, this pre-work could be used to show due diligence and to minimize any liability.

Further on in the present AIA, proportionality is also addressed via the use of harmonized standards, the alignment with the New Legislative Framework, and by allowing the conformity assessment based on self-assessment for the vast majority of all high-risk systems. Considering the breadth of the requirements for these standards, even with the existing language, a high variation of interpretations can be expected. The use of harmonized standards is presumed to place providers in conformity with the requirements the standards cover. In addition, systems that would otherwise require third-party conformity assessment can follow a self-assessment process. Considering the factors summarized in table 2, we believe this approach has all the ingredients to improve both governance quality and efficiency.

Table 2: Standardization as a means for governance quality and efficacy. Though see also notes on proportionality, above, including concerns regarding regulatory capture.

- Typically wide representation in the standardization process from industry, researchers, NGOs etc., including persons from varying disciplines.
- The response to AI Act's requirements will likely come from several standards, allowing a wide range of expert contributions in the process (compared to an individual provider's AI team size and expertise profiles).
- Standards development follows an established and well-documented methodology including critical assessment before being approved.
- For safeguarding against gaps or needed additional expert contribution on safety or fundamental rights, EC has laid down a system of Common Specifications (Art 41) as follows:

  > "Where harmonised standards referred to in Article 40 do not exist or where the Commission considers that the relevant harmonised standards are insufficient or that there is a need to address specific safety or fundamental right concerns, the Commission may, by means of implementing acts, adopt common specifications in respect of the requirements set out in Chapter 2 of this Title."

- Market surveillance mechanisms will feed in surveillance data of all types of systems in the market. This data should reveal if it would appear that systems that have gone through the standards path are not actually in conformity with Chapter 2.

Accountability of AI supply chain, i.e. providers and deployers, not the end-users. Another less discussed but incredibly important characteristic of the proposal is how it creates grounds for significant improvements in the supply chain transparency and accountability. Let us be clear: no end user can take full responsibility for evaluating the trustworthiness of complex technology products such as AI products. In order to do so, one would need a good level of transparency to the workings of the system and the technical skills necessary for meaningful evaluation. From this perspective, we want to acknowledge the EC's choice to focus on the accountability of providers, developers, and deployers, even if it may have led to some compromises on the end-user transparency obligations. This provider-deployer dualism is also important taking into consideration that 60 percent of organizations report "Purchased software or systems ready for use" as their sourcing strategy for AI.[29]

The AIA does not suggest mechanisms that allow individual persons to submit complaints about their concerns and harm caused by AI. This has raised concerns by some. However, the choice seems logical considering that proper evaluation of system conformity would require much more information and technical evaluation skills than what will be available to end users.

The solution the AI Act proposes is the following: Providers are required to set up a post-market monitoring system for actively and systematically collecting, documenting, and

[29] European Commission, Ipsos Survey, European enterprise survey on the use of technologies based on artificial intelligence, 2020, p.53.

analyzing data provided by deployers or collected otherwise on the performance of high-risk AI systems on the market. Deployers of such systems are obliged to monitor and report potential situations presenting risks. To support this mechanism's function, it would be sensible (and seems likely) that providers and deployers implement feedback channels or contact points also for the end users. This solution should probably though be encouraged in revisions to the AIA. In addition, similar feedback channels may be expected from national market surveillance authorities to support their role in identifying potential incidents outlined in Article 65.

We believe this intended mechanism, together with the EC's planned civil liability regime for AI,[30] rightly allocates the monitoring responsibility to providers, deployers, and market surveillance authorities, and incentivizes these to opening feedback channels without direct enforcement. Nevertheless, making the expected channels for end-user feedback more explicit might ensure faster convergence to best practice, as well as defraying some present criticism.

Meaningful documentation requirements aligned with engineering best practices. The documentation requirements should be evaluated on the basis of whether they are capable of revealing whether an AI system aligns with the requirements set out in Chapter 2. These requirements are: Risk management system; Data and data governance; Technical documentation; Record-keeping; and Transparency and provision of information to deployers.

Based on our analysis, the requirements are detailed enough to enable proper conformity assessment as well as proper oversight of systems with AI, and align reasonably well with the transparency research and best practices. We provide an overview of the documentation requirements in table 3 as the adoption of these documentation guidelines is the first practical step in adopting AIA as a code of conduct. Every company, even the smallest SME can help with regulation just by demonstrating understanding of the requirements for transparency and compliance. Again, mandated levels of compliance with these requirements should be suitably proportionate.[31] It should be clear that for lower-risk, small applications a much more abstracted and limited level of documentation is allowable. With these practices in place, the malfeasant can no longer claim either that documentation is impossible, or that "AI is necessarily opaque,"[32] nor that they didn't understand the regulations. We need to build up a culture demonstrating that good practice in documentation is easily knowable, and that ignorance is negligence.

