

Zacklad, M. (2014). «Humanités numériques et digitalisation de la science», dans Actes du XIX congrès de la SFSIC, Toulon, 4-6 juin, <http://sfsic2014.sciencesconf.org/31853/>

Humanités numériques et digitalisation de la science

Manuel Zacklad

Dicen-IDF EA7339, CNAM

manuel.zacklad@cnam.fr

Pour comprendre les effets de la numérisation sur les pratiques scientifiques, nous considérons qu'il est nécessaire d'analyser les mutations actuelles en référence à la notion de document. La digitalisation s'accompagne d'une crise de la documentalité, d'une nouvelle manière de faire document dans les trois étapes que sont : (1) la construction du matériel empirique (données/corpus), (2) la production des résultats à partir de celui-ci et (3) la présentation et publication de ces résultats. Pour illustrer notre thèse, nous présenterons deux exemples relevant des étapes (2) consolidation participative d'un corpus sur la biodiversité dans le cadre du projet e-recolnat et (3) la mise en place d'un événementiel transmédia pour l'organisation d'une conférence dans le cadre des ENMI¹.

Introduction

Dans les nombreuses publications s'inscrivant dans le courant des Digital Humanities (DH ou Humanités Numériques) la plupart visent à présenter des cas d'applications dans des disciplines spécifiques en faisant une part plus ou moins importante à la conception de dispositifs informatiques originaux. D'autres travaux articulent le potentiel de certaines technologies numériques avec des démarches méthodologiques dans certaines disciplines comme nous l'avons fait en présentant les apports du web socio-sémantique à la recherche en SHS (Zacklad 2007). Enfin, un troisième type de contribution cherche à cerner la spécificité des démarches de DH en proposant des cadres conceptuels ou théoriques permettant de définir ce qui relève ou non des DH. C'est notamment le cas du livre de Lunefeld et al. 2012 (cf. infra) dont nous discuterons certaines propositions dans la première partie. C'est également la perspective qui est la nôtre dans le cadre de cette contribution dans laquelle nous chercherons à proposer une définition originale de la spécificité des projets de DH et de leur dynamique basée sur les notions de digitalisation et de documentalité.

Définition du DH selon Lunefeld et al. et nouvelles hypothèses pour la définition du DH et de la science en environnement numérique

Comme l'expliquent clairement Lunefeld, Burdik, Drucker, Presner et Schnapp (2012) dans un des traités les plus aboutis sur les humanités numériques, il n'est plus possible de se contenter aujourd'hui d'une définition de ce champ qui ferait seulement référence à l'utilisation de l'informatique dans l'activité académique.

“Two decades ago, working with digital documents was the exception. Today it is the norm, the “natural” environment for carrying out research, teaching, and reading. Wireless networks have consigned the off-the-grid, off-line classroom to the dustbin of history. If the novelty of Digital Humanities work has already been absorbed into day-to-day business, then what claims can it make for introducing new insights or methods

¹ Cette recherche a bénéficiée du financement de l'Infrastructure Nationale en Biologie et Santé e-ReColNat dans le cadre des investissements d'avenir.

into research for the broader fields of the humanities? The practice of Digital Humanities cannot be reduced to “doing the humanities digitally”; both criticality and experimentation must shape its future development.” (Lunefeld et al. 2012)

A l'issue d'une argumentation stimulante et provocatrice, ces auteurs finissent par considérer que c'est essentiellement la pratique du « design » pour la conceptualisation et la production de la connaissance qui doit être au cœur d'une définition renouvelée du DH, celui-ci constituant la nouvelle rhétorique des humanités. La maîtrise d'une « littératie multimédia » appuyée sur un savoir-faire en conception, implique des compétences diverses combinant la maîtrise pointue d'une sous-discipline dans le champ des humanités avec d'autres compétences relevant du design, de l'informatique, de la curation et des sciences de l'information (library science).

Tout en rejoignant ces auteurs dans plusieurs de leurs thèses, nous pensons néanmoins qu'ils échouent à définir de manière entièrement satisfaisante la spécificité des humanités numériques et plus généralement de la science en environnement numérique généralisé. Dans cette présentation, nous défendons donc de nouvelles hypothèses permettant de définir les DH et plus largement la science en environnement numérique.

