L’utilisation de ChatGPT temporairement restreinte en Italie : une décision motivée par la pizza hawaïenne ou les pâtes au ketchup ?

Livio di Tria, le 1er avril 2023

Le Président de l’autorité italienne de protection des données, le Garante per la protezione dei dati personali, a pris une décision le 30 mars 2023 qui limite temporairement le traitement de données personnelles par le service « ChatGPT » d’OpenAI.

Garante per la protezione dei dati personali, Provvedimento del 30 marzo 2023

De quoi parle-t-on ?

Le *Presidente del Garante per la protezione dei dati personali* (GPDP), soit le Président de l’autorité italienne de protection des données, a prononcé le 30 mars 2023 une décision enjoignant à OpenAI de limiter temporairement le traitement de données personnelles via son – désormais célèbre – service ChatGPT, ce qui implique que le service n’est plus accessible depuis l’Italie. Parallèlement, une enquête administrative a été ouverte.

Nous soulignons qu’un communiqué de presse a également été publié par le GPDP le 31 mars 2023. Celui-ci mentionne expressément la faille de sécurité qui a touché le service ChatGPT le 20 mars 2023 lors de laquelle les conversations des utilisateurs et les informations de paiement des abonnés au service payant ont été brièvement exposées. Cette faille pourrait d’ailleurs être le déclencheur initial de l’affaire.

Remarques liminaires

Préalablement à l’analyse de la décision (dont le titre de la contribution humoristique ne reflète pas l’importance de la décision, mais se veut provocateur)[1], nous nous devons de formuler plusieurs remarques liminaires nécessaires à la compréhension du cadre dans laquelle la décision a été rendue.

Premièrement, la décision s’étend à toutes les données personnelles des personnes concernées établies sur le territoire italien. Elle a été rendue en conformité avec l’art. 58 par. 2 let. f RGPD qui permet à chaque autorité d’adopter une mesure correctrice permettant d’imposer une limitation temporaire ou définitive, y compris une interdiction, du traitement.

Deuxièmement, la décision a été rendue par le Président du GPDP comme le lui permet l'art. 5 par. 8 du Règlement 1/2000 relatif à l'organisation et au fonctionnement du Bureau du Garante de la protection des données. Ce dernier prévoit que dans les cas d'urgence particulière et d'impossibilité de report qui ne permettent pas de convoquer le GPDP en temps utile, le Président peut adopter les mesures relevant de la compétence de l'organe. Ces mesures doivent toutefois être ratifiées par le GPDP qui, en tant qu'organe collégial, est apte à prendre les décisions fondées sur le RGPD. À cet égard, la décision doit être ratifiée lors de la première réunion utile, à convoquer au plus tard le trentième jour suivant le prononcé de la décision.

Troisièmement, OpenAI ne dispose pas d'un établissement au sein de l'Union européenne. C'est la raison pour laquelle la décision a été notifiée à son représentant (art. 27 RGPD) au sein de l'Union européenne. Conformément à la politique de confidentialité d'OpenAI, le représentant désigné pour les utilisateurs européens est l'entreprise VeraSafe Ireland Ltd

Quatrièmement, la décision est temporaire. OpenAI dispose d'un délai de 20 jours pour communiquer les initiatives prises pour mettre en œuvre les prescriptions de la décision et fournir tout élément jugé utile pour justifier les violations identifiées par le GPDP. À noter que l'absence de réponse à la demande du GPDP est passible d'une amende administrative pouvant s'élever jusqu'à EUR 20'000'000 ou, dans le cas d'une entreprise, jusqu'à 4% du chiffre d'affaires annuel mondial (art. 83 par. 5 let. e RGPD).

Le dernier point, mais non des moindres, concerne l'impact pratique que pourra avoir la décision et son effet boule de neige. Si la décision dont il est question ne déploie des effets que sur le territoire italien, d'autres autorités européennes ont déjà pris contact avec leur homologue italien afin d'échanger sur les constats qui ont pu être faits. C'est par exemple le cas de la Commission nationale de l'informatique et des libertés (CNIL), soit l'autorité française de protection des données. Il ne serait donc pas impossible que d'autres décisions soient rendues prochainement.

