

Humanités numériques et historiens analogiques

Pieter Lagrou, Delphine Lauwers et Ornella Rovetta

La prophétie d'Emmanuel Le Roy Ladurie en 1967, voici un demi-siècle, selon laquelle « l'historien de demain sera programmeur ou il ne sera plus », montre à quel point l'expertise sur les choses à venir n'est pas le point fort des historiens⁴³. Dix ans avant la commercialisation du premier *Personal Computer*, l'informatique qui servit de référence à Le Roy Ladurie se trouvait dans les sous-sols des Centres de Calcul avec leurs stocks de cartes perforées et leur papier d'imprimante en accordéon. En 1965 déjà, Gordon Moore avait formulé sa *Moore's Law* sur le développement exponentiel de la puissance de calcul des ordinateurs, mais ni Moore et encore moins Le Roy Ladurie ne pouvaient imaginer que la révolution informatique serait caractérisée par les usages totalement imprévus de cette puissance de calcul, non pas tant pour les *Big Data*, mais pour le partage de photos et de documents, pour des jeux vidéo et, finalement, surtout, sans doute, pour le développement de logiciels intuitifs et *user friendly*. Tout comme, au début de l'invention de l'automobile, il fallait être mécanicien pour conduire une voiture, au début de l'informatique, il fallait être *programmeur* pour utiliser un ordinateur. Depuis, la technologie a évolué et l'utilisateur peut heureusement se débrouiller sans devoir impérativement savoir démonter un carburateur ou apprendre à écrire en MS-DOS ou FORTRAN.

Sur le développement de l'informatique, Le Roy Ladurie avait donc à peu près tout faux et on ne peut pas lui en tenir rigueur. Sur le développement de la discipline historique, il était plus encore « à côté de la plaque ». Ce ne sont pas les grands chantiers sur l'évolution du prix du blé sur les foires de Champagne ni les promesses de l'histoire quantitative qui ont constitué l'horizon d'avenir de la discipline dans les années 1967-2017, mais bel et bien l'approche qualitative – les *small data* de la microhistoire et la méthode analogique des jeux d'échelle et des histoires croisées. La prophétie de l'auteur de *Montailou* sonne donc souvent comme un avertissement quand il s'agit de proclamer une fois de plus l'aube dorée des humanités numériques, les opportunités illimitées des *Big Data* et, par la même occasion, le crépuscule de tous ceux qui rateront leur test de *digital literacy*, ces analphabètes, ignares et borgnes de demain. Les humanités numériques sont-elles alors une nouvelle discipline réservée aux initiés, armés de compétences techniques comme la maîtrise de logarithmes de probabilité inaccessibles aux plus communs des historiens et autres chercheurs mortels en sciences sociales? L'avenir appartient-il, cette fois-ci pour de bon, aux *programmeurs* et à leurs petits-enfants *geeks*?

De quoi parlons-nous exactement, quand nous parlons de *digital literacy*? Les sciences humaines et sociales sont massivement passées à l'ère digitale, le plus souvent sans s'en rendre compte. La plupart des sources sur lesquelles nous travaillons se présentent déjà sous forme numérique et cela transforme

43. Émilien Ruiz, "Les historiens seront-ils finalement programmeurs?", en ligne in *La boîte à outils des historiens*, 22 septembre 2011 <http://www.boiteaoutils.info/2011/09/les-historiens-seront-ils-finalement/>, consulté le 27 septembre 2017.

notre façon de travailler, ainsi que notre rapport avec ces sources.

Trouver la petite bête

Il s'agit d'abord d'une question d'accessibilité. Nous utilisons plus facilement le livre intégralement accessible sur *Google Books* et l'article en ligne que celui qui nécessite qu'on se déplace physiquement vers une bibliothèque, qui l'aurait entreposé en magasin fermé et pourrait nous le remonter une heure, un jour ou un voyage d'études plus tard. Il en va de même, de plus en plus, pour les documents d'archive. Même dans le cas de documents originaux que l'on consulte physiquement, plus leur lieu de conservation est éloigné, plus on aura tendance à les photographier au lieu de les lire et de les retranscrire sur place, générant ainsi une archive numérique parallèle. La photographie digitale – le plus souvent à l'aide d'un *smartphone* – combinée à la révolution économique des missions de recherche avec vols *low cost* et logements *AirBNB* font en sorte qu'il est aujourd'hui souvent plus aisé pour un étudiant en master de croiser des sources provenant d'archives de différents pays qu'il ne l'était il y a une dizaine d'années encore pour un aspirant FNRS dans le cadre d'un séjour de recherche de longue durée. La politique du *Bundesarchiv*, par exemple, qui permet graduellement aux chercheurs de photographier certaines collections, plutôt que de les obliger à en confier la reproduction à l'entreprise *Selke GmbH* – qui, en bon monopoliste, n'est ni pressée ni bon marché –, participe à une révolution digitale qui

