Sébastien GAYA

OSINT :
La pluridisciplinarité
au-delà du cyber

Promotion 2024-2025

« L’OSINT c’est avant tout un état d’esprit »
– Jean-Marc MANACH

L’OSINT (Open Source Intelligence), ou Renseignement d’Origine Source Ouverte en français. Depuis quelques années, ce phénomène prend de plus en plus d’ampleur, dans le domaine judiciaire par les enquêtes où le renseignement en fait une partie intégrante, dans le domaine journalistique où les auteurs font leurs recherches pour étoffer leurs sujets du jour, ou encore n’importe qui d’assez curieux pour être tombé sur cette pratique, qui peut s’y adonner par l’intermédiaire de plateformes comme Osintopia qui propose des challenges pour s’exercer, ou alors se saisir d’un sujet d’actualité et mener son enquête pour espérer y débusquer une information qui serait encore dissimulée aux yeux de toutes et tous.

Dans un contexte mondial où on parle de désinformation, de post-vérité, où les évolutions de règles de modération sur certains réseaux facilitent les discours conspirationnistes, l’OSINT devient un élément critique dans la lutte pour la vérité. Par l’expansion du monde cyber et des données qui y sont stockées, on pourrait y lier la croissance de l’OSINT, et ainsi supposer un lien fort entre cette discipline et les connaissances techniques que peuvent avoir des développeurs, des ingénieurs réseau, des administrateurs système, des experts en cybersécurité, etc.

Cependant, derrière cet acronyme se cache une multitude de disciplines et de compétences qui peuvent être mises à contribution pour collecter, analyser et dévoiler des informations. Que l’exercice d’investigation en source ouverte n’est pas tant la démonstration d’as du clavier qui par leurs algorithmes et leurs codes peuvent débusquer n’importe quelle information, mais cache un monde bien plus vaste. Dans cet article, nous allons donc essayer de vérifier que l’OSINT s’étend au-delà de compétences techniques issus du milieu cyber. Pour cela, nous allons explorer cette pluridisciplinarité par trois axes : tout d’abord par une approche des sciences humaines et sociales, suivi du journalisme d’investigation, et enfin la data science et la visualisation.

I. Sciences humaines et sociales : un croisement de contextes, de cultures et de récits

Avant de commencer toute investigation, il est important de poser un cadre dans lequel travailler. Une donnée peut être qualifiée de pertinente uniquement en fonction du contexte dans lequel elle est collectée. Il est donc important de comprendre l’humain dans son environnement. Pour cela, nous allons décliner la thématique des sciences humaines et sociales en trois pistes : la géopolitique et science politique, la piste sociologique et anthropologique, et enfin l’approche psychologique et cognitive.

Géopolitique et science politique : contextualiser l’information

Il arrive, lorsque l’on travaille sur un sujet de conflit tel que la guerre entre l’Ukraine et la Russie, ou le conflit israélo-palestinien, de devoir travailler sur des images satellites ou des

images et vidéos issus des réseaux sociaux. Cependant, afin que les données analysées soient pertinentes, il faut prendre en compte le conflit en lui-même et ses différentes caractéristiques. Que ce soit la position des lignes de front, les alliances entre différents pays ou régions, ou alors les différents mouvements de population, tout ceci doit être pris en compte. Il faut également se renseigner sur la dimension politique de la région étudiée, des potentiels conflits engendrés, du fonctionnement du régime au pouvoir, du niveau de contrôle sur son territoire, et par conséquent des zones à risque.

Prenons l'exemple d'une manifestation qui a eu lieu en août 2020 en Colombie, plus précisément à Corinto dans le sud-ouest du pays. Abelardo Liz, un journaliste indigène a été tué de plusieurs balles alors qu'il couvrait cet événement. Le collectif bénévole Bellingcat, spécialiste en renseignement de données issues de sources ouvertes, avec Cerosetenta, un média indépendant national ont travaillé ensemble sur un ensemble de médias issus de vidéos, certaines provenant de réseaux sociaux, l'une venant de la caméra d'Abelardo lui-même, etc. Ils ont procédé à une analyse du son des balles pour en déterminer la distance et la provenance, mais pour que cette information soit réellement pertinente, il faut également savoir qui se situe où pendant le conflit, que ce soient les manifestants, les policiers ou encore les soldats déployés. C'est dans ce cadre qu'il a été possible de définir que la balle provenait du front des soldats et des policiers.

