

# Visite des plate-formes Pharos et Thesee

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 3 octobre 2023

<https://www.bortzmeyer.org/pharos-thesee.html>

---

J'ai visité aujourd'hui les locaux du ministère de l'Intérieur (SDLC, Sous-Direction de Lutte contre la Cybercriminalité, à Nanterre) où sont installées les plate-formes Pharos (signalement de contenus illégaux) et Thesee (gestion centralisée des plaintes pour escroquerie). La visite avait été organisée par le député Renaissance Éric Bothorel.

C'était une visite très cadrée, évidemment. On n'a finalement pas vu la plate-forme Pharos mais elle nous a été décrite comme très proche de celle de Thesee (un grand "open space" avec des PC : rien de spectaculaire). Donc, pas de démonstration de l'informatique de Pharos.

Pharos est le service derrière . Leur rôle est de recevoir des signalements de contenus illégaux par les internautes (identifiés ou pas), de les qualifier (est-ce vraiment illégal), de déclencher une procédure judiciaire si pertinent, d'identifier les auteurs et éventuellement de faire suivre à un autre service. La plate-forme travaille 24x7 et traite parfois des cas urgents (annonce d'un suicide, Pharos prévient la police/gendarmerie la plus proche et intervention immédiate). Pour identifier les auteurs, Pharos utilise diverses méthodes, notamment de réquisitions auprès des FAI/hébergeurs/VPN/etc. Évidemment, leur rapidité et leur efficacité à répondre est très variable. Notez que certains opérateurs de VPN répondent bien, donc l'utilisation d'un VPN ne vous rend pas forcément anonyme. Mais on n'a pas eu de chiffres précis sur le pourcentage de délinquants non identifiés car utilisant un VPN qui ne répond pas.

Tout contenu illégal est concerné : racisme, terrorisme, appel à la violence, pédopornographie, escroquerie (voir plus loin le rôle de Thesee), etc.

Les signalements peuvent être faits par des citoyens ordinaires, ou par des « signaleurs de confiance », les 84 personnes morales qui ont signé un partenariat avec Pharos (la liste ne semble pas publique). Ces signaleurs de confiance peuvent signaler plusieurs URL d'un coup, et mettre une pièce jointe (ce qui n'est pas possible pour les signaleurs ordinaires, entre autres pour des raisons de sécurité).

Pour enquêter, les gens de Pharos utilisent les outils classiques de l'Internet plus des logiciels du marché et des développements locaux. Il ne semble pas qu'il existe un logiciel unique, regroupant tout

ce dont l'enquêteur a besoin. (Mais je rappelle qu'on n'a pas eu de démonstration.) Pour ne pas être identifié comme la police (un site d'escroquerie pourrait leur servir un contenu différent), ils passent par un abonnement Internet ordinaire, qui n'est pas celui du ministère, et éventuellement par des VPN.

Pharos ne publie pas actuellement de rapport d'activité, avec des chiffres. Toutes les données chiffrées qui circulent viennent d'informations ponctuelles (par exemple de questions écrites à l'Assemblée nationale). Un amendement législatif récent (dû à Bothorel, au nom de l'"*open data*") impose la publication d'un rapport d'activité régulier, avec des données chiffrées. La première édition du rapport est d'ailleurs prête, n'attendant que sa révision RGAA. En 2022, il y a eu 175 000 signalements (on ne sait pas pour combien de contenus, plusieurs signalements peuvent concerner le même), 390 procédures judiciaires ouvertes, 89 000 demandes administratives de retrait, 354 blocages exigés (leur liste n'est pas publique contrairement à, par exemple, celle de l'ANJ).

Le nombre de signalements varie beaucoup avec l'actualité, par exemple lors des troubles suivant la mort de Nahel Merzouk à Nanterre. C'est notamment le cas pour la rubrique « terrorisme », très liée à l'actualité. Seule la pédo-pornographie a un rythme constant.

Pharos ne travaille que sur du contenu public (tweet, site Web), afin de pouvoir le vérifier et le qualifier, donc une insulte raciste sur WhatsApp, par exemple, n'est pas de leur ressort.

Les signalements sont évidemment de qualité variable. Ils peuvent concerner du contenu légal, ou être franchement incompréhensibles. (Les responsables disent qu'ils n'ont pas vu beaucoup de signalements malveillants; les erreurs sont nombreuses, mais de bonne foi.) Par exemple un type avait fait une vidéo très bizarre (« ambiance et bruitages ») sur la disparition du petit Émile, et elle avait été signalée comme « c'est peut-être lui le coupable ». Pharos ne tient pas compte du nombre de signalements pour un contenu donné, chaque signalement compte.

