Médiations documentaires / Utopies documentaires

# Plan (du PPT 1)

1. Médiation documentaire
2. Traitement du document
3. Utopies documentaires
4. Grandes questions : google book, google print, etc.

# Médiation/médiation documentaire

## Introduction

Le mot « médiation » dans différents domaines comme :

* Au sens large du terme,
* Au sens social,
* Au sens documentaire,
* Au sens numérique,
* Etc.

On peut donc se demander, « La médiation : concept-clé ou mot-valise ? ». Une référence à Bertrand Calenge, auteur et conservateur de bibliothèque mais également directeur de l’institut de formation des bibliothécaires de Villeurbanne, dans son livre « La médiation des connaissances » qui est un ouvrage réflexif (suite à toutes ses années d’expérience) sur cette notion et l’évolution de la fonction de la bibliothèque en elle-même.

## Définition : qu’est-ce que la « médiation » ?

De son origine étymologique latine « mediare », « servir de médiateur ou « mediatio » renvoie à « médiation ou entremise ». Il s’agit en fait d’expliciter la nature de cette mise en lien afin de la qualifier au vu des domaines dans lesquels elle se déploie.

* **Définition générique :** le dictionnaire Trésor de la langue française définit la médiation comme étant « le fait de servir d'intermédiaire entre deux ou plusieurs choses ou personnes ».

Plusieurs acceptions sont citées dans le dictionnaire :

* **Définition philosophique** : elle traduit un espace, une trajectoire, il est mis en avant ici l’idée d’un processus de transformation qui s’opère et qui connecte deux points : « action de servir d'intermédiaire entre un terme ou un être duquel on part, et un terme ou un être auquel on aboutit ». C’est donc dans cette trajectoire que se joue le processus de médiation, de mis en lien.
* **Définition en psychologie** : on évoque plutôt un « processus par lequel une connaissance sensorielle se transforme en une donnée intellectuelle », il s’agit typiquement ici d’une médiation des savoirs. Ce qu’on entend, on l’intègre dans notre mémoire pour devenir une donnée intellectuelle.
* **Définition en droit** : il est question d’une « entremise destinée à concilier ou à faire parvenir à un accord, à un accommodement des personnes ou des parties ayant des différends ». Dans le cadre de médiation sociale, on va renouer le dialogue et revenir à un espace d’échange, de négociation ou de discussion.

A travers ces facettes disciplinaires, un élément constant est présent, celle de lien ou relation entre deux entités, ces éléments sont aussi présents en ethnologie, mais aussi en histoire voire en religion pour qui, il s’agit de « l’entremise d'un génie protecteur ou d'un prêtre auprès des humains pour les rapprocher des divinités ».

**SAUF** quelques exceptions que l’on vient de voir en psychologie ou en droit mais on peut aller beaucoup plus loin (sciences religieuses ou ethnologie) **MAIS** **on remarque qu’il y a des constantes à partir du moment on prend la médiation comme processus.**

## Définition en SIC (sciences de l’information et de la communication)

*« Le concept de médiation documentaire au sein des sciences de l’information et de la communication (SIC), renvoie à la notion d'intermédiaire, de lien entre le singulier et le collectif. La médiation documentaire concerne une médiation des savoirs mettant en place, grâce à un tiers, des interfaces qui accompagnent l’usager et facilitent les usages. Elle permet de concilier deux choses jusque-là non rassemblées pour établir une communication et un accès à l’information. C’est par sa capacité à lier information et communication qu’elle peut être qualifiée de médiation documentaire. Elle s’appuie sur des composants humains ou matériels qu’on peut distinguer en « médiateurs sociaux « naturels » (normes, valeurs…), médiateurs humains (négociateurs, chefs…), dispositifs complexes (agencements matériels et géographiques, organisationnels et techniques…) »1. La médiation est donc fortement liée à la question du dispositif étudié en SIC comme objet matériel médiateur qui désigne « l’ensemble des substrats matériels de la communication* ». **Par I. Fabre, C. Gardiès.**

Tout d’abord, on peut imaginer que le système d’information, étant dans une organisation, peut être visible depuis l’extérieur. C’est-à-dire, diffuser l’information à travers un système de diffusion à échelle mondiale (ce qu’est le WEB), qui permet ainsi son accessibilité. **DONC** la dimension ou le dispositif complexe doit être compris au sens large avec son agencement (composition de matériel ou de lieu, d’une ou plusieurs organisations s’appuyant sur des dispositifs sociotechniques qui vont les aider à accéder à la fois à l’information et la communication à distance). **Exemple** : dispositif sociotechnique comme les médias sociaux, avec une notion de plateforme qui va modifier les rapports entre les individus.

