



HAL
open science

Vers une architecture du savoir

Arthur Perret

► **To cite this version:**

Arthur Perret. Vers une architecture du savoir. CiDE.23 : Document et archivage : pratiques formelles et informelles, Oct 2023, Grenoble, France. hal-04294175

HAL Id: hal-04294175

<https://hal.science/hal-04294175v1>

Submitted on 19 Nov 2023

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.



Distributed under a Creative Commons Attribution 4.0 International License

Vers une architecture du savoir

Towards knowledge architecture

Arthur Perret

ELICO, Université Jean Moulin Lyon 3
arthur.perret@u-bordeaux-montaigne.fr

Résumé. Ce texte examine le concept d'archive chez Foucault sous l'angle du design documentaire. Foucault proposait une « archéologie du savoir » visant à décrire les conditions de possibilité d'existence des énoncés. Nous proposons une « architecture du savoir » dont l'objet est de concevoir ces conditions, en amont de l'ingénierie documentaire. Nous illustrons cette idée avec l'exemple du graphe documentaire personnel, une forme de documentation qui implique des choix en matière de support, d'écriture et d'organisation des connaissances.

Mots-clés. Archive, graphe, hypertexte.

Abstract. In this paper, we apply Foucault's concept of archive to document design. Foucault invented an “archaeology of knowledge” with the goal of describing the conditions of possibility for discourse. We imagine an “architecture of knowledge” whose aim is to design these conditions. To illustrate this idea, we use the example of personal documentary graphs, describing the architectural choices they involve, from markup to knowledge organization systems.

Keywords. Archive, graph, hypertext.

1 Introduction

Quelles sont aujourd'hui les grandes questions relatives au concept d'archive en lien avec celui de document ? Marie-Anne Chabin a montré récemment que les termes d'archive et d'archivage ont vu leur acception évoluer, depuis une « définition absolue » vers une « appréhension relative » :

« Fondamentalement, sont archives les documents que l'on regarde comme des archives, auxquels on attribue le statut d'archives. La plupart des acceptions du mot archive s'expliquent finalement par l'intention de la personne [...] c'est le passage d'une définition absolue des archives (des objets présentant telles caractéristiques intrinsèques) à une appréhension relative des archives » (Chabin, 2021, p. 15).

Ainsi le terme d'archive ne concerne plus seulement les « traces adjacentes de l'activité » (Chabin, 2021, p. 4) mais également des documents qui constituent leur propre finalité. Ces archives sont des documents par intention (Meyriat, 1981/2001) ; ils sont documentaires au sens de la documentarité, c'est-à-dire de ce qui « fait document » à nos yeux (Perret & Le Deuff, 2019). Chabin évoque par exemple l'usage du terme archivage pour désigner la création de documents à des fins de capture ou de sauvegarde d'information, notamment dans une démarche

personnelle, individuelle (Chabin, 2021, pp. 6, 13-15). Sur ce dernier point, elle écrit ceci :

« La finalité n'est plus ici [...] la constitution d'une documentation de savoir, la finalité est la prise de conscience de ces informations à l'instant T [...] une acquisition de connaissance immédiate et non une possibilité de connaissance dans un futur non défini. On assiste à une sorte de transfert de connaissance de l'objet archivé vers l'archiviste au moment même de la création de l'archive » (Chabin, 2021, pp. 14-15).

L'archive, c'est donc aussi ce qu'on produit lorsqu'on *écrit pour penser*. Cette forme d'archivage est proche de la notion d'information comme processus (*information-as-process*, Buckland, 1991). Il nous semble toutefois que cet archivage à la finalité immédiate peut coexister dans la pratique avec un archivage pensé pour le temps plus long, notamment dans les formes de documentation personnelle qui mélangent la prise de notes avec une logique encyclopédique. C'est ce type de pratique que nous proposons d'examiner dans ce texte. Plus spécifiquement, nous allons nous intéresser à la documentation personnelle hypertextuelle, suite à l'apparition ces dernières années d'une nouvelle génération d'outils de gestion des connaissances personnelles, qui appliquent la logique des wikis à la prise de notes.

Notre intention ici est d'ouvrir de nouvelles pistes de réflexion pour l'ingénierie documentaire appliquée au contexte de l'activité individuelle, en tenant compte des problématiques de préservation et de la question du « cycle de vie » des documents. Nous souhaitons montrer que l'archivage n'est pas juste la partie distale de ce cycle mais qu'il peut s'intégrer pleinement à la construction d'une documentation personnelle, notamment dans un mouvement cyclique entre « latence » (terme utilisé par Eco, 2010) et émergence. L'enjeu est d'imaginer des méthodes pour concrétiser ces concepts : comment concevoir un système de manière à faire ressurgir l'information archivée ?

Notre point d'entrée sera le concept d'archive, objet de tensions entre une approche philosophique ancrée dans la théorie et une approche archivistique beaucoup plus pragmatique ; c'est un « exemple assumé de non-congruence entre les catégories des uns et les catégories des autres » (Ogilvie, 2017, p. 133). Nous souhaitons réconcilier ces deux approches le temps d'une proposition à la fois théorique et pratique, autour d'une méthode que nous appellerons ici « architecture du savoir ». La première partie du texte sera consacrée à présenter le cheminement théorique vers cette idée. La seconde partie sera consacrée à son application dans le contexte de la documentation personnelle hypertextuelle.

