

# Reconstruire les conditions d'intelligibilité du document numérique patrimonial : mobilisations documentaire et sémiotique des *Linked Open Data*

*Reconstruct digital heritage conditions of intelligibility: mobilize the Linked Open Data through a semiotic and documentary approach*

*Reconstruir las condiciones de comprensión del documento numérico patrimonial: movilización documental y semiótica del Linked Open Data*

Article inédit, mis en ligne le 28 octobre 2015

## Lénaïk Leyoudec

*Lénaïk Leyoudec est doctorant CIFRE en sciences de l'information et de la communication à l'Université de technologie de Compiègne, au sein de l'équipe d'accueil Costech (Connaissance, Organisation et Systèmes Techniques). Il compose un travail de recherche doctoral en partenariat avec la société Perfect Memory SA sur la problématique d'intelligibilité de l'archive numérique familiale.*

### Plan

Introduction : le fossé d'intelligibilité propre au document numérique

Mobiliser le web des données dans un cadre patrimonial

Le document patrimonial comme matière première d'une expérimentation sémiotique

*Présentation du corpus*

*Autour de la méthodologie construite*

*Étude sémiotique du film Vues de Biarritz et de Bayonne (Inconnu, 1950)*

*Bilan du terrain*

Implémenter les effets de sens dans un dispositif éditorial

Conclusion : mode d'existence de l'artefact construit

Références bibliographiques

## Résumé

Le document numérique patrimonial est soumis avec le temps à un double hiatus : non seulement son intégrité documentaire s'affaïsse mais également son intelligibilité culturelle. Considérant la redocumentarisation comme une méthode de recontextualisation du document, un dispositif annotationnel mobilisant les *Linked Open Data* comme une source d'informations contextuelles est construit. Ce dispositif est le fruit de l'instrumentation éditoriale des effets de sens – spécifiques à l'interprétation du document – identifiés au sein d'un corpus de documents. Ces effets émergent par la mobilisation de la sémantique différentielle en tant que démarche interprétative. La redocumentarisation de l'archive par l'expérience du dispositif est à terme interrogée quant à la nature de l'intelligibilité construite.

## Mots clés

Document, patrimoine, interprétation, web des données, sémiotique.

## Abstract

Numerical (documentary) heritage suffers from a double hiatus over time: both its documentary integrity and its cultural intelligibility erode. Considering redocumentation as a method for recontextualisation, we build an annotational artifact using *Linked Open Data* as a source of contextual information. Effects of senses are specific to the interpretation of documents. By identifying these effects on a corpus of documents using differential semantic, we chose an interpretative approach to create our artifact. Then we question the nature of document intelligibility obtained through the use of the redocumentation artifact.

## Keywords

Document, heritage, interpretation, web of data, semiotics.

## Resumen

El documento digital patrimonial presentado con tiempo un doble hiato: no sólo su integridad documental debilita sino también su inteligibilidad cultural. Considerando redocumentarización como método de re-contextualización del documento, un dispositivo anotacional movilizándolo el *Linked Open Data* como una fuente de información contextual se construye. Estos efectos emergentes por la movilización del diferencial semántico como ejercicio interpretativo. La redocumentarización del archivo por la experiencia del dispositivo se analiza en última instancia en función del carácter de la inteligibilidad construida.

## Palabras clave

Documento, patrimonio, interpretación, web de datos, semiótico

## Introduction : le fossé d'intelligibilité propre au document numérique

Le numérique est depuis une décennie un support majeur pour l'ensemble des contenus, en particulier patrimoniaux à l'instar des programmes de numérisation des collections muséales. Nous entendons l'acception de patrimoine comme tout objet auquel est attribuée une valeur mémorielle. Cette valeur peut être individuante, à l'instar d'un objet n'évoquant quelque chose que pour notre seule personne, ou identifiante : une communauté exprimant la nécessité de préserver ledit objet. L'accès à un contenu numérique n'est possible qu'après un double procédé d'instrumentation et de manipulation. En effet, aux médias ne nécessitant pas de la part du lecteur d'instrumentation supplémentaire, tel que le livre, s'opposent les médias technologiques soumis à une double manipulation : « à l'incontournable médiation culturelle, il faut donc ajouter une médiation technologique pour l'accès au contenu » (Mattelart, 1994). Le numérique, par définition, impose un décodage du contenu archivé afin d'en permettre son visionnage ou son utilisation. L'archive en tant que telle est une ressource codée et l'outil informatique propose à l'utilisateur une reconstruction du

contenu. Le média technologique est soumis avec le temps à deux problèmes qui impactent son utilisation : le fossé d'obsolescence – correspondant à la préservation des formats associés et plus largement aux moyens technologiques d'accès au contenu archivé – et le fossé d'intelligibilité (Bachimont, 2010, p. 22). Ce dernier incarne l'effet du temps sur la lisibilité culturelle du contenu. Décontextualisée, l'archive perd son rôle de réceptacle de la mémoire : son intelligibilité doit de fait être restaurée.

