



**creative  
commons**



CC - BY - NC - ND

**Ce mémoire est mis à disposition sous un contrat Creative Commons  
"Paternité-Pas d'Utilisation Commerciale-Pas de Modification 2.0 France"**

**Les détails de ce contrat sont disponibles à l'adresse suivante :**

**<http://creativecommons.org/licenses/by-nc-nd/2.0/fr/>**

<b>Synthèse .....</b>	<b>1</b>
<b>Introduction .....</b>	<b>4</b>
<b>1. Contexte de la création multimédia et surgissement du web sémantique 6</b>	
<b>1.1. Evolution des contenus multimédia.....</b>	<b>6</b>
1.1.1. Diversification des contenus multimédia.....	6
1.1.2. Evolution de la création de contenu multimédia.....	9
1.1.3. Evolution de la consommation de contenu multimédia.....	10
<b>1.2. L’annotation et l’indexation : pré-requis à la recherche .....</b>	<b>12</b>
1.2.1. Les méthodes actuelles d’indexation du contenu multimédia.....	12
1.2.2. Le concept de folksonomie.....	13
<b>1.3. Le web sémantique : Une autre vision du contenu .....</b>	<b>16</b>
1.3.1. Le web sémantique, un concept historique.....	16
1.3.2. Définition de la couche basse du Web sémantique.....	17
1.3.3. Définition des couches élevés du Web sémantique .....	18
1.3.4. Avantage du web sémantique sur une l’annotation actuelle du contenu.....	19
1.3.5. Les problématiques et faiblesses du Web sémantique.....	20
<b>2. Une tendance actuelle vers l’enrichissement par les métadonnées .....</b>	<b>22</b>
<b>2.1. Le Web sémantique remis au gout du jour.....</b>	<b>22</b>
2.1.1. Apparition d’outils d’annotation sémantique par les acteurs du web. ....	22
2.1.2. Comparaison des outils d’annotations des contenus actuels .....	23
2.1.3. Fragmentation de la lecture des contenus multimédia sur internet. ....	25
2.1.4. Une fragmentation qui favorise l’annotation.....	26
<b>2.2. Sémantiser et annoter les contenu multimédia.....</b>	<b>27</b>
2.2.1. « Ligne de temps » : outil dédié à l’annotation sémantique. ....	27
2.2.2. Polémic Tweet : utiliser les réseaux sociaux pour annoter .....	29
2.2.3. Synote et Méta-data Player : outils de jointure entre le contenu et l’annotation.....	31
<b>2.3. Définir les usages de ces méta données .....</b>	<b>34</b>
2.3.1. Favoriser l’interopérabilité des contenus .....	34
2.3.2. Vers un Mash-up temporel des contenus multimédia ?.....	35
2.3.3. Quelles informations pour ces annotations ? .....	37
<b>3. Vers une indexation socio-sémantique du Web.....</b>	<b>38</b>
<b>3.1. Passer du Web 2.0 au Web 3.0.....</b>	<b>38</b>
3.1.1. Avantage du web sémantique sur l’indexation classique .....	38
3.1.2. Capter et transformer les données du web 2.0.....	39
3.1.3. Faciliter l’ajout de ces tags.....	41
<b>3.2. Construction d’une page à contenu multimédia enrichis.....</b>	<b>42</b>
3.2.1. L’indexation sémantique favorise la sérendipité.....	42
3.2.2. Intégration de tags sémantiques dans la fragmentation des contenus .....	43
3.2.3. Présentation du lecteur multimédia enrichi.....	45
3.2.4. Objectif du lecteur multimédia sémantique.....	46
<b>3.3. Applications et réserves .....</b>	<b>46</b>
3.3.1. Les Obstacles à surmonter .....	47
3.3.2. Intégration de ces modules dans les TV connectées ? .....	47
<b>Conclusion .....</b>	<b>49</b>
<b>Bibliographie.....</b>	<b>i</b>
<b>Annexes.....</b>	<b>v</b>



## Synthèse

En mai 2012 Google lançait une brique particulièrement inattendue et intéressante le « Google Knowledge Graph »<sup>1</sup>, Ce module présent dans les listes de résultats de Google n'est pas un tri par pertinence comme c'était le cas jusqu'à maintenant, il s'agit d'une recherche d'un nouveau type : la recherche sémantique. Cette technologie n'est pas étrangère n'est pas étrangère à la firme de Mountain View qui avait déjà acheté un moteur de recherche sémantique – Freebase – en 2010. Pourtant un pas à été franchi puisque cette fois-ci ce n'est plus seulement une amélioration de l'algorithme qui est en jeu mais bien un changement de stratégie. Les moteurs de recherches et Google particulièrement avaient fait de timides approches sur ce contenus (2007 : intégration des vidéos dans le protocole Sitemap, 2011 : création du consortium schema.org avec Bing et Yahoo!) mais il semblerait bien que cette fois-ci, la machine sémantique se soit mise en route.

Parallèlement à cela, les contenus multimédias, jusqu'alors assez marginaux sur le Web se sont mis à croître, à évoluer et à occuper une part de plus en plus importante dans notre façon de naviguer sur le web. YouTube est aujourd'hui le troisième site le plus consulté au monde après Facebook et Google. Pourtant, ce site racheté par le moteur de recherche le plus puissant du monde peine à indexer ses contenus et à retrouver des résultats pertinents pour ses utilisateurs. Une étude sur les effets d'étranglement des plateformes de vidéos (Cha, et al. 2007) montrait à ce sujet que cela avait un impact fort sur les visionnages de la plateforme. Au lieu de se retrouver dans une loi de puissance classique qui aurait pu présager de la présence d'une longue traîne, la plus grosse base de données du monde souffrait d'une troncature de la partie la plus éloignée de la courbe de visionnage pouvant s'expliquer par l'incapacité de la plateforme à exploiter le contenu et les effets de niches de ces vidéos isolées.

Le contenu vidéo à en effet la particularité de ne pouvoir être exploité, lu, ou parcouru par les acteurs classique du web : Les robots<sup>2</sup>. Ces contenus sont à la fois complexe (encodage, son et image non compréhensible) riche en informations (format, taille de la vidéo, auteurs, acteurs, lieux, thèmes etc.) et semblent symboliser l'aporie de l'indexation et de la recherche classique tant il est difficile de les exploiter. Le web sémantique qui avait dès ses débuts présagé ces difficultés dans l'indexation des contenus multimédias semble représenter une piste de déblocage de l'indexation du contenu

---

<sup>1</sup> <http://www.youtube.com/watch?v=mmQl6VGvX>

<sup>2</sup> Nous entendons par robots les logiciels développés par les moteurs de recherche pour parcourir les pages du web et ainsi indexer les contenus existants.

multimédia. Parce que sa puissance est justement dans sa capacité à décrire un contenu quelque soit l'objet en question.

Nous assistons d'ailleurs au développement d'outils d'annotations sémantique de ces contenus comme Schéma.org, Facebook Open Graph, ainsi qu'à de nouvelles technologies visant à éclaircir ce rapport difficile du web aux contenus riches : les médiafragments. Les premiers outils ajoutent de la finesse dans la description des contenus, le second rajoute enfin une véritable dimension temporelle aux musiques et aux vidéos appréhendables cette fois-ci par les robots.

Ces différentes recherches ont conduit à une nouvelle génération de lecteurs multimédia non plus seulement capable de lire les vidéos mais aussi d'y rajouter des axes de lectures basés sur le déroulement de la vidéo. Ces solutions permettent de joindre la temporalité du contenu à une annotation sémantique créée par les utilisateurs. De cette manière, le web fait ainsi sa transition du Web 2.0 guidé par les communautés (O'reilly 2005) vers le web 3.0 qui y rajoute cette vision de liaison entre les contenus par la force de la communauté. Des chercheurs comme (Passant, Technologies du Web Sémantique pour l'Entreprise 2.0 2009) ont ainsi cherché à démontrer qu'il était possible de passer de l'un à l'autre, nous nous attacherons à montrer que ce passage est aussi possible sur ces contenus multimédia.

C'est dans ce contexte que nous sommes donc amenés à réfléchir à l'alternative suivante : Quel intérêt présente le web sémantique pour l'indexation des contenus multimédia par rapport aux technologies existantes d'annotation et d'indexation.

Maintenant que la preuve technique qu'une utilisation massive de ces données est possible, il faut définir si ce type d'information est capable de résoudre les problématiques que pose le contenu multimédia. Nous montrerons qu'il s'agit d'abord d'améliorer l'expérience d'utilisation des plateformes de visionnage pour les internautes par une meilleure suggestion de contenu. Le système présente aussi un intérêt pour les WebTV car ce système ajoute une dimension interactive jusqu'ici assez peu présente dans les solutions proposées. Le web sémantique permet aussi un point particulièrement intéressant sur la promotion de contenus non connus : la navigation sérendipitaire : Grâce à l'ajout de ces liaisons et parce qu'il favorise une meilleure indexation, le web sémantique favorisera une navigation à l'aveugle entre les contenus plus pertinents.

Mais cette meilleure suggestion de contenu est aussi capable de diminuer les effets de niches que l'indexation classique n'avait pas dépassés et permettra donc une augmentation pour l'éditeur des vues de l'ensemble des vidéos sur son site.

## Synthèse

Toutefois cette nouvelle vision du contenu n'est pas non plus exempte de critique et possède elle aussi des difficultés intrinsèques qui demanderont à être traité si l'on veut pouvoir proposer un système pertinent de traitement des vidéos. On peut ainsi parler des difficultés à définir de véritable couche supérieures d'organisation du contenu pour l'indexation. L'indexation ne sera donc pas par des moteurs de recherche sémantiques mais par les moteurs de recherche traditionnels qui auront ajoutés une brique sémantique à leur solution.

Nous verrons aussi que l'accès à l'annotation est un point particulièrement sensible, qu'il n'y a pas d'indexation sans annotation et que celui-ci doit donc être la prochaine étape sur lequel réfléchir si l'on veut créer des liaisons pertinentes entre les contenus.

## Introduction

Depuis le milieu des années 2000, une nouvelle tendance se dessine autour des contenus multimédias audio et vidéo. Avec la montée en puissance du stockage de masse, la création et la consommation de ces données jusqu'alors inabordables<sup>3</sup> sont devenues accessibles à tous.

Cet accès à un nouveau contenu sur internet a très vite poussé les grands acteurs du marché à s'implanter durablement dans le monde du web multimédia. En 2005, Yahoo! rachetait Flickr pour suivre Google et son service Picasa. Google, lui, rachetait YouTube en 2006 pour 1.5 milliard de dollars. Enfin, Facebook, s'est illustré l'année dernière en prenant le contrôle de la start-up Instagram pour 1 milliard de dollars lui laissant prendre un peu plus d'avance en tant que importante base de données photographique sur internet<sup>4</sup>.

Tous ces grands investissements concédés par ces acteurs majeurs du web actuel illustrent à quel point la maîtrise des contenus multimédia est devenue cruciale. Face à cet engouement la problématique de l'indexation du contenu se pose : il est de plus en plus difficile pour les internautes de trouver un contenu pertinent tant il est noyé dans cette masse de contenus multimédias à perte de vue.

Jusqu'alors le contenu du web était principalement verbeux et compréhensible par les robots des moteurs de recherche et, même avec cela, seul 5% du web était indexé<sup>5</sup>. Or à la différence de ce web historique, ces contenus multimédias ne sont lisibles que par les humains car les robots n'analysent que le texte. Pour palier à cet état de fait, ces contenus avaient été initialement enrichis par des informations réservées aux robots et invisibles pour les utilisateurs - les attributs de balises images `alt`<sup>6</sup> - mais ces informations n'ont permis qu'une indexation pauvre et partielle des contenus.

Le web 2.0, le web des réseaux et des communautés a tenté de résoudre ce problème par la participation des utilisateurs à l'indexation de ces contenus. Il a été dès lors possible de qualifier ces contenus par des mots-clés signifiants : les tags. Pour affiner les résultats, ces tags ont ensuite été agglomérés par ces utilisateurs pour créer des folksonomies, système de classement collaboratif basé

---

<sup>3</sup> Le prix d'un Mégaoctet de données est passé de 6.25\$ en 2000 à 0.035\$ en 2011 (Vasilev 2011)

<sup>4</sup> Dès 2010, Facebook avait pris la tête en terme de contenu exporté vers ses serveurs, ils totalisaient alors une base de donnée de 15 Milliard de photos. ([http://articles.timesofindia.indiatimes.com/2010-07-09/internet/28297234\\_1\\_facebook-active-users-user-base](http://articles.timesofindia.indiatimes.com/2010-07-09/internet/28297234_1_facebook-active-users-user-base) 2010)

<sup>5</sup> On parle ainsi du Web profond (Deep Web) qui n'est pas accessible aux moteurs de recherche. (Kepeklian et Lequeux 2008)

<sup>6</sup> Il s'agit là d'une volonté du W3C sur la mise en place de normes d'accessibilité, la balise `alt` permet ainsi aux déficients visuels d'appréhender le contenu de la page. Les robots des moteurs de recherches les utilisent eux aussi pour appréhender les images.



sur la répétition de mêmes tags sur un contenu. Toutefois cette nouvelle solution ne semblait pas non plus totalement satisfaisante car le rangement ainsi créé restait chaotique.

La question de l'indexation de ces contenus multimédias est donc toujours posée alors qu'elle prend une résonance de plus en plus forte<sup>7</sup>. C'est dans ce contexte qu'une ancienne solution d'indexation resurgit, le web sémantique.

Tim Berners-Lee, fondateur du web tel que nous le connaissons aujourd'hui<sup>8</sup>, se posait déjà la question du traitement de ces données non textuelles. Il avait en particulier développé l'idée d'un web sémantique, lisible et compréhensible par les machines. Si l'époque avait alors mis l'idée de côté pour se concentrer sur d'autres priorités<sup>9</sup>, ce concept refait surface et semble être aujourd'hui une piste sérieuse pour une meilleure compréhension de ces contenus riches. Contrairement aux métadonnées utilisées actuellement, le web sémantique qualifie ces informations et les lie entre elles<sup>10</sup>. Un contenu ainsi enrichi permettra à un moteur de recherche de comprendre la requête de l'internaute et d'y rapprocher une information pertinente c'est-à-dire proche du concept sur lequel l'internaute posait une question.

Dans ce contexte actuel nous sommes donc amenés à réfléchir à l'alternative suivante : Quel intérêt présente le web sémantique pour l'indexation des contenus multimédia par rapport aux technologies existantes d'annotation et d'indexation.

Nous reviendrons tout d'abord sur les conditions d'apparition de ce contenu et les efforts qui ont été fait pour l'indexer.

Nous montrerons ensuite que des solutions d'annotations par les métadonnées apparaissent et forme une tendance intéressante pour le Web sémantique.

Enfin nous montrerons que cet enrichissement par les métadonnées n'est qu'une première étape, que des solutions sémantiques plus souples et ouvertes à la communauté devraient permettre le développement d'un véritable vecteur de recherche et d'indexation par le Web sémantique des contenus multimédia.

---

<sup>7</sup> A ce jour, les utilisateurs de YouTube téléchargent vers la plateforme une heure de vidéo par seconde (upload) (<http://www.onehourpersecond.com> 2012)

<sup>8</sup> Ce sont ses travaux au CERN qui ont créé la première base du World Wide Web (W3). Il a en particulier créé la première version du langage HTML (HyperText Mark-up Language) en 1989 et le protocole d'échange associé HTTP (HyperText Transfer Protocol). Il est actuellement à la tête du consortium W3C chargé de l'évolution des langages utilisés sur le web. (Berners-Lee, Tags used in HTML 1992)

<sup>9</sup> La question de la standardisation des langages du Web à en particulier pris beaucoup de temps aux groupes de travail du W3C et de l'IETF (Internet Engineering Task Force) (Harold 2004)

<sup>10</sup> L'objectif à terme d'un système sémantique est de regrouper et de hiérarchiser l'ensemble des liens entre les concepts pour construire une ontologie.

## 1. Contexte de la création multimédia et surgissement du web sémantique

Si les contenus ont fait une apparition tardive dans le monde de l'internet, la création et la consommation de ces contenus à connu un boom dès le début des années 2000. Depuis ce moment, nous assistons à une explosion de ce contenu qui semble ne pas avoir de limite.

### 1.1. Evolution des contenus multimédia

Les raisons de cette brusque évolution sont d'ailleurs des plus intéressantes car elles permettent de mieux appréhender la situation actuelle et les problématiques que nous devons adresser.

#### 1.1.1. Diversification des contenus multimédia

Avant de rentrer plus en détail dans le fonctionnement du web et de ces contenus multimédia, il faut témoigner d'une réalité du Web : La toile n'a jamais été uniquement verbeuse, aride, composée uniquement de texte. Même si les premières pages créées par Tim Berners Lee se focalisaient sur le contenu et son indexation, les écarts vers le multimédia se fera dès 1993. Le premier véritable navigateur, NCSA Mosaic, suivi de près par Netscape, intégreront une balise Html permettant l'affichage d'images en format GIF et XMB.

Il y aura d'ailleurs un schisme entre l'école académique centrée sur le balisage sémantique et les nouveaux acteurs du web comme les navigateurs internet. Lors de la première conférence du World Wide Web, Tim Berners Lee précisait sa volonté de créer un web ordonné et rangé, un outil de collaboration scientifique entre les universités du monde entier. De leur côtés, dans ce web qui se démocratisait et sortait du domaine académique, les nouveaux acteurs du web comme les navigateurs internet<sup>11</sup> poussaient eux une vision plus humaine, plus visuelle et plus ergonomique du web.

Netscape et Mosaic iront jusqu'à bouleverser les standards implémentés par le W3C en créant des balises non standards plus en rapport avec les besoins et les volontés des utilisateurs. La balise Image () sera ajoutée en dehors du cadre des documents de travail fournie par l'IETF et le W3C<sup>12</sup>.

---

<sup>11</sup> Le créateur de Netscape, Marc Andreessen, déclarera d'ailleurs , dès 1993 « Je pense que s'occuper du SGML en général est une complète perte de temps, et que nous en serions aujourd'hui beaucoup plus loin si nous n'étions pas encombrés avec cet héritage SGML que nous continuons à porter. 99,99 % des gens avec qui je parle veulent mettre en ligne des documents riches, veulent contrôler leur apparence, et se contre-fichent totalement du balisage sémantique ou des différences entre la structure et le rendu d'un document »

<sup>12</sup> Cela impliquera d'ailleurs un éclatement du standard HTML caractérisé par les balises spécifiques à chaque navigateur appelé aussi la période de la soupe de balise (Tag soup). La version 2.0 du langage HTML en 1995 sera une tentative de formaliser et d'unifier ces différentes avancées.

