DIRITTI COMPARATI
Comparare i diritti fondamentali in Europa

ACCOUNTABILITY, TRANSPARENCY, AND FAIRNESS TO ASSESS GENERATIVE AI SOLUTIONS WITH THE LENSES OF DATA PROTECTION LAW

Posted on 14 Giugno 2023 by ***Denise Amram***

At the end of March 2023, the Italian Data Protection authority imposed OpenAI LLC a temporary limitation in Italy of the ChatGPT services, including a ban on processing under article 58, §2, sub f) of the EU Regulation 2016/679 on General Data Protection Regulation (hereinafter GDPR), because of several critical aspects impacting on their users – especially children- 's fundamental rights protection.

The decision considered the following grounds of assessment. Firstly, the lack of a privacy information policy explaining users and data subjects how OpenAI LLC would have collected and processed data in the platform. Secondly, the lack of a legal basis to process personal data for algorithms training purposes. Thirdly, the lack of safeguards to assess users' age in order to avoid minors under 13 years old to use the software in alignment with the included terms and conditions, that -in any case- have been considered misleading in some parts related to data processing activities.

That decision opened an international debate on risks and opportunities of OpenAI applications, providing a *domino effect* in other EU systems: for example, on April 13th, the EDPB launched a task force on ChatGPT, while on April 23rd the French CNIL published a *Dossier* on the generative AI (https://linc.cnil.fr/dossier-ia-generative-chatgpt-un-beau-parleur-bien-entraine) to explain the chatbot mechanisms and their effects from an ethical-legal perspective.

In April 2023, the limitation has been overruled since the US company implemented a series of technical and organisational measures able to mitigate any risks of fundamental rights compromission. In particular, the platform implemented a more easily accessible procedure for opting-out from data processing either for users or non-users, it published in the webpage a detailed privacy policy also for the data processed to train the algorithm; it developed mechanisms to allow users to erase and or correct possible inaccuracies. These safeguards have been considered first essential measures to make that chatbot service available again.

Thus, the Italian *affaire* on ChatGPT is particularly interesting to be analysed under several perspectives.

It shall be considered a worthy sample of joint collaboration with the competent data protection authority to identify proper organisational and technical measures to mitigate the impact of a given data processing activity on the users' fundamental rights, even if stimulated by an investigation instead of being promoted *by design*. In fact, the interaction between the data controller and the competent authority is stated by article 36 GDPR, that establishes the conditions for a "prior consultation" mechanism in case the results of the "self" data protection assessment referred to a high-risk processing activity for the data subjects' fundamental rights and liberties or if none of the implemented safeguards could ensure an appropriate level of mitigation.

Moreover, the ChatGPT case highlighted the role of the data protection law in addressing the standards of accountability for AI-based solutions. In fact, since the AI package has not been approved yet, the lack of a specific setting of obligations for AI developers is covered by the GDPR at least as far as personal data is concerned. Thus, the GDPR could protect only direct or potential users of a given AI-based application, as long as they could be considered as data subjects of a given data processing. However, any further implication on fundamental rights not included in a data protection impact assessment is still not enforceable by a data protection authority at this stage. For this reason, many data protection authorities decided to open dedicated departments and task forces in order to specifically address data protection issues related to AI-based solutions. In fact, despite of the decision to restore the ChatGPT services, the monitoring and risk assessment activities on generative AI have just started in light of the principles of accountability, transparency, and fairness.

In this regard, the grounds analysed by the Italian Data Protection authority are shaping a minimum standard suitable to be applied to analyse the effects of AI solution, including the generative ones with the lenses of data protection law. In particular, it appeared that the developer data controller has at least: i) to address the risks considering the different categories of data subjects/users and their tailored vulnerabilities; ii) to ensure a transparent and clear information policy; iii) to ensure easily mechanisms to opt out from the data processing activities.

As far as the sub i), children are *per se* considered as vulnerable users and specific technical and organisational measures have been required to assess their age in order to establish an aware and lawful contractual relationship between the service providing the chat bot and the user. However, individual digital vulnerabilities emerging from the digital divide or to the possible consequences on the human oversight are not protected by the data protection law. They could be addressed under a trustworthy assessment of the given AI-based application (as envisaged by the High Level Group on AI) but they cannot in the context of pure data protection impact assessment under article 35 GDPR.

The ground related to the transparency (ii) is limited as well. In fact, according to the GDPR some information on the data processing are mandatory, including those related to possible profiling activities – that are the ones undertaken by the algorithms. However, the so-called privacy policy does not require any details on what could happen to data once they are anonymised. Even if it could be essential to understand the implications to use a given application. As a consequence, to ensure a user-friendly mechanism to opt out (iii) from the data processing activities undertaken by the given application is a limited obligation under the data protection law. In fact, it is applicable as long as the data are personal (including the pseudonymised ones), but the data subject will completely loose the information control once that the data have been made sufficiently anonymous. In addition, the data protection law cannot solve possible issues related to technical *bias* that could bring as a result a discriminatory decision for a category of persons, but only if the decision is directly impacting on the data subject (see the *Deliveroo case* - Tribunal of Bologna 31.12.2020). Therefore, also the fairness achieved through the GDPR compliance is limited respect to the range of possible risks of fundamental rights compromission.

Such a short analysis aimed to highlight how the ChatGPT Italian *affaire* has been essential to develop a remarkable assessment on AI applications with the lenses of the privacy and data protection compliance. It opened a serious debate on the urgent necessity to extend the analysis on all possible ethical and legal implications of a given solution through a consolidated methodology based on the principles of accountability, transparency, and fairness by specifying the limitations that could be met in the context of a data protection investigation by the competent authority.

Acknowledgement

This contribution has been partially supported by Programma Operativo Nazionale Ricerca e Innovazione 2014-2020, PON “Il danno alla persona e la giustizia predittiva”.