



Pour une approche info-documentaire mixte numérique/non-numérique

Chloé Girard

► **To cite this version:**

Chloé Girard. Pour une approche info-documentaire mixte numérique/non-numérique. Balisages, Enssib, 2020, 10.35562/balisages.292 . hal-02860936

HAL Id: hal-02860936

<https://hal-bnf.archives-ouvertes.fr/hal-02860936>

Submitted on 10 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

POUR UNE APPROCHE INFO-DOCUMENTAIRE MIXTE NUMÉRIQUE/NON-NUMÉRIQUE

Chloé Girard

Doctorante, Université Paris-Nanterre; Laboratoire Dicen-IDF Cnam (EA 7339)
chloe.girard@parisnanterre.fr

Nous proposons de revenir sur le présupposé selon lequel des «objets numériques instables» devraient ou pourraient être pensés ensemble en tant qu'«objets documentaires numériques». D'une part, le flux, l'instabilité ne sont pas exclusifs au numérique. La documentation des objets instables, êtres vivants comme logiciels, consiste en capture d'états et en descriptions de différentes natures. Ce caractère n'est donc pas un argument fort qui justifierait de séparer objets numériques et non numériques dans une approche info-documentaire. D'autre part, si certaines opérations telles que la collecte ou la mise à disposition sont communes à certains objets parce que numériques, elles ne suffisent pas à les distinguer radicalement des objets dits «traditionnels». Les problématiques se recoupent d'un «espace» à l'autre sans fracture nette. Mais c'est sur le plan de la classification qu'il semble le moins pertinent de considérer ensemble différents objets sur la base de leur espace technique. Nous proposerons au contraire ici d'intégrer objets numériques et non-numériques dans une classification mixte. Une telle approche implique, sur le plan de la classification info-documentaire, d'abandonner l'idée d'«objets numériques instables», et jusqu'à celle d'«objets numériques».

Mots-clés: document, espace technique, stabilité, instabilité, classification

We propose to revisit the assumption that “unstable digital objects” should or could be thought of together as “digital documentary objects”. On the one hand, the flow, the instability, are not exclusive to digital. The documentation of unstable objects, both living beings and software, consists of state captures and descriptions of various kinds. This character is therefore not a strong argument that would justify separating digital and non-digital objects in an info-documentary approach. On the other hand, while some operations such as collection or provision are common to certain digital objects, particularly in technical terms, they are not sufficient to distinguish them radically from so-called “traditional” objects. The problems overlap from one “technical space” to another without a clear fracture. But it is in terms of classification that it seems least relevant to consider different objects together on the basis of their technical space. On the contrary, we will propose here to integrate digital and non-digital objects in a mixed classification. Such an approach implies, in terms of info-documentary classification, abandoning the idea of “unstable digital objects”, and even that of “digital objects”.

Keywords: document, technical space, stability, instability, classification

L'appel à contributions de ce premier numéro de la revue *Balisages* propose de se pencher sur la manière d'approcher, sur le plan info-documentaire, des objets numériques dits instables, des « productions numériques non stabilisées, qui s'insèrent dans une logique de flux »¹. « Quels enjeux posent ces objets numériques [...] par rapport aux objets documentaires traditionnels ? »² Ces questions suggèrent, d'une part, que les « objets numériques instables » peuvent constituer une catégorie info-documentaire et, d'autre part, que « les objets numériques » peuvent, ou doivent, se penser séparément des « objets documentaires traditionnels ». Ces hypothèses sous-tendent la constitution d'une liste d'objets aussi disparates que la suivante, que nous pourrions penser ensemble : application Smart City, visualisation 3D, flux de tweet, logiciel, données numériques, code informatique, documents numériques, données de la recherche, codes sources, simulations, données issues de réseaux sociaux, services de vidéo à la demande, données issues d'objets connectés, pages Web, Wikipédia, données de la recherche, produits culturels, objets scientifiques, objets patrimoniaux.

Ce sont ces deux prémisses que nous proposons ici de discuter, voire de renverser. Les objets numériques instables ne sont pas moins documentarisables que les objets non numériques instables, et l'espace technique ne justifie pas *a priori* de séparation dans la classification des objets documentaires.

L'INSTABILITÉ COMME ARGUMENT INFO-DOCUMENTAIRE ?

La documentation des objets instables commence précisément en en stabilisant des états : captures photographiques, dessins, cartels de musée, captations vidéos, etc. Au-delà de ce point, qu'il s'agisse d'une antilope ou d'une partie de jeu d'échecs, le traitement des objets documentaires, des dites captures, est similaire. Si l'on accepte qu'un document est un non-événement [Escarpit, 1991], alors ce que l'on appelle des « productions numériques non stabilisées » n'en sont nécessairement pas. L'antilope n'est pas un document mais des captures de celle-ci en sont [Briet, 1951, § 5]. C'est donc aussi du côté de ces « captures » qu'il faut se tourner quant aux objets numériques. Prenons le cas d'un objet logiciel. Le 10 juillet 2019, WordPress déclarait 359 *releases*

1. « Appel à articles n° 1, Les objets nativement numériques : transformations et nouveaux enjeux documentaires ? ». < <https://publications-prairial.fr/balisages/index.php?id=266> >.