À la pratique orale de transmission des savoirs, « l'ancestral art de la mémoire » (Yates, 1975), viennent s'ajouter les pratiques documentaires (Briet, 1951), s'incarnant aujourd'hui en autant de traces numériques, tags et métadonnées, associées au document. De fait, l'intelligibilité de ce dernier est rendue possible par le biais d'un panel, ici caricatural, de pratiques professionnelles : de la médiation culturelle, documentaliste égrainant son savoir au fur à mesure d'entretiens, du système expert en passant par le cas de figure du documentaliste couplé à une solution logicielle spécifique. Nous inscrivons cette étude dans l'environnement de cette dernière pratique professionnelle en convoquant le questionnement suivant : le web des données, entendu comme un réseau global d'informations, peut-il participer de la chaîne de transmission des savoirs décrite ? Plus spécifiquement : peut-on mobiliser le web des données dans le cadre de l'exercice de l'intelligibilité du document patrimonial ?

## Mobiliser le *Web des données* dans un cadre patrimonial

Incarnant une composante majeure de cette étude, l'appellation *Linked Open Data*, que l'on pourrait traduire en français par « données liées ouvertes », mérite d'être décomposée. D'abord, les *Linked Data* désignent des données reliées entre elles. Cette liaison est permise par la conjonction d'une publication sur le Web et l'association à chaque donnée d'un identifiant unique : l'*Uniform Resource Identifier* (URI). Par le biais de combinaisons « sujet-prédicat-objet » baptisées « triplets », ces données peuvent être mises en relation avec d'autres données tandis que leur relation peut être qualifiée à l'instar d'une équivalence. Le respect d'un standard commun - le RDF - au niveau du formatage des triplets valide l'intégration des données liées dans le réseau global baptisé « Web des données » (Gandon, 2011).

Dans un second temps, l'*Open Data* (données ouvertes) désigne le mouvement d'ouverture des données des collectivités, des services publics et des entreprises, qui émerge depuis la fin des années 2000. L'objectif est d'arriver à une meilleure transparence de ces acteurs en partant du principe que ces données représentent un bien commun selon l'acception d'Elinor Olstrom [Référence ? n.d.l.r.] et devraient donc être accessibles par tous.

Afin d'être appréhendées comme ouvertes, ces données doivent respecter plusieurs critères : technique (fourniture des données dans un format non propriétaire), économique (liberté d'utilisation des données) et juridique - données sous licence ouverte - (Chignard, 2012). De fait, l'appellation *Linked Open Data* illustre la conjonction, à l'échelle des données, d'une structure (données liées) et d'un statut (données ouvertes). Ce contexte spécifique crée une interopérabilité de plusieurs sortes : technique (portée par le format RDF), juridique (absence de copyright) mais également économique (liberté d'accès) qui ouvre un large éventail de possibilités d'utilisation.

Historiquement, l'initiative *Linked Open Data* revient à Tim Berners-Lee, l'inventeur du *World Wide Web*. Celle-ci visait à simplifier les principes du web sémantique afin d'en favoriser le développement (Monmin, 2013). Celui-ci publie en 2006 quatre principes centrés autour de la notion

d'URI, encadrant la publication des données en ligne et l'interconnexion des jeux de données (Berners-Lee, 2006). Mettant en application ces principes de référencement et de nomination des ressources, le projet *Linking Open Data* est créé en 2007. Il est initialement porté par des chercheurs universitaires en informatique ainsi que des développeurs évoluant dans diverses entreprises privées, le tout accompagné par le W3C, avant d'être rejoint par des institutions telles que la BBC, Thompson Reuters ou la *Library of Congress* (Heath, 2009). En investissant cet espace de publication de contenus, ces acteurs favorisent l'émergence d'un écosystème encyclopédique qui constitue l'environnement faisant l'objet d'une expérimentation dans le cadre de cette recherche sur l'intelligibilité de l'archive. Le projet *Linked Open Data Cloud*, initié la même année, a pour principe de cartographier les données libres liées. En 2014, le graphe actualisé accueille désormais 570 silos de données, c'est-à-dire 570 acteurs privés ou publics ayant produit et rendu accessibles un ensemble de données sous une forme adéquate aux principes écrits par Tim Berners Lee (Cyganiak, 2014). Ensemble, ces différents silos forment le « Web des données » pensé par le même chercheur américain comme une reformulation du concept délicat de « Web sémantique » (Théodule, 2007). Au total, en 2014, plus de 900.000 documents décrivant plus de huit millions de ressources ont été déployés et liés. Des silos spécifiques coexistent avec des silos généralistes à l'instar de DBpedia, construit comme l'équivalent sémantique de Wikipedia. Créé en 2007, ce projet porté par l'Université libre de Berlin et l'Université de Leipzig comporte en 2014 dans sa version anglaise plus de 4,58 millions d'« entités » (*things*) avec 583 millions de « faits » (*facts*) (Thibodeau, 2015). Par son caractère générique, cette base de triplets sémantiques incarne un des acteurs privilégiés d'interrogation dans le cadre de l'étude présente.