2 Du concept d'archive à une architecture du savoir

Dans *L'archéologie du savoir*, Michel Foucault utilise le mot « archive » pour désigner le mode d'existence des choses dites : « le système général de la formation et de la transformation des énoncés » (1969/2008, p. 179). Ce concept vient appuyer une proposition : renouveler l'histoire des idées avec une nouvelle méthode, l'« archéologie », redéfinie comme l'étude du « déjà-dit » (Foucault, 1969/2008, pp. 180-184). Mais sa portée est en fait plus large. Revenons sur les termes précis dans lesquels Foucault le définit :

« L'archive, c'est d'abord la loi de ce qui peut être dit, le système qui régit l'apparition des énoncés comme événements singuliers [...] [qui] se groupent en figures distinctes, se composent les unes avec les autres selon des rapports multiples, se maintiennent ou s'estompent selon des régularités spécifiques [...] une pratique qui fait surgir une

multiplicité d'énoncés comme autant d'évènements réguliers, comme autant de choses offertes au traitement et à la manipulation » (Foucault, 1969/2008, pp. 177-178).

Cette définition établit un mode d'existence des énoncés en cinq points : suivant des conditions de possibilité ; comme entités discrètes ; ayant des relations les uns aux autres ; suivant des régularités ; pouvant faire l'objet de traitements. Il y a là un outil théorique générique, qui n'a peut-être pas été identifié comme tel par d'autres auteurs. Dans *Mal d'archive* (1995/2008) par exemple, Jacques Derrida s'intéresse lui aussi au processus d'archivage (qu'il nomme « archivation ») sous l'angle des conditions de possibilité des énoncés :

« L'archive, comme impression, écriture, prothèse ou technique hypomnésique en général, ce n'est pas seulement le lieu du stockage et de conservation d'un contenu archivable *passé* qui existerait de toute façon [...] Non, la structure technique de l'archive *archivante* détermine aussi la structure du contenu *archivable* dans son surgissement même » (Derrida, 1995/2008, p. 34).

Si la question posée est similaire, l'approche est différente. Cette réflexion sur l'archive intervient dans le contexte d'un travail sur la psychanalyse. Derrida parle notamment de l'« ardoise magique », cette tablette inscriptible et effaçable qui inspira à Freud des réflexions sur la mémoire, et qui paraît aujourd'hui « un jouet d'enfant » à côté des ordinateurs. Dans ce contexte, Derrida se demande ceci à propos de la psychanalyse :

« En quoi l'ensemble de ce champ a-t-il été déterminé par un état des techniques de communication et d'archivage ? On peut rêver ou spéculer sur les secousses géo-techno-logiques qui auraient rendu méconnaissable le paysage de l'archive psychanalytique depuis un siècle si [...] Freud, ses contemporains, collaborateurs et disciples immédiats, au lieu d'écrire des milliers de lettres à la main, avaient disposé [...] du courrier électronique (Derrida, 1995/2008, p. 33) ».

Comme Foucault, qui introduit un concept (l'archive) plus large que son sujet (l'histoire des idées), Derrida problématisé la question de l'archivage d'une manière qui dépasse le sujet de la psychanalyse. Son questionnement initial porte sur la part de refoulement dans l'archivage et l'historiographie. Mais la manière dont il le développe en fait un problème d'épistémologie : il s'agit du rapport entre les conditions de l'énonciation et la nature d'un champ de connaissance.

Cependant, là où Foucault nous livre un outil théorique influencé par le structuralisme, Derrida préfère emprunter à la phénoménologie : la façon dont il reconstitue les conditions de l'énonciation ne met pas à jour un « système » (le terme utilisé par Foucault) mais une « impression » (le sous-titre de *Mal d'archive* est « Une impression freudienne »). L'« archive » foucauldienne est un concept défini de façon précise ; c'est un système qui définit par principe ce qui est énonçable ou non-énonçable, et qu'on peut reconstituer par une archéologie minutieuse. L'« archive » derridienne est une notion indéterminée, tributaire de mouvements de l'esprit tels que la répression et le refoulement, et dont on peut seulement avoir une impression (Derrida, 1995/2008, pp. 50-51). Si les deux approches s'avèrent intellectuellement stimulantes, il nous est plus aisé de repartir de Foucault que de Derrida pour réfléchir à l'archive dans une perspective d'ingénierie documentaire.

Dans sa définition de l'archive, Foucault commence par évoquer la question des conditions de possibilité. Sur ce point, notre hypothèse est la suivante : dans l'environnement technique qui est aujourd'hui le nôtre, les conditions de possibilité des énoncés consistent essentiellement en une forme d'écriture. Nous appliquons ici le concept d'architexte. Initialement, l'architexte est un outil théorique relatif aux

genres textuels : Gérard Genette définit l'architextualité comme la « relation d'inclusion qui unit chaque texte aux divers types de discours dont il ressortit » (Genette, 1979, pp. 87-88). Mais Yves Jeanneret et Emmanuël Souchier (1999) lui empruntent le terme architexte et le redéfinissent dans le contexte des sciences de l'information et de la communication (SIC) pour désigner quelque chose qui est à la fois écrit et qui définit des conditions d'écriture. Jeanneret lui donne une portée très générale :

« [L'architexte est] l'appareillage technique d'une énonciation éditoriale [...] Il y a toujours eu en amont du texte un format du texte [...] L'architexte relève de cela mais potentialisé, amplifié dans le fait qu'il y a cette inscription écrite, durable et dynamique de ces objets [...] Une écriture de l'écriture, une écriture en amont de l'écriture, une écriture des conditions de l'écriture » (Jeanneret, 2005).

