Global Network
on Extremism & Technology

Recherche relative aux contenus à caractère extrémiste sur les médias sociaux : enjeux et opportunités de la protection des données et de l'éthique de la recherche

Manjana Sold, Julian Junk

Le GNET est un projet spécial du Centre international d'étude de la radicalisation du King's College, à Londres.

Les auteurs du présent rapport sont Manjana Sold et Julian Junk.

Le Global Network on Extremism and Technology (Réseau mondial sur l'extrémisme et la technologie – GNET) est une initiative de recherche universitaire bénéficiant du soutien du Forum mondial de l'Internet contre le terrorisme (GIFCT), une initiative indépendante mais financée par le secteur qui vise à mieux comprendre et lutter contre l'utilisation des technologies par les groupes terroristes. Le GNET est formé et dirigé par le Centre international d'étude de la radicalisation (ICSR), un centre de recherche universitaire basé dans les locaux du Département d'étude des guerres du King's College, à Londres. Les opinions et conclusions exprimées dans ce document sont celles des auteurs et ne doivent en aucun cas être interprétées comme représentant les opinions et conclusions, expresses ou implicites, du GIFCT, du GNET ou de l'ICSR.

COORDONNÉES

Pour toute question, demande d'information et demande de copies supplémentaires du présent rapport, contacter :

ICSR
King's College London
Strand
Londres WC2R 2LS
Royaume-Uni

T. **+44 20 7848 2098**
E. **mail@gnet-research.org**

Twitter : **@GNET_research**

Ce rapport peut, comme toutes les autres publications du GNET, être téléchargé gratuitement à partir du site Internet du GNET : www.gnet-research.org.

Résumé exécutif

Les rapports entre terrorisme et technologie sont plus pertinents que jamais, tant sur le plan social que politique. La quasi-totalité des processus de mobilisation et de radicalisation, et la totalité des attentats violents, qu'ils soient menés à terme ou empêchés, comportent un volet virtuel. Les chercheurs, notamment ceux qui travaillent avec le GNET, sont directement confrontés au défi que représente l'analyse de ces processus dans des environnements virtuels empiriques difficiles. Même en tenant compte de l'évolution constante des plateformes et des politiques, ainsi que de la migration vers des espaces de plus en plus fermés et chiffrés, les données empiriques abondent. Cette abondance suscite ses propres défis et crée ses propres possibilités[1], mais ce qu'un chercheur s'intéressant aux contenus à caractère extrémiste en ligne a la possibilité et le droit de faire connaît aussi des limites strictes et des zones d'ombre, ce qui amène à réfléchir aux questions d'ordre éthique et relatives à la protection des données. Les débats sur ces questions ont pris beaucoup d'ampleur ces dernières années, et sont particulièrement animés au sein des consortiums de recherche internationaux.

Tout en faisant une synthèse de l'état d'avancement de ces débats sur l'éthique et la confidentialité, le présent rapport du GNET met en exergue les limites pour la recherche et les possibilités qui s'offrent à elle, et émet, à destination des chercheurs et des sociétés technologiques, des recommandations à ce propos. Les auteurs procèdent pour cela en trois étapes : ils résument dans un premier temps certaines des principales questions éthiques qu'un chercheur travaillant dans ce domaine doit garder à l'esprit ; ils fournissent ensuite un aperçu des principes majeurs de la protection des données qu'il convient de respecter et mettent en lumière les possibilités qui s'offrent aux chercheurs et les talents d'équilibriste dont ceux-ci devront faire preuve ; enfin, ils traitent des interactions entre les chercheurs, les sources de données et les politiques des différentes plateformes, tout en résumant certaines recommandations essentielles pour les chercheurs, les sociétés technologiques et les législateurs. Les principaux points abordés ici sont les suivants : tout d'abord, la multiplication des points d'accès et des bases de données multiplateformes permettrait d'élargir et de dynamiser le champ de recherche ; ensuite, une collaboration internationale en matière de recherche dans le domaine de l'analyse des contenus à caractère extrémiste en ligne, qui fait cruellement défaut aujourd'hui, tirerait parti d'une plus grande harmonisation internationale et d'un rapprochement des règles de protection des données ; troisièmement, les régimes de protection des données ne devraient pas être perçus comme des inconvénients mais plutôt comme des éléments qui facilitent la recherche en

1 Abdullah Alrhmoun, Shiraz Maher, Charlie Winter, *Décrypter la haine : emploi de l'analyse de texte expérimentale aux fins de classification des contenus à caractère terroriste*, ICSR King's College, Londres (2020).

posant des limites claires à ce qu'il est possible et pas possible de faire ; et, enfin, ce champ empirique dynamique nécessite la mise en place de mécanismes d'échanges réguliers entre les sociétés technologiques, les chercheurs et les législateurs afin d'adapter les politiques, les habitudes et les cadres juridiques d'une façon qui prenne pleinement en considération la pertinence sociale et politique du rapport entre extrémisme et technologie.

Table des matières

1 Introduction

L'espace numérique a joué un rôle central dans les processus de radicalisation de nombreux auteurs d'attentats[1] : des extrémistes tels qu'Anis Amri (Berlin, Allemagne), Brenton Tarrant (Christchurch, Nouvelle-Zélande) et Stephan Balliet (Halle, Allemagne) ont profité des médias sociaux non seulement pour recueillir et diffuser des informations, étoffer leur réseau et s'organiser, mais aussi pour échanger des idées avec d'autres personnes partageant leur point de vue voire, parfois, diffuser en direct leur attentat auprès de milliers de spectateurs. Ce sont ces communications d'acteurs radicalisés et extrémistes qui nous permettent d'en apprendre davantage sur les processus de radicalisation qui se déroulent sur la toile. Le contenu et sa présentation, de même que la façon dont ces acteurs communiquent, revêtent une importance capitale à cet égard, et peuvent servir de repères à partir desquels élaborer les mesures de prévention et de démobilisation les plus adaptées.

Les données tirées des réseaux sociaux ont, naturellement, acquis une importance croissante dans le cadre de ce sujet d'étude[2]. De nombreuses publications scientifiques fondées sur ce type de données, tirées de Facebook[3], Twitter[4], YouTube[5] ou encore Instagram[6], illustrent ces propos. Un ensemble de données très fourni peut aujourd'hui être consulté et utilisé pour développer et tester des hypothèses[7]. Ces opportunités s'accompagnent toutefois de certaines limites et écueils, liés à d'éventuelles questions éthiques et de protection des données et qui constituent certes des défis pour les chercheurs, mais leur offrent aussi de nombreuses possibilités. S'il est essentiel de faire preuve de transparence et de suivre la ligne de conduite visant à « maximiser les avantages tout en réduisant au

---

1 Un grand merci à Sebastian Golla pour ses commentaires sur les versions antérieures de ce rapport et pour les conseils juridiques pertinents qu'il nous prodigue depuis plusieurs années dans le cadre de nos travaux de recherche. Nous remercions également Clara-Auguste Süß pour ses commentaires, ainsi que Leo Bauer et Klara Sinha pour l'aide apportée dans la finalisation de ce rapport.

2 Sebastian J. Golla, Henning Hofmann et Matthias Bäcker, « Connecting the Dots: Sozialwissenschaftliche Forschung in Sozialen Online-Medien im Lichte von DS-GVO und BDSG-neu », *Datenschutz und Datensicherheit – DuD* 42, n° 2 (2018) : 89, http://link.springer.com/10.1007/s11623-018-0900-x ; Manjana Sold, Hande Abay Gaspar et Julian Junk, Designing Research on Radicalisation using Social Media Content: Data Protection Regulations as Challenges and Opportunities, 2020.

3 Agata Błachnio, Aneta Przepiórka et Patrycja Rudnicka, « Psychological Determinants of Using Facebook: A Research Review », *International Journal of Human-Computer Interaction* 29 (2013), https://doi.org/10.1080/10447318.2013.780868 ; Ralf Caers *et al.*, « Facebook: A Literature Review », *New Media & Society* 15 (2013), https://doi.org/10.1177/1461444813488061 ; Stefania Manca et Maria Ranieri, « Is It a Tool Suitable for Learning? A Critical Review of the Literature on Facebook as a Technology Enhanced Learning Environment », *Journal of Computer Assisted Learning* 29 (2013), https://doi.org/10.1111/jcal.12007 ; Ashwini Nadkarni et Stefan G. Hofmann, « Why do People Use Facebook? », *Personality and Individual Differences* 52, n° 3 (2012), https://doi.org/10.1016/j.paid.2011.11.007 ; Robert E. Wilson, Samuel D. Gosling et Lindsay T. Graham, « A Review of Facebook Research in the Social Sciences », *Perspectives on Psychological Science* 7 (2012), https://doi.org/10.1177/1745691612442904.

4 Jytte Klausen, « Tweeting the Jihad: Social Media Networks of Western Foreign Fighters in Syria and Iraq », *Studies in Conflict & Terrorism* 38, n° 1 (2015) ; Amandeep Dhir, Khalid Buragga et Abeer A. Boreqqah, « Tweeters on Campus: Twitter a Learning Tool in Classroom? », *Journal of Universal Computer Science* 19 (2013) ; Shirley Ann Williams, Melissa Terras et Claire Warwick, « What Do People Study When They Study Twitter? Classifying Twitter Related Academic Papers », *Journal of Documentation* 69 (2013).

5 Chareen Snelson, « YouTube Across the Disciplines: A Review of the Literature », *MERLOT Journal of Online Learning and Teaching Journal of Qualitative Methods* 7, n° 14 (2011), http://jolt.merlot.org/vol7no1/snelson_0311.htm ; Raphael Ottoni *et al.*, « Analyzing Right-wing YouTube Channels: Hate, Violence and Discrimination », (2018) ; Kostantinos Papadamou *et al.*, « Understanding the Incel Community on YouTube », (2020).

6 Tim Highfield et Tama Leaver, « A Methodology for Mapping Instagram Hashtags », *First Monday* 20, n° 1 (2015) ; Asunción Bernardez-Rodal, Paula Requeijo Rey et Yanna G. Franco, « Radical right parties and anti-feminist speech on Instagram: Vox and the 2019 Spanish general election », *Party Politics* (2020) ; Lena Frischlich, « #Dark inspiration: Eudaimonic entertainment in extremist Instagram posts », *New Media & Society* (2020).

7 Golla, Hofmann et Bäcker, « Connecting the Dots: Sozialwissenschaftliche Forschung in Sozialen Online-Medien im Lichte von DS-GVO und BDSG-neu », 89.

minimum les dommages » tout au long du processus de recherche, il existe d'autres principes et lignes de conduite à prendre en compte. Dans les deux premières sections, nous résumerons certaines des principales questions éthiques qu'un processus de recherche mené dans ce domaine doit prendre en compte et donnerons un aperçu des principes fondamentaux à observer en matière de protection des données[8]. Nous mettrons ensuite en lumière les possibilités qui s'offrent aux chercheurs et les talents d'équilibriste dont ceux-ci devront faire preuve. Dans la troisième partie, nous discuterons des interactions entre les chercheurs, les sources de données et les politiques des différentes plateformes, et émettrons certaines recommandations essentielles.

8 Dans l'article de Sold, Abay Gaspar et Junk, « Designing Research on Radicalisation using Social Media Content: Data Protection Regulations as Challenges and Opportunities », nous approfondissons certains de ces éléments, comme le font de Koning *et al.* dans les chapitres « On Speaking, Remaining Silent and Being Heard: Framing Research, Positionality and Publics in the Jihadi Field » et « Ethics in Gender Online Research: A Facebook Case Study » de l'ouvrage de Günther et Pfeifer *Jihadi Audiovisuality and its Entanglements. Meanings, Aesthetics, Appropriations* (Édimbourg : Edinburgh University Press, 2020).

2 Questions éthiques fondamentales

Les questions éthiques se posent dans la plupart des projets de recherche. L'éthique de la recherche « cherche [ainsi] à protéger les personnes concernées, et non pas seulement à garantir le respect des aspects légaux »[9]. Les données à caractère personnel occupent une place prépondérante dans l'analyse du contenu des médias sociaux[10]. S'ils ne soulèvent pas de questions éthiques complètement nouvelles et ne remettent pas en cause les « normes et valeurs reconnues de l'éthique de la recherche »[11], l'accessibilité souvent relativement aisée de ces données, leur volume même, ainsi que la vitesse à laquelle les plateformes, les contextes et les événements évoluent font qu'il est nécessaire, mais aussi difficile, d'accorder un espace suffisant aux questions éthiques et de les adapter à des plateformes, perspectives et politiques en pleine mutation. Ainsi, comme pour tout projet de recherche, il convient de trouver un équilibre entre les intérêts sociétaux et scientifiques et le droit des individus au respect de leur vie privée. L'exploitation des données issues des réseaux sociaux soulève des enjeux particuliers et représente « un véritable terrain miné »[12].

S'il n'existe pas de règle universellement reconnue affirmant que les principes éthiques doivent revêtir un caractère obligatoire, la littérature a, à maintes reprises, souligné la pertinence de certains d'entre eux. Nous pouvons classer ces principes en trois catégories : ceux qui ont trait aux rapports entre le chercheur et les sujets de la recherche ; ceux qui se rapportent à la dimension sociétale ; et enfin, les principes introspectifs, qui concernent les chercheurs eux-mêmes. Nous résumons ces conclusions ci-dessous dans le but de les rendre plus facilement accessibles pour les recherches futures sur les rapports entre extrémisme et technologie.

Principes éthiques concernant le sujet de la recherche

Les principes concernant le sujet de la recherche renvoient à la *confidentialité* ou au *respect des personnes*, ainsi qu'à la *bienfaisance*. Pour garantir la *confidentialité*, les chercheurs qui connaissent l'identité d'un sujet de recherche doivent prendre toutes les mesures nécessaires pour s'assurer que cette identité n'est révélée à/par personne. Le consentement de la personne concernée doit être obtenu, dans la mesure du possible, lorsque ses données sont

9 National Committee for Research Ethics in the Social Sciences and the Humanities (NESH), *A Guide to Internet Research Ethics* (2019), 3, https://www.forskningsetikk.no/en/guidelines/social-sciences-humanities-law-and-theology/a-guide-to-internet-research-ethics/.

10 Les données à caractère personnel se définissent comme toute information se rapportant à une personne physique identifiée ou identifiable (la « personne concernée ») ; voir l'article 4(1) du Règlement général sur la protection des données (RGPD).