Considérant le potentiel d'instrumentation du web des données dans le cadre spécifique de la gestion de l'intelligibilité du document patrimonial, les silos comme DBpedia ou Freebase incarnent autant de sources informationnelles mobilisées dans un procédé de recontextualisation de l'archive, reprise d'un objet culturel dans un nouveau contexte médiatique (Treleani, 2013). La spécificité des Linked Open Data - les données sont liées et ouvertes - permet d'envisager la mobilisation de manière générique de ces silos. Ces ensembles d'informations contextuelles sont ici intégrés comme première [étape ? n.d.l.r.] de l'action d'éditorialisation, processus consistant à enrôler des ressources pour les intégrer dans une nouvelle publication, réalisée sur l'archive numérique. Cette éditorialisation, contribuant au processus de valorisation de l'archive, prend la forme spécifique d'une redocumentarisation : l'injection au sein du document d'un certain nombre d'informations visant à valoriser le document (Pédauque, 2007). L'expérimentation réalisée sur l'archive numérique s'inscrit dans le cadre d'un dispositif éditorial spécifique, incorporant en son sein une visualisation de l'archive et des annotations associées. Déclinons le processus d'enrichissement sémantique du document : dans un premier temps, il s'agit de réaliser une sémantisation des métadonnées associées à l'archive. Dans un second temps, par l'acte annotationnel l'interrogation des bases de connaissances ouvertes est réalisée pour chaque entité inscrite dans le document. Si la requête aboutit alors la notice associée est rapatriée dans le dispositif.

Ce processus d'enrichissement sémantique est le moteur de l'artefact technologique annotationnel - conceptualisé et implémenté dans le cadre de ce travail de recherche - destiné à la restauration de l'intelligibilité culturelle de l'archive. Celui-ci prend la forme d'une interface homme-machine permettant à l'utilisateur d'y déployer ses archives, de les annoter en mobilisant les notices contextuelles des *Linked Open Data* et d'y faire l'expérience d'un parcours interprétatif inédit.

En cherchant à permettre la recontextualisation de l'archive au sein d'une pratique professionnelle documentaire, la méthodologie envisagée est bipartite : l'instrumentation sémiotique d'un corpus expérimental de documents afin de déduire les marqueurs de sens puis l'instrumentation de ces

résultats dans le squelette d'un dispositif éditorial expérimental destiné à la pratique de recontextualisation.

## **Le document patrimonial comme matière première d'une expérimentation sémiotique**

### *Présentation du corpus*

Dans l'intention de mener une étude sémiotique d'archives numériques, un ensemble de documents a été sélectionné parmi les fonds d'institutions patrimoniales :

- le fonds de films amateurs de l'Établissement de communication et de publication audiovisuelle de la Défense (ECPAD),
- un fonds de la Cinémathèque de Bretagne,
- un fonds du Forum des images
- et enfin le fonds *Mémoires partagées* de l'Institut national de l'audiovisuel.

### *Autour de la méthodologie construite*

Première des deux phases expérimentales de cette recherche, le corpus constitué est interrogé par une méthodologie d'analyse sémiotique du film construite autour de la sémantique différentielle (Rastier, 2001, 2009) : on vient ici interroger l'image pour en extraire les unités minimales de sens.

Le projet d'une sémantique différentielle est présenté ainsi : « nous souhaitons donner à la sémantique linguistique toute la place qui lui revient, en unifiant la description du lexique, de la syntaxe profonde, et des structures textuelles. À chacun des trois paliers traditionnels de la description linguistique (mot, phrase et texte) nous faisons alors correspondre trois paliers de la théorie sémantique (micro-, méso-, et macrosémantique) en unifiant leur conceptualisation » (Rastier, 2001, p. 25).

En guise de préambule à son approche, l'auteur précise : « le sens est un niveau d'objectivité qui n'est réductible ni à la référence, ni aux représentations mentales. Il est analysable en traits sémantiques qui sont des moments stabilisés dans des parcours interprétatifs »; « la problématique interprétative dépasse les textes et peut s'étendre à d'autres objets culturels, comme les images (susceptibles des mêmes méthodologies : recueil de corpus, détermination des genres, indexation par des traits de l'expression) » (Rastier, 2009, p. VI).

Les concepts de parcours interprétatif – suite d'opérations cognitives permettant d'assigner une signification à une séquence linguistique – et de trait sémantique sont convoqués dans le cadre de cette expérimentation. Le constat également réalisé par l'auteur est la compatibilité de son approche micro-sémantique avec d'autres objets culturels à l'image du document audiovisuel, matière première de l'étude présente.

Fort de ce bagage théorique, l'objectif de cette phase expérimentale est de définir les limites des parcours interprétatifs propres au visionnage de documents patrimoniaux. Un parcours interprétatif se compose de deux temps : l'assignation, phase de qualification de l'objet puis la sémantisation, marquant l'établissement d'une signification par élaboration de formes et fonds du contenu (Rastier, 2011, p. 9). Il s'agit par le biais d'une méthodologie sémiotique d'appréhender les effets de sens qui

se dégagent du film et d'y affecter un parcours interprétatif. L'effet de sens désigne les éléments, présents dans le film, identifiés intuitivement par un spectateur et participant du procédé sémiotique d'interprétation du sens associé à l'objet culturel.

### ***Étude sémiotique du film Vues de Biarritz et de Bayonne (Inconnu, 1950)***

Considérant l'approche interprétative présentée précédemment, trois passages ont été définis et interrogés : respectivement les plans n°13, n°14 et n°15 du film *Vues de Biarritz et de Bayonne* (INA Mémoires partagées). L'acception de « passage » est définie par Rastier de cette manière : *Les unités textuelles élémentaires ne sont pas des mots mais des passages. Un passage a pour expression un extrait et pour contenu un fragment. Sur le plan sémantique, les traits pertinents sont organisés pour composer des formes sémantiques, comme les thèmes, qui se détachent sur des fonds sémantiques, les isotopies notamment.* (Rastier, 2009, p. VI)

En guise de signifiant, les trois passages sont des extraits et interviennent précisément aux *times-codes* suivants : 1'55"-2'00", 2'01"-2'06" et 2'07"-2'22". Au niveau du signifié, les passages sont des fragments pointant vers des contextes latéraux : le premier passage est précédé d'un plan d'ensemble mettant à l'image une foule de baigneurs en bord de mer, tandis que le troisième passage est suivi d'un nouveau plan d'ensemble mettant à l'image le bord de mer et différents baigneurs.