Concrètement, les architextes peuvent être des logiciels mais aussi des langages et des formats, qui conditionnent l'énonciation. Comme nous l'avons mentionné en introduction, archivage et création de documents se confondent désormais dans certains contextes ; les conditions dans lesquelles les documents sont créés, leur architextualité, constitue donc déjà en partie l'archive au sens foucauldien.

Dans la définition de Foucault, l'architextualité est indissociable de la « régularité ». Le terme procède d'une idée plus générale :

« L'archive, c'est aussi ce qui fait que toutes ces choses dites ne s'amassent pas indéfiniment dans une multitude amorphe, ne s'inscrivent pas non plus dans une linéarité sans rupture, et ne disparaissent pas au hasard d'accidents externes [...] Entre la tradition et l'oubli, elle fait apparaître les règles d'une pratique qui permet aux énoncés à la fois de subsister et de se modifier régulièrement » (Foucault, 1969/2008, pp. 177-178).

Derrière l'étymologie commencement-commandement de l'archive au sens foucauldien, il y a donc l'idée d'un ordre, d'un ensemble de règles, d'où découlent aussi bien la nature des énoncés comme entités discrètes que la possibilité de leur mise en relation et de leur traitement. Dit autrement, Foucault ne voit pas l'archive comme un simple processus de sélection de choses existantes, un sous-produit d'une activité comme le dit Chabin, mais comme un processus organisé et organisant de l'activité elle-même.

Quelles sont ces règles de l'archive foucauldienne ? On peut d'abord penser aux différentes dimensions du processus d'archivage tel qu'il s'entend classiquement : pérenniser, gérer les risques et s'appropriier l'information (Chabin, 2021, pp. 9-14). Elles sont structurées par des règles, souvent érigées en normes, par exemple concernant la durée légale de conservation des documents, l'emploi de certaines métadonnées, ou encore le recours à certains supports. Ces règles interviennent durant le travail archivistique, au moment de la sélection ou de la description par exemple.

Mais comme nous l'avons évoqué en introduction, les pratiques d'archivage se sont diversifiées : aux archives désignées comme telles a posteriori s'ajoutent désormais des archives pensées comme telles a priori, de manière intentionnelle. Dès lors, la question des règles de l'archive se déplace aussi : ces règles n'organisent plus seulement le traitement des énoncés mais leur construction même, ainsi que leur transformation. Cela signifie que l'archive implique aussi des règles d'ingénierie documentaire et des processus d'organisation des connaissances (tels que le classement, l'indexation, la mise en relation). Dans l'archive foucauldienne, c'est toute

la documentation (comme ensemble organisé de techniques) qui participe potentiellement aux conditions de possibilité des énoncés.

Reprenons donc le mode d'existence des énoncés défini par Foucault à travers le concept d'archive : suivant des conditions de possibilité ; comme entités discrètes ; ayant des relations les uns aux autres ; suivant des régularités ; pouvant faire l'objet de traitements. Nous avons qualifié cette définition d'outil théorique générique. En effet, Foucault en a tiré une « archéologie » parce qu'il avait pour objectif de renouveler l'histoire des idées. Mais ce n'est pas la seule utilisation possible de ce concept. Si on adopte une posture prescriptive plutôt que descriptive, on peut tirer de cette même définition une « architecture », c'est-à-dire une méthode de conception.

Nous imaginons cette « architecture du savoir » comme une branche appliquée de l'organisation des connaissances, influencée par le design, au croisement théorique entre document et archive. Elle peut être considérée comme voisine de l'architecture de l'information, laquelle regroupe un ensemble de questionnements sur l'organisation spatiale et logique des données dans les documents (Wurman, 1997). La distinction entre les deux porte sur le fait que l'architecture de l'information intervient au niveau de l'énonciation – il s'agit généralement d'une forme de design documentaire ou informationnel appliqué au Web –, alors que l'objectif d'une architecture du savoir est d'intervenir au niveau des conditions de cette énonciation.

Cette architecture du savoir peut également être située en amont de l'ingénierie documentaire. Celle-ci concerne la mise en place des dispositifs informationnels, et notamment des chaînes de production documentaires (Guyot, 2011). L'architecture du savoir peut orienter la phase de conception de ce travail ; son but n'est pas de structurer des contenus mais de déterminer la manière dont ils peuvent être structurés, ce qui inclut potentiellement la modélisation conceptuelle des données mais ne s'y réduit pas.

Pour illustrer cette proposition, nous présentons dans la seconde partie une étude de cas portant sur la documentation personnelle hypertextuelle, avec une mise en pratique de la définition de Foucault via le modèle du graphe.

3 Étude de cas : le graphe documentaire personnel

Notre objectif dans cette seconde partie est de donner un exemple concret de conception de l'archive au sens foucauldien. Pour cela, nous nous appuyons sur une recherche-conception menée dans le cadre du programme ANR HyperOtllet, qui a consisté à imaginer un système de documentation personnelle numérique inspiré par la tradition de la fiche érudite et le modèle du graphe, mis en œuvre via l'hypertexte.