2. *Ibid.*

et 39789 *commits*³ sur son dépôt GitHub⁴. Son histoire a dû en connaître plus, la première *release* du logiciel datant de mai 2003⁵ et GitHub n'ayant été fondé qu'en 2008⁶. L'œuvre logicielle WordPress n'est donc pas stabilisée, c'est un objet « en flux ». Cependant, chacun de ses états, ou manifestations, publiés en tant que *release* est stable. L'œuvre, abstraction de tous ses états, doit évoluer à mesure que les logiciels, librairies et équipements dont elle dépend évoluent, mais la raison d'être de chacune de ses *releases* est précisément la stabilité, sur laquelle fonder toute la documentation d'usage et de déploiement. Ce schéma éditorial peut être comparé à celui du livre, dont l'analyse cache souvent une confusion entre *titre* et *état*, entre *objet titre* (œuvre) et *objet documentaire* (manifestation). *L'origine des espèces*, de Charles Darwin, possède une histoire éditoriale riche, une forte dynamique : 6 éditions en anglais du vivant de l'auteur, entre 1859 et 1876, 9 en comptant les éditions augmentées et corrigées. De nombreuses traductions, de son vivant également. 635 éditions sont déclarées pour ce titre sur le site Darwin Online⁷, dont la dernière date de 2018. Dire que le livre est stable, c'est confondre l'œuvre et la manifestation. L'œuvre *L'origine des espèces* est instable, et n'est pas un document. Chacun de ses états, chacune de ses « captures », imprimée ou numérique, est, quant à elle, stable et constitue un document. *L'origine des espèces* n'est pas un document mais “*On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*”, London: John Murray, 24 November 1859, en est un.

En ce qui concerne la question de stabilité du livre, œuvre et manifestation sont souvent confondues. Avec l'arrivée des objets numériques, nous avons en conséquence comparé une œuvre (un logiciel, un site web, etc.) avec une manifestation (un livre imprimé), concluant naturellement que le premier, numérique, était instable alors que le second était stable, jusqu'à en conclure que les objets numériques sont essentiellement instables au contraire des objets traditionnels. Mais si l'on compare une œuvre logicielle avec une œuvre littéraire, et un exemplaire imprimé avec une *release* logicielle (imprimée ou numérique), on s'aperçoit du parallèle entre ces deux types d'objets. Tous les deux sont instables (selon des fréquences qui peuvent être très différentes) et leurs auteurs choisissent de fixer des états stables, qui font document. En termes info-documentaires, si fracture il y a entre ces différents objets elle

3. Les *releases* sont des sorties officielles, numérotées et documentées, les *commits* sont des modifications régulières entre deux *releases* officielles.

4. Dépôt officiel WordPress sur GitHub. < <https://github.com/WordPress/WordPress> >.

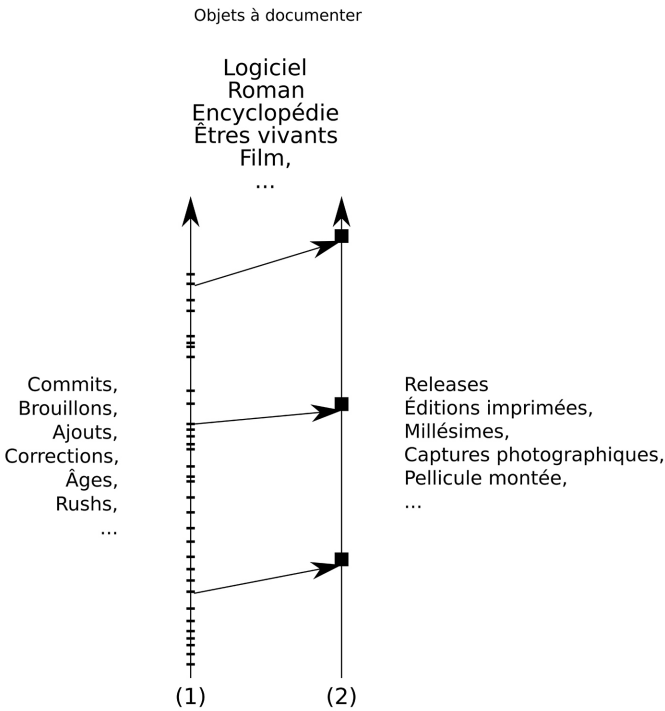
5. Article WordPress, Wikipédia. < <https://en.wikipedia.org/wiki/WordPress> >.

6. Article GitHub, Wikipédia. < <https://en.wikipedia.org/wiki/GitHub> >.

7. Site Darwin Online. < <http://darwin-online.org.uk/> >.

n'est pas basée sur leur soi-disant instabilité : un livre imprimé et une *release* logicielle sont tous les deux des objets stables issus d'objets instables.