La description réalisée - synthétique, plan par plan - mérite d'être mentionnée en guise de présentation de ces passages.

- Plan n°13, de 1'55" à 2'00" : plan large, architecture urbaine spécifique (Sud de la France), présence de végétation et de véhicules au premier et second plan, passants traversant le champ au second plan, personne avec chapeau traversant le premier plan, panoramique latéral mettant à l'image un jardin urbain.
- Plan n°14, de 2'01" à 2'06" : plan américain, caméra fixe, deux femmes et un homme (vêtus avec distinction : costume, tailleur, chapeau, cravate et canne) avancent vers la caméra dans une rue piétonne et quittent le champ par la gauche, véhicules, architecture et végétation urbaines à l'arrière-plan.
- Plan n°15, de 2'07" à 2'22" : plan d'ensemble, bord de mer avec plage, nombreux vacanciers avec matériel et jeux de plage, panoramique latéral, architecture balnéaire à l'arrière-plan, phare et bord de mer à l'arrière-plan.



Figure 1 : *Vues de Biarritz et de Bayonne*, Extrait du plan n°13



Figure 2 : *Vues de Biarritz et de Bayonne*, Extrait du plan n°14



Figure 3 – *Vues de Biarritz et de Bayonne*, Extrait du plan n°15

Appliquant un parcours interprétatif inductif, il convient de s'intéresser successivement à chacun des passages de manière séparée (approche micro-sémantique), puis d'interpréter la réunion des passages en formalisant les genres et d'éventuelles isotopies (approche méso-sémantique) avant enfin de confronter le corpus à l'intertexte : les métadonnées associées au document (approche macro-sémantique).

Localement plusieurs traits sémantiques sont identifiables au sein du premier passage (plan n°13) : la présence d'une architecture spécifique, de fait pouvant être localisée. Le second passage (plan n°14) met à l'image des personnes vêtues avec distinction : l'étude du vêtement permet de positionner la localisation générale de la captation (Europe de l'Ouest) ainsi qu'une datation approximative du document par le biais des usages et modes vestimentaires. Enfin, le troisième passage (plan n°15) met à l'image un paysage à la fois naturel (plage) et humain (foule de baigneurs) suggérant un emplacement relatif de la captation (bord de mer, côte Atlantique) et une datation (postérieure aux années 1940 par l'étude du matériel, des installations balnéaires et du vêtement de plage).

De manière secondaire, l'interprétation croisée des trois passages permet de mettre en lumière différents genres - « le genre se définit par un type d'interaction entre composantes au sein de deux plans du contenu et de l'expression, ainsi qu'entre ces deux plans » (Rastier, 2008, p. 5) - auxquels se rapproche le document. Le grain de l'image suggère le genre du document patrimonial tandis que la technique audiovisuelle identifiée ainsi que la présence d'une bande altérant l'image suggèrent le genre cinématographique du film amateur dans lequel s'inscrirait le document présent. Le choix des images et la technique contemplative utilisée suggère également un genre documentaire qui néanmoins semble neutralisé par l'absence de bande sonore. La présence d'enfants à l'image suggère l'inscription dans le film de famille, qui peut être nuancée par l'absence de regards caméra, typique du registre familial. Enfin, plusieurs similarités dans les traits sémantiques suggèrent la présence d'isotopies sémantiques - « effet de la récurrence syntagmatique d'un même sème » (Rastier, 2009, p. 276). L'architecture urbaine balnéaire est récurrente au sein des passages et incarne un trait central du corpus tandis que le champ matériel peuplant l'image (voitures, objets du quotidien, matériel de plage) inscrivent la captation dans une temporalité spécifique (les années 1940-1950). Enfin, le vêtement - de manière évidente un élément commun aux représentations humaines - incarne un moyen efficace de datation de l'archive, il permet d'inscrire les trois passages dans une même temporalité.

Dans un troisième temps, une approche macro-sémantique permet d'envisager le document par le prisme de l'intertexte : la conjonction des passages et des métadonnées associées au document.

« [Image amateur]. Balade dans Biarritz : le front de mer, la ville, la plage, le centre ville où des personnes distinguées se promènent. A Bayonne, une procession religieuse marche dans les rues de la ville. Vue de l'église Saint Michel au loin filmée depuis l'actuel Parc des Sports Saint Michel. Ambiance festive dans un quartier où des couples dansent dans la rue au rythme d'un air joué par un groupe de musiciens (accordéon, batterie et guitare) » (INA, 2015).

Précisions que l'intertexte est ici « un corpus qui entretient des relations fonctionnelles avec un texte déterminé » (Rastier, 2008, p. 11). La confrontation des passages avec ce texte additionnel permet d'actualiser - « opération interprétative permettant d'identifier un sème en contexte » (Rastier, 1987, p. 275) - plusieurs traits sémantiques. Tandis que le statut amateur des images est renforcé, la localisation précise des passages étudiés est précisée - Bayonne, sa plage et son centre-ville - de même que la datation de la captation est indiquée : 1950. En rapprochant le corpus avec le texte supplémentaire, on obtient la vérification d'un certain nombre d'interprétations stéréotypiques à l'instar de l'énoncé : la technique est moyenne et le scénario absent, il s'agit d'un film amateur.