promet la gratuité, l'immédiateté et donc la démocratisation de la recherche. D'autres initiatives favorisent la consultation numérique, comme les collections digitales déclassifiées mises en ligne par le Département d'État américain dans une *Virtual Reading Room* ou les archives de procès des juridictions internationales⁴⁴. Ces archives en ligne placent le chercheur face à un enjeu de taille : le référencement de ces sources et leur positionnement dans un ensemble documentaire plus large.

Il s'agit ensuite d'une question de modalités d'usage et d'exploration de ces sources. Une bonne partie des documents accessibles en ligne permettent la reconnaissance textuelle et donc le *full text search*. Il est cruel de penser aux générations d'étudiants en histoire qui ont passé de longs mois dans les bibliothèques à feuilleter, jour par jour, mois par mois et année par année la presse quotidienne pour recenser les occurrences de tel ou tel mot ou sujet, afin de rédiger un mémoire sur la représentation de XYZ dans la presse entre 1952 et 1954. Ce travail peut aujourd'hui être fait en quelques heures seulement par une recherche dans la version digitale des mêmes sources, souvent en limitant les risques d'omettre l'un ou l'autre article par inadvertance ou simple ennui. Même pour le chercheur en bibliothèque ou en archive, les outils techniques existent pour scanner et OCRiser (transformer l'image en un fichier texte) instantanément un grand volume de texte afin d'y retrouver une occurrence, grâce à des logiciels pour *smartphone* proposés par *Scan-*

44. Voir les plus de 175.000 documents actuellement disponibles sur la plateforme du Département d'état en application du *Freedom of Information Act* <https://foia.state.gov/Search/Search.aspx>; pour les juridictions internationales, voir les bases de données des documents judiciaires publics du TPIR http://jrad.unmict.org/index_fr.html et du TPIY <http://icr.icty.org/default.aspx> consultés le 28 septembre 2017.

bot, *Officelens*, *ScanWitr*, *CamScanner*, *Evernote* et même *Dropbox Scans*⁴⁵. Plus la peine de lire les 500 pages d'un livre si vous voulez seulement savoir s'il mentionne Jean Dupuis: il suffit de le scanner et de faire une recherche textuelle dans le pdf avec reconnaissance des caractères, généré en temps réel. Il faudrait être bête pour s'en priver et point besoin de logarithmes: il suffit de savoir se servir d'un téléphone. La digitalisation, avant d'être une révolution épistémologique, permet tout simplement un énorme gain de temps. Pour ces usages, plus vite n'équivaut pas forcément à moins bien. Pouvoir brasser en peu de temps beaucoup de sources permet d'ouvrir son spectre d'analyse, d'identifier rapidement les sources potentiellement intéressantes pour ensuite pouvoir s'y attarder et les analyser avec d'autant plus de temps et profondeur. Marnix Beyen nous rappelle que la chouette de Minerve, image de la sagesse qui, comme l'historien, ne s'envole qu'à la tombée de la nuit, est bel et bien un rapace⁴⁶. Elle a besoin de survoler de vastes surfaces à quelques dizaines de mètres du sol, mais quand elle aperçoit sa proie, elle descend en flèche pour la capturer et prendre tout son temps pour la dévorer. Les possibilités de recherche textuelle offertes par le document digital, pour les sources en ligne comme pour la numérisation *in situ*, permettent de prendre de la hauteur pour gagner en efficacité et pour multiplier les prises. Les charognards, dont les historiens, auraient donc tort de rechigner devant les méthodes de chasse de rapaces comme la chouette de Minerve.