Figure 1. Objets flux et objets documentaires stables, numériques ou traditionnels



Chaque état, qu'il soit de la branche 1 ou 2 est stable. Seul l'objet-titre évolue.

Source : Chloé Girard.

Sur la figure 1, les habitués reconnaîtront la schématisation propre aux travaux versionnés dans un logiciel de gestion et de fusion de versions de fichiers tel que Git. Ce type de logiciels et leur logique de branches, *commits* et *releases*, initialement destinée au développement logiciel, montre ce que l'on ne met pas si souvent en image : le flux, discrétisé en états multiples, de développement (les brouillons, *commits*, âges, etc.) ou de référence (les « sorties » ou manifestations)⁸. Une version, n'est finalement qu'un état du développement

8. On notera avec intérêt l'émergence de nouveaux services tels que GitBook (< <https://www.gitbook.com/> >). Cette plateforme s'empare heureusement de cet outil pour la gestion de versions d'écrit de type « livres » alors que dans l'esprit général, il est encore associé exclusivement au code logiciel.

que les auteurs distingueront parmi d'autres pour faire référence. Ce schéma, flux éditorial et sorties documentaires, s'applique à de nombreux objets sinon à tous, antilope comprise. Une antilope, ou quelque être vivant que ce soit, est un objet instable dont les versions sont les différents âges (pour donner une échelle arbitraire dans un flux continu) et les manifestations documentaires quelques captures que ce soit. Une partie d'échecs, qu'elle soit numérique ou non, est une œuvre instable dont les différentes manifestations stables sont les différents coups documentés. Le jeu d'échecs lui-même est une œuvre instable dont les manifestations sont les différentes règles rédigées au cours des époques et des zones géographiques. Un document est toujours un instantané, final ou intermédiaire, dans l'évolution d'un objet, qu'il soit œuvre de l'esprit ou de la nature.

La question de la stabilité ou de l'instabilité des objets, numériques ou non numériques, n'est donc pas un argument de classification info-documentaire.

UN « ESPACE TECHNIQUE » COMME ARGUMENT DE CLASSIFICATION ?

Au-delà de ce point de la capture, de la manifestation, qui place tous les objets au même niveau quant à la question de la stabilité, il s'agit maintenant de penser ensemble ou séparément *manifestations numériques* et *manifestations non-numériques*, *release* logicielle par exemple et volume papier. Devons-nous les penser séparément du fait de leur appartenance à deux espaces techniques différents ? Est-ce qu'un espace technique et d'usage commun [Epron et Vitali-Rosati, 2018], l'espace numérique, définit un espace documentaire distinct ? Rien ne le suggère *a priori*. Tous les objets documentaires dits « traditionnels » ne sont d'ailleurs eux-mêmes pas traités ensemble sur la base de leur « espace technique ». On ne réfléchit pas ensemble une chaîne de télévision et une antilope, un livre et un spectacle vivant sur la base seule de leur « espace technique commun » qui serait l'espace « non numérique ». Devons-nous, et pouvons-nous penser les *objets numériques* sur cette base et ainsi les opposer aux *objets non numériques* ? Cette hypothèse implique que les opérations info-documentaires telles que la conservation, collecte, publication/rendu, concernant les objets numériques seraient propres à ceux-là au point qu'elles les rendraient incommensurables aux objets non numériques.

La conservation des objets techniques, supports d'inscription ou outils de rendu, qu'ils soient analogiques ou numériques, demande dans les deux cas du soin et de la portabilité. L'expérience avec de nombreux supports et standards analogiques (Pal, Secam, VHS, Betacam, etc. sont exemplaires pour le cas des bandes magnétiques audiovisuelles) le démontre. L'évolution des formats et

logiciels de traitement s'accélère sans doute mais la migration successive de support n'est pour autant pas propre au seul numérique [Jacobson, Larrousse, et Massol, 2014]. En ce qui concerne la collecte, les objets dits traditionnels, du livre au jeu de plateau, de la céramique au film, sont pour beaucoup d'entre eux, produits en de nombreux exemplaires, de nombreuses copies identiques, et distribués. Les exemplaires sont répartis géographiquement et socialement par le jeu du commerce, du don, etc. La collecte en est d'autant plus aisée et les occurrences conservées plus nombreuses. Pensons aux exemplaires papier de journaux quotidiens dans de très nombreuses bibliothèques. Or, l'une des particularités des objets numériques est, pour beaucoup d'entre eux, la possibilité d'être mis à disposition en réseau et d'être accessibles depuis quelque point du globe que ce soit, sauf censure, qu'il s'agisse d'accès gratuit ou non.

