



HAL
open science

Interpréter la voix de Deleuze

Samuel Szoniecky

► **To cite this version:**

Samuel Szoniecky. Interpréter la voix de Deleuze. *Frontières d'archives Recherches, mémoires, savoirs, Editions des archives contemporaines*, p. 165-177, 2015, 9782813001771. <hal-01278875>

HAL Id: hal-01278875

<https://hal.science/hal-01278875v1>

Submitted on 24 Jan 2018

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.



HAL Authorization

Interpréter la voix de Deleuze

Exemple de jardinage des connaissances

SAMUEL SZONIECKY, LABORATOIRE PARAGRAPHE

A partir d'une exploration des cours de Gilles Deleuze, une application Web pour interpréter les enregistrements audio a été développée. La conception de cet outil a fait émerger des problématiques tant techniques, qu'ergonomiques ou éthiques. Le présent article propose de les exposer et de présenter les solutions adoptées pour réaliser cet outil en montrant comment associer les modèles arborescent et rhizomatique pour concevoir une interface d'indexation et de diffusion de fragments multimédias.

1. Introduction

Des siècles de traitement documentaire et d'élaboration d'un discours scientifique aboutissent aujourd'hui à des pratiques relativement stables dans l'utilisation des ressources bibliographiques. Il reste encore de nombreuses recherches et améliorations à faire dans ce domaine (Moed, 2005), mais la stabilité de ces pratiques permet des traitements de bibliométrie consistant par exemple à évaluer l'impact de telle ou telle publication suivant le nombre de fois où elle a été citée, par qui, à quelle période... Il n'en va pas de même avec les contenus audiovisuels malgré leur disponibilité de plus en plus importante et leur intérêt scientifique indiscutable.

Avant même de s'interroger sur la primauté de l'écrit sur l'image et le son qui ferait privilégier le livre par rapport aux autres médias, le problème de la citation des contenus audiovisuels fait qu'il est encore délicat aujourd'hui d'en faire référence dans une production scientifique ou lors d'un cours. En effet, s'il est facile de faire référence à la page d'un livre pour indiquer précisément d'où vient une citation, il n'existe pas de pratique standard équivalente qui permette de retrouver facilement tel extrait sonore ou de film. Ne pouvant citer précisément les sources d'une référence, les chercheurs ou les enseignants préfèrent se cantonner aux livres ou à une transcription écrite de l'extrait plutôt que de tenter des manipulations délicates pour extraire et mettre à disposition la citation.

L'utilisation scientifique des contenus multimédias fait l'objet de nombreux travaux de recherche. Par exemple, celles entreprises dans le cadre de l'ANR ASA-SHS qui ont pour objectifs avec « L'atelier de sémiotique audiovisuelle », de fournir les bases théoriques, techniques et pratiques d'un moteur de recherche capable de facilement retrouver dans un flux audio ou vidéo des extraits pertinents par rapport à une interrogation à base de mots-clés (Prié, 2011 ; Stockinger, 2011). De même les travaux autour de Ligne de temps (Grison, 2009), de Canal-i-média (Amato et Péreny, 2011) de Vannotea (Hunter et Schroeter, 2011) ou de DiVA (Grunewaldt *et al.*, 2006) qui en optimisant le traitement des flux multimédias permettrons sans doute d'optimiser la citation de flux audio ou vidéo dans des travaux scientifiques.

Dans cette perspective, nous présentons une expérimentation dont l'objectif est le développement d'un écosystème d'information pour citer des fragments de la voix de Gilles Deleuze dans les cours qu'il donna à Paris 8 entre 1980 et 1985. Notre ambition n'est pas ici de rivaliser avec l'ampleur des programmes de l'ANR mais de proposer un point de vue sur ces questions afin d'alimenter la recherche avec des résultats pratiques. Pour ce faire, nous décrivons les contraintes initiales puis nous présenterons l'architecture technique de l'écosystème d'information mis en place. Enfin nous justifierons pourquoi nous avons choisi de nous orienter vers la récolte des interprétations, nous analyserons les résultats de cette expérience et décrivons les voies qu'ils restent encore à explorer.

2. Transformer « la voix de Deleuze » : vers une navigation rhizomatique

Le site « la voix de Gilles Deleuze » propose une série de cours au format audio et leur transcription au format texte faites par l'association « Siècle Deleuzien » avec l'aide de bénévoles et l'autorisation des ayants droits. Il permet de naviguer de cours en cours et d'en prendre connaissance gratuitement par la

lecture et l'écoute. Les séances correspondant à une journée de cours sont organisées dans une arborescence de huit thèmes et découpées entre une et quatre page Web possédant chacune un fichier mp3 et sa transcription. La gestion éditoriale de l'ensemble se fait avec SPIP un CMS (Content Management System) grand public.

Simple d'utilisation, ce site permet d'accéder facilement à l'information essentielle mais manque toutefois, à l'heure du rapprochement entre le Web sémantique et le Web social (Bachimont *et al.*, 2011), d'une indexation socio-sémantique pour enrichir la navigation de parcours croisés à la fois par rapport au thème et aux individus ayant consulté le site. Par l'intermédiaire de « capteurs sémantiques » (Zacklad, 2010) on peut créer une indexation à partir d'un « tagging sémantique » (Limpens, 2010) mettant en relation des documents, des acteurs et des concepts. Ainsi transformé, le site permettrait non seulement d'accéder à une connaissance linéaire des cours de Deleuze mais aussi à une connaissance des rapports entre des tags, des documents et des acteurs. Plus encore, ce dispositif de lecture-écriture permet l'expression d'une connaissance par la création de nouveaux rapports entre des tags et des documents voir l'invention de nouveaux tags et l'ajout d'acteurs. Sur la base de ces indexations sociaux-sémantiques, une modélisation fragmentaire des contenus voit le jour au fur et à mesure de la consultation des contenus pour atteindre un niveau de détail et une pertinence que les traitements automatiques de sont pas capable d'obtenir. Muni d'un tel dispositif, l'utilisateur des cours de Deleuze pourra documenter les traces de ces explorations et les partager facilement.