Pour une démocratie numérique

Ces deux formidables avantages de la digitalisation des sources – accessibilité et modalités de recherche textuelle – ont bien sûr des implications profondes sur nos pratiques de recherche, qui méritent réflexion. Commentons par l'accessibilité. Les choix de digitaliser telle ou telle collection ou tel ou tel fonds produisent, par la facilité d'accès et la visibilité que confère le numérique, un effet mécanique déformateur. Les chercheurs pressés par des contrats précaires ou tout simplement casaniers privilégieront ces fonds-là à d'autres dans leurs choix de thématiques de recherche et, au sein d'une même thématique, les sources digitales seront automatiquement davantage sollicitées que les sources en attente de digitalisation⁴⁷. Les bibliothèques, les archives et les bailleurs de fonds pour de grands projets de numérisation disposent ainsi d'un pouvoir inédit et parfois inconscient d'orientation de la recherche. Le débat sur ces choix doit donc être un débat ouvert, citoyen et peut-être surtout pragmatique. Il doit impliquer les chercheurs, les enseignants et les historiens amateurs. Il doit privilégier les collections les plus sollicitées pour lesquelles l'accès numérique constitue une mesure de protection de la source physique, qui risque la détérioration par une manipulation répétée. Il doit privilégier également les collections les plus susceptibles d'intéresser un public international, pour lequel le coût d'une consultation physique peut être prohibitif. Car le fléchage numérique induit par les choix de digitalisation peut aussi être

45. Les auteurs remercient Yoanna Alexiou, Karsten Marhold et Margaux Van Uytvanck de les avoir initiés à ces outils. 46. Marnix Beyen, "Thick Description beyond the Digital Space", in *Humanities*, 2016, 5, 2, doi: 10.3399/H5010002. 47. Voir Kiran Klaus Patel, "Zeitgeschichte im digitalen Zeitalter", in *Vierteljahrsheft für Zeitgeschichte*, 3, 2011, p. 1-21 et Pieter Lagrou, "De l'histoire du temps présent à l'histoire des autres. Comment une discipline critique devient complaisante", in *Vingtième siècle. Revue d'histoire* n° 118, Avril 2013, p. 101-119.

vu sous un autre angle: celui d'une histoire globale, transnationale, partagée et ouverte à tous. On peut débattre de la question de la restitution des Archives Africaines à la République Démocratique du Congo, à l'image des frises du Parthénon du *British Museum*, mais cela n'empêche pas de déjà envisager une digitalisation prioritaire de ces fonds afin de permettre à tout chercheur disposant d'une connexion internet, en Afrique ou ailleurs, de les consulter. On peut avoir des débats plus difficiles sur l'impact social de l'ouverture de telle ou telle collection par rapport à une autre. Le fonds de la Police des Étrangers en est un exemple, si le but poursuivi est de démocratiser les recherches sur les trajectoires migratoires dans le creuset belge en créant une sorte de *Ellis Island* virtuel. Les archives de la justice militaire et du Service des victimes de la guerre aussi, pour répondre à une demande sociale de connaissance des aspects les plus sombres de l'histoire récente⁴⁸. Ce qui pourrait, par exemple, porter un coup de frein à la digitalisation des registres paroissiaux, au grand dam des nombreux généalogistes. Les ressources étant limitées, voire très limitées, il faudra des arbitrages permanents, des ordres de priorités ajustables qui demandent à être débattus en toute transparence. La digitalisation doit être conçue comme un moyen de démocratiser l'accès aux sources, dans un processus qui garantit l'équité, l'authenticité et la traçabilité. Elle doit donc faire l'objet d'un processus de prise de décision démocratique lui aussi, ce qui implique que pour les archives publiques, c'est l'État qui doit en être l'initiateur par l'action des Archives Générales du Royaume. Le recours à des opérateurs pri-

vés doit aussi faire l'objet d'une veille démocratique. Les partenariats exclusifs entre *Universiteit Gent* ou le Mundaneum et *Google* permettent un gain de temps important, mais sont-ils vraiment un modèle durable de développement d'une démocratie digitale?

Lost in cyberspace

Vient ensuite la question des modalités de recherche textuelle. Les moteurs de recherche – d'une simple recherche générique consistant à *googler* un mot jusqu'aux bases de données les plus élaborées – sont autant de raccourcis qui permettent souvent d'accéder à un document sans aucune indication de son contexte documentaire: une page d'un livre, un document d'un dossier ou d'un fonds d'archives. Or, souvent, seul ce contexte permet de saisir le sens de ce qu'on en extrait. La page en question peut donner le contrepoint de tout l'argumentaire que l'auteur développe dans son ouvrage. Le document d'archive peut être tout à fait exceptionnel ou tout à fait représentatif de son « carton » d'origine. Qui est le producteur du document? S'agit-il d'un procès-verbal de police établi sur le lieu du crime peu de temps après sa commission ou d'un témoignage produit lors du procès en appel, huit ans après les faits? La recherche textuelle sélective détruit des collections « organiques », produites par une accumulation lente, où la proximité d'un document avec un autre est en soi un indice de qui l'a écrit, utilisé, reclassé. C'est l'association d'un terme de recherche et d'un document dans une logique d'accès aléatoire qui précipite le chercheur dans un paysage virtuel. Même un document mis en ligne sur un site qui préserve son contexte