C'est précisément la logique du *cloud*. Cette mise à disposition se passe très bien de la distribution, voire organise la non-distribution, et induit en retour une certaine centralisation des sources, avec ce que cela implique en matière de difficulté de collecte [Girard, à paraître en 2020]. Des objets dits « traditionnels » sont également produits en très petites quantités ou à l'unité, qu'il s'agisse d'une œuvre d'art, d'un manuscrit ou d'une bague célébrant la victoire d'une équipe de basket. Mais ces objets circulent, changent de main, se lèguent, s'offrent, se vendent aux enchères, alors que de plus en plus d'objets numériques ne se présentent qu'en accès, ne changent pas de mains, et peuvent se trouver remplacés sans aucune forme de procès. Il y a dans la collecte, ou l'absence de possibilité de collecte, des objets numériques un risque info-documentaire réel, auquel n'échappe aucun des objets mentionnés plus haut et qui risque d'en faire des « éphémères » au même titre que nos prospectus, billets de spectacles, cartes postales ou bulletins de vote papier [Derrot, 2016]. C'est peut-être ici la seule caractéristique info-documentaire que l'on pourrait envisager comme exclusive aux objets numériques, leur potentielle non-distribution, ni dans l'espace (sous différentes autorités à un temps donné) ni dans le temps (objets non héréditaires). Les questions de la lisibilité des fichiers sont également un aspect info-documentaire extrêmement lié à la technique. Tous les fichiers numériques peuvent être cryptés. Aussi distribués seraient-ils, ils risquent de rester illisibles. L'ouverture ou la fermeture des sources, leur accessibilité en lecture ou non, est un second facteur de risque info-documentaire commun à tous les objets numériques. Ce risque est sans doute moindre pour les objets non numériques. Quoique l'on puisse citer la machine Enigma ou le manuscrit de Voynich, la cryptographie avant l'ère numérique restait une affaire de spécialiste. Il fallait être spécialiste pour crypter et décrypter, quand il ne faut plus aujourd'hui qu'être

spécialiste pour décrypter. Les outils de cryptage sont à la disposition de chacun, quoique sans doute pas les plus forts. Le rendu d'un objet documentaire numérique est reproduit, recalculé, lors de chaque consultation. Mais tous les objets numériques n'ont pas un rendu complexe, voire pas de rendu du tout. C'est également le cas de certains objets non numériques qui ne sont perceptibles que par leurs traces, comme dans le cas d'objets stellaires ou subatomiques, objets théoriques non directement visibles. Le rendu de certains autres change régulièrement, qu'il s'agisse d'être vivants ou de cours d'eau. D'autre part, tous les objets non numériques n'ont pas un rendu stable. Les objets projetés (film, sons) sont tributaires des outils de projection (projecteur, écran, lumière, enceintes).

La question du rendu est complexe, dans les deux systèmes techniques. Elle demande à être traitée au cas par cas, certains objets numériques se rapprochant de la problématique documentaire du spectacle vivant, d'autres du film ou de l'écrit. En termes de sélection et d'oubli, les questions à la base de notre documentation, quoique réactivées, restent : doit-on tout documenter parce que nous le pouvons ? Ramener la documentation du flux à quelques *releases* ou captures peut sembler réducteur, voire simpliste, alors que nous pouvons éventuellement tout capter⁹. Documenter tous les états d'un flux comme on le fait avec un logiciel comme Git semble raisonnable en période de préproduction. Git permet de ne rien perdre, de revenir en arrière, de corriger ses erreurs. Mais qu'en est-il pour un projet d'archivage tel que le *Software Heritage Project* qui conserve *releases* et *commits* ? Dans le même ordre d'idées, notons que le slogan de la plateforme GitBook est *Document everything* (Documentez tout). Pour un documentaliste, et pour un auteur, cela ressemble sans doute à une forme de monstruosité, alors que c'est peut-être un rêve d'historien. Et encore. Sur le plan éditorial, c'est un non-sens puisque certaines opérations indexées seront intégrées dans la *release*, alors que les autres auront été positivement rejetées par décision d'auteur (développeurs, écrivains, etc.). Aurons-nous à ce point à l'avenir de généticiens du logiciel ou de n'importe quel autre type d'objet qu'il faille ainsi conserver chaque mouvement, comme l'on conserverait chaque mouvement de chaque partie d'échecs jouée ? Même si, il faut bien l'admettre, un logiciel de différentiel est en mesure de nous produire cette analyse philologique « en un clic ». Renoncer à documenter l'ensemble du flux et remettre l'accent sur la part de stabilité de ces objets numériques n'est pas une échappatoire. Nous avons déjà accepté face à différents types d'objets, y compris non numériques, d'être

9. Mais le pouvons-nous vraiment ? Rappelons que la Bibliothèque du Congrès a enregistré tous les tweets américains entre 2006 et 2017 avant de passer au « tri sélectif » [Noisette, 2017].

« condamné[s] à ne saisir que ce qui change[ait] moins vite que le reste », selon les termes d'un archiviste du web [La Porte, 2016].