Ce premier enrichissement du web par du contenu multimédia est donc concomitant avec son développement et sa démocratisation. Les images seront ainsi la première pierre de l'histoire du web multimédia. Les autres types de contenu ne se feront pas attendre très longtemps : alors le web développe ses fondements et sa logique de liaison hypertexte, un autre domaine connaît une révolution, la numérisation des contenus photo audio et vidéo.

Le disque compact inventé en 1982 avait sonné le glas des supports analogiques pour favoriser le numérique. Le véritable tournant des années 90 se fera sur la compression des informations numériques pour répondre aux contraintes de transfert de l'information<sup>13</sup>. Grâce aux nouveaux formats numériques compressés, les images, les vidéos ou bandes sonores vont pouvoir tenir dans un espace numérique plus restreint. Le format JPEG, utilisé dans le monde de la photographie permettra de diviser le poids des images avec un rapport de 1 à 10. De la même manière le format MPEG-1 Layer 3 plus connu sous l'extension de fichier .mp3 va permettre de compresser de manière importante les fichiers sans que ceux-ci perdent en qualité audible (Di Virgilio 2006).

Autre point important, ces formats de compression ont été réfléchis comme des standards. Le « Moving Picture Expert Group » a ainsi rassemblé les experts de l'audiovisuel et du numérique avec comme objectif de construire des formats de compression performants, utilisés par le plus grand nombre (FUCHS, et al. 2012). Il est alors intéressant de constater que ces normes ont très vite été appliquées par les navigateurs, acteurs les plus importants au début du web. Dès les années 1995, les formats GIF, JPEG, et PNG, seront intégrés en natif et donc directement lisible<sup>14</sup> dans les premiers navigateurs IE4 et Napster 1.01.

Le sort des contenus audio sera plus long et plus complexe car cela demandera de passer par une plateforme tiers : le lecteur multimédia. La démocratisation des fichiers audio va se faire par le biais de la plateforme d'échange Napster qui encodera les fichiers dans le format MP3, Le succès rencontré facilitera l'utilisation du format mais celui restera utilisable seulement dans les logiciels présents sur l'ordinateur de l'internaute.

Un grand changement se produira avec l'arrivée de lecteurs multimédia directement intégrable dans les navigateurs. Ces lecteurs vont ainsi permettre l'accès à ce contenu multimédia sans pour autant que l'internaute télécharge le contenu. Le lecteur appellera directement le serveur qui lui transmettra le fichier multimédia. QuickTime, lancé par Apple en 1999 mais surtout « Adobe Flash

---

<sup>13</sup> Liaison satellite, conférence par VOIP et bien évidemment le web, vont ajouter une contrainte à ces contenus riches pour qu'ils soient plus facilement transportable et que leurs stockage prennent moins de place.

<sup>14</sup> Lorsque qu'un format est déclaré en natif dans un navigateur, cela veut dire que l'on peut ouvrir le type de fichier. A l'inverse certains formats nécessitent un décodage et l'on est alors obligé de passer par un plug-in logiciel.

Player » développé par Adobe vont révolutionner la consommation de ce contenu et lui donner une véritable place sur le Web.

La lecture de ces fichiers audio et vidéo par des plug-ins tiers n'est toujours pas satisfaisante<sup>15</sup>. Des travaux autour du nouveau standard HTML 5 vont ainsi se pencher sur la création d'une nouvelle balise <vidéo> pour permettre aux navigateurs une lecture native des contenus multimédia<sup>16</sup>. Toutefois cette balise est encore jeune, et pose un certain nombre de problèmes par rapport aux lecteurs flash ou Quicktime. YouTube illustre d'ailleurs très bien la transition en cours, puisqu'il propose les deux formats. Pour la plateforme la balise <vidéo> est très prometteuse mais ne répond pas encore à l'ensemble des critères nécessaires pour justifier l'abandon des lecteurs externes. Les arguments qui laissaient l'avantage aux lecteurs sont les suivants<sup>17</sup> :

- Pas de formats standards vidéo lisibles pour l'ensemble des navigateurs :

A la différence d'un lecteur flash qui fonctionne sur tous les navigateurs pourvu que le plug-in flash soit installé, la balise <vidéo> serait dépendante des formats supportés par le navigateur<sup>18</sup>.

- Difficultés à protéger les contenus propriétaires :

Le lecteur flash dispose d'une sécurité forte dans ses appels au serveur, il est ainsi plus difficile d'accéder au contenu en vue de le copier. Cela est particulièrement important lorsqu'il s'agit d'appeler le serveur depuis un site tiers (lors d'encapsulation de vidéo par exemple.)

- Accès aux fonctionnalités annexes :

Anecdotique mais bloquant pour une plateforme comme YouTube, l'HTML 5 ne permet pas encore d'accéder à des fonctionnalités comme l'affichage en plein écran comme le permet le lecteur flash.

Etape importante dans la formalisation des contenus multimédia sur le Net, les formats de compression ont permis au web d'intégrer plus facilement ce contenu. Ces formats ont aussi permis de répondre à ces internautes qui demandaient l'accès à ce contenu. Dès le milieu des années 2000, cette demande se transforme en phénomène qui n'a pas faibli jusqu'à ce jour.

---

<sup>15</sup> Des raisons de propriété intellectuelle et de standardisation sont d'ailleurs le moteur de cette mouvance vers la lecture en natif des vidéos

<sup>16</sup> La balise <vidéo> sera proposée en 2007 par l'ingénieur Anne van (Kesteren 2007) lorsqu'il travaillait pour le navigateur Opéra.

<sup>17</sup> "The <video> tag certainly addresses the basic requirements and is making good progress on meeting others, but the <video> tag does not currently meet all the needs of a site like YouTube" (<http://apiblog.youtube.com/2010/06/flash-and-html5-tag.html> 2010)

<sup>18</sup> C'est d'ailleurs pour cela que les différents navigateurs se sont regroupés en un groupe de travail pour définir un format standard de vidéo libre et gratuit. A ce jour le format le plus complet est le MPEG-4 qui lui est payant mais seulement supporté en version dégradé sur le navigateur Chrome.

### 1.1.2. Evolution de la création de contenu multimédia

La numérisation du contenu photo et vidéo à aussi eu un effet non prévu sur la création de ces contenus. Avant les années 80, la création de contenu était majoritairement tenue par des professionnels du secteur. La difficulté de réalisation, de production et même du seul développement impliquait trop de connaissances matérielles ou techniques pour qu'une personne non professionnelle puisse produire un contenu de qualité par elle-même<sup>19</sup>. Pourtant, en trois décennies, les principales barrières techniques et technologiques ont été dépassées, laissant le champ libre à la création par les particuliers.

Premières victimes de cette mutation de la création de contenu multimédia, les appareils argentiques, jusqu'alors prédominants, vont être mis à mal par de nouveaux arrivants. A partir de 1997 Les appareils photos numériques vont bouleverser le marché de la photo<sup>20</sup> pour remplacer totalement leurs ancêtres argentiques<sup>21</sup>. Ils seront rejoints dès 2004 par un autre moyen de prise de vue numérique : les téléphones portables. Derniers éléments de cette montée en puissance des outils de captures numériques. Les reflex vont permettre aux particuliers d'accéder à des outils de qualité professionnelle. Avec leur capacité à filmer en haute définition, ils permettent l'accès à la production d'un contenu de bonne voire de très bonne qualité<sup>22</sup> achevant ainsi ce qu'avait déjà commencé la Super 8 en 1965 c'est-à-dire l'accès à la création vidéo amateur<sup>23</sup>.

A cette facilité de la prise de vue se rajoute la plus grande interopérabilité qu'ont apportée les formats de compressions présentés précédemment. Les outils de prise de vue actuels de basse qualité (téléphone portable) et de moyenne qualité (compact ou bridge numériques) intègrent maintenant les standards de compression en format de sortie<sup>24</sup>. Il est ainsi plus simple pour l'utilisateur de transférer son contenu depuis son appareil vers les plateformes en ligne comme YouTube, Dailymotion ou Flickr<sup>25</sup>. Pour la plupart des Smartphones de nouvelle génération, le partage peut s'effectuer directement depuis le téléphone.

L'ensemble de ces moyens de création a donc modifié la façon et la tranche de population qui traditionnellement créait les contenus multimédia. Nous sommes passés d'une création faites par des

---

<sup>19</sup> On pourrait ainsi considérer que le marché de la photographie était protégé par de très fortes barrières à l'entrée

<sup>20</sup> Les ventes d'appareils photo numérique (APN) vont passer de 40.000 unités en 1997 à 1.092.000 en 2002 avec un rythme de croissance quasiment égale à 100% (SIPEC 2002).

<sup>21</sup> En 2010, la vente d'appareils photo argentiques n'est plus représentée (Observatoire des professions de l'Image 2011)

<sup>22</sup> Les reflex numériques représentent aujourd'hui une alternative aux caméras professionnelles. Avec leur résolution HD, les reflex ont déjà été utilisés par des cinéastes pour tourner leurs films

<sup>23</sup> Il est d'ailleurs intéressant de voir que cette démocratisation de l'outil Super 8 s'est aussi faite par la mise en place d'un nouveau format plus adapté aux usages des particuliers (moins couteux et plus solide).

<sup>24</sup> Le format JPEG est particulièrement utilisé pour les photos ainsi que le standard MPEG-4 pour ce qui est des vidéos

<sup>25</sup> Celles-ci proposent d'ailleurs des outils de montage gratuit intégrés qui permette à un utilisateur de bénéficier de toutes les fonctionnalités techniques qu'aurait pu avoir un professionnel.

professionnels à une création faite par des particuliers. Tendance actuelle du web, l'utilisateur n'est plus seulement consommateur mais devient aussi créateur de ce qu'on appelle « Contenu généré par les utilisateurs » ou « **User Generated Content (UGC)** », génération spontanée de contenus par ces nouveaux acteurs. Causé par la chute des obstacles à la création, cette évolution de la création implique deux conséquences importantes :

1. Multiplication des acteurs de la création et hétérogénéisation de la création :

D'une création structurée par les professionnels qui proposent un média de masse, l'UGC agrège des contenus hétérogènes créés par des centaines de milliers de personnes (Cha, et al. 2007).

2. Volatilité de la création :

D'une création avec un temps de création en moyenne plus long les vidéos faites par les majors doivent en général avoir une durée de vie plus longue que les vidéos amateurs qui deviennent populaires ou s'oublie en une fraction de seconde.

De ces caractéristiques nous pouvons en conclure que la création de contenu multimédia va se scinder en deux catégories : d'une part, les contenus professionnels à forte valeur ajoutée. D'autre part, la création par les utilisateurs d'un contenu très important mais extrêmement hétérogène en qualité et en durée de vie. Nous sommes passés en quelques années d'une vision classique de répartition du contenu vidéo selon une courbe de Pareto<sup>26</sup> qui prônait la maximisation d'un nombre limité de contenu à forte audience vers une vision impactée par l'UGC, où il n'y a pas de limite au contenu produit car celui-ci est gratuit. La création de contenu change le modèle économique audiovisuel : d'une vision verticale s'est substituée une vision horizontale apparenté à la notion de longue traîne telle que la définie Anderson<sup>27</sup> (Anderson 2006).

Mais la création n'est pas la seule facette qui a été impactée par cette explosion du contenu multimédia. il y a aussi eu une modification de la consommation de ce contenu par les internautes.

**1.1.3. Evolution de la consommation de contenu multimédia.**

Grâce à l'amélioration des lignes de communications<sup>28</sup> et à la création de plateforme de visionnage<sup>29</sup> la consommation de contenu riches s'en est trouvée facilitée. Si l'usage était encore embryonnaire dans les années 90, le visionnage de vidéo a pris une place de plus en plus importante

---

<sup>26</sup> La courbe de Pareto est associée à la règle des 20/80, où 20% des vidéos représentent 80% des visionnages (et donc des revenus).

<sup>27</sup> On parle de longue traîne lorsqu'il n'y a pas de limite au catalogue proposé par le service. Cette vision s'oppose à la vision d'un magasin à l'étalage fini qui doit donc maximiser son gain par produit.

<sup>28</sup> Aux États-Unis, la part de la population connectée passe de 35% en 1999 à 69% en 2004 et les pays européens suivent la même tendance (Banque Mondiale 2011)

<sup>29</sup> YouTube et Dailymotion sont créés en 2005, Vimeo en 2004.

dans les usages du Web. Aujourd'hui cette consommation est totalement intégrée dans les mœurs puisque 75% des internautes visionnent des vidéos lorsqu'ils sont en ligne (Scheid, et al. 2012).

Mais à cette nouvelle habitude va aussi se greffer une nouvelle façon de consommer ce contenu vidéo. Par la limitation du temps maximum des vidéos par les principales plateformes<sup>30</sup>, la consommation vidéo sera donc plus fragmentée car celle-ci ne sera plus dorénavant composée de longs programmes mais plutôt d'une suite de programmes plus courts. Les vidéos vont aussi se détacher de la plateforme vers lequel l'utilisateur les a téléchargées (Upload). On va ainsi parler de version embarquable (Embedded) de la vidéo. Le lecteur flash sera alors directement présent dans la page du site client et appellera par lui-même le serveur de la plateforme qui contient la vidéo en question : le contenu ne sera plus limité à la seule plateforme et pourra être présent sur l'ensemble des sites qui la relaieront.

La propagation virale d'une vidéo (buzz) a beaucoup fait parler d'elle mais elle n'est qu'une conséquence de la modification des outils de prescription des vidéos. Grace aux fonctionnalités de partage (URL, version embarquable), et par le biais des réseaux sociaux, des blogs ou des forums, l'utilisateur n'est plus simple consommateur de contenu, il peut aussi donner son avis et partager les contenus qu'il a appréciés. Une vidéo pourra ainsi potentiellement bénéficier d'une audience dépassant le cercle des connaisseurs proches pour bénéficier de celui de ses prescripteurs si elle présente un intérêt pour lui. Les prescripteurs traditionnels de contenu (chaines de télévision) se voient donc concurrencés par les consommateurs eux-mêmes qui deviennent prescripteurs particuliers de contenu. La promotion verticale de contenu fait donc place à une propagation horizontale des contenus avec des consommateurs qui relaieront le message s'ils se sentent concernés. De même que l'UGC fragmentait la création, la consommation va se déporter d'une prescription faite par des acteurs institutionnels vers une prescription de niche par des particuliers. Or il faut bien se rendre compte qu'à la différence des acteurs traditionnels, les prescripteurs<sup>31</sup> ne sont pas forcément les producteurs de ce contenu, nous assistons à une dichotomie entre ces deux fonctions<sup>32</sup>. Il est donc crucial pour eux d'avoir des outils les plus fins pour trouver le contenu le plus riche.

Pourtant les solutions de recherche actuelles mises en place par les principales plateformes ne semblent pas permettre de répondre totalement à ces objectifs. Une étude Isobar de 2007 montre d'ailleurs que les internautes passent directement par les plateformes de visionnage plutôt que par les

---

<sup>30</sup>« Compared with non-UGC, the UGC video length is shorter by two orders of magnitude. Then median movie length in *Lovefilm* is 94 minutes » (Cha, et al. 2007).

<sup>31</sup> On parle aussi de la notion de curation de contenu, c'est-à-dire, la capacité d'un individu à agréger un contenu de qualité.

<sup>32</sup> Il n'est donc pas étonnant que des outils comme Scoop-It se développe, il y a aujourd'hui une vraie problématique sur l'accès à un contenu riche provenant d'une multitude de source.

moteurs de recherches classiques pour trouver du contenu vidéo. Il est donc tout à fait possible de se questionner sur la capacité de ces moteurs à fournir un résultat pertinent sur ces contenus<sup>33</sup>.

## 1.2. L'annotation et l'indexation : pré-requis à la recherche

Nous ferons donc un tour d'horizon sur l'existant ainsi que sur les solutions proposées par les acteurs spécialisés dans l'indexation sociale comme Deli.cio.us ou Flickr qui ont trouvé une façon innovante de chercher d'ordonner et le contenu.

### 1.2.1. Les méthodes actuelles d'indexation du contenu multimédia

Avec la démocratisation du web, l'explosion des sites et des contenus en avait déjà fait un espace où les informations ne se retrouvaient pas facilement. Dès 1995<sup>34</sup>, des dizaines de moteurs de recherche comme Yahoo! ou Altavista verront le jour pour combler ce manque.

Ces premiers moteurs de recherche travaillaient par indexation automatique des pages internet sur leur chemin. Ce travail était effectué par des robots appelés aussi « crawler » qui parcouraient le Web à l'aveugle par le biais des liens hypertextes. Leur mission était de trouver un maximum de pages puis de stocker les informations textuelles de celles-ci en cache sur les serveurs du moteur de recherche. Des algorithmes triaient alors ces résultats pour définir les pages internet les plus significatives pour l'internaute en fonction des occurrences des mots clés par rapport sa requête dans l'index général des pages stockées.

Lors du passage à un contenu multimédia, un certain nombre de ces critères qui paraissaient évident lors d'une recherche sur ce web classique sont devenu difficiles à appréhender dans le cadre de la recherche de contenu multimédia. L'annotation et l'indexation, pré-requis d'une recherche, ne fonctionnaient plus de la même manière. L'annotation, qui consiste à apporter des données de manière naturelle (texte) ou bien artificielle (métadonnées et tags à destination des robots) était directement identifiable et les robots pouvaient ainsi indexer les sites internet de manière efficace<sup>35</sup>. Dans le cas de contenu multimédia, ni l'annotation ni l'indexation ne sont directement possible. Jusqu'à aujourd'hui, l'indexation ne se fait jamais sur le contenu en lui-même mais sur la page du site qui lui est associée<sup>36</sup>.

---

<sup>33</sup> En effet, l'accès au contenu vidéo ne se fait que dans 47,6 % des cas par les moteurs de recherche, tandis que 68,2 % des internautes déclarent aller directement sur les plateformes de diffusion ou de partage de vidéo (ISOBAR 2007).

<sup>34</sup> Il s'agit là des moteurs de recherche de première génération, Google apparaîtra ensuite à la fin des années 90

<sup>35</sup> Et même dans ce cas là, la recherche n'est pas évidente, seul 5% du web est indexé, le reste considéré comme le web profond, c'est-à-dire l'ensemble des contenus non indexés par les moteurs de recherche.

<sup>36</sup> Google a fait une note pour les Webmaster sur son blog à propos des vidéos. Seuls quelques tags sont permis (<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=80472&topic=10079&ctx=topic#4> 2012)

Pour ce qui est des photos, Google ne prend en compte que la balise Alt pour indexer le contenu en lui-même.