48. Voir Koen Aerts, Dirk Luyten, Bart Willems, Paul Drossens, Pieter Lagrou (eds.), *Papy était-il un nazi? Sur les traces d'un passé de guerre/Was opa een nazi? Speuren naar het oorlogsverleden*, Racine/Lannoo, 2017.

archivistique et garantit son authenticité peut ensuite être repris sur d'autres sites bien mieux référencés, perdant ainsi toute trace de leur origine. Le *cyberspace* a son lot de *space litter*, débris de satellites documentaires mis en orbite et ensuite désintégrés à force de collisions dans l'espace encombré des archives numériques. Qui n'a jamais téléchargé de document pdf passionnant qui s'avère ensuite parfaitement inutilisable car, malgré le logiciel bibliographique *Zotero*, on ne trouve plus la trace de sa localisation ou de son producteur ?

Il importe donc pour les chercheurs que les collections numériques soient les plus *analogues* possibles aux collections originales, dans leur structure, dans les liens de proximité qui existent entre documents d'un même fonds et dans leur filiation matérielle au producteur d'archives. La meilleure archive virtuelle concevable est celle qui reproduit l'expérience d'une consultation physique. Ceci est d'autant plus vrai qu'il n'est possible ni d'anticiper les questions de recherche des générations futures d'historiens ni de deviner les usages futurs d'archives du passé. Même le moteur de recherche le plus sophistiqué ou la base de données la plus avant-gardiste risque d'être dépassé avant d'être opérationnel. Les concepteurs des infrastructures de recherche de demain doivent faire preuve de réalisme et de modestie, car l'avant-gardisme technologique ne garantit pas la pérennité de l'accès à ces sources, tant sur le plan des outils que de leur hébergement. Les gadgets futuristes du passé sont touchants pour ce qu'ils révèlent sur les fantasmes d'avenir d'une époque donnée mais ils sont rarement utiles. Trop de projets prestigieux d'humanités numériques débouchent sur des outils éphémères, dont la technologie

est presque instantanément dépassée, dont les mises à jour ne sont pas anticipées et l'hébergement pas garanti.

Vie privée et anonymat sur la toile

La facilité de la recherche textuelle et la production, parfois inconsciente, de bases de données parallèles, soulèvent une autre question pour les utilisateurs et les producteurs de collections numériques, mais aussi pour tous ceux qui publient le résultat de leurs recherches en ligne : celle de la protection de la vie privée. Pour revenir sur le cas de Jean Dupuis : sa vie privée se trouvait relativement bien protégée, enfouie dans les 500 feuilles de papier d'un livre. Ce n'est plus le cas si le même contenu se trouve sous forme numérique sur internet. Dans le premier cas, seul le chercheur le plus acharné pouvait retrouver sa trace. Dans le second, n'importe quel internaute qui saisit son nom dans la fenêtre d'un moteur de recherche le débusque.

Les décisions de justice illustrent bien ces nouvelles dynamiques de recherche. Elles sont, par définition, publiques et ne font jurisprudence qu'après publication sélective. On connaît alors les décisions par le nom des justiciables concernés : l'arrêt « Bosman » pour les transferts de joueurs de football dans l'Union Européenne, ou l'arrêt « Roe v. Wade » pour la dépénalisation de l'avortement aux États-Unis. Or, la plupart des justiciables ne tiennent pas à une telle notoriété. La jurisprudence internationale en matière de mise en ligne d'informations judiciaires est unanime sur le fait qu'il y a une différence fondamentale entre la publication papier, dans les grands volumes de Pasicrisie conservés sur les étagères des bibliothèques de droit et des cabinets d'avo-

cats, et une publication en ligne⁴⁹. Car la facilité et la rapidité de la recherche textuelle dans les bases de données jurisprudentielles permettent à des opérateurs privés d'en faire un usage détourné: ils peuvent constituer des bases de données parallèles et vendre leurs services de *background check* à des institutions financières ou à des employeurs. Les risques liés aux détournements rapides, aisés et systématisés des bases de données judiciaires à des fins de contrôle (octroi de crédits, contrats de travail, secret médical) ont conduit à préconiser l'anonymisation du justiciable lorsque son nom ne constitue pas une donnée fondamentale en matière de jurisprudence. Mais la question se pose très différemment dans les cas de procès dits « historiques », notamment pour crimes de guerre, génocide ou crimes contre l'humanité. La publicité de l'identité des auteurs de ces crimes, ainsi que de leurs victimes, est alors au cœur même de la décision de justice, car elle porte reconnaissance et réparation. L'anonymisation de l'information rendrait ces documents inutilisables pour la recherche historique, et elle escamoterait une dimension fondamentale de ces procès: l'objectif de constituer des archives pour l'avenir. Plusieurs décennies plus tard, ces procès continuent toutefois de produire des effets en sur les processus écriture historique et de mémoire familiale ou collective. Les enjeux liés à l'usage de l'information contenue dans ces archives, à la manière dont elle est exploitée et publiée sont donc aigus.