Ne pas confondre informatisation et digitalisation

A la suite de Milad Doueïhi (2008), nous considérons que les transformations introduites par le numérique induisent une évolution culturelle majeure, une conversion pour reprendre les termes de cet auteur, dont les implications anthropologiques sont multiples. De ce fait, il critique une certaine vision des DH qui verrait uniquement dans le numérique des outils de traitement qu'il faut mettre au service des SHS (Doueïhi 2011), une vision qui corollairement peut tendre à mettre trop l'accent sur la dimension technologique des DH. Car la digitalisation n'est pas l'informatisation entendue comme la mobilisation des nouvelles technologies au service de tâches pour en améliorer l'automatisation et l'efficacité. Comme nous l'expliquerons, la digitalisation correspond à une présence des supports numériques comme un « milieu » qui ne concerne pas des tâches professionnelles spécialisées mais dans lequel un ensemble de pratiques quotidiennes sont réalisées. En retour, c'est le secteur d'activité dans son ensemble qui est impacté par le milieu numérique, du point de vue des caractéristiques des ressources qui l'alimentent, de la réalisation des activités productives et des relations avec les bénéficiaires. La même analyse s'applique dans le contexte de la digitalisation de la science. La numérisation ne désigne pas simplement l'utilisation d'une application informatique dans une procédure de recherche mais le fait que la totalité de la recherche, la constitution du matériel empirique, les formes de collaboration entre chercheurs, la diffusion des résultats, est impactée par le milieu numérique.

Digitalisation et documentalité

Pour comprendre les effets de la numérisation sur les pratiques, nous considérons qu'il est nécessaire d'analyser les mutations actuelles en référence à la notion de document dans la ligne de la stratégie initiée par le collectif Pédaque (2006). La digitalisation s'accompagne selon nous d'une crise de la documentalité, d'une nouvelle manière de faire document qui impacte un très grand nombre de pratiques. Rappelons que selon la définition de M. Zacklad (2004), la notion de document désigne tout support d'écriture ou d'enregistrement qui a fait l'objet d'un travail de documentarisation, d'une mise en forme spécifique visant à permettre la circulation du support dans l'espace et le temps, c'est-à-dire à le constituer en support de mémoire et en médium pour la coopération à distance. Dans un travail récent, M. Zacklad (2015) définit cinq dispositifs de médiation documentaire diffusionnelle, rédactionnelle, contributive, attentionnelle en mode flux et enfin transmédia combinant une diversité de supports et de plateformes étroitement imbriqués dans le cadre de la documentarisation coordonnée

d'un patrimoine public ou privé, d'un processus ou d'un projet, d'une conférence ou d'un évènement...

Pour la science, dont les enjeux de mémoire et de coopérativité sont particulièrement importants, ces mutations de la documentalité sont susceptibles d'introduire des transformations radicales dans les trois étapes stratégiques que sont :

- La construction du matériel empirique (données/corpus) : la documentation des sources empiriques de la recherche qui diffèrent entre le domaine des humanités (prédominance du corpus) et des sciences sociales (importance des données) avec des rapprochements forts dans le domaine des recherches qualitatives ;
- La production des résultats à partir du matériel empirique : les processus de documentation de la construction individuelle et/ou collective des résultats, faits, interprétations, conjectures théoriques, particulièrement influencés par les modalités d'organisation des connaissances et les modes de raisonnement qu'ils induisent ;
- La présentation et publication des résultats : les processus de documentation associés à la diffusion, communication, mise en débats de ces résultats dans différentes communautés plus ou moins hétérogènes.

Dans le cadre de cette contribution nous donnerons deux exemples illustrant de manière partielle les différentes activités associées aux étapes de production et de présentation des résultats. Le premier exemple de production des résultats est le projet E-Recolnat qui relève en partie des sciences citoyennes et qui correspond à la consolidation collaborative d'un matériel empirique déjà constitué pour en faciliter l'exploitation collective. Le deuxième exemple, présentation et publication, emprunté à L. Merzeau (2013), décrit les formes de présentation et de mise en débat originales dans le cadre des Entretiens du Nouveau Monde Industriel (ENMI).

Consolidation participative d'un corpus sur la biodiversité et station de « lecture savante »

Le projet e-Recolnat (Pignal et Perez 2013, Zacklad et al. 2014) vise à constituer des collections en ligne à partir de la numérisation des collections naturalistes de plusieurs institutions françaises et notamment de celles du le Muséum National d'Histoire Naturelle (MNHN). Nous nous intéresserons principalement ici aux collections d'herbiers qui sont les premières à faire l'objet de la numérisation. Différents objectifs sont poursuivis. A la mission de conservation de spécimens pour les besoins scientifiques, s'ajoute un double enjeu patrimonial : les herbiers constituent des témoignages de la biodiversité et de son évolution mais aussi de l'histoire de la botanique.