**Conclusion** : en regardant cette définition on a,

* La notion d’espace évoquée par la philosophie,
* De transmission et transformation des connaissances de la psychologie,
* Mais aussi une mise en lien nécessitant la mise en place d’un espace de discussion et de négociation entre celui qui offre ses services d’information et de documentions à distance et l’usagers qui est la cible de ces unités documentaires.

**CE QUI SIGINFIE QUE toutes les spécificités et les disciplines travaillant la notion de « médiation » ont des points communs et des rapports entre**

* Lieux de médiations,
* Objets de médiations (ce pourquoi il y a des liens et pourquoi la médiation se créé)
* Et des acteurs de médiations (qui vont jouer des rôle d’accompagnement/médiateurs).

*Finalement, la communication pourrait avoir une part de plus en plus conséquente dans l’espace documentaire.*

# Traitement documentaire

## Definitions

* **L'analyse documentaire :**

*‘’…opération ou l'ensemble des opérations qui visent à formuler le contenu d'un document sous une forme différente de sa forme originelle afin de faciliter le repérage et la consultation dans un stade ultérieur’’*. (Source J. Chaumier, « Les techniques documentaires »).

* Cette définition se base sur celle de la « médiation », on parle de taches permettant de constituer un fond documentaire.
* Ici, les opérations sont les ensembles d’activités professionnelles qui vont reposer sur la chaine documentaire (feuille de route du professionnel de l’information et qui va consister à décrire l’ensemble de ces activités, de la politique d’acquisition jusqu’à la mise à disposition de l’usager).
* **Dans la partie analyse de la chaine documentaire, nous trouvons 3 sous-tâches :**
  + **La classification** : système de référence permettant d’organiser le savoir (cf. DEWEY et la classification à son nom par exemple),
  + **La condensation** : synthèse des idées (par exemple : permettant de faire la différence entre des ouvrages traitant la même question)
  + **Et l'indexation** : système de référence également de représentation des connaissances (par exemple : le thésaurus, système en se développant différemment du modèle encyclopédique de DEWEY, de manière générique ou thématique).
* **L'indexation documentaire :**

On va se poser une question : « si on indexe un document depuis l’émergence du Web s’agit-il encore d’une indexation documentaire ou de l’information ? »

**Mode de description des documents ou de l'information :**

* *« …opération de traitement des documents ou de l'information par laquelle on choisit manuellement ou de façon automatique les termes les plus appropriés pour représenter objectivement le contenu d'un texte … »* (source site URFIST Toulouse).
* Ces termes sont exprimés dans la plupart des cas dans un langage documentaire.

La nuance importée va amorcée une réflexion sur la question du traitement automatique, qui est l’outil de recherche d’information (exemple : l’interface de la recherche s’appuyant sur l’utilisation et les choix de termes en amont pour préparer l’accès à des ressources).

**La recherche d’information = traitement documentaire réalisé + méthode d’indexation.**

Indexation manuelle VS indexation automatique :

* **Manuelle** : intervention d’un professionnel de l’information qui va mettre en correspondance et procéder à l’analyse du document, mais également à son traitement et à l’intégration des informations auquel il renvoie. **Tout ça pour permettre une recherche qui elle est AUTOMATIQUE.**
* **Automatique** : à la fin des années 90 et début années 2000, les informations en ligne sont de plus en plus importantes et nombreuses. « Quel système va pouvoir traiter cette quantité d’informations que l’humain ne peut gérer ? »

## Introduction: indexation automatique

* Les nouvelles technologies de l'information et de la communication ont permis :
  + La création d'outils d'analyse automatique de l'information,
  + Le repérage de l’information de façon rapide,
* Comment exploiter de façon pertinente des volumes toujours croissants d'information en des temps et à des coûts réduits ?