11 NESH, *A Guide to Internet Research Ethics*, 2.

12 Sold, Abay Gaspar et Junk, « Designing Research on Radicalisation using Social Media Content: Data Protection Regulations as Challenges and Opportunities », 52 ; voir également : Farina Madita Dobrick *et al.*, *Research Ethics in the Digital Age: Ethics for the Social Sciences and Humanities in Times of Mediatization and Digitization*, dir. Farina Madita Dobrick, Jana Fischer et Lutz M. Hagen (Wiesbaden : Springer VS, 2018), 1.

utilisées à des fins de recherche[13]. Le consentement éclairé permet de s'assurer que les droits individuels d'une personne et le droit de cette dernière à l'autodétermination informationnelle sont garantis. En vertu de ce principe, les chercheurs sont donc tenus de faire preuve de transparence et de vérifier que leurs sujets savent qu'ils font l'objet de recherches, qu'ils ont été informés de façon compréhensible sur le projet de recherche à venir et qu'ils ont eu la possibilité d'accepter ou de refuser d'y participer. Il est néanmoins très difficile de demander un consentement éclairé aux sujets de la recherche lorsque celle-ci porte sur des contenus à caractère extrémiste, puisque cela pourrait compromettre les recherches menées : le fait de savoir que leurs données sont observées, voire analysées, amène en effet les individus à modifier leur comportement. Par exemple, s'ils savent qu'ils (ou leurs échanges, publications ou commentaires en ligne) sont observés, ils sont susceptibles d'agir différemment, de communiquer par d'autres moyens, d'arrêter d'exprimer leurs opinions sur la toile ou de les adapter.

D'autres problèmes se posent lorsque les données d'un grand nombre de personnes différentes sont utilisées à des fins de recherche[14]. Il est par exemple guère réaliste de croire que les utilisateurs de 100 000 comptes Twitter donneront, en temps voulu, leur consentement aux chercheurs concernant l'utilisation de leurs données. Dans ce cas de figure, les chercheurs peuvent proposer aux sujets de retirer leur consentement à n'importe quelle étape du projet de recherche. Cette démarche présente l'avantage pour les chercheurs de ne pas avoir à obtenir le consentement des individus en amont. Cette solution peut également être employée si le volume de données à caractère personnel est relativement faible ou si les données sont anonymisées. Les chercheurs doivent toujours traiter de façon confidentielle les données collectées dans le cadre d'une étude et après son achèvement. Que les individus aient ou non donné leur consentement à l'analyse n'a aucune incidence à cet égard. Ils doivent, dans tous les cas, pseudonymiser ou anonymiser les données. Il est toutefois souvent difficile d'anonymiser les données[15], et généralement possible d'identifier les personnes, même après anonymisation[16]. Les chercheurs doivent donc se demander où et comment les données sont conservées, si le logiciel utilisé est fiable et si, par exemple, il est nécessaire d'utiliser un programme de chiffrement, et s'interroger sur le degré d'exhaustivité de la politique de confidentialité d'un éditeur de logiciels.

S'applique également le principe de *bienfaisance* : les chercheurs doivent s'assurer qu'aucun tort n'est causé aux participants et que les avantages de l'étude sont optimisés. Par exemple, dans le cadre d'une étude sur les combattants étrangers, le chercheur doit garantir que toutes les données sont anonymisées de telle façon que la personne concernée ne pourra être identifiée de nouveau, puisque cela pourrait aboutir à des poursuites ou à une condamnation publique. Si une telle anonymisation ne peut être garantie (par exemple, en raison d'une surveillance constante pesant sur le chercheur pendant ses rencontres avec le participant à la recherche, ou parce que l'omission

13 NESH, *A Guide to Internet Research Ethics*, 2.
14 Elizabeth Buchanan, « Considering the ethics of big data research: A case of Twitter and ISIS/ISIL », *PLoS ONE* 12, n° 12 (2017) : 2, https://doi.org/10.1371/journal.pone.0187155.
15 Buchanan, « Considering the ethics of big data research: A case of Twitter and ISIS/ISIL », 4.
16 Matthew J. Salganik, *Bit by Bit. Social Research in the Digital Age* (New Jersey : Princeton University Press, 2018), 40. Idéalement, et d'après la Raison 26 du RGPD, il doit être complètement impossible d'identifier les personnes.

des données à caractère personnel d'un participant rendrait impossible la vérification d'hypothèses), il pourra être nécessaire de mettre fin à l'étude, ou de la restructurer.

L'optimisation des avantages de la recherche et, plus particulièrement, la minimisation des risques sont des entreprises complexes[17]. Toutefois, les avantages de la collecte de données en ligne peuvent être optimisés à moindre effort dans le cadre des recherches relatives à des sujets numériques. Citons à cet effet les outils à seuil minimal offrant la possibilité de fournir des données et des codes en vue de leur reproductibilité (p. ex., Harvard Dataverse ou GitHub) ou les revues de qualité en libre accès pouvant atteindre un large public (comme la toute nouvelle revue *Global Studies Quarterly* d'ISA ou le *Texas National Security Review* de l'Université du Texas à Austin).

Principes éthiques relatifs à la dimension sociétale

Les principes relatifs à la dimension sociétale d'un projet de recherche renvoient à la *justice* et au *respect de la loi et de l'intérêt public*. Le principe de *justice* renvoie au fait que les scientifiques doivent assurer l'équilibre entre les coûts et les avantages pour les différents groupes sociaux concernés par un projet de recherche donné. Les minorités et groupes vulnérables ne doivent pas payer le prix fort alors que la majorité et les groupes fortunés en bénéficient[18].

Selon le principe de *respect de la loi et de l'intérêt public*, les lois et les politiques des sites (p. ex., des sociétés de médias sociaux) applicables à la recherche doivent, en règle générale, être respectées[19]. L'un des principaux problèmes posés par la recherche numérique concerne les nombreuses responsabilités qui doivent être observées, par exemple lors de la collecte de données sur les extrémistes politiques dans différents pays. Dans de très rares cas, toutefois, il est possible de contrevenir aux conditions d'utilisation. Par exemple, l'Université de New York a sciemment décidé de contrevenir aux conditions d'utilisation de Facebook pour collecter des données sur la stratégie de publicité politique de la société, sans doute parce que cette dernière continue de refuser de fournir ces données aux chercheurs[20]. Les publicités politiques et la mésinformation dans le monde numérique étant des problématiques importantes pour l'intégrité électorale et l'amélioration de la démocratie, et les données ne devant être utilisées que pour le bien commun, les conditions d'utilisation de la société ont pu être transgressées dans ce cas particulier. En outre, les chercheurs doivent impérativement discuter de leurs décisions en public et en toute transparence pour satisfaire l'intérêt public[21]. Ce n'est qu'à partir de ce moment-là que le public sera en mesure de mener à bien des débats éthiques sur ce que font les scientifiques et que les opinions issues de ces débats pourront être prises en compte dans les plans de recherche, assurant ainsi une meilleure redevabilité des projets de recherche et un plus grand ciblage de leur contenu. La *transparence* se définit à la fois comme

17 Salganik, *Bit by Bit. Social Research in the Digital Age*, 298.
18 Salganik, *Bit by Bit. Social Research in the Digital Age*, 298 ; NESH*A Guide to Internet Research Ethics*, 5-6 ; British Psychological Society, *Ethics Guidelines for Internet-mediated Research* (2017), 17, www.bps.org.uk/publications/policy-and-guidelines/research-guidelines-policy-documents/research-guidelines-poli.
19 Salganik, *Bit by Bit. Social Research in the Digital Age*, 300.
20 « Facebook to researchers: Stop using our data », 2020, https://edition.cnn.com/2020/10/24/tech/facebook-nyu-political-ad-data/index.html.
21 Salganik, *Bit by Bit. Social Research in the Digital Age*, 300-01.

le fait de présenter et expliquer le projet de recherche aux participants et comme le fait de faire preuve de franchise quant aux méthodes de collecte et de traitement des données utilisées lors de la présentation ou de la publication des résultats de la recherche.

Principes éthiques applicables aux chercheurs eux-mêmes

Les chercheurs font évidemment partie intégrante du processus de recherche – et les environnements universitaires et institutionnels dans lesquels ils évoluent doivent se soucier de leur bien-être. Il est question ici de la *sécurité du chercheur*. Plus particulièrement, lorsque la recherche traite d'un sujet aussi délicat que celui de la radicalisation, elle doit tenir compte dès sa conception de la protection des chercheurs concernés. Ceux-ci pourraient en effet être victimes de menaces physiques ou d'intimidations, mais leur sécurité comporte également un volet psychologique, puisque le fait de devoir analyser des contenus potentiellement éprouvants peut avoir ses limites. Ces aspects doivent être pris en considération avant le lancement du projet, mais sont beaucoup trop souvent omis ou insuffisamment garantis par les institutions engagées dans la recherche.

La question de la *confiance* se pose également, du point de vue à la fois des sujets de la recherche et des chercheurs eux-mêmes. Par exemple, les chercheurs doivent se demander si un profil est authentique. Il y a des limites à ce qui peut être vérifié et à la transparence de l'identité affichée (il peut être nécessaire de dissimuler son identité pour assurer sa sécurité). Nombreux sont les utilisateurs qui emploient des pseudonymes, fournissent des données de localisation inexactes ou choisissent de s'exprimer dans une autre langue. Souvent, les contributions sont écrites en anglais, ce qui complique l'attribution d'une nationalité aux utilisateurs.

De plus, les passages d'une plateforme à une autre, dans le cadre desquels un fil de discussion sur une plateforme est relié à un autre sur une autre plateforme, s'avèrent problématiques pour les chercheurs. Les abréviations, les néologismes, l'utilisation de plusieurs langues et les structures de phrases incomplètes sont caractéristiques des conversations en ligne et représentent des défis supplémentaires[22]. L'analyse automatisée de contenu dirigée par des programmes est donc moins aisée et crée un nouvel ensemble de difficultés[23]. La recherche intégrée est l'un des moyens permettant de faire face à ces subtilités. Le rôle actif ou passif qu'adoptent les chercheurs dans le cadre du processus de collecte de données peut avoir de graves conséquences pour la validité interne d'un plan de recherche, et soulève un autre sous-groupe de questions éthiques. Si les chercheurs réussissent à endosser un rôle complètement passif/d'observation à chaque étape du processus de collecte de données, ils n'influenceront probablement pas les processus de communication observés – ce qui peut revêtir une importance capitale pour la validité

22 Albert Bifet et Eibe Frank, « Sentiment knowledge discovery in Twitter streaming data », in *Discovery Science*, dir. Bernhard Pfahringer, Geoff Holmes et Achim Hoffmann, Lecture Notes in Computer Science (Heidelberg : Springer VS, 2010) ; Simon Carter, Wouter Weerkamp et Manos Tsagkias, « Microblog language identification. Overcoming the limitations of short, unedited and idiomatic text », *Language Resources and Evaluation* 47, n° 1 (2013).

23 Voir Alrhmoun, Maher et Winter, *Décrypter la haine : emploi de l'analyse de texte expérimentale aux fins de classification des contenus à caractère terroriste*.

des résultats de leurs recherches. Mais l'observation comporte souvent des limites (p. ex., questions ciblées sur le profil du chercheur), et la frontière entre observations intrusives et non intrusives est ténue.

D'un point de vue éthique, les paramètres de confidentialité sont, eux aussi, importants pour la mise en œuvre d'un projet de recherche. Si les paramètres choisis rendent le contenu public, l'analyse des données sera alors considérée comme moins attentatoire à la vie privée du sujet que s'il a choisi de ne partager ses données qu'avec ses « amis » ou un sous-ensemble plus restreint encore de personnes triées sur le volet. Les questions éthiques qui se posent dans le cadre des recherches menées sur les données issues des médias sociaux vont de pair avec des questions juridiques. Puisque les éventuels préjudices ne peuvent ni être totalement évités ni pleinement anticipés, l'objectif à la fois de la réflexion éthique et des exigences juridiques est de rechercher à maintenir l'équilibre entre les avantages escomptés de la recherche et les intérêts relatifs à la vie privée[24]. Si les exigences juridiques et les questions éthiques sont interdépendantes et ne peuvent être comprises que comme un tout, elles doivent toutefois être abordées séparément par les chercheurs. Nous examinons ci-dessous les recommandations juridiques extraites de la littérature.

24 Anne Lauber-Rönsberg, « Data Protection Laws, Research Ethics and Social Sciences », in *Research Ethics in the Digital Age. Ethics for the Social Sciences and Humanities in Times of Mediatization and Digitization*, dir. Farina M. Dobrick, Jana Fischer et Lutz M. Hagen (Wiesbaden : Springer VS, 2018), 41.

3 Principes fondamentaux de la protection des données – limites légales, enjeux et possibilités pour les chercheurs

Les individus (et les groupes) divulguent souvent beaucoup d'informations sur eux-mêmes sur les médias sociaux, et plus particulièrement sur les réseaux sociaux. Ils peuvent donner des informations personnelles relatives à leur origine ethnique, leurs opinions politiques, leurs croyances religieuses et idéologiques, leurs habitudes et orientation sexuelles, leur état de santé et leurs affiliations, à un syndicat par exemple. Certaines de ces informations personnelles peuvent avoir un intérêt pour les chercheurs travaillant dans différents domaines. Quel que soit le contexte, les règles de protection des données doivent être respectées en cas de collecte ou d'exploitation de données à caractère personnel. Nous présentons ci-dessous le cadre juridique de protection des données applicable à la recherche sociale empirique menée par observation sur les médias sociaux, fondé sur le Règlement général sur la protection des données (RGPD)[25].

Si la réglementation accorde quelques privilèges importants à la recherche scientifique, elle ne prévoit cependant aucune exception particulière pour le traitement des données. La légalité du traitement à des fins de recherche est souvent appréciée au cas par cas, en mettant en balance les différents intérêts. Certaines données à caractère personnel sont particulièrement sensibles et font donc l'objet d'une protection renforcée (citons, par exemple, les croyances religieuses ou les opinions politiques, particulièrement pertinentes pour l'examen des contenus à caractère extrémiste sur les médias sociaux). Les chercheurs peuvent aussi, en fonction de leur projet de recherche, faire face à certaines limites, difficultés ou possibilités, que nous étudierons ci-dessous, en distinguant les cas où le consentement a été obtenu et ceux où il n'a pas été recherché, cet aspect revêtant une importance capitale.

Règles juridiques applicables à la recherche portant sur les données à caractère personnel menée avec le consentement des personnes concernées

Bon nombre des données affichées sur les médias sociaux revêtent un caractère personnel. Si ces informations sont relativement facilement accessibles en ligne et ont été publiées délibérément par les personnes concernées, elles demeurent tout de même protégées

25 Le RGPD est entré en vigueur le 25 mai 2018 dans tous les États membres de l'Union européenne. Il vise à harmoniser les lois relatives à la confidentialité des données sur le continent européen.

par le RGPD de par leur nature de données à caractère personnel. La collecte et l'analyse de données à caractère personnel est inévitable dans le cadre de nombreux projets de recherche. La protection juridique de ces informations personnelles est assurée, au sein de l'Union européenne, par le Règlement susmentionné[26]. Celui-ci n'autorise pas expressément le traitement des données à caractère personnel à des fins de recherche scientifique, et les conditions applicables au traitement desdites données sera donc régi par les dispositions des articles 5, 6 et 9. En vertu du RGPD, le traitement des données à caractère personnel est généralement interdit sauf si la loi l'autorise expressément ou si la personne concernée a donné son accord.