2.1. De l'arbre au rhizome

Dans l'introduction à *Mille Plateaux* (Deleuze et Guattari, 1980), les auteurs mettent l'accent sur les limites du modèle arborescent et sur la nécessité de le concilier avec le modèle rhizomatique :

« Ce qui compte, c'est que l'arbre-racine et le rhizome-canal ne s'opposent pas comme deux modèles : l'un agit comme modèle et comme calque transcendants, même s'il engendre ses propres fuites ; l'autre agit comme processus immanent qui reverse le modèle et ébauche une carte, même s'il constitue ses propres hiérarchies, même s'il suscite un canal despotique. » (Deleuze et Guattari, 1980, p. 31)

La richesse des écrits de Deleuze et Guattari demanderait un article à part entière pour décortiquer leurs propositions et voir dans quelles mesures elles s'appliquent aux sciences de l'information et plus particulièrement au design des écosystèmes d'informations dans un rapprochement entre Web et philosophie. Notre objectif ici consiste à proposer une hypothèse de travail qui s'inspire de nos précédents travaux sur l'ontologie-éthique de Spinoza vue par Deleuze (Bootz et Szoniecky, 2008), sur une méthode graphique de filtrage des flux

d'information et leur mise en application dans un langage allégorique pour le Web (Szoniecky, 2011) et sur la modélisation d'existence dans les écosystèmes d'information numérique (Szoniecky et Hachour, 2014). Dans cette optique, nous interprétons la proposition d'associer l'arbre et le rhizome comme la mise en rapport par un acteur ici et maintenant, d'une dimension physique-métrique-transcendante-extérieure : l'arbre ; avec une dimension conceptuelle-topologique-immanente-intérieure : le rhizome. Ce que nous pouvons représenter par le diagramme suivant qui montre comment ces trois dimensions s'articulent pour représenter une potentialité de connaissances :

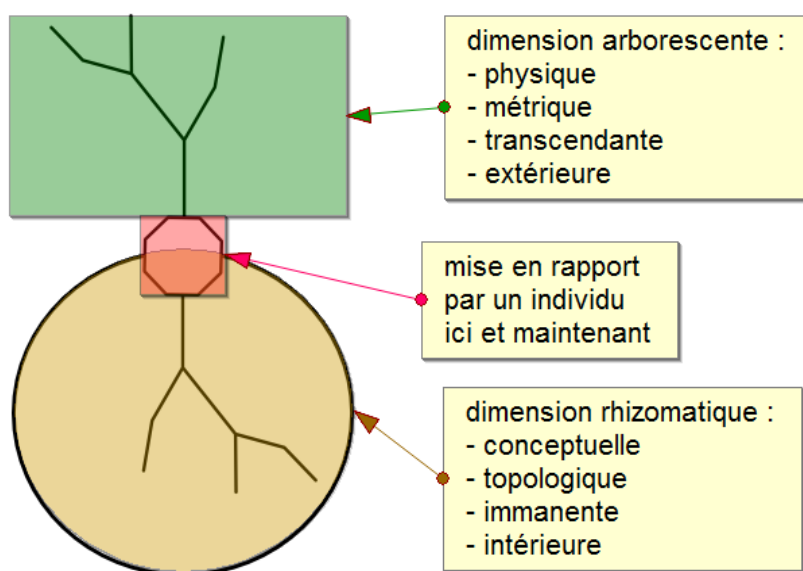


Figure 1. Agent allégorique arbre <-> rhizome

Dans la version actuelle du site, les rapports entre les documents physiques et les concepts sont très simples. En effet, si on prend comme base les liens hypertextes la hiérarchie des documents est à quatre niveaux : le site - la page du thème - la page du cours - le cours au format audio (mp3). De même concernant l'organisation des concepts qui elle aussi est à quatre niveaux : la voix de Gilles Deleuze - le nom du thème - le numéro, la date du séance et l'indice de la transcription - le cours en lui-même. Les acteurs participant à la vie du site sont eux aussi organisé en quatre ensembles différents mais sans positionnement hiérarchiques des uns par rapport aux autres : les responsables de l'association « Siècle Deleuzien », le comité scientifique, les personnes

remerciées et les auteurs de transcriptions. On remarquera que dans le site aucun formalisme de type micro-format ne permet de différencier la mention de ces individus, mis à part les responsables de l'association dont les noms sont exprimés sous la forme d'un lien hypertexte renvoyant à leur adresse email.

A l'énoncé de cette description, on peut dire que le site propose un point de vue univoque accessible par une navigation arborescente. Un seul parcours entre les documents et les concepts est disponible du point de vue d'un seul acteur (le site) qui exprime des faits bruts dans leur simplicité et leur évidence. Le site privilégie une approche documentaire arborescente pour présenter les cours de Deleuze le plus simplement possible en laissant au lecteur le soin de découvrir la richesse des contenus et de la pensée qu'ils véhiculent. En prenant les principes de l'agent allégorique que nous avons présenté plus haut, nous avons réalisé le diagramme du site « la voix de Deleuze » en mettant en relation l'arbre physique des documents, les rhizomes des topologies conceptuelles et la mise en rapport des ces deux dimensions par un ensemble d'acteurs.