On peut également prendre l'exemple de DFRLab, une équipe de recherche affiliée au think tank américain Atlantic Council, qui a étudié la présence au Soudan de membres du groupe Wagner, une organisation paramilitaire russe, qui traduit un position stratégique du pays en Afrique notamment pour un accès à des ressources naturelles en plus d'une présence militaire, une motivation à la fois politique et économique.

Enfin, l'investigation de l'Institut australien de stratégie politique autour des camps de rééducation dans la région du Xinjiang en Chine, qui en plus de la démarche technique de la recherche, vient mettre un éclairage sur le régime autoritaire chinois, la construction d'un Etat de surveillance, et les dynamiques de pouvoir. Une clé de lecture par l'étude des régimes politiques et du fonctionnement du gouvernement à travers le pays était nécessaire pour que cette investigation et les informations récoltées puissent être pertinentes.

Sociologie et anthropologie : lire les dynamiques sociales

Après avoir apporté du contexte à une investigation à base de données en source ouverte, il est également important d'identifier les codes et représentations dans le domaine dans lequel on travaille. En effet, en prenant un angle sociologique et anthropologique, les contenus récoltés peuvent être interprétés afin de donner sens à des comportements collectifs, et les rapports sociaux entre les personnes, que ce soit dans des groupes Telegram, des bulles de discussion sur X (ex-Twitter) ou des forums spécialisés.

Prenons l'exemple de Moonshot, une start-up qui combat et analyse les mouvements extrémistes, en travaillant sur différents domaines tels que le terrorisme, le trafic d'êtres humains ou les violences de genre. Ils ont publié en 2021 un guide sur les symboles et la terminologie autour de la communauté des incels, ces "célibataires involontaires" qui mettent la faute de leur incapacité à avoir des relations avec des femmes, qu'elles soient simplement amoureuses ou sexuelles, sur le dos de ces-dites femmes qui seraient des manipulatrices et

sur la société en général. Ce guide dégage de nombreux symboles autour de ce mouvement, comme les archétypes Chad, représentant l'homme alpha qui serait l'homme idéal pour les femmes, Stacy, la femme au sommet de la hiérarchie sociale, la cible n°1 des incels, et Becky, représentant les femmes moyennes qui seraient redevables d'avoir des rapports avec eux. En listant certaines figures de ce mouvement, comme Elliot Rodger, l'auteur de la tuerie de masse d'Isla Vista, on peut découvrir certains discours et codes comme la phrase "Go ER", en référence aux initiales d'Elliot, signifiant un appel à commettre un meurtre, motivé par la haine des femmes. Sur les réseaux sociaux, forums et autres moyens d'échange, le guide mentionne « une terminologie contrôlée et établie, [...], renforçant une dynamique du "nous contre eux" ».

On y retrouve également le concept de « blue pill » et « red pill », en référence au film Matrix, pour distinguer dans un cas les hommes continuant leur vie ignorant la réalité du monde, et dans l'autre cas ceux ayant pris conscience de la vraie nature de la société, dans laquelle les femmes ne seraient pas opprimées mais où leur vrai désir est d'entretenir les rapports traditionnels, et par cette connaissance, leur donnent la possibilité de manipuler les femmes à leur avantage. Enfin, le concept plus extrême de « black pill », allant au-delà de la pilule rouge, avec une vision plus pessimiste des rapports humains, qui tend à la radicalisation sur la haine des femmes et potentiellement des passages à l'acte.

Psychologie et sciences cognitives : anticiper les intentions

Enfin, la clé de lecture par l'aspect psychologique et cognitif peut être primordiale dans le cadre d'investigations OSINT. Dans un monde de post-vérité, où nous sommes envahis de fake news et de propagande, il est important de savoir par quels biais ces discours se propagent, et quels sont les mécanismes d'influence. Comment un tel message peut être perçu, comment peut-il être diffusé aussi rapidement, et quelle est la cible principale.