Car si les nouvelles normes européennes pour la protection de la vie privée s'appliquent exclusivement à des personnes vivantes⁵⁰, les atteintes à la réputation et aux intérêts des ayants droit ou descendants relèvent quant à elles de la diffamation. Ce n'est alors pas le principe de l'accès à l'archive ou à la décision judiciaire qui pose problème, mais bien l'éventuelle intention diffamatoire de l'utilisateur. La simple disponibilité de ces informations ne devrait engager ni les responsables des collections numérisées dans le cadre de projets de recherche ni les chercheurs qui en font un usage scientifique. C'est une mise au point précieuse, car la recherche textuelle dans les publications scientifiques leur donne aussi une visibilité inédite, pour tout utilisateur d'un moteur de recherche. Gare donc quand même aux Jean Dupuis et leurs descendants soucieux de leur réputation.

Deep data mining et les procès belges pour crimes de guerre

Au plan international, les projets de numérisation et de mise en ligne de collections documentaires relatives aux crimes de masse et à leur jugement se sont multipliés. On peut penser ici aux procès qui ont eu lieu après la Seconde Guerre mondiale ou aux archives des juridictions locales rwandaises *gacaca* qui ont été saisies du contentieux de génocide dans les années 2000. En outre, la justice dite transitionnelle inclut de plus en plus dans ses objectifs la diffusion à un large public

⁴⁹. Voir notamment: Cécile de Terwangne, "Diffusion de la jurisprudence via internet dans les pays de l'Union européenne et règles applicables aux données personnelles", in *Petites Affiches*, 194, 2005, p. 4-55. ⁵⁰. Règlement (UE) 2016/679 du Parlement Européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données), ainsi que de la Loi du 8 décembre 1992 relative à la protection de la vie privée à l'égard des traitements de données à caractère personnel ainsi que l'arrêté royal du 13 février 2001 portant exécution de cette loi.

du double digital des documents produits. Ainsi, le développement des plateformes numériques des tribunaux internationaux a accompagné en temps réel l'activité de cette justice. Conçus au départ pour les praticiens et les personnels de ces institutions, ces outils ont subi diverses mutations et sont devenus aujourd'hui des sources incontournables pour écrire une histoire de ces conflits. Au-delà de leur usage par les chercheurs en sciences sociales et par les juristes, ces collections sont aussi des ressources pour les victimes et peuvent cristalliser des tensions politiques lorsque des États les réclament comme part de leur patrimoine. Il en est ainsi pour le Rwanda vis-à-vis des archives du Tribunal Pénal International pour le Rwanda créé par les Nations Unies en 1994⁵¹.

Sans s'inscrire dans un passé aussi proche, la numérisation de fonds qui documentent les initiatives successives de la Belgique en matière de poursuite des crimes internationaux commis dans le contexte des deux Guerres mondiales, pose une série de questions bien vivaces sur l'accès aux archives, l'identification des victimes et des responsables et la place que peuvent prendre ces sources dans le débat public. Le projet Jusinbellgium « *Un siècle de jurisprudence pionnière. Une base de données digitales de précédents belges en matière de justice internationale, 1914-2014* » place ainsi la digitalisation à l'intersection d'une réflexion sur les modalités techniques, la recherche, les usages et les usagers des archives judiciaires⁵².