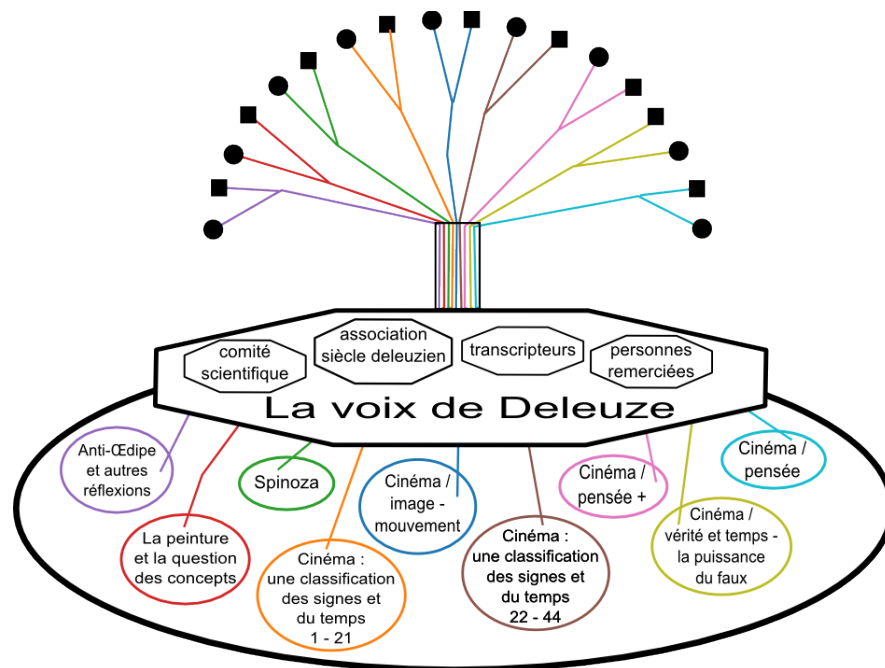


Figure 2. Diagramme du site "La voix de Deleuze"

2.1. Du rhizome au crible

Comme le propose Deleuze et Guattari dans la citation que nous avons placée en introduction de ce chapitre, essayons d'associer à cette organisation arborescente documentaire, une organisation rhizomatique de l'information. Pour ce faire, il convient dans un premier temps d'exprimer la dimension conceptuelle en rapprochant « un arbre syntaxique et un graphe conceptuel » comme l'approche classique de représentation des « arbres conceptuels » le propose (Amardeilh, 2007, p. 65). Cependant, nous nous détachons de cette approche qui minimise la dimension du rapport entre syntaxe et concept en concevant cette relation comme unique. En posant comme postulat de la modélisation la multiplicité de la relation, notre démarche se place d'avantage dans les perspectives définies par Mugur-Schächter qui introduit explicitement la dimension du rapport dans sa méthode de conceptualisation relativisée (Mugur-Schächter, 2006) dont la forme la plus condensée « DM/G,meG,VM/ » montre comment la description d'un micro-état sémantique est le résultat d'une adéquation entre une entité objet de l'ordre du document (meG), une grille de qualification de l'ordre du conceptuel (VM/) et un générateur de vue de l'ordre du rapport (DM) :

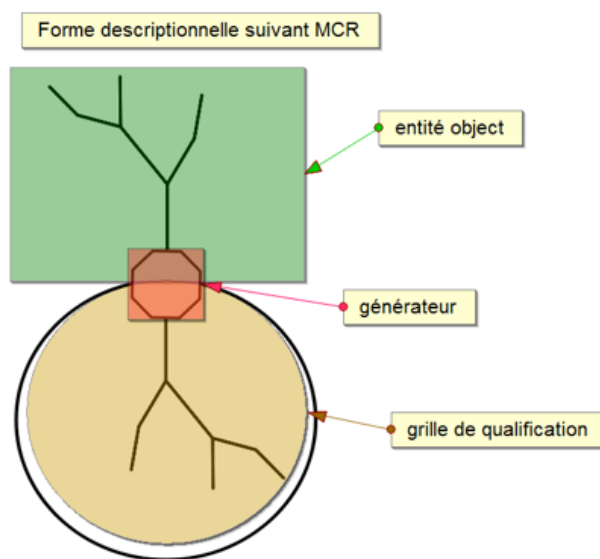


Figure 3. Diagramme Méthode de Conceptualisation Relativisée

Pour développer une organisation rhizomatique sur le site, nous avons donc besoin d'enrichir la structure documentaire existante de deux couches informationnelles supplémentaires. D'une part, une représentation rhizomatique de concepts et d'autre part le réseau des acteurs qui effectueront la mise en relation des potentialités documentaires et conceptuelles ; le tout constituant un système dynamique de représentation des connaissances socio-sémantique. Dans cette optique, nous concevons les rhizomes conceptuels à la manière d'un « crible » ayant une double fonction à la fois comme représentation des connaissances et comme générateur de l'événement les produisant :

« L'événement se produit dans un chaos, dans une multiplicité chaotique, à condition qu'une sorte de crible intervienne. » (Deleuze, 1988, p. 103)

Dés lors, la représentation des connaissances associe lecture et écriture pour créer une potentialité de connaissances. En ce sens nous suivons l'interprétation par Deleuze du « crible » comme étant la forme diagrammatique (Batt, 2005) de la chôra si chère à Platon (Berque, 2009, p. 30-35) et Derrida (Derrida, 1997) et dont on peut donner la définition suivante :

« Polysémique, le terme chôra évoque l'entrelacement de "l'aspect constitutif" et de "l'aspect spatial", "ce en quoi" apparaissent les choses sensibles et "ce de quoi" elles sont constituées. » (Zamora, 2003, p. 22)

Mais pour que cette potentialité générative de la chôra, génère effectivement quelque chose, il faut l'associer à une dynamique particulière, à un pouvoir d'agir. En effet, le crible du chercheur d'or n'est efficace que si on le remue. De même le champ donnera du blé seulement si on y plante des graines de blé dont la dynamique interne transformera les potentialités offerte par la terre en végétal. Dans notre conception de la navigation rhizomatique, la dynamique particulière est apportée par l'acteur qui, seul ou en réseau, humain ou machine, officiera à la manière d'une graine pour générer des connaissances par la mise en relation des rhizomes conceptuels et des arbres documentaires. Pour que l'analogie soit plus compréhensible, on peut dire que l'acteur sera comme le lierre qui pousse sur l'arbre en suivant l'organisation des branches existantes et en se nourrissant du même terroir. Ainsi, l'acteur en discernant des rapports entre des documents-branches et des concepts rhizomes construira un point de vue particulier qu'il pourra ensuite faire évoluer, partager ou faire disparaître.