Des technologies d'analyse de l'image existent mais elles ne permettent pas l'indexation de vidéo. Leur utilisation est très spécifique et les informations qui peuvent être recueillies sont assez basiques (recherche d'images identiques, couleur globale de la photographie). Ainsi ces technologies basées directement sur l'image permettent de définir si une vidéo répond correctement à un critère<sup>37</sup> mais ne permet pas d'aller plus loin, c'est-à-dire de tirer une information signifiante sur le sens de la vidéo. Seuls quelques outils comme ceux de reconnaissances faciales permettent d'ajouter des informations sur Picasa (2008) ou Facebook (2011). Ils restent toutefois sujet à un certain nombre de critique tant sur le coté déontologique de leur utilisation que sur sa véritable capacité à rajouter du sens.

Car le problème principal de ces moteurs de recherche classique vient de leur incapacité à saisir le sens des requêtes posées. Par l'indexation sur l'occurrence des mots clés, les moteurs ne prennent pas en compte le champ sémantique de la requête : les synonymes, les formes conjuguées ou les expressions peuvent être compris comme des éléments de la requête alors qu'ils n'ajoutent pas de sens à la requête mais plutôt du bruit nuisant à la pertinence des résultats retournés.

Autre point intéressant à analyser, les métadonnées, seules informations permettant de donner un sens à ce contenu ont toujours été considérées avec défiance par les moteurs de recherche. Ces techniques d'annotations artificielles du contenu avaient conduit à des dérives comme la remontée de contenus non pertinents (pornographiques en particulier). Cela avait obligé les moteurs de recherche à vérifier si ces métadonnées étaient cohérentes avec l'ensemble des annotations naturelles du site, difficilement falsifiables.

Une partie du modèle économique des plateformes de vidéos est donc remis en cause par l'incapacité des moteurs de recherche à indexer pleinement ces données. Selon Chris Anderson, c'est pourtant la capacité à exploiter cette longue traîne qui détermine la santé économique de l'entreprise. Or on note que cette exploitation est limitée dans le cadre de la recherche de contenu vidéo. (Cha, et al. 2007) notent ainsi que dans le cas de Netflix, les vidéos non populaires ne suivent plus une courbe de puissance mais que la queue de cette longue traîne semble tronquée par la difficulté à indexer ces contenus de niches.

Il a donc fallu chercher d'autres solutions que ces seuls moteurs de recherche classiques pour indexer ces contenus. Une des solutions viable a été une des modifications phare du web 2.0 : le taggage et la folksonomie.

### 1.2.2. Le concept de folksonomie

---

<sup>37</sup> YouTube en particulier des technologies associées à l'analyse de vidéos pour filtrer les contenus sous droits d'auteur (youtube content Id : <http://www.youtube.com/t/contentid/>)

Les métadonnées, invisibles à l'utilisateur, avaient été créées pour pallier à l'incapacité des robots à comprendre ces contenus multimédia. Le tag est au contraire une information visible volontairement mise en place par le créateur du site internet pour catégoriser les informations et donc favoriser l'indexation de ses articles<sup>38</sup>. Ce concept est ainsi Parce que le tag est visible et parce que son impact sur le référencement est plus beaucoup plus facile à appréhender. La folksonomie consistera à mettre ce taggage en commun.

Dans le cadre des contenus vidéo, ce système du taggage est utilisé à la fois par les moteurs de recherche (c'est un ajout de contenu pertinent participant à l'indexation de la page) mais aussi par les utilisateurs qui peuvent ainsi trouver plus facilement un contenu proche de ce qu'ils ont précédemment vu (en continuant sa recherche sur le même tag ou sur les autres tags qui sont associés à ce contenu). Le taggage est une annotation qui facilite l'indexation classique des moteurs de recherche classique ainsi que le trajet des internautes d'un contenu à un autre.

Ce principe de taggage du contenu peut aussi s'appliquer à un contenu que l'utilisateur va classer personnellement. Dans ce cas de figure, ce n'est plus le créateur de contenu qui ajoute du sens à sa création, c'est bien les utilisateurs qui vont organiser par eux-mêmes les contenus qu'ils consomment. Les sites de Musique à la demande (Deezer, Spotify ou Grooveshark) permettent ainsi à leurs utilisateurs de classer leur propre musique par la création de playlists qui peuvent être rapproché d'un taggage de la chanson. Cette classification personnelle par taggage a été définie principalement par Thomas Vander Wal (Le Deuff 2006) sous le nom de folksonomie étroite (Narrow Folksonomie).

Mais le véritable intérêt de la folksonomie réside dans les folksonomies générales (Broad Folksonomies), stade d'organisation où l'ensemble des acteurs, professionnels ou non ont la possibilité de tagger les contenus. Cela permet de bâtir une organisation pertinente, décentralisée et surtout totalement assumée du contenu par les utilisateurs. L'utilisateur remplace donc ici le robot des moteurs de recherche en étant acteur de l'indexation du contenu. Contrairement au seul taggage (annotation du contenu et indexation pour soi), la folksonomie permet aussi une indexation consultable et modifiable par les autres utilisateurs : certain site comme delicious.com ou Flickr recensent ainsi leurs objets de cette manière, laissant les utilisateurs organiser l'ensemble des contenus sur le site.

A la différence des moteurs de recherche classique les folksonomies, animé par des humains, possèdent la capacité d'appréhender ces contenus multimédia. Grâce à la notion de réseau et de communauté les folksonomies sont capables de traiter une somme de contenus importants. Ces

---

<sup>38</sup> Un blog pourra ainsi définir un certain nombre de mots-clés signifiants et les assigner à ses différents articles pour y gagner en cohérence. Le nuage de tags qui en ressort est un bon descriptif du champ de discussion du blog

solutions d'indexation par les utilisateurs sont d'autant plus attrayantes qu'elles sont moins chères à mettre en place et à entretenir que les solutions d'indexation classiques. Ce n'est d'ailleurs pas un hasard si l'ensemble des plateformes de visionnage utilisent ces tags sur leurs réseaux comme un socle à l'annotation et à l'indexation.

Pourtant le bilan des folksonomies n'est pas entièrement satisfaisant. Tout comme les solutions d'indexations classiques, les folksonomies ne traitent pas non plus les ambiguïtés du langage. Des tags peuvent reprendre des mots-clés ou des abréviations ambiguës à même de fausser la recherche<sup>39</sup>. Ce problème se répète lorsque le sujet aborde des environnements techniques. Les mots-clés définissant une même réalité mais avec un degré de technicité différent ne seront pas rapprochés par le taggage du contenu<sup>40</sup>.

Le taggage et la folksonomie répondent donc mieux au principe de la longue traîne, du moment que la communauté du site internet soit assez nombreuse pour annoter et indexer le contenu. De ce fait ils ont permis l'augmentation de la performance des outils de veille ou de suivi de l'information<sup>41</sup>. Le contenu est maintenant indexé sur un ou plusieurs mots clés spécifiques qu'il est possible de rassembler. L'utilisateur peut ainsi suivre un sujet sur plusieurs plateformes différentes du moment qu'elles utilisent même tags pour indexer un contenu similaire. Grâce à la technologie d'agrégation des flux, l'ensemble de ces mots clés ainsi être regroupé au sein d'un flux RSS unique qui représente la réalité d'un domaine de connaissance.

Pourtant reste le sentiment qu'un système pourrait répondre de manière plus efficace et même nativement à cette indexation de ces contenus aussi divers que la photo la vidéo ou les clips audio. Tim Berners Lee avait d'ailleurs prophétisé en 1994 que le rangement des contenus hétéroclites allait être un des défis majeurs des décennies suivantes. Il avait à ce sujet développé le concept de web sémantique, web intelligent, fait de robots et de moteurs de recherche capable de comprendre et de répondre aux requêtes faites en langage naturel. Nous allons donc nous pencher sur ce courant pour définir si les promesses et le concept développé seront capables de répondre à la problématique de l'indexation de ce contenu multimédia.

---

<sup>39</sup> Un même mot-clé peut faire ainsi référence à plusieurs domaines de connaissances différents (Passant 2009)

<sup>40</sup> Un objet et sa référence technique feront par exemple l'objet de deux tags différents alors qu'ils reflètent la même réalité (Passant 2009).

<sup>41</sup> « *L'avantage des folksonomies n'est pas qu'elles sont meilleures que les vocabulaires contrôlés, mais -qu'elles sont meilleures que rien - car les vocabulaires contrôlés ne sont pas adaptables dans la majorité des cas où un tag est requis. Construire, maintenir et améliorer un vocabulaire contrôlé revient très cher par rapport aux folksonomies, à la fois en temps de développement et pour l'utilisateur, notamment l'utilisateur amateur dans l'utilisation d'un système normé.* » Traduction par Olivier Le Doeuff (Shirky 2005)

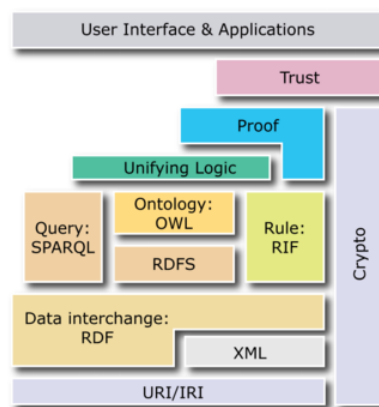
### 1.3. Le web sémantique : Une autre vision du contenu

La notion du web sémantique a été créée par Tim Berners Lee en 1999. Pour le créateur du web, l'organisation des contenus était un sujet important. Il considérait que le web n'était pas complet et qu'il manquait un système sémantique indispensable aux robots pour la compréhension des contenus hétérogènes (Amar et Menon 2011). Son but, soutenu par le W3C, était de faire du web un espace qui ne soit plus seulement lisible par les humains, mais aussi par les robots.

#### 1.3.1. Le web sémantique, un concept historique.

Ce concept a beaucoup évolué et a changé de nom plusieurs fois au cours des 20 dernières années. On le retrouve sous de nombreux termes, qui représentent autant de facettes du projet : les plus courants sont le Web 3.0, le mouvement d'Open Data, de Linked Data. Dans l'ensemble, ces différents concepts visent le même but, enrichir les contenus par des métadonnées sémantiques pour les rendre plus lisibles.

Pour atteindre cet objectif le web sémantique s'est basé sur un modèle en couches, similaire au modèle OSI (fig.1).



**Fig. 1 Organisation en couche du web sémantique (Wikipedia 2012)**

Cette configuration en couches va permettre de séparer les éléments en fonction de leurs complexités :

- les couches basses du web sémantique vont définir le langage RDF<sup>42</sup>, les formats de transports de ces informations sémantiques (XML) ainsi que le langage de requête sur ces métadonnées (SparQL). Les couches se serviront aussi de nouveaux identificateurs de ressources : les URI (Uniform Ressource Identifier) plus complets que les URL (Uniform Ressource Locator).

---

<sup>42</sup> Resource Description Framework:

- Les couches plus hautes représentent les langages de formalisation des couches basses. Grâce à celles-ci, les métadonnées vont pouvoir être regroupées en un espace structuré et homogène autour de concepts hiérarchisés. Cette couche comprendra les langages RDFS, OWL.
- Les couches de preuves et de confiance ne sont pas encore finalisées. Elles font encore l'objet de recherche et ne seront pas plus détaillées dans cette partie.

### 1.3.2. Définition de la couche basse du Web sémantique

Première étape de ce web sémantique, la définition des données va être particulièrement importante. Les URLs utilisées jusqu'à maintenant ne suffisent plus. En effet, comme leur dénomination le précise, celles-ci n'indiquent que l'emplacement de la ressource appelé (Uniform Resource Locator). Il faudra maintenant les identifier grâce un emplacement sous la forme d'URI (Uniform Resource Identifier) afin de distinguer plusieurs types de ressources (Passant 2009) :

- Les données existantes sur le web (documents, vidéos)
- Un élément du monde extérieur (un pays, une personne identifiée)<sup>43</sup>
- La relation entre deux éléments (visionner une vidéo)

Le Web sémantique va partir de ces ressources et les relier entre elles par le biais de ces URI. Le langage RDF aura besoin de trois éléments qui formeront de l'information sémantique (W3C-Recommendation 2004):

- Le **sujet** de l'action (défini par une URI)
- Le **prédicat** ou action du sujet sur l'objet là-aussi défini par une URI ou par une chaîne de caractère.
- Un **objet** qui sera l'élément sur lequel l'information portera (défini par une URI)

La réunion de ces trois éléments sera appelé un **triplet RDF**.

Pour que le code suivant rentre dans les spécifications du W3C sur le langage RDF, il nous faudrait le décrire sous sa forme XML :

```
<foaf:name rdf:about="Luka Magnotta">
  <foaf:travel>
    <rdf:description http://www.example.org/place/Paris>
    </rdf:description>
  </foaf:travel>
</foaf:name>
```

---

<sup>43</sup> L'objet réel sera ainsi décrit comme une classe au sein d'une ontologie sur lequel les triplets RDF pourront se baser.

Toutefois le XML n'est pas le seul moyen de décrire les triplets RDF qui peuvent tout aussi bien s'écrire dans leur version HTML.

```
<div xmlns:dc="http://purl.org/dc/elements/1.1/">  
<span property="dc:title">freakonomics</span>  
</div>
```

Nous avons ainsi une première couche d'information qui permet de relier entre elles des informations hétéroclites qui sont rapprochées par leurs URI. Pourtant, il est difficile de comprendre en quoi les robots peuvent mieux lire la donnée « *foaf:name* » que notre façon de décrire l'information naturellement : « *le nom est* ». Les deux sont aussi incompréhensibles pour les robots. La première couche est une couche de formalisation de l'information, il manque encore la couche de sens qui sera assurée par les ontologies.

### 1.3.3. Définition des couches élevés du Web sémantique

Pour que ces données reliées prennent du sens, il faut donc les rapprocher d'un ensemble à même de définir les liens entre ces différents triplets. Lorsque dans l'exemple précédent, les ressources étaient appelées, un préfixe suivi de deux points était présent, cet élément est la caractéristique de l'appel d'une ontologie. Dans l'exemple suivant `<foaf:name rdf:about="... ">`, le sujet du triplet a fait appel à l'ontologie « Friend of a friend » (FOAF), qui est l'ontologie principale de description entre les individus. Toutes indexations qui font appel à des classes d'une ontologie doivent mentionner la racine de l'ontologie pour que le robot puisse s'y référer :

```
xmlns:foaf="http://xmlns.com/foaf/0.1/".
```

On parle alors d'une instance (un objet singulier) qui fait appel à une classe (représentation général d'un concept, dans notre exemple, il s'agit du nom).

Les ontologies seront aussi composées de trois éléments qui ne doivent pas être rapprochés des triplets RDF :

- Classes : Concepts ou ensemble abstrait qui définissent un groupe d'individu. L'abstraction des classes dépendra de leur position hiérarchique. Des concepts hiérarchiquement élevé seront plus abstrait que leurs cousins qui seront reliés directement avec des instances.
- Propriété : il s'agit là des liens entre les classes qui définissent le type de relation qu'une classe entretient avec une autre classe ou avec des instances. Ces propriétés pourront être

des relations de hiérarchie, ou bien de paramétrage (la classe équipe de foot ne peut pas contenir plus de onze joueurs)

- Les axiomes : ils sont utilisés dans la définition de concepts ou de propriétés afin d'affiner celles-ci. ils permettent d'établir de nouveaux faits à partir des connaissances de base ou de vérifier la consistance d'un ensemble d'assertions.

#### 1.3.4. Avantage du web sémantique sur une l'annotation actuelle du contenu

Pour bien saisir le potentiel que représente le web sémantique, il faut revoir la question des systèmes de classification actuels. Comme nous l'avons mentionné, la finalité d'une ontologie est d'ordonner l'ensemble des liens qui lie ses éléments. Les classifications hiérarchiques classiques avaient déjà cette ambition. Les taxonomies utilisées par exemple en biologie, permettent seulement un classement d'ordre hiérarchique<sup>44</sup>. A la différence de ces dernières, les ontologies disposent d'un panel de liaisons beaucoup plus important<sup>45</sup> capable de décrire plus fidèlement la réalité du domaine de l'ontologie (Sapientis 2010).

Cette hiérarchisation des concepts va permettre de définir et surtout de qualifier les relations entre les concepts présents au sein d'une même ontologie. On peut par exemple définir non seulement le degré de séparation hiérarchique de deux concepts, mais aussi d'établir un score de similarité sémantique entre deux éléments de l'ontologie<sup>46</sup>. La structure sémantique permet donc par sa construction une vision plus juste des relations directes et surtout indirectes entre les concepts.

Enfin, le Framework RDF constitue une recommandation du W3C. Il a été construit pour permettre une interopérabilité maximum entre les différents fournisseurs de données sémantiques. Un site peut ainsi faire référence à plusieurs ontologies ou ajouter la sienne en complément, tant que celles-ci seront décrites par une URI, un triplet RDF sera lisible et interprétable par les robots. Ce format a aussi été beaucoup utilisé dans les mouvements de libération des données publiques. La transcription en Triplet RDF de ces données jusqu'alors illisible pour les robots montrent toute la puissance et la flexibilité du format qui a permis de rendre exploitable ces données par tous<sup>47</sup>.

---

<sup>44</sup> On peut rapprocher les taxonomies d'une ontologie composé de classe reliée uniquement par des liens de subsumption.

<sup>45</sup> Le langage d'ontologie Web OWL peut par exemple lier ses classes par équivalence, symétrie, contraire, cardinalité, transitivité, inclue des attributs avec des cardinalités et des valeurs restreintes

<sup>46</sup> On parle de degré de relation sémantique (i.e. semantic relatedness) lorsque la mesure calcule si deux concepts sont reliés sémantiquement, au sens où ils sont liés dans leur fonction. Par exemple, voiture et essence auront un fort degré de relation sémantique (Mazuel et Sabouret 2007)

<sup>47</sup> Pour Alexandre Monnin il s'agit d'ailleurs d'un des arguments les plus en faveur du web sémantique car « *c'est un facteur [...] essentiel pour la réutilisation des données, notamment dans la perspective de l'open Data dont la majorité des jeux de données (Excel, ou autres) sont inexploitable.* » entretien du 14 aout 2012

Le potentiel lié à l'annotation et l'indexation par le web sémantique de ces contenus multimédia paraît donc extrêmement prometteur et semble répondre en partie aux écueils dont souffraient les systèmes de recherche précédents. Toutefois, le web sémantique est encore un système jeune qui peine à trouver une pleine place dans le web actuel.

### 1.3.5. Les problématiques et faiblesses du Web sémantique.

Si le principe paraît prometteur du côté théorique, la mise en place de ces solutions peinent à voir le jour.<sup>48</sup> Nous dissociérons la partie ontologique du Framework RDF car ces deux parties présentent des problématiques différentes.