On peut prendre l'exemple de la campagne de vaccination lors de la pandémie du Covid-19, qui a eu droit à son florilège de détracteurs, plus communément appelés les "antivax". Des réseaux sociaux comme X ont permis la propagation rapide de fausses informations. Au-delà des faits en eux-mêmes, on peut identifier des biais cognitifs ayant facilité le partage de ces informations. François Sellal, Geoffroy Hautecloque et Guido Ahle en parlent dans la revue de Neuropsychologie, tel que le biais d'évaluation des probabilités. La disponibilité d'un vaccin qui a été plus rapide que prévu a rapidement posé question, « en arguant qu'on n'avait pas pris le temps de les évaluer et surtout, en signalant des effets indésirables graves ». De nombreux décès suivant les prémices de la vaccination ont encouragé à pointer du doigt ce remède, alors que les autopsies mettaient en lumière des problèmes de santé chez les personnes. Le cas de Tiffany Dover est aussi un exemple marquant : cette infirmière dans un hôpital du Tennessee est devenue une icône suite à son évanouissement devant des caméras de télévision après une injection du vaccin Pfizer. Les conspirationnistes ont tout de suite propagé l'idée qu'elle était morte suite à celle-ci, alors que Tiffany allait très bien. Les auteurs parlent également de l'effet Dunning-Kremer, selon lequel on sous-estime les compétences qu'on détient, mais on surestime celles que nous manquons. Par cet effet, de nombreuses figures politiques s'exprimaient dans les médias sur ce sujet dont ils ne sont clairement pas experts, à tous bords politiques, que ce soit Jean-Luc Mélenchon à gauche, qui déclare « ne pas pouvoir être un cobaye », ou Nicolas Dupont-Aignan à droite, qui déclare que les laboratoires pharmaceutiques nous privent du «

droit d'étudier des médicaments qui dans le monde entier soignent », en citant par exemple l'ivermectine, dont des études prouvent son inefficacité dans la lutte contre le Covid.

Tout ceci contribue à des propagations rapides et massives dans des bulles ancrées dans une croyance où le biais de confirmation facilite l'adhésion à ces informations et par conséquent les idées qui en découlent. On a pris ici l'exemple du Covid, mais ce sont des biais qui se retrouvent ailleurs, comme le conflit israélo-palestinien, qui est à la fois militaire et informationnel, avec la propagation de fausses informations, également relayées par des responsables politiques sans vérification préalable. Le journal du Monde a, à ce sujet, une page spéciale sur son site avec la liste de fausses images et vidéos depuis l'attentat perprétré par le Hamas le 7 octobre 2023, afin de désamorcer à leur niveau la propagande, comme par exemple une fausse vidéo présentée comme un reportage de la BBC se fondant sur une enquête de Bellingcat, une affirmation démentie par les 2 acteurs de l'information.

Ainsi, on peut en déduire que les sciences humaines et sociales font partie intégrante de l'investigation OSINT, en apportant des clés de lecture et un cadre culturel, politique et social, des éléments primordiaux pour définir la pertinence des données qui seront récoltées.

II. Journalisme d'investigation : l'art de la recherche, de la vérification et de l'éthique

Une fois le cadre posé, il est temps de passer à la récolte de données et d'informations. Les méthodes employées dans le cadre d'investigations OSINT se rapprochent beaucoup de celles qu'on trouve dans le monde journalistique : recherches documentaires, vérification de faits ou encore l'éthique lors de la publication, où les informations doivent être vérifiables par le public ciblé. Ce sont ces points communs que nous allons aborder dans cette partie. Tout d'abord par les différentes techniques de recherche et d'enquête, suivi de la vérification et du recoupement de l'information récoltée, et enfin l'aspect éthique et l'enjeu de responsabilité quant à la publication des résultats.

Techniques de recherche et d'enquête en source ouverte

Aujourd'hui, les journalistes d'investigation sont de plus en plus à même d'utiliser des sources ouvertes, Internet facilitant l'accès à une gigantesque quantité de données, il n'est plus forcément utile d'aller sur le terrain. Tout comme les OSINTers, ils usent de bases de données publiques, des réseaux sociaux, des médias, archives, forums, pour trouver les informations dont ils ont besoin.

Par exemple, en janvier 2023, les journalistes du Monde ont mis en évidence comment des appareils de surveillance de Frontex, l'agence européenne des garde-frontières et garde-côtes, sont utilisés par les garde-côtes de Libye pour intercepter des migrants. En effet, des images de surveillance ont été publiées sur Facebook entre 2018 et 2022 par les autorités libyennes. Les journalistes ont ensuite recoupé avec des données en source

ouverte ainsi que des informations provenant d'ONG présentes en Méditerranée, ce qui a permis de dévoiler que certaines images venaient de l'agence européenne, alors que la Libye ne devrait pas y avoir accès.