En donnant ou en refusant son accord, une personne peut décider des modalités de divulgation et d'utilisation de ses données à caractère personnel. Elle doit avoir la possibilité de décider, au cas par cas, si ses données peuvent être traitées, et dans quelle conditions. Pour cela, les règles relatives au consentement portant sur le traitement des données à caractère personnel précisent les conditions, à la fois formelles et relatives au contenu, à remplir pour obtenir le consentement. En vertu du RGPD, ces conditions sont les suivantes : la finalité visée par le traitement des données doit être expliquée, la personne concernée doit recevoir suffisamment d'informations sur le traitement envisagé, le consentement doit être donné volontairement et l'utilisateur doit avoir la possibilité de révoquer son consentement à tout moment.

Les chercheurs ont aussi le droit d'exploiter les données sensibles lorsqu'elles ont été consciemment rendues publiques par la personne concernée. Dans ce cas de figure, l'article 9(2)(e) du Règlement lève l'interdiction de traitement prévue au paragraphe 1, et la personne concernée n'a pas particulièrement besoin d'être protégée. La publication consciente des données par la personne concernée peut être considérée comme une exception à la protection spéciale prévue à l'article 9. Il convient de remarquer néanmoins que, même si les données sont publiées par la personne concernée, elles ne se soustraient pas complètement à la protection du RGPD[27]. L'article 6 notamment est, quant à lui, toujours applicable, et le traitement des données nécessite toujours une base légale, même en cas de dérogation à l'article 9 (1)[28].

Ce qui nous amène à la question suivante : qu'entend-on par données « rendues publiques » ? Les données sont réputées avoir été rendues publiques lorsqu'elles sont mises à disposition du public par un nombre indéterminé de personnes sans entrave particulière. Un autre aspect central de la protection des données concerne donc les types de médias sociaux sur lesquels les données ont été publiées à l'origine. Proviennent-elles de (parties de) médias sociaux ouverts ou fermés, ou de médias sociaux créés spécialement à des fins de recherche ? Le principal critère de démarcation ici est « la difficulté d'accès par inscription et ouverture de session »[29]. Les utilisateurs

26 Le champ d'application du règlement porte sur tous les types et toutes les formes de « manipulation » des données, y compris la collecte, le stockage, la structuration, etc. Voir l'article 4(2) du RGPD.
27 Golla, Hofmann et Bäcker, « Connecting the Dots: Sozialwissenschaftliche Forschung in Sozialen Online-Medien im Lichte von DS-GVO und BDSG-neu », 92.
28 Il convient de noter à ce stade que les article 6 et 9 du RGPD sont aussi applicable en parallèle l'un de l'autre. Ces articles figurent tous deux dans le règlement et doivent tous deux être respectés.
29 Golla, Hofmann et Bäcker, « Connecting the Dots: Sozialwissenschaftliche Forschung in Sozialen Online-Medien im Lichte von DS-GVO und BDSG-neu », 96.

ont, sur certaines plateformes, la possibilité de limiter les destinataires de leur contenu. Certains médias sociaux sont spécialement mis en place à des fins de recherche, et la recherche du consentement des utilisateurs relatif au traitement de leurs données à caractère personnel s'avère être, dans ce cas, réalisable. Mais ce n'est pas le cas pour tous les médias sociaux. C'est là que les exigences légales relatives au traitement des données à caractère personnel prennent toute leur pertinence. C'est aussi le cas si les données collectées proviennent de sources disponibles publiquement[30]. En outre, « des données de nature publique peuvent relever de la vie privée lorsqu'elles sont, d'une manière systématique, recueillies et mémorisées dans des fichiers tenus par les pouvoirs publics »[31]. Les individus jouissent par ailleurs du droit à la protection de la vie privée même lorsqu'ils entrent volontairement sur la scène publique. Les données issues d'espaces de communication semi-publics ou fermés ont encore plus besoin d'être protégées que les données tirées des espaces publics.

Cependant, comme nous l'avons mentionné plus haut, le consentement des personnes peut compromettre la recherche, ou est tout simplement impossible à obtenir compte tenu du nombre de personnes qu'il faudrait solliciter. Ce problème ne se pose pas que dans le domaine de la recherche sur les individus ou groupes radicalisés ou extrémistes, mais aussi dans d'autres domaines de recherche sensibles. C'est pourquoi nous étudions ci-dessous les règles juridiques applicables à la recherche portant sur les données à caractère personnel menée sans le consentement des personnes concernées.

Règles juridiques applicables à la recherche portant sur les données à caractère personnel menée sans le consentement des personnes concernées

Puisque la recherche doit souvent s'appuyer sur des données à caractère personnel pour atteindre ses objectifs, le législateur a établi des critères d'admissibilité qui lui sont propres. Celles-ci autorisent d'imposer des limites au droit à l'autodétermination informationnelle à des fins de recherche scientifique. Si les données sont utilisées pour des recherches n'appartenant à aucune des catégories définies par l'article 9 du RGPD, la licéité de leur traitement est exclusivement régie par l'article 6. Au moins un fondement visé par cet article doit donc s'appliquer lors du traitement des données à caractère personnel. Par conséquent, le traitement de données à caractère personnel en l'absence de consentement n'est autorisé que dans certaines circonstances : par exemple, si les intérêts légitimes de la personne concernée ne sont pas lésés du tout ou si l'intérêt public du projet de recherche l'emporte sur les intérêts légitimes de la personne concernée et si le but de la recherche ne peut être atteint d'une autre façon, ou seulement moyennant des efforts disproportionnés. Si ces conditions sont remplies, la recherche peut être menée sans le consentement de la/des personne(s) concernée(s). La licéité du traitement sans consentement dépend souvent d'un équilibre entre le droit au respect

30 Ian Brown et Josh Cowls, *Check the Web. Assessing the Ethics and Politics of Policing the Internet for Extremist Material*, Oxford Internet Institute (2015), 46. Les espaces publics se caractérisent par un accès illimité et la coprésence d'inconnus. En revanche, les « espaces privés se caractérisent par un accès limité … et l'absence d'inconnus ». Nicolas Legewie et Anne Nassauer, « YouTube, Google, Facebook: 21st Century Online Video Research and Research Ethics », *Forum: Qualitative Social Research* 19, 32, n° 3 (2018).

31 Cour européenne des droits de l'homme, « Rotaru c. Roumanie, requête n° 28341/95 », (2000) : § 43.

de la vie privée et les avantages de la recherche. Dans tous les cas, il est nécessaire d'évaluer les intérêts de la recherche à la lumière des intérêts légitimes des personnes concernées.

En vertu de l'article 9(2)(j) du RGPD, certaines dispositions légales autorisent également le traitement de certaines catégories de données à caractère personnel aux fins de la recherche : celles-ci exigent, en premier lieu, l'existence *d'un sujet et d'un concept de recherche spécifiques*. Aux fins de la recherche scientifique, les chercheurs doivent prouver que la structure et le contenu de leur projet de recherche répond à des exigences scientifiques. Deuxièmement, les chercheurs doivent prouver que le *projet ne peut aboutir* sans les données à caractère personnel recherchées. Ils doivent ainsi expliquer en détail pourquoi il leur est impératif de collecter les données à caractère personnel en question pour leur projet de recherche. Ils devraient par exemple se demander s'ils peuvent mener à bien leurs recherches avec moins de données ou avec d'autres types de données. Troisièmement, les intérêts doivent de nouveau être évalués, par exemple à la lumière du volume de données et des circonstances spéciales des personnes concernées. Ainsi, pour légitimer le traitement des données à des fins de recherche sans le consentement de la personne concernée, les chercheurs devront prouver pourquoi les intérêts de la recherche l'emportent (largement) sur l'intérêt qu'a la personne concernée de protéger ses données.

À cette fin, les *principes de nécessité, de pertinence et de proportionnalité* du traitement des données à caractère personnel doivent être observés, et des règles d'accès doivent être adoptées afin de garantir que ces données soient exploitées dans le respect des règles de protection des données : les chercheurs doivent tout d'abord prouver que leur projet poursuit un *objectif légitime*. Il est vrai que la recherche en général peut être considérée comme un objectif légitime en soi[32]. Néanmoins, le traitement des données à caractère personnel dans le cadre d'un projet de recherche en l'absence de consentement des personnes concernées ne doit être envisagé que si l'objectif de la recherche ne peut être atteint d'une autre façon[33]. La *nécessité* doit être comprise comme une autre condition préalable à la proportionnalité. Une mesure est nécessaire si aucune autre mesure plus modérée – moins intrusive – ne permet d'atteindre le même objectif. Le test de nécessité doit être vu comme la première étape à franchir pour une mesure proposée impliquant le traitement de données à caractère personnel. Si la mesure en question ne franchit pas cette étape, il devient inutile d'en examiner la proportionnalité. S'il n'est pas démontré qu'une mesure est nécessaire, elle devra être modifiée de façon à le devenir.

Le traitement des données obtenues en ligne sans consentement doit également être *pertinent*. Le principe de pertinence suppose que le contenu et la forme de l'action n'excèdent pas le niveau nécessaire pour atteindre les objectifs. Pour déterminer le degré de pertinence de leur intervention, les chercheurs doivent évaluer les raisons justifiant sa réalisation sur le plan juridique (souvent fondées sur les bénéfices sociétaux perçus de la recherche) à la lumière de l'obligation de protéger l'individu dont le droit à la vie privée sera bafoué par

32 Golla, Hofmann et Bäcker, « Connecting the Dots: Sozialwissenschaftliche Forschung in Sozialen Online-Medien im Lichte von DS-GVO und BDSG-neu », 90.
33 Sold, Abay Gaspar et Junk, « Designing Research on Radicalisation using Social Media Content: Data Protection Regulations as Challenges and Opportunities », 62-63.

cette intervention. Ils doivent par la suite démontrer le respect des mesures et garanties visant à protéger les personnes concernées – p. ex., la pseudonymisation –, conformément à l'article 89 du RGPD.

Un autre aspect pertinent de l'examen du projet de recherche du point de vue de la protection des données renvoie au rôle actif ou passif qu'endosse le chercheur. La question de savoir si les chercheurs sont des observateurs passifs lorsqu'ils collectent les données – en d'autres termes, s'ils ont opté pour une méthode de collecte non réactive – a également des effets sur les exigences en matière de protection des données. Dans le cadre d'une telle procédure, le chercheur n'endosse à aucun moment un rôle actif et ne participe pas au débat. Bien qu'une telle observation passive écarte également la recherche du consentement dès le départ[34], l'intervention est réduite au minimum dans la mesure où aucune influence n'est exercée sur les commentaires ou les publications. En parallèle, l'adoption d'une approche active de la recherche signifie qu'il est possible d'obtenir l'accord des internautes en vue de la collecte et de l'analyse de leurs données à caractère personnel, mais introduit également le risque de produire des contenus parasites, de faire avancer (davantage) le débat ou d'influencer le comportement des autres en matière de publication.

Il convient d'accorder une meilleure protection aux individus dont le consentement ne peut être obtenu. Selon l'article 89(1) du RGPD, des mesures techniques et organisationnelles doivent être prises pour garantir, en particulier, le respect du principe de minimisation des données. La réduction du volume de données collectées et la limitation de la portée du traitement à la seule mesure nécessaire au but poursuivi, la détermination de la durée de conservation, et une réglementation sur l'accessibilité des données constituent certains aspects importants à cet égard. L'article 89(1) du RGPD dispose par exemple, dans ses troisième et quatrième phrases, que les données devront être anonymisées ou pseudonymisées dans la mesure où les finalités de la recherche peuvent être atteintes de cette manière. En ce qui concerne l'archivage des données, les concepts de rôles et les solutions d'accès sécurisé sont évidents[35].

Enfin, même si les données à caractère personnel peuvent être traitées (en vertu d'un consentement ou d'une disposition légale), des mesures techniques et organisationnelles devront être prises pour garantir que les objectifs de la protection des données sont satisfaits. On pourrait par exemple parvenir à ce résultat en stockant les identifiants et les données séparément. Par ailleurs, les données devraient être réservées à une affectation précise. Les informations ne seront ainsi conservées et examinées que dans le but pour lequel elles ont été collectées.

34 Kerstin Eppert *et al.*, *Navigating a Rugged Coastline: Ethics in Empirical (De-)Radicalization Research*, core-nrw Netzwerk für Extremismusforschung in Nordrhein-Westfalen (Bonn, 2020), 9, https://www.bicc.de/fileadmin/Dateien/Publications/other_publications/Core-Forschungsbericht_1/CoRE_FP_1_2020.pdf.

35 Golla, Hofmann et Bäcker, «Connecting the Dots: Sozialwissenschaftliche Forschung in Sozialen Online-Medien im Lichte von DS-GVO und BDSG-neu», 94.

4 Sources de données, politiques des plateformes et chercheurs – présentation, interactions et recommandations

Outre les principes éthiques et les réglementations en matière de protection des données, les utilisateurs et les chercheurs doivent respecter les accords juridiques des plateformes et les prendre en compte, de même que les autres limites individuelles accompagnant l'utilisation de programmes tiers. Le fait que les principales plateformes du secteur aient adopté différentes politiques et que ces dernières soient souvent trop longues ou difficiles à comprendre constitue un défi en soi. Nous fournissons ci-dessous un bref aperçu des politiques les plus importantes des principales sociétés technologiques, et formulons quelques recommandations générales.