Tous les jardiniers savent que le lierre est un parasite et que si on n'y prend pas garde, il peut tuer l'hôte qui l'accueille. En enrichissant le site d'une navigation rhizomatique, on pourrait croire que le danger est équivalent et qu'à force de multiplier les points de vue on finisse par perdre l'essence même de ce site à savoir : diffuser la pensée de Gilles Deleuze. Toutefois, les écosystèmes

d'information Web offre l'avantage d'être hyper-plastique, c'est-à-dire de pouvoir se transformer en n'importe quelle forme et de revenir à une forme d'origine. En effet, à tout moment l'utilisateur peut choisir d'afficher ou non une couche d'information ou de retrouver la version originale du site. Placé ainsi face à un choix, l'utilisateur devient acteur de sa connaissance et non plus seulement spectateur, il doit interpréter, se situer et discerner ce qu'il faut prendre en compte ou occulter, ce qui pour lui est important ou non :

« Or la pratique interprétative s'avère offrir une voie particulièrement propice pour approcher la question de savoir ce qui rend quelque chose important (ou non). Même si cette question peut difficilement être abordée de front, l'effort de jugement réfléchissant et d'explicitation des pertinences qui va de pair avec l'expérience interprétative - en particulier dans les domaines littéraire et philosophique - conduit naturellement à se demander ou nom de quoi quelques chose peut apparaître comme important. » (Citton, 2010, p. 82)

Mettre en place une navigation rhizomatique consiste à multiplier le potentiel des relations croisées entre une arborescence documentaire et un réseau de mot-clefs, et en même temps impliquer un tiers qui instanciera des relations particulières par des choix cohérents. En multipliant ces choix, le tiers modélise un réseau de discernement qui constitue son point de vue, ce qu'on peut aussi appeler son identité informationnelle ou pour reprendre les mots de Leibnitz son « tissu de l'âme ». La dimension numérique des documents que nous traitons, laisse envisager à la fois un traitement automatique de cette mise en relation mais aussi un traitement manuel dans le sens où le choix d'une relation est effectué par un humain. Dans la suite de cet article nous allons montrer comment mettre en place une collaboration entre les traitements automatiques et manuels et comment cette collaboration peut aboutir à la modélisation d'une identité informationnelle dédiée à la présentation d'un contenu multimédia pour un usage scientifique.

3. L'écosystème d'information mis en place

Plutôt que le concept d'architecture de l'information qui impose une vision rigide, figée et monumentale, nous préférons utiliser ici celui d'écosystème d'information mieux adapté au contexte du Web de part sa dimension constamment évolutive, modulaire et interconnecté. La notion d'écosystème est d'avantage en adéquation avec l'idée que le développement d'une application Web consiste aujourd'hui à créer des composants informatiques qui constitueront autant d'agents autonomes qui seront mis en relation avec d'autres agents dans des espaces informationnels proposant des ontologies c'est

à-dire des « modes d'existences » bien définies tant d'un point de vue technique que sociologique (Ferber, 2009 ; Latour, 2012).

Dans ce chapitre nous décrivons l'écosystème d'information que nous avons mis en place en précisant tout d'abord les contraintes documentaires qui nous étaient imposées puis les choix techniques que nous avons adoptés. En effet, afin de réduire les temps de traitements pour les expérimentation que nous avons menées, nous nous sommes focalisé uniquement sur les cours consacrés à Spinoza car il nous semblait que le projet de Spinoza de mettre en relation éthique et ontologie résonnait pleinement dans le domaine des sciences de l'information et de la communication à l'heure où les ontologies deviennent des normes pour l'organisation de l'information et où les questions d'éthique (identité, usage, droit...) focalisent autant les préoccupations des utilisateurs que l'attention des chercheurs. L'autre contrainte documentaire qui nous semblait importante de respecter consistait à ne pas modifier le site existant mais à développer une couche informationnelle autonome que nous définissons comme une niche éco-informationnelle.

3.1. De l'arborescence documentaire aux rhizomes conceptuels

Le premier problème que nous avons abordé dans ce projet porté sur l'extraction des mots clés de l'arborescence documentaire du site pour constituer les premiers rhizomes. Pour ce faire, il existe une multitude d'outils accessibles via des service Web¹ ou des applications à installer sur un ordinateur. Un travail d'analyse de ces différents outils serait sans doute utile à la communauté scientifique mais cela dépasse le cadre du propos que nous souhaitons développer ici. Parmi ces outils nous avons choisi d'utiliser le plugin Firefox KGEN² et le moteur de recherche Lucene³. Le scénario expérimental que nous proposons ici n'a pas pour vocation d'être un modèle pour le développement d'une modélisation rhizomatique, il propose juste la description d'une manière de faire sans rentrer dans le détail des problématiques d'extraction de l'information pour renseigner des bases de connaissances (Knowledge Base Population KBP) (Ludovic, 2011).

KGEN est le premier outil que nous avons utilisé. Il permet d'extraire très facilement des mots clés à partir d'une page Web et de récupérer ces informations sous la forme d'un tableau de valeurs. Il utilise pour cela un algorithme classique pondérant les mots du texte en comptant leurs occurrences, éliminant ceux qui sont trop court ou présent dans une liste des mots à exclure. De plus, l'algorithme ajoute une pondération supplémentaire

1 voir par exemple <http://www.mkbergman.com/sweet-tools-simple-list/> :

2 <http://kgen.elitwork.com/accueil.html>

3 <http://lucene.apache.org/core/>

correspondant à la structure documentaire de la page HTML, par exemple les mots d'un titre de niveau 1 (<h1>) ont une pondération supplémentaire de 10 alors que les mots d'un titre de niveau 2 (<h2>) ont une pondération supplémentaire de 8. Nous avons appliqué l'algorithme à chaque page des cours de Deleuze sur Spinoza en utilisant le paramétrage par défaut de KGEN.