CLASSIFICATION

Ainsi, (in)stabilité, collecte, conservation, accessibilité en lecture, modalités de rendu, sélection, ne permettent pas de caractériser en tant que tels les objets numériques dans leur diversité au point de les séparer des objets non numériques et de les penser ensemble dans une analyse info-documentaire. Quelle intelligibilité, alors, pour un ensemble d'objets aussi disparates que celui mentionné en introduction de cet article ? Il nous semble, dans un premier temps, nécessaire d'en écarter la classe des « données ». Qu'elles soient dites « de la recherche », issues « d'objets connectés » ou « de réseaux sociaux », « données » est un rôle logique attribué de manière contextuelle. Il n'est rien, objet numérique ou non, qui soit *a priori* une donnée ou qui n'en soit pas. Un logiciel peut être aussi une *donnée de la recherche*. Nous ne pouvons donc comparer « données » et « logiciels » sans confondre logiquement un ensemble et l'un de ses sous-ensembles et sans tomber dans un questionnement paradoxal du type « qu'est-ce qu'ont en commun des fichiers de bases de données et des données de la recherche ? ». Il en va de la même manière avec les « objets scientifiques », « objets patrimoniaux » ou « produits culturels » qui peuvent être à la fois « produits culturels » et « cartes et plans interactifs » ou « œuvres de l'art numérique ». Au-delà de ces deux points, les obstacles, recoupements ou intersections, continuent de surgir : Wikipédia est aussi un ensemble de pages web, une carte interactive peut être une œuvre d'art ; une application *smart city* est bel et bien une application, etc. Tenter de séparer les *logiciels* et leurs *produits*¹⁰, selon la vision schématique considérant qu'un logiciel est un ensemble de procédures destiné à fabriquer des produits à partir de données et de commandes, ne tient pas non plus. Les logiciels peuvent en effet également produire des logiciels. Il n'y a pas de généalogie en matière d'objets numériques qui, à l'instar de la taxonomie zoologique, permettrait de définir des classes et sous-classes bien distinctes. Face à la difficulté de penser une telle liste d'objets, le refus s'impose. Nous choisirons plutôt de les considérer séparément et d'évaluer pour chacun s'il n'est pas plus fertile, en matière de classification, de les penser non plus sur la base de leur « espace technique » commun mais sur d'autres bases : fonctionnelle, d'usage, etc. Plutôt que de nous demander comment penser un ensemble hétéroclite d'objets numériques, demandons-nous si une *partie de jeu numérique* doit d'abord être pen-

10. Dans le sens du produit d'une opération algébrique. Un objet numérique est un objet calculé.

sée comme une partie de jeu ou comme un objet numérique. L'objet «partie de jeu numérique» gagne-t-il à être pensé en regard des objets «logiciels» ou «site web», ou bien en regard d'une partie d'échecs ou de jeu de rôle non numériques ?

Le logiciel

Nous avons vu plus haut comment distinguer l'objet logiciel de ses sources, *objets documentaires logiciels*. Que le logiciel considéré soit un moteur de jeu, un système d'exploitation ou Git, la question des sources reste la même. Mais un logiciel ne consiste pas seulement en ses sources. Un logiciel se déploie (s'installe) et se lance (s'ouvre). La question se pose alors de la documentarisation d'un logiciel déployé et lancé. On pensera surtout à l'aspect graphique du logiciel. Or, tous les logiciels ne bénéficient pas, n'ont pas besoin, d'une interface graphique complexe. Un grand nombre d'entre eux (le plus grand nombre ?) opèrent par commandes émanant d'un autre logiciel, donc sans aucune interface graphique. Git, quant à lui, est livré sans interface graphique bien que l'on puisse lui en associer une *ad hoc*. On le manipule au moyen de simples (sur le plan graphique) lignes d'écriture dans un terminal (un lanceur). Documentariser un tel logiciel consiste à en rédiger toutes les commandes et options. D'autres logiciels intègrent au contraire des interfaces graphiques très riches. Elles sont aux logiciels ce que les tableaux de bord sont aux voitures. Les commandes et leurs options se trouvent derrière les boutons et champs graphiques. Il peut être nécessaire de documenter ces interfaces, témoin de leur époque, du degré d'évolution du logiciel, etc. Un tel rendu peut être extrêmement difficile à documenter dans son ensemble. À la différence du tableau de bord d'une voiture pré-numérique, l'interface logicielle se présente en effet comme un espace d'action replié, il ne peut être appréhendé en une image unique. Nous aurons donc une multitude de captures. Les objets documentaires issus de l'objet numérique en flux *logiciel* peuvent donc être de différents ordres : fichiers source, captures d'écran, cahier de fonctionnalités, documents d'architecture, documentation utilisateur (commandes). À quel type d'objet semble-t-il fertile de comparer un tel objet sur le plan info-documentaire ? Un logiciel est un calculateur qui produit des objets numériques ou guide des objets non numériques. Le terme de moteur est souvent employé pour désigner ces «outils» que sont les logiciels : moteur de blog, moteur de jeu, etc. La signature de WordPress est intéressante : «Fièrement propulsé par WordPress» ou, en anglais, *Proudly powered by WordPress*, la propulsion étant précisément le rôle des moteurs. S'agit-il d'une simple métaphore ou d'une supposée analogie fonctionnelle ? Faut-il penser l'objet logiciel parmi les objets «moteurs» et considérer sa documentation au même titre

que celle d'un moteur thermique (dessins techniques, brevets, plans de montage, manuel d'utilisation, mesures de performances, etc.) ou vaut-il mieux penser le logiciel en regard d'une émission de web TV ?