Or, le processus numérisation implique non seulement de scanner les images mais également d'être en mesure de traiter les données qui y sont associées, écrites à la main, en abrégé le plus souvent, sur des étiquettes collées à côté des spécimens naturels. Ces inscriptions échappent au processus de reconnaissance optique de caractères (OCR) mais sont indispensables pour permettre l'exploitation de la collection numérique dans un contexte où les inventaires informatisés des planches d'herbier disponibles sont très peu nombreux. Le recours aux amateurs est donc apparu donc comme une opportunité de premier plan pour effectuer ce travail d'inventaire qui est aujourd'hui hors de portée financière pour la quasi-totalité des herbiers. Il vise à encourager les pratiques de sciences citoyennes dans un contexte où les outils du Web 2.0 donnent un nouveau souffle aux réseaux d'amateurs, dont la participation à la collecte et à la documentarisation de spécimens est ancienne.

Dans l'optique de notre projet, le travail de redocumentarisation implique trois types d'opérations de nature différentes : (1) une redocumentarisation « littérale », dans laquelle les amateurs vont essentiellement procéder à une retranscription d'informations scripturales dans un format alphanumérique, (2) une redocumentarisation « inférentielle » dans laquelle ils actualiseront les informations en fonction du contexte actuel en fournissant par exemple le nom

moderne d'une localité à partir de sa dénomination historique, et enfin (3) une redocumentarisation créative, correspondant à une documentarisation inédite des planches selon des points de vue associés à des savoirs vernaculaires absents de la planche initiale (Zacklad et al. 2014).

Indépendamment de la dimension patrimoniale que nous n'examinons pas dans le cadre de cet article, la collection ainsi numérisée et redocumentarisée offrira de nombreux avantages aux scientifiques. La redocumentarisation, en enrichissant les systèmes d'organisation des connaissances disponibles, leur permettra d'effectuer des recherches plus précises et rapides dans la collection numérisée. Une fois les planches sélectionnées, ils disposeront d'un ensemble d'outils d'analyse et d'annotation leur permettant d'étudier les spécimens et de confronter leurs analyses à celles d'autres chercheurs travaillant sur les mêmes objets (outils de zoom, de mesure en ligne, d'annotation collaborative, etc.). Mais malgré l'intérêt que présentera l'accès aux collections numérisées, celles-ci ne pourront pas remplacer l'accès aux spécimens physiques. En effet, dans le domaine des collections naturalistes, cet accès reste indispensable du fait des caractéristiques biologiques du spécimen, dont l'ADN est le meilleur exemple.

Le projet E-Recolnat montre bien à quel point, avec les projets de numérisation, les sciences naturelles rentrent de plein pied dans la science digitale et en particulier dans le régime de documentalité du transmédia documentaire caractérisé par la coexistence d'une diversité de supports numériques et non numériques relevant chacun de plateformes et de systèmes d'organisation des connaissances hétérogènes qui doivent être coordonnés voire enrichis au cours des activités individuelles et collectives. Cette diversité conditionne bien sur les raisonnements scientifiques qui s'appuient sur les associations possibles entre sources empiriques, éléments de bibliographie, chercheurs du réseau.

Cet exemple manifeste bien plusieurs caractéristiques du transmédia documentaire dans la gestion et l'exploitation d'un corpus de collections naturaliste :

- Existence du corpus sur deux supports, physique et virtuel, relevant de modalités d'organisation des connaissances hétérogènes mais devant être coordonnées dans le cours de l'activité scientifique ;
- Dispositif de sciences citoyennes pour la documentarisation du corpus faisant évoluer le Système d'Organisation des Connaissances (SOC) sur une plateforme différente de celle utilisée pour la consultation scientifique avec passerelle logicielle entre les systèmes ;
- Evolution en continu du SOC sous la double influence de l'activité scientifique de systématique et de la redocumentarisation citoyenne, une indexation sociale contrôlée portant sur les « métadonnées » (redocumentarisation inférentielle et créative).

Conformément à notre vision transactionnelle des documents comme artefacts médiateurs, qui articulent les dimensions communicationnelles et informationnelles, cet exemple montre également à quel point l'enrichissement de l'organisation des connaissances qui sera rendu possible par les plateformes transmédia est lui-même la conséquence de la richesse des transactions entre des acteurs hétérogènes suscitée par le dispositif : transactions entre scientifiques et amateurs via une plateforme pouvant être assimilée à un jeu sérieux connectée à une base de données scientifique pour la redocumentarisation des planches et transactions entre scientifiques permettant de confronter des hypothèses scientifiques en débat via des outils d'annotation collaboratives.