Pour contextualiser cette expérimentation, nous examinerons d'abord l'état de l'art scientifique sur la question de la dimension personnelle dans le travail de la connaissance. Puis, nous verrons en quoi le modèle du graphe documentaire est une manière de concrétiser le concept d'archive.

3.1 Dimension personnelle du travail de la connaissance (état de l'art scientifique)

La dimension personnelle du travail de la connaissance fait l'objet de plusieurs champs de recherche relativement cloisonnés.

La gestion des connaissances personnelles (*personal knowledge management*, PKM) désigne est une extension du *knowledge management* à la perspective individuelle (Zhang, 2009, p. 241). Le PKM prend place dans un cadre organisationnel mais

contrairement au *knowledge management* classique, il concerne une relation entre individu et collectif qui est plutôt montante (*bottom-up*) que descendante (*top-down*), et correspond à un travail de la connaissance qui s'autonomise vis-à-vis de l'emploi (Le Deuff, 2012, p. 4).

La gestion de l'information personnelle (*personal information management*, PIM) désigne la gestion des informations rencontrées par un individu dans son espace personnel d'information (Jones, 2008 ; Jones et al., 2017). Le PIM s'émancipe plus encore que le PKM du cadre organisationnel et donne à voir une grande variabilité dans les pratiques (Jacques, 2016, p. 139). Comme le montrent par exemple Lindley et al. (2013), cette diversité des pratiques de PIM rend difficile l'émergence d'un modèle unifié.

La documentation personnelle (*personal documentation*) est la documentation créée et utilisée par un individu (Stibic, 1980). À la différence du PKM et du PIM, qui se focalisent sur la gestion, l'accent est mis sur le versant matériel et technique du dispositif, à travers le document et les processus d'organisation des connaissances (tels que le classement ou l'indexation). Notons que certains travaux utilisent l'expression « *personal documentation* » pour désigner la documentation qui porte sur un individu, dans un but administratif (documents d'identité) ou archivistique (documents biographiques). Ces travaux ne correspondent pas au périmètre de notre étude, de même que les recherches en informatique sur le *personal data management* (PDM), c'est-à-dire la gestion des données personnelles dans une logique de maîtrise du privé (*privacy*).

PIM et PKM tendent à reléguer le document au second plan, éludant du même coup la question du rapport entre document et archive. Certains travaux en PIM mobilisent la perspective archivistique pour réfléchir aux rapports entre documentation et préservation (Williams et al., 2009). Mais ce « PIM archivistique » est le fruit d'un point de vue extérieur : celui de l'ingénieur (qui conçoit le système de PIM) ou bien de l'archiviste (qui cherche à sélectionner et préserver *a posteriori*). Dans ce contexte, l'archivage est présenté comme la partie distale d'un « cycle de vie » du PIM qui est en réalité fort linéaire et qui ne clarifie pas le rapport entre document et archive.

Les recherches sur les archives personnelles (*personal archives*) ouvrent d'autres pistes. L'expression a d'abord été utilisée pour attirer l'attention sur l'individualité exprimée dans les archives au sens classique du terme, c'est-à-dire comme traces de l'activité (c'est le cas par exemple chez Hobbs, 2001). Plus récemment, cette approche personnelle a été associée à un nouveau courant de recherche (Douglas et al., 2022), qui s'inscrit dans le phénomène décrit par Chabin (2021), à savoir une appréhension relative de l'archive : la focale se déplace du document d'archives à la personne qui a produit ce document. Ces travaux expriment notamment un intérêt marqué pour les affects et les relations qui entourent la production ainsi que le traitement des archives. Ce mouvement est représentatif d'une tendance à inscrire des pratiques documentaires dans ce que nous appelons un « milieu de savoir » d'après Le Deuff (2018), c'est-à-dire l'ensemble des acteurs et actants impliqués dans un travail de la connaissance, y compris individuel.

Dans tous ces champs de recherche, un élément interroge : la question des conditions qui permettent aux connaissances d'émerger et de se transformer ne semble pas être considérée comme une problématique dont l'utilisateur peut se saisir de façon idiosyncratique. Le PKM s'intéresse plus aux transferts entre connaissance personnelle et organisationnelle qu'à la manière dont les individus organisent leurs connaissances pour eux-mêmes. Le PIM s'intéresse plus au besoin et à la recherche d'information qu'à la création et à l'organisation ; par ailleurs, les

travaux dans ce domaine restent souvent à un niveau conceptuel (ainsi que l'admettent Jones et al., 2017). Enfin, les recherches sur la dimension personnelle des archives correspondent surtout à un renouvellement sociologique de l'archivistique classique. Aucun de ces champs ne s'intéresse finalement à la possibilité pour un individu de déterminer lui-même ce que Foucault appelle l'archive, c'est-à-dire les conditions de possibilité des énoncés à l'intérieur de sa propre pratique. PIM et PKM en particulier semblent considérer que les formats, outils et méthodes relèvent d'un « donné », déterminé a priori dans un contexte organisationnel, et que la variabilité se situe essentiellement dans les pratiques, qui consisteraient à s'ajuster à l'existant.

