

Des écosystèmes d'intelligence collective au service des archives numériques

1.1. Les archives numériques

La gestion des archives numériques constitue un des enjeux majeurs aujourd'hui et pour les années qui viennent. On estime que tous les deux jours, l'humanité produit autant d'information numérique que pendant les deux millions d'années qui nous précèdent. A cette production humaine, il faut ajouter les informations que les machines produisent en continu. Le coût de la mémoire informatique étant de plus en plus bon marché, la plupart de ces informations sont stockées dans d'immenses bases de données. L'ensemble de ces « Big Data » représentera en 2025 près de 8 zettaoctets (mille milliards de gigaoctets) [SAD 15]. A notre époque, il ne reste que très peu d'activités humaines qui ne génèrent pas d'archives numériques, chaque jour nous alimentons les flux de production numérique même en dehors de nos usages des ordinateurs, des téléphones ou autres dispositifs numériques. Il nous suffit d'allumer la lumière, de faire des courses, de prendre les transports publics ou de regarder la télévision pour produire des traces numériques qui, pour la plus grande partie, ne nous seront jamais accessibles mais qui se trouvent compilées, indexées, calculées dans des fermes de serveurs et des centres de gestion.

Le statut de ces archives numériques n'est, bien évidemment, pas le même lorsqu'il s'agit du *tweet* envoyé automatiquement par une vache, de la numérisation d'un cours de Gilles Deleuze ou de la modélisation en 3D de la Citadelle Ferrière de Milot du Cap Haïtien. Même si *in fine*, ces archives se composent d'un ensemble de 0 et de 1,

et sont donc formellement comparables les unes avec les autres, leur importance n'est pas équivalente et surtout elle varie suivant les contextes d'espaces, de temps et d'acteurs qui sont confrontés à ces informations. Le *tweet* envoyé par un dispositif numérique en relation avec les activités d'une vache¹ n'a sans doute pas beaucoup d'importance pour la plupart d'entre nous mais pour le producteur de lait qui veut suivre les déplacements de son troupeau pour corrélérer la composition du lait avec les herbages broutés, il est important de savoir que tel herbage a une influence sur la quantité de lipides dans le lait. De même, tel passage des cours de Gilles Deleuze où il parle de l'importance comme un critère fondamental, apparaît pour certains comme une phrase quasi insignifiante alors qu'elle prend une très grande importance pour le chercheur qui s'intéresse aux rapports entre éthique et ontologie, mais aussi pour le lecteur de ces lignes qui en cet instant pense à ce concept par le juste fait de le lire :

« Qu'est ce que ça veut dire ça, cette catégorie ? L'important. Non, on est d'accord, ça c'est embêtant, mais ce n'est pas important. Qu'est ce que c'est ce calcul ? Est-ce que ce n'est pas ça ? Est-ce que ce n'est pas la catégorie du remarquable ou de l'important qui nous permettrait de faire des proportions entre les deux sens irréductibles du mot partie ? Ce qui dépend et ce qui découle de la part intensive de moi-même et ce qui renvoie au contraire aux parties extensives que j'ai. »²

Ces proportions entre l'être-intérieur et l'avoir-extérieur sont tout à fait transposables dans le domaine des archives numériques. De par leurs caractères dynamiques, évolutifs et interactifs, les archives numériques sont des écosystèmes où chaque élément peut être analysé en termes d'existences composées de « parties intensives » et de « parties extensives ». L'exemple de la numérisation du fort du Cap Haïtien apporte un éclairage sur l'importance des archives numériques qui illustre cette double dimension « intensive/extensive » que souligne Deleuze pour montrer la corrélation entre une dimension extérieure liée à l'avoir et au matériel, et une dimension intérieure liée à l'être et à l'immatériel. Dans le cas de ce monument historique classé au patrimoine de mondial de l'UNESCO, l'archivage numérique est l'occasion, dans un des pays les plus pauvres du monde, de développer un patrimoine à la fois matériel et immatériel. La création d'un programme de recherche international sur les problématiques de réalités augmentées, l'enseignement et la formation d'étudiants sur ces questions ainsi que la mobilisation des artistes pour un usage innovant de ces technologies sont trois exemples de développement du patrimoine immatériel. Parallèlement, ces activités permettent d'envisager le développement du patrimoine matériel à travers la mise en place d'une économie utilisant ces archives numériques pour créer de nouveaux services à destination des

1. Voir : criticalmedia.uwaterloo.ca/teattweet/.

2. Voir : www2.univ-paris8.fr/deleuze/article.php?id_article=24.

touristes en croisière qui passent à proximité de ce pays. Ici l'impact de l'archive numérique dépasse le cadre d'une entreprise ou celui de la connaissance en ayant des répercussions sur l'économie globale d'un pays par un développement conjoint du patrimoine matériel et immatériel.