On peut également prendre l'exemple de CNN qui a mis en lumière le bombardement russe sur la maternité de Marioupol en mars 2022. Du côté de la Russie, 2 versions, l'une selon laquelle aucun bombardement n'a eu lieu d'après le ministère de la Défense, et l'autre selon laquelle celui-ci serait justifié par la présence de forces armées ukrainiennes. Or la chaîne américaine a rendu publique son enquête sous forme de chronologie pour retracer les différents événements autour de cette attaque pour démentir les propos du côté russe. On y retrouve des publications sur des réseaux sociaux, des médias de civils quittant le bâtiment, des images satellites, ou encore des analyses d'experts comme Chris Cobb-Smith dans le domaine de l'armement sur la nature de la munition utilisée.

Ce sont des techniques de recherche qu'on retrouve également chez Evidence Lab, l'équipe d'investigation d'Amnesty International, spécialisée dans la recherche en source ouverte, ou encore le collectif Bellingcat, mentionné dans la partie précédente.

En juillet 2024, Bellingcat travaille sur l'identification d'un missile ayant frappé un hôpital pour enfants à Kiev, capitale de l'Ukraine. En effet, des sources telles que le Ministère russe des Affaires étrangères prétendaient que le missile était américain, et qu'il était parti d'un site ukrainien. Or, les enquêteurs du collectif, accompagnés d'un expert en missiles et armes nucléaires, le Dr. Jeffrey Lewis, ont pu, grâce à une vidéo trouvée sur les réseaux sociaux, réaliser un modèle 3D du missile avec des caractéristiques précises permettant de l'identifier comme un missile exclusivement russe, le Kh-101. De plus, les débris avaient des points communs avec d'autres missiles Kh-101 utilisés dans des attaques passées.

Vérification et recoupement de l'information

En plus de récolter l'information, il est aussi indispensable de la vérifier et de la recouper avec d'autres informations afin d'en assurer la fiabilité. Savoir qui a publié, avec quelle intention, dans quel contexte et avec quelle logique de diffusion.

Un premier exemple est une vidéo datant de mars 2022, du président ukrainien Volodymyr Zelensky dans laquelle il exprime sa reddition à la Russie. Il s'agissait bien entendu d'une deepfake, une vidéo générée par intelligence artificielle. Le montage pouvant apparaître un peu grossier pour les connaisseurs, un public non averti pourrait y croire. En analysant la vidéo, on peut distinguer une mauvaise synchronisation entre les lèvres et le discours, en croisant avec de précédents discours, le ton était bien trop différent, puis une réaction rapide du gouvernement a permis de dénoncer la désinformation. L'amélioration continue des IA tend malheureusement à la diffusion de plus en plus accessible de ces deepfakes, ce qui rend les enquêtes de fact-checking d'autant plus nécessaires.

Un autre exemple est l'investigation menée par Bellingcat pour vérifier la véracité d'attaques chimiques dans la ville syrienne de Douma en avril 2018. En partant de rapports venant d'ONG telles que le Réseau syrien pour les Droits humains ou le Centre de documentation des violations, ainsi que d'analyses médicaux, le collectif a ensuite analysé plusieurs vidéos, les deux premières montrant un nombre important de victimes dans un même bâtiment

ayant des symptômes qui recoupent les analyses des médecins récoltées par les ONG. Trois autres vidéos ont permis de noter des éléments caractéristiques visibles sur les images, afin de localiser le bâtiment dans lequel toutes ces vidéos ont été tournées. Enfin une sixième vidéo met en lumière au même endroit une bouteille de gaz comprimé qui a traversé le toit. Une analyse de la bouteille a également permis de faire le lien avec d'autres attaques chimiques comme en février 2018 à Saraqeb, une autre ville syrienne. Toute cette investigation est relatée sur leur site.

Éthique de la publication et responsabilité

Enfin, pour terminer les points communs avec le journalisme d'investigation, il faut également aborder la dimension éthique de l'OSINT. Il y a une responsabilité à publier les rapports des enquêtes en ligne, qui plus est quand on traite des sujets sensibles comme les conflits armés ou de trafics d'êtres humains. Publier une vidéo avec du contenu violent, révéler l'identité d'un auteur présumé ou des lieux d'importance stratégique peut avoir des répercussions importantes, et un investigateur OSINT ne peut arbitrer sur ces sujets par de simples compétences techniques.