Pour la partie basse, c'est-à-dire l'enrichissement des pages par des métadonnées de type RDF, deux paramètres bloquent une partie des développeurs. Le premier point concerne la difficulté d'appréhension du concept de métadonnées sémantiques et de leurs constructions. Les métadonnées demandent une compréhension poussée de ce qu'est le web sémantique (maîtrise du format RDF/XML, interaction avec les ontologies) comme la tendance réelle du web sémantique est récente, les développeurs n'ont pas forcément développés de compétences dans ce domaine pour se servir de ces métadonnées.

Le deuxième point concerne l'utilisation des données sémantiques par les moteurs de recherches. Jusqu'à peu, Google ne prenait pas en compte les indications sémantiques dans ses critères d'indexation, il n'y avait donc pas d'avantage à développer des métadonnées puisqu'elles n'apportaient pas de valeur ajoutée à la SEO des sites internet<sup>49</sup>. Mais des efforts ont été faits pour ajouter aux CMS des blocs d'enrichissement sémantique qui ne nécessitent pas de compétence technique car le système de gestion de contenu devient capable de les générer de manière automatique (Passant 2009). Ce point-là est donc en passe d'être surmonté d'autant plus que les moteurs de recherches s'intéressent de plus en plus à l'indexation sémantique du contenu.

La grande difficulté du web sémantique concerne plutôt les ontologies. Du fait de la finesse de leur analyse, les ontologies gèrent assez mal les incohérences au sein de leur système<sup>50</sup> et l'ajout d'éléments peut ébranler l'ensemble de l'édifice.

---

<sup>48</sup> Nevertheless, after more than 80000 research articles about the 'Semantic Web' and 200000 about 'ontologies', the feasibility of the upper layers still seems unclear (Benel, Zhou et Cahier 2010)

<sup>49</sup> Search Engine Optimisation : Travail fait sur le code par les développeurs pour rendre le site internet le plus lisible et le plus facilement indexables pour les moteurs de recherche.

<sup>50</sup> : « *les machines sont mauvaises pour repérer les changements de version et peuvent s'embrouiller assez vite. En outre, ce sont des conceptualisations très fines, un changement quelque part peut déstabiliser l'édifice entier* » Alexandre Monnin entretien du 14 août 2012.



Enfin dernier élément plus philosophique, la vision ontologique des données part de l'idée que chaque concept représente une réalité objective, le système n'a pas la possibilité d'exprimer des concepts flous ou subjectifs : A vouloir remplacer l'ambiguïté d'une instance par le point de vue non-critique de l'ontologue, on perd toute la question du point de vue<sup>51</sup>. L'ontologie classique permet donc facilement une représentation des connaissances lorsque celles-ci sont objectives<sup>52</sup>. Elle ne permet pas la même finesse lorsque les contenus parlent de sujets discordants. Cela se retrouve en particulier dans les contenus appelant à l'interprétation, la négociation et les affaires ou bien dans tout simplement dans la collaboration (Benel, Zhou et Cahier 2010).

Il semble donc que le web sémantique permette une annotation plus précise et plus juste des contenus et en particulier des contenus difficilement indexables et annotables aujourd'hui.

Nous mesurerons la tendance autour de la nouvelle annotation des contenus qui s'oriente vers une grammaire sémantique. Il semble que la vision actuelle du web sémantique ne s'appuiera pas totalement sur les ontologies existantes et que l'annotation soit aujourd'hui l'utilisation majoritaire du web sémantique. A la différence de la vision de Tim Berners Lee et des ontologues bibliothécaires, le web sémantique ne pourra se construire sur des ontologies objectives mais d'abord sur une meilleure compréhension des données par leur annotation. Ce phénomène se produira à un niveau plus local et donc plus subjectif

---

<sup>51</sup>François Rastier définit les ontologies comme des : « Concrétisation d'un point de vue non critique » elles ne permettent pas une représentation d'un savoir fluctuant. (PhiloWeb 2010)

<sup>52</sup> C'est d'ailleurs pour cela qu'elles sont fortement utilisées dans les systèmes logiques de type circuits électronique.

## 2. Une tendance actuelle vers l'enrichissement par les métadonnées

On assiste donc aujourd'hui à une montée en puissance des triplets RDF sur le web et à de plus en plus d'initiatives personnelles d'utilisation de ces données. La quantité de triplets en circulation est la preuve que l'enrichissement des données commence à être pris en compte par les acteurs du secteur qui poussent les usagers et les développeurs à se servir des outils sémantiques qu'ils ont créés.

### 2.1. Le Web sémantique remis au gout du jour.

Nous montrerons ici que de plus en plus d'acteurs du marché s'intéressent et créent des outils d'annotation sémantique du contenu et que cette annotation en devient ainsi de plus en plus intéressant pour toutes les personnes qui veulent promouvoir leurs contenus.

#### 2.1.1. Apparition d'outils d'annotation sémantique par les acteurs du web.

Le développement de ces outils ne se fait pas de manière désintéressée de la part des grands moteurs de recherche et réseaux sociaux. On voit depuis quelques années le développement d'outils basés sur cette technologie pour améliorer la pertinence de leurs outils de recherche respectifs. Que ce soit Google, Bing, Yahoo !, ou bien encore Facebook, la tendance va vers la prise en compte de ces données. Ces acteurs ont publié leurs propres ontologies et les ont intégrés à leurs moteurs de recherche, créant ainsi un intérêt à l'annotation sémantique pour les sites internet indexés.

En 2007, Google poursuivait cette stratégie sur les contenus riches. Après le rachat de YouTube et la fusion entre la plateforme de visionnage et le service Google Vidéo, l'entreprise californienne allait s'attaquer au référencement des vidéos externes.

Fin 2007, Google proposera un outil de référencement des vidéos sur son site Webmaster Tools. Ce service appelé Sitemap vidéo, facilitera l'accès aux vidéos présentes sur les sites internet<sup>53</sup> pour les robots du moteur de recherche. Ce fichier XML dont la construction ressemble étrangement aux triplets RDF/XML rassemblera l'ensemble des données techniques des vidéos du site (durée de la vidéo, dimension du cadre, format de compression) ainsi que des données descriptives (image de présentation, tags, description textuelle) qui serviront à l'indexation de ces vidéos (Google Webmaster Central Blog 2007).

En 2008, Bing rachète le moteur de recherche Powerset suivi en 2010 par Google avec le rachat d'un autre moteur Freebase. Ces rachats illustrent l'ambition de ces acteurs majeurs d'utiliser

---

<sup>53</sup> Un fichier écrit en XML de type Sitemap sera directement présent à la racine du site, Google donne l'ensemble de ses spécifications techniques à l'adresse suivante : <http://support.google.com/webmasters/bin/answer.py?hl=fr&answer=80472>

l'information sémantique pour améliorer leurs recherches (ZDnet.fr 2010). Les moteurs de recherche l'ont bien compris, la prochaine étape du web sera sémantique. Mais ils ne sont pas les seuls à porter un intérêt à l'enrichissement des vidéos par les métadonnées. Facebook, réseau social aux 850 millions d'utilisateurs, a lui aussi lancé un projet d'enrichissement des contenus par une annotation sémantique.

Créé en 2010, le « Facebook Open Graph » permet aux partenaires d'enrichir eux-aussi les informations transmises à Facebook. Au travers de l'API du réseau social, ces partenaires transmettront ces mêmes types de métadonnées techniques et descriptives que l'outil Sitemap. Lorsque les personnes inscrites au réseau social utiliseront les boutons de type « like » ou « partager » ces métadonnées seront intégrées par Facebook dans des snippets pour mieux qualifier le contenu en question.

En 2011, les principaux moteurs de recherches Yahoo!, Bing et Google décideront finalement d'un outil commun d'enrichissement sémantique des vidéos. Le consortium à ainsi pour objectif de rendre l'accès aux contenus audio et vidéo plus facile pour leurs robots respectifs. Sous le nom de Schema.org, l'outil d'annotation est une évolution du format MicroData auquel on a rajouté une vision ontologique.

#### 2.1.2. Comparaison des outils d'annotations des contenus actuels

Les outils d'annotations sémantiques des contenus multimédias sont donc maintenant à disposition de tous. Pour autant la question de l'harmonisation est encore posée. Que ce soit les moteurs de recherche, le W3C ou bien Facebook, chacun défend sa plateforme d'annotation et veut en faire un standard. Comme les outils ne fonctionnent ni sur la même technologie, ni avec les mêmes objectifs, un tableau comparatif (fig. 1 : Tableau comparatif des solutions d'annotation) permettra d'avoir une vision plus claire sur les possibilités et avantages que chacune des solutions offrent.

Nous nous attacherons tout d'abord à décrire :

- les capacités techniques de chacune des solutions d'annotation (capacité à décrire les contenus)
- la capacité des outils à évoluer
- leur intégration dans les différentes plateformes CMS<sup>54</sup>
- leur utilité dans l'indexation par les différentes plateformes (moteurs de recherches et réseaux sociaux).

---

<sup>54</sup> CMS : Content Management System ou système de gestion de contenu (SGC) est une famille de logiciels destinés à la conception et à la mise à jour dynamique de sites Web ou d'applications multimédia.

Outils d'annotation sémantique actuels			
Informations générales	Video SiteMap	Schema.org	facebook OpenGraph
Date de création du format	2007	2011	2010
Format de référence pour les informations	XML / txt	XML/XHTML/HTML	HTML
Format d'écriture	XML	MicroData	RDFa
appel de l'ontologie	xmlns:video="http://www.google.com/schemas/sitemap-video/1.0"	"http://schema.org/VideoObject"	http://ogp.me/ns/video#
Informations techniques			
emplacement de la ressource	loc	itemprop="url"	og:video / og:video:secure_url
hauteur*largeur	pas d'information	itemprop="width" / itemprop="height"	og:video:width / og:video:height
durée des vidéos	video:duration	itemprop="duration"	video:duration
Richesse sémantique de l'ontologie	faible (intégré dans le protocole sitemap.org)	forte (plus de 2000 classes)	faible (10 types d'actions, peu de classes)
Capacité à décrire le contenu			
Nom du contenu	video:title	itemprop="name"	og:title
auteur du contenu	pas d'information	itemprop="author"	og:video:director
Thumbnails	video:thumbnail_loc	itemprop="thumbnail" / itemprop="thumbnailUrl"	og:image
Mots-clés / tags	video:tag	itemprop="tag"	video:tag
Description	video:description	itemprop="description"	og:description
Utilisation par les différentes plateformes			
Utilisation sur les moteurs de recherche	Google seulement	Oui	Partiellement (ne permet pas la création de snippets)
Utilisation par Facebook	Non	Non	Oui
Facilité d'utilisation & Evolutivité			
Interface de création d'annotation (pas de passage par le code)	Non	Non	Oui (possibilité d'ajouter ces données directement depuis le "Facebook Dashboard")
Evolutivité de l'ontologie	Non	Possibilité de proposition de nouveaux éléments	Administré par Facebook
Compatibilité avec les microData	Non	Oui	Non
Popularité de l'outil au 1 août 2012 (cf. Annexe 3)	4	19	33
intégration sur les principales plateformes CMS			
Wordpress (en occurrence google et en modules disponibles)	170 000 17 (source index WP)	214 000 43 (ibid)	614 000 118 (ibid)
Drupal (en modules disponibles)	20	3	1299
Joomla	34 100	56 500	164 000

**fig. 2 : Tableau comparatif des solutions sémantiques Apport personnel**

Les trois solutions sont assez similaires sur leur capacité à annoter le contenu vidéo, on toutefois un léger avantage à Schema.org qui possède un vocabulaire légèrement plus riche et plus étendu ainsi qu'une capacité à voir son vocabulaire s'enrichir par les utilisateurs, plutôt intéressante<sup>55</sup>.

Pour ce qui est de l'intégration sur les différentes plateformes CMS, on note un avantage fort pour le format Facebook Open graph, l'ensemble des plateformes ont vu se développer un nombre important de modules permettant d'enrichir automatiquement les sites internet qui les ont adoptés. Ce sentiment est corroboré par une tendance plus forte pour le Facebook Open Graph plutôt que pour le

<sup>55</sup> Schema.org propose ainsi un outil assez complet pour ajouter ses propres classes à l'ensemble : <http://www.schema.org/docs/extension.html>

protocole Schema.org. On note également la perte de vitesse du Sitemap qui semble être absorbé peu à peu par les deux autres outils.

L'outil Facebook d'annotation du contenu semble être moins puissant que son concurrent Schema.org. Cependant, on remarque que l'interopérabilité entre les deux systèmes semble faible. C'est donc le Facebook Open Graph qui aura les faveurs des développeurs portés par le web social. L'ontologie Schema.org sera particulièrement conseillée pour les référenceurs de contenus (SEO, e-commerce en particulier).

Dans les tout les cas, c'est bien la volonté d'améliorer la visibilité du site, le taux de clic ou le trafic venant des moteurs de recherche ou des réseaux sociaux<sup>56</sup> qui doit guider le choix entre les deux solutions : Le passage à une annotation de type Open Graph et/ou Schema.org permet seulement de définir sur quel canal on veut jouer. Rien n'empêche d'ailleurs d'utiliser les deux systèmes.

Cette étape se retrouvera dans la mise en avant de résultats riches. D'une liste simple les moteurs de recherche rajoutent des aperçus<sup>57</sup> sous forme d'image ainsi qu'un texte descriptif. La recherche devient donc un mélange de pages et d'éléments plus riches appelé snippets (cf. Annexe 2. Annotation automatique des vidéos par YouTube et création de Snippets Google et Facebook)

A cette nouvelle façon d'annoter le contenu proposé par les moteurs de recherche vient aussi se greffer une nouvelle façon de considérer le contenu : le médiafragment.

### 2.1.3. Fragmentation de la lecture des contenus multimédia sur internet.

De nouvelles idées sont venues se greffer sur le format <vidéo> pour le rendre aussi riche et attrayant dans ses fonctionnalités que les lecteurs multimédia classiques et cela concerne en particulier la fragmentation de la vidéo directement par l'URL<sup>58</sup>. Jusqu'à maintenant la vidéo était un espace privé dans une page HTML, inaccessible aux robots des moteurs de recherche. Ceux-ci devaient se fier cette page pour lire les informations essentielles de la vidéo. L'idée de ce format est donc de faire en sorte que la vidéo augmente sa capacité d'interaction avec la page : les informations de cette dernière ne sont plus liées avec l'ensemble de la vidéo mais directement avec un moment particulier de la vidéo.

---

<sup>56</sup> Tumblr a ainsi multiplié son trafic venant de Facebook par 2.5 depuis l'ajout de balise Open graph, Bestbuy's allègue une augmentation de 30% de CTR au passage aux Snippets sur les moteurs de recherche. (Brown 2012)

<sup>57</sup> On parle aussi de thumbnails qui sont une de ces nouvelles balise de Schema.org <meta itemprop="thumbnailUrl" content="thumbnail.jpg" /> ou de la feuille de Google vidéo Sitemap en XML <vidéo:thumbnail\_loc> <http://www.example.com/thumbs/123.jpg> </vidéo:thumbnail\_loc> </vidéo:vidéo>

<sup>58</sup> Il s'agit en particulier des travaux sur les médiafragments qui sont devenu une recommandation de la part du W3C en mars 2012 : Media Fragments URI 1.0 (basic) <http://www.w3.org/TR/2012/PR-media-frags-20120315/>

Les liaisons entre la vidéo, la page et son adresse URL sont ainsi facilitées. L'information circulera directement dans l'URL sous le format historique de l'ancre (Anchor ou bien #) pour indiquer un moment précis de la vidéo. Ainsi si l'on veut regarder le passage de la conférence TechCrunch situé à 1min 45 secondes du début de la vidéo, il suffira d'ajouter le suffixe suivant : #t=1m45s à l'URL :

<http://www.youtube.com/watch?v=1bibCui3IFM#t=1m45s> (TechCrunch 2008)

Cette fragmentation de la vidéo permet donc à l'utilisateur de transmettre à ces amis une vidéo non plus seulement dans son ensemble mais directement l'extrait qui lui semblait important. On passe d'une conception où la vidéo était un contenu continu à celle d'une vidéo grammatisée : c'est-à-dire découpée en unités de sens signifiantes (fragments)<sup>59</sup>. Cette possibilité de fragmenter le contenu représente la capacité de l'individu à ajouter du sens à un contenu trop indigeste en le séparant en des ensembles temporels cohérents.

Cette fragmentation de la vidéo va permettre d'y intégrer plus facilement le concept du lien hypertexte. Comme le contenu est maintenant organisé en unité de sens, il peut plus facilement interagir avec des sources extérieures.

Arte a d'ailleurs développé un lecteur pour ses web-documentaires qui illustre cette nouvelle vision de la vidéo relié avec le contenu : Si un concept ou un personnage important apparaît dans la vidéo, le spectateur peut mettre le document en pause et suivre un lien qui lui en apprendra davantage sur le sujet.

Ces deux modifications de la consommation de la vidéo par la fragmentation de sa ligne de temps pourraient d'ailleurs améliorer l'annotation sémantique de ces contenus. Nous allons donc creuser cette innovation technique pour lui donner un caractère sémantique.

#### 2.1.4. Une fragmentation qui favorise l'annotation

L'ajout d'un caractère temporel est fondamental dans la description des contenus multimédia, ils peuvent être maintenant non seulement définis par une description globale mais aussi par une description qui peut varier au cours du temps. Dans le cas de contenus hétérogènes cela améliore grandement la qualité de l'annotation qui se fie maintenant à des fragments de vidéo plus précis que l'ensemble de la vidéo<sup>60</sup>.

---

<sup>59</sup> « En passant de l'appareil analogique à l'appareil numérique, des parties du signal m'apparaissent en tant que discrètes, et c'est ce qui rend possible ce qu'on appelle l'interactivité : je peux alors agir sur l'information, la transformer, et non seulement la subir ». (Ars Industrialis 2011)

<sup>60</sup> Youtube avait d'ailleurs permis en 2009 d'augmenter le temps maximum d'une vidéo à l'upload à 15 minutes et à 2 heures pour les comptes débloqués. Cela implique mécaniquement des contenus plus hétérogènes.

Youtube à d'ailleurs commencé à rajouter des fonctionnalités de ce type au sein de sa plateforme. Les utilisateurs peuvent maintenant renvoyer directement à un moment de la vidéo commentée en inscrivant le moment sous le format MM:SS dans leurs commentaires. La navigation non sémantique à travers ces médiafragment est donc déjà avérée depuis 2011, il s'agit de déterminer tous les atouts que pourrait apporter ces média fragments s'ils étaient sémantiques. Ces médiafragments ne font pas que rajouter du sens, ils placent ce sens dans un contexte temporelle ce qui est tout aussi important. La liaison des deux serait donc d'autant plus fructueuse.