Dès lors, la problématique fondamentale des archives numériques consiste à s'interroger sur leur importance tant sur le plan matériel qu'immatériel afin d'estimer leur pertinence en termes d'adéquation entre la finalité du processus de numérisation et les usages qui en sont faits. Vu l'ampleur que prennent aujourd'hui les archives numériques et leurs impacts dans nos vies, nous devons interroger l'importance de ces archives tant au niveau personnel que collectif. Ces questionnements ne peuvent se concevoir que dans un travail collectif sur le long terme qui passe nécessairement par la mise en commun d'analyses et la constitution d'une intelligence collective capable de donner à l'humanité les moyens de ne pas laisser aux machines l'entière responsabilité des choix sémantiques nécessaires à l'interprétation des archives [CIT 10]. Des solutions existent déjà ou sont en cours de constitution, comme les initiatives menées par le W3C pour harmoniser les pratiques de gestion de l'information, d'autres restent à inventer tant d'un point de vue technique, qu'épistémologique, politique ou éthique.

1.2. L'intelligence collective

Il est assez trivial d'expliquer ce qu'est l'intelligence collective à travers l'analogie de la fourmilière [FER 97] ou de toute autre société d'insectes [PEN 06]. Cette conception entraîne une vision très partielle du phénomène d'intelligence collective et occasionne un positionnement éthique discutable dans le cas d'organisations humaines. La conception d'une intelligence collective à la manière des sociétés d'insectes tend à faire de l'humain participant à cette intelligence un être simple et basique, dont il faut soustraire toute complexité pour faire réagir chacun à la manière de tout le monde. Comme le remarque Bernard Stiegler, il y a là les enjeux d'une guerre pour le contrôle des sociétés à travers les symboles [STI 04]. C'est d'ailleurs une des critiques récurrentes vis-à-vis de l'intelligence collective qui n'aurait d'intelligent que le nom et ne servirait qu'à centraliser la mémoire pour mieux la contrôler sans permettre l'émergence de nouvelles connaissances [MAY 06].

Ce qui différencie l'humain de la fourmi, c'est sa capacité à réfléchir les flux d'information dans son intériorité et ainsi exprimer une conscience réflexive [LEV 11]. Comme le précise Yves Prié, la réflexivité est la faculté de rentrer en relation avec soi-même pour construire à partir de sa mémoire des représentations permettant la régulation de son action [PRI 11]. Cette définition qui place la réflexivité dans un contexte individuel peut toutefois être aussi comprise dans un cadre collectif où des

individus mettent en commun leur réflexivité pour travailler collectivement dans le respect des consciences de chacun. On trouve là les principes de base d'une science qui a pour vocation l'élaboration d'un consensus et qui nous permet de définir l'intelligence collective comme la mise en commun de la réflexivité afin de finaliser une action qui ne pourrait pas l'être par une seule personne.

Mais avant de pouvoir bénéficier du « pouvoir d'agir » [RAB 05] collectif, les acteurs doivent accepter de diriger leurs intérêts personnels vers une mise en commun altruiste de leur volonté. Ceci est possible en formalisant et en partageant des connaissances tout en acceptant leur validation par des contraintes collectives pour rendre le travail interopérable et réutilisable par une communauté. Toute la difficulté de l'intelligence collective réside dans cette capacité des individus à accepter de contraindre leurs expressions par un formalisme car il remet bien souvent en cause des habitudes de travail réflexif. Il ne faut pas se leurrer sur les motivations premières des humains qui ne vont pas nécessairement dans le sens du développement éthique d'une collaboration harmonieuse. Comme le remarque Yves Citton, il faut parfois utiliser des artifices pour faire évoluer les pratiques et les ancrer dans de nouvelles organisations sociales [CIT 08]. Il est assez symptomatique de voir que les recherches menées par Anita Woolley pour définir un facteur d'intelligence collective confirment que l'abandon des intérêts égoïstes au profit d'une démarche altruiste augmente la capacité d'intelligence collective d'un groupe. En effet, elle montre que l'intelligence de chaque individu est bien moins impactante que la sensibilité sociale des membres d'un groupe leur permettant notamment d'écouter et de ne pas monopoliser la parole [WOO 10].