[30] EU rules to address liability issues related to new technologies, including AI systems (last quarter 2021-first quarter 2022), source: A European approach to Artificial intelligence, available at https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence.

[31] Article 8 (2): "The intended purpose of the high-risk AI system and the risk management system referred to in Article 9 shall be taken into account when ensuring compliance with those requirements."

[32] Bryson, J.J. & Theodorou, A., 2019. How society can maintain human-centric artificial intelligence. In Human-centered digitalization and services (pp. 305-323). Springer, Singapore.

Table 3: Technical documentation requirements as outlined in the AIA for systems of at least high-risk level. (Presumably, if "unacceptable" risk systems continue to be permitted in exceptional circumstances, these too will require transparency.)

General description of the system	Intended purpose Accountable persons Version of the system Hardware and software infrastructure Photographs or illustrations Instructions of use (see table 4)
Elements of the system and its development process	Development methods, incl. use of third-party technologies Key design choices and assumptions System architecture Use of computing Datasheets for datasets Human oversight Changes and change management Validation and testing procedures, incl. accuracy, robustness, cybersecurity, bias
Monitoring, functioning, and control	Capabilities and limitations in performance Expected level of accuracy Foreseeable sources of risks to health and safety, fundamental rights, and discrimination Human oversight measures Specifications on input data
Risk management and risks	Risk identification and analysis Continuous iterative evaluation of the risks Risk management measures Residual risks
Change management	A description of any changes made to the system
Standards	List of harmonized standards applied List of other relevant standards and technical specifications applied
Declaration of conformity	A copy of the EU declaration of conformity
Post-market monitoring plan	A system to evaluate the performance in the post-market phase

A particularly interesting and important piece of documentation required is the "Instructions of use": the documentation attached to a high-risk system by the provider and also available for the public via (at a minimum) a specialized EU Database. We anticipate this requirement will play a highly influential role in facilitating supply-chain transparency of AI, and will quickly find its way to AI technology contracts between various parties. It is very clear by the requirements, and validated in the EC's impact assessment, that the document is designed in a way that provides valuable information of the key characteristics of the system while safeguarding companies' intellectual property ("IP"). We suggest including the input data specifications in all instructions of use. We therefore advise removing a small but potentially deteriorating condition in the current draft: "when appropriate."

Table 4: Instructions of use as outlined in the AIA

Provider contact details	Identity and the contact details of the provider
Characteristics, capabilities, and limitations of performance of the system	Intended purpose Level of accuracy, robustness, and cybersecurity Foreseeable circumstances which may lead to risks to health and safety or fundamental rights Performance as regards the persons on which the system is intended to be used Specifications on input data
Pre-determined changes	Any required or implemented changes to the system and its performance already recognized by the provider from initial conformity assessment.
Human oversight measures	Human oversight measures, incl. technical measures to facilitate the interpretation of the outputs
Expected lifetime and necessary maintenance measures	Expected lifetime of the system Necessary maintenance and care measures

For the detailed interpretation of the required documentation, industry practices and standards[33] will play an important role in helping companies operationalize the requirements in their everyday processes. At the same time, no AI providers or deployers should use missing standards as an excuse not to pay attention to good documentation practices in developing high-risk systems. The best way to prepare is to gradually take into use practices that are aligned with the proposed requirements.

Note that transparency information should ultimately ground out in the system itself – its code, development (revision control) history, data, and hardware realization. This is good practice for allowing developers to understand, maintain, and improve their own system, as well as for carrying out in-house checks on everything from cybersecurity to the efficacy of developer staff. Ideally, developers would feel neither the need nor the possibility to "Volkswagen" the documentation of their system into separate, irreconcilable pathways for regulators rather than real-world use. Rather, we should want them to develop or deploy tools that, in a lightweight manner, allow the same information to serve multiple purposes. These can and should include cybersecurity defenses to ensure corporate secrets are only revealed in-house or to trusted (and intended) auditors.

Contextual transparency reporting to AI end users. While the main focus of the proposal is in setting specific requirements for high-risk AI systems, what is laid down in the Article 52 regarding transparency obligations of systems that interact with natural persons is definitely worth mentioning. Positively thinking, this short article is addressing what has become a major challenge with the GDPR informing practices (privacy policies): they're out of context. The requirement of the AIA is focused on the actual use context. It simply requires that an end-user is made aware of interacting with an AI system. This may well mean that industry standards around labelling AI products will finally start to emerge as providers begin to mark their end-user interfaces accordingly. Moreover, the AIA requires the deployers of emotion intelligence, biometric categorization, and deep fake systems to inform natural persons of their exposure to such AI systems.

