::Savoirs LE FIL DES IDÉES

Cataloguer, indexer, encoder

Ou comment les données prennent vie

Revue d'anthropologie des connaissances, 7, n°4, 2013

Tiziana Nicoletta Beltrame and Christine Jungen

Abstract

L'article se penche sur la production et la manipulation de dispositifs de gestion de l'accumulation. Il entre dans les infrastructures conçues pour mettre en ordre et rendre accessible l'accumulation des connaissances par la clé de la matérialité et s'interroge sur les productions, configurations et manipulations in situ d'informations, de données, d'inscriptions. Les problèmes et enjeux de ces dispositifs se situent à la croisée de plusieurs champs de réflexion, notamment la taxinomie, la fabrication des données et de la production de standards, l'écologie des infrastructures informationnelles et la coopération entre communautés de pratique. Les technologies de l'information et de la communication sur laquelle reposent les économies de la société de l'information se fondent sur des équipements qui amplifient la diffusion, le stockage et l'organisation d'agrégats hétéroclites. Ces équipements sont des structures invisibles mais éminemment politiques de la production des connaissances. La perspective adoptée hérite de l'anthropologie des cultures matérielles autant que de l'attention portée aux ontologies et aux modes de présence des humains et non humains. Il s'agit de réinscrire les techniques numériques de stockage des connaissances dans un ensemble de pratiques de gestion documentaire dont la liste, l'inventaire ou le catalogue, la fiche ou l'imprimé. Ces technologies numériques sont dotées d'une matérialité dont il s'agit d'explorer les propriétés et les spécificités. Elles constituent des lieux d'exercice de politiques de contrôle en même temps que les données constituent des environnements dotés d'une texture spécifique qui agissent en

retour sur les structures normatives dans lesquels elles s'inscrivent.

En quoi est-il *compliqué* de traiter numériquement des êtres humains ? Par quels indices est *instaurée* une population de loups ? Comment des tags *agissent*-ils sur une photographie ? Ce sont quelques-unes des questions abordées dans ce dossier, dont l'objet est d'examiner la production et la manipulation de dispositifs de gestion de l'accumulation. Le parti adopté ici est simple : entrer dans les infrastructures conçues pour mettre en ordre et rendre accessible l'accumulation des connaissances – catalogues, banques de données, inventaires, index – par la clé de la matérialité ; s'interroger sur les productions, configurations et manipulations *in situ* d'informations, de données, d'inscriptions.

Les problèmes et enjeux des dispositifs de gestion des ressources documentaires se situent à la croisée de plusieurs champs de réflexion : citons, en vrac, celui de la taxinomie et des catégorisations opérant sur le monde ; celui de la fabrication des données et de la production de standards ; celui de l'écologie des infrastructures informationnelles, et plus largement, des « objets-frontières » comme lieu d'articulation et de coopération entre communautés de pratique. Ces sujets et les débats qui s'y rattachent ayant très largement été exposés au cours de ces dernières années dans la *Revue d'anthropologie des connaissances* [1], nous focalisons ici le propos sur les lignes directrices qui ont présidé à la constitution de ce dossier, et proposons quelques clés de lecture – non exhaustives bien sûr – des contributions qui le composent[2].

Thésauriser, classer et redistribuer des données à des échelles de plus en plus grandes est une des caractéristiques des pratiques de savoir dans le monde contemporain. Les technologies de l'information et de la communication sur laquelle reposent aujourd'hui les économies de cette « société de l'information » se fondent sur des équipements qui génèrent une amplification à grande échelle de la diffusion, de la capacité de stockage, mais aussi, plus largement, d'organisation d'agrégats souvent hétéroclites. Rendus possibles par l'articulation et la coordination d'une grande variété de dispositifs et d'opérations, par la catégorisation et la normalisation de matières diverses en « données » et en « informations » redistribuables, ces équipements font partie de ces structures, aussi invisibles qu'éminemment politiques, de la production des connaissances (Bowker, 2000 ; Bowker & Star, 1999).