Objets spectaculaires

Des objets comme le *stream live*, la web TV ou une chaîne YouTube nous paraissent très éloignés du logiciel. Ils ont en commun de ne pas être opératifs mais purement spectaculaires, sans pour autant procéder d'un même rapport au temps et au public. Le *stream live* consiste en une captation et une diffusion en temps réel. C'est un objet multimédia (audio, vidéo) mais à la différence d'une chaîne de télévision il dispose d'un aspect interactif écrit (*chat*, commentaire), et ne se présente pas sous la forme d'émissions, il n'est pas séquencé. La web TV l'est et s'apparente directement aux chaînes classiques de télévision, si ce n'est qu'elles se déploient nativement dans le web. Devons-nous aborder ces objets comme l'on aborderait un objet numérique du type logiciel ou plutôt comme l'on aborderait d'autres objets spectaculaires traditionnels en flux ? En matière de collecte, l'INA archive déjà web TV et web radios et donne en cela une réponse. Peut-être y ajoutera-t-elle un jour le *stream live*. Ce dernier type d'objets est particulièrement lourd, quoique l'on puisse arguer que l'archivage de toutes les émissions d'une web TV peut l'être également. Pour l'instant, Google s'en charge si votre *live stream* YouTube dure moins de 12 heures¹¹ et vous conseille de toute façon de créer des *highlight clips*, des extraits, qu'il archivera volontiers¹². Quoi qu'il en soit, le mode de collecte est assez proche. Une visualisation 3D (ou une simulation) présente de grandes différences avec les objets spectaculaires précédents. Les sources d'une visualisation 3D consistent en images statiques et en fichiers de scripts (chaque séquence compte un nombre donné d'images par seconde de projection, à l'instar d'une pellicule de film). Par contre, le déroulement, le parcours et le montage des séquences sont tributaires des commandes que sont le déplacement ou l'orientation du spectateur/personnage : nous sommes face à un montage en temps réel, ce qui n'a pas d'équivalent dans l'espace non numérique si ce n'est dans le spectacle vivant interactif. Il en diffère cependant car il est fini. Une visualisation 3D est un ensemble de fichiers statiques délivré de manière interactive mais tous les points de vue possibles sont pré-produits. Le déroulé de la projection est potentiellement innombrable mais il y a un seul et même objet technique. Peut-on parler ici d'instabilité ? Nous l'appellerions volontiers *objet-parcours déterminé*. Nous sommes ici face à un

11. Aide YouTube. < <https://support.google.com/youtube/answer/6247592?hl=en> >.

12. Aide YouTube. < <https://support.google.com/youtube/answer/4522163> >.

objet spectaculaire mais très éloigné de la web TV ou du spectacle vivant. Il nous semble par contre très proche des jeux, qui exploitent d'ailleurs considérablement la visualisation 3D. Peut-être une visualisation 3D doit-elle être pensée avec les jeux finis ? Peut-être ces deux catégories, spectacles et jeux, sont-elles mitoyennes sur le plan info-documentaire ? Le spectacle, vivant ou non, est de longue date documenté par captures et captations. Y associer certains objets spectaculaires numériques, ou un objet-parcours tel que la visualisation 3D semble pertinent mais les éloigne sans doute d'autres objets numériques. Nous sommes loin des moteurs, bien que ces objets soient eux-mêmes toujours propulsés par des moteurs. De la même manière, nous n'étudions pas non plus ensemble une émission de télé et le réseau de câbles qui la propage sur le plan documentaire bien que l'on ne puisse les séparer selon une autre discipline telle que la médiologie.

Jeux et œuvres d'art numériques

Le jeu, numérique ou non, mêle deux objets distincts ; règles du jeu et parties. Nous laisserons ici de côté les règles du jeu, le logiciel, qui sera envisagé comme tout autre logiciel. Nous nous pencherons par contre sur l'objet « partie ». Le jeu est un univers complexe qui peut aller de la plus grande stabilité à un aléatoire volontaire. On ne peut donc parler de « jeu numérique » de manière générique, il faudrait l'envisager titre à titre. Sur le plan info-documentaire *Mario Bros*, mono ou bi-joueurs, linéaire et fini, ne peut pas être appréhendé comme *Eve Online*, aléatoire, ouvert, massivement multi-joueurs (MMOG ou MMO, de l'anglais *massively multiplayer online game*). Certains titres ont un impératif de stabilité éditoriale. Il n'est alors pas question d'évolution mais bien plutôt d'*opus*. Ces opus sont portés vers de nombreux environnements mais se doivent de rester le même au fil des années, comme un roman que l'on rééditerait dans d'autres collections, il est bon à ce sujet de savoir que lors d'une réédition d'un livre dans une autre collection, il n'est pas rare que l'auteur se permette quelques réécritures. Vous ne voulez pas que votre fille joue au même *opus* de *Mario Bros* que celui auquel vous jouiez à 20 ans mais que Mario soit tout à coup équipé d'un lance-roquettes plutôt que d'étoiles magiques ! Quoique stable sur le plan éditorial, et avec des règles et parcours finis, que ces parcours soient linéaires ou ramifiés, il n'existe pas deux parties identiques. La notion de partie se complexifie si l'on considère que certains jeux numériques intègrent une part d'aléatoire. Ils sont dits jeux « Rogue-like », du nom du premier jeu de ce type. Et il n'y a sans doute pas de moyen simple de documenter l'aléatoire. Plus encore, certains logiciels de jeu offrent aux joueurs d'apporter des données non prévues par les développeurs, des objets graphiques aussi bien que de nouvelles règles. Les joueurs peuvent