Ideally, the AIA might become a new vanguard for transparency more generally. Again, taking proportionality into account, companies and other organizations may choose to expose not only the minimal amount of transparency required by the law (e.g. whether the system deploys AI) but also other aspects of their transparency documentation. This should probably be done in a hierarchical way so that ordinary end-users are not overwhelmed by complexity, nor are small companies required to maintain multiple different types of documentation (which would almost certainly soon fall out of synchronization). But where providers are comfortable exposing the capacity to "drill down" into the same documentation used for regulatory and self-documentation purposes, they may find that they facilitate trust in or engagement with their AI systems. Some public authorities have already started to implement such transparency via public AI registers, as also recommended by the EC in the coordinated plan for AI.[34]

[33] European Commission, Joint Research Centre, Nativi, S., De Nigris, S. & AI Watch, AI standardisation landscape state of play and link to the EC proposal for an AI regulatory framework, Publications Office, 2021, https://data.europa.eu/doi/10.2760/376602.

[34] Coordinated Plan on Artificial Intelligence 2021 Review by the European Commission, April 21, 2021.

The EU transparency database – likely to become a key vehicle for public oversight. The system presently known as the "EU database for stand-alone high-risk AI systems" is as we have said a key concept. It is mandatory for high-risk systems, but we recommend it should be made available – on a voluntary and proportionate basis to all AI systems. It should also be consolidated with the transparency requirements of the DSA. Right now, the concept of this transparency database is another well-hidden, golden secret of the proposal. In short, all stand-alone high-risk systems (Annex III) that are made available, placed on market, or put on service in the EU will be searchable via a centralized database controlled by the EC. Presently in our opinion, the EC's thinking around objectives for the role of the database is not made sufficiently clear. While the potential uses for such a database are many, we would like to envision a few in order to understand the nature of the net impact.

Table 5: Anticipated impacts of an EU Transparency Database (Article 60)

Impacts to	Positive impacts	Both positive and negative impacts	Negative impacts
Providers developing AI products for sale	Gain competitive insights about available products, their workings, governance and contacts	Expose systems for wider visibility among potential customers, end-users, potential competitors, researchers, journalists and activists	Submit and maintain data in EU database (note: this cost would be minimal due to no additional documentation beyond what's required for conformity assessments is required for EU database).
Providers developing AI products for their own use	Gain market insights about available products, their workings, governance and contacts	Expose systems for wider visibility among potential end-users, potential competitors, researchers, journalists and activists	Submit and maintain data in EU database (see note above)
Deployers	Gain market insights about available products, their workings, governance and contacts Verify conformity to law of the third-party systems		
End users	Verify conformity to law of systems one is interacting with		
Researchers, journalists, activists, general public	Gain market insights about available products, their workings, governance and contacts Gain market insights about providers and their product portfolios Gain market insights about the product market developments Verify conformity to law of systems in the market Source material for information services to potentially connect AI incidents to similar systems in the market		
Supervisory authorities, market surveillance authorities, European Commission etc.	Gain market insights about available products, their workings, governance and contacts Gain market insights about providers and their product portfolios Gain market insights about the product market developments		

Based on this short analysis, the EU Transparency Database is likely to have both positive as well as negative impacts on the providers of the high-risk systems. Even where there are negative costs such as those associated with extra documentation, these may be ameliorated by unification with standard development and operations practices within the firm. For this reason, it seems quite likely that firms and governments may choose, and indeed insurance organizations may advise, that the database be used well beyond the “certainly high-risk AI” classification. We might for example imagine a small firm having run-away success and becoming concerned about the larger user base and wider range of applications than they originally anticipated asking to go through the exercise of checking compliance for the documentation of their system even before being required to do so. Such a choice should certainly be rewarded by the courts as evidence of good practice should an unanticipated outcome of the system's deployment prove to be socially costly.

The main source of potentially negative impact, therefore, is via increased competitive and critical civil society exposure to systems, increasing thus the competitive and brand reputation risks of providers. The incremental administrative effort of submitting the data to the EU Database after the conformity assessment seems minimal. For all other parties, including deployers, the impact is clearly positive and would deserve an even more deliberate separate analysis.

Finally, we briefly outline the likely impacts to companies' and public organizations' everyday AI development, when they ensure compliance with the new EU requirements. To start with, many AI providers will face, and may already be facing, the impacts of the proposed AI Act through new incoming requirements in procurement.[35] We believe this mechanism will have a significant transformative impact on industries even prior to the regulation being fully in place. Moreover, we foresee specific contractual clauses being established between the AI providers and deployers, to limit the use of providers' technologies to the ones defined in the contracts and instructions of use, as well as securing proper oversight and maintenance measures by deployers.