La nécessité de produire des outils permettant de gérer des masses de données de plus en plus considérables a conduit à des mutations profondes dans les modes d'organisation et d'accès des ressources documentaires, et plus largement du savoir. De la liste, au catalogue, à

la base de données, les dispositifs pour gérer l'assemblage d'objets, de connaissances, de données, d'informations se sont ainsi progressivement multipliés, transformés, complexifiés. Index, inventaires, banques de données et collections numériques, répertoires, moteurs de recherche, tables, logiciels de gestion forment ainsi aujourd'hui un ensemble dialogique de structures physiques et numériques répondant à un objectif commun : accommoder des éléments au sein d'ensembles, les ranger, les retrouver.

Dans la continuité des travaux qui ont souligné, à rebours de l'appréhension du classement comme opération relevant exclusivement de l'intellect, le rôle central de la matérialité dans les pratiques de rassemblement, d'agencement et d'organisation des ressources documentaires et informationnelles (Baratin & Jacob, 1996 ; Denis & Pontille 2012a ; Gardey, 2008 ; Jacob, 2011 ; Ruhleder, 1995), ce dossier aborde la fabrication et la manipulation de ces dispositifs comme des artefacts à part entière, dotés de matérialités singulières qui agissent sur le travail d'accumulation, de classement, de catégorisation et de configuration de ces ressources. La perspective adoptée ici doit beaucoup, également, au renouvellement récent de l'anthropologie des cultures matérielles (Ingold, 2007, 2011) autant qu'à celui d'une attention portée aux paradigmes ontologiques et aux modes de présence des humains et non humains, leur manière d'« être-au-monde » (Ingold, 2000 ; Piette, 2009, 2010 ; Stengers & Latour, 2009) : en choisissant d'aborder les dispositifs de gestion de l'accumulation par leurs textures matérielles, en cadrant la focale sur les opérations empiriques d'accommodement, d'attachement, de tissage de matériaux effectués par des techniciens, praticiens et concepteurs, le propos est, ici, de saisir des objets – et des *êtres* – plutôt que des données, des *accommodements* plutôt que des classifications, des *environnements* plutôt que des catégories.

De la liste à la base de données : la domestication technique de l'accumulation

De la liste, forme d'énumération méthodique, à la base de données fondée sur la mise en relation, les modalités d'agencement de l'information se sont profondément transformées au cours du temps. Il n'est pas inutile, dans ce cadre, de revenir sur la diversité des dispositifs matériels qui ont précédé l'avènement, dans l'histoire occidentale, de l'ère informatique.

Qu'est-ce qu'exactement une liste ? Jack Goody, dans son examen des techniques d'écriture en Mésopotamie ancienne, en avait rappelé quelques principes :

> La liste implique discontinuité et non-continuité. Elle

> suppose un certain agencement matériel, une certaine disposition spatiale ; elle peut être lue en différents sens, latéralement et verticalement, de haut en bas comme de gauche à droite, ou inversement ; elle a un commencement et une fin bien marqués, une limite, un bord, tout comme une pièce d'étoffe. Elle facilite, c'est le plus important, la mise en ordre des articles par leur numérotation, par le son initial ou par catégories. Et ces limites, tant externes qu'internes, rendent les catégories plus visibles et en même temps plus abstraites. (Goody, 1986, p. 150)

Comme le souligne Goody, la liste se distingue par un formatage visuel et matériel des informations – la disposition sur la tablette, la finitude, mais aussi la mise en ordre et la hiérarchisation auxquelles donne lieu le dispositif de l'énumération – intimement lié au traitement des données ainsi agencées : *discontinuité* des informations – on découpe, on retranche, on ordonne en séparant – et *linéarité* constituent ainsi dans l'opération de fabrication d'une liste les principes formels des catégories et de ce qu'elles contiennent.

L'histoire des techniques de gestion de l'accumulation est indissociable de celle de ses supports matériels. Aux listes et premiers catalogues sur tablettes d'argile de la Mésopotamie ancienne succèdent les catalogues sous forme de rouleaux de la Grèce antique. De même que la liste agence et contraint des modes précis d'organisation du savoir, le format du livre, rouleau puis codex à la fin de l'Antiquité, implique des modes d'écriture, de lecture, et d'« arts de mémoire » précis, comme le rappelle Christian Jacob :