La valorisation de sources longtemps méconnues, car conservées dans les dépôts des archives russes, mais également difficiles d'accès, car soumises à des démarches longues et fastidieuses, participe d'une double démarche de visibilité archivistique et historique. Les enquêtes sur les violations des lois de la guerre et du droit des gens commises en 1914-18 et en 1940-45 ont produit des collections documentaires jumelles mais aux trajectoires archivistiques bien différentes. Au sortir de la Première Guerre mondiale, un effort massif de documentation et de récolte de témoignages est mené auprès des autorités locales et de la population au sujet de l'invasion et de l'occupation. Ce n'est que six ans plus tard et au terme d'une série de rebondissements et de revirements que la justice belge ouvre des instructions devant les juridictions militaires. C'est ainsi qu'une partie de la masse documentaire produite par la *Commission d'enquête sur la violation des règles du droit des gens, des lois et des coutumes de la guerre (1914-1926)*, de même que par des enquêtes ouvertes devant les juridictions civiles (tribunaux de première instance, justice de paix), s'est trouvée redistribuée dans les dossiers des Conseils de guerre belges. En 1924 et 1925, les procès *in absentia* menés contre 200 criminels de guerre allemands – une part minime du nombre de dossiers ouverts – sont donc le fruit de la circulation et de la superposition des documents et des éléments de preuve générés et collectés par diverses institutions. Saisies par les Allemands en 1940, puis par l'armée soviétique en 1945 à Berlin,

51. Colette Braeckman, "Kigali accuse la justice internationale de 'trahison' et réclame les archives", in *lesoir.be*, 16 avril 2013. <http://blog.lesoir.be/colette-braeckman/2013/04/16/kigali-accuse-la-justice-internationale-de-trahison-et-reclame-les-archives/>, consulté le 29 septembre 2017. 52. Voir le blog du projet Jusinbellgium <http://jusinbell.hypotheses.org/>, consulté le 5 octobre 2017.

ces archives dorénavant dites de Moscou ont vu leur périple prendre fin en 2002, lors de leur retour en Belgique. C'est donc un pan entier de la préhistoire du jugement des criminels de guerre en Belgique entre 1947 et 1951 qui s'ouvre à la recherche grâce à la numérisation et à la mise en ligne de cet ensemble de sources.

La machine est, d'après *Moore's Law*, exponentiellement plus intelligente que tous les sous-sols des Centres de Calcul de l'époque de Le Roy Ladurie réunis, mais elle demeure toujours incapable de distinguer un témoignage produit dans le cadre de l'enquête judiciaire de celui produit en audience; la décision produite en première instance de celle de l'appel ou du renvoi en cassation ou des documents produits dans le cadre d'une demande de grâce. Elle n'est pas même capable de distinguer un dossier d'enquête classé sans suite d'un dossier ayant abouti à un jugement. C'est bien au chercheur, et à l'archiviste, qu'il appartient de dénouer les fils des catégories et des contenus des documents. Ainsi, les sources judiciaires répondent à une logique de production propre qu'il faut reproduire ou transmettre le plus fidèlement possible lors de leur numérisation et de leur mise en ligne. Ce défi se pose lors du processus de digitalisation lui-même, mais aussi lorsqu'il s'agit de déterminer la localisation digitale d'une collection – dans quelle (sous-) catégorie documentaire de la base de données a-t-elle sa place – et sa description. Les métadonnées, précieuses alliées de la recherche transversale, suscitent pourtant des questions en matière de compréhension historique. L'exemple concret de l'archive judiciaire montre clairement la difficulté de trou-

ver une correspondance entre les catégories utilisées par les chercheurs (victimes, auteurs, massacres, violence de masse), par le droit de l'époque (violations des conventions de La Haye, Code pénal belge), et par les qualifications juridiques auxquelles se réfère la justice pénale internationale aujourd'hui (Statut de Rome, définitions des infractions internationales et Conventions). Sans chercher une concordance parfaite entre ces trois dimensions, les humanités numériques permettent justement de réfléchir à ces différents regards posés sur des épisodes de guerre, de violence et de tentatives de justice.

Le projet Jusinbellgium a choisi la stratégie du coucou comme réponse au défi de l'hébergement durable. Nous avons en effet pondu notre œuf dans un autre nid: celui de la base de données de jurisprudence internationale de la Cour Pénale Internationale (CPI), profitant du fait que les jugements belges intéressent tout autant les historiens que les praticiens du droit international. Or, là où les juristes s'intéressent principalement aux jugements, les historiens s'intéressent aussi et peut-être surtout à l'enquête. Le défi (mais aussi l'opportunité) qui se présente alors à nous est de susciter l'intérêt de chaque utilisateur pour un volet documentaire qu'il ne cherchait a priori pas.

