GÖTEBORGS UNIVERSITET

IT-fakulteten
Cecilia Ihse, kanslichef
E-post: Cecilia.ihse@itfak.gu.se
Tel: 031-786 9037

Dnr:
GU 2021/1541
I2021/01304

Till Infrastrukturdepartementet

Göteborg 2021-06-23

Yttrande från Göteborgs universitet avseende Remiss till Europeiska kommissionens förslag till förordning om harmoniserade regler för artificiell intelligens (dnr I2021/01304)

IT-fakulteten vid Göteborgs universitet har tagit del av Remiss gällande Europeisk kommissionens förslag till förordning om harmoniserade regler för artificiell intelligens, Proposal for a Regulation laying down harmonised rules on artificial intelligence (artificial intelligence act) and mending certain union legislative acts (COM (2021) 206). Vid fakulteten finns en flervetenskaplig och mycket internationell forskningsmiljö där vi arbetar med allt från abstrakt matematik till högteknologiska tillämpningar och samhällsvetenskapliga studier av digitaliseringens konsekvenser.

Detta är IT-fakultetens och Göteborgs universitet yttrande avseende remissen. Då remissen som helhet är på engelska har vi valt att besvara den på engelska.

Yttrande

The IT Faculty is, in general, positive towards the regulation of using AI systems. The current state of the development of AI systems and its wide usage brings many risks and therefore it is of outmost importance to start to regulate the usage of AI. High-risk AI systems have started to influence our politics (via social media content direction), business models (via customer data analytics) and our ability to act freely (via biometric identification and categorization).

It is very good that the proposal addresses the challenges related to data and data governance. However, **we would like to see more elaborative regulation in terms of licensing the data** and potential regulation on which and how data should and could be licensed between parties.

In the current proposal, **we would like to point out that the current proposal only deals with physical and psychological harm, explicitly excluding the economic harm**. However, the recent events related to hacker attacks, like the attack of the oil pipeline in the US, show that the economic aspects have equal (if not larger) social impact then the physical and psychological. Therefore, we would like the proposal to include even the economical and societal aspects in defining high-risk AI systems.

More specific points:

- Title II, article 5, point 1d) (p. 43-44): We should be able to specify how these activities are to be monitored in practice. For example, who will be responsible for monitoring and then distinguishing between the approved and disapproved use, given the fact that AI is not always correct (probabilistic nature of AI).
- Title II, article 5, point 2 (p. 44): How will the assessment of "seriousness, probability and scale of those consequences" be done in practice?
- In Title II, children are identified as a special group, in several points. How about other vulnerable groups, e.g. based on their ethnicity, diseases, or other properties.
- In relation to the definitions in Article 3 definitions, point (33) (p. 42): We have one observation/question. The statement ”behavioural characteristics of a natural person” is not exemplified. Does this include also digital traces in the form of input patterns when interacting with digital platforms? Such as keystrokes, pauses, cursor behaviour etc.?
- Transparency obligations: Under GENERAL PROVISIONS, Article 1 Subject matter c) (p. 38) “harmonised transparency rules for AI systems intended to interact with natural persons, emotion recognition systems and biometric categorisation systems, and AI systems used to generate or manipulate image, audio or video content;” However, in the text that follows, this general formulation is substantially narrowed down to **systems used to generate or manipulate content that appreciably resembles existing persons.** But what about textual content? It is also being manipulated on the web. Information we get is often aggregated and processed by AI. Existing information is not only changed and generated but also omitted.
- TITLE IV TRANSPARENCY OBLIGATIONS FOR CERTAIN AI SYSTEMS:
    - Here we find formulations that consider direct contact and immediate consequence: (point 70, p. 34): ”interact with natural persons or to generate content”, ”natural persons should be notified that they are interacting with an AI system, unless this is obvious from the circumstances and the context of use”. It is very good to have “transparency obligation” in the direct contact with an AI (including the web & voice agents). But what is the position about the indirect (implicit) interaction?
    - How about being informed that the lawyer is an AI, or that the political decision is based on AI, or that the job applications are reviewed by AI? The direct contact can be with a human, but indirectly AI may be making decisions.

    - In case of media moderation: shall the citizens be able to find out that the content is AI generated or moderated? Based on what criteria/values?
    - In the school context: Shall students be informed that the exam is corrected by AI? (It should not be controversial; we have already programs like URKUND and students know their texts are sent for control to that program. But if the program will grade exams, without teacher's involvement, would that be acceptable?)
- Biometrics and surveillance: "content that appreciably resembles existing persons" (p. 34), *appreciably resemble* will be a matter of judgement as it depends on the interpreter.
    - Perhaps a culture grown around AI usage can help people to get a feeling what is decent or not. Initially, it will take time to establish.
    - What about intelligent automated surveillance systems? People are typically informed about the surveillance cameras in the city, shops etc. Is there a law about transparency of surveillance of citizens (not criminals)? Or a surveillance of labor force to measure productivity?
    - Related to school context: how about automated intelligent surveillance of students? Will that be allowed (under transparency obligation)?
    - The above examples of high-risk AI systems which can present risks for human rights and democracy are listed in Annex III and shall also be considered as impotent high-risk, substantial for democratic society. So, the above questions about transparency obligations in case of high-risk for democracy and human rights should be added to the list of transparency obligations. It is in the first place central that stakeholders clearly understand their hazards.
- The following high-risk AI systems from ANNEX III HIGH-RISK AI SYSTEMS REFERRED TO IN ARTICLE 6 (2) Should be mentioned in the main text, just as categories, because they are so central, and can be presented in detail in the Annex.:
    - 1. Biometric identification and categorization of natural persons
    - 2. Management and operation of critical infrastructure
    - 3. Education and vocational training
    - 4. Employment, workers management and access to self-employment
    - 5. Access to and enjoyment of essential private services and public services and benefits
    - 6. Law enforcement
    - 7. Migration, asylum and border control management
    - 8. Administration of justice and democratic processes
  - HIGH-RISK: To the criterion: "whether an AI system poses a risk of harm to the health and safety or a risk of adverse impact on fundamental rights" (p. 45), one might add threat to democracy and cultural values.
  - TITLE IX, CODES OF CONDUCT, Article 69, "codes of conduct intended to foster the voluntary application to AI systems other than high-risk AI systems of the requirements set out in Title III" (p. 80): The AI field is in fast and intense transformation, and we may safely assume that we do not yet know all possible high-risk systems. We will be discovering as we go. Therefore, the AI code of

ethics should be available to stakeholders such as students of different fields that develop AI, engineers and designers, practitioners, and other groups of stakeholders. The culture of responsible AI development is developed over time and the best way to prevent people from creating and deploying high-risk AI systems.

- REGARDING THE LEGISLATION AS A WHOLE
  - The process: How is the process of continuous update of regulation envisaged? The technology will continue developing quickly and pose new challenges.
  - How will new insights from the use of regulation be incorporated?

Enligt uppdrag

Prof. Miroslaw Staron, prodekan, IT-fakulteten

Prof. Gordana Dodig Crnkovic, institutionen för data- och Informationsteknik

Prof. Jonas Ivarsson, institutionen för tillämpad informationsteknologi

Kopia

Registrator Göteborgs universitet