Pour rendre plus accessible les extractions de mots clef obtenus et dans l'optique d'utiliser pleinement les avantages des services en ligne (Cloud Computing) nous avons créé une identité informationnelle dédiée au projet : mie.voixdeleuze@gmail.com. A partir du compte Google lié à cette identité, nous avons créé des tableurs (spreadsheet) pour conserver et partager les mots clef de chaque cours⁴. Ainsi les mots clefs pour chaque cours sont accessibles et modifiables par toute personnes disposant des droits pour le faire. On peut alors envisager de multiplier les tableurs pour par exemple mener le travail de comparaison entre les différents outils d'extraction que nous avons mentionné plus haut. De plus, pour faciliter encore l'intégration dans un écosystème Web et contribuer au développement du mouvement « Open Science », le résultat de cette extraction a été enregistré dans une base MySQL dont les données sont accessibles au format JSON (JavaScript Object Notation) soit pour toutes les cours⁵, soit pour une seule transcription⁶, soit pour un mot clef particulier⁷.

Le deuxième outil que nous avons utilisé pour extraire des mots-clefs est le moteur de recherche Lucene qui permet d'indexer des documents pour ensuite les retrouver par rapport à des requêtes multicritères. Pour optimiser le déploiement de ce moteur de recherche dans un environnement Web-PHP, nous avons utilisés dans ce projet le Framework Zend qui propose une implémentation de Lucene⁸. Dans un premier temps nous avons indexé les pages du site grâce à un algorithme⁹ qui nous a permis de récupérer dans le code source des pages HTML : l'url du cours, le titre du cours, le lien vers le fichier mp3 du cours, la transcription du cours. Ces informations ont été compilées dans l'index Lucene est rendu accessible par une page Web dédiée à sa consultation¹⁰.

4 Voir par exemple : <http://goo.gl/TwIe8u>

5 <http://gapai.univ-paris8.fr/jdc/public/deleuze/>

6 Par exemple : <http://goo.gl/8ubyN9>

7 Par exemple : <http://gapai.univ-paris8.fr/jdc/public/deleuze?tag=ontologie>

8 <http://framework.zend.com/manual/fr/zend.search.lucene.overview.html>

9 Le code de l'algorithme est accessible ici : <http://goo.gl/ortRDj>

10 Résultat HTML : <http://gapai.univ-paris8.fr/jdc/public/chercher>

Résultat JSON : <http://goo.gl/94qE91>

Grâce à KGEN et Lucene nous avons pu créer des relations automatiques entre les documents HTML des cours et une liste de mots-clefs. Toutefois, les réponses obtenues se situent à l'échelle de la page Web or celle-ci correspond en moyenne à une heure de cours et entre 17 000 et 65 000 caractères, ce qui ne répond pas à notre objectif premier qui est de pouvoir citer un cours de Deleuze à la page près (+/- 3 000 caractères) voir même à la phrase prêt. Il convient donc d'augmenter l'échelle de consultation du document en diminuant la granularité de l'extraction. Pour consulter directement un fragment du document, nous avons développé un algorithme qui permet de découper le texte d'un document par rapport à une requête multicritère. Cet algorithme se décompose en deux fonctions. La première qui découpe la totalité du texte par rapport à la requête multicritère¹¹ et la deuxième qui extrait un fragment textuel pour chaque élément découpé¹², le tout renvoyant pour chaque document trouvé par le moteur de recherche, un tableau de fragment avec le mot-clef recherché, sa position, le début du fragment, la fin du fragment¹³. Pour dimensionner les fragments, nous avons fait le choix de la simplicité en coupant le texte suivant un maximum de caractère (par défaut 100) et en respectant l'intégrité des mots¹⁴.

Cette première étape du projet nous a permis d'extraire des documents constituant les cours de Deleuze, une liste des mots-clefs qui composent les premiers rhizomes de concepts et leur relation avec un fragment précis du document. Il nous faut maintenant répondre à la question de savoir comment synchroniser ces fragments textuels avec les fragments sonores qui leur correspondent.

3.2. Synchronisation du texte et du son

Comme le remarque Thierry Bazillon dans sa thèse sur le traitement de la parole pour la reconnaissance automatique :

« La reconnaissance automatique de la parole est en effet un défi qui passionne la communauté scientifique depuis de nombreuses années. Mais le mythe d'une machine qui, à défaut de les comprendre, retranscrirait parfaitement les propos d'un être humain s'est peu à peu fissuré. » (Bazillon, 2011, p. 14)

En effet, dans cet thèse, l'auteur met l'accent sur les progrès importants fait par les outils de traitement automatique notamment grâce à des programmes de recherche ambitieux comme ceux du projet EPAC mais il insiste aussi sur

11 <http://goo.gl/Cgu8vz>

12 <http://goo.gl/tqLNZ2>

13 Pour un tableau HTML : <http://goo.gl/pez7Wd>

Pour un JSON : <http://goo.gl/K9vuq3>

14 <http://goo.gl/M8SJe5>

l'incontournable dimension humaine et sur le travail fastidieux que représente la transcription d'un document audio en texte même quand celui-ci consiste à corriger les erreurs de la machine.

Dans le projet qui nous occupe, ce travail de transcription a déjà été fait par les membres de l'association « Siècle deleuzien », mais l'absence de synchronisation entre les documents audio et les documents textes ne permet pas d'entendre la voix correspondant à un fragment de texte ni l'inverse. Pour pouvoir citer précisément un fragment audio du cours de Deleuze, un travail d'alignement du texte et du son est donc nécessaire. Si pour faire ce type de travail des outils existent depuis près de 20 ans (Bazillon, 2011, p. 139), nous n'en avons trouvé aucun qui permette de le faire directement sur le Web même si les prémices de tels outils sont en train d'apparaître. En effet, depuis l'arrivée de la version 5 du HTML, de nouvelles balises permettent d'afficher et de manipuler des documents multimédias : `<audio>`, `<video>`. Parallèlement, des bibliothèques Javascript sont apparues pour faciliter la gestion de l'interactivité avec l'utilisateur et entre les autres éléments documentaires de la page Web comme par exemple la bibliothèque PopCorn qui fait partie d'un projet plus vaste de la fondation Mozilla pour mettre en place « An expanding ecosystem » pour la création et de développement des interactivités entre des documents multimédias et des données Web comme celle de Wikipedia, Twitter, des cartes géographiques, des sites de partages de photo (Flickr...) ou de musique (last.fm...). Parmi les nombreux exemples disponibles aucun ne correspondait pleinement à notre projet, nous avons donc décidé de développer un outil d'alignement du texte et de l'audio accessible en ligne et permettant de citer facilement les fragments de texte et de son.