> Le livre médiéval, sous forme de codex, a pu donner lieu à des arts de mémoire sophistiqués, reposant parfois sur une analogie étroite entre la configuration matérielle d'une page précise et son image intériorisée, sa "photocopie mentale" (différents espaces d'écriture à l'intérieur de la page, miniature, couleurs, lettre initiale enluminée, etc.). Le rouleau de papyrus alexandrin n'offrait que la succession linéaire de colonnes d'écritures sans séparation des mots, avec des repères visuels beaucoup moins prégnants [...] (Jacob, 1996, p. 57)

L'avènement de la mécanisation des techniques d'écriture, en premier lieu celui de l'imprimerie induit de nouvelles formes de rationalisation des connaissances : l'imprimerie fixe physiquement les écrits du passé, standardise les textes, permet leur diffusion dans l'espace public (Eisenstein, 1991). Cette profonde transformation que connaît l'Europe moderne dans ses pratiques de l'écrit est suivie, au XIX^{e} siècle, par ce que Delphine Gardey a appelé la « révolution de papier », au cours de laquelle

> [...] un monde s'invente dans cette frénésie mécanicienne, la multiplication des accessoires, les transferts et correspondances de matériels et de formats autorisant une profusion nouvelle de lettres et de chiffres. Un monde s'invente parce que le rapport au monde, le mode d'agir, les *possibilités d'intervention sur le monde* sont transformés : ces médiations mécaniques et de papier, ces ordonnancements, ce tramage chiffré, ces quadrillages reconfigurent les formes d'activité, permettent de les redéployer et produisent une grandeur nouvelle du capitalisme et du gouvernement. (Gardey, 2008, p. 9)

Mise en liste, quadrillage, standard : ces « économies de l'écrit » qui produisent autant de modes d' « intervention sur le monde » ont ainsi fait l'objet d'une sédimentation successive, à partir de laquelle a pu se déployer la révolution numérique et les changements que celle-ci a amenés dans nos pratiques de gestion de l'information.

De ce point de vue, le souci de ce dossier est double : réinscrire, d'une part, les techniques digitales de stockage des connaissances dans un ensemble de pratiques de gestion documentaire dont la liste, l'inventaire ou le catalogue, la fiche ou l'imprimé constituent à la fois des formes complémentaires ou alternatives aux technologies informatiques. Restituer, d'autre part, à ces technologies et aux univers propres qu'elles font émerger leur matérialité : nous partons ici du principe que les bases de données numériques, au même titre que les tablettes en argile sumériennes, les catalogues en rouleau de papyrus, les imprimés, formulaires ou fiches papier sont, à rebours des discours sur la « dématérialisation », dotés d'une matérialité dont il s'agit d'explorer les propriétés et les spécificités.

Des propriétés de la matière

Une première remarque s'impose à ce stade : une multiplicité de formats et supports matériels cohabitent aujourd'hui, sont manipulés concurremment les uns avec les autres. La constitution de dispositifs de gestion des données donne ainsi lieu à un travail incessant de coordination entre différents supports, de va-et-vient, à des opérations de traduction, de translation d'un support à l'autre. Plusieurs contributions font état des chaînes opératoires dans lesquelles s'enchaînent et s'alternent, de manière plus ou moins fluide, des séries d'inscriptions : de celles portées manuellement sur un formulaire ou une fiche jusqu'aux codes entrés dans une base de données, qu'il s'agisse d'herbiers ou de matériaux biologiques à faire « entrer » dans une base de données (Lorna Heaton et Florence Millerand), de symptômes de maladie à modeler en code d'un logiciel de gestion (Pierre-André Juven), ou de traces animales à muter en

outil statistique (Antoine Doré). De ce point de vue, les infrastructures de gestion de collections et de données, si elles peuvent se lire comme l'aboutissement d'un travail d'extraction, de traduction et de réduction (Latour, 1996), forment également, dans le même temps, des collectifs de coordination de matériaux multiples dont le dispositif final – la base de données – ne constitue que la composante la plus visible.