## 2.2. Sémantiser et annoter les contenu multimédia

L'annotation des contenus devient donc maintenant une réalité du web. Pourtant cette nouvelle norme n'a pas encore impliquée de gestion native de cette annotation. Il faut donc maintenant s'interroger sur la capacité de nouveaux outils à sémantiser le web et en particulier les contenus multimédias.

### 2.2.1. « Ligne de temps » : outil dédié à l'annotation sémantique.

« Ligne de temps » est un outil d'annotation pour les long-métrages développé par l'Institut de recherche et d'Innovation (IRI). Il a été conçu à l'origine pour fragmenter un film en lui donnant des unités de sens annotées sémantiquement. L'outil peut créer plusieurs lignes d'annotation sur un même film en fonction des axes d'analyse que l'utilisateur souhaite créer. Chaque personne peut ainsi construire sa propre interprétation du film en fonction des axes qu'il a créé. Le logiciel permet ensuite de récupérer les différentes analyses sous la forme d'un fichier .ldt. Ce fichier enregistre les données dans un format XHTML suivant la hiérarchie qui suit (cf. fig. 2.) :

The screenshot shows the 'Lignes de temps' (Timeline) tool interface. At the top, there is a header with the text 'l'analyse et sa description : conteneur<iri><head>métadescriptions</head></iri>'. Below this is a timeline for the video 'Romanzo Criminale' with a duration of 02:26:00. A 'Nouveau découpage' (New cut) button is visible. The timeline is divided into segments, with one segment labeled 'Fragment n°X'. A callout box provides a detailed view of the XML metadata for this fragment:

```
<element id="eID_X"
begin="4567" dur="65"
author="DDC-IRI" date="06/2006"
title="axe d'analyse, fragment n°X"
abstract="description du fragment"
src="iri_IDfilm/thumbs/0_in.jpg"
/>
```

Fig. 3. : Description des fichiers de sauvegardes Ligne de temps (Apport personnel)

On retrouve dans le fichier le nom du film et l'analyse associée auxquels sont ajoutées des métadonnées qui décrivent l'ensemble du projet. L'axe d'analyse fragmentera le film en éléments signifiants. Ces fragments délimités dans le temps pourront ensuite être enrichi avec des descriptions (abstract), images (src) ou encore des tags (tag).

L'outil qui partage de nombreuses caractéristiques avec les logiciels de montage permet de mélanger le film avec ses propres textes et/ou avec d'autres extraits de film. On peut ainsi s'en servir pour construire une version documentaire de l'analyse filmique. Il est à ce titre utilisé pour associer ou comparer des œuvres et leur construction directement dans l'outil.

On peut d'ailleurs rapprocher son utilisation d'une folksonomie étroite, c'est-à-dire vers une indexation du film à usage personnel. Il s'agit d'un outil qui ne permet d'annoter le contenu que par une seule personne. Il n'y a pour l'instant pas de moyen de continuer le travail de manière communautaire comme pourrait le permettre un wiki<sup>61</sup>. Il serait pourtant intéressant de pouvoir confronter les analyses sur un film.

D'autre part, l'annotation filmique demande un engagement fort de la part des personnes qui utilisent l'outil. L'ajout d'une ligne de temps demande entre quatre à cinq heures sinon plus selon la finesse de l'analyse. L'effort que cela implique n'est pas forcément à la portée de la majorité des consommateurs. De la même manière, si l'on considère la quantité de film sortant chaque année, cette annotation par des agents humains rémunérés impliquerait un coût trop important par rapport au bénéfice apporté par l'annotation. L'outil est tellement puissant et demande des compétences filmographiques tellement poussées que son utilisation n'est pas à la portée du néophyte. La portée infinie de l'annotation qu'implique la liberté d'action de l'outil peut aussi freiner une mise en œuvre d'outils d'indexation<sup>62</sup>. Depuis, l'IRI a recentré l'usage du logiciel sur un autre type de contenu qui peut intégrer une quantité importante d'annotation : les conférences. L'outil permet ainsi plus facilement de scinder les sujets ou les différentes interventions.

Le modèle proposé semble donc permettre une annotation optimale mais cette annotation implique à l'heure actuelle un coût en termes de temps ou d'argent trop important pour que cette solution puisse être généralisée sur le web. Il faut donc trouver un moyen de favoriser l'annotation par les utilisateurs en intégrant cette création aux usages déjà existants<sup>63</sup>.

---

<sup>61</sup> Un wiki est un site web dont les pages sont modifiables par les visiteurs afin de permettre l'écriture et l'illustration collaboratives des documents numériques qu'il contient. (Wikipedia 2012)

<sup>62</sup> En particulier parce que la distance sémantique augmente lorsque beaucoup d'éléments sont rajouté

<sup>63</sup> « *Le challenge de l'annotation sémantique c'est d'arriver à augmenter le sens sémantique des usages déjà existants des utilisateurs. Il faut partir de leur manière de pratiquer la lecture de vidéo pour trouver des outils capables de générer des triplets sémantiques* » Samuel Huron, entretien du 13 août 2012



### 2.2.2. Polemic Tweet : utiliser les réseaux sociaux pour annoter

Pour que ces annotations aient un sens économique, il faut favoriser l'implication des utilisateurs dans la création d'une organisation sémantique de l'information. Dès lors, il s'agit de définir les contenus qui sont les plus à même de recueillir les commentaires et témoignages des utilisateurs. Si ce sont des usages déjà existant, ils ne sont pas beaucoup utilisés sur les plateformes de contenu multimédia : les taux de mention constatés sur les vidéos sont de manière générale plus faible que sur d'autres types de contenus. Les blogs et les forums restent les espaces de discussion alors que les plateformes de visionnage souffrent d'un plus faible taux d'engagement des utilisateurs<sup>64</sup>

Toutefois, on remarque que les contenus vidéos de type débats ou interviews drainent des utilisateurs plus engagés. On passe alors d'un taux de commentaires de 0.36 pour mille pour les contenus génériques à 3.4 pour mille soit un taux dix fois meilleurs pour ces contenus (cf. Ann. 1). Les vidéos porteuses d'un débat ou en tous cas d'une vision du monde sont donc plus à même d'accueillir les réflexions des utilisateurs. Cela reprend un axe d'étude sur la démocratie directe : plus le débat est centré sur un point précis délimité dans le temps et dans l'espace, plus il sera riche et porteur de solutions ou d'idées, apportant avec cela de nouveaux participants à même d'annoter ces contenus (Blondiaux 2008).

Pourtant il est étrange de ne pas considérer les réseaux sociaux comme des vecteurs d'annotation alors même qu'ils permettent de commenter de manière plus efficace et plus large ces contenus<sup>65</sup>. L'Institut de Recherche et de l'Innovation (IRI) a ainsi développé un outil appelé Polemic Tweet. Cet outil a été mis en place pour permettre aux auditeurs d'une conférence de commenter sémantiquement celle-ci en direct. L'outil répond ainsi aux trois problématiques qui sont posées par l'annotation des contenus sémantiques :

- Ajouter le sens par les utilisateurs et favoriser l'annotation,
- Annoter dans une temporalité,
- Annoter les contenus porteurs de potentiels.

#### Ajouter le sens par l'utilisateur (User Generated Interlinking)

---

<sup>64</sup> Dans une étude Orange sur le suivi des matchs lors de la coupe du monde 2010, le taux de commentaire moyen était de 0.36 par 1000 vues sur Youtube (Beuscart, et al. 2011). L'engagement est aussi assez faible sur les outils de classement des vidéos, les votes comptabilisent 0.22% des vues totales (Cha, et al. 2007).

<sup>65</sup> Facebook représente 150 millions de clics pour Dailymotion, il serait intéressant de savoir combien de message le premier réseau social créé sur ces contenus partagés (Agence Mode d'emploi 2012).

Le web sémantique a souvent été représenté comme une technologie qui allait révolutionner le web et à ce titre modifier totalement les usages des internautes. Depuis, cette manière de voir les contenus sémantiques a évolué. Aujourd'hui, avec le passage du concept de Linked Data, il s'agit plutôt de faire correspondre l'ajout de métadonnées avec les usages en place pour faciliter la navigation et les façons de faire de l'utilisateur<sup>66</sup>.

Dans le cadre de Polemic Tweet, l'objectif était de transcrire l'avis de l'utilisateur (positif, négatif, interrogatif ou simplement informatif) ainsi que les mots-clés importants (tags) de leurs message grâce à Twitter (fig.3). Pour cela les personnes pouvaient utiliser l'interface dédiée ou leur propre compte Twitter et envoyer les messages avec le hashtag<sup>67</sup> de la conférence en question<sup>1</sup>.

Dans les deux cas, ils pouvaient donner leur avis sur la conférence en y ajoutant les mentions correspondantes : ++ -- ?? ==<sup>2</sup>.

Le Tweet était ensuite associé temporellement avec le moment de la conférence<sup>3</sup> par l'heure à laquelle l'utilisateur twittait le message (IRI 2012). Les tweets étaient enfin retranscrits sous formes de diagrammes en bâtonnets placés de manière chronologiques le long de la conférence.<sup>4</sup>



**Fig. 4. Chemin de l'annotation par Polemic Tweet (Institut de recherche et d'innovation)**

Toutefois les données ainsi récoltées ne sont pas directement exploitables, il faut faire le lien entre les métadonnées qui ont été créées et le contenu multimédia sur lequel les internautes ont participé.

<sup>66</sup> Pour Samuel Huron, designer de l'interface de l'outil Sémantico Tweet, il s'agissait d'inclure cette annotation sémantique dans une pratique déjà existante du commentaire de conférence de manière à ce que l'ajout de l'information soit le plus facile possible.

<sup>67</sup> Hashtag : symbolisé par un # il permet d'identifier le mot qui y est accolé comme étant signifiant du Tweet, cette notion peut être rapproché de la notion de tag.

Cette fusion ne pourra se faire que dans le cadre d'un lecteur capable d'ajouter cette couche de métadonnées sémantiques à la seule lecture du contenu.

### 2.2.3. Synote et Metadata Player : outils de jointure entre le contenu et l'annotation

A cause des lecteurs multimédia, les contenus sont toujours dans cette problématique d'emplacement et de lecture externe à la page. A ce problème initial se rajoute l'annotation sémantique qui répond à une logique temporelle de l'annotation et qui ne peut donc pas être présente en tant que telle sur la page<sup>68</sup>Cette dernière en effet est le socle de l'annotation générale, c'est-à-dire commune à l'ensemble des parties (médiafragments) du contenu aux. Nous prendrons pour exemple deux lecteurs multimédia qui comprennent et permettent une navigation par les média fragments :

- Metadata Player, créé lui aussi par l'IRI pour sa navigation temporelle remarquable.
- Synote, lecteur développé par les chercheurs de l'« University of Southampton » qui se focalise plutôt sur la capacité à indexer ces méta-données par le biais de Schema.org et les contraintes que l'indexation temporelle des contenus implique.

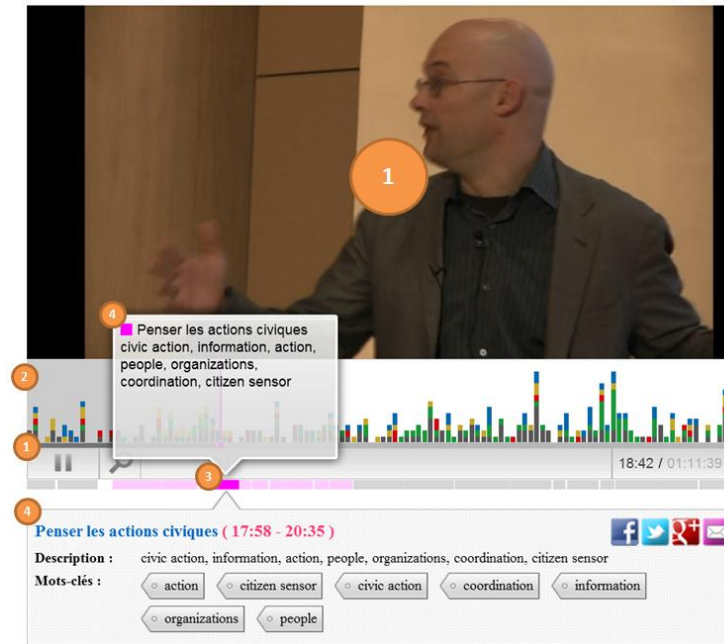
#### Présentation de metadata Player.

Comme nous l'avons précédemment mentionné, l'objectif de Metadata Player est de parvenir à rassembler le contenu avec l'ensemble des annotations qui y sont affiliées. Ce lecteur rassemblera les caractéristiques d'un lecteur vidéo classiques, c'est-à-dire, le bloc de visionnage et le barre de lecture <sup>1</sup> (cf. fig. 4 ci-dessous). Mais à ces attributs normaux vont se rajouter de nouveaux axes de lecture de la vidéo grâce à une brique JavaScript qui appellera dynamiquement les tags et annotations reliées au médiafragment en question <sup>3</sup> :

- La lecture polémique de la conférence tirée de l'outil Polemic Tweet <sup>2</sup>
- la fragmentation de vidéo en unités de sens <sup>3</sup>. Ici cette fragmentation se fait sur les sujets abordés par les conférenciers <sup>4</sup>.
- Enfin, les tweets des utilisateurs qui ont permis de récolter les mots-clés signifiants sur le fragment concerné <sup>4</sup>

---

<sup>68</sup> Les tags temporels dépendent de fragments de la vidéo, ils ne répondent pas au taggage global de la vidéo et doivent donc être séparés des métadonnées générales de la vidéo.



**Fig. 5 : Présentation de métadonnées Player.**

On notera aussi que les outils métadonnées Player ne permettent pas nativement une annotation sémantique rigoureuse du contenu<sup>69</sup>. Cela empêche donc la création de snippets directement liés au fragment regardé plutôt qu'à la vidéo en entier. De la même manière il serait intéressant de créer une évolution des boutons de partage dont l'URL serait directement reliée au médiafragment sur lequel il se trouve et qui renverrait un snippet correctement borné via le protocole Facebook Open Graph.

### Présentation de Synote

Synote est un lecteur de contenu multimédia qui s'apparente à Métadonnées Player présenté ci-dessus. Il permet la visualisation des contenus de la même manière mais son véritable atout concerne surtout le stockage des annotations en vue d'une indexation des fragments par les moteurs de recherche<sup>70</sup>. En effet, les fragments sont actuellement annotés avec des ancres html, le contenu qui leur est associé est donc appelé dynamiquement via des appels JavaScript. Les moteurs de recherche n'indexant pas les contenus dynamique ni les ancres<sup>71</sup> il a fallu trouver une solution de repli pour que ces données soient

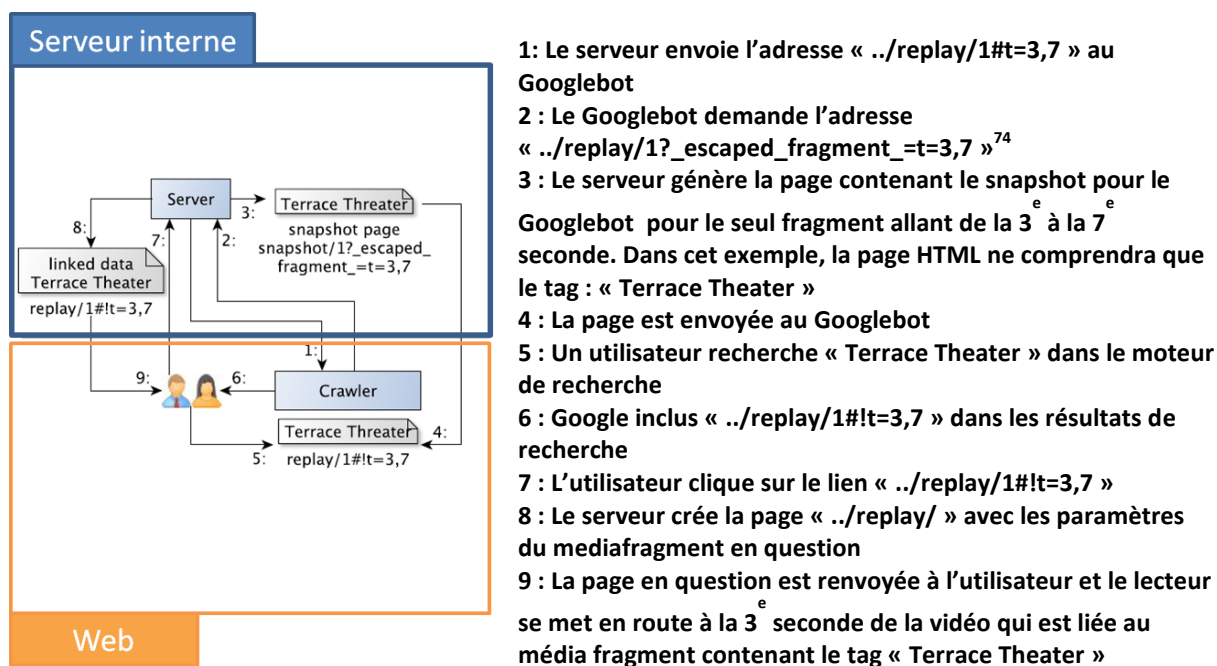
---

<sup>69</sup> Les tags de métadonnées par exemple sont encapsulés dans des listes mais pourrait avoir un contenu plus riche par le biais de schema.org.

<sup>70</sup> L'élément vidéo n'est plus l'élément central du contenu, le fragment devient une entité (URI propre) accessible directement par le moteur de recherche. « *As the use of schema.org is based on the HTML and search engine indexing infrastructure, we need to make the Web pages, which contain media fragments, search-friendly.* » (Li, Wald et Omitola, et al. 2012)

<sup>71</sup> Les ancres ne renvoient actuellement qu'à une partie de la page et ne sont pas considérées par les robots comme représentant un contenu unique (doublement avec les informations de la page qui contient l'ancre).

quand même lisibles pour les robots<sup>72</sup>. En utilisant le Framework (devant la pauvreté de la traduction du terme nous nous en tiendront au terme anglais) de visualisation des pages dynamiques générées via Ajax<sup>73</sup>, Synote a pu contourner le problème sans y sacrifier l'ergonomie (les données sémantique et le lecteur sont sur la même page) et permettre en même temps une indexation des contenus par les moteurs de recherche classique sans perte d'information (un fragment aura donc son snippet attiré et les tags spécifique du fragment)



**Fig. 6 : Génération de snapshot dynamique par Synote (Li, Wald et Omitola, et al. 2012)**

La solution ainsi proposé par Synote est alléchante, elle permet de concilier à la fois la vision sémantique, la vision temporelle des videos en y ajoutant une réelle capacité à les faire indexer par les moteurs de recherche classique. Ce dernier point est tout de même à regarder attentivement, le Framework utilisé pour concilier affichage dynamique et indexation reste spécifique à Google et les ancrages spécifiques aux appels dynamiques via Ajax (#!) intégrés aux URL sont de plus en plus critiqués<sup>75</sup>. Cette solution doit donc être considérée comme une solution de secours qui fixe un cadre à la direction que devrait prendre l'annotation et l'indexation des contenus multimédia par le Web sémantique.