L'accumulation de parcours interprétatifs à plusieurs échelles permet de déconstruire ces stéréotypes qui accompagnent le visionnage de documents audiovisuels. De même que l'on s'attend à un certain type de contenus en allumant la télévision - la promesse de genre de François Jost (Jost, 1997) - le visionnage de contenus patrimoniaux fait également l'objet de raccourcis interprétatifs, qu'il s'agit de déconstruire afin d'en tester la véracité.

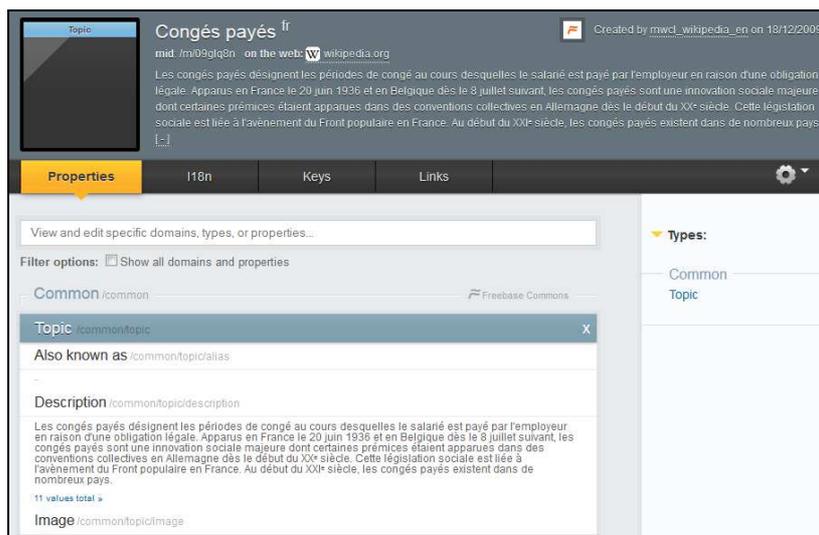
### Bilan du terrain

Ce cas d'étude nous permet de dresser une conclusion préliminaire sur les parcours interprétatifs typiques du visionnage d'un document patrimonial audiovisuel. Après l'étude des différents traits sémantiques identifiés lors des parcours interprétatifs associés aux films, deux catégories d'effets de sens apparaissent : les informations contextuelles et les « marqueurs de proximité historique ».

Intuitivement, en observant une image non familière, plusieurs questions émergent et nous tentons d'y répondre en interprétant l'image. Le spectateur tente intuitivement de comprendre le contexte du document visionné : où se passe l'action - Bayonne dans l'exemple -, qui sont les protagonistes et quel type d'action se déroule-t-il, entre autres. Ces informations contextuelles nécessaires pour l'interprétation forment un premier ensemble d'informations à intégrer dans le dispositif annotationnel. Proposer différentes catégories d'annotations peut constituer une manière de retranscrire les multiples aspects de l'information contextuelle pouvant s'incarner à la fois dans un lieu de tournage, dans un personnage apparaissant à l'écran, un objet, une activité comme le jeu de plage mais également d'une catégorie plus théorique - « sujet » - illustrant des concepts immatériels comme « vacances ». L'examen des propriétés associées à chaque notice du *Linked Open Data Freebase* mentionne un certain nombre de propriétés pouvant se rapprocher des catégories générales d'annotations construites. Pensé comme un silo généraliste, *Freebase* incorpore des notices relatives à un nombre important de domaines, allant de la notice géographique - Bayonne (figure n°4) - à la notice biographique en passant par des notices historiques et théoriques comme « les congés payés » (figure n°5). En conclusion, cet écosystème informationnel semble vertueux pour le cadre expérimental présent, les différentes notices de *Freebase* sont compatibles avec le premier type d'attentes annotationnelles identifié.

The screenshot shows a Freebase notice for 'Bayonne'. At the top, there is a small image of a coastal town. Below it, the title 'Bayonne' is followed by a brief description in French: 'Bayonne est une commune du sud-ouest de la France et est l'une des deux sous-préfectures du département des Pyrénées-Atlantiques...'. The notice is structured with a 'Properties' tab selected, showing a list of domain-specific properties like 'Location', 'Ville/Village', 'Lieu', 'Région statistique', 'Date du lieu', 'Periodicals', and 'Newspaper circulation area'. There is also a 'Types' section on the right. The interface includes search filters, a 'View and edit specific domains, types, or properties...' field, and a 'Filter options' section with a checkbox for 'Show all domains and properties'. At the bottom, there are links for 'Terms of Service', 'How to Attribute to Freebase', and 'Feedback', along with a copyright notice for 2015 Google.

Figure 4 - Notice *Freebase* en français sur Bayonne

Figure 5 – Notice *Freebase* en français sur les congés payés

Derrière les informations contextuelles nécessaires à la compréhension générale du film, l'étude sémiotique met en avant une autre catégorie d'effets de sens, que l'on pourrait qualifier de « marqueurs de proximité historique ». Certaines évidences comme les stéréotypes, les vêtements ou les coiffures ont été intériorisées par le spectateur contemporain au film ancien ; aujourd'hui ces évidences ne le sont plus à l'instar du costume bourgeois des années 1950 présent dans l'exemple. L'image est par nature stéréotypée : en regardant ce type de film ancien, on s'attend à être confronté à tel type d'image, à tel stéréotype. Au contraire, ces stéréotypes sont contrecarrés par l'entreprise interprétative.