Les procès belges des années 1920 et des années 1940 peuvent ainsi être intégrés dans la structure généalogique et géographique mise en place par la base de données de la CPI⁵³. Généalogique d'abord, car elle a l'ambition de couvrir une jurisprudence relative aux crimes internationaux remontant bien en amont des années 1990. Géographique ensuite, car la dimension internationale que confère la base

53. Voir la plateforme en ligne www.legal-tools.org, consulté le 28 septembre 2017.

de données à la collection numérisée traduit un contexte de création lui-même fondamentalement international : les deux Guerres mondiales et le rapport entretenu avec les autres initiatives de justice pénale au niveau international (Nuremberg) et national (les tentatives à Leipzig en 1921, les procès alliés en Allemagne après 1945 et les processus judiciaires nationaux successifs dans les années 1940, 1960, 1980, voire plus récents).

Décoder la grammaire des fonds

L'histoire particulière des enquêtes et procès de l'après-Première Guerre mondiale qui ont généré une archive hybride, pose aujourd'hui la question d'une « restitution » des documents à leur producteur d'origine, en l'occurrence la Commission d'enquête, la justice de paix et les juridictions militaires provinciales. L'archive digitalisée n'existerait alors plus que comme collection digitale virtuelle sans reflet analogique. Or, ces sources numérisées n'ont de sens que s'il y a correspondance avec une réalité « papier », que celle-ci soit préservée ou qu'elle soit détruite (comme c'est le cas lorsque la digitalisation sert à libérer de l'espace de conservation). L'environnement numérique, outre les possibilités d'accès à distance et les outils plus ou moins poussés qu'il peut proposer, doit aussi permettre de traduire la trajectoire singulière des archives et l'histoire de leur constitution.

La plateforme publique et en ligne de la CPI, si elle évacue la difficulté liée à l'hébergement durable, ne résout pas la question du contexte archivistique. Car, une fois la numérisation réalisée et l'organisation des fichiers par juridiction et par dossier menée à bien, ces unités extraites d'une collection de procès plus vaste

ayant ses propres logiques, peuvent elles aussi se désintégrer dans le *cyberspace*. C'est donc à cette étape que l'identification et l'authentification du document digital entrent en jeu. Outre les possibilités qu'offre le numérique en matière d'expérience d'utilisation, il doit aussi prévoir l'adjonction au document digital d'une carte d'identité permettant d'en retrouver la trace physique et de le référencer de manière complète et précise. Cette marque de traçabilité est ainsi à la fois une indication à l'attention des chercheurs mais aussi de tous les autres publics susceptibles de s'intéresser à ces collections, et un moyen de produire une description de la source.

La réalisation d'une fiche de présentation de chaque dossier ayant mené à un jugement situe en bref les faits, le déroulé de l'enquête et du procès, les catégories juridiques mobilisées dans les archives, l'identité des accusés – et dans la mesure du possible celle des témoins et victimes –, et renvoie vers le référencement de l'archive physique. Ce faisant, elle inscrit les sources dans les fonds et sous-collections documentaires correspondantes, ce qui en dit autant sur le plan archivistique que sur la trajectoire historique des documents. Cette fiche est donc pensée en parallèle aux métadonnées, qui décrivent l'archive de façon beaucoup plus technique, et accompagne chaque document pdf. Là où les procès représentent un ensemble de plusieurs pdf, un filigrane indique la provenance de chaque dossier et signale l'ampleur du fichier digital : 1/3, 4/6, etc. Cela permet à l'utilisateur d'apprécier l'étendue matérielle et digitale de chaque dossier.

L'archive digitale n'est pas une représentation unidimensionnelle figée d'une collec-

tion fictive, mais bien un outil qui reflète la grammaire particulière des fonds et de leurs producteurs, permettant même de jeter des ponts entre des collections documentaires liées. Dans le même temps, l'archive brute ne raconte, comme on le sait, rien tant qu'elle n'est pas interrogée et confrontée aux débats qui animent les historiens et, au-delà, l'espace public. La nouvelle « topographie » numérique n'échappe pas aux « angles morts » de la recherche⁵⁴. En ce sens, les choix de numérisation ne sont pas neutres. Dans un contexte où l'environnement digital se transforme à une très grande vitesse, il est difficile d'anticiper les formes qu'il prendra dans cinq ou dix ans.