Cette question de la coordination et de la distribution, lors de la constitution des bases et catalogues, de matériaux hétérogènes appelle à revenir sur le problème de la saisie de la matérialité et de ses enjeux. Tim Ingold émet une critique corrosive d'une tradition d'anthropologie de la matérialité qui, remarque-t-il, s'est tant préoccupée de *matérialité* qu'elle n'a, finalement, que très peu à dire sur la *matière* proprement dite (Ingold, 2007). Le reproche principal qu'il fait à cette littérature est de s'être nourrie – et d'avoir nourri en retour – d'une série de partitions, telle celle entre l'esprit et la matière, ou encore entre les objets naturels et les artefacts. À la suite des travaux de Gibson sur les *affordances*(Gibson, 1986), Ingold propose quant à lui de considérer les matières non comme des objets physiques stables, mais comme des constituants actifs du « monde-en-formation » (*world-in-formation*). Prises comme produits d'histoires (*stories*), en continuel renouvellement, les matières ne sont pas tant nanties d'« attributs » fixes que susceptibles d'être transformées en permanence dans les interactions auxquelles elles se trouvent soumises dans l'environnement. Dans une perspective qui rejoint celle de la « matérialité relationnelle » proposée par John Law et Annemarie Mol (Law & Mol, 1995), les composants matériels de l'environnement sont dotés de « propriétés relationnelles » ; il convient d'en examiner non pas des qualités intrinsèques, mais les trajectoires et les mutations, les *affordances* dans un environnement qui se « déplie » (*unfold*) au fur et à mesure des interactions.

Cette approche propose des outils en premier lieu pour saisir des formes de matérialisations particulièrement évanescentes, à l'instar de celles qui nous préoccupent ici : inscriptions, codes, chiffres sont autant d'artefacts, qui, si stables et solides soient-ils, sont dotés de « modes de présence » particulièrement « minimaux », pour emprunter à Albert Piette (Piette, 2009), sur le traitement desquels vient régulièrement buter la description empirique. C'est une des raisons, sans doute, aussi, pour lesquelles les techniques informatiques sont si souvent appréhendées sous l'angle de la dématérialisation : on aurait ainsi d'un côté un monde physique, tangible, peuplé d'objets dotés d'une tactilité immédiate, patente – des objets qui tombent, littéralement, sous la main, offrent des prises et des adhérences évidentes ; de l'autre, un monde fait de surfaces parcourues par des lignes, traits, chiffres, codes, et dont la matérialité beaucoup plus fugace, subreptice appelle à des modes d'engagement singulièrement autres ; la transformation et reconfiguration de matières entre ces

mondes et ces formats, semble, de la sorte, se subsumer sous une disparition, magique presque, de la matière au profit de la représentation symbolique. Comme le remarquent Jérôme Denis et David Pontille, ces opérations sont pourtant, bien au contraire, des moments d' « instanciation » matérielle de l'information (Denis & Pontille, 2012b) dont il convient d'examiner les modalités.

En second lieu, l'approche par la matière et ses dépliements peut également être particulièrement riche pour saisir une des conditions centrales de la production de ces « infrastructures-frontières » (Bowker & Star, 1999) que constituent les dispositifs de gestion de l'accumulation, celle du formatage et de la standardisation. Comment, en effet, se *comportent* les standards et ces « mobiles immuables » (Latour, 1989) dont le caractère unifié et stabilisé permet la circulation à grande échelle ? Les contributions dans ce dossier montrent, à des échelles diverses, comment la « mise au standard » s'accompagne, localement, d'un travail de manipulation et de mise en forme(s) de la matière dont l'analyse met en valeur la diversité, précisément, de ces « propriétés relationnelles » : Dominique Boullier et Maxime Crépel notent ainsi, à propos de l'intégration de photographies dans la banque d'images Flickr que l'effort de standardisation d'un format unifié des images lors de leur mise en ligne dans la base s'accompagne d'opérations de mises en relation par des dispositifs tels le « tagging » qui remodèlent en profondeur le statut de la photo. Sophie Houdart souligne la difficulté, dans les catalogues d'éléments de composition de dessins d'architectures, à trouver le « bon format » de l'être humain, et plus encore, à le faire voisiner dans un même espace digital avec d'autres éléments au format numérique pourtant identique.

