SHERPA

Shaping the ethical dimensions of smart information systems– a European perspective (SHERPA)

**Feedback to the
European Commission on its**

**Inception Impact Assessment
for a legal act on the regulation of artificial intelligence**

Based on research from the SHERPA project[1]

September 10, 2020

This project has received funding from the

**European Union's Horizon 2020 Research and Innovation Programme
Under Grant Agreement no. 786641**

SHERPA (Shaping the ethical dimensions of smart information systems (SIS) – a European perspective) is an EU-funded project that focuses on ethical and human rights aspects of smart information systems (artificial intelligence and big data analytics). SHERPA aims to ensure that its recommendations and findings help to move the AI ecosystems in the desirable ethical and human-rights respectful direction. The final recommendations from SHERPA are due in April 2021.

The SHERPA consortium welcomes the opportunity to provide feedback to the European Commission on its Inception Impact Assessment on the proposal for a legal act of the European Parliament and the Council laying down requirements for AI. This feedback is a follow-up to previous comments by the SHERPA consortium in response to the public consultation on the European Commission's White Paper on AI.

Feedback on 'A. Context, Problem definition and Subsidiarity Check'

The SHERPA consortium welcomes the language in section A that prioritises protection of human rights and recognises both the benefits and harms of AI. We have highlighted some strong statements below, and we recommend the Commission keep these statements and concerns as a central part of the proposed regulatory framework.

We strongly agree with the articulation that the "***ultimate objective is to foster the development and uptake of safe and lawful AI that respects fundamental rights*** *across the Single Market by both private and public actors while ensuring inclusive societal outcomes.*" The protection of human rights and safety is a societal priority, and not only a means to create (or the outcome of creating) trustworthy AI. To make this clearer, the regulation should explicitly acknowledge the potential for misuse and abuse of artificial intelligence, and hold developers, deployers, and users to the ethical principle of non-maleficence.

We strongly agree that existing EU law needs more effective enforcement, particularly related to fundamental rights and freedoms. An EU approach should build on existing legislation and enforcement mechanisms. A comprehensive gaps analysis is needed to identify & prioritise required regulatory reform. SHERPA has analysed various regulatory options and specific measures that support the ethical and/or responsible development of AI and big data which might be useful in this context.[2]

We strongly recommend the creation and/or promotion of effective redress mechanisms, particularly as "[*AI*] *may make it difficult for persons having suffered harm to obtain compensation under the current EU product liability legislation.*" This should be an explicit part of the regulatory framework.

We strongly agree that we must be concerned about intended and unintended negative outcomes. An EU approach must be based on comprehensive assessment of risks; relying on self-reported risks related to intended use will not capture realistic concerns and potentially significant negative outcomes.

---

[2] SHERPA, Report on Regulatory Options, December 2019, https://doi.org/10.21253/DMU.11618211

Feedback on 'B. Objectives and Policy options'

Comments on objectives
- Aim (a): We recommend clarification on the definition of 'illegal discrimination' to ensure it is in line with the principle of non-discrimination. Additionally, the meaning of "avoid illegal discrimination by ensuring the relevant documentation for the purposes of private and public enforcement of EU rules" is unclear. Discrimination is avoided by actors taking appropriate actions to prevent unjust or prejudicial treatment, not by 'relevant documentation'.
- Aim (c): We recommend removing 'where possible', as that significantly dilutes the potential impact.
- Aim (e): We recommend further consideration of the creation of a central coordinating institution at the EU level. SHERPA is currently developing Terms of Reference for an EU Agency for AI, in consultation with experts and stakeholders, which should inform the Commission. A central body with sufficient funding and legal and technical competences will be critical to the effectiveness of a network of national competent authorities.
- Aim (f): We recommend removing "ensuring a level playing field". This is not practically possible and may indirectly encourage measures that undermine the ultimate objective of the regulatory framework to protect human rights and safety.

Comments on policy options
The SHERPA project strongly supports Option 4, which would be a combination of Option 2 (EU legislative instrument setting up a voluntary labelling scheme) and Option 3 (EU legislative instrument establishing mandatory requirement for all or certain types of AI). The SHERPA project might support limited regulation for high-risk only AI (sub-option 3(b)), depending on how risk is determined, by whom, and which criteria are used.

We believe that it is absolutely critical that the EU develop a mandatory regulatory framework of ex-ante and ex-post enforcement mechanisms to ensure that AI systems are safe and do not violate fundamental rights and ethical principles. This regulatory framework must have a "higher degree of detail and specificity" in order for developers and users to understand their legal obligations, for EU and national authorities to monitor compliance, and for the general public to have trust in the regulatory framework. There is already a proliferation of largely ineffective "high-level principles and obligations … completed by industry-led norms such as in the form of standards or codes of conduct." Furthermore, these initiatives are not always consistent and are therefore confusing for researchers, developers and users.

A complementary voluntary labelling scheme would be useful, but only if robust, clearly articulated, and not symbolic.

We strongly recommend against the baseline option (no action) and Option 1 (EU 'soft law'). While soft law initiatives may be useful in some context and could complement legislation, they do not address legal issues of liability and other ensuring harms, which is critically needed. Additionally, Member States will continue to develop and implement their own regulatory frameworks if no action is taken at the EU-level, contributing to further fragmentation and confusion. A strong legal standard must be set at the EU-level that establishes a baseline and encourages high standards of protection of fundamental rights and societal values.

Feedback on 'C. Preliminary Assessment of Expected Impacts'

The SHERPA project strongly recommends further assessment of expected impacts. An impact assessment should also evaluate the risks associated with failure to act, and not only the impacts of a potential regulatory framework. The social impacts of the proposed options have not been addressed at all (e.g., impacts on how society understands trust or how the options may create and/or promote a false sense of trustworthiness).

SHERPA Contributions to the AI Ethics and Human Rights Ecosystem

SHERPA, itself part of an ecosystem consisting of other Swafs projects (e.g., SIENNA, PANELFIT) and numerous research groups and projects, is working on the following activities:

1. Contribution to knowledge base and curricula

SHERPA has done extensive empirical and conceptual research[3] allowing the consortium to build up capacity to contribute to the definition of a required body of knowledge and model curricula. In addition to factual knowledge, SHERPA has also produced audiovisual material and artistic representations of key aspects of AI and ethics to allow reaching out to a broader audience.

2. Terms of Reference for regulator

SHERPA has undertaken a review of regulatory options[4] and is working on the terms of reference for a regulatory body that could be at the heart of the framework of AI regulation.

3. Ethics by design

SHERPA (in collaboration with SIENNA) has developed guidelines for AI developers and AI users that are based on the principles of ethics by design.[5] It has provided an initial training session for about 80 Project and Policy Officers of the EC (December 2019). SHERPA can contribute to the design and delivery of further relevant training.

4. Standardisation

SHERPA members are involved in ISO SC 42 and are contributing to the development of ethically-aware standardisation.

SHERPA aims to ensure that its recommendations and findings help to move the AI ecosystems in the desirable direction and is requesting feedback from stakeholders to help it further shape recommendations and outputs. For more information, please visit: https://www.project-sherpa.eu

---

[3] SHERPA, SIS Workbook, https://www.project-sherpa.eu/workbook/.
[4] SHERPA, Report on Regulatory Options, December 2019, https://doi.org/10.21253/DMU.11618211.
[5] SHERPA, Guidelines for the Ethical Use of AI and Big Data Systems, https://www.project-sherpa.eu/wp-content/uploads/2019/12/use-final.pdf.