Twitter

Avec sa nouvelle politique de confidentialité conforme au RGPD, entrée en vigueur en mai 2018, Twitter a octroyé à ses utilisateurs plus de contrôle sur leurs données. Puisque cette politique s'applique à tous les utilisateurs quel que soit la région du globe où ils se trouvent, la protection fournie par le RGPD semble s'étendre aux utilisateurs du monde entier. Twitter collecte des informations sur l'adresse IP et le type d'appareil utilisé auprès de ses utilisateurs dès l'instant où ils consultent la plateforme. Bien entendu, des données sont aussi générées et collectées lorsqu'un utilisateur envoie des Tweets, échange avec d'autres utilisateurs, retweete, aime un Tweet, etc. En vertu de la politique de confidentialité de Twitter, le contenu des messages directs est exclus de la collecte et du traitement des données. Les données collectées sont utilisées pour suggérer des Tweets, suivre des comptes et envoyer des publicités ciblées. Twitter offre donc à ses utilisateurs un certain contrôle sur les types de données pouvant être recueillies. Par exemple, les utilisateurs peuvent rendre leur compte public ou privé et activer ou désactiver la possibilité pour d'autres utilisateurs de les identifier dans des photos. Ils peuvent également télécharger les informations qu'ils ont partagées sur Twitter. Par exemple, outre les Tweets publics, qui « peuvent immédiatement être vus et faire l'objet de recherches par qui que ce soit dans le monde entier », les utilisateurs ont la possibilité d'utiliser « des moyens non publics de communiquer sur Twitter par le biais des Tweets protégés et des Messages Directs »[36]. Il est par ailleurs possible d'utiliser un pseudonyme sur Twitter et les « données sont conservées

36 Twitter, *Politique de confidentialité de Twitter* (2020), https://cdn.cms-twdigitalassets.com/content/dam/legal-twitter/site-assets/privacy-page-gdpr/pdfs/PP_Q22018_April_FR.pdf

pendant 18 mois maximum ou jusqu'à la suppression du compte »[37]. Avec le lancement de la v.2 de son API en août 2020, « Twitter aide les entreprises, les universitaires et les développeurs tiers à tirer le meilleur parti de sa plateforme »[38] : la société offre aux développeurs tiers un accès à des fonctionnalités depuis longtemps refusées à ses clients, tels que « les fils de discussion, les résultats de sondage dans les Tweets, les Tweets fixes, les filtres antispams, et un filtrage de flux et un langage de requête de recherche plus puissants »[39]. L'accès aux flux de Tweets en temps réel est également désormais possible.

L'accès à l'API a été réorganisé par Twitter à trois niveaux : pendant la période d'accès anticipé, il était possible « d'écouter et d'analyser la discussion publique »[40]. Mais, compte tenu du fait que seul le niveau d'accès gratuit et de base a été lancé, qui limite le nombre d'appels API que peuvent passer les développeurs, il reste à voir quels changements et opportunités s'offriront aux chercheurs. Twitter présente un avantage non négligeable par rapport aux autres réseaux sociaux : sa communication ouverte. Les Tweets individuels, de même que des conversations entières, peuvent faire l'objet de recherches et être vus par tout un chacun, qu'ils soient ou non utilisateurs de Twitter ou suivent ou non la personne concernée. Les chercheurs ont ainsi accès non seulement à des données exhaustives et non filtrées, mais aussi à la protection des données. L'une des limites posées à la réalisation de recherches sur Twitter est que l'utilisation des données ne doit pas nuire aux intérêts économiques de l'entreprise. La création, l'enrichissement et la diffusion de vastes bases de données de Tweets sont interdits, même à des fins non commerciales[41]. Les résultats ne peuvent être pondérés ou comparés parce que les chercheurs n'ont aucune information sur l'activité globale sur Twitter.

Facebook

Comme Twitter, Facebook a révisé sa politique de protection des données pour la mettre en conformité avec le RGPD. Cette politique est désormais applicable à ses utilisateurs du monde entier. Facebook, Instagram, Messenger et les autres produits et fonctionnalités proposés par Facebook collectent différents types d'informations en fonction des interactions des utilisateurs avec les produits de la société, comme les informations et contenus téléchargés sur les plateformes par un utilisateur, les données de ses réseaux (comptes, groupes, hashtags etc. avec lesquels il interagit), ses informations d'utilisation et les données relatives aux achats effectués sur les plateformes, ou encore les données relatives aux interactions d'autres utilisateurs avec ses contenus et profils. Par ailleurs, des données sont collectées sur les appareils connectés à Facebook ou Instagram, y compris les propriétés de l'appareil, les mouvements du curseur, le fournisseur d'accès Internet, la société de téléphone et les paramètres de l'appareil. Facebook utilise ces données pour affiner ses produits et personnaliser ses recommandations de contenus et comptes. La société met également ces données à la disposition de

37 Identity Guard, « What You Need to Know About Twitter's Privacy Policy », (2018), https://www.identityguard.com/news/twitter-privacy-policy.
38 « Twitter launches new API as it tries to make amends with third-party developers », 2020, https://www.theverge.com/2020/8/12/21364644/twitter-api-v2-new-access-tiers-developer-portal-support-developers.
39 « Twitter ändert API zugunsten von Third-Party-Entwicklern », 2020, https://onlinemarketing.de/technologie/twitter-api-third-party-entwicklern.
40 « Twitter API v2: Early Access », 2020, https://developer.twitter.com/en/docs/twitter-api/early-access.
41 Michael Beurskens, « Legal questions of Twitter research. Twitter and society », in *Digital Formations*, dir. Katrin Weller (New York *et al.* : Peter Lang, 2014).

clients tiers. En outre, les données sont non seulement partagées avec des annonceurs, mais également avec des tiers qui exécutent des applications sur Facebook ou utilisent ses services d'une autre façon. Comme sur les autres médias sociaux, les utilisateurs peuvent limiter la collecte de données en modifiant leurs paramètres, de même que télécharger et accéder aux données utilisateur collectées à leur propos. Certaines données sont sujettes à des protections spéciales : les utilisateurs peuvent choisir de fournir à Facebook des informations sur leurs opinions religieuses ou politiques, leur santé, leurs origines raciales ou ethniques, leurs croyances philosophiques ou leur adhésion à un syndicat[42]. Bien que Facebook se soit récemment amélioré sur le plan du respect de la vie privée[43], l'interface utilisateur n'est toujours pas suffisamment transparente. De plus, s'ils peuvent limiter les publicités personnalisées, les utilisateurs ne peuvent guère en revanche restreindre la collecte de données à leur sujet. Facebook donne à ses utilisateurs la possibilité d'accéder à leurs informations Facebook, y compris leurs photos, leurs publications, leurs réactions et leurs commentaires à l'aide de l'outil « Accéder à ses informations ». Enfin, les utilisateurs peuvent télécharger une copie de leurs informations Facebook grâce à l'outil « Téléchargez vos informations ».

Facebook met aussi à la disposition des chercheurs et universitaires des informations et des contenus pour mener leurs recherches[44]. En réponse au fiasco de Cambridge Analytica en 2018, Facebook a promis de mettre sur pied une initiative de recherche pour donner aux universitaires un accès à ses données tout en assurant la confidentialité des informations des utilisateurs. Malgré le lancement d'une nouvelle plateforme d'accès aux données sensée donner aux chercheurs l'occasion de voir tous les ensembles de données Facebook disponibles, la société continue d'être critiquée[45] pour le manque d'assistance fournie au monde de la recherche.

Google

Google, contrairement à Facebook et Twitter, semble avoir jusqu'à présent s'être refusé d'appliquer sa politique de confidentialité, conforme au RGPD, aux régions extérieures à l'UE. Il a par exemple été signalé que les utilisateurs situés au Royaume-Uni allaient perdre la protection que leur fournissait le RGPD et devaient accepter que leurs données seraient désormais conservées aux États-Unis – contrairement à celles des utilisateurs de l'UE, qui doivent être conservées sur des serveurs conformes aux règles du RGPD. Cela signifie que les niveaux de protection des données varient en fonction de directives que les prestataires de service ont globalement établies eux-mêmes. YouTube n'étant qu'une partie de l'empire Google, qui se compose de dizaines d'applications et de services et d'un système d'exploitation mobile, l'entreprise est donc susceptible de collecter plus de données sur ses utilisateurs que, par exemple, Twitter ou Facebook. YouTube recueille notamment des données sur les interactions entre utilisateurs, les commentaires, les téléchargements de vidéos sur la plateforme, la consommation de vidéos et bien d'autres choses.

42 « Politique d'utilisation des données », 2020, https://fr-fr.facebook.com/policy.php.
43 « Mit mehr Kontrolle über die eigene Privatsphäre ins neue Jahrzehnt », 2020, https://about.fb.com/de/news/2020/01/mehr-kontrolle-uber-die-eigene-privatsphare/.
44 Pour plus de détails, voir Facebook Research. « Supporting exciting and innovative research through meaningful engagements », 2020.
45 Voir par exemple « Facebook needs to share more with researchers », World View, 2020, https://www.nature.com/articles/d41586-020-00828-5.

Si YouTube partage les données utilisateurs avec les tiers qui publient des annonces publicitaires sur le site et fournit une API, il est explicitement indiqué que la plateforme ne vend pas ces données à des tiers, comme d'autres médias sociaux. YouTube offre de multiples possibilités aux utilisateurs qui le souhaitent d'examiner, voire de supprimer, leurs données.

TikTok

Contrairement à de nombreux médias sociaux, TikTok a choisi de segmenter sa politique de confidentialité sur le plan régional. En Europe, par exemple, elle a mis en place une directive qui tient compte de certaines exigences du RGPD. D'autres principes s'appliquent aux États-Unis et dans d'autres pays. Outre les données habituelles (activités d'utilisation, informations sur l'appareil, données de localisation, répertoire lorsque l'utilisateur en fournit l'accès, informations sur les contenus tiers publiés sur la plateforme), le contenu est également collecté et analysé. TikTok ne semble pas avoir donné accès aux chercheurs à un API, ni leur avoir fourni d'autres moyens de collecter des données légalement. Des spécialistes informatiques ont en revanche trouvé le moyen de créer des API non officiels pour collecter des données sur les utilisateurs, les vues et les interactions[46].

Telegram

À l'instar de TikTok, Telegram a mis en place une politique de confidentialité distincte pour ses utilisateurs européens. En tant que plateforme de communication, Telegram ne conserve que des informations de base sur ses utilisateurs (numéro de téléphone, adresse e-mail, nom d'utilisateur, etc.). Les discussions normales (appelées « discussions cloud ») entre les utilisateurs, de même que les discussions de groupe, sont également conservées. Les discussions secrètes sont soi-disant chiffrées de bout en bout et uniquement visibles pour les utilisateurs concernés. Telegram ne fournit pas non plus aux chercheurs de moyens de collecter et d'analyser les données, comme un API. Certains chercheurs ont toutefois créé leur propre outil de scraping pour accéder aux chaînes, interactions et messages publics à des fins de recherche[47]. Si le scraping est un outil attrayant pour évaluer les réseaux sociaux à des fins de recherche, il représente, sur les plans de la légalité et de l'éthique, une méthode particulièrement contestée de récupération des données[48].

Recommandations générales

L'image fournie par cet aperçu est, au mieux, diffuse : les chercheurs ont accès à certaines données, en fonction de la plateforme concernée. Généralement, les plateformes se réservent les droits sur les données, et le droit de les traiter et de les transmettre. Toutes les plateformes n'ont cependant pas clairement explicité les points

46 « How to Collect Data from TikTok », 2020, https://towardsdatascience.com/how-to-collect-data-from-tiktok-tutorial-ab848b40d191.
47 Jason Baumgartner *et al.*, *The Pushshift Telegram Dataset* (2020).
48 Sebastian J. Golla et Max von Schönfeld, « Kratzen und Schürfen im Datenmilieu – Web Scraping in sozialen Netzwerken zu wissenschaftlichen Forschungszwecken », *Kommunikation und Recht* (2019).

d'accès et termes de référence applicables à l'utilisation scientifique de ces données. Une plus grande ouverture de nombreuses sociétés technologiques à la science serait également souhaitable, avec la mise en place d'API clairs, durables et harmonisés, ainsi que dans le domaine des recherches (par exemple, d'ensembles de données créés à l'aide d'un critère de recherche). Sur Twitter, par exemple, les Tweets reliés par des réponses ne figurent pas dans les résultats de recherche. Data Grants[49] est un programme pilote visant à donner aux chercheurs un accès aux données publiques et historiques, mais cet accès est limité à quelques projets choisis par Twitter.

La recherche et les enquêtes sur l'extrémisme politique souvent violent en ligne figurent parmi les principales priorités de nombreuses institutions politiques et sociétales, mais aussi des sociétés technologiques. Les bases de données contenant les données des utilisateurs sont, comme nous l'avons déjà vu plus haut, soumises aux réglementations nationales respectives de protection des données, qui limitent, pour d'excellentes raisons, le partage de données existantes avec d'autres scientifiques dans le pays et, surtout, à l'étranger. Dans ce contexte, il y a lieu de s'interroger sur l'utilisation potentielle par les chercheurs des données collectées pour leurs analyses et projets. Tandis que le RGPD s'applique à l'ensemble des États membres de l'UE, les règles de collaboration avec des partenaires extérieurs sont moins évidentes, même si les normes partout dans le monde semblent converger de plus en plus et si les sociétés technologiques comme Facebook mettent en œuvre et prônent l'adoption de règles internationales.

S'il existe encore pas mal d'obstacles, la tendance est prometteuse. De plus, la plupart des réglementations de protection des données, y compris le RGPD, octroient certains privilèges aux chercheurs. Si certains principes sont mis en balance de façon systématique et transparente dans les stratégies de protection des données pour certains projets de recherche et en concertation étroite avec les responsables de la protection des données (et, dans certains cas, avec les plateformes), les analyses nécessaires et l'accès aux conclusions pour d'autres chercheurs sont possibles dans la quasi-totalité des cas. Pourtant, il existe encore certaines limites à la reproduction des résultats si les données sont tirées d'espaces chiffrés et sous condition de pseudonymisation et d'anonymisation. Cette situation est aggravée par le fait que, de plus en plus, les contenus à caractère extrémiste sont supprimés rapidement. Si les contenus à caractère extrémiste supprimés étaient hébergés en toute sécurité, les chercheurs pouvant y accéder seraient susceptibles de mener des analyses plus approfondies. À cet égard, il y a beaucoup d'éléments à aborder avec les éditeurs universitaires en ligne et papier, par exemple la garantie d'un niveau élevé de validité externe pour une étude publiée sans prévoir de mesures incitant à violer les règles de protection des données et l'éthique de la recherche. Si cet équilibre n'est pas atteint, la recherche produira trop peu de résultats pertinents.

Une coopération étroite entre les sociétés technologiques et les chercheurs en matière de partage de connaissances, de collaboration technique et de recherche partagée bénéficierait aux deux parties. Les chercheurs pourraient signaler davantage les contenus

49 Voir « Introducing Twitter Data Grants », 2014, https://blog.twitter.com/engineering/en_us/a/2014/introducing-twitter-data-grants.html.

problématiques, par exemple, mais devraient s'intéresser de manière critique aux effets du signalement conformément aux normes éthiques présentées plus haut. Les sociétés technologiques doivent faire preuve de transparence sur les mécanismes mis en place pour traiter le contenu signalé et être conscientes des difficultés éthiques et pratiques auxquels sont confrontés les chercheurs à cet égard. Une solution pourrait être de fournir une option de traiter le contenu signalé par les chercheurs différemment des autres contenus : les sociétés pourraient surveiller ce contenu de près sans le supprimer. Un exemple de coopération réussie entre les sociétés technologiques et le monde de la recherche est le Forum mondial de l'Internet contre le terrorisme (GIFCT). L'un des objectifs principaux du GIFCT est de « permettre aux chercheurs d'étudier le terrorisme et la lutte contre le terrorisme, notamment en mettant en place et en évaluant les meilleures pratiques de coopération pluripartite et la prévention de l'utilisation abusive des plateformes numériques »[50]. Le GNET est financé par le GIFCT et les dialogues – à la fois critiques et ouverts – que cela engendre sont extrêmement précieux et doivent être maintenus de façon permanente pour aller plus loin.