C'est dans cet espace que l'idée d'une architecture du savoir trouve une application possible. Son objet est justement la conception des formats, outils et méthodes du travail de la connaissance : concevoir comment peuvent émerger des énoncés comme entités discrètes dans des conditions déterminées. Pour cela, il faut un ensemble de techniques, qui doivent non seulement permettre d'exprimer des connaissances mais aussi de façonner les conditions de cette expression. L'objectif est également de pouvoir mettre en relation les énoncés, ainsi que leur appliquer différents traitements – notamment le fait de mettre en latence des informations et de les faire ressurgir. Ceci peut passer aussi bien par des mécanismes classiques de recherche d'information (*information retrieval*) que par des mécanismes que nous qualifions d'émergence informationnelle (*information surfacing*), c'est-à-dire une remontée d'information autonome, passive, non sollicitée.

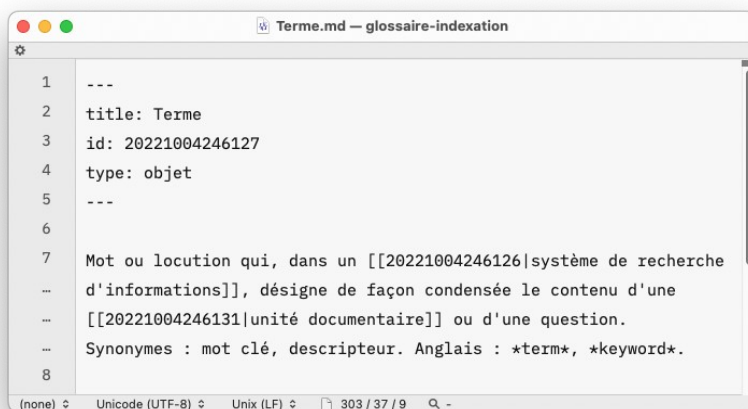
Nous proposons donc une application de cette idée à la documentation personnelle, que nous allons détailler maintenant.

3.2 Le graphe documentaire personnel comme archive

La documentation personnelle est la documentation faite par un individu pour lui-même. Elle constitue un objet hybride, entre documentation et archive. En effet, son existence peut se superposer à une carrière, voire à une vie entière, ce qui incite à y intégrer une logique de préservation. Par ailleurs, plus la durée d'utilité de la documentation personnelle s'allonge, plus elle est susceptible de se structurer suivant une logique encyclopédique, c'est-à-dire de prendre la forme d'un réseau de connaissances formalisées et organisées. Elle emprunte alors des caractéristiques au modèle de l'archive ou de la bibliothèque comme « conteneur d'un savoir toujours actualisable », non pas oublié mais latent (Eco, 2010, p. 131).

Comment construire une telle documentation ? Dans le cadre du programme ANR HyperOlet, nous avons mené un travail de recherche-conception pour répondre à cette question, autour de la réalisation d'objets que nous appelons des graphes documentaires personnels. Nous avons repris le concept de graphe documentaire proposé par Thibaut Arribé (2014) pour en donner une définition générale : un graphe documentaire est un mode d'organisation des connaissances qui se présente sous la forme de documents ou de fragments de documents interreliés (Perret, 2022, p. 268). Un graphe documentaire personnel peut alors être défini comme une forme de documentation personnelle mise en œuvre suivant le modèle du graphe. C'est à ce concept que nous avons appliqué une démarche d'« architecture du savoir ». D'abord, nous avons traduit concrètement le modèle du graphe documentaire : nous avons utilisé des fichiers au format texte, reliés de manière hypertextuelle, avec différents langages de balisage pour structurer les contenus. Et ensuite, nous avons défini un système d'organisation des connaissances (SOC) : nous avons établi les types d'entités sur lesquelles porterait cette documentation, ainsi que les types de relations possibles entre ces entités.

Une telle démarche implique des choix, qui peuvent être influencés par une culture commune de l'information mais qui procèdent aussi de préférences individuelles. Dans notre expérimentation, le choix des fichiers texte a été influencé par la tradition de la fiche : le format texte est robuste et compact, ce qui confère durabilité et mobilité aux fichiers ; et le fichier texte constitue une unité matérielle que l'on peut faire coïncider avec une unité intellectuelle – ce que Paul Otlet appelle le principe monographique (Otlet, 1934/2015, pp. 385-386). Pour structurer les contenus, nous avons choisi des langages de balisage dits « légers » (*lightweight markup languages*), qui rendent l'écriture hypertextuelle plus accessible : Markdown pour le contenu, YAML pour les métadonnées, la syntaxe MediaWiki pour les liens. Enfin, nous avons choisi des types d'entités (concepts, œuvres, personnes...) et des types de relations (parenté, proximité, opposition...) correspondant à la fois au sujet de cette documentation et à des préférences subjectives. La figure 1 donne un exemple de fichier illustrant tous ces choix.



```
1 ---
2 title: Terme
3 id: 20221004246127
4 type: objet
5 ---
6
7 Mot ou locution qui, dans un [[20221004246126|système de recherche
... d'informations]], désigne de façon condensée le contenu d'une
... [[20221004246131|unité documentaire]] ou d'une question.
... Synonymes : mot clé, descripteur. Anglais : *term*, *keyword*.
8
```

Figure 1. Un exemple de fichier texte produit dans le cadre de notre expérimentation, avec en-tête en YAML, corps en Markdown et liens wiki.

C'est l'ensemble de ces choix qui produisent l'archive. Le modèle du graphe, décliné sous la forme de fichiers textes, définit le mode d'existence des énoncés comme entités discrètes et interreliées. Le SOC, qu'il s'agisse des types d'énoncés ou des types de relations, permet de les faire émerger et évoluer suivant des régularités.