ainsi créer des déclinaisons à un jeu donné, des *mods*. Les combinatoires entre ces différentes architectures de parties sont possibles, complexifiant encore l'approche info-documentaire « du » jeu numérique. Certains parcours sont linéaires et relativement reproductibles, d'autres sont uniques, non reproductibles. Il en va de la même manière avec les œuvres d'art numériques interactives, très comparables à des parties, bien qu'elles n'aient pas la même fonction. Elles peuvent être linéaires, ramifiées, dépendre de données extérieures ou non, etc. Une partie de jeu ou une œuvre d'art numérique sont donc des objets très divers, y compris dans le cadre de règles très stables. Ce sont des *objets-parcours*, à l'instar de la visualisation 3D que nous avons décrite ci-dessus, objets-parcours déterminés ou indéterminés selon les cas, finis ou non finis. Comment les aborder sur le plan info-documentaire ? Cette question amène inexorablement les suivantes : comment traiter une partie de jeu d'échecs sur plateau ? Peut-on dire d'une partie de jeu, numérique ou non, parce qu'elle évolue de son ouverture à sa résolution, qu'elle est instable ? Le jeu d'échecs, parce qu'il permet une infinité de parties, est-il instable ? Est-ce que documenter le jeu d'échecs consiste à documenter toutes les parties jouées ? Que dire d'une partie qui n'accepte pas de résolution, de ces parties qui n'ont pas de fin ? Et des parties persistantes, durant lesquels votre personnage continue d'évoluer, et ses biens avec, alors que vous n'êtes pas connectés ? La partie de jeu, numérique ou non, est l'archétype de l'objet interactif et, partant, de la diversité. Le joueur est un acteur. Les règles limitent parfois les possibilités de cheminement mais nous nous approchons, sur le plan info-documentaire, du spectacle vivant. Faut-il documenter toutes les représentations ? Faut-il capturer la totalité de chacune ou quelques moments (états) de chacune ? Faut-il sélectionner certaines mises en scène parmi les nombreuses qui sont montées ? Jeu numérique, jeu non numérique, spectacle vivant, visualisation 3D, voilà de part et d'autre de la « fracture » technique numérique des objets qui posent des défis similaires à la documentation, défis très éloignés de ceux du logiciel ou du *stream web*.

Wikipédia, site web

Si le web est un objet d'intérêt en soi, il est aussi un espace de productions de différentes natures. Les ressources en ligne sont propulsées par différents moteurs, ou logiciels, avec pour résultat des caractéristiques très différentes sur le plan documentaire. Considérons un site-catalogue de recettes de cuisine propulsé par WordPress. Il existe dans ce CMS (*content management system*, système de gestion de contenu) une fonctionnalité d'export des sources du « site », lesquelles pourraient être traitées comme des *releases* logicielles et mises à disposition, voire collectées dans un projet similaire au *Software*

Heritage Project. De tels projets, concernant tel ou tel type éditorial, pourraient s'avérer très pertinents. Les historiens ont eu largement recours aux guides de voyage millésimés. L'objet moteur logiciel du type wiki, qui propulse notamment Wikipédia, est conçu pour conserver tous les états de chaque article. Sur le plan éditorial il s'agit bien d'une encyclopédie mais elle rompt avec ses alter ego papier en ne publiant pas de sortie millésimée, d'éditions, de *releases*. C'est un parti pris documentaire mais cela n'en fait pas un objet non documentable car ses états successifs sont tous durablement conservés. C'est à la fois un objet éditorial d'une grande dynamique et un objet documentaire d'une grande stabilité, à ranger peut-être au côté de ces titres littéraires dont nous possédons tous les brouillons. Les objets propulsés par un wiki ou par un CMS comme WordPress sont tout à fait distincts en termes documentaires. L'un publie et conserve automatiquement tous les brouillons (il existe quelques possibilités de réglages sur cet aspect automatique), l'autre demande un export volontaire et un archivage individuel (des extensions existent pour WordPress et sans doute pour tous les grands CMS aujourd'hui afin de gérer au mieux ces opérations et notamment l'archivage régulier et automatique des bases de données). La question se posera de choisir son mode de documentarisation selon le type éditorial. De grands quotidiens indiquent par exemple que des modifications ont été apportées à tel ou tel article sans être en mesure d'en donner le contenu¹³ alors que tout l'historique des modifications de n'importe quel article est bien consultable sur Wikipédia.