L'application Web d'alignement de l'audio et du texte que nous avons conçue a été développée avec : la bibliothèque Javascript D3 (Data Driven Document) qui permet de manipuler des données pour créer des graphiques SVG interactifs, la bibliothèque Javascript Popcorn dont nous avons parlé ci-dessus et la plateforme PHP Zend avec laquelle nous avons développé une application MVC (Modèle Vue Contrôleur) pour gérer les données des transcriptions, des alignements, des citations et des utilisateurs connectés à l'application... Les sources informatiques de l'application sont accessibles sur la plateforme de développement collaboratif code.google.com comme sous partie d'un projet plus vaste : le jardin des connaissances. Nous ne rentrerons pas plus dans les détails du développement informatique, précisons simplement que pour des raisons de compatibilité avec les fichiers de type Mp3 l'application ne fonctionne que sur le navigateur Chrome.

Le principe de l'interface que nous avons développée est très simple, il consiste à mettre en parallèle deux barres horizontales représentant pour l'une la totalité du son, exprimé en nombre de seconde, et pour l'autre la totalité du texte, exprimé en nombre de caractère. Ces barres réagissent au survol par la souris. Dans le cas de la barre du son on affiche la position en heure, minute,

seconde et dans le cas de la barre du texte on affiche les trois mots précédents et suivant le caractère survolé par la souris. Dans ce contexte, une citation du cours de Deleuze correspond à un fragment textuel délimité par un intervalle entre la position d'un caractère de départ et d'un caractère de fin auquel est associé un fragment de son délimité par un intervalle entre deux positions en seconde définissant le début et la fin d'un fragment. Chaque fragment possède sa propre représentation sur la barre de texte et de son. Cette représentation est dynamique et interactive pour permettre à l'utilisateur de déplacer le fragment en entier ou grâce à des « poignets » modifier les limites de début et de fin du fragment soit librement soit mot à mot pour le texte ou seconde à seconde pour le son. A chaque déplacement d'un fragment, le son est déclenché et le texte correspondant est affiché linéairement et sous la forme d'un nuage de mots-clés calculé automatique. Ainsi, il est possible de définir très simplement l'alignement d'un fragment de son et d'un fragment de texte.

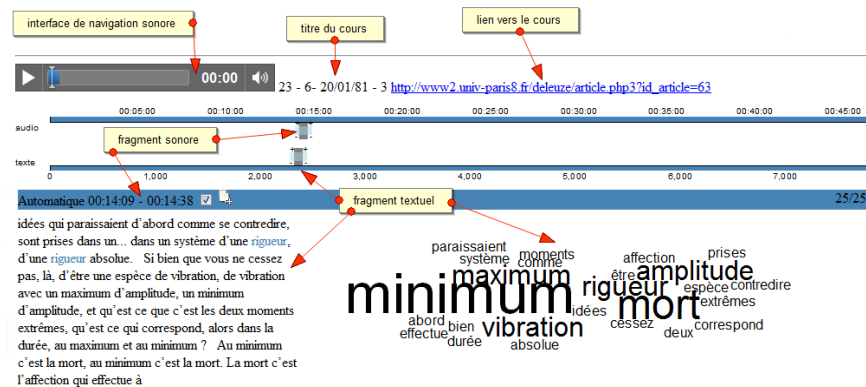


Figure 4. Interface d'alignement du son et du texte

L'exemple que nous venons de présenter est le résultat de l'alignement automatique d'un fragment textuel trouvé à partir d'une requête sur le mot clef « rigueur » (en bleu dans le texte) avec un fragment audio calculé par une simple extrapolation du nombre de caractère avec le nombre de seconde (nb. de seconde total / nb. de caractère total * position du caractère sélectionné). Il va sans dire que cet alignement automatique est une approximation qui génère de nombreuses erreurs car un discours oral n'a jamais la régularité d'un tel calcul. Nous n'avons pas poussé plus avant nos recherches sur une analyse algorithmique plus précise du débit du locuteur qui permettrait sans doute d'affiner grandement cette approximation automatique préférant nous focaliser sur les capacités humaines d'analyses et les moyens de la faciliter. Nous avons

donc élaboré un scénario où l'humain est au cours d'un processus d'interprétation stimulé par la machine.

3.3. Synchronisation du texte et du son

Pour fournir aux utilisateurs une matière sur laquelle exercer leur interprétation nous avons tout d'abord ajouté à notre application un double moteur d'interrogation. Le premier est un simple formulaire permettant de saisir une requête booléenne qui sera soumise au moteur Lucene. Le second est une liste de mots-clés définie par une existence humaine (collective ou individuelle) ou machinique (ex. KGEN). Cette liste est présentée sous la forme d'un nuage où chaque mot-clé est interactif et permet de lancer une requête Lucene. En référence aux interprétations de Deleuze sur la chôra que nous avons présentées plus haut, nous appelons ce deuxième moteur un crible. Ces deux moteurs d'interrogation renvoient la liste des documents et des fragments correspondant au résultat de la requête avec la possibilité d'afficher ou de masquer un fragment et surtout de se l'approprier pour en donner sa propre interprétation.