Qu'est-ce qu'on reproche à OpenAI et à son service ChatGPT ?

En raison du caractère urgent, et du fait que l'enquête en est à ses prémisses, la décision du GPDP n'est que partiellement motivée. À ce stade, seuls trois points peuvent être mis en exergue.

Premièrement, le GPDP est d'avis qu'aucune information n'est fournie aux utilisateurs et aux personnes concernées dont les données sont collectées par OpenAI via ChatGPT. Le GPDP fait

ici référence au devoir d'informer à charge du responsable du traitement tel que fixé par les art. 13 et 14 RGPD. Nous relevons qu'OpenAI dispose d'une politique de confidentialité abordant la question des données collectées et traitées (cf. par. 1 de la politique). Si les informations collectées lors de la création d'un compte – processus obligatoire pour utiliser ChatGPT – sont pour leur part détaillées, les informations qui sont collectées lors de l'utilisation de ChatGPT ne sont pas aussi claires. Cela soulève en effet des questions dans la mesure où, lorsqu'une personne utilise ChatGPT, elle peut non seulement révéler des données personnelles qui la concernent, mais aussi des informations qui concernent des tiers (p. ex. lorsqu'on demande à ChatGPT de rédiger une lettre à l'attention d'une personne précise). De telles données peuvent en outre être inexactes, ce qui est également un problème dès lors que le service utilise ces données en vue d'entraîner les algorithmes sur lesquels il s'appuie.

Deuxièmement, le GPDP relève qu'aucune base juridique au sens de l'art. 6 RGPD ne semble sous-tendre la collecte et le traitement massifs de données personnelles en vue d'entraîner les algorithmes sur lesquels le service s'appuie. Ce point est particulièrement problématique puisqu'en droit européen, le traitement doit nécessairement reposer sur l'une des bases juridiques prévues par l'art. 6 RGPD. En l'espèce, force est d'admettre que la politique de confidentialité d'OpenAI ne détaille pas sur quel(s) fondement(s) les traitements reposent. La critique du GPDP ne prête donc pas le flanc à la critique.

Troisièmement, le GPDP relève que le service est interdit aux enfants de moins de 13 ans (cf. par. 6 de la politique). Malgré cette interdiction, le GPDP relève l'absence de tout mécanisme de vérification de l'âge, ce qui expose les enfants à recevoir des réponses inappropriées à leur âge et à leur sensibilité, et ce contrairement à l'art. 8 RGPD. À ce titre, OpenAI aurait dû mettre en œuvre des mesures techniques et organisationnelles appropriées par défaut et dès la conception du traitement.

Au vu de ce qui précède, le GPDP conclut que le service ChatGPT est contraire aux art. 5, 6, 8, 13 et 25 RGPD.

Intelligence artificielle et protection des données : je t'aime moi non plus

Ces dernières années, le Parlement européen et le Conseil ont explicitement demandé des mesures législatives concernant les systèmes d'intelligence artificielle (IA) afin de garantir un cadre pouvant permettre à la fois la sécurité et l'innovation.

En 2018, la Commission européenne a publié sa « stratégie européenne en matière d'IA », ainsi qu'un « plan coordonné dans le domaine de l'intelligence artificielle », a mis en place un

groupe d'experts de haut niveau sur l'intelligence artificielle et a publié sur cette base, en 2019, des « lignes directrices en matière d'éthique pour une IA digne de confiance ».

En 2020, la Commission européenne a publié dans ce contexte son « livre blanc sur l'IA » qui développait pour la première fois une approche spécifique de la réglementation de l'IA. Sur cette base, la Commission européenne a finalement présenté le 21 avril 2021 une première proposition de règlement sur l'intelligence artificielle (également connu sous le nom de Règlement sur l'IA, ou *AI Act*), qui constitue la première réglementation horizontale juridiquement contraignante au monde sur les systèmes d'IA (et qui prévoit un champ d'application extraterritorial).