Colette Cujipers, une maîtresse de conférences en droit et technologie, à l'Université de Tilburg, dit « le fait que les données soient librement accessibles ne signifie pas qu'elles peuvent être traitées sans tenir compte des exigences légales et éthiques. En d'autres termes, le simple fait que des données soient accessibles au public n'implique pas l'absence de restrictions à leur recherche ».

Bellingcat parle également sur son blog de Jennefer Harper, une chercheuse indépendante impliquée dans des investigations autour de l'invasion de l'Ukraine par la Russie en 2022. A partir de sources ouvertes, elle a commencé à documenter chaque aspect de la guerre, et a travaillé sur la construction d'une base de données remplie des informations personnelles des soldats russes impliqués dans l'invasion et la perpétration de crimes de guerre. Ce travail a été rendu d'autant plus nécessaire suite au massacre de Boutcha, mais celle-ci questionnait encore l'éthique de sa manœuvre, car il est possible que nombre de soldats n'aient pas eu le choix de rejoindre la guerre, forcés par un gouvernement autoritaire.

En somme, l'OSINT présente nombre de similitudes avec le journalisme d'investigation, que ce soit sur le plan méthodologique ou éthique. Que ce soit dans la récolte d'informations, leur recoupement et vérification, la transparence des sources et la responsabilité de leur publication, notamment celles comportant des points sensibles, ce sont des principes hérités de cette discipline et adaptés à l'utilisation des sources ouvertes disponibles en ligne.

III. Data science et visualisation : donner du sens à la masse d'information

Dans cette partie, nous allons nous pencher sur comment structurer et interpréter les données. Comme précisé en introduction, la quantité de données n'est plus vraiment un problème, grâce à la multiplicité des sources qu'apporte Internet. Mais cette saturation

d’informations, notamment dans le cadre d’enquêtes pouvant s’étaler sur des années, doit pouvoir rester intelligible et exploitable pour les acteurs impliqués, ce qui peut demander des compétences issus de la visualisation et de la data science.

Structuration et traitement des données massives

Pour mener une investigation avec une grande masse de données, il faut être capable de structurer et traiter celles-ci. Pour cela, il faut des compétences semblables à la documentation scientifique.

Par exemple, depuis 2014, l’ACLED, acronyme pour Armed Conflict Location & Event Data, une organisation à but non lucratif, travaille sur la collecte, l’analyse et la cartographie des différents conflits armés dans le monde. Un travail conséquent a été fait sur la normalisation et la structure des données pour définir les événements violents. La date, la localisation géographique, les acteurs impliqués, le type d’événement ou encore le nombre de victimes sont nombre de caractéristiques standardisés permettant l’interopérabilité des données. Pour rejoindre la partie sur la vérification des informations, des métadonnées permettent également de sourcer chaque attribut des événements. L’organisation permet également d’exporter les différentes données dans des formats tels que CSV ou Microsoft Excel afin de permettre leur traitement dans d’autres systèmes d’information. On se rapproche ici de compétences techniques, mais il est important de noter que la structuration de conflits est un travail demandant bien plus que du code, c’est pourquoi l’ACLED possède en son sein, des experts venant de différents univers, des analystes, des data scientists, des coordinateurs de méthodologie qui s’occupent d’évaluer la qualité, la rigueur et l’exactitude des données, de multiples chercheurs issus de milieux comme l’anthropologie, le sciences politiques ou la linguistique.

Un rapport de l’organisation portant sur l’opération Iron Wall, menée par Israël en Cisjordanie depuis le début de l’année 2025, montre à travers le jeu de données récoltées, l’évolution d’événements violents impliquant l’armée de l’Etat hébreu. La standardisation des événements et la collecte massive permet de mettre en évidence l’explosion de la quantité d’attaques sur des civils palestiniens depuis le lancement de l’opération, alors que la majorité des événements depuis l’attentat du 7 octobre 2023 n’étaient que des conflits armés.

C’est pourquoi la structuration de données massives est importante afin de pouvoir les manipuler à posteriori pour en dégager ensuite des visuels.