Intuitivement, les différents éléments forment autant de marqueurs sémiotiques qui vont permettre à la personne de dater l'archive qu'elle visionne. Ces effets de sens se rapprochent d'une conception intime ou empathique de notre rapport au document ancien : l'interprétation s'incarne dans notre éloignement historique par rapport aux éléments à l'image. Ce décalage permet en réalité de renouer un continuum : le spectateur tend à se mettre à la place de la personne qui filme dans le cas d'un document audiovisuel. Il s'agit d'un point de vue éditorial de retranscrire cette catégorie d'effets de sens au sein du dispositif éditorial.

### Implémenter les effets de sens dans un dispositif éditorial

Reprenant les éléments mis en lumière dans l'étude sémiotique, la seconde expérimentation consiste en la conception d'un artefact technologique éditorial destiné à la recontextualisation de l'archive. L'étude du corpus a fait émerger plusieurs éléments éditoriaux qu'il convient de considérer. Dans un premier temps, la nécessité d'une catégorisation des annotations semble nécessaire, elle est censée offrir à l'utilisateur une granularité d'inscriptions, propice pour la description d'évènements sensibles ou difficiles à appréhender dans l'environnement patrimonial. Les champs annotationnels « lieu », « personne », « objet », « activité » et « sujet » sont ainsi considérés et semblent couvrir la dimension descriptive encyclopédique identifiée. Dans un second temps, l'entrée dans l'archive s'effectue également par la construction d'un rapport sensible au passé. Cette dimension observée lors de l'étude sémiotique n'est pas appréhendée par le dispositif d'annotations typées. Il convient dès lors d'intégrer cette recommandation éditoriale dans le cahier des charges fonctionnel du dispositif technique : assister l'utilisateur dans la formalisation de son rapport au passé à travers un design et une expérience utilisateur spécifiques.

De fait, la conjonction au sein de l'artefact de deux types de fonctionnalités éditoriales semble nécessaire : les annotations typées permettant la convocation des éléments contextuels et les éléments graphiques ou fonctionnels favorisant le rapport au passé. La conception de l'artefact technologique à vocation patrimoniale est réalisée dans un cadre précis incarné par une convention CIFRE entre l'Université de technologie de Compiègne et le partenaire industriel PERFECT MEMORY SA, start-up en ingénierie des connaissances. La première phase expérimentale de cette étude a pour objectif de faire émerger des recommandations fonctionnelles et éditoriales en plus du bilan sémiotique associé à l'intelligibilité de l'archive. Celles-ci sont transmises à plusieurs collaborateurs - un concepteur d'interface graphique, un architecte de plateforme et un développeur d'interface Web. Cette collaboration qui positionne l'expérimentation sémiotique réalisée au cœur d'un projet technologique est pensée comme une co-construction.

Au niveau de l'ancrage théorique revendiqué par cette phase expérimentale, les conclusions de Jack Goody - le dispositif technique de l'écriture a fait émerger chez l'homme de nouvelles structures conceptuelles telles que la liste, la formule et le tableau - nourrissent la réflexion. Poursuivant le constat fondateur d'une raison graphique (Goody, 1979), Bruno Bachimont met en avant une raison computationnelle, cristallisant les apports cognitifs et phénoménologiques du numérique à la connaissance (Bachimont, 2000). Trois structures fondamentales de pensée émergent de cette interrogation : le programme, le réseau et la couche. La notion de réseau permet de dépasser l'espace d'inscription historique de la tablette d'argile afin de penser et d'appréhender des dispositifs pluri-sémiotiques à l'instar du dispositif construit intégrant dans le même plan un document, des inscriptions et des « petites formes » (Candel, 2012).

De fait, dans le cadre de la conception de ce dispositif d'organisation des connaissances, les « mondes possibles » éditoriaux semblent larges et à la fois en partie cartographiés à l'instar du logiciel Ligne de temps conçu par l'Institut de recherche et d'innovation (cf. figure n°6). Celui-ci met en avant la représentation graphique spécifique de la *timeline* (ligne de temps), permettant d'éditorialiser le film numérique par le biais d'une expérience temporelle de l'objet. D'ores et déjà, la fonctionnalité d'une annotation temporelle offerte à l'utilisateur pour l'annotation des documents temporels est expérimentée. L'articulation entre un espace d'annotation comme la *timeline* avec les différentes catégories d'annotations émergées par l'étude sémiotique est un dispositif éditorial fonctionnel (cf. figure n°7). La coexistence de cet espace d'inscription incarné par la *timeline* avec l'espace de visionnage du film semble primordiale, l'acte d'annoter le film est une activité conjointe au visionnage, l'utilisateur documente le film en fonction de son interprétation en même temps qu'il le visionne.