The screenshot displays a web application interface for text and audio synchronization. At the top, there is a search bar and a 'crible sélectionné' (selected filter) dropdown. Below the search bar, a 'nuage de mots-clés représentant le crible' (word cloud representing the filter) is shown, with the word 'abstrait' (abstract) being the most prominent. Other words in the cloud include 'abstraction', 'rigueur', 'fiction', and 'perception'. A 'mot-clé sélectionné' (selected keyword) is also indicated. The interface includes a timeline for audio and text, with a play button and a progress bar. The text section shows a paragraph with highlighted fragments and a word cloud of related terms. The word cloud includes words like 'générales', 'autres', 'faire', 'précis', 'peut', 'impression', 'tous', 'fait', 'tendus', 'considérations', 'dis', 'triste', 'parce', 'rapports', 'tout', 'étendue', 'abstraites', 'notions', 'très', 'tout', 'communes', 'choses', 'corps'. The interface also includes a 'bouton pour s'approprier un fragment' (button to appropriate a fragment) and a 'fragments masqués' (masked fragments) section.

Figure 4. Résultats d'une interrogation

Lorsqu'un utilisateur décide de s'appropriier un fragment en cliquant sur le bouton dédié à cet effet, une copie du fragment est effectuée et enregistrée dans la base de données en lien avec cet utilisateur. La liste des fragments s'enrichit alors d'un nouveau fragment qui possède des fonctionnalités supplémentaires pour pouvoir le supprimer, le mettre à jour, obtenir le lien vers ce fragment et surtout pour interpréter ce fragment en augmentant ou en diminuant le poids des mots clefs :

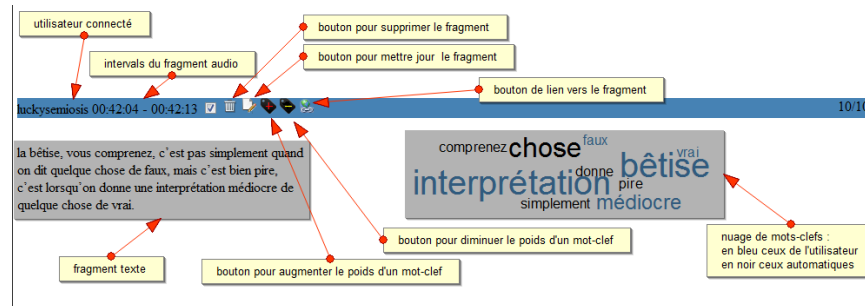


Figure 6. Interprétation d'un fragment

Dans l'exemple ci-dessus, l'utilisateur a sélectionné un fragment puis la aligné pour que le texte et le son correspondent. Ensuite il a sélectionné le bouton avec un « + » rouge pour, en cliquant sur les mots-clefs du nuage, augmenter le poids de « bêtise », « interprétation », « médiocre », « vrai », « faux » et en sélectionnant le bouton avec le « - » jaune, il a supprimé « quelque ». Outre la mise à disposition d'un lien direct vers le fragment qui permet de le citer facilement, l'utilisateur, grâce au dispositif d'augmentation et de diminution de la puissance d'un mot-clef, va pouvoir exprimer par un simple clic sur un mot-clef ce qu'il trouve d'important dans les cours de Deleuze. De plus, il pourra retrouver d'autres fragments à partir de ces mots-clefs qui constitueront dès lors le « crible » de cet utilisateur ou pour reprendre la définition de la chôra que nous avons donnée plus haut : « ce en quoi » apparaissent les connaissances et « ce de quoi » elles sont constituées.

Ainsi se met en place une boucle vertueuse où les efforts d'interprétation de l'utilisateur pour définir précisément les fragments des cours qui l'intéressent et ce en quoi ils sont importants, vont enrichir sa découverte des cours de nouvelles pistes d'exploration et ainsi lui permettre d'affiner de plus en plus sa propre interprétation des documents. C'est ce processus continu de sélection documentaire et d'affinage conceptuel que nous appelons le jardinage des connaissances et qui nous semble correspondre au projet de « Milles plateaux » proposé par Deleuze et Guattari. Dans ce processus, la machine agit comme un

« moteur écosystémique » qui va créer automatiquement des rapports entre les documents et les concepts, c'est elle qui va générer de nouvelles possibilités de connaissances et ainsi stimuler l'interprétation humaine.

4. A suivre...

Avec cette application, nous avons rempli notre premier objectif qui consistait à pouvoir citer simplement et précisément un fragment d'un contenu audio. Toutefois, il reste beaucoup de travail à faire pour que cette proposition d'interface puisse remplir pleinement son rôle d'organisation de l'intelligence collective et de mise en commun des efforts individuels d'analyse. Il faudra sans doute rendre encore plus simple l'utilisation de cette application pour que le travail d'indexation et d'alignement du texte et du son, fait par des explorateurs curieux, puisse être utile à tous. Il semble viable de continuer le développement de cette application dans le cadre des ateliers laboratoires CreaTIC¹⁵ qui permettent une approche multi disciplinaires et donnent les moyens pour des expérimentations pratiques. Avec notamment des enseignants et des étudiants venus des domaines de l'hypermédia, de l'histoire, des archives, de l'informatique, de l'art... les théories et les technologies utilisées pour cette application offrent un terrain très fertile pour développer les compétences qu'un étudiant de ces domaines doit pouvoir maîtriser. Parallèlement aux questions d'amélioration de l'ergonomie de cette application et de développement de nouvelles fonctionnalités, nous souhaitons à travers ce projet mettre en place une plate forme d'évaluation des outils d'extraction de l'information pour apprendre aux étudiants à se servir de ces outils et comparer leurs résultats. De même concernant les techniques d'indexation (Huynh-Kim-Bang, 2009), de tagging (Crepel, 2011) et d'adressage sémantique (Lévy, 2011) que nous aimerions là aussi transmettre aux étudiants en leur donnant les moyens de les tester en grandeur nature.