Pharos emploie entre 40 et 50 personnes (selon les fluctuations, le "*burn out*" est un problème, voir plus loin), tous policiers et gendarmes. Il y a un peu de spécialisation : un pôle est spécialisé dans l'ouverture des procédures judiciaires, un autre dans l'expression de la haine en ligne, jugée assez différente pour avoir son propre pôle.

Lors d'un interview à France Inter, la semaine dernière, la présidente du Haut Conseil à l'Égalité, en réponse au journaliste qui lui demandait si elle prônait l'augmentation des effectifs de Pharos pour les nouvelles règles sur la pornographie, avait répondu que ce n'était pas nécessaire car « [le problème] est très facile, il suffit d'utiliser les techniques d'Intelligence Artificielle ». Nos interlocuteurs ont affirmé que Pharos n'utilisait pas du tout d'IA, que peut-être dans le futur, mais se sont montrés plutôt sceptiques, vu la difficulté à qualifier exactement les cas auxquels elles sont confrontés. L'humain reste irremplaçable, à leur avis.

Sur le problème psychologique, puisque les enquêteurs voient passer pas mal d'horreurs (par exemple les vidéos de Daech) : la politique de Pharos est qu'il n'y a pas de jugement de valeur si un employé n'y arrive plus. « Pas de culture viriliste », « Même les grands et forts peuvent avoir des problèmes ». Les employés ont droit à une aide psychologique.

Un autre pôle à Pharos est celui chargé des mesures administratives (ne passant pas par le contrôle d'un juge). Pharos peut demander à un hébergeur le retrait d'un contenu, et si ce n'est pas fait, ordonner aux grands FAI de bloquer (presque toujours par résolveur DNS menteur <<https://www.bortzmeyer.org/resolveur-dns.html>>). C'est une « erreur humaine » qui a été la cause du blocage de Telegram <<https://www.bortzmeyer.org/blocage-telegram-france.html>>.

Et Thesee? Cette plate-forme, plus récente (mars 2022), et que nous avons visitée (comme dit plus haut : un grand "open space" avec des PC, rien d'extraordinaire), est consacrée à la centralisation de certaines plaintes sur l'escroquerie. (Vous pouvez enregistrer votre plainte via Thesee <<https://www.service-public.fr/particuliers/vosdroits/N31138>>.) L'idée est que beaucoup d'escroqueries portent sur des petites sommes et que, si on regarde individuellement les plaintes, elles peuvent sembler peu importantes et être mises en bas de la pile de dossiers à traiter. Or, ces « petites » escroqueries peuvent être l'œuvre de gens qui travaillent en grand et escroquent de nombreuses personnes, donc le montant total peut être important. Mais, sans centralisation et regroupement, on peut ne pas voir qu'il ne s'agit pas de manœuvres isolées. Thesee est donc une plate-forme humaine d'analyse de ces "clusters". Par exemple, si dix escroqués ont porté plainte contre un faux vendeur et que l'IBAN du vendeur est toujours le même, on a trouvé un "cluster". Un logiciel aide à effectuer ces regroupements. (Mais la démonstration de ce logiciel n'a pas fonctionné.) De toute façon, aussi bien pour Thesee que pour Pharos, les enquêteurs ont insisté sur le fait que la technique ne peut pas tout, qu'une enquête est surtout un travail humain.

Thesee compte 32 personnes. Déjà 100 000 plaintes.

Ces escroqueries touchent souvent des personnes vulnérables, et les enquêteurs de Thesee sont souvent confrontés à la détresse des victimes, quand on leur apprend qu'ils ne reverront jamais leurs économies, ou que la jolie femme qui, depuis l'autre bout du monde, leur affirmait qu'elle était tombée amoureuse d'eux et avait juste besoin de 4 000 € pour venir les rejoindre, n'existait pas. Bref, il faut être bienveillant lors des contacts avec les victimes, d'autant plus que l'entourage de celles-ci ne l'est pas toujours (« mais comment as-tu pu être aussi bête? »).

Pendant la visite. De gauche à droite, Jeanne Bouligny, Éric Bothorel, Pierre Beyssac et moi :