Qu'est-ce que cette archive nous permet de dire, et qu'est-ce qu'elle ne nous permet pas de dire ? L'écriture hypertextuelle permet de décrire des choses et les relations qu'elles entretiennent. C'est un modèle souple : ce qui n'est pas énonçable dans un graphe documentaire particulier peut le devenir si on modifie l'architecture de ce graphe – ses catégories d'entités, ses règles de mise en relation. Les applications peuvent sembler infinies : on pourrait dire avec Grandjean (2014) que « la connaissance est un réseau », ou avec Bertin (1977/2017) que « l'information est une relation », et donc que toute connaissance est potentiellement modélisable sous forme de graphe documentaire. Mais ce serait oublier les phénomènes continus, difficiles à discrétiser (l'image, le son) ; les objets infiniment complexes, difficiles à

restituer (le cerveau humain, l'univers). Le graphe est donc une instance d'archive particulièrement expressive ou permissive, dans la limite toutefois de notre capacité à analyser et synthétiser les choses connaissables.

Cette archive évolue de manière itérative : Bernard Stiegler parle d'un « idiotexte », c'est-à-dire une mémoire prothétique régulièrement « remuée » par l'écriture (Stiegler, 2010, p. 52). On peut dès lors véritablement parler de « cycle de vie » des documents : création, mise en relation, latence, ré-émergence, transformation, et ainsi de suite jusqu'au « désherbage » qui est autant nécessaire dans ce type d'environnement que dans une bibliothèque au sens classique.

Cette pratique itérative nécessite toutefois des interfaces appropriées. Un inconvénient de l'écriture hypertextuelle réside dans la complexité des structures produites. Dans ses premières expérimentations sur la mécanisation de la documentation personnelle hypertextuelle, Engelbart (1962) évoque la difficulté qu'il éprouve à maîtriser mentalement l'écheveau de relations qu'il établit entre différents documents. La visualisation de données peut alors intervenir comme une technique permettant de gérer la complexité. Concrètement, il s'agit de fournir une interface visuelle permettant d'appréhender la structure hypertextuelle mais également d'y naviguer, grâce à différentes techniques qui créent des niveaux de contexte au sein de la visualisation (Hofmann & Langendörfer, 1990 ; Mayr et al., 2022).

Dans le cadre de l'expérimentation, nous avons contribué à élaborer un logiciel de visualisation appelé Cosma (voir figure 2). Ce logiciel interprète un graphe documentaire, c'est-à-dire un ensemble de documents interreliés (par exemple des fiches de lecture, des notices terminologiques, des notes relatives à des personnes, œuvres ou événements, etc.). À partir de cela, il fournit une visualisation interactive, sous la forme d'un fichier HTML unique et autonome, ce qui facilite le partage des informations. Cosma exploite notamment les catégories et les liens présents dans les documents, qu'il traduit visuellement. Les liens entrants, ou rétroliens, sont mis en évidence avec leur contexte d'énonciation, permettant de savoir où une fiche a été mentionnée et comment. Toutes ces fonctionnalités ont pour but de faire de la visualisation un aide-mémoire augmenté, qui favorise l'émergence informationnelle et la sérendipité (au sens de Catellin & Loty, 2013).

Si ce logiciel peut jouer un rôle utile dans le processus documentaire, ses fonctionnalités ne remplacent cependant pas le travail intellectuel. Cela reste un outil d'aide à la décision : l'évolution itérative du graphe dépend de notre attitude réflexive.

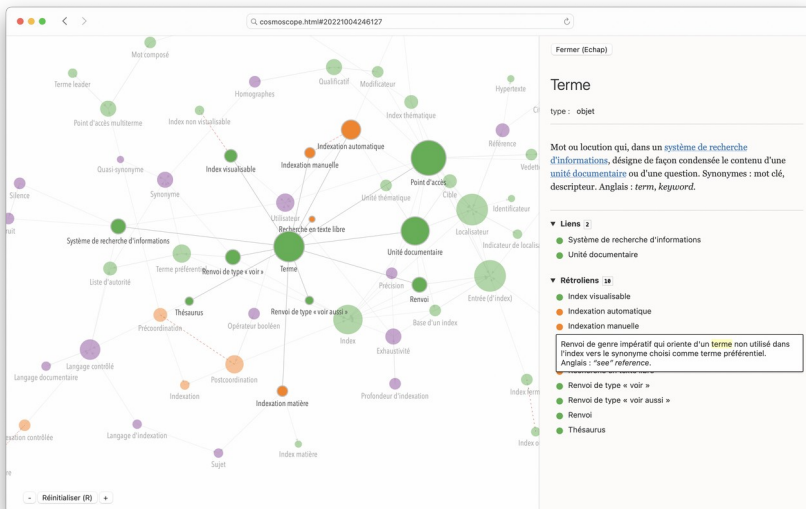


Figure 2. Une visualisation produite avec le logiciel Cosma. Sur la droite, le contexte d'énonciation d'un rétrolien apparaît au survol.

4 Conclusion

Dans ce texte, nous avons examiné l'hybridité entre document et archive, à l'intersection entre architexte, hypertexte et idiotexte. À la suite de Foucault, nous nous sommes intéressés au concept d'archive comme mode d'existence des énoncés. Nous avons alors suggéré une approche pratique, l'architecture du savoir, dont le but est de définir ce mode d'existence. Influencés par le concept d'architexte, nous avons proposé un exemple orienté vers l'écriture, en montrant quelques possibilités offertes par l'hypertexte et la visualisation.