Différentes plateformes ont mis sur pied plusieurs outils, ou tout au moins plusieurs initiatives, qui revêtent un intérêt pour les chercheurs travaillant sur les contenus publics des médias sociaux. CrowdTangle[51] en est un, qui permet d'analyser les contenus publics sur les médias sociaux et de les compiler sous forme de rapports. Grâce à CrowdTangle, on peut avoir accès à l'heure à laquelle une publication a été postée, au type de contribution (vidéo, image, texte), à des informations sur la page, le compte public ou le groupe public sur lesquels cette publication a été publiée, au nombre d'interactions (p. ex., informations sur les « J'aime », réactions, commentaires, nombre de partages de la contribution) ou de vues qu'elle a généré, et aux autres pages ou comptes publics qui l'ont partagé. C'est un bon début, mais il est possible de faire mieux et d'aller plus loin. Entre autres, les détracteurs de CrowdTangle estiment que cet outil n'est pas particulièrement utile pour les chercheurs, puisqu'il est difficile de rechercher des tendances non identifiées à l'avance[52]. Par ailleurs, de nombreux projets de recherche requièrent des données spécifiquement non publiques. D'autres initiatives parallèles à CrowdTangle, ou une révision de ce que ce dernier propose, sont les bienvenues. Puisque les utilisateurs passent d'une plateforme à l'autre, puisque la présence des réseaux extrémistes s'étend sur différentes plateformes, et puisque les contenus passent de plus en plus d'une plateforme à l'autre, des outils transversaux permettraient de stimuler la recherche et favoriseraient la présence de plus de disciplines et de chercheurs analysant les différents enjeux sociaux et politiques qui découlent des dynamiques en ligne de l'extrémisme : d'autres initiatives sont donc nécessaires et appréciées.

50 « Global Internet Forum to Counter Terrorism: Evolving an Institution », 2020, https://www.gifct.org/about/.
51 L'accès total à CrowdTangle n'est ouvert qu'à certaines entreprises et organisations qui satisfont aux critères établis. Toutefois, l'extension CrowdTangle Link Checker, disponible sur Chrome, est accessible pour toutes les parties intéressées. Cette extension montre le nombre de fois où une URL a été partagée, quelles pages publiques ou comptes l'ont partagé et les données d'interaction pour ces publications.
52 Hegelich, « Facebook needs to share more with researchers ».

5 Remarques finales

Certains chercheurs évitent encore de travailler avec des données tirées de médias sociaux ou s'engagent dans des projets de recherche en n'accordant que peu ou pas d'attention aux questions de protection des données et aux principes éthiques. Pour vaincre les hésitations des chercheurs, ce rapport a apporté un éclairage sur les principales questions éthiques et exigences relatives à la protection des données que rencontrent les scientifiques lorsqu'ils travaillent avec des données à caractère personnel tirées des médias sociaux, tout en présentant les difficultés et limites de ces travaux. Malgré ces obstacles, nous ne pouvons ni ne devons éviter d'analyser les données de l'univers numérique. Le monde virtuel et le monde réel ont toujours été étroitement reliés. Il est donc inévitable de les prendre tous deux en compte pour mieux comprendre certains phénomènes. Notre objectif est ainsi d'encourager d'autres chercheurs à travailler avec des données tirées des médias sociaux. C'est pourquoi ce rapport a également mis en lumière les opportunités existantes.

Les chercheurs doivent, lorsque c'est possible, s'acquitter de leurs obligations et de leurs responsabilités et atténuer tous les risques pesant sur les sujets de recherche. Ils doivent par ailleurs obtenir un consentement éclairé lorsque cela est possible, supprimer les informations facilement identifiables et conserver les informations relatives à l'acquisition du consentement jusqu'à la diffusion du projet. Ils doivent tenir compte des règles d'éthique et de protection des données à toutes les étapes d'un projet de recherche (depuis sa conception jusqu'à la diffusion des résultats et lors de la manipulation des données une fois le projet terminé).

Les données qui intéressent les chercheurs sont issues de différentes plateformes. Il existe autant de politiques de confidentialité appliquées par les différentes sociétés technologiques que de plateformes. Si l'on retrouve certaines similitudes – par exemple, Facebook et Twitter appliquent les dispositions du RGPD où que se trouvent leurs utilisateurs dans le monde – il existe aussi des différences, dont certaines ont été abordées dans cet article. Bien que l'accent ait été mis ces dernières années sur l'adoption de politiques plus ouvertes, ne serait-ce qu'en raison d'une multiplication des exigences et des pressions exercées sur les sociétés exécutant les plateformes, les opportunités offertes aux chercheurs ne sont pas toujours clairement formulées. Il est urgent d'améliorer encore les droits et l'accès des chercheurs sur l'ensemble des plateformes. Les sociétés technologiques doivent accorder de nouvelles concessions aux chercheurs, même s'il y a déjà eu des évolutions positives. Ces derniers doivent en même temps exploiter davantage ce que les sociétés technologiques leur proposent déjà – et ce, dans le respect des principes juridiques et éthiques de base décrits dans ce rapport, qui constituent un moteur plus qu'un frein lorsqu'ils sont utilisés correctement dans les plans de recherche.

Contexte politique

Cette section a été rédigée par Armida van Rij et Lucy Thomas, toutes deux adjointes de recherche au Policy Institute du King's College, à Londres. Elle fournit un aperçu du contexte politique dans lequel s'inscrit ce rapport.

Introduction

La recherche sur les contenus à caractère terroriste et/ou extrémiste pose depuis plusieurs dizaines d'années des questions difficiles sur la légalité, la moralité et les aspects pratiques pour les chercheurs, les gouvernements, les militants et les forces de l'ordre. Nous avons d'une part des lois qui régissent la protection des données et des obstacles auxquels doivent faire face les chercheurs lorsqu'ils manipulent des données à caractère personnel. D'autre part, nous avons des lois relatives à la lutte contre le terrorisme et à l'usage qui peut être réservé aux données à caractère terroriste et extrémiste à des fins de recherche. Tout cela donne naissance à un secteur de plus en plus complexe pour les chercheurs, parsemé de risques pour eux-mêmes et pour autrui.

L'approche adoptée ici diffère légèrement de celle adoptée dans les rapports précédents : nous présenterons, dans un premier temps, le contexte législatif relatif à la protection des données à caractère personnel dans huit pays. Nous donnerons ensuite une vue d'ensemble de la lutte antiterroriste dans le neuvième pays, le Royaume-Uni, puis aborderons certaines des questions complexes que pourront se poser les chercheurs intéressés par la recherche sur le terrorisme.

Protection des données sur les médias sociaux : relever les défis, évaluer les nouvelles évolutions

Canada

Le Commissariat à la protection de la vie privée du Canada est chargé de protéger et de promouvoir le droit à la confidentialité des données des individus. Il a pour mandat d'assurer le respect de la loi sur la protection des renseignements personnels, qui régit les pratiques de traitement des renseignements personnels mises en œuvre par les organismes fédéraux, et de la loi sur la protection des renseignements personnels et les documents électroniques (LPRPDE), qui régit la protection des renseignements personnels dans le secteur privé. La LPRPDE est une loi fédérale, mais les provinces d'Alberta, de Colombie-Britannique et de Québec ont adopté leurs propres lois sur la confidentialité des données, dont le fond est similaire[53].

53 « Survol de la LPRPDE », Commissariat à la protection de la vie privée du Canada. Disponible à l'adresse : https://www.priv.gc.ca/fr/sujets-lies-a-la-protection-de-la-vie-privee/lois-sur-la-protection-des-renseignements-personnels-au-canada/la-loi-sur-la-protection-des-renseignements-personnels-et-les-documents-electroniques-lprpde/lprpde_survol/

Globalement, la LPRPDE oblige les organisations privées à « obtenir le consentement des personnes lorsqu'elles recueillent, utilisent ou communiquent des renseignements personnels les concernant », et à respecter les exigences légales relatives à la protection de ces données. En vertu de la LPRPDE, pour s'assurer de la protection des droits relatifs aux données, les entreprises doivent suivre dix « principes relatifs à l'équité dans le traitement de l'information », tels que la responsabilité, le consentement, la limitation de la collecte, la limitation de l'utilisation, de la communication et de la conservation, l'exactitude et les mesures de sécurité[54].

En vertu de la LPRPDE, le Commissariat peut mener des tours d'horizon pour examiner les nouvelles technologies et leurs effets sur les droits relatifs aux données des Canadiens[55], mais a également des pouvoirs coercitifs en cas de violation des règles en matière de protection des données. Il détient ainsi des pouvoirs d'enquête et peut prendre des mesures de dissuasion financière – les entreprises qui ne déclarent pas les violations de données au Commissariat peuvent se voir imposer une amende allant jusqu'à 100 000 dollars. Comme celle applicable en Nouvelle-Zélande, cette amende est beaucoup moins élevée que dans d'autres régions, à l'instar de l'amende de 20 millions d'euros (ou jusqu'à 4 % du chiffre d'affaires annuel total) prévue par le RGPD.

En novembre 2020, le ministre canadien de l'Innovation, des Sciences et de l'Industrie a proposé une nouvelle loi visant à protéger les données personnelles. Dans un communiqué de presse, le ministère a cité la pandémie de coronavirus comme contexte dans lequel moderniser et mettre à jour la législation sur la protection de la vie privée, compte tenu de la multiplication du nombre d'individus utilisant les technologies pour communiquer entre eux[56].

La loi proposée, intitulée loi sur la mise en œuvre de la Charte du numérique, impose au secteur privé – qui englobe les plateformes de médias sociaux – de nouvelles règles applicables au respect de la vie privée. Elle prévoit un renforcement des pouvoirs de surveillance et d'exécution en cas de non-respect – jusqu'à 5 % des recettes ou 25 millions de dollars – et exige des entreprises qu'elles fassent preuve de transparence concernant leur usage des algorithmes et de l'intelligence artificielle. En vertu de cette loi, « les entreprises devront faire preuve de transparence au sujet de l'utilisation qu'elles font de tels systèmes dans le but de faire des prédictions et des recommandations ou de prendre des décisions ayant des répercussions importantes sur telle ou telle personne. Ces personnes auront le droit de demander que l'entreprise explique de quelle manière elle en est venue à ces prévisions, recommandations ou décisions en s'appuyant sur un système automatisé de prise de décisions, et aussi comment cette information a été obtenue »[57]. La loi relative à la protection des données adoptée en 2012 par le Ghana contient des clauses similaires (voir ci-dessous).

54 *Ibid.*
55 « Recherche », Commissariat à la protection de la vie privée du Canada. Disponible à l'adresse : https://www.priv.gc.ca/fr/mesures-et-decisions-prises-par-le-commissariat/recherche/
56 « Nouveau projet de loi pour protéger la vie privée des Canadiens et accroître leur contrôle sur leurs données et leurs renseignements personnels », Gouvernement du Canada, 17 novembre 2020. Disponible à l'adresse : https://www.canada.ca/fr/innovation-sciences-developpement-economique/nouvelles/2020/11/nouveau-projet-de-loi-pour-proteger-la-vie-privee-des-canadiens-et-accroitre-leur-controle-sur-leurs-donnees-et-leurs-renseignements-personnels.html
57 « Fiche de renseignements : Loi de 2020 sur la mise en œuvre de la Charte du numérique », Gouvernement du Canada. Disponible à l'adresse : https://www.ic.gc.ca/eic/site/062.nsf/fra/00119.html

Commission européenne

Dans le cadre de son initiative visant à « préparer l'Europe à son entrée dans l'ère numérique », la Commission européenne s'est employée à réguler les nombreuses facettes des services numériques, y compris les aspects relatifs à la confidentialité et à la protection des données à caractère personnel. La confidentialité des données est régie, au sein de l'Union européenne, par le Règlement général sur la protection des données (RGPD). Celui-ci est entré en vigueur en 2016. Il « protège le droit fondamental des citoyens à la protection des données les concernant lorsque ces données sont utilisées par les services répressifs à des fins répressives » et « garantit notamment la protection des données à caractère personnel des victimes, des témoins et des suspects et facilite la coopération transfrontière dans la lutte contre la criminalité et le terrorisme »[58]. Il est essentiel de signaler qu'il est opposable aux sociétés qui opèrent sur le marché européen, où que se trouve leur siège social. Cela signifie que les sociétés comme Google doivent elles aussi respecter les principes du RGPD, au risque de recevoir une amende et/ou de faire l'objet de poursuites.

Parallèlement au RGPD, un Contrôleur européen de la protection des données a également été mis sur pied. Il s'agit d'un organe indépendant de l'Union européenne qui veille au respect du RGPD et traite toutes les plaintes y afférentes[59].

En plus d'assurer la protection des droits des citoyens européens, le RGPD donne aux autorités compétentes les moyens d'assurer le respect de ses dispositions et vise à renforcer la responsabilité des acteurs manipulant des données à caractère personnel. Depuis 2018, date à laquelle tous les États membres de l'UE ont dû mettre en application le RGPD, des milliers de plaintes ont été déposées et des centaines d'amendes ont été infligées pour non-conformité au règlement. L'amende de 50 millions d'euros infligée à Google par la France pour « manque de transparence, information insatisfaisante et absence de consentement valable pour la personnalisation de la publicité » est probablement l'une des affaires les plus médiatisées à cet égard[60].

Il existe, parallèlement au RGPD, une Directive en matière de protection des données dans le domaine répressif. La Directive 2016/680 porte sur le traitement par les services répressifs des données à caractère personnel d'un suspect, d'un témoin ou d'une victime[61]. Toutefois, la frontière entre les champs d'application du RGPD et de la Directive 2016/680 est ténue, le risque étant qu'une opération de traitement des données soit considérée comme relevant du RGPD dans certains États membres de l'UE et de la Directive dans d'autres[62].

Citons enfin la Directive européenne sur la sécurité des réseaux et des systèmes d'information dans l'Union, qui prévoit des mesures législatives pour stimuler la cybersécurité[63]. Celle-ci

58 https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_fr
59 Voir https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_fr
60 Voir https://www.bbc.com/news/technology-46944696
61 Voir https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_fr
62 Voir https://www.tandfonline.com/doi/pdf/10.1080/13600869.2017.1370224?needAccess=true, p. 253
63 Voir https://ec.europa.eu/digital-single-market/en/network-and-information-security-nis-directive

prévoit notamment le renforcement de l'état de préparation des États membres, une coopération accrue entre les États membres, et le renforcement des infrastructures essentielles au sein de l'UE[64].