Visualisation : rendre intelligible l’invisible

En effet, pour rendre intelligible ces différents sujets à des lectrices et lecteurs qui ne sont pas forcément des experts de la géopolitique ou des professionnels de l’analyse de données, il est nécessaire de synthétiser la masse d’informations sous forme de graphiques, de cartes, de chronologies interactives. Nous avons vu précédemment l’exemple de l’enquête de CNN sur le bombardement de l’hôpital pour enfants de Marioupol. La chaîne télévisée a rendu les résultats de l’enquête disponibles en ligne sous forme de chronologie relatant les différents faits, avec des cartes, des modélisations 3D, etc. Cela permet

également aux personnes en charge de l'investigation de mettre en valeur des anomalies ou des patterns plus facilement que de simplement analyser ces données à leur format brut.

Un exemple de data visualisation est une carte de chaleur produite par Airwars, une organisation non gouvernementale londonienne. Celle-ci l'a conçue dans le cadre de la bataille de Kharkiv, une ville ukrainienne, ayant eu lieu de février à mai 2022 au début de l'invasion russe. Cette carte permet de mettre en valeur les quartiers ayant été les plus touchés par les conflits, mettant en valeur les endroits avec le plus d'incidents impliquant des dommages civils.

Il existe également DeepStateMap, une carte interactive encore plus complète regroupant une grande quantité de données liées au conflit entre la Russie et l'Ukraine. On y retrouve des positions de brigades russes, de quartiers généraux, la position des lignes de défense, la direction des attaques, etc. L'ONG à l'origine de cette carte et s'occupant de la maintenance, Deep State UA, a été fondée par deux ukrainiens, l'un diplômé en droit et l'autre en marketing. Partant au départ de messages sur Telegram sur des sujets internationaux tels que la pandémie du Covid-19 ou le mouvement #BlackLivesMatter, ils ont dérivé sur la création d'une carte interactive dans le cadre de l'offensive des talibans en 2021 afin de mettre à jour la ligne de front tout au long du conflit. En effet, cela rend l'information plus intelligible d'une part avec des éléments visuels clairs, et plus accessible en utilisant un site web à la place de canaux Telegram. En parallèle, l'ONG prévient que certaines informations ne sont publiées qu'après un délai, étant un récapitulatif de données en sources ouvertes, il est également important de ne pas mettre en danger les soldats ukrainiens, et donc de ne pas forcément dévoiler leurs positions, rejoignant l'axe éthique et responsable vu dans la partie sur le journalisme d'investigation.

Collaboration humain-machine : la complémentarité des expertises

Le cas de DeepStateMap ou de l'ACLED peut mettre en évidence l'utilisation de compétences techniques, notamment les traitements algorithmiques et la mise en œuvre des data visualisations des données récoltées. Cependant, cela ne remplace pas l'analyse humaine. En effet, c'est lui qui va orienter les recherches, interpréter les résultats et poser les bonnes questions autour d'un sujet. La machine va permettre de mettre en évidence certaines tendances, mais pour valider ou au contraire infirmer une hypothèse, l'analyse critique est nécessaire.

Le collectif Bellingcat organise des hackatons pour réunir des personnes issus de milieux hétérogènes, que ce soit de la technique, du journalisme, de la recherche et plus encore. En juin 2023, ils publient à propos de leur 3e édition, centrée sur l'accessibilité des outils pour les investigations. En effet, quasiment un an plus tôt, ils publient les résultats de sondages sur le profil des personnes faisant de la recherche en source ouverte. Plus de 500 personnes ont répondu, dont 35 % indiquent faire des recherches en ligne occasionnellement mais sentent un besoin de progresser. Et sur un aspect plus technique, une question interroge sur l'utilisation d'outils provenant de GitHub, un site hébergeant du code source, comme pour des scripts automatiques. Sur celle-ci, 30 % ne connaissent pas GitHub, et 21 % disent connaître l'existence de ces outils mais ne savent pas les utiliser. C'est là qu'on distingue par la multiplicité des disciplines qui se croisent dans l'OSINT, certains acteurs n'ont pas de profil technique, mais restent néanmoins des éléments

importants. Le hackathon faisait alors collaborer ces différents corps de métiers pour créer des outils accessibles à des profils moins techniques. On peut citer le projet Sentinel qui a gagné le premier prix, une interface simple pour effectuer des recherches dans les documents du site du Armed Forces Tribunal en Inde, un tribunal militaire.