## Texte et information

* Dans un système d'information :
  + Texte : véhicule le plus immédiat de l'information, un contenant de connaissances.
* Manipuler (dans une approche manuelle) un texte exige que :
  + Transformation de l’information en données qui font sens pour la machine, car la machine ne peut appréhender un ensemble de lettres formant un bloc de texte.

Donc lorsqu’il y a transformation en données factuelles = information indexée.

## Le principe de l’indexation automatique

* Comme l'indexation manuelle le principe de **l'indexation automatique** est de :
  + Identifier un document par un ensemble de mots-clés représentatif de son contenu,
  + Génération automatisée des index (fichier de type texte où il y aura tous les mots qui sembleront avoir du sens, par ordre alphabétique, qu’on va mettre de côté afin qu’ils puissent être connectés avec un système permettant de les reconnaître en fonction de la requête de l’usager et des informations présentes)
  + Mise en correspondance de manière automatique.

On a donc une forme de médiation documentaire car ce processus d’indexation permet de faire le lien possible entre le contenu d’un document et les recherches potentielles des usagers.

Pour rappel, lors d’une saisie automatique dans un moteur de recherche, on nous suggère en premier lieu des recherches déjà faites ce qui facilite en partie le travail de l’indexation.

* **Indexation assistée** :
  + Les processus intellectuels (analyse, synthèse, etc.) mis en œuvre ne sont ni parfaitement ni totalement automatisables,
  + L’indexation assistée c’est faire effectuer par la machine des propositions d'indexation qu'un humain validera ou corrigera.
  + Evidemment, nous sommes dans un cas de communication d’humains à machines. Signifiant que ce système est confronté à pas mal de limites (dans les bases de données ou dans les informations) car il y déjà pas le même langage avec des adaptations des deux côtés.
* **Système indexation automatique** :
  + Ensemble des traitements automatisés permettant identification et représentation sous une forme normalisée homogène et réduite les informations pertinentes contenues dans les textes
* **Historiquement** :
  + *Années 40*: premiers ordinateurs.
  + *1950-1960* : édition des index de mots-clés dans des titres (pas encore dans des textes) par des ingénieurs de grandes entreprises d’informatiques (ex : IBM). Ayant à l’époque seulement des machines à écrire, ces débuts de l’informatique vont donc lancer les débuts de la micro-informatique et de l’information de « masse »
  + *1960-1970* : élaboration des index des articles (dans les années 70, mise en place de bases de données documentaires reposant sur un système de bases de données classiques).
  + *1970-1990* : gestion informatisée, bdd, texte intégral (opposition aux références bibliographiques), etc.
  + *1990 – 2000* : « ère des réseaux », avec la création du langage HTML permettant la connexion de plusieurs pages dites ‘web’ ou ‘électroniques’. On va donc travailler la notion de « connexion », celles à l’ordinateur mais également celles des pages WEB entre elles. Avec la création d’une entité nommé « site web » et du changement total au niveau des techniques et des modalités de recherche d’informations, l’effort intellectuel qui se faisait sur la recherche d’informations à l’époque, va aujourd’hui se concentrer sur l’analyse de l’information trouvé parmi une masse conséquente d’informations.
* **Distinction en informatique documentaire:**
  + BDD/base de données bibliographique : champ documentaire important (titre, auteur, résumé, mots-clés, côte, etc.)
  + VS. Textuelle (titre et résumé)
* **Les différentes techniques et méthodes :**
  + Approches mathématiques (méthodes statistiques et probabilistes), en calculant le nombre de fois que le terme apparait.
  + Approches linguistiques (TALN, traitement automatique du langage naturel), va associer le correcteur orthographique/syntaxique pour développer des dictionnaires assez imposants, venant appuyer les approches mathématiques
  + Approches basées sur les connaissances (systèmes experts, bases de connaissances)‏, anthologie, on va dissocier un domaine de connaissance qu’on va stocker MAIS tout en lui créant des liens au sein de ce domaine. Système basé sur de l’inférence, on fait donc des déductions car les deux autres systèmes ont plus de limites que celui-ci.