In organizations with established data protection practices, the existing structures can be relatively effectively adjusted to respond to the expectations of the AIA. For some organizations, the AIA will become the driver to finally deploy risk management that is long overdue. While such processes can be effectively reused, organizations will need to establish systematic documentation practices across AI portfolios, e.g. via AI registers. The main challenge for organizations will be: who to assign the responsibility, and how to systematize keeping the documentation up to date over the lifecycle of their AI product? Again, these challenges are ones faced by all organizations delivering complex, engineered products, regardless of legal requirements. Further, for digital products, the potential for automated tools for both capturing and then simplifying or distilling such information are both high.

The costs of implementing the AIA requirements obviously depend on the risk level of a given system, as well as an organization's preparedness prior to the new regulation. We provide here a short overview of the costs as anticipated in the EC's Impact Assessment.

> ***The main source of potentially negative impact, therefore, is via increased competitive and critical civil society exposure to systems, increasing thus the competitive and brand reputation risks of providers***

The EC addresses the costs of compliance for individual organizations and verification costs. Focusing on high-risk systems to which the AIA requirements are mostly addressed, the EC's rough estimate for an organization's first compliance cost i.e. fulfilling the requirements outlined in Chapter 2 of the AIA, is around 6000-7000€ for a typical AI project (170 000€) or ca. 4-5 percent. Those providers who would need to go through a conformity assessment process by a third party, would face an additional 3000-7500€ or 2-5% per system assuming the provider has an existing Quality Management System ("QMS") in place and audited. Finally, deployers of the high-risk systems would face an additional human oversight cost of around €5000 – €8000 per year. While the bulk of these estimates look reasonable and in line with our practical experience, a deeper analysis reveals that potentially even unproportionately-high cost implications could occur if the scope of third-party verification would be extended beyond its current scale. We encourage a review on the cost impacts for all parties to ensure any suggestions are rooted on solid understanding of financial impacts.

[35] See, e.g. reports by the City of Amsterdam, Telstra, and the World Economic Forum. See https://www.amsterdam.nl/innovatie/digitalisering-technologie/algoritmen-ai/contractual-terms-for-algorithms/, https://www.itnews.com.au/news/telstra-creates-standards-to-govern-ai-buying-use-567005 and https://www.weforum.org/whitepapers/ai-government-procurement-guidelines, respectively.

Table 6: Compliance costs of providers and deployers

Compliance costs	**Providers**	**Deployers**
Compliance costs	6,000 - 7,000€[36]	5,000 - 8,000€[37]

Table 7: Verification costs depending on conformity process

Verification costs	**Provider**	
	Enterprise	**SME**
Verification costs based on third-party assessment[38]	3,000 - 7,500€	3,000 - 7,500€
Verification costs based on internal control	0€	0€

Finally, we shouldn't underestimate the importance of the proposed structures enabling public scrutiny. We believe both the EU Database as well as the end-user transparency requirements will have a significant impact on enabling democratic oversight by citizens, civil society activists, journalists, and researchers. Providers of AI should prepare for welcoming such public discourse as a source for continuous feedback and faster identification of potentially harmful impacts. No doubt such public interest will also increase organizations' brand risk associated with AI, but this only calls for better preparedness, which is of course the goal of the regulation.

With its proposal, the EC has shown a way to manage AI-related risks to health, safety, fundamental rights, and even social stability in a way that has all the means to incentivize the industry to take appropriate action. This is of fundamental importance, offering an opportunity to governance efficiency in regulating technologies the influences and impacts of which will be significant, and are already substantial though perhaps under-recognized. We have in this document highlighted and amplified a few open concerns that need to be addressed in the refinement of the AIA. But the bulk of our article is aimed to defend the act against assaults from those who, whether out of misplaced concern, or perhaps overestimating costs, will try to shirk these obligations. Those who see the AIA as too much government interference are perhaps underestimating the importance and value of high-quality regulatory oversight, even to their own endeavor. ■

[36] AIA Impact Assessment, p.70.

[37] AIA Impact Assessment, p.71.

[38] AIA Impact Assessment, p.71.

CPI SUBSCRIPTIONS

CPI reaches more than **35,000 readers** in over **150 countries** every day. Our online library houses over **23,000 papers**, articles and interviews.

Visit **competitionpolicyinternational.com** today to see our available plans and join CPI's global community of antitrust experts.

CPI COMPETITION POLICY™ INTERNATIONAL