Éviter de sortir du virage numérique

Si nous voulons éviter que l'historien de demain ne soit plus, quels sont alors les principaux défis à relever ? Ne pas rater notre *digital turn* ? Nous avons tous déjà pris le virage numérique depuis quelque temps, même s'il nous en faudra encore pour découvrir et comprendre le nouveau paysage qui s'ouvre à notre vue. Les humanités numériques sont les humanités tout court, plutôt qu'une nouvelle sous-discipline. Pour prendre le contre-pied de la prophétie de Le Roy Ladurie, on peut aujourd'hui poser un diagnostic réaliste sur l'avenir de la discipline de l'histoire contemporaine. Il concerne le devenir de quelques-unes des collections-clefs pour l'histoire du vingtième siècle, actuellement en cours de transfert de leurs administrations d'origine vers les Archives générales du Royaume : les archives du Ministère des Affaires Étrangères, dont les Archives Africaines ; les archives de la Justice Militaire ; les archives du Service des victimes

de la guerre ; les archives du Ministère des Finances. Ces transferts représentent probablement le plus important versement d'archives publiques aux Archives générales du Royaume dans un laps de temps aussi court depuis la création de l'État belge. Ils ouvrent des perspectives excitantes de transparence, de professionnalisation de la gestion documentaire, de facilité d'accès, de raccourcissement des délais de consultation. Ils sont aussi accompagnés de risques considérables. Le premier est celui d'une perte de savoirs internes aux administrations sur l'organisation et le fonctionnement de leurs collections. Si on transfère des dizaines de kilomètres d'archives sans en transférer aussi les codes d'accès et les savoir-faire, dont certaines administrations ont cultivé la confidentialité comme un secret de famille et comme un petit monopole cognitif, ce sont les utilisateurs, au premier rang desquels les historiens, qui risquent d'être durablement pénalisés. Dit plus prosaïquement : si l'État Fédéral voit le transfert massif des archives comme un simple moyen de réaliser des économies de personnel dans les administrations les ayant produites sans doter les Archives générales du Royaume de moyens humains et techniques au moins équivalents, l'historien de demain risque d'être paralysé au point où il ne sera plus.

Il faut donc aujourd'hui une politique ambitieuse et urgente de construction de nouveaux silos d'archives, de recrutement d'archivistes compétents et évidemment d'élaboration d'outils numériques performants. Les choix opérés au moment des transferts dans la réalisation d'inventaires et dans la numérisation des outils

54. Voir la thèse développée par Lara Putnam : Lara Putnam, "The Transnational and the Text-Searchable: Digitized Sources and the Shadows They Cast", in *The American Historical Review*, v. 121, n° 2, 2016, p. 379.

de recherche des administrations d'origine (les centaines de fichiers cartothèques, les tableaux de correspondance et autres registres et recueils de circulaires) détermineront en grande partie le succès de ces transferts exceptionnels⁵⁵.

L'avant-gardisme technologique de systèmes permettant l'exploitation de *Big Data*, qui doivent encore très largement faire leurs preuves dans la recherche historique, constitue dans ce contexte un risque à ne pas

prendre. Pour la construction d'infrastructures de recherche pérennes qui évitent d'anticiper les questions de recherche de l'historien de demain, parions sur des choix numériques inscrits dans un esprit conservateur et analogique. Il s'agit de construire des environnements de recherche virtuels qui respectent au mieux la structure d'origine de l'archive, dans un souci de garantir l'authenticité et la traçabilité de l'information et de l'ouvrir au plus grand nombre de chercheurs.

Pieter Lagrou enseigne l'histoire contemporaine à l'Université Libre de Bruxelles. Ses recherches portent sur l'histoire des conflits et de la justice, notamment dans le cadre du projet Jusinbellgium (Brain-Belspo) ainsi que sur l'histoire comparée européenne et les questions de souveraineté populaire, langues nationales et nationalisme.

Delphine Lauwers réalise un post-doctorat aux Archives générales du royaume, dans le cadre du projet Brain Jusinbellgium. Ses recherches doctorales ont porté sur les conflits de mémoire de la Première Guerre mondiale dans le Saillant d'Ypres entre 1914 et 2014. Actuellement et dans le cadre du projet Jusinbellgium (Brain-Belspo), elle s'intéresse particulièrement au jugement des crimes de guerre commis durant les deux conflits mondiaux en Belgique, et à la manière dont ces épisodes ont jalonné l'histoire de la justice pénale internationale.

Ornella Rovetta est post-doctorante en histoire contemporaine à l'Université libre de Bruxelles (Centre de recherche Mondes Modernes et Contemporains), dans le cadre du projet de recherche Jusinbellgium (Brain-Belspo). Ses recherches portent sur l'histoire de la justice pénale internationale, en particulier le cas du génocide des Tutsi rwandais, et sur la poursuite des violations des lois de la guerre commises durant la Première Guerre mondiale en Belgique.

55. Voir, par exemple le projet néerlandais TRIADO (Tribunaalarchieven als Digitale Onderzoeksfaciliteit) <https://www.oorlogsbronnen.nl/triadovoorstel>, consulté le 29 septembre 2017.