Le Règlement sur l'IA – qui n'est qu'au stade de proposition – répartit les systèmes d'IA en catégories et les classe en fonction du degré de risque qu'ils présentent pour la santé, la sécurité et les droits fondamentaux. Il en résulte les quatre catégories de risques suivantes :

- système d'IA interdit ;
- système d'IA à haut risque ;
- système d'IA à risque limité ;
- système d'IA à risque minimal.

Dans la première catégorie (système d'IA interdit), tout ce qui est considéré comme une menace évidente pour les citoyens de l'UE est interdit. Cela concerne par exemple l'évaluation du comportement social par les autorités (ce que l'on appelle le « *social scoring* ») aux jouets avec assistant vocal qui incitent les enfants à adopter un comportement à risque.

La deuxième catégorie (système d'IA à haut risque) comprend d'une part tous les systèmes qui sont des produits couverts par certains autres règlements de l'UE ou qui sont utilisés comme composants de sécurité dans un tel produit lorsque celui-ci est soumis à une évaluation de conformité par un tiers (annexe II). Ensuite, la liste de l'annexe III comprend des systèmes d'IA dont les risques se sont déjà concrétisés ou dont on peut prévoir qu'ils se concrétiseront, dans les domaines de l'identification et de la catégorisation biométriques, des infrastructures critiques, de l'éducation, du lieu de travail, de l'accès aux services, de l'application de la loi, de la migration et de l'asile, ainsi que de l'administration de la justice et des processus démocratiques.

La catégorie à haut risque est la pièce maîtresse de la proposition. Il prévoit, pour les systèmes à haut risque, certaines exigences qui doivent être contrôlées dans le cadre d'un système de gestion des risques. Ces systèmes doivent notamment être développés sur la

base de données répondant à certains critères de qualité et atteindre un niveau défini de précision et de sécurité. Des systèmes de gestion des risques et une documentation technique doivent être tenus à jour et une journalisation automatique doit être assurée.

La troisième catégorie à risque limité comprend les systèmes d'IA qui interagissent avec des personnes physiques (comme les *chatbots*). Ils sont soumis à des obligations de transparence, c'est-à-dire que les personnes doivent être informées qu'elles interagissent avec un système d'IA si cela ne ressort pas clairement du contexte.

La dernière catégorie comprend les systèmes d'IA qui ne présentent pas de risque nécessitant une réglementation et qui ne sont finalement pas concernés par les dispositions du Règlement sur l'IA. La Commission européenne part du principe que la grande majorité des systèmes d'IA entrent dans cette catégorie et cite comme exemple des applications telles que les jeux vidéo basés sur l'IA ou les filtres anti-spam.

Conclusion

La décision du Président du GPDP a été prise dans l'urgence. Elle est certainement motivée par la faille de sécurité dont a été victime OpenAI. Bien que temporaire, la décision du GPDP est importante en raison de la portée qu'elle pourrait avoir (notamment auprès des autres autorités) et en raison de l'augmentation du nombre d'utilisateurs et d'utilisatrices de ChatGPT. Nous nous réjouissons donc des conclusions de l'enquête menée par le GPDP.

La décision confirme également à quel point le développement de l'IA peut représenter un problème en matière de protection des données, raison pour laquelle un encadrement juridique est souhaitable. Seul ce cadre pourra permettre de garantir à la fois la sécurité et l'innovation.

1. L'auteur s'excuse pour le titre de mauvais goût, mais réaffirme par-là que tous les mélanges culinaires ne sont pas nécessairement acceptables. Lorsqu'on lui pose la question de savoir si une pizza hawaïenne ou des pâtes au ketchup sont socialement acceptables, ChatGPT y répond par l'affirmatif. Ni le Chef Alfredo Linguini, ni le Chef Rémy ne sauraient valider cette affirmation, pas plus que l'auteur de ces quelques lignes dont les racines italiennes bouillonnent à la lecture de la réponse comme une marmite prête à exploser.

_swissprivacy.law

Proposition de citation : Livio DI TRIA, L'utilisation de ChatGPT temporairement restreinte en Italie : une décision motivée par la pizza hawaïenne ou les pâtes au ketchup ?, 1er avril 2023 *in* www.swissprivacy.law/213