1. Utopies documentaires (définition, enjeux et ambitions)

## Introduction

**Bibliothèques numériques** : perçues comme entités globalisantes d’un accès unique à l’ensemble des savoirs, constitutives de l’histoire de la documentation dans toutes ses modalités (bibliothéconomie, classifications, catalogage, recherche documentaire …).

**Utopies documentaires :**

* H. G. Wells avec World Brain : (à 72ans) dans cet essai/discours, nous retrouvons “The idea of a permanent Encyclopedia”, qui renvoie à une vision similaire du Wikipédia d’aujourd’hui, dès 1938. A l’issu de sa vie, il est arrivé à imaginer un monde où un cerveau mondial prime. Cette métaphore veut exprimer les connaissances que chaque individu a mais en se référant aux connaissances de l’ensemble des individus de cette terre.
* Bush avec Memex (« memory extended ») : étant le conseiller scientifique et avec toutes les fiches d’informations qu’il a dû préparer pour le président Roosevelt pendant la WW2. Il va imaginer un système où ses fiches vont avoir des liens entre elles à partir de notions, de mots, ou de concepts dans un fichier. On peut comparer cette pratique à celle utilisée lors de l’émergence du web. C’est-à-dire, la création d’une matrix de connaissances extrêmement riche à partir d’une aide à la consultation des fiches par des associations.
* Nelson avec Xanadu : dans les années 70, il va penser à une bibliothèque répondant à la question, « comment l’informatique peut venir en aide à la gestion automatique des données et de l’information ? ». C’est un précurseur sur l’appréhension de la notion de d’hyperlien/d’hypertexte.
* Otlet avec Mundaneum. Avocat du 20eme siècle, il est à l’origine de l’émergence d’une nouvelle discipline aux côtés de la bibliothéconomie, nommé « Documentation ». Ayant vécu la WW1, il pensait que si les gens se faisaient la guerre c’est parce qu’ils ne se connaissent pas et ne connaissaient pas ce que les écrits du voisin. D’après lui, si on passait par la connaissance et la culture et si on constituait une base de données bibliographique ainsi qu’un système qui va rassembler toutes les connaissances produites par toutes populations mondiales : on va diminuer les conflits. L’histoire va lui donner tord mais sa manière de voir le monde va être reprise dans beaucoup de conférences et de colloques au 21ème siècle, on le considère un peu comme un visionnaire.
* Après toutes ces évolutions, on remarque qu’elles se basent sur des idéologies et sur d’autres modèles existants. Ce prisme n’est qu’un début mais on voit qu’aujourd’hui, il y a l’apparition d’une quête permanente d’un modèle idéal qui nous permettrait d’être indépendant de toutes ces sur-sollicitassions technologiques.

Au fil des siècles, les contenants de l’information changent et évidemment, cela change le contenu en son sein ainsi que la manière d’y accéder.

**Définir le document numérique** : « *La représentation d’un document sous la forme d’une structure de données informatique entreposable dans la mémoire d’un ordinateur et transmissible d’un ordinateur à un autre. Sa nature dépend du format auquel se conforme le document électronique* ». Source IFLA (international federation librery association)

Dans cet espace informatisé, on ne parle pas du document en lui-même mais de sa représentation et cela fait qu’il est structuré et constitué de tel sorte pour que la machine puisse le visualiser, le traiter, le gérer, etc. L’information stockée dans ce contenant va donc être transférée dans un espace physique de l’ordinateur appelé « mémoire » et qui peut aussi être diffusé et envoyé. La nature de ce document va donc dépendre du format.

« Mais comment peut-on expliquer pourquoi on est arrivé à numérique aujourd’hui alors que l’on parlait beaucoup d’électronique auparavant ? » Dématérialisation de l’information + électronique étant le côté technique/électrique et de la dimension de codage + numérique étant le côté visuel/visualisation et accessible pour un individu lambda.

////////////////////////////////////////////////////////////////////////////////////////////