<sup>72</sup>En particulier, le Framework Schema.org ne prend pas en compte les médiafragments (Li, Wald et Omitola, et al. 2012).

<sup>73</sup><https://developers.google.com/webmasters/ajax-crawling/docs/getting-started?hl=fr>

<sup>74</sup>Voir les spécifications techniques : Mapping from #! to \_escaped\_fragment\_ format (Google developpers 2012)

<sup>75</sup>“Twitter recently also announced that they would remove all the hashbang URL” (Li, Wald et Wills, Let Google Index Your Media Fragments 2012)

Ces lecteurs multimédias permettent néanmoins un pas de géant dans la visualisation et l'utilisation de la donnée temporelle dans l'annotation. Il appartient donc maintenant aux acteurs actuels de développer des solutions à même de prendre ces données en compte dans leur indexation.

### 2.3. Définir les usages de ces méta données

Mais ces métadonnées sémantiques temporelles présentent d'autres moyens d'être utilisée que par la seule indexation. Cette façon d'annoter est beaucoup plus lisible pour les agents machine et cela ne concerne donc pas uniquement les robots des moteurs de recherche. La libération des données par le mouvement Open data a poussé les gouvernements à pousser des données publiques sur le web. Le caractère indigeste et extrêmement technique de ces données ont poussé les développeurs à ne pas les utiliser en tant que tel mais à les mélanger avec d'autres jeux de données pour leur faire prendre un sens parfois insoupçonné<sup>76</sup>.

#### 2.3.1. Favoriser l'interopérabilité des contenus

Grâce aux éléments qui nous avons précédemment expliqué, les contenus peuvent maintenant être annotés par des triplets RDF indexés dans la temporalité de la vidéo. Mieux encore, ces données sont compréhensibles par les machines qui les explorent car ces données respectent la grammaire de la recommandation Médiafragment ainsi qu'une ou plusieurs ontologies.

Mais les spécifications techniques du web font aussi que ces données doivent s'inscrire dans un format de transfert formalisé pour permettre une interaction entre des applications tierces. Les formats actuels pour cette encapsulation de l'information sont principalement le JSON et le XML<sup>77</sup>. Ces deux formats ont l'intérêt de créer une structure compréhensible par les robots car celles-ci sont balisées de manière stricte<sup>78</sup>. Les robots peuvent ainsi comprendre les interactions et les relations entre les éléments. Ces couches de formalisation vont ensuite être utilisées par des protocoles comme le JavaScript, l'Ajax et le jQuery pour modifier dynamiquement la page sur laquelle l'utilisateur se trouve (Tran 2011). A la différence des systèmes aux formats propriétaires, ces protocoles ne demandent pas de couches de traduction (ETL ou EAI) entre applications tiers. L'interopérabilité entre

---

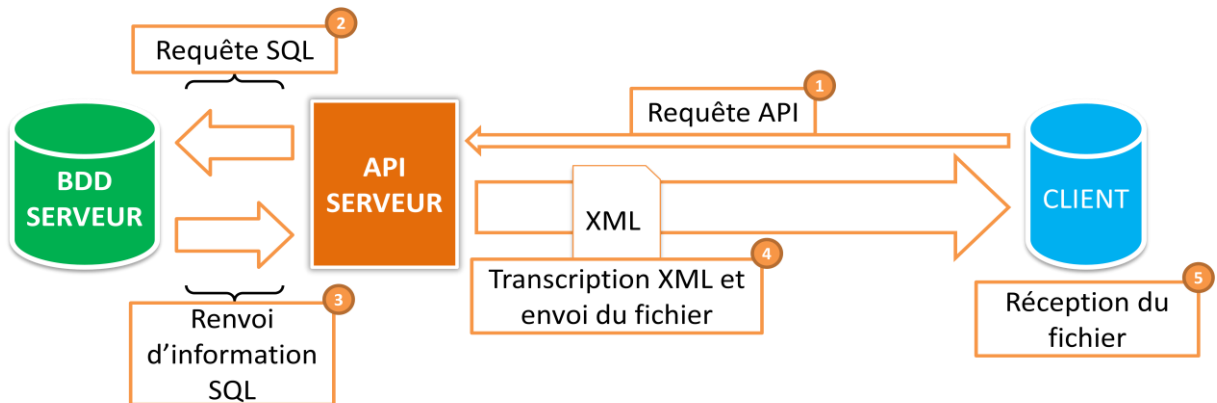
<sup>76</sup> Des outils de visualisation ont par exemple été mis en place pour suivre l'évolution spatiale de maladies à travers le monde. L'outil Google basée sur les recherches du mot « flu » et de ses traductions avait permis de définir les aires d'effets de la maladie ainsi que sa saisonnalité et ceux-ci, avec un niveau de précision similaire à celle d'études médicales. <http://www.google.org/flutrends/>

<sup>77</sup> Le premier se rapproche dans sa syntaxe du JavaScript alors que le XML est beaucoup plus lié au RDF de part la recommandation du W3C sur le RDF/XML (W3C-Recommendation 2004).

<sup>78</sup> Un document est valide lorsqu'il est bien formé, et respecte les règles syntaxiques du langage XML et obéit à une structure type définie explicitement dans un DTD (document type définition). Un document doit impérativement respecter la syntaxe du XML. On dira alors que le document est "bien formé" [Well-formed].

les systèmes et donc beaucoup plus grandes et l'on parle alors d'une approche SOA de l'architecture entre ces différents clients (Pétine 2012).

Enfin dernière étape de cette communication entre les acteurs, il faut maintenant appeler ces informations par des web services via les interfaces de programmation (Application Programming Interface ou API) des acteurs en question. Ces interfaces permettent d'interroger la base de données selon les critères définies par l'acteur qui la détient. Cela lui permet d'avoir à la fois un contrôle plus important sur les données de son site mises à disposition de savoir comment et en quelle quantité les informations de sa base sont utilisées. Pour l'utilisateur, passer par l'API, c'est avoir la garantie que les informations seront à jour et disponibles facilement, cela limite aussi la compétence technique nécessaire pour accéder aux informations car la couche API permet directement l'interrogation des données<sup>79</sup>.



**Fig. 7 : processus d'appel de lignes de temps dynamique (Apport personnel)**

Cette solution n'est pas nouvelle, la liaison entre Polemic Tweet et MétaData Player fonctionne sur ce principe, l'appel JavaScript reprend la variable de temps du Médiafragment et recherche les tweets référencés sur cette période de temps. Il nous faut juste définir comment étendre ce principe à d'autres systèmes d'annotation.

### 2.3.2. Vers un Mash-up temporel des contenus multimédia ?

Si l'on reprend les facteurs limitant de cette annotation des contenus multimédias, on remarque que ces deux facteurs sont liés à la plateforme de visionnage :

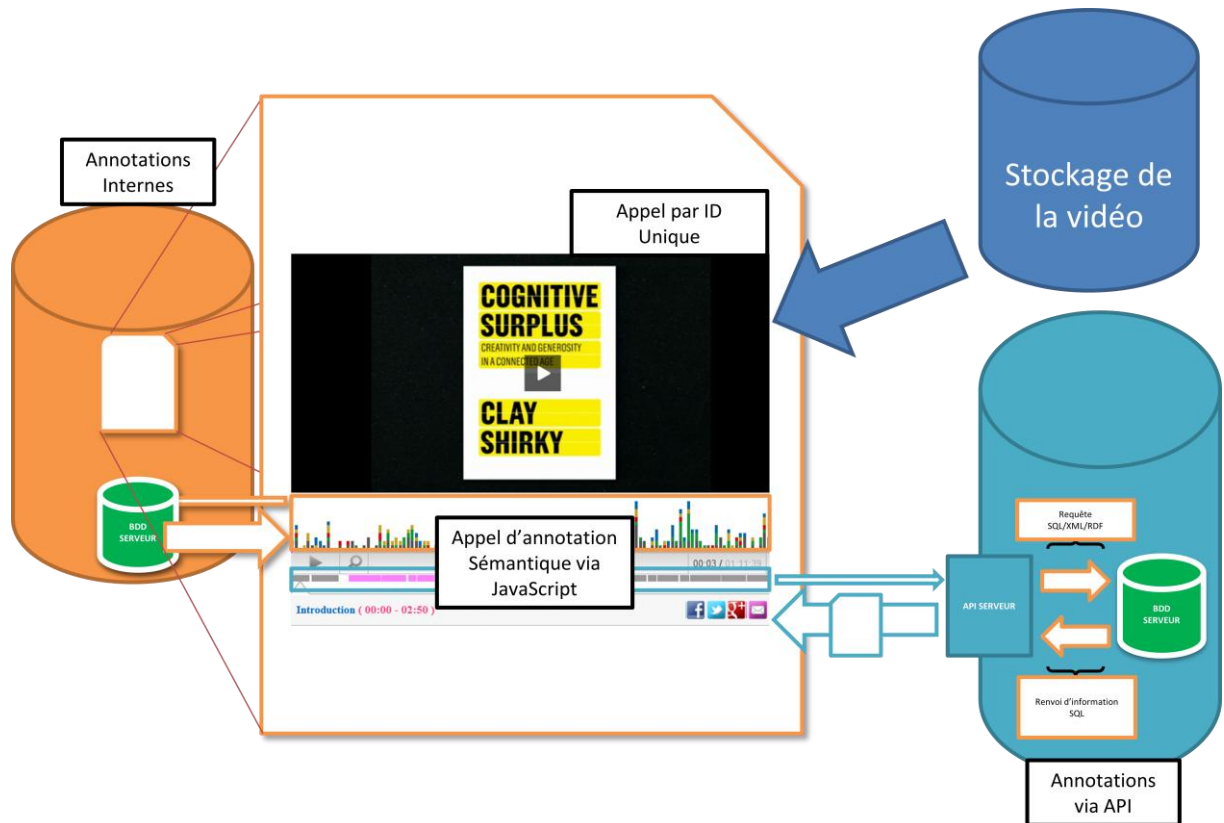
- Faible engagement des personnes lorsqu'il regarde le contenu.
- Mauvaise indexation des contenus

<sup>79</sup> On évite ainsi le passage de requêtage SQL ou affilié pour se focaliser sur les requêtes dynamiques de type Ajax.



L'outil Métadata Player permet déjà de faire appel à plusieurs descriptions du contenu, il s'agit de reproduire cette annotation du contenu à partir de web service ou d'API tiers. On pourrait ainsi utiliser des informations d'autres plateformes d'annotation et les lier au contenu en question grâce aux deux clés : l'ID de la vidéo (par exemple de type : oHg5SJYRHA0 pour YouTube) et l'ID du médiafragment (ancre ou appel Ajax).

De cette manière on pourrait appeler un nombre illimité d'axe d'annotation sur le même lecteur, chaque axe étant lié à sa source d'annotation, soit interne, soit externe via une API.



**Fig. 8 : Intégration modulaire des lignes de temps sémantiques sur un lecteur multimédia intégrant les média fragments (Apport personnel)**

Ce système pourrait aussi favoriser des plateformes d'annotations différentes des plateformes de visionnage (i.e. L'annotation lors de la conférence contre le visionnage sur Métadata Player). C'est-à-dire un endroit spatial et/ou temporel différent de celui de la vidéo qui serait plus propice à cette annotation.

Un des avantages de cette vision du contenu est qu'elle permet à l'ensemble des utilisateurs de créer leur vision de la vidéo à la fois pour eux (narrow) et à la fois pour les autres (broad) créant ainsi une indexation de ces vidéos. Cette façon de faire permet d'enrichir chaque fragment de la vidéo



d'autant de vision de la vidéo qu'elle a d'axe de lecture. Cela permet d'avoir à la fois la lecture brute par le contenu simple mais aussi de permettre l'utilisation de cette vidéo dans un concept.

D'autre part, ce système permet d'esquiver les problématiques de droits d'auteurs des contenus multimédia. Comme les vidéos ne sont pas présentes sur le serveur propriétaire, rien n'empêche de l'afficher du moment que celles-ci sont en ligne. L'annotation elle est permise et légale car elle reste la propriété de celui qui l'a mise en place. Nous allons donc nous intéresser aux possibilités de Mash-up qu'entraîne cette nouvelle vision et comment favoriser leur usage.

### 2.3.3. Quelles informations pour ces annotations ?

Premières étapes à ce recoupement d'informations, définir qu'elles seraient les axes d'analyse que l'on pourrait faire définir automatiquement par des algorithmes. Le découpage en plan par exemple est facilement décryptable par des programmes et cela permettrait de la créer une première structure autour de ces contenus. Youtube tente depuis peu d'ajouter des « transcripts » : traduction automatique de la voix humaine dans la vidéo. Ce concept aussi pourrait être mis en relation de manière dynamique avec la vidéo. Cela permettrait de faciliter la vision des contenus multimédia par les malentendants et permettrait d'annoter les vidéos avec les discours ou les paroles des protagonistes.

L'autre avantage de cette nouvelle formalisation des contenus est qu'il est plus simple d'intégrer un contenu vidéo à partir de requêtes automatique. Les précisions sur les auteurs, les lieux<sup>80</sup> sont autant de paramètres accessibles depuis l'API de la plateforme de visionnage. Il s'agit donc maintenant de trouver des intérêts à cette annotation et des domaines qui tireraient avantage de ces Mash-up.

D'autre comme Bachimont diront que l'annotation est une finalité en soi, qu'il faut donc donner aux utilisateurs la possibilité de se servir des outils pour qu'ils s'en servent. Ce phénomène s'est produit lors de la libération des données par le mouvement Open Data et il y a donc de forte chance que cela se reproduise avec la libération d'annotation sur ces contenus multimédia.

Grâce à ces nouvelles visions de la consommation vidéo, on ajoute ce contenu dans les contenus lisibles et annotables par le web. La multiplication des axes de lectures représentent autant de facettes de la vidéo qui ne demande qu'à être présenté aux utilisateurs. Ils représentent tous à la fois une mine d'information pour l'indexation de ces contenus par les robots classiques des moteurs de recherches mais aussi pour le web sémantique et les mises en relations qu'il représente. La problématique restante est donc : comment utiliser et enrichir cette information produite par les utilisateurs ?

---

<sup>80</sup> Youtube intègre par exemple un module Google maps dans les outils d'annotations des vidéos

### 3. Vers une indexation socio-sémantique du Web

Les outils développés permettent maintenant de franchir cette première étape qu'était l'annotation par des technologies sémantiques. Il reste néanmoins que ces vidéos peinent à dépasser ce stade et que les solutions d'indexation par les moteurs de recherches ne semblent pas permettre pleinement de jouer ce rôle. Les Mash-up permettaient de s'orienter vers cette vision de liaison des données mais de manière incomplète : l'accès direct aux contenus n'est toujours pas accompli.

#### 3.1. Passer du Web 2.0 au Web 3.0

A la différence des moteurs de recherche classique et de leur indexation, le concept de web sémantique possède dans sa conception des atouts à l'indexation des contenus multimédia et plus encore lorsqu'on ajoute la couche médiafragment. Il faut seulement arriver à faire passer les informations du stade « 2.0 » au stade 3.0 c'est-à-dire, au stade sémantique

##### 3.1.1. Avantage du web sémantique sur l'indexation classique

Le premier avantage c'est que le système de compréhension des contenus multimédia se base sur la logique sémantique. Alors que les moteurs de recherche actuels s'adaptent à ces nouvelles façons de décrire les contenus, les robots sémantiques pourront parcourir et indexer ce contenu sans avoir besoin de Framework et autres béquilles. Nous aurons ainsi des banques de données RDF sur lesquelles il sera possible de faire directement des requêtes SparQL.

La compréhension des données donc au bénéfice du web sémantique. La fragmentation du contenu est aussi plus facilement appréhendable du côté du web sémantique : alors que les moteurs de recherche vont poser problème sur cette notion liée à des contenus dynamique, l'appel des fragments par une URL unique va permettre de les définir plus facilement et de les considérer comme des URI c'est-à-dire comme un identifiant unique pour un fragment unique.

Ces fragments disposent aussi d'une structure extensible sur laquelle un certain nombre d'objets sémantiques peuvent être identifiés (Cf. Schema.org et Facebook Open Graph). Ils sont aussi assez souples pour accueillir d'autres types d'annotations ou d'ontologies.

Mais le web sémantique dispose aussi d'une réserve d'information héritée de la vision 2.0 qui ne n'ont toujours pas été traduits : les tags. Ces éléments sont les seuls qui pour l'instant ne sont pas encore reliés à une ontologie. Jusqu'alors ces contenus générés par les utilisateurs étaient considérés par les acteurs du web comme peu strict, souvent doublonnés et donc porteurs d'un sens limité. Pourtant c'est justement sur ce contenu la qu'une vision ontologique du contenu à le plus d'intérêt.

### 3.1.2. Capturer et transformer les données du web 2.0

Il y aurait presque une ironie à ne pas utiliser l'indexation des utilisateurs pour comprendre le contenu qu'ils ont créés car il s'agit bien du point fort qu'a permis le web 2.0<sup>81</sup>. C'est d'ailleurs parce que cette annotation était simple et peu coûteuse et représentative des avis des utilisateurs qu'elle a si bien marché<sup>82</sup>. Si aujourd'hui cette annotation gérée par les utilisateurs ne suffit plus, il faut tenter de reproduire ces qualités dans les nouveaux outils si l'on veut les rendre performant et attrayant. Lorsque nous avons défini les folksonomies et leur capacité à décrire le contenu, un certain nombre de critiques s'élevaient contre le caractère chaotique de ce rangement de l'information. Nous allons démontrer que ce web sémantique a la capacité d'éclaircir ces zones d'ombres causées par les folksonomies et qu'il s'agit là d'une de ses plus grandes forces.