Évènementiel transmédia pour l'organisation d'une conférence

Les Entretiens du nouveau monde industriel (ENMI) sont organisés tous les deux ans par l'IRI, le pôle de compétitivité Cap digital et l'École nationale supérieure de création indus-

trielle (ENSCI)². Ils ont été récemment étudiés par Louise Merzeau (2013) comme relevant de l'Éditorialisation collaborative d'un événement et ils constituent comme nous allons le voir un excellent exemple du régime de documentalité du transmédia documentaire appliqué à la présentation de résultats de nature scientifique. Toutes les conférences scientifiques avec actes relèvent du transmédia en ce qu'elles articulent un dispositif de rencontre social et de présentation orale avec la mise à disposition sur un support écrit des textes en version longue et du programme des sessions. Mais l'exemple des ENMI témoigne d'un enrichissement considérable du dispositif de transmédia documentaire (contenus consignés sur support pérenne) tant du point de vue de la diversité des systèmes d'organisation des connaissances mobilisés, de l'hétérogénéité des plateformes, que de la multiplicité des agencements transactionnels associés.

Comme le décrit Louise Merzeau, les ENMI sont articulés autour de deux ensembles applicatifs, le premier précédent et prolongeant les événements, le second les accompagnants pendant leur durée. Le premier ensemble est constitué par :

- « un graphe *Pearltrees*³ [...], dont la structure épouse le découpage du colloque en sessions, rassemble une documentation sur chaque conférencier (biographie, abstract et ressources de la conférence, controverse, sites personnels...) ;
- un glossaire⁴ interactif [...]
- une carte heuristique *Freemind*⁵ [...] spatialise les cheminements argumentatifs de chaque conférence.
- un *Storify*⁶ du séminaire préparatoire organisé quelques mois avant le colloque, lui-même relié à une interface *PolemicTweet*⁷ [...] synchronisant le live tweet des participants sur le streaming vidéo des interventions. » (Merzeau 2013)

Le second ensemble est constitué par :

- « une plateforme *PolemicTweet*, [...]
- une plateforme *Unishared*⁸ de prise de notes collaborative en temps réel ;
- une plateforme *Sharypic*⁹ (fig. 5) accueillant la couverture photographique de l'événement sur scène, dans la salle ou en coulisses ;
- un *Storify*¹⁰ où chaque session est résumée. » (Merzeau 2013).

Ce deuxième exemple montre comment le régime de documentalité transmédia est susceptible de transformer les formes de diffusion des résultats scientifiques et au-delà, bien sûr, d'amplifier les processus de production coopérative des connaissances durant les événements, ce qui est toujours un des objectifs poursuivis lors de l'organisation de ces manifestations. Le dispositif des ENMI possède ainsi plusieurs caractéristiques :

- Avant la conférence, il y a production de documents numériques originaux qui visent à faciliter les interactions entre les participants par un travail préalable de structuration de leur biographie, d'organisation de mots-clefs et de production de glossaire. Ces do-

² <http://www.iri.centrepompidou.fr/actualites/entretiens-du-nouveau-monde-industriel-2012-digital-studies-organologie-des-savoirs-et-technologies-industrielles-de-la-connaissance/>

³ <http://enmi12.org/pearltrees-enmi12/>

⁴ <http://www.enmi12.org/glossaire/>

⁵ <http://enmi12.org/freemind/enmi12.html>

⁶ <http://storify.com/IriResearch/preparatory-workshop-for-enmi-2012>

⁷ Sur l'outil *PolemicTweet*, voir <http://www.iri.centrepompidou.fr/outils/polemic-tweet/>

⁸ <http://unishared.com/enmi/>

⁹ <https://www.sharypic.com/enmi12/>

¹⁰ <http://enmi12.org/livestorify/>

cuments produisent une sorte de cartographie du champ, une organisation des connaissances sémiotique et terminologique, qui est utile pour les orateurs comme pour les auditeurs ;