Si la réglementation de la confidentialité des données présente certaines limites en matière de compétence, la Commission européenne cherche également à mettre en œuvre son Règlement « vie privée et communications électroniques » (en remplacement de la Directive du même nom)[65]. Ce règlement vise à protéger la vie privée des citoyens sur les plateformes en ligne comme les applications de messagerie. Le Parlement européen a adopté ledit Règlement, mais les débats se sont enlisés au niveau du Conseil de l'UE[66]. Certains soutiennent que l'accent porté sur la protection des données est contraire à la législation de l'UE en matière de lutte contre le terrorisme[67].

France

Le RGPD et la Directive européenne sur la sécurité des réseaux et des systèmes d'information dans l'Union sont respectivement entrés en vigueur en France, État membre de l'UE, en mai 2018 et en 2019. Lorsque le Conseil de l'UE sera en mesure de conclure les négociations sur le règlement « vie privée et communications électroniques », s'il y parvient, ce texte régira la protection des données des citoyens européens parallèlement au RGPD et à la Directive sur la sécurité des réseaux et des systèmes d'information.

Cette dernière impose la création d'une autorité chargée de la protection des données. En France, cette autorité est la Commission Nationale de l'Informatique et des Libertés (CNIL), qui a déjà infligé des amendes à Google et à d'autres acteurs pour violation du RGPD.

Ghana

Les règles phares du Ghana relatives à la confidentialité des données figurent dans la loi de 2012 relative à la protection des données. À l'instar des textes de loi canadien et néo-zélandais, la loi ghanéenne met en place une Commission de protection des données (DPC) dotée de pouvoirs de surveillance et d'exécution pour garantir le respect des responsabilités émises dans la loi[68].

La loi de 2012 relative à la protection des données traite à la fois des organes des secteurs public et privé contrôlant les données, et les oblige à respecter huit principes de protection des données, dont la responsabilité, la fourniture d'informations sur le but recherché et la transparence.[69]. Tout comme dans d'autres pays, la DPC a le pouvoir d'infliger des amendes aux organes contrôlant les données qui ne respectent pas les responsabilités qui leur incombent en vertu de la loi.

64 *Ibid.*
65 Voir https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2020:0568:FIN:FR:PDF
66 Voir https://www.europarl.europa.eu/legislative-train/theme-a-europe-fit-for-the-digital-age/file-jd-e-privacy-reform
67 Voir https://www.coe.int/en/web/commissioner/-/human-rights-in-europe-should-not-buckle-under-mass-surveillance
68 Voir https://www.dataprotection.org.gh/
69 « The Data Protection Principles », Commission de protection des données. Disponible à l'adresse : https://www.dataprotection.org.gh/data-protection/data-protection-principles

L'un des aspects les plus innovants de la loi relative à la protection des données du Ghana, compte tenu notamment du fait qu'elle est entrée en vigueur en 2012, est une clause donnant aux individus le droit de ne pas être soumis à un processus décisionnel automatisé. Cette clause signifie que « les décisions importantes vous concernant fondées sur vos informations personnelles doivent être prises dans le cadre d'une intervention humaine et ne doivent pas être générées automatiquement, sauf consentement de votre part »[70]. Ce modèle moderne, fondé sur le consentement au traitement automatisé et algorithmique des données, pourrait avoir des effets considérables sur l'accès des chercheurs aux données des médias sociaux et le traitement desdites données par logiciel. Bien que cette clause porte actuellement sur les informations ayant « une incidence significative sur l'individu concerné »[71], la prise éventuelle de mesures par le gouvernement ghanéen pour la renforcer pourrait rendre difficile l'usage de logiciels de scraping automatisé par les chercheurs.

Toutefois, la version actuelle de la loi de 2012 relative à la protection des données porte atteinte aux droits des citoyens ghanéens en matière de données, en disposant que « les données personnelles qui sont traitées à des fins de recherche … peuvent être conservées pour une durée illimitée »[72]. De plus, si « les données sont traitées dans le respect des conditions applicables », alors « les données à caractère personnel qui ne sont traitées qu'à des fins de recherche sont exclues du champ d'application de la présente loi »[73]. Ces dispositions portent atteinte aux droits des individus en matière de données, puisque les chercheurs peuvent traiter les données d'une manière contraire à l'éthique tout en satisfaisant aux exigences minimales en matière de protection des données. La définition large et vague de la « recherche » signifie également que les droits en matière de données peuvent être assez facilement menacés.

Japon

Au Japon, la protection des données est régie par la loi sur la protection des informations personnelles de 2003 (LPIP). La Commission de protection des informations personnelles (CPIP), créée en 2016 pour centraliser des autorités de régulation auparavant isolées, est chargée de faire respecter la LPIP.

La CPIP dispose de pouvoirs de surveillance et d'exécution inférieurs à la moyenne : les violations des données peuvent donner lieu à des amendes ou à des peines de prison, mais les amendes sont très faibles, puisqu'elles ne dépassent pas les ¥300 000 (un peu plus de £2 000, soit environ $2 800)[74]. Par ailleurs, la LPIP n'insiste pas sur les obligations directes incombant aux organismes qui traitent des données à caractère personnel, mais impose plutôt

70 « Data Protection for Individuals », Commission de protection des données. Disponible à l'adresse : https://www.dataprotection.org.gh/data-protection/data-protection-for-individuals
71 Loi relative à la protection des données de 2012, a. 41. Disponible à l'adresse : https://www.dataprotection.org.gh/index.php/resources/downloads/data-protection-act/38-data-protection-act-2012-act-843
72 *Ibid.*, a. 65.
73 *Ibid.*
74 Loi sur la protection des informations personnelles de 2003, a. 56. Disponible à l'adresse : https://www.cas.go.jp/jp/seisaku/hourei/data/APPI.pdf

des mesures de contrôle et d'orientation allégées. Ce point est particulièrement important pour la recherche universitaire, puisque la LPIP s'applique au-delà des frontières japonaises, de façon que toute personne manipulant les données de citoyens japonais – même si cela a lieu hors du pays – n'a l'obligation que de respecter ces mesures allégées.

La LPIP a été révisée et amendée en 2020, ce qui a eu des conséquences importantes. Contrairement à la tendance mondiale au renforcement des droits en matière de données, les amendements de 2020 assouplissent les obligations incombant aux responsables du traitement des données. Pour les informations traitées sous pseudonyme, le but de l'utilisation des données peut être modifié par rapport à l'objectif d'origine, l'obligation d'informer la CPIP d'une violation des données n'existe plus, et les individus n'ont plus le droit d'accéder à leurs données, de les corriger ou de demander à ce qu'il soit mis fin à leur utilisation[75].

Autre revers pour les droits relatifs aux données, les chercheurs ne sont pas tenus de respecter les dispositions de la LPIP, puisqu'elle « ne s'applique qu'aux personnes et entités qui manipulent des informations personnelles dans le cadre de leur entreprise »[76]. En réalité, cela signifie que les citoyens japonais dont les données personnelles sont consultées et traitées par des chercheurs n'ont que peu de droits sur leurs données.

Nouvelle-Zélande

En Nouvelle-Zélande, la protection des informations et données à caractère personnel incombe au Commissariat à la protection de la vie privée (OPC). Ce bureau a été créé en 1993 en vertu de la loi sur la protection de la vie privée adoptée la même année, le premier acte législatif de fond à régir les données à caractère personnel dans le pays. Cette loi régit la façon dont les renseignements personnels sont « collectés, utilisés, divulgués et conservés », de même que les modalités d'accès à ces données[77]. L'OPC exerce à la fois des fonctions réactives et préventives : il enquête sur les plaintes relatives à des cas de violation de la vie privée et s'assure du respect de la loi sur la protection de la vie privée, mais le Commissaire contrôle également les évolutions des technologies émergentes pour comprendre quel impact elles peuvent avoir sur la confidentialité des renseignements personnels[78].

En décembre 2020, une nouvelle loi de protection des renseignements personnels est entrée en vigueur en Nouvelle-Zélande : la loi sur la protection de la vie privée de 2020. Cette nouvelle loi a été proposée « pour répondre à la façon dont les technologies ont révolutionné la manipulation des données

75 « Japan – Data Protection Overview », Data Guidance. Disponible à l'adresse : https://www.dataguidance.com/notes/japan-data-protection-overview
76 *Ibid.*
77 « What is personal information and the Privacy Act? », Data.govt.nz. Disponible à l'adresse : https://www.data.govt.nz/manage-data/privacy-and-security/what-is-personal-identifiable-information-and-the-privacy-act/
78 « What we do », Office of the Privacy Commissioner. Disponible à l'adresse : https://www.privacy.org.nz/about-us/what-we-do/

à caractère personnel »[79], la nature et le volume des données de ce type ayant changé du tout au tout depuis 1993. Malgré ce constat, les changements apportés à la loi de 1993 sont remarquablement peu nombreux ; d'après le Commissaire actuel, ceci s'explique par le fait que « la loi sur la protection de la vie privée est une mesure législative neutre sur le plan technologique, ce qui en assure la résilience face aux évolutions technologiques »[80].

Le principal changement apporté à la nouvelle loi concerne la protection des données à caractère personnel des Néo-zélandais à l'étranger : les informations ne peuvent désormais plus être « divulguées à l'étranger, sauf s'il existe des protections comparables à celles fournies par la loi néo-zélandaise »[81]. La loi de 2020 produit également un « effet extraterritorial » explicite, selon lequel toute entreprise opérant en Nouvelle-Zélande sera soumise aux obligations relatives à la protection des données, même si elle n'est pas physiquement présente sur le territoire[82]. Ces éléments concernant la compétence territoriale sont intéressants, puisque la plupart des sociétés technologiques et de médias sociaux sont établies à l'étranger, principalement aux États-Unis – dont les lois de protection des données sont moins strictes. Avec la multiplication du nombre de pays adoptant des législations de ce type, les États-Unis sont soumis à une pression internationale croissante pour renforcer leurs propres lois sur la protection de la vie privée et respecter ainsi les obligations leur incombant à l'étranger.

La loi sur la protection de la vie privée donne également à l'OPC des pouvoirs d'exécution étendus, en faisant notamment passer l'amende maximale pour violation des principes de confidentialité de $2 000 à $10 000. Cette amende est beaucoup moins élevée que dans d'autres régions, à l'instar de l'amende de 20 millions d'euros (ou jusqu'à 4 % du chiffre d'affaires annuel total) prévue par le RGPD ou l'amende de $10 000 000 maximum applicable en Australie. Par ailleurs, la nouvelle loi de Nouvelle-Zélande ne reflète pas le « droit à l'oubli » du RGPD, grâce auquel les individus peuvent demander la suppression de leurs renseignements personnels[83]. Ce droit est particulièrement important concernant l'éthique des données reliée à la recherche, puisque les utilisateurs publiant des contenus à caractère extrémiste sur les médias sociaux – des contenus qui peuvent être utilisés à des fins de recherche – ont le droit de les supprimer.

79 « Input of the New Zealand Human Rights Commission: OHCHR Report on the Right to Privacy in the Digital Age », Haut-Commissariat des Nations Unies aux droits de l'homme, 10 avril 2018. Disponible à l'adresse : https://www.ohchr.org/Documents/Issues/DigitalAge/ReportPrivacyinDigitalAge/HRC_NewZealand.pdf

80 « Media Release: Privacy Act turns 25 », Office of the Privacy Commissioner, 19 février 2018. Disponible à l'adresse : https://www.privacy.org.nz/assets/Uploads/2018-02-19.pdf

81 « Privacy Act 2020: One Small Step for New Zealand, but No Giant Leaps in Sight », Equal Justice Project, 31 août 2020. Disponible à l'adresse : https://www.equaljusticeproject.co.nz/articles/37tbkho3ex74g87sw2n6yz6beyso4a2020

82 « Privacy 2.0: Key changes in the Privacy Act 2020 », Office of the Privacy Commissioner, 16 juin 2020. Disponible à l'adresse : https://www.privacy.org.nz/blog/key-changes-in-the-privacy-act-2020/

83 « Privacy Act 2020 », Equal Justice Project, 31 août 2020. Disponible à l'adresse : https://www.equaljusticeproject.co.nz/articles/37tbkho3ex74g87sw2n6yz6beyso4a2020

Direction exécutive du Comité contre le terrorisme des Nations Unies

En ce qui concerne le système des Nations Unies, la protection des données s'inscrit dans le cadre du travail de la Conférence des Nations Unies sur le commerce et le développement (CNUCED). La CNUCED s'est penchée sur la nécessité de trouver un équilibre entre protection des données et surveillance, ainsi que sur les difficultés que cela engendre. Elle a décrit comment la décision prise par la Cour européenne de justice dans une affaire très médiatisée « vise à imposer des conditions et restrictions à la surveillance dans tout régime de protection des données en Europe, ce qui pourrait avoir des répercussions sur toutes les régions qui suivent de près le droit européen »[84].

La compétence territoriale est un autre domaine posant des difficultés incroyables, en particulier lorsqu'il s'agit de la protection des données en ligne. La CNUCED note la présence d'une clause d'extraterritorialité dans le RGPD : l'article 3 tente en effet d'assurer une « protection locale des données » qui cible les résidents locaux, où que se situe le siège de l'entreprise concernée[85].

États-Unis

Contrairement à de nombreux autres pays, les États-Unis ne se sont pas dotés d'une loi fédérale centrale en matière de protection de la vie privée. Il existe à la place plusieurs lois qui portent sur des aspects distincts de la confidentialité des données – par exemple, les données de santé sont protégées par la loi de 1996 sur la portabilité et la responsabilité en assurance santé, et les données à caractère personnel détenues par le gouvernement sont soumises à la loi américaine sur la protection de la vie privée de 1974.

De manière décisive, les données à caractère personnel et la confidentialité des données sur Internet aux États-Unis ne sont pas réglementées à l'échelle fédérale. Aux États-Unis, Internet est une sorte de no man's land réglementaire, dans lequel les individus, les groupes, les organisations et les entreprises peuvent accéder aux données et les traiter sans réglementation spécifique des droits y afférents.

À l'heure actuelle, la seule façon de protéger les droits relatifs aux données individuelles sur les plateformes de médias sociaux est au travers de la Commission fédérale du commerce (FTC). Par exemple, en 2019, la FTC a pu imposer à Facebook une amende gigantesque de 5 milliards de dollars pour violation de la vie privée dans le cadre du scandale de Cambridge Analytica[86]. Après enquête, la FTC a condamné Facebook à une amende conformément aux pouvoirs qui lui sont conférés par l'article 5 relatif aux « agissements et pratiques malhonnêtes ou mensongers ». Facebook a partagé les informations personnelles de ses utilisateurs avec des applications

84 Voir https://unctad.org/system/files/official-document/dtlstict2016d1_en.pdf, p. 16.
85 *Ibid.*, p. 20.
86 Julia Carrie Wong, « Facebook to be fined $5bn for Cambridge Analytica privacy violations – reports », *The Guardian*, 12 juillet 2019. Disponible à l'adresse : https://www.theguardian.com/technology/2019/jul/12/facebook-fine-ftc-privacy-violations

tierces téléchargées par les « amis » desdits utilisateurs, ce qui a été considéré comme un agissement malhonnête ou mensonger, puisque de nombreux utilisateurs n'avaient ni conscience de ces pratiques ni la possibilité de ne pas y participer[87]. Ce point juridique est important, puisqu'il signifie que si une entreprise ne divulgue pas d'informations sur le traitement ou la manipulation qu'elle réserve aux données, elle ne peut être tenue responsable en vertu de la clause relative aux « agissements ou pratiques malhonnêtes ou mensongers ».