Derrida se demande dans *Mal d'archive* comment les « machines à archiver » informatiques pourraient tenter de représenter l'appareil psychique. L'hypertexte est une réponse possible : c'est une modalité d'organisation des connaissances (Ridi, 2018), imaginée en référence au fonctionnement associatif de l'esprit humain (Bush, 1945 ; Engelbart, 1962 ; Nelson, 1973). Et parce que l'hypertexte suit la logique du graphe, c'est-à-dire une structure faite d'entités interreliées, il s'agit aussi d'un moyen de concrétiser l'archive au sens de Foucault – ce mode d'existence des énoncés comme entités discrètes, ayant des relations les unes aux autres.

Il existe différentes manières de mettre en œuvre un système hypertexte. Dans le contexte du PIM, Jones (2008) propose par exemple d'utiliser RDF, XML et des bases de données. Mais ces approches sont réputées complexes, coûteuses et difficiles d'accès (Poupeau, 2018). Or plusieurs innovations récentes changent la donne et facilitent la mise en œuvre d'un dispositif de documentation personnelle hypertextuelle : ce sont d'une part les langages de balisage dits « légers », comme Markdown et AsciiDoc ; et d'autre part une nouvelle génération de logiciels d'écriture non-linéaire, comme Notion (2016), Zettlr (2017) et Obsidian (2020), parfois qualifiés d'« outils pour penser » (*tools for thought* selon l'expression de Matuschak & Nielsen, 2019). Ces innovations abaissent la barrière à l'entrée de l'écriture hypertextuelle pour les concepteurs comme pour les utilisateurs. Par

ailleurs, elles répondent à une évolution des attentes en matière de maîtrise des données, sur des questions de propriété et d'interopérabilité.

Cette nouvelle accessibilité de l'écriture hypertextuelle ouvre des possibilités en matière de prise de notes et d'accumulation encyclopédique mais aussi d'idéation. En effet, la mise en relation constitue un processus d'organisation des connaissances (POC) qui remplit différentes fonctions : classer, indexer mais aussi et surtout composer, c'est-à-dire créer quelque chose de nouveau par assemblage de différents éléments (Perret, 2022, voir tableau p. 257).

L'« archive » au sens de Foucault constitue le niveau méta de cette pratique. Mais le sens classique du « document d'archives » ne disparaît pas pour autant : dans une documentation personnelle, les choix qui jalonnent le processus d'architecture sont aussi influencés par des problématiques de préservation.

Foucault s'est intéressé aux formations discursives à une échelle très large, celle des disciplines et des sociétés. Nous avons travaillé dans un périmètre beaucoup plus restreint, avec une étude de cas qui concerne les pratiques documentaires individuelles. Mais on peut remonter à l'échelle organisationnelle, et c'est d'ailleurs là que se situent des défis intéressants, qui restent à aborder dans le cadre d'autres travaux : comment opérationnaliser l'« architecture du savoir » dans un dispositif où les points de vue sont démultipliés ?

5 Références

- Arribe, T. (2014). *Conception des chaînes éditoriales : documentariser l'activité et structurer le graphe documentaire pour améliorer la maîtrise de la rééditorialisation*. Thèse de doctorat. Université de technologie de Compiègne, Compiègne. <https://tel.archives-ouvertes.fr/tel-01127476>
- Bertin, J. (2017). *La graphique et le traitement graphique de l'information*. Bruxelles : Zones sensibles. (Ouvrage original publié en 1977).
- Buckland, M. (1991). Information as thing. *Journal of the American Society for Information Science*, 42(5), 351-360. DOI : 10.1002/(SICI)1097-4571(199106)42:5<351::AID-ASI5>3.0.CO;2-3
- Bush, V. (1945). As We May Think. *The Atlantic*, 101-108. <https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>
- Catellin, S., & Loty, L. (2013). Sérendipité et indisciplinarité. *Hermès*, (67), 32-40. DOI : 10.4267/2042/51882
- Chabin, M.-A. (2021). Archive(s) et archivage(s). *Signata*, (12). DOI : 10.4000/signata.2992
- Derrida, J. (2008). *Mal d'archive: une impression freudienne*. Paris : Galilée. (Ouvrage original publié en 1995).
- Douglas, J., Ballin, M., & Lapp, J. (2022). Toward Person-Centred Archival Theory and Praxis. Introduction. *Archivaria*, (94), 5-21. DOI : 10.7202/1094874ar
- Eco, U. (2010). *De l'arbre au labyrinthe. Études historiques sur le signe et l'interprétation*. Paris : Grasset.