Nous sommes donc ici, avec la plupart des publications web, dans une approche documentaire qui s'analyse de manière assez proche de celle du logiciel : export et distribution régulière des sources. Mais comme pour le logiciel la question du rendu reste, et elle est complexe. Entre texte et paratexte, reproduction permanente, agrégation de ressources éparées et indépendantes, la question du rendu d'une page web est bien connue. Il faut penser séparément l'archivage du web et l'archivage de telle ou telle ressource dans le web, où le web serait cet agrégat toujours changeant et une ressource donnée du web un objet pouvant être documenté précisément par ses auteurs.

13. Entretien avec Bénédicte Dumont, responsable de la documentation du journal *Libération*, 18 mai 2018.

VERS UNE CLASSIFICATION INFO-DOCUMENTAIRE INTER ESPACES TECHNIQUES

Notre revue rapide de quelques objets numériques sur le plan info-documentaire, sur le plan notamment de leur rapport au temps, à leur environnement et à leur capture, vise à montrer que certains de ceux-là se rapprochent parfois plus d'objets documentaires traditionnels que de certains autres objets numériques. Plutôt donc que de penser objets numériques, d'une part, et objets traditionnels d'autre part, nous tentons volontiers une classification mixte telle que la suivante :

moteurs

- thermiques
- calculateurs numériques (logiciels)

...

écrits « littéraires »

- romans
- essais
- encyclopédies
 - éditions imprimées millésimées
 - *releases* millésimées
 - différentiels wiki

...

objets spectaculaires

- montés
 - TV hertzienne/numérique
 - chaîne YouTube
 - web TV
 - film analogique ou numérique
- montés en temps réel, en direct (objets parcours).
 - visualisation 3D/simulation (objets parcours finis non aléatoires)
 - jeux
 - finis/non finis
 - aléatoires/non aléatoires
 - œuvres d'art interactives, numériques ou non.
- direct
 - TV
 - stream live
 - spectacle vivant : théâtre, danse, etc.
 - interactifs
 - non interactifs

À la question des « enjeux que posent ces objets numériques [...] par rapport aux objets documentaires traditionnels » peut-être faut-il finalement répondre que le principal est précisément de ne pas les considérer d'abord en tant qu'objets numériques. Penser les objets numériques, qu'ils soient instables ou non, comme une classe info-documentaire en elle-même, nous apparaît impossible sinon contre-productif ou inintéressant. C'est la classe elle-même, et notre tentation de traiter les objets numériques comme un ensemble, qu'il faut questionner. Rien n'empêche *a priori* de penser ensemble objets numériques et traditionnels : ni leur instabilité, supposée ou réelle, ni leur conservation, leur collecte ou leur rendu. De part et d'autre de l'espace technique censé les caractériser, la simplicité comme la complexité de ces opérations les unissent plus qu'elles ne les séparent. L'info-documentation pourrait ainsi gagner à penser une classification mixte des objets et non basée d'abord sur quelque espace technique que ce soit, calculé, mécanique, organique, etc. Les intersections info-documentaires semblent trop nombreuses pour poser cette fracture, ou les disjonctions trop peu démontrées.

BIBLIOGRAPHIE

Briet, S. (1951). *Qu'est-ce que la documentation ?* Paris, France : Éditions documentaires, industrielles et techniques. < <http://martinetl.free.fr/suzannebriet/questcequeladocumentation/> >.

Derrot, S. (2016). « L'archivage des éphémères sur le Web : les paradoxes d'une mission Patrimoniale ». < <https://hal-bnf.archives-ouvertes.fr/hal-01270147> >.

Epron, B., et Vitali-Rosati M. (2018). « I. Pour une définition de l'édition à l'ère numérique » *Reperes*, May, 5-33. < <http://www.cairn.info/l-edition-a-l-ere-numerique--9782707199355-page-5.htm> >.

Escarpit, R. (1991). *L'information et la communication : théorie générale*. Version orig. Hachette Université Communication. Paris, France : Hachette.

Girard, C. (À paraître en 2020). « L'impensé de l'infrastructure Internet ». In Robert P. (éd.), *L'impensé numérique, Tome 2*. Paris, France : Éditions des archives contemporaines.

Jacobson, M., Larrousse N., et Massol M. (2014). « La question de l'archivage des données de la recherche en SHS (sciences humaines et sociales) ». July. < <https://halshs.archives-ouvertes.fr/halshs-01025106> >.

La Porte, X. de. (2016). « Archiver le Web, c'est être condamné aux vides ». *L'Obs*. < <https://www.nouvelobs.com/rue89/rue89-nos-vies-connectees/20161208.RUE5970/archiver-le-web-c-est-etre-condamne-aux-vides.html> >.

Noisette, T. (2017). « Archivage de Twitter : la Bibliothèque du Congrès passe au tri sélectif ». *L'Obs*. < <https://www.nouvelobs.com/societe/20171227.OBS9812/archivage-de-twitter-la-bibliotheque-du-congres-passe-au-tri-selectif.html> >.