La problématique de contraindre l'intelligence individuelle au profit de la finalisation d'une action collective passe aujourd'hui par des dispositifs techniques et particulièrement par des interfaces graphiques qui vont formaliser des systèmes sémiotiques dont l'objectif est de faciliter l'expression individuelle en corrélation avec les contraintes nécessaires au partage de cette expression. L'utilisation d'un langage informatique comme le Wiki est un exemple flagrant de ce passage par la contrainte pour faciliter l'interopérabilité d'une expression individuelle et finaliser le projet d'une encyclopédie. Ces projets d'intelligence collective ne s'arrêtent pas au seul langage informatique, ils entraînent avec eux tout un écosystème de connaissances au sein duquel ces projets pourront se développer par la finalisation successive des actions individuelles.

1.3. Les écosystèmes de connaissances

Ce sont les solutions à ces problématiques que nous allons analyser en prenant des exemples concrets dans des domaines aussi différents que l'innovation en entreprise ou les archives personnelles mais qui ont en commun l'usage de l'intelligence collective

pour l'exploitation des archives numériques. Pour donner une cohérence forte aux différents exemples et traiter toute la complexité des problématiques qu'elles présentent, nous analyserons les solutions suivant l'analogie des écosystèmes. Nous verrons dans ces solutions qui mettent en place des démarches d'intelligence collective en relation avec les usages des archives numériques, comment ces pratiques peuvent être analysées en concevant les informations non pas comme des objets inertes mais comme des êtres autonomes qui développent des modes d'existences particuliers [LAT 12].

L'objectif du modèle que nous proposons consiste à développer une méthode générique pour l'analyse des écosystèmes de connaissances qui composent un univers complexe de relations à la fois complémentaires et antagonistes entre une multitude d'existences humaines, mécaniques, institutionnelles, conceptuelles, documentaires... Avec ce modèle, nous souhaitons fournir aux chercheurs les moyens de décrire leurs champs d'investigation et les arguments qu'ils défendent par la modélisation d'existences informationnelles. Le but est de pouvoir rendre interopérables des analyses par la comparaison automatique de ces existences. Pour parvenir aux analyses comparatives de ces écosystèmes, nous modélisons les existences informationnelles en croisant les principes ontologiques spinoziens de Gilles Deleuze [DEL 68] avec ceux de Philippe Descola [DES 05]. En ce qui concerne Deleuze, nous reprenons les trois dimensions d'existence (parties extensives, rapports, essences) corrélées à trois genres de connaissance (chocs, logiques, intuitions). Chez Descola, nous utilisons les matrices ontologiques qui caractérisent les rapports entre physicalités et intériorités. Plus particulièrement, nous nous focalisons sur l'ontologie analogiste qui correspond tout à fait au cas des archives numériques et de l'intelligence collective, de par la capacité de transformation illimitée des physicalités numériques et la multiplicité des rapports d'intériorité que propose l'intelligence collective :

« Cette lutte continuée entre un océan vertigineux et des réseaux de relations toujours en train de multiplier leurs connexions définit en rigueur l'analogisme, mot qui résume et peint à merveille notre monde objectif, nos travaux cognitifs, nos rêves subjectifs ainsi que les collectifs qui naissent aujourd'hui et feront la politique du futur. » [SER 09, p. 85]

A l'aide de ces principes, nous composons des représentations uniques que nous qualifions de monades. Elles se composent de quatre ensembles : les documents, les acteurs, les concepts et les rapports. A l'intérieur de chaque ensemble, les éléments entretiennent des relations de sémantique différentielle [BAC 07, p. 142] suivant la position relative d'un élément par rapport à deux axes, celui du père et du frère dans un arbre.