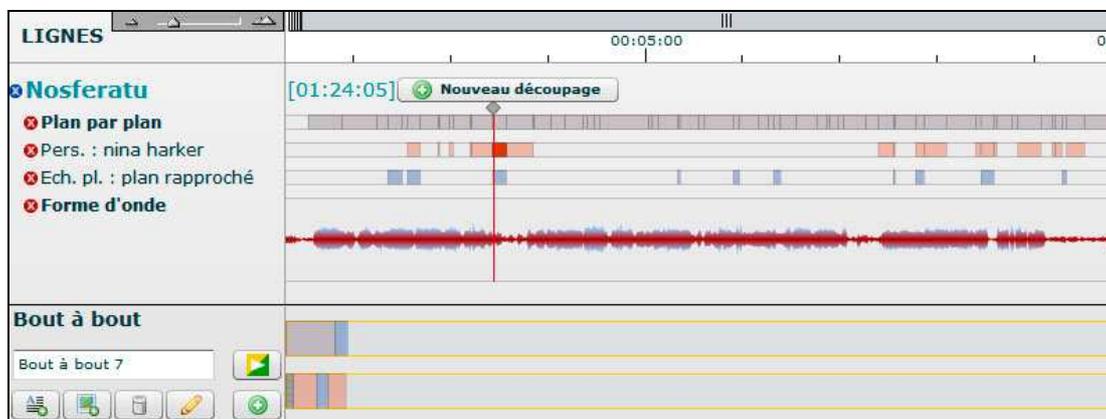


Figure 6 - Capture d'écran du logiciel *Lignes de temps* (Crédits iri.centrepompidou.fr, 2015)

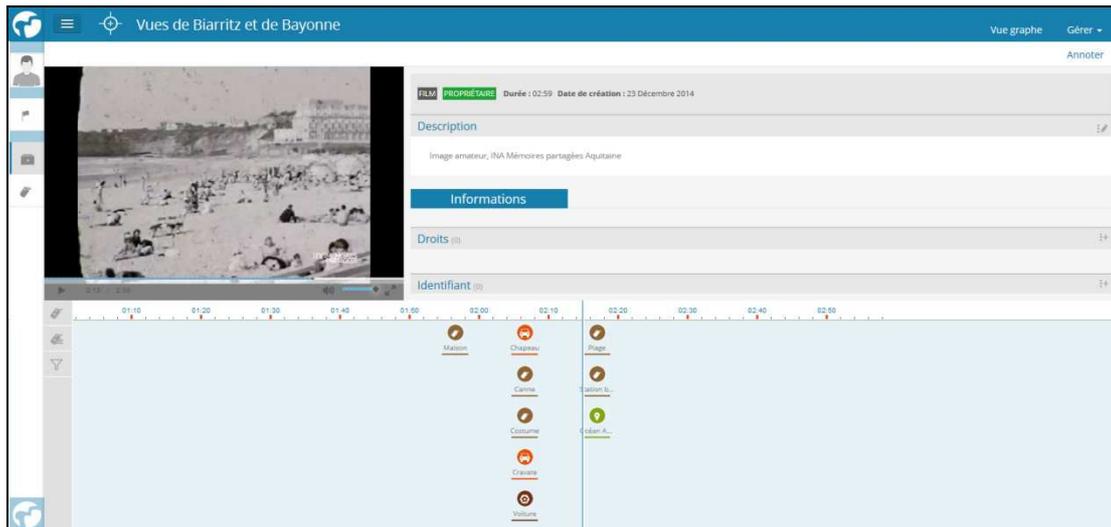


Figure 7 – Capture d'écran d'une « vue *timeline* » de l'artefact construit (Crédits Perfect Memory, 2015)

D'un point de vue documentaire, l'artefact en cours de construction est fonctionnel, l'archive peut être recontextualisée par le truchement des notices des *Linked Open Data*. Par contre d'un point de vue sémiotique, aucune certitude à propos de la pertinence de cette redocumentarisation quant à l'intelligibilité de l'archive ne peut être formulée. L'environnement de signes ainsi construit peut se révéler hostile quant à une démarche interprétative du document patrimonial. Il convient de préciser que la question de l'usage, large volet de toute construction d'un dispositif, est délibérément positionnée à l'issue de cette recherche, afin de laisser une plus grande marge de manœuvre à l'appréhension sémiotique du phénomène d'intelligibilité par le prisme de la sémantique différentielle.

### Conclusion : mode d'existence de l'artefact construit

Le Web des données semble une source d'informations contextuelles efficace pour le processus de recontextualisation qui est au cœur de la présente étude. Les caractéristiques « ouvertes » et « liées » des données favorisent leur réutilisation dans le cadre de dispositif tiers comme celui construit ici-même. Nous rappelons que nous avons fait le choix de ne discuter ici que d'un seul film du corpus. Les résultats obtenus pour un second film, avec les trois paliers d'interprétation, permet de construire et d'affiner un modèle théorique de recontextualisation de l'archive. Néanmoins, ce projet dépasse le cadre de cet article où nous nous limitons à formuler le processus d'une possible reconstruction d'une intelligibilité de l'archive sans chercher à établir un modèle global.