Plusieurs États, en particulier la Californie, ont adopté des lois visant à protéger la confidentialité des données des consommateurs. Puisque bon nombre des principales sociétés technologiques et de médias sociaux sont basées en Californie, la réglementation relative à la protection des données y revêt une très grande importance. La loi californienne de protection de la vie privée en ligne de 2004 était la première loi exigeant la publication par les sites Internet de leurs politiques de confidentialité, ce qui, de façon décisive, s'étend à tout site Internet auquel ont accès les Californiens. En réalité, cela oblige pratiquement tous les sites Internet américains à s'y conformer.

La loi californienne de protection de la vie privée des consommateurs (CCPA) est entrée en vigueur le 1er janvier 2020. Cette loi marque un tournant pour la protection des données aux États-Unis, puisqu'elle s'applique aux « entreprises à but lucratif qui mènent des activités en Californie » ou qui remplissent d'autres conditions liées aux recettes et aux données des Californiens. En pratique, cela signifie que cette loi est opposable aux grandes sociétés technologiques et de médias sociaux. La CCPA garantit aux individus le droit de savoir quels renseignements personnels sont collectés à leur propos, le droit de supprimer ces informations et le droit de refuser la vente de leurs informations personnelles. Les entreprises sont tenues d'informer les consommateurs sur leurs pratiques en matière de confidentialité[88].

L'adoption de la CCPA et l'amende imposée par la FTC à Facebook témoignent d'une volonté politique de protéger les droits relatifs aux données des individus aux États-Unis. En février 2020, la sénatrice Kirsten Gillibrand a proposé une loi générale de protection des données instaurant une agence fédérale indépendante compétente dans ce domaine[89]. Bien que cette initiative ne suffise pas à garantir les droits et obligations spécifiques de protection de la vie privée de tous les Américains, elle laisse penser que les États-Unis se dirigent peut-être vers la mise en place d'une législation fédérale.

87 « FTC Imposes $5 Billion and Sweeping New Privacy Restrictions on Facebook », Federal Trade Commission, 24 juillet 2019. Disponible à l'adresse : https://www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions

88 « California Consumer Privacy Act », ministère de la Justice de l'État de Californie. Disponible à l'adresse : https://oag.ca.gov/privacy/ccpa

89 « A run-down of US Sen. Gillibrand's proposed Data Protection Act », International Association of Privacy Professionals, 21 février 2020. Disponible à l'adresse : https://iapp.org/news/a/an-run-down-of-sen-gillibrands-proposed-data-protection-act/

Recherche sur les contenus à caractère extrémiste au Royaume-Uni : stratégie Prevent, législation antiterroriste et évolution des politiques

Au lendemain des attentats terroristes du 11 septembre 2001 à New York et au Pentagone, aux États-Unis, de nombreux pays occidentaux ont resserré leurs mesures de sécurité intérieure pour tenter d'éviter d'autres attentats sur leur territoire. La politique antiterroriste occidentale est devenue de plus en plus obsédée par la notion de radicalisation, sur laquelle elle s'est fortement concentrée – le fait que des individus s'identifient peu à peu avec des valeurs terroristes, finissent par les adopter, voire mènent des attaques violentes à des fins terroristes. Ce processus de radicalisation a été attribué à un large éventail de facteurs sociaux et individuels : exposition à l'idéologie, victimisation, aliénation, socialisation, réseaux sociaux, Internet, défaillance des liens familiaux, traumatismes, dénuement social et économique relatif, et « cultures de la violence »[90]. Compte tenu du seul nombre de « voies menant à la radicalisation », les gouvernements en sont venus à « croire qu'ils pouvaient anticiper les attentats terroristes futurs au moyen d'un ensemble d'interventions dans la vie quotidienne »[91].

En 2003, le ministère britannique de l'Intérieur a lancé une stratégie de prévention intitulée « Prevent » dans le cadre de sa stratégie de lutte antiterroriste, CONTEST. Prevent a été révisée et réintroduite en 2011, dans le but de cibler les individus « vulnérables » à la radicalisation[92], en particulier au sein des institutions civiques comme les écoles, les garderies agréées, l'enseignement supérieur, les prisons, les services de probation, les services de santé, les services sociaux et les services d'immigration. La stratégie Prevent occupe « l'espace précriminel »[93] – elle intervient avant qu'une activité criminelle ne soit commise dans l'espoir d'interrompre la radicalisation[94].

Prevent vise à « fournir une assistance et réorienter les individus exposés au risque d'être radicalisés/formés à la commission d'actes terroristes ou en cours de formation/radicalisation, avant qu'ils ne commettent une infraction »[95]. Prevent, qui se définit davantage comme une mesure d'assistance que comme une mesure d'incrimination, est dépeinte comme un programme de protection et non de répression. La mise en application de la stratégie incombe par conséquent aux institutions civiques. Ces institutions, p. ex. les universités, sont obligées d'anticiper, de surveiller et d'intervenir dans les cas éventuels de radicalisation dans le cadre de leur devoir de diligence. Cela signifie également que les employés et les employeurs sont à l'affût d'une liste infiniment vaste, complexe et vague d'indicateurs montrant qu'une personne est vulnérable à la radicalisation. Dans ce contexte, les institutions adoptent tout naturellement une attitude excessivement prudente.

90 Katherine E. Brown et Tania Saeed (2015), « Radicalization and counter-radicalization at British universities: Muslim encounters and alternatives », *Ethnic and Racial Studies*, vol. 38, n° 11, p. 1952–68.
91 *Ibid.*
92 Gouvernement britannique, « Prevent Strategy », juin 2011. Disponible à l'adresse : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/97976/prevent-strategy-review.pdf
93 David Goldberg, Sushrut Jadhav et Tarek Younis (2017), « Prevent: What is Pre-Criminal Space? », *British Journal of Psychology* Bulletin, vol. 41, n° 4, p. 208–11.
94 Il est intéressant de noter que le terme « précrime » a été inventé par Philip K. Dick, auteur de la nouvelle de science-fiction *Minority Report*. Voir : Goldberg, Jadhav et Younis, p. 208–11.
95 Charlotte Heath-Kelly et Erzsébet Strausz, « Counter-terrorism in the NHS: Evaluating Prevent Duty Safeguarding in the NHS », Université de Warwick. Disponible à l'adresse : https://warwick.ac.uk/fac/soc/pais/research/researchcentres/irs/counterterrorisminthenhs/project_report_60pp.pdf

À ses débuts, la stratégie Prevent dans les universités était principalement axée sur les communautés d'étudiants, en particulier musulmans. Les établissements ont commencé à scruter les présentations, conférences et événements organisés par les associations d'étudiants afin de se conformer à la stratégie Prevent et éviter ainsi toute ambiguïté quant à savoir si des croyances extrémistes étaient glorifiées sur les campus. Les exemples d'étudiants musulmans ciblés et interrogés de manière disproportionnée sur les campus sous l'égide de Prevent foisonnent[96] ; un étudiant a par exemple été envoyé devant l'équipe de sécurité de l'Université du Staffordshire pour avoir lu un manuel dans le cadre de son cursus de troisième cycle sur le terrorisme, la criminalité et la sécurité internationale[97]. Près de 2 500 événements dans 300 universités ont été annulés ou modifiés (p. ex., désinvitation d'intervenants) en 2017-18.

Le tableau se complique lorsque l'on s'intéresse aux chercheurs universitaires qui étudient l'extrémisme et le terrorisme. L'exposition aux contenus et valeurs extrémistes et terroristes est beaucoup plus évidente et directe, puisque la recherche suppose souvent d'avoir accès à ces types de contenu et de les collecter, tels que les déclarations officielles publiées par des groupes terroristes, la propagande terroriste (y compris les médias visuels), les publications soutenant des points de vue extrémistes sur les médias sociaux, les forums en ligne, etc. Plus particulièrement, la recherche mettant l'accent sur la collecte de données « issues du terrain », telles que des entretiens avec des terroristes condamnés ou des individus radicalisés, suppose que le chercheur est en contact constant avec des individus identifiés comme ayant des croyances extrémistes ou terroristes.

Ceci ouvre la porte à des questions intéressantes sur la nature du risque dans la recherche : les chercheurs universitaires peuvent-ils et doivent-ils être considérés comme vulnérables à la radicalisation ? Quelles implications cela peut-il avoir d'un point de vue juridique et politique ? Quels effets cela peut-il avoir sur la recherche et les chercheurs ?

Les principales mesures législatives du Royaume-Uni en matière de lutte antiterroriste, pour ce qui concerne la recherche relative à l'extrémisme, sont les lois de 2000 et 2006 sur le terrorisme. Les articles 57 et 58 de la loi de 2000 concernent la possession de matériel laissant « raisonnablement penser que le matériel concerné est détenu en vue de commettre, préparer ou fomenter un acte de terrorisme »[98], ou que les informations concernées « sont susceptibles d'être utiles à une personne commettant ou préparant un acte de terrorisme »[99]. En d'autres termes, la possession d'informations ou de matériel en lien avec l'extrémisme ou le terrorisme est illégale,

96 « The Impact of Prevent on Muslim Communities: A Briefing to the Labour Party on how British Muslim Communities are Affected by Counter-Extremism Policies », The Muslim Council of Britain, février 2016. Disponible à l'adresse : http://archive.mcb.org.uk/wp-content/uploads/2016/12/MCB-CT-Briefing2.pdf ; Barbara Cohen et Waqas Tufail, « Prevent and the normalization of Islamophobia », *Islamophobia: Still a challenge for us all*, Runnymede Trust. Disponible à l'adresse : https://core.ac.uk/download/pdf/161895664.pdf

97 Randeep Ramesh et Josh Halliday, « Student accused of being a terrorist for reading book on terrorism », *The Guardian*, 24 septembre 2015. Disponible à l'adresse : http://www.theguardian.com/education/2015/sep/24/student-accused-being-terrorist-reading-book-terrorism

98 *Loi de 2000 sur le terrorisme*, a. 57. Disponible à l'adresse : https://www.legislation.gov.uk/ukpga/2000/11/section/57

99 *Ibid.*, a. 58.

en particulier si ces informations sont susceptibles d'aider des individus ou des groupes à recruter ou radicaliser d'autres personnes ou à mener des attentats violents.

La loi de 2006 sur le terrorisme s'appuie sur l'infraction de possession créée dans la loi de 2000 et l'étend pour y inclure la diffusion de ce type de matériel (article 1), et crée l'infraction d'apologie du terrorisme (au travers notamment de la possession et de la diffusion de ce type de matériel, article 2). Le premier article renvoie aux individus ou groupes qui cherchent à « encourager, directement ou indirectement, ou inciter de toute autre manière [autrui] à commettre, préparer ou fomenter des actes de terrorisme »[100], y compris en faisant des déclarations qui « font l'apologie de la commission ou de la préparation […] de tels actes »[101]. De plus, tous les citoyens britanniques, y compris les chercheurs, sont susceptibles de faire l'objet de sanctions pour ces infractions, même lorsqu'ils sont situés hors du territoire national[102]. En d'autres termes, un chercheur pourrait être à l'étranger dans le cadre d'une bourse de recherche ou de recherches sur le terrain, et être accusé, en vertu de la loi britannique, d'encourager le terrorisme. Le deuxième article traite de la diffusion de publications à caractère terroriste. Plus précisément, il punit la distribution, la diffusion, le don, la vente, le prêt, l'offre, l'envoi électronique de publications à caractère terroriste ou la fourniture de services permettant à autrui d'obtenir, de lire, d'écouter, de visionner, d'acquérir, d'acheter ou d'emprunter ce type de publications[103].

Les problèmes que cela engendre pour les chercheurs qui enseignent et effectuent des recherches dans les domaines de l'extrémisme et du terrorisme sont évidents. Un professeur qui diffuse une vidéo de propagande de l'État islamique dans le cadre de son cours magistral, par exemple, pourrait être accusé de commettre plusieurs infractions : possession de matériel à caractère terroriste, incitation indirecte à commettre des actes de terrorisme et diffusion de publications à caractère terroriste.

L'affaire du « duo de Nottingham » illustre parfaitement ces affirmations. En mai 2008, Rizwaan Sabir, étudiant en master à l'Université de Nottingham, envoie un e-mail à son conseiller pédagogique, Hicham Yezza, pour préparer sa proposition de recherche de doctorat sur le terrorisme islamique. M. Sabir a auparavant consulté le site Internet du ministère américain de la Justice et téléchargé un document du gouvernement intitulé « Études militaires dans le cadre du djihad contre les tyrans : manuel de formation d'Al-Qaïda » (utilisé dans le cadre du procès contre un groupe responsable d'attentats en Afrique de l'Est)[104]. Ce document est mis à disposition des étudiants via le système de bibliothèque de l'Université, et peut être acheté dans n'importe quelle librairie du Royaume-Uni, comme Waterstones[105]. Un collègue remarque le document sur l'ordinateur de M. Yezza et le signale à l'université,

100 *Loi de 2006 sur le terrorisme*, a. 1.2 (b)(i). Disponible à l'adresse : https://www.legislation.gov.uk/ukpga/2006/11/section/1
101 *Ibid.*, a. 1.3 (a).
102 *Ibid.*, a. 17.
103 *Ibid.*, a. 2.
104 Rizwaan Sabir, « Damages for my unjust "terror" arrest », *Al Jazeera*, 21 septembre 2011. Disponible à l'adresse : https://www.aljazeera.com/opinions/2011/9/21/damages-for-my-unjust-terror-arrest/
105 Voir https://www.waterstones.com/book/military-studies-in-the-jihad-against-the-tyrants/anonymous/9781907521249

qui prévient la police. MM. Sabir et Yezza sont tous deux arrêtés sans mandat en vertu de la loi de 2000 sur le terrorisme, et M. Sabir est placé à l'isolement pendant sept jours[106].