Il faut considérer que les tags ont de fortes similarités avec les instances des concepts dans les ontologies. Les tags « sémantiques » « sémantique » « semantique » peuvent être ainsi considérés comme faisant partie de la classe « tag : Semantic » de l'ontologie *Tag Ontology* (Gruber 2007) elle-même *PartOf* (lien de subsomption) de la classe *skos:Concept* :

```
<rdf:description "sémantique">
  < tag:Semantic>
    <skos: http://example.org/tag/Semantic>
  </tag>
</rdf>
<rdf:description "semantique">
  < tag:Semantic>
    <skos: http://example.org/tag/Semantic>
  </tag>
</rdf>
<rdf:description "sémantiques">
  < tag:Semantic>
    <skos: http://example.org/tag/Semantic>
  </tag>
</rdf>
```

Grace à cet état d'esprit on peut donc considérer les tags comme des triplets RDF : C'est donc une personne (**sujet**) qui effectue cette action de taggage (**prédicat**) sur un contenu ou un fragment de vidéo (**objet**). Néanmoins, comme le signale Alexandre Passant, la notion de signification des tags n'est pas

---

<sup>81</sup> "Chief among those rules is this : build applications that harness network effects to get better the more people use them" (O'Reilly 2006)

<sup>82</sup> "All of these technologies [(Google Page Rank), directories (DMOZ, also used in Google). NdR] use simple structures created by human authors or readers; none uses content formalized by knowledge engineers" (Benel, Zhou et Cahier 2010)

encore transmise par cette reconnaissance des tags de manière sémantique<sup>83</sup>. Il faut donc y rajouter une quatrième information qui rapporte la notion de contexte d'utilisation du tag (Passant 2009) :

*Tagging (Utilisateur, Ressource, Tag, Signification)*

Pour étayer cette signification de manière non ambiguë, nous utiliserons l'ontologie MOAT (Meaning of a tag). Celle-ci se base sur des triplets de base de connaissance comme DBpédia (ontologie basée sur les articles Wikipédia) pour permettre la dissociation de sens des tags<sup>84</sup>

Cette capacité de l'ontologie à dépasser le problème d'ambiguïté peut ainsi être dépassé et se combiner avec les points forts des autres ontologies :

Ontologie	Format	Supporte la modélisation de					
		Tag	Tagging (simple)	Tagging (tripartite)	Agent	Nuage de tags	Signification
Gruber	N/A	✓	×	✓	✓	×	×
Tag Ontology	OWL-Full	✓	✓	✓	×	×	×
SCOT	OWL-Full	✓	×	×	×	✓	×
NAO	RDFS	✓	✓	×	✓	×	×
TagOnt	OWL	✓	×	✓	✓	×	×
Echarte	OWL-DL	✓	×	✓	✓	×	×
SKOS Core	OWL-Full	✓	✓	×	×	×	×
SIOC	OWL-Lite	✓	✓	×	✓	×	×
Annotea	RDFS	✓	✓	×	×	×	×
Taxonomy	RDFS	×	✓	×	×	×	×
rel-tag	Microformat	×	✓	×	×	×	×
MOAT	OWL-DL	✓	✓	×	×	×	✓
Tag Ontology + SCOT + SIOC + MOAT	OWL-Full	✓	✓	✓	✓	✓	✓

**Fig. 9 : description des caractéristiques des différentes ontologies d'annotations sociales (Passant, Using Ontologies to Strengthen Folksonomies and Enrich Information Retrieval inWeblogs 2007)**

En combinant ainsi ces quatre ontologies, l'ensemble des possibilités offertes par les outils de folksonomies classique sont offerte aux possibilités du web sémantique :

- Pour l'action de Tagging : *Tag Ontology* tel que présenté ci-dessus

<sup>83</sup> L'ontologie devient alors un support à la folksonomie, permettant d'associer souplesse de l'annotation par tag et langage formel pour représenter sans ambiguïté et de manière interprétable les significations associées à ces tags. Ce processus nous permet ainsi d'établir un lien fort entre ces différents degrés de formalisation, offrant chacun des perspectives différentes en termes d'annotation et de recherche d'information comme le souligne [Zacklad, 2007] (Passant 2009)

<sup>84</sup> L'exemple repris par A. Passant concerne l'ambiguïté qu'il y a entre le fruit Apple et l'entreprise Apple. Grâce aux URI des ressources DBpedia, on peut distinguer l'entreprise (dbpedia:Apple\_Inc) du fruit homonyme (dbpedia:Apple) (Passant, Using Ontologies to Strengthen Folksonomies and Enrich Information Retrieval inWeblogs 2007)

- Pour la reconnaissance des utilisateurs et leurs liaison : SIOC (Semantically-Interlinked Online Communities)
- Pour la capacité à définir des nuages de mots-clés : l'ontologie SCOT
- Enfin pour la capacité à supprimer l'ambiguïté des tags ; l'ontologie MOAT

L'intégration de l'ontologie MOAT permet donc de faire entrer plus profondément le web sémantique dans les pratiques actuelles d'annotation par les utilisateurs. Ce modèle ne permet pas de représenter l'ensemble des informations<sup>85</sup> que représentent ces tags (Monnin 2009) mais ajoute une brique supplémentaire à l'indexation sémantique de ces contenus multimédia.

Cette façon de concevoir les tags représente tout de même une avancée et le concept de désambiguïsation des contenus par les ontologie de représentation de la connaissance à été ajouté au Framework Common tag<sup>86</sup> développé par le consortium Yahoo, Zemanta Freebase (moteur de recherche sémantique de Google) mais aussi par Syndice, organe de rangements des connaissance de l'université de Southampton, les créateurs de Synote.

Les effets de l'ontologie sur l'ambiguïté des tags ne semblent pas être le seul avantage de la solution, l'effet sur l'hétérogénéité des tags pour un même contenu semble être nettement limité grâce à cette solution<sup>87</sup>.

### 3.1.3. Faciliter l'ajout de ces tags

Dès lors il s'agit toujours de ne pas brusquer l'utilisateur dans ses usages et de mettre en adéquation ces pratiques avec les nouvelles solutions proposées par cette Ontologie. L'objectif consiste donc à la fois à présenter la nouvelle annotation sans que celle-ci ne soit considérée par les utilisateurs comme trop complexe.

Un certain nombre de critère que nous allons lister devrons être conservé afin de faciliter l'appropriation de l'outil par les utilisateurs :

- Mise en place d'une architecture de participation :

Une telle architecture doit permettre le partage des significations au sein d'une communauté, de la même manière que les plates-formes classique de systèmes à base de tags permettent à tous de bénéficier des apports de chacun (auto-complétion, suggestion...).

---

<sup>85</sup> Une des limitations qui pour l'instant n'a pas été surmontée prend la forme d'appel de ressources externes au contexte dans ces tags (allusion, références et autres usages non-dénotationnels)

<sup>86</sup> <http://commontag.org/>

<sup>87</sup> Dans l'étude de cas effectué par A. Passant : « 205 URIs sur les 715 recensées sont sujettes à des problèmes d'hétérogénéité. 96 instances ont ainsi été associées à deux tags, 70 à trois d'entre eux et 39 à quatre tags ou plus. » Il note en particulier la capacité de MOAT à supprimer les synonymes et les traductions de thème dans les tags.

- Facilité de l'ajout de nouveaux tags :

Un aspect qui semble important quant à l'utilisation de MOAT dans ce contexte d'écosystème sémantique est l'utilisation d'instances créées par les wikis sémantiques afin de définir les significations associées aux tags

- Génération automatique des triplets RDF :

Une fois le choix validé, le client produit automatiquement l'ensemble des annotations RDF relatives à l'indexation sémantique du contenu annoté.

- Facilité l'appel aux tags :

Mettre les annotations sur un serveur spécifique qui va stocker l'ensemble des tags utilisés au sein d'une communauté donnée ainsi que les significations globales associées à ceux-ci, i.e. les URIs des concepts signifiants ;

- Ajout de tag guidé par le serveur :

L'utilisateur va choisir parmi cette liste le concept correspondant à l'utilisateur va choisir parmi cette liste le concept correspondant à son tag dans ce contexte particulier d'annotation. Si rien ne convient, il a la possibilité de définir une nouvelle signification ;

### **3.2. Construction d'une page à contenu multimédia enrichis**

Nous allons maintenant organiser les différentes solutions que nous avons étudié en vue de construire un outils capable d'adresser une partie importante des problématiques que rencontre le contenu multimédia dans son indexation.

#### **3.2.1. L'indexation sémantique favorise la sérendipité**

En septembre 2010, Eric Schmidt avait mis en branle la blogosphère pour ses déclarations sur la capacité de Google à fournir le contenu désiré ainsi que le contenu non désiré mais pertinent (Techcrunch.com 2010). Cette dernière caractéristique appelée aussi sérendipité consiste à proposer aux utilisateurs un contenu plus proche de leurs centres d'intérêts alors qu'ils cherchaient des informations au hasard. Avec le développement des nouvelles technologies, cette capacité à fournir devient d'ailleurs une composante importante de la puissance des outils. Le cross-marketing et autres outils de suggestion permettent de proposer aux utilisateurs un contenu pertinent non connu et la capacité à rapprocher ces éléments des consommateurs est d'ailleurs la clé de voute du succès d'Amazon.com.

Ce système de reconnaissance de contenu et de création de chaîne fonctionnait déjà correctement sur les sites de musique pour proposer de nouvelles chansons aux utilisateurs : Comme les styles de musique sont limités en nombre (Spotify compte ainsi 26 genres de musique, Grooveshark 23 et Deezer 12<sup>88</sup>). Ce faible nombre de catégories permet de proposer facilement des contenus similaires et il nous faut donc nous interroger ce web sémantique sur sa capacité à ajouter cette composante au visionnage des vidéos.

Cette capacité à fournir des contenus inconnus et pertinents est un point important sur lequel la technologie sémantique sera jugée. On peut préjuger qu'avec l'ajout de liaison et la disparition des ambiguïtés, cette composante sera nativement exploitable grâce aux annotations. Des organismes comme la Bibliothèque nationale française ou le centre George Pompidou considèrent qu'il s'agit là de la meilleure façon de rapprocher l'utilisateur du contenu et ainsi d'augmenter le temps passé sur le site (lemondeinformatique.fr 2012).

Nous intégrerons cela de deux manières de deux manières différentes : Une navigation active de l'utilisateur en fonction d'un contenu suggéré et sémantiquement lié (cf. fig. 6). L'utilisateur aura ainsi la possibilité de choisir le contenu qu'il veut continuer de regarder ou bien de le laisser guider par l'outil de suggestion de contenu.

On pourrait ainsi penser au module de suggestion de vidéos qui pourrait se baser sur les tags génériques de la vidéo et lier les contenus proposés avec les moments de la vidéo les plus opportuns. En fonction des classes (concepts) de l'ontologie appelée par les tags dans les fragments, on pourrait définir les vidéos et même les fragments vidéo les plus proches du contenu. Ces concepts pourraient être aussi utilisés pour définir l'appel de publicité non plus sur des mots-clés mais sur des concepts de l'ontologie et ces extraits vidéo pourraient ainsi être intégrés entre deux fragments (entre deux unités de sens) de la même manière que le propose la vidéo classique.

### 3.2.2. Intégration de tags sémantiques dans la fragmentation des contenus

Nous repartirons ici de la structure telle que celle-ci a été décrite dans la partie (2.3.2). Nous avons donc un contenu multimédia hébergé sur un serveur tiers ainsi que des axes de lectures définis par des médiafragments.

Le conteneur d'information le plus ancien et le plus général sera donc la page en elle-même. Celle-ci contiendra donc la présentation de la vidéo dans sa description la plus classique. Pour faciliter la recherche classique et sémantique, nous indexerons la vidéo par le biais des langages Schema.org ainsi

---

<sup>88</sup> Score basés sur les propositions de radio automatiquement générées

que les mêmes balises mais cette fois-ci pour Facebook par le biais de l'Open Graph. Cette partie là est intéressante car il permet déjà de se mettre au niveau de plateforme comme Youtube en termes d'indexation de contenu pour les moteurs de recherches externe (cf. Annexe 2. Annotation automatique des vidéos sur YouTube).

```
<div id="watch-container" itemscope itemtype="http://schema.org/VideoObject">
  <meta itemprop="name" content="Chaton trop mignon">
  <meta itemprop="description" content="Un chaton qui écarte les pattes quand on
arrête de lui caresser le ventre. Trop mignon :) Et pour celles et ceux qui se poseraient
la question, non il ne s&#39;a...">
  <meta itemprop="duration" content="PT0M17S">
  <meta itemprop="unlisted" content="False">
  <meta itemprop="paid" content="False">
  <span itemprop="author" itemscope itemtype="http://schema.org/Person">
    <link itemprop="url" href="http://www.youtube.com/user/virginviola">
  </span>
  <meta itemprop="playerType" content="Flash">
  <meta itemprop="width" content="480">
  <meta itemprop="height" content="360">
  <meta itemprop="isFamilyFriendly" content="True">
  <link itemprop="embedURL"
href="http://www.youtube.com/v/eddYjWSZvp8?version=3&autohide=1">
```

Nous nous intéresserons donc maintenant aux informations liés aux fragments de temps, qui représentent une véritable avancée dans la compréhension des contenus par les robots. Alors que la page représente la vision globale des thèmes de la vidéos les prochains éléments, sous-classe de cette vision, affinerons la qualification du contenus par les spécificités du moment de la vidéo ou de la bande son. Nous ajouterons aux lignes de temps définies précédemment les tags de l'ontologie MOAT décrit dans la partie (3.1.2). Nous nous baserons en particulier sur la formalisation Schema.org pour intégrer ces nouveaux formats de tags nous focalisant sur les balises dont le contenu varie en fonction la vidéo. Nous laisserons donc de coté les détails techniques qui seront défini dans les méta-données de la page elle-même. Il en sera de même pour la description que demanderai trop de temps de rédaction à l'utilisateur pour être indexée sur un moment précis de la vidéo. Ces informations seront générées à la volée grâce à un appel d'API :

```
<link itemprop="url"
href="http://www.youtube.com/watch?v=eddYjWSZvp8H!#t=00:00:01.000,00:00:14.000">
  <link itemprop="thumbnailUrl"
  <span itemprop="keywords" itemscope itemtype="moat:"http://moat-project.org/ns#">
  <link itemprop="tag" content="Cat">
```



```

<link itemprop="moat:Meaning" content="http://dbpedia.org/resource/Cat">
<link itemprop="moat:meaningURI" content="http://dbpedia.org/resource/Cat">
</span>
href="http://i2.ytimg.com/vi/eddYjWSZvp8!#t=00:00:01.000,00:00:14.000/hqdefault.jpg">
  <span itemprop="thumbnail" itemscope itemtype="http://schema.org/ImageObject">
    <link itemprop="url"
href="http://i2.ytimg.com/vi/eddYjWSZvp8!#t=00:00:01.000,00:00:14.000/mqdefault.jpg">
    <meta itemprop="width" content="320">
    <meta itemprop="height" content="180">
  </span>
  <link itemprop="embedURL"
href="http://www.youtube.com/v/eddYjWSZvp8!#t=00:00:01.000,00:00:14.000?version=3&auto
hide=1">

```

Nous avons de cette manière le regroupement des avantages de l'utilisation des MicroData de Schema.org, la capacité d'indexation des fragments par le Framework Ajax défini par Google ainsi que l'indexation sémantique du contenu par les tags enrichis.

### 3.2.3. Présentation du lecteur multimédia enrichi

Le lecteur multimédia présentera donc la vision du contenu classique, la capacité à naviguer parmi plusieurs axes temporels et à commenter la vidéo en fonction de ces derniers (Ces éléments sont présents en orange sur la description ci-dessous, on notera qu'à la différence de Polemic Tweet cette annotation se fera pendant le temps de visionnage de la vidéo en question). Enfin, le module de suggestion ainsi que l'ajout d'une publicité ciblée sont représenté en bleu ci-dessous :

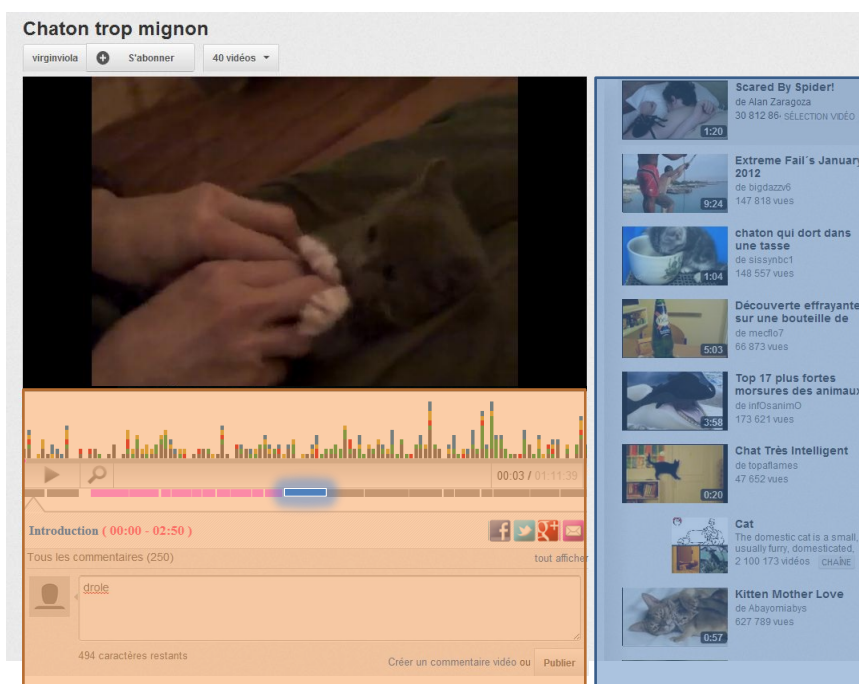


Fig. 6 : organisation des blocs au sein du nouveau lecteur multimédias (Apport personnel)

#### 3.2.4. Objectif du lecteur multimédia sémantique

Pour finir sur la solution proposé, nous repasserons les objectifs à atteindre et sur les solutions que ce lecteur propose. Premier point et non des moindre, l'objectif de ce lecteur est de s'intégrer dans les solutions déjà existantes avec le moins de modifications possible à faire. L'architecture modulaire - bloc lecteur basé sur Metadata Player, bloc d'annotation sémantique et bloc de suggestions par l'annotation sémantique sont indépendants et peuvent être mis en place selon le degré de technicité et de détails souhaités. De la même manière cette solution se base sur un contenu existant et une annotation (taggage et folksonomie) déjà présent et ne cherche pas à travailler ex nihilo.

Deuxième objectif, la favorisation de la recherche par les moteurs de recherches classiques : grâce à l'intégration des informations sous la forme de méta description Schema.org ou Facebook Open Graph, les moteurs de recherches ont désormais plus d'accroche, de contenu pour indexer les vidéo, si l'on ajoute à cela la tentative d'intégrer l'indexation des médiafragments dans ces résultats de recherches, cela aura à terme un impact fort sur ces contenus qui passeront d'une utilisation encore fortement liée aux plateformes de stockage vers une utilisation et une consommation directement liée aux moteurs de recherche ou au site ayant développé les outils spécifiques d'utilisation de ces métadonnées .