Les transformations, traductions, formatages, le rééquipement des matières et matériaux font ainsi souvent l'objet à chaque moment de procédures aussi laborieuses que fragiles, comme le montrent Lorna Heaton et Florence Millerand : intégrer des planches d'herbier ou des échantillons de planctons à une base de données numérique ne va, *de facto*, jamais de soi. Les procédures de transformations matérielles et les nouvelles relations que celles-ci produisent aboutissent à des échafaudages toujours précaires et instables : comme si « l'effacement des modalités » (Latour & Woolgar, 1988) nécessaires à la standardisation s'accompagnait de nouvelles dispositions de la base de données et des éléments qui la composent à fragiliser la fermeté de l'infrastructure et son aptitude à faire circuler correctement les nouvelles matières-en-relation. Cette capacité de la matière à résister à la standardisation, et à produire *quelque chose d'autre*, se retrouve également dans le travail de catalogage de manuscrits arabes, analysée par Christine Jungen. Ici, c'est la matérialité floue, incertaine de copies microfilmées de manuscrits originaux qui constitue le point de départ d'une série de réagencements matériels et intellectuels dans la production de catalogues : la construction des notices du catalogue – et la résistance que celui-ci incarne à toute forme d'uniformisation et

d'anonymisation – vient alors s'inscrire dans des interactions multiples entre le microfilm, le catalogueur, et les aspérités et prises de l'environnement dans lequel celui-ci travaille.

Recomposer les corps

Entasser des données à partir du monde, concevoir et manipuler des outils pour gérer celles-ci, s'effectue par la fragmentation et le réassemblage d'entités, le travail de réduction de la singularité, par des opérations de catégorisation, de redistribution et de mise en réseau dans et en dehors des différents dispositifs destinés à gérer l'accumulation de matériaux. Dans ce cadre, décomposer, façonner, recomposer ou recompiler sont, de la part des praticiens de la ressource documentaire et informationnelle, des opérations souvent parfaitement anodines et routinières. Elles n'en produisent pas moins des occurrences matérielles particulièrement efficaces pour « intervenir sur le monde », pour reprendre la formule employée par D. Gardey. Les dispositifs de gestion de collections sont, de fait, également des espaces par excellence à partir desquels s'exercent un pouvoir, un contrôle, une domestication. La capacité à accumuler la matière informationnelle, à construire des lieux où celle-ci est stockable, manipulable, et trouvable, peut, de fait, faire tenir des empires[3].

Ces dispositifs constituent des lieux d'exercice d'un pouvoir, en premier lieu dans l'action que les classements et catégorisations ont sur le réel (Fradin, Quéré & Widmer, 1994), de même que, comme l'a montré Alain Desrosières, leur « naturalisation » à travers la constitution d'instruments de gouvernance (Desrosières, 1993). L'opération d'encodage du réel qui se manifeste dans la réduction, le chiffrage, la division en « plus petits dénominateurs communs », en unités minimales, et leur réarticulation dans de nouveaux assemblages permet non seulement de manier et recomposer à grande échelle ; ce sont aussi des lieux où sont reconstitués, voire fabriqués de nouvelles formes de la réalité, de nouveaux corps sur lesquels peuvent venir s'exercer des politiques de contrôle et de surveillance.

C'est à ce travail de décomposition et recomposition matérielles que s'intéressent respectivement Frédéric Keck et Antoine Doré dans des contextes où il s'agit, avant tout, de saisir et surveiller des entités aussi fuyantes que menaçantes (le virus, le loup). Frédéric Keck examine la double gestion, bio-informatique et écologique, des virus et des oiseaux qui les portent. Produire des occurrences matériellement analysables et saisissables du virus passe dans ce cadre moins par un travail de réduction que celui d'un *dédoublement* qui rejoue ou suspend les partitions entre espèces : entre la culture de cellules et leur transposition bio-informatique, entre le vivant et le code. Les bases de

données constituent alors « un espace codé » aux relations « suspendues », où l'on peut « "jouer" avec les données, c'est-à-dire essayer des mouvements possibles afin de deviner l'avenir ».