Enfin dans la mesure où l'application sera devenu stable et que nous aurons l'accord des contributeurs qui ont permis que le site « La Voix de Deleuze » soit en ligne, nous aimerions que cette application permette à tous de découvrir la navigation rhizomatique pour expérimenter le fait que la connaissance sur le Web n'est pas uniquement un choc affectif à la manière d'un « J'aime » ou une mise en relation complexe et raisonnable mais aussi l'intuition d'une essence.

15 <http://idefi-creatic.net/ateliers-laboratoires/>

Remerciements

Nous tenons à remercier Stravoula Bellos, Stéphane Malolepsa pour leur aide sans laquelle nous n'aurions pas réalisé cette application.

Bibliographie

- Amardeilh, F. (2007). Web Sémantique et Informatique Linguistique□ : propositions méthodologiques et réalisation d'une plateforme logicielle. Université de Nanterre - Paris X, Paris. Retrieved from <http://tel.archives-ouvertes.fr/tel-00146213/en/>
- Amato, É. A., & Perény, É. (2011). Audiovisuel interactif. *Communications*, 88(1), 29–36. doi:10.3406/comm.2011.2582
- Bachimont, B., Gandon, F., Poupeau, G., Vatant, B., Troncy, R., Gandon, F., ... Zacklad, M. (2011). Enjeux et technologies□ : des données au sens. *Documentaliste-Sciences de l'Information*, 48(4), 24–41. doi:10.3917/docs.484.0024
- Batt, N. (2005). L'expérience diagrammatique□ : un nouveau régime de pensée. In *Penser par le diagramme□ : de Gilles Deleuze à Gilles Châtelet*. Saint-Denis: Presses universitaires de Vincennes.
- Bazillon, T. (2011). Transcription et traitement manuel de la parole spontanée pour sa reconnaissance automatique. Université du Maine. Retrieved from <http://tel.archives-ouvertes.fr/tel-00598427/en/>
- Berque, A. (2009). *Ecoumène□ : Introduction à l'étude des milieux humains*. Belin.
- Bootz, P., & Szonieczky, S. (2008). Toward an Ontology of the Field of Digital Poetry. Presented at the Electronic Literature in Europe, Bergen.
- Citton, Y. (2010). *L'avenir des humanités□ : Economie de la connaissance ou cultures de l'interprétation□ ?*. Editions La Découverte.
- Crepel, M. (2011, November 30). Tagging et folksonomies□ : pragmatique de l'orientation sur le Web. Université Rennes 2, Rennes. Retrieved from <http://tel.archives-ouvertes.fr/tel-00650319>
- Deleuze, G. (1988). *Le pli*. Editions de Minuit.
- Deleuze, G., & Guattari, F. (1980). *Mille plateaux*. Paris: Éditions de minuit.
- Derrida, J. (1997). *Chora L Works: Jacques Derrida and Peter Eisenman*. Monacelli Press.
- Ferber, J. (1997). *Les systèmes multi-agents□ : vers une intelligence collective*. Paris: InterÉd.
- Grizon, X. (2009). Lignes de temps□ : l'analyse filmique en classe de français. *Le français aujourd'hui*, 165(2), 71–77. doi:10.3917/lfa.165.0071

Grunewaldt, L., Moller, K., Morisse, K., & Sempf, T. (2006). Collaborative Video Annotation Environments: Implemented Features and a new Approach for Object Detection and Recognition. In 2010 Fifth International Conference on Digital Telecommunications (Vol. 0, p. 14). Los Alamitos, CA, USA: IEEE Computer Society. doi:10.1109/ICDT.2006.25

Huynh-Kim-Bang, B. (2009). Indexation de documents pédagogiques : fusionner les approches du Web Sémantique et du Web Participatif. PhD thesis, Université Henri Poincaré-Nancy 1.

Latour, B. (2012). Enquêtes sur les modes d'existence : Une anthropologie des Modernes. Editions La Découverte.

Lévy, P. (2011). La sphère sémantique : Tome 1, Computation, cognition, économie de l'information. Hermes Science Publications.

Limpens, F. (2010, October 25). Multi-points of view semantic enrichment of folksonomies (Informatique). Université de Nice - sophia Antipolis - Ecole doctorale STIC. Retrieved from <http://tel.archives-ouvertes.fr/tel-00530714/fr/>

Ludovic, J.-L. (2011, December 15). Approches supervisées et faiblement supervisées pour l'extraction d'événements et le peuplement de bases de connaissances. Université Paris Sud - Paris XI, Paris. Retrieved from <http://tel.archives-ouvertes.fr/tel-00686811>

Moed, H. F. (2005). Citation analysis in research evaluation. Springer.

Mugur-Schächter, M. (2006). Sur le tissage des connaissances. Hermes Science Publications.

Prié, Y. (2011). Vers une phénoménologie des inscriptions numériques. Dynamique de l'activité et des structures informationnelles dans les systèmes d'interprétation. (HDR). Université Claude Bernard - Lyon I, Lyon. Retrieved from <http://tel.archives-ouvertes.fr/tel-00655574>

Schroeter, R., Hunter, J., & Kosovic, D. (2011). Vannotea: A collaborative video indexing, annotation and discussion system for broadband networks. In Knowledge capture (pp. 1–8).

Stockinger, P. (2011). les archives audiovisuelles description indexation et publication traite des sciences et techniques. Hermes Science Publications.

Szowiecky, S. (2011). Le langage du Web du symbolique à l'allégorique, vers une représentation de la connaissance en train de se faire. In ISKO - Magreb 2011. Hammamet, Tunisie.

Szowiecky, S., & Hachour, H. (2014). Monades pour une éthique des écosystèmes d'information numériques. Presented at the Digital Intelligence, Nantes. Retrieved from <https://hal-univ-paris8.archives-ouvertes.fr/hal-01070113/document>

Zacklad, M. (2010). Introduction aux appareils de capture sémantique dans les TIC et les SI. In Technologies de l'information et intelligences collectives. Hermes Science Publications.

Zamora, J. M. (2003). La chôra après platon. In Symboliques et dynamiques de l'espace (pp. 16–32). Publication Univ Rouen Havre.