- Pendant la conférence, il y a mise à disposition de plateformes d'annotations qui permettent d'échanger des commentaires et des questions de manière à faciliter « *un processus d'individuation collective où s'élabore un « nous » simultanément énonciatif et réceptif* » (Merzeau 2013). Dans les ENMI, les auditeurs ne sont pas incités à contribuer de manière directe comme dans les BarCamp mais la mise en visibilité des commentaires dans des outils de type « *polemic tweet* » peut inciter à rebondir de manière constructive ;
- Du point de vue de l'architecture des ensembles applicatifs mobilisés, l'hétérogénéité des supports utilisés durant les ENMI illustrent bien le régime du transmédia documentaire. Comme le souligne Louise Merzeau « *chaque application privilégie un outil, une temporalité (avant, pendant, après), une forme sémiotique (image, texte, oralité) et une modalité participative spécifique (commentaire, annotation, documentation, témoignage, archivage...)* » (Merzeau 2013). Cette richesse à simultanément pour corollaire une dépendance vis-à-vis d'applications tierces dont la pérennité n'est pas nécessairement garantie. Mais il s'agit d'une caractéristique qui n'est pas nécessairement problématique dans le contexte d'un transmédia évènementiel centré sur une conférence qui n'appelle pas le même travail dans la durée que l'usage d'un corpus transmédia a vocation patrimoniale ou de preuve scientifique.

Analyser la digitalisation comme une succession d'innovations incrémentales et d'usage essentiellement ascendantes

Ces exemples, auxquels nous n'avons pas pu associer ici la problématique de la construction transmédia des données empiriques, montrent à quel point la digitalisation et les évolutions documentaires associées induisent des mutations épistémologiques dans le champ scientifique et particulièrement dans celui des SHS que nous analysons. En transformant les conditions de production et de diffusion des savoirs, ces évolutions sont aussi importantes que d'autres changements paradigmatiques récents dont les sciences cognitives, la biologie moléculaire, ou le développement des sciences informatiques, que nous différencions de la digitalisation, ont constitué des exemples récents.

Mais alors que ces changements précédemment évoqués s'appuyaient sur des théories fortes susceptibles de donner lieu à la construction de grands programmes ou de grands équipements, la digitalisation relève d'une forme de transformation paradigmatique ascendante et opportuniste qui la rend beaucoup plus difficile à saisir et qui ne permet pas d'en anticiper aisément les évolutions. En effet, comme dans les autres secteurs d'activité où elle se diffuse, la digitalisation commence par prendre appui sur des pratiques quotidiennes, en s'appuyant sur des outils qui relèvent autant de la sphère domestique que de la sphère professionnelle. Les nouveaux supports mobilisés pour « faire document » impactent les modalités de mémorisation individuelle et collective comme les formes de socialisation qui sont en interaction constante avec des artefacts externes et dont les évolutions conditionnent lourdement l'activité des chercheurs.

Bibliographie

Ian Hacking, 2001, *Entre science et réalité. La construction sociale de quoi?*, Paris, La Découverte, p. 72-74

Doueïhi M. (2008). *La Grande conversion numérique*, Paris: Le Seuil.

Doueïhi, M. (2011). *Pour Un Humanisme Numérique*. Paris: Le Seuil.

- Lunefeld, P., Burdick, A., Drucker, J., Presner, T., & Schnapp, J (2012). *Digital_Humanities*. The MIT Press. <http://mitpress.mit.edu/books/digitalhumanities-0>.
- Merzeau, L. (2013). Éditorialisation collaborative d'un événement, *Communication et organisation*, 43 | 2013, 105-122.
- Pédauque, R.T. (2006). *Le document à la lumière du numérique*, C&F Éditions Septembre
- Pignat, M., Perez, E. (2013). Numériser et promouvoir les collections d'histoire naturelle, *BBF*, 5: 27-30.
- Zacklad, M., (2004). Processus de documentarisation dans les Documents pour l'Action (Do-pA), in *Le Numérique : impact sur le cycle de vie du document*, Colloque ENSSIB. <http://www.enssib.fr/bibliotheque-numerique/document-1209>
- Zacklad, M. (2007), Démarche de caractérisation des collaboratoires en SHS et apports du web socio-sémantique, présentation au séminaire « Economie politique de l'édition numérique » de l'ENS Lyon du 4 décembre 2007 - <https://www.academia.edu/5439110/>
- Zacklad, M., Chupin, L., Bertin, G., Payeur, C. (2014) La documentarisation participative au service de la patrimonialisation des collections naturalistes, *Cahiers de la SFSIC*
- Zacklad, M., (2015). Genre de dispositifs de médiation numérique et régimes de documentalité, dans L. Gagnon-Arguin, S. Mas, D. Maurel (dir.), *Les genres de documents dans les organisations, Analyse théorique et pratique* (p. 145-183), PUQ, Québec.