Dans le cas analysé par Antoine Doré – la gestion biopolitique d'une population animale par la détection de leur présence –, il s'agit, de manière similaire, d'obtenir à travers la collecte et le travail d'encodage et de décodage d'indices, d'échantillons, des cristallisations matérielles d'une entité à la fois potentiellement menaçante et rebelle à la saisie. Les pratiques de détection des loups dans la constitution de bases de données à partir des indices de présence qu'ils laissent sur le territoire ne sont pas, de fait, très éloignées des « pratiques divinatoires » décrites par F. Keck : l'instauration arithmétique d'une population animale aussi insaisissable qu'invisible passe par la détection, les recompilations, codages et décodages, alignements statistiques d'indices et de traces – crottes, urines, poils. Les différentes combinaisons et les modes de consomption de ces indices aboutissent ainsi à recréer, au bout de la chaîne, une « population loup » comme référent ontologique stable, dont on peut gérer et prévoir à distance les développements.

Contenir l'incertitude par des opérations de cristallisation, de « précipitation » pour user d'une métaphore chimique, aboutit ainsi à produire de nouveaux référents, de nouveaux corps à partir desquelles peuvent désormais s'exercer d'autres logiques de gestion, voire de nouveaux *êtres* : c'est la conclusion à laquelle aboutit Pierre-André Juven dans son examen du travail de codage des maladies dans le cadre des gestions hospitalières : le corps du malade est transformé en un code gestionnaire suite à une série d'extractions et de transformations qui produisent des entités susceptibles d'êtres inscrites dans des catégories de coûts. Des malades à l'hôpital deviennent ainsi, par le travail d'encodage de la maladie, des « êtres de gestion » poursuivant désormais une nouvelle trajectoire dans les bases de données de l'Assurance maladie : les dispositifs de gestion et de contrôle d'entités mouvantes et instables – la maladie, le virus, le loup – sont aussi des lieux où, par le travail d'encodage, de recomposition, de dédoublement, ces entités sont dotées de matérialités et d'ontologies alternatives (ou parfois suspendues) sur lesquelles peuvent alors s'exercer les logiques de gestion, de surveillance, ou d'endiguement.

Manières d'habiter et modes de présence

Plusieurs contributions de ce dossier le signalent, l'architecture matérielle des collections digitales prolonge souvent et amplifie, plus qu'elle ne transforme radicalement, des dispositifs élaborés à l'extérieur de l'univers informatique. Les techniques numériques de

gestion des collections ont, en revanche, opéré un basculement dans un autre domaine : celui du point de vue à partir duquel sont effectuées les manipulations de données, et celui du lieu d'où se construit l'engagement à ces dispositifs. Ainsi, un catalogue imprimé, une série de fiches ou une liste sur une feuille de papier s'appréhendent par une vue « d'en haut ». Les informations inscrites dans ces supports sont « tenues sous le regard », pour reprendre la formule de Bruno Latour (Latour, 1996), saisissables par l'œil dans leur ensemble – et dans leur finitude : la liste comporte un bord, l'imprimé une reliure, le fichier un panneau. On les parcourt comme on parcourt une surface, mais c'est une surface déterminée et délimitée par ses contours tangibles. À l'inverse, une base de données informatique, une banque d'images, une bibliothèque numérique ont la capacité à s'agrandir de manière quasiment infinie ; elles se naviguent bien plus qu'elles ne s'arpentent ; et le praticien d'une banque numérique ne voit, littéralement, que des parcelles, des bribes de la base complète dans laquelle il se déplace en suivant des liens.

De ce point de vue, le passage du support papier au support digital, de la surface à l'environnement digital n'est pas sans rappeler le mouvement, inverse, que Tim Ingold souligne dans l'histoire de l'imagerie européenne du globe terrestre : celui qui, dans l'Europe moderne, a vu le déplacement de l'homme vers l'extérieur de la sphère terrestre, désormais représentée visuellement sous forme de globe. Ingold remarque à ce propos :

> Thus the movement from spherical to global imagery is also one in which 'the world', as we are taught it exists, is drawn ever further from the matrix of our lived experience. It appears that the world as it really exists can only be witnessed by leaving it, and indeed much scientific energy and resources have been devoted to turning such an imaginative flight into an achieved actuality. One consequence is the alleged discrepancy between what, in modern jargon, are called 'local' and 'global' perspectives. Insofar as the latter, afforded to a being outside the world, is seen to be both real and total, the former, afforded to beings-in-the-world (that is, ordinary people) is regarded as illusory and incomplete. (Ingold, 2000, p. 211)

Si l'on applique le raisonnement d'Ingold au (re)basculement qu'ont opéré les techniques informatiques dans la gestion des ressources documentaires, on peut alors formuler le questionnement dans les termes suivants : comment la perspective *locale* – celle de celui, qui, d'un terminal, a accès à une vision parcellisée de l'ensemble de la collection à laquelle il contribue – s'articule-t-elle à la perspective *englobante*, totale, le « rêve panoptique » (Bowker, 2000, p. 645) que paradoxalement les bases numériques briguent comme horizon ultime [4] ? Quelles sont, plus largement, les formes d'expérience particulière

que ces environnements suscitent ?