- Engelbart, D. C. (1962). *Augmenting Human Intellect: A Conceptual Framework* (Rapport No. AFOSR-3223). Stanford : Stanford Research Institute. <https://www.dougenelbart.org/content/view/138/>
- Foucault, M. (2008). *L'archéologie du savoir*. Paris : Gallimard. (Ouvrage original publié en 1969).
- Genette, G. (1979). *Introduction à l'architexte*. Paris : Seuil.
- Grandjean, M. (2014). La connaissance est un réseau. Perspective sur l'organisation archivistique et encyclopédique. *Les Cahiers du numérique*, 10(3), 37-54. DOI : 10.3166/LCN.10.3.37-54
- Guyot, B. (2011). introduction à l'ingénierie documentaire et aux sciences de l'information. Institut national des sciences et techniques de la documentation. https://archivesic.ccsd.cnrs.fr/sic_00665267
- Hobbs, C. (2001). The Character of Personal Archives: Reflections on the Value of Records of Individuals. *Archivaria*, (52), 126-135. <https://www.archivaria.ca/index.php/archivaria/article/view/12817>
- Hofmann, M., & Langendörfer, H. (1990). Browsing as Incremental Access of Information in the Hypertext system CONCORDE. Dans *H2PTM'89 : Communication interactive*. Paris. [https://wicri-demo.istex.fr/Wicri/Sic/H2PTM/fr/index.php/H2PTM_\(1989\)_Hofmann](https://wicri-demo.istex.fr/Wicri/Sic/H2PTM/fr/index.php/H2PTM_(1989)_Hofmann)
- Jacques, J. (2016). *Définition des compétences propres à l'organisation des collections d'informations personnelles numériques* (Thèse de doctorat). Université catholique de Louvain, Louvain-la-neuve.
- Jeanneret, Y. (2005). Sémiotique de l'écriture. *Canal-U*. https://www.canal-u.tv/video/fmsh/semiotique_de_l_écriture.28783
- Jeanneret, Y., & Souchier, E. (1999). Pour une poétique de l'écrit d'écran. *Xoana*, (6/7).
- Jones, W. (2008). *Keeping found things found: the study and practice of personal information management*. Boston : Morgan Kaufmann.
- Jones, W., Dinneen, J. D., Capra, R., Diekema, A. R., & Pérez-Quiriones, M. A. (2017). Personal Information Management. Dans J. D. McDonald & M. Levine-Clark (Éds), (pp. 3584-3605). Boca Raton : CRC Press, Taylor & Francis. DOI : 10.1081/E-ELIS4-120053695
- Le Deuff, O. (2012). Le réseau personnel de gestion des connaissances et la redéfinition du travail. *Terminal*, (110), 39-54. DOI : 10.4000/terminal.1227
- Le Deuff, O. (2018). *Les humanités digitales : historique et développements*. Londres : ISTE Editions.
- Lindley, S. E., Marshall, C. C., Banks, R., Sellen, A., & Regan, T. (2013). Rethinking the web as a personal archive. Dans *Proceedings of the 22nd international conference on World Wide Web* (pp. 749-760). Rio de Janeiro : ACM. DOI : 10.1145/2488388.2488454

- Matuschak, A., & Nielsen, M. (2019). How can we develop transformative tools for thought? <https://numinous productions/ttft/>
- Mayr, E., Liem, J., & Windhager, F. (2022). Reasoning with Knowledge Graph Visualizations. Dans *6th International Conference Graphs and Networks in the Humanities 2022. Technologies, Models, Analyses, and Visualizations*. https://graphentechnologies.hypotheses.org/files/2022/01/Reasoning_with_Knowledge_Graph_Visualizations_A_Mental_Models_etc-Mayr_Liem_Windhager.pdf
- Meyriat, J. (2001). Document, documentation, documentologie. Dans V. Couzinet (Éd.), *Jean Meyriat, théoricien et praticien de l'information-documentation* (pp. 143-159). Paris : ADBS Éditions. (Ouvrage original publié en 1981).
- Nelson, T. H. (1973). As We Will Think. Dans *Selected Papers, 1977*.
- Ogilvie, D. (2017). Paradoxes de « l'archive ». *Sociétés & Représentations*, 43(1), 121-134. DOI : 10.3917/sr.043.0121
- Otlet, P. (2015). *Traité de documentation. Le livre sur le livre*. Bruxelles : Les Impressions nouvelles. (Ouvrage original publié en 1934).
- Perret, A. (2022). *De l'héritage épistémologique de Paul Otlet à une théorie relationnelle de l'organisation des connaissances*. Thèse de doctorat. Université Bordeaux Montaigne, Bordeaux. <https://theses.hal.science/tel-03954967>
- Perret, A., & Le Deuff, O. (2019). Documentarité et données, instrumentation d'un concept. Dans *12ème Colloque international d'ISKO-France : Données et mégadonnées ouvertes en SHS : de nouveaux enjeux pour l'état et l'organisation des connaissances?* <https://hal.archives-ouvertes.fr/hal-02307039>
- Poupeau, G. (2018, 6 octobre). Les technologies du Web sémantique, entre théorie et pratique. *Les petites cases*. <http://www.lespetitescases.net/les-technologies-du-web-semantique-entre-theorie-et-pratique>
- Ridi, R. (2018). Hypertext. *Knowledge Organization*, 45(5), 393-424. DOI : 10.5771/0943-7444-2018-5-393
- Stibic, V. (1980). *Personal documentation for professionals: means and methods*. New York : Elsevier.
- Stiegler, B. (2010). Le concept d'idiotexte : esquisses. *Intellectica*, 53(1), 51-65. DOI : 10.3406/intel.2010.1178
- Williams, P., Leighton John, J., & Rowland, I. (2009). The personal curation of digital objects: A lifecycle approach. *Aslib Proceedings*, 61(4), 340-363. DOI : 10.1108/00012530910973767
- Wurman, R. S. (1997). *Information architects*. New York : Graphis.
- Zhang, Z. (2009). Personalising organisational knowledge and organisationalising personal knowledge. *Online Information Review*, 33(2), 237-256. DOI : 10.1108/14684520910951195