Si la question de la perte de l'intelligibilité associée à l'archive, la méthode sémiotique, l'outil (les *Linked Open Data*) et le livrable (un artefact technologique annotationnel) semblent cohérents, une autre question demeure saillante : l'usage. Ce dispositif d'organisation des connaissances dans un cadre patrimonial mérite d'être interrogé quant à sa dimension prescriptive : celui-ci va-t-il avoir une influence sur les usages professionnels d'un public averti tels que les documentalistes ? Le grand public en tant que communauté alternative visée par le dispositif pourra-t-il et devra-t-il se réapproprier l'artefact avant de se réapproprier son patrimoine culturel numérisé ? Comment positionner un « curseur éditorial » au bon endroit entre des fonctionnalités et une expérience utilisateur souples ou plus rigides ? Les enjeux ouverts via l'expérience du dispositif annotationnel construit méritent en effet d'être formalisés bien qu'ils débordent de la présente problématique

sémiotique et documentaire de reconstruction d'une intelligibilité de l'archive par la mobilisation du Web des données. Le choix méthodologique du positionnement d'un « témoin standard », faisant écho au « lecteur modèle » (Eco, 1985), nous amène à considérer l'intelligibilité du document comme un phénomène dépassant l'« encyclopédie » (Eco, 1985) de chacun. Si la dimension individuelle n'est pas appréhendée dans le projet – étude du corpus puis construction du dispositif – elle constitue à terme une démarche supplémentaire heuristique d'interrogation du dispositif en particulier selon les modalités de cette recherche amenant à co-concevoir un outil de réappropriation patrimoniale pour le grand public. Un échantillon de bêta-testeurs viendra interroger le dispositif tandis que les comportements ainsi observés fourniront une matière scientifique relative au problème de la dimension prescriptive du dispositif patrimonial et au statut des annotations produites.

## Références bibliographiques

Bachimont, Bruno (2000), « L'intelligence artificielle comme écriture dynamique : de la raison graphique à la raison computationnelle » (p. 290-319), in Petitot, Jean et al. (coord.), *Au nom du sens*, Paris : Grasset.

Bachimont, Bruno (2010), « La présence de l'archive : réinventer et justifier », *Intellectica*, n°53-54, p. 281-309.

Berners-lee, Tim (2006), *Linked Data*, W3C, [en ligne] <http://www.w3.org/DesignIssues/LinkedData.html>

Briet, Suzanne, (1951), *Qu'est ce que la documentation ?*, Paris : Editions documentaires industrielles et techniques EDIT.

Candel, Etienne *et al.* (2012), « Petites formes, grands desseins. D'une grammaire des énoncés éditoriaux à la standardisation des écritures » (p. 135-166), in Davallon, Jean (coord.) (2012), *L'économie des écritures sur le web*, Paris : Hermès-Lavoisier, coll. ingénierie représentationnelle et construction de sens.

Chignard, Simon (2012), *L'open data : comprendre l'ouverture des données publiques*, Limoges : Fyp Editions, coll. Entreprendre.

Cyganiak, Richard et al. (2014), *The Linking Open Data cloud diagram*, [en ligne] <http://lod-cloud.net/>

Eco, Umberto (1985), *Lector in fabula*, Paris : Grasset.

Gandon, Fabien et al. (2011), *Le Web sémantique : comment lier les données et les schémas sur le Web ?*, Paris : Dunod, coll. InfoPro.

Goody, Jack (1979), *La raison graphique, la domestication de la pensée sauvage*, Paris : Éditions de Minuit, coll. Le sens commun.

Heath, Tom et al. (2009), "Linked Data - The Story so far", in *International Journal on Semantic Web and Information Systems*, Special Issue on Linked Data.

Institut national de l'audiovisuel (2015), *Vues de Biarritz et de Bayonne*, [en ligne] <http://www.ina.fr/video/AMX13000074/vues-de-biarritz-et-de-bayonne-video.html>

Jost, François (1997), « La promesse des genres », *Réseaux*, vol. 15 n°81, p. 11-31.

Mattelart, Armand (1994), *L'invention de la communication*, Paris : Éditions La Découverte, coll. La Découverte Poche.

Monnin, Alexandre (2013), *Vers une Philosophie du Web. Le Web comme devenir-artefact de la philosophie (entre URIs, Tags, Ontologie(s) et Ressources)*, thèse de doctorat, Université Paris 1 Panthéon-Sorbonne.

Pédaque, R.T. (2007), *La redocumentarisation du monde*, Paris : Cepaduès.

Rastier, François (2009), *Sémantique interprétative*, Paris : PUF, 3e éd. (collection « formes sémiotiques »).

Rastier, François (2011), « Objets et performances sémiotiques : l'objectivation critique dans les sciences de la culture », *Texto!*, vol. XVI n°1, p. 1-31.

Rastier, François (2008), « Passages et parcours dans l'intertexte », *Texto!*, vol. XIII n°4, p. 1-20

Rastier, François et al. (2001), *Sémantique pour l'analyse*, Paris : Masson (collection « sciences cognitives »).

Rastier, François (1987) [Référence manquante. François Rastier, *Sémantique interprétative*, Paris, PUF, 1987 ? n.d.l.r.]

Théodule, Marie-Laure (2007), "Tim Berners-Lee : « Le Web va changer de dimension »", *La Recherche*, n°413 novembre 2007, [en ligne] <http://www.larecherche.fr/savoirs/dossier/2-tim-berners-lee-web-va-changerde-dimension-01-11-2007-75889>

Thibodeau, Ted Jr (2015), "The DBpedia Data Set (2014)", *DBpedia*, [en ligne] <http://wiki.dbpedia.org/Datasets>

Treleani, Matteo (2013), « Recontextualisation : ce que les médias numériques font aux documents audiovisuels », *Réseaux*, vol. 31, n° 177, p. 233-258.

Yates, Frances A. (1975), *L'art de la mémoire*, Paris : Gallimard (collection « Bibliothèque des Histoires »).