Il s'agit donc aussi, pour aller plus loin, de prendre ici au pied de la lettre la proposition, formulée par Geoffrey Bowker et Susan Star, d'examiner comment des données « habitent » physiquement des contextes donnés (Bowker & Star, 1999). De fait, si les dispositifs de gestion de l'information constituent des interfaces de coopération entre de multiples communautés de pratiques, entre différentes formes de localité, ils constituent également des environnements physiques, dotés d'une texture spécifique, et éminemment diverse : les bases de données numériques recouvrent des réalités extrêmement hétérogènes, dont les contributions à ce dossier se font l'illustration : une base de données bio-informatique n'est pas de même nature qu'un catalogue d'images ou une collection d'herbiers numérisés, ni ne propose les mêmes agencements relationnels. Les items qui circulent et/ou sont stockés dans les infrastructures informatiques proposent des modes d' « être » dans leur environnement numérique à chaque fois uniques.

La contribution de Sophie Houdart éclaire cette perspective dans son analyse des catalogues numériques d'êtres humains à usage des agences d'architectures, en interrogeant les modes de présence, de voisinage et la bonne manière d' « être réel » des éléments du catalogue ; de fait, les figures humaines rassemblées dans ces banques d'éléments de figuration ne sont pas de simples représentations, mais bien une « manière d'être » très spécifique dans l'environnement numérique. Se posent alors des problèmes d'équivalences – du point de vue de son format informatique, un ciel est ainsi équivalent à une figure humaine bien qu'il ne le soit pas ontologiquement – mais aussi de « récalcitrances » des êtres humains à entrer dans l'univers digital : l'environnement numérique, du fait qu'il traite de la même manière les différentes composantes d'un dessin d'architecture, se trouve par là même confronté à la cohabitation malaisée, problématique, des êtres humains avec d'autres entités numériques, en dépit des différentes procédures mises en œuvre pour extraire les figures humaines de leur contexte référentiel d'origine.

En d'autres termes, importer des êtres, humains ou non humains, dans un catalogue ou dans une base de données impose de devoir composer avec leur ontologie et leur manière propre d'habiter et de circuler dans un environnement. Cette perspective amène donc à examiner comment humains et non-humains « peuplent » le monde (Houdart & Thiery, 2011), ici celui de l'univers produit par les bases digitales. Les techniques numériques permettent non seulement de traduire et de réduire des collections en inscriptions, mais également d'importer des matériaux sous le format informatique : les banques de données ne contiennent ainsi plus seulement des inscriptions, mais permettent de véritables coalescences du monde et de ce qu'il contient sous forme numérique : des planctons et des humains, des photographies et des

malades, des livres, des virus et des loups, pour reprendre les cas traités ici, soit une diversité d'éléments qui poursuivent, sous leur identité informatique, une trajectoire de vie distinctive dans le nouvel environnement que constituent la base, le catalogue, la collection dans laquelle ils sont désormais pleinement intégrés.

Comment rendre compte de ces trajectoires, de ces biographies singulières ? C'est à un véritable pari épistémologique que se livrent dans leur contribution Dominique Boullier et Maxime Crépel, en examinant les parcours de photographies dans la banque d'images Flickr. La photographie, prise ici comme sujet à suivre au sein de la banque d'images, est au centre de l'examen des *agencies*qui sont à l'œuvre dans Flickr : les différentes entités – l'utilisateur, les tags, les photographies, la base de données – possèdent ainsi des *agencies* distribuées à travers lesquelles la photographie acquiert ses catégories, se transforme en permanence, et génère autour d'elles des « associations vivantes ». Le changement de point de vue que les auteurs adoptent, et la lecture tardienne de la base de données à laquelle ils se livrent – la photographie « et ce qu'elle emmène avec elle » – permet d'éclairer la banque de données non plus comme une structure fixe dans laquelle sont rangés des éléments, mais comme un habitat à l'architecture lâche, où les « sociations » s'élaborent et se recomposent au fur et à mesure des circulations et des reclassements des photographies.