Un autre exemple de cette collaboration entre l'humain et la machine, est l'outil PulseSatellite. Issu d'UN Global Pulse, le laboratoire d'innovation rattaché au secrétaire général des Nations Unis. Présenté dans le contexte de catastrophes naturelles qui requièrent des réponses humanitaires rapidement, les chercheurs se sont penchés sur l'analyse d'images satellite en temps réel par intelligence artificielle. Celui-ci est présenté comme utilisant une boucle IA-humain, où des interventions manuelles venant d'analystes permettent d'affiner les modèles.

On se rapporte à un principe assez courant dans le monde de l'intelligence artificielle nommé « human in the loop » ou humain dans la boucle. Comme l'intitulé l'indique, il s'agit d'introduire une intervention humaine dans un processus automatisé, un modèle de coopération qui croise des compétences techniques articulées avec des capacités d'analyse, de culture et d'éthique. Car au-delà du principe de l'investigation OSINT qui produit de la donnée, il s'agit également de produire du sens.

Ainsi, on aura pu voir que la data visualisation est une étape importante pour structurer une masse importante de données, ainsi que les rendre intelligible pour les enquêteurs d'une part afin de mettre en évidence des hypothèses, ou parfois détecter des incohérences, et d'autre part pour le grand public afin que les informations communiqués soient compréhensibles, y compris pour des personnes non expertes dans les domaines concernés. Enfin, nous avons vu que les compétences techniques peuvent être mises à contribution dans le cadre des investigations ou dans la création d'outils utilisables, mais toujours conjointement à des capacités issues de multiples domaines, car la technologie n'est qu'un moyen à la disposition des enquêteurs et enquêtrices.

Conclusion

L'OSINT, cette discipline qui pouvait être réservée au renseignement militaire ou cyber, puis rendue populaire par l'accessibilité des données, s'impose aujourd'hui comme une stratégie d'investigation à part entière, en mouvement constant. Au-delà de l'accroissement de la quantité des données disponibles, sa popularité peut également pointer du doigt la diversification des compétences nécessaires pour mener à bien des enquêtes, en collectant, analysant et interprétant avec pertinence et fiabilité, et cela avec des disciplines allant bien au-delà de l'univers cyber. C'est ce qui a permis de guider le fil de cet article, sortir de la façade de l'OSINT, pleine de techniques et d'algorithmes, en passant par les sciences humaines et sociales, le journalisme d'investigation, et la data science.

Dans un premier temps, les sciences humaines et sociales sont indispensables pour contextualiser les traces numériques. Pour évaluer la pertinence des données qui seront récoltées, il faut comprendre les dynamiques culturelles, sociales et politiques, pour les situer dans un récit, un territoire ou une temporalité. Identifier une image, un discours, un

compte sur un réseau social est une chose, mais encore faut-il en décrypter les codes culturels, le poids symbolique ou le sens qui en découle. La sociologie, l'anthropologie, la science politique ou les sciences cognitives sont autant de disciplines dont la contribution permet de transformer des données brutes en indices intelligibles. Ce sont des compétences cruciales pour désamorcer le risque de biais culturels et être préparé pour la production d'analyses robustes, nuancées et contextualisées.

Ensuite, nous avons observé les similitudes entre l'OSINT et le journalisme d'investigation, qui apporte une méthodologie éprouvée de techniques diverses de recherche, de recoupement et de vérification des informations et enfin d'éthique et de responsabilité éditoriale. Diverses pratiques comme la géolocalisation ou la datation croisée, issues de cette discipline, renforcent la rigueur des enquêtes en source ouverte. Encore plus important, la culture de la transparence des sources, de la documentation des rapports et de l'éthique de la publication, est fondamentale dans un monde de post-vérité, où la viralité, les fausses informations et les manipulations peuvent avoir des effets importants sur la géopolitique ô combien fragile aujourd'hui. L'enquête OSINT n'est pas seulement une opération technique mais également un acte de publication responsable, en croisant des expertises, de l'intégrité et des stratégies d'influence.