La loi de 2000 a été modifiée au fil du temps et en réponse à plusieurs évolutions politiques et sociales. La première modification de taille a eu lieu en 2015, avec l'adoption de la loi relative à la lutte antiterroriste et à la sécurité, qui a renforcé l'obligation pour les institutions de se conformer à la stratégie Prevent. Les universités ont désormais l'obligation légale « de tenir dûment compte de la nécessité d'empêcher les individus de basculer dans le terrorisme »[107] et doivent se doter de politiques et de procédures précises pour les chercheurs travaillant dans ce domaine. Les amendements de 2015 sont fondés sur une approche basée sur le risque, c'est-à-dire que les institutions sont tenues d'assurer un suivi et une évaluation constants des activités de recherche et de prendre des mesures pour atténuer les risques que ces activités pourraient engendrer. Dans la pratique, de nombreuses universités ont désormais absorbé l'évaluation des risques de la stratégie Prevent dans leurs procédures d'éthique de la recherche[108]. Dans la pratique, les conseils d'évaluation éthique semblent avoir élargi la notion de risque, plaçant la réputation de l'établissement au cœur de leurs préoccupations. On peut considérer que la stratégie Prevent a permis aux conseils d'évaluation des établissements d'enliser les demandes d'évaluation éthique de la recherche – portant sur les « recherches risquées et "politiquement délicates" »[109] – dans un bourbier bureaucratique pesant et complexe dans le but de « frustrer et décourager les menaces potentielles pesant sur la réputation de l'établissement »[110], soulevant, à son tour, de graves inquiétudes sur la liberté académique.

Une deuxième modification majeure a été introduite en avril 2019, avec l'adoption de la loi relative à la lutte antiterroriste et à la sécurité des frontières. Cette loi a élevé les peines maximales pouvant être prononcées pour toutes les infractions mentionnées dans les lois de 2000 et 2006 sur le terrorisme ; par exemple, la peine maximale pour la diffusion de publications à caractère terroriste ont plus que doublé, passant de sept à 15 ans de réclusion[111].

Quatre nouvelles mesures visées par la loi de 2019 ont un impact décisif sur la recherche académique relative à l'extrémisme et au terrorisme :

1. La loi érige en infraction le fait d'obtenir ou de visualiser du matériel à caractère terroriste en ligne[112] ;
2. Elle exclue explicitement les individus exerçant la fonction de journaliste ou menant des travaux de recherche académique

106 Rizwaan Sabir et Hicham Yezza ont été remis en liberté sans inculpation. En 2011, M. Sabir a engagé une procédure judiciaire contre la police du Nottinghamshire pour séquestration et discrimination raciale, affaire qui a par la suite été réglée à l'amiable. Voir Sabir, « Damages for my unjust "terror" arrest ».

107 « Statutory guidance: Revised Prevent duty guidance for England and Wales », ministère britannique de l'Intérieur, mis à jour le 10 avril 2019. Disponible à l'adresse : https://www.gov.uk/government/publications/prevent-duty-guidance/revised-prevent-duty-guidance-for-england-and-wales#c-a-risk-based-approach-to-the-prevent-duty

108 Voir, par exemple, « Oversight of security-sensitive research material in UK universities », Universities UK, novembre 2019. Disponible à l'adresse : https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2019/Oversight-security-sensitive-research-material-guidance-3.pdf

109 Adam Hedgecoe (2015), « Reputational Risk, Academic Freedom and Research Ethics Review », *Sociology*, vol. 50, n 3, p. 495.

110 *Ibid.*

111 *Loi de 2019 relative à la lutte antiterroriste et à la sécurité des frontières*, a. 7. Disponible à l'adresse : https://www.legislation.gov.uk/ukpga/2019/3/section/7

112 *Ibid.*, a. 3.

du champ d'application de l'infraction de collecte d'informations (y compris sur Internet) (article 58 de la loi de 2000 sur le terrorisme)[113] ;

3. Elle érige en infraction le fait pour les citoyens d'entrer ou de rester dans une « zone désignée » en dehors du territoire britannique[114]. Le ministre de l'Intérieur est compétent pour désigner ces zones au cas par cas, aux « fins de protéger le public contre la menace terroriste »[115] ;
4. Elle ajoute un article à la loi de 2006 sur le terrorisme, afin d'inscrire, au titre des infractions, la diffusion de publications à caractère terroriste à l'étranger (cela ne concernait auparavant que l'apologie du terrorisme).

Le point 2 ci-dessus – l'exclusion des chercheurs de l'infraction de collecte de matériel à caractère terroriste (y compris sur Internet) – semble, à première vue, être une évolution positive, qui rétablit la liberté académique d'effectuer des recherches dans les domaines du terrorisme et de l'extrémisme sans crainte d'éventuelles répercussions légales. Toutefois, il est essentiel de noter que, si les chercheurs peuvent désormais explicitement se soustraire à l'article 58 de la loi de 2000 sur le terrorisme (possession de matériel à caractère terroriste), ils ne jouissent pas de protection légale explicite vis-à-vis des articles 1 (apologie du terrorisme) ou 2 (diffusion de matériel à caractère terroriste) de la loi de 2006 sur le terrorisme[116].

En pratique, cela signifie que les chercheurs accédant ou collationnant du matériel à caractère extrémiste en ligne à des fins de recherche ou d'enseignement bénéficient clairement d'une défense juridique. Mais s'ils ont extrait ces documents d'articles de revues ou d'ouvrages universitaires, ou s'ils les ont montrés en classe comme des exemples de propagande extrémiste sans dénoncer explicitement les groupes qui en sont à l'origine, ils pourraient se retrouver dans une situation juridique difficile. De plus, en vertu de la loi de 2000 sur le terrorisme, ils peuvent être arrêtés sans mandat et détenus pour une période de 28 jours pendant que des accusations sont portées contre eux, comme ça a été le cas pour Rizwaan Sabir et Hicham Yezza.

De la même manière, les chercheurs effectuant des travaux sur le terrain ou collectant des données à l'étranger peuvent être visés par cette nouvelle législation. Si un chercheur effectue des travaux sur le terrain à l'étranger, ou prévoit de le faire, dans une zone déclarée par le ministre de l'Intérieur comme une « zone désignée », il pourrait se rendre coupable d'une infraction s'il entre dans ladite zone ou décide d'y rester.

113 *Ibid.*, a. 7.
114 *Loi de 2000 sur le terrorisme*, a. 58(b). Disponible à l'adresse : https://www.legislation.gov.uk/ukpga/2000/11/section/58B
115 « Counter-Terrorism and Border Security Bill: Supplementary Delegated Powers Memorandum », ministère britannique de l'Intérieur, 5 septembre 2018. Disponible à l'adresse : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/739267/Supplementary-Delegated-Powers-Memo-designated-area-offence.pdf
116 « Les articles 2 et 3 de la loi de 2006 contre le terrorisme proscrivent également la diffusion de publications à caractère terroriste, y compris par voie électronique, et définissent très largement les « publications à caractère terroriste » et les « déclarations » qui peuvent être interprétées comme encourageant ou incitant à la commission, à la préparation ou à l'instigation d'actes de terrorisme. *La recherche académique ne peut être considérée comme une défense en vertu de la loi de 2006 contre le terrorisme* [accentuation des auteurs] ». « Oversight of security-sensitive research material in UK universities », Universities UK, novembre 2019. Disponible à l'adresse : https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2019/Oversight-security-sensitive-research-material-guidance-3.pdf

Globalement, la situation juridique des chercheurs travaillant dans la recherche sur le terrorisme et l'extrémisme reste floue. Si la loi adoptée l'année dernière signale que le gouvernement comprend que les chercheurs seront en possession de matériel compromettant, d'autres textes de loi sont encore opposables à ces derniers. La loi, les politiques et les recherches reflètent et ancrent toutes le climat politique actuel ; en cette époque de montée de l'islamophobie et de soutien généralisé à la surveillance et au maintien de l'ordre par anticipation, la stratégie Prevent et les lois sur le terrorisme sont très révélatrices de ces deux phénomènes.

Un facteur important à prendre en compte lorsque nous évaluons la probabilité que les chercheurs soient touchés par la gouvernance et la législation antiterroristes du gouvernement britannique (comme la stratégie Prevent) est l'effet disproportionné que ces dernières ont sur les musulmans. L'« extrémisme islamiste » représente 65 % des signalements effectués en vertu de Prevent ; cela signifie que les « musulmans ont une chance sur 500 d'avoir été signalés dans le cadre de Prevent l'année dernière, soit 40 fois plus qu'une personne non musulmane ». De même, plus de la moitié (54 %) des arrestations pour terrorisme effectuées en 2017 au Royaume-Uni ont concerné une personne « d'apparence sud-asiatique »[117]. La réalité statistique est la suivante : les étudiants et les chercheurs racialisés et minorisés comme musulmans sont beaucoup plus menacés – signalés via Prevent ou incriminés – par la zone d'ombre juridique.

Jusqu'à présent, nous avons vu que les musulmans avaient été injustement ciblés par les législations antiterroristes sur les campus. Toutefois, en novembre 2020, le plus grand nombre de signalements liés à l'extrême droite a été enregistré : 43 %, contre 30 % pour l'extrémisme islamiste[118]. Cette évolution pose des questions intéressantes sur le profilage racial et la recherche sur l'extrémisme et le terrorisme : les chercheurs non musulmans seront-ils considérés comme « vulnérables » et « exposés au risque de radicalisation » lorsqu'ils effectueront des recherches sur le terrorisme suprémaciste blanc ? Si oui, quelles réponses sociales et politiques cette situation va-t-elle déclencher ? Un nombre croissant de détracteurs considèrent la stratégie Prevent et la législation antiterroriste comme un mécanisme visant à surveiller et contrôler les communautés musulmanes sur les campus et ailleurs[119]. Si cette fonction est désormais établie, quel rôle la stratégie va-t-elle jouer maintenant, le cas échéant ?

117 « Operation of police powers under the Terrorism Act 2000 and subsequent legislation: Arrests, outcomes, and stop and search, Great Britain, financial year ending 31 March 2017 », ministère britannique de l'Intérieur, juin 2017. Disponible à l'adresse : https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/619016/police-powers-terrorism-mar2017-hosb0817.pdf

118 Jamie Grierson et Dan Sabbagh, « Largest number of Prevent referrals related to far-right extremism », *The Guardian*, 26 novembre 2020. Disponible à l'adresse : https://www.theguardian.com/uk-news/2020/nov/26/just-one-in-10-prevent-referrals-found-at-risk-of-radicalisation

119 Fahid Qurashi (2018), « The Prevent strategy and the UK "war on terror": embedding infrastructures of surveillance in Muslim communities », *Palgrave Communications*, vol. 4, n° 17 (2018) ; « Liberty's written evidence to the JCHR's Inquiry on Freedom of Expression in Universities », Liberty, décembre 2017. Disponible à l'adresse : https://www.libertyhumanrights.org.uk/wp-content/uploads/2020/02/Libertys-Evidence-to-the-JCHRs-Inquiry-into-Freedom-of-Expression-in-Universities-Dec-2017.pdf

Conclusion : un paysage mondial en pleine mutation

Les enjeux éthiques et juridiques auxquels sont confrontés les chercheurs souhaitant accéder à des données personnelles et les traiter sont complexes et variés. Face aux évolutions juridiques et politiques rapides à l'échelle nationale et internationale, les perspectives pour les chercheurs travaillant dans le domaine de l'extrémisme et le terrorisme sont marquées par le changement et l'incertitude.

En ce qui concerne l'accès aux données à des fins de recherche, il existe une tendance générale au renforcement des lois de protection des données visant à mieux préserver les droits des individus en matière de données (à quelques exceptions près, comme le Japon). Cela signifie que les chercheurs sont susceptibles, à l'avenir, d'être confrontés à un nombre croissant d'obstacles concernant les données à leur disposition et les façons dont ils peuvent les traiter et les utiliser. Alors que les entreprises cherchent à se maintenir au courant du patchwork de lois nationales et supranationales en vigueur, les plateformes de médias sociaux doivent constamment mettre à jour et réviser leurs politiques de confidentialité. Puisque les conséquences d'un manquement à ces obligations – comme en témoigne l'amende de 5 milliards de dollars imposée à Facebook par la Commission fédérale du commerce américaine – s'aggravent de plus en plus, il est possible que les plateformes optent pour une approche plus conservatrice pour préserver leur sécurité financière et leur réputation.

En même temps, les perspectives juridiques et politiques pour les chercheurs travaillant dans les domaines de l'extrémisme et du terrorisme sont également incertaines. Au Royaume-Uni, un climat de maintien de l'ordre par anticipation justifié par des menaces à la sécurité nationale a produit un environnement politique dans lequel les chercheurs risquent d'être considérés comme des criminels en raison de leur seule proximité avec certains types de matériels. Alors que la « guerre contre le terrorisme » se poursuivait dans les années 2000, le contexte législatif du Royaume-Uni reflétait une approche de la lutte contre le terrorisme fondée sur l'ordre public qui a donné naissance à plusieurs évolutions juridiques limitant le matériel auquel les chercheurs pouvaient accéder, dont ils pouvaient discuter et sur lequel ils pouvaient écrire, enseigner et publier. Toutefois, alors que l'attention mondiale se détourne de la prétendue « menace islamiste » pour se porter davantage sur la suprématie blanche violente, les cadres politique et juridique existants, conçus pour cibler une minorité, deviennent problématiques. Les mécanismes actuels de dénonciation par les collègues et les pairs dans les universités reposaient en grande partie sur le profilage racial ; comment ces approches fonctionneront-elles au moment de considérer des chercheurs occidentaux travaillant dans les domaines de la suprématie blanche ?

La nature et la portée de la recherche relative à l'extrémisme et au terrorisme en Occident peuvent changer considérablement dans les années à venir, compte tenu de l'évolution actuelle du contexte mondial. Il sera par exemple peut-être plus difficile de mener des analyses quantitatives de grande échelle si les lois relatives à la confidentialité des données et les politiques de confidentialité des sociétés sont renforcées, ou d'accéder aux individus impliqués dans des groupes ou des actes à caractère terroriste. Cela pourrait signifier une évolution des méthodologies à la disposition

des chercheurs en matière d'extrémisme, par exemple vers des méthodes plus ciblées sur le plan qualitatif, à plus petite échelle ou centrées sur l'ethnographie numérique[120]. Bien que ces modifications soient alarmantes, elles pourraient aussi fournir des avantages considérables : des face-à-face plus intimes et nuancés avec l'extrémisme et le terrorisme qui peuvent mieux refléter les complexités et contradictions des individus ayant des croyances extrémistes en ligne.

120 Voir, par exemple : Sarah Pink *et al.* (dir.), *Digital Ethnography: Principles and Practice* (2015), SAGE Publications Ltd.

Global Network
on Extremism & Technology

COORDONNÉES

Pour toute question, demande d'information et demande de copies supplémentaires du présent rapport, contacter :

ICSR
King's College London
Strand
Londres WC2R 2LS
Royaume-Uni

T. **+44 20 7848 2098**
E. **mail@gnet-research.org**

Twitter : **@GNET_research**

Ce rapport peut, comme toutes les autres publications du GNET, être téléchargé gratuitement à partir du site Internet du GNET : www.gnet-research.org.