* * *

Concluons ce tour d'horizon par un constat : un effort important a été porté, dans la littérature sur les infrastructures informationnelles comme en sociologie des sciences, sur la déconstruction des « boîtes noires » qui supportent la construction des connaissances, sur le dévoilement des structures invisibles qui organisent nos pratiques, de savoir et au-delà. Dans ce cadre est notamment, régulièrement, soulignée la force d'inertie de ces structures informationnelles une fois qu'elles ont été stabilisées et rendues invisibles, leur forme d'irréversibilité (Bowker, 2000 ; Bowker & Star 1999 ; Callon, 1991). Ce que suggèrent les contributions à ce dossier, c'est aussi le penchant du monde et des êtres qui le composent à agir en retour sur les structures normatives dans lesquels ils sont inscrits, à s'y soustraire, à faire problème ; à remodeler, également, ces dispositifs de l'intérieur : une capacité à *repeupler*, en somme, des infrastructures qui sont, de toute évidence, moins inertes qu'il n'y paraît.

Notes

1. On renvoie ici aux dossiers qui ont été consacrés à ces différents

sujets : Faugère & Mauz, 2013 sur la taxinomie ; Milanovic, 2011 sur les ressources biologiques ; Trompette et Vinck, 2009 et 2010 sur l'objet-frontière ; Denis & Pontille, 2012a sur les « petites mains » de l'information.

2. Ce dossier est le fruit d'une réflexion qui a été engagée lors des Journées d'études « Manipuler, accommoder, ranger. Catalogues, inventaires et bases de données » (1-2 mars 2010, Musée du Quai Branly) qui se sont tenues dans le cadre du programme ANR « Archiver. Les pratiques historiographiques dans le Moyen-Orient contemporain » (IISMM, EHESS).

3. C'est ce que montre par exemple Thomas Richards à propos de l'Empire britannique, qui s'est appuyé pour contrôler ses territoires sur un usage extensif de la collecte d'informations : bibliothèques et musées deviennent les lieux dans lesquels existent, véritablement, l'Empire et son administration (Richards, 1993) ; plus proche de nous, citons l'exemple de Google, qui a construit son empire aussi puissant que diversifié à partir d'un des impératifs les plus basiques de la gestion des connaissances : trouver facilement une information.

4. G. Bowker donne pour exemples les entreprises de bases de données génétiques (tel le Human Genome Initiative), et plus largement biologiques, qui ont pour objectif de collecter et « mettre en banque » l'ensemble de la « vie ». Ces projets entrent également en écho avec le « mal d'archive » et la frénésie mémorielle disséquée par Jacques Derrida (Derrida, 1995) ; un des exemples les plus récents en est le projet Internet Archive (www.archive.org), qui a pour ambition de collecter et stocker toutes les pages web et plus largement toute production digitale depuis la création d'internet…

Nos partenaires

Le projet *Savoirs* est soutenu par plusieurs institutions qui lui apportent des financements, des expertises techniques et des compétences professionnelles dans les domaines de l'édition, du développement informatique, de la bibliothéconomie et des sciences de la documentation. Ces partenaires contribuent à la réflexion stratégique sur l'évolution du projet et à sa construction. Merci à eux !

L'ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES

haStec
Laboratoire d'Excellence
Histoire et anthropologie
des savoirs, des techniques
et des croyances

fmsh fondation maison des sciences de l'homme

PDN
Pôle Document Numérique
Maison de la Recherche en Sciences Humaines
CNRS - UNIVERSITÉ DE CAEN

métopes
méthodes et outils
pour l'édition structurée

EPFL

- CONCEPTION : ÉQUIPE *SAVOIRS*, PÔLE NUMÉRIQUE RECHERCHE ET PLATEFORME GÉOMATIQUE (EHESS).

- DÉVELOPPEMENT : DAMIEN RISTERUCCI, IMAGILE, MY SCIENCE WORK. DESIGN : WAHID MENDIL.