Pour terminer, la data science et la visualisation permettent de donner une forme à la masse d'informations pour la rendre lisible et compréhensible. Loin d'être un domaine réservé aux spécialistes de la programmation, ce prisme mobilise également des compétences issues de la cartographie, parfois de la facilitation graphique et de la scénarisation des données. Mais il est vrai que pour accélérer ce travail, des outils de traitement automatique ou des scripts peuvent être mis à contribution lors des investigations. Cependant, il faut garder en mémoire que leur pertinence n'est avérée que lorsque lesdits outils s'inscrivent dans une coopération entre la machine et l'humain, qui est capable d'interpréter, corriger voire enrichir les modèles produits. La technique ne remplace pas, elle complète.

L'avènement des modèles d'intelligence artificielle renforcent encore plus cette dynamique, d'un côté les IA génératives qui demanderont aux enquêteurs de redoubler d'efforts dans la bataille contre les fausses informations, mais de l'autre, les IA qui peuvent servir de véritables assistants aux analystes dans le cadre de reconnaissance d'images, de transcription et traduction de discours pour les enquêteurs ne maîtrisant pas certaines langues, ou plus largement en première ligne de détection d'anomalies dans les flux d'informations. C'est une opportunité pour traiter des volumes de données inédits qui ne cesseront d'augmenter avec le temps, et réduire le risque de passer à côté d'éléments invisibles pour l'œil humain. Il faut cependant garder en tête l'encadrement de ces outils, ou la transparence, la validation humaine et l'inclusion des contextes culturels et politiques sont des piliers nécessaires à leur bon usage.

Nous sommes donc face à une pratique véritablement pluridisciplinaire, même au-delà du cyber, en associant des experts techniques à des chercheurs en sciences sociales, sciences politiques, journalistes d'investigation, analystes du renseignement, cartographes, et bien plus encore. Une conjugaison entre la puissance de traitement des machines avec la finesse de l'analyse et de l'interprétation de l'être humain. Une discipline qui, plus loin que l'extraction de données, construit du sens, éclaire les discours et événements, aussi tragiques soient-ils, partout dans le monde, pour contribuer à une compréhension

stratégique de la réalité, à l'heure où la prolifération de fausses informations et la manipulation viennent entraver le bon fonctionnement de l'espace public et menacent la confiance du grand public, l'OSINT serait alors la garantie d'un renseignement ouvert par tous et pour tous.

Bibliographie

Le Deuff, O. (2021). L'Open Source Intelligence (OSINT) et les enquêtes numériques : la panoplie du maître d'armes. I2D - Information, données & documents, 1(1), 92-95. https://doi-org.ezproxy.u-bordeaux-montaigne.fr/10.3917/i2d.211.0092

Gibson, Steyyn D. "open source intelligence" in Dover, R., Goodman, M. S., & Hillebrand, C. (2013). Routledge Companion to Intelligence Studies. Routledge. p. 123-13.

Roumanos, R. et Le Deuff, O. (2022). Les défis de l'investigation journalistique en sources ouvertes. Multitudes, 89(4), 67-74. https://doi-org.ezproxy.u-bordeaux-montaigne.fr/10.3917/mult.089.0067

Mielcarek, R. (2022). Journalisme : l'enquête en sources ouvertes, entre mirage et opportunité. Hérodote, 186(3), 43-55. https://doi-org.ezproxy.u-bordeaux-montaigne.fr/10.3917/her.186.0043

Lewis, M. (2019). Don't F**k with Cats: Hunting an Internet Killer [série documentaire]. Netflix.

Higgins, E. (2021). We Are Bellingcat: An Intelligence Agency for the People. Bloomsbury.

Sellal, F., Hautecloque, G. et Ahle, G. (2021). Les aléas de la vaccination contre la Covid-19 à la lumière des biais cognitifs. Revue de neuropsychologie, 13(2), 129-132. https://www.jle.com/10.1684/nrp.2021.0671

Cuijpers, C. (2013). Legal aspects of open source intelligence – Results of the VIRTUOSO project. Computer Law & Security Review, 29(6), 642-653. https://doi.org/10.1016/j.clsr.2013.09.002

Vuarin, L. et Steyer, V. (2025). La résilience organisationnelle face à l'intelligence artificielle : Examen critique du concept de human in the loop en action. Innovations, 76(1), 239-276. https://doi.org/10.3917/inno.pr2.0183

Logar, T., Bullock, J., Nemni, E., Bromley, L., Quinn, J. A., & Luengo-Oroz, M. (2020). PulseSatellite: A tool using human-AI feedback loops for satellite image analysis in humanitarian contexts. arXiv. https://doi.org/10.48550/ARXIV.2001.10685