Troisième objectif, cette solution s'appuie sur le constat d'une indexation lacunaire des contenus en particulier vidéo qui a comme principale conséquence une exploitation partielle de l'effet de longue traîne. Par une annotation et un taggage plus précis limitant les ambiguïtés de langages, par une suggestion plus précise des contenus, les éléments indexés devraient être mieux exploités dont ces éléments qui étaient jusqu'alors dans l'ombre des grands comptes.

Enfin, point le plus important, l'objectif de l'outil est de favoriser et d'améliorer l'expérience visiteur lorsqu'il visionne et écoute ces contenus. Parce que ce contenu est riche, les visites de ces utilisateurs sont à la fois longues et répétées<sup>89</sup>.

### 3.3. Applications et réserves

Si le web sémantique est une idée ancienne, son adoption et son utilisation sont encore récente et son avenir est encore flou. Si de récente annonce comme la création du Facebook Open Graph ou bien encore du Google Open graph, il reste encore un long chemin avant que cette façon d'appréhender le contenu se démocratise totalement.

---

<sup>89</sup> Le temps moyen d'une visite sur Youtube est de 25 minutes par visites sachant qu'un visiteur passe en moyenne 14 fois par mois sur la plateforme (5,8 h par mois) (Pingdom.com 2011) contre 1.8 h pour Dailymotion.

### 3.3.1. Les Obstacles à surmonter

Comme l'a montré Facebook, un des biais pour cela est de faciliter l'accès techniques aux ressources et surtout à des outils de création d'annotation sémantique automatique (API, Plug-in) de manière à ce que les utilisateurs fasse une transition indolore du web 2.0 à ce Web 3.0 tant attendu<sup>90</sup>.

Autre point intéressant, il n'y a à ce jour qu'un seul moteur de recherche sémantique réellement performant : Wolfram Alpha. Si ce moteur de recherche possède des atouts indéniables pour ce qui est de la recherche en langage naturel, il souffre encore de faiblesse dans les requêtes flous. La période de transition vers ce web sémantique se fera donc avec les moteurs de recherche actuels, dans la dépendance de leur acceptation de ces informations<sup>91</sup>. Pour faire la pirouette, l'indexation de ces contenus se fera pas le web sémantique mais par pour ce web sémantique.

L'annotation est aussi un point important à surveiller, les informations qui peuvent être ajoutées tout au long de la mise en ligne de la vidéo représentent un enjeu crucial pour les créateurs de ce contenu et les moteurs de recherche car sans cela, leurs outils ne permettent pas de résultat pertinent. Or il est très compliqué de pousser les gens à s'exprimer et il s'agira donc d'un défi au web social et au web sémantique de pousser les gens à commenter les contenus.

Enfin, cette vision fragmentée doit aussi être acceptée en tant que format pour les utilisateurs. Ce nouvel attribut peut bouleverser notre vision du contenu qui était jusqu'à un et indivisible. L'idée de fragmentation facilite la navigation mais détruit une partie de la narration. Il faudra donc définir si cette nouvelle navigation peut s'appliquer à l'ensemble des contenus ou si elle peut être utilisée que dans des cas spécifiques.

### 3.3.2. Intégration de ces modules dans les TV connectées ?

Il faut voir aussi voir des avantages à cette nouvelle façon de comprendre et de trier le contenu offre de multiples nouveautés. On a vu apparaitre depuis quelques années d'importantes avancées dans le domaine audiovisuel et surtout de la gestion des contenus directement sur le téléviseur. Qu'il s'agisse de l'arrêt du direct, de la possibilité d'enregistrer le contenu ou bien encore de passer d'un programme à l'autre, la télévision est en train de devenir intelligente. Il y a donc de forte chance pour

---

<sup>90</sup> Facebook est la seule plateforme à proposer nativement l'enrichissement sémantique via son outil pour les développeurs (cf. fig.2)

<sup>91</sup> Il faudrait par exemple mesurer la capacité de nuisance de Google Panda sur ces contenus fragmentés car il s'agit de répétition d'information la plupart du temps.

que les chaînes de télévisions s'intéressent d'ici peu à l'organisation des contenus par le web sémantique.

Le zapping avait commencé à fragmenter la ligne de temps des chaînes de télévision et donc d'affaiblir leur capacité de prescription du contenu, ces nouveaux outils de navigation ne vont pas forcément changer les habitudes audiovisuelles (les heures de grande écoute resteront les mêmes par exemple) mais cela va permettre d'ajouter ce côté interactif à ce média qui était resté jusqu'alors très figé dans sa relation aux contenus télévisuels alternatifs (plateforme de visionnage et web en général).

Pourtant si l'on rapproche le temps passé par les internautes sur les plateformes de contenus multimédia du temps passé sur la télévision, on pourra remarquer que les courbes se rapprochent et finiront par se croiser. On peut donc envisager pour demain des chaînes de télévision qui seraient gérées par des utilisateurs présentant un contenu spécifique selon des mots clés spécifiques et à partir desquels l'utilisateur pourrait son mode de navigation : continuer à regarder le contenu choisi pour lui par cet organisateur de contenu, revenir au contenu défini par la chaîne ou bien encore passer directement par le mode automatique et se laisser guider au gré de ce contenu sérendipitaire

### Conclusion

La création d'un environnement promouvant le web sémantique est donc techniquement faisable, l'annotation par Facebook Google ou une annotation sémantique des tags sont réalisable voire très souhaitable dans une vision à court terme pour la visibilité des contenu. Ces ajouts sont aussi bénéfiques à long terme car la vision sémantique se développe et créer de plus en plus d'outil capable de l'exploiter. Les lecteurs multimédias capables de comprendre les médiafragments en sont ainsi un exemple. Grâce à ces nouveaux outils les contenus s'ouvrent à de nouvelles formes d'enrichissement du contenu qui prennent en compte la dimension temporelle de ces contenus. Dorénavant ces derniers seront plus facilement compréhensibles, ils auront changé de dimension : statique et sans liens, nous sommes passés à une contenu temporellement dynamique et surtout hypertextuel. Le contenu multimédia prend sa véritable place dans le web car il est maintenant lié de lui-même aux autres contenus (lien dans les médiafragments dans ligne de temps, modules de suggestions sémantique qui le relie aux autres vidéos par ses attributs).

L'indexation par le web sémantique que nous avons présenté répond donc à la fois à la faiblesse du web classique dans sa compréhension du contenu et elle permet tout à la fois d'améliorer la liaison intrinsèque de ces vidéos. Ces deux raisons devraient permettre ainsi de diminuer les effets de niches que l'indexation classique n'avait pas dépassées et permettra donc une augmentation de la visibilité des vidéos. Preuve de ces effets de niches, la vidéo d'animation : « Le General » créé par J. Grandjonc (ogHNXWSom3Y) totalise 402 vues sur YouTube contre 34.800 sur Vimeo. Dans un autre registre, les vidéos du testeur Usul (UsulMaster) comptabilise 20.000 vues en moyenne contre 200.000 sur jeuxvideo.com. Parce que YouTube n'a pas les capacités d'une indexation correcte pour des résultats pertinents, ces contenus à forte valeur ajoutée ne sont pas suffisamment exploités. La création d'un outil d'indexation sémantique permettrait donc de limiter la troncature de la longue traîne tout en aplatissant la courbe de cette dernière<sup>92</sup>.

Les objectifs vont donc vers une amélioration des outils d'annotation de ces contenus car ces informations sont le ciment d'une indexation claire des contenus et de leurs relations. Il serait ainsi intéressant de définir l'efficacité des outils d'annotation existants ainsi que les paramètres à prendre en compte pour favoriser cette variable vitale.

Autre point important sur lequel pourrait se former un groupe de travail du W3C, l'intégration de la pensée sémantique dans la mise en place des médiafragments, ceux-ci présentent de nombreux

---

<sup>92</sup> (Anderson 2006) montre ainsi qu'en améliorant l'accès aux contenus, la part des vidéos les plus regardées décroît proportionnellement à la part du contenu qui est mis en valeur.

## Conclusion

avantages mais comme l'avait montré le logiciel Synote. Mais le système actuel d'adressage des fragments par les ancres est mal compris par les moteurs de recherche et demande à être éclairci. Avec la remise en cause des solutions dynamique, il faudrait se pencher sur un modèle qui puisse satisfaire les exigences du HTML 5 tout comme celui de l'annotation sémantique.

## Bibliographie

### Ouvrages & périodiques:

Amar, Muriel, et Bruno Menon. *Bienvenue dans la « gigantesque base de données »*. CAIRN.INFO, 2011.

Anderson, Chris. *The Long Tail: Why the Future of Business is Selling Less of More*. 2006.

Banque Mondiale. «Rapport sur le développement des télécommunications/TIC dans le monde et base de données de l'Union internationale des télécommunications et estimations de la Banque mondiale.» 2011.

Benel, Aurelien, Chao Zhou, et Jean-Pierre Cahier. *Beyond Web 2.0... And Beyond the Semantic Web*. Université de technologie de Troyes, 2010.

Berners-Lee, Tim. *A Representation of Textual Information and MetaInformation for Retrieval and Interchange*. 1993.

*Tags used in HTML*. World Wide Web Consortium., 1992.

Beuscart, Jean-Samuel, Couronné, Thomas Beauvisage : Thomas, et Kevin Mellet. « *La fin de la télévision* », vraiment ? *Synchronisation et recomposition des programmes par les publics en ligne de la Coupe du monde de football*. 2011.

Blondiaux, Loïc. *Le nouvel esprit de la démocratie. Actualité de la démocratie participative*. Seuil, 2008.

Bos, Bert, et Wium Lie Håkon. *Cascading style sheets: designing for the web*. Addison Wesley, 1999.

Brown, Matthew. *Catching up with Structured Data*. Matthew Brown, 2012.

Cha, Meeyoung, Haewoon Kwak, Pablo Rodriguez, Yong-Yeol Ahn, et Sue Moon. *I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System*. 2007.

Chaumond, Julien. *social commerce : Quand le e-commerce rencontre le Web d'aujourd'hui*. Digital Mammouth Editions, 2010.

Di Virgilio, Marion. *La circulation de la musique à l'ère du média numérique en ligne : innovations technologiques et impacts sur la médiation des artistes*. Université Paul Verlaine Metz, 2006.

Eurostat. *Broadband penetration rate*. Eurostat, the statistical office of the European Union, 2010.

Fuchs, Philippe, Jean-Noël gouyet, Jacques Jouhaneau, François Le Chevalier, et Simon Richir. *Formats et gestion des données audio et vidéo*. Techniques de l'ingénieur, 2012.

Gruber, Thomas. *Ontology of Folksonomy : A Mash-up of Apples and Oranges*. International Journal on Semantic Web and Information Systems, 2007.

- Hammond, David. *Web standards and maturing technologies*. 2005.
- Harold, Elliotte Rusty. «Java Network Programming 3rd edition.» 2004: 44-46.
- IPPOLITA. *La face cachée*. Payot & Rivages, 2008.
- ISOBAR. *Panorama 2007 de la vidéo sur Internet", au mois de mai 2007*. 2007.
- Kepeklian, Gabriel, et Jean-Louis Lequeux. *Déployer un projet Web 2.0 : Anticiper le Web sémantique (Web 3.0)*. Editions d'Organisation, 2008.
- Le Deuff, Olivier. *Folksonomies*. BBF, 2006.
- Lee, Tim Berners. *Tags used in HTML*. World Wide Web Consortium., 1992.
- Lefebvre, Alain, et Laurent Poulain. *Cow-boys contre chemin de fer ou que savez-vous vraiment de l'histoire de l'informatique ?* 2011.
- Li, Yunjia, Mike Wald, et Gary Wills. *Let Google Index Your Media Fragments*. 2012.
- Li, Yunjia, Mike Wald, Tope Omitola, Nigel Shadbolt, et Gary Wills. *Synote: Weaving Media Fragments and Linked Data*. 2012.
- Mazuel, Laurent, et Nicolas Sabouret. *Degré de relation sémantique dans une ontologie pour la commande en langue naturelle*. Laboratoire d'Informatique de Paris 6 (LIP6), 2007.
- Monnin, Alexandre. *Qu'est-ce qu'un tag ? Entre accès et libellés, l'esquisse*. 2009.
- Nelson, Ted. *Literary Machines: The report on, and of, Project Xanadu concerning word processing, electronic publishing, hypertext, thinkertoys, tomorrow's intellectual revolution, and certain other topics including knowledge, education and freedom*. Mindful press, 1981.
- Observatoire des professions de l'Image. *Les chiffres officiels 2010 du marché de la photo et de l'image en france et dans le monde*. 2011.
- O'Reilly, Tim. *Web 2.0 Compact Definition: Trying Again*. 2006.
- Oreskovic, Alexei. «Exclusive: YouTube hits 4 billion daily video views.» *Reuters* (Reuters), 2012.
- Passant, Alexandre. *Technologies du Web Sémantique pour l'Entreprise 2.0*. Université Paris IV - Sorbonne & École doctorale V - Concepts et Langages, 2009.
- Pétine, Patricia. *Les principes généraux et les différentes approches de l'urbanisation des SI*. ISC Paris, 2012.
- PhiloWeb . *Entretien avec François Rastier*. 2010.
- Pickering, Jeanne M., Rebecca E. Grinter, et John Le. *Grandeur et décadence d'Arpanet. La saga de Netville, cité champignon du cyberspace*. : Réseaux, 1996, volume 14 n°77, 1996.
- Raymond, Eric S. *The Cathedrale & the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. O'reilly , 2001.



Scheid, François, Grégoire De Montaigu, Renaud Vaillant, et François Scheid. *Le marketing digital*. 2012.

SIPEC. *Bilan et Analyses de l'année 2002*. 2002.

Tran, Sébastien. *Une revue littéraire du Web 2.0*. M-Lab, 2011.

Vasilev, Mike. *History of digital Storage*. Mashable Infographics, 2011.

W3C/MIT, Tim Berners-Lee, Robert Fielding, Day Software, L Masinter, et Adobe Systems. *RFC 3986 Uniform Resource Identifier (URI): Generic Syntax*. 2005.

W3C-Recommendation. «RDF/XML Syntax Specification (Revised)» 2004.

### **Articles internet :**

Agence Mode d'emploi. *Quelle est la part de Facebook dans le trafic des sites média ?* (<http://www.agence-modedemploi.com/buzz/fr/quelle-est-la-part-de-facebook-dans-le-trafic-des-sites-media/>). 2012.

Ars Industrialis. *Grammatisation : techniques de reproduction* (<http://www.arsindustrialis.org/grammatisation>). 2011

Google developers. *Making AJAX Applications Crawlable* (<https://developers.google.com/webmasters/ajax-crawling/?hl=fr>). 2012.

Google Webmaster Central Blog. *Introducing Video Sitemaps* (<http://googlewebmastercentral.blogspot.fr/2007/12/introducing-video-sitemaps.html>). 2007.

YouTube API Blog «Flash and the HTML5<video> tag.», (<http://apiblog.youtube.com/2010/06/flash-and-html5-tag.html>.) 2010.

Times of india, *Facebook beats YouTube, Flickr*. ([http://articles.timesofindia.indiatimes.com/2010-07-09/internet/28297234\\_1\\_facebook-active-users-user-base](http://articles.timesofindia.indiatimes.com/2010-07-09/internet/28297234_1_facebook-active-users-user-base)). 2010.

Google Webmaster Tools : *Creating a Video Sitemap (Video-specific tag definitions)*(<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=80472&topic=10079&ctx=topic#4>). 2012.

Google *Une heure par seconde : Upload record pour Youtube*. (<http://www.onehourpersecond.com>)., 2012.

IRI. *Polemic Tweet : un dispositif tirant parti des dynamiques sociales* (<http://polemictweet.com/about.php>). 2012.

Journal-du-net. *Les tags, cette idée simple que redécouvre le Web*. ([http://www.journaldunet.com/solutions/0502/050208\\_tags.shtml](http://www.journaldunet.com/solutions/0502/050208_tags.shtml)), 2005.

Kesteren, Anne van. «<video> element proposal : <http://lists.whatwg.org/pipermail/whatwg-whatwg.org/2007-February/009702.html>.» [*whatwg*], 2007.

lemondeinformatique.fr. *La BNF et le Centre Pompidou misent sur le web sémantique* (<http://www.lemondeinformatique.fr/actualites/lire-la-bnf-et-le-centre-pompidou-misent-sur-le-web-semantique-48856.html>). 2012.

O'reilly, Tim. «What Is Web 2.0.» (<http://oreilly.com/web2/archive/what-is-web-20.html>), septembre 2005.

Pingdom.com. *Facebook, YouTube, our collective time sinks (stats)* (<http://royal.pingdom.com/2011/02/04/facebook-youtube-our-collective-time-sinks-stats/>). 2011.

Sapientis. *Introduction au Web sémantique et aux ontologies* ([http://www.sapientis.fr/conseil/?page\\_id=2680](http://www.sapientis.fr/conseil/?page_id=2680)). 2010.

Shirky, Clay. *Folksonomies + controlled vocabularies*. [http://many.corante.com/archives/2005/01/07/folksonomies\\_controlled\\_vocabularies.php](http://many.corante.com/archives/2005/01/07/folksonomies_controlled_vocabularies.php), 2005.

TechCrunch. *YouTube Enables Deep Linking Within Videos* (<http://techcrunch.com/2008/10/25/youtube-enables-deep-linking-within-videos/>). 2008.

Techcrunch.com. *Eric Schmidt On The Future Of Search — A Move Towards A "Serendipity Engine"* (<http://techcrunch.com/2010/09/28/eric-schmidt-future-of-search/>). 2010

Wikipedia. *Web sémantique*. 2012.

—. *Wiki*. 2012.

ZDnet.fr. *Google rachète Metaweb, spécialiste des données sémantiques* (<http://www.zdnet.fr/actualites/google-rachete-metaweb-specialiste-des-donnees-semantiques-39753238.htm>). 2010.

## Annexes

### Annexe 1. Taux de participation et d'annotation en fonction du type de contenu

Tiré de « La fin de la télévision », vraiment ? Synchronisation et recomposition des programmes par les publics en ligne de la Coupe du monde de football. »

video_title	audience	commentaires	Commentaires pour 1000 vues
Doit-on utiliser la vidéo pour aider les arbitres ?	6 338	70	11,0
Les français jugent l'équipe de France !	1 255	6	4,8
Revoir le Mag, épisode 08 du 17 juin en intégralité	2 099	9	4,3
La Roja loin de la furia	2 892	11	3,8
Patrice Evra, l'interview exclusive : le capitaine de l'équipe de France nous répond !	30 577	104	3,4
Ribéry dans Téléfoot : les coulisses de son intervention inopinée !	3 653	11	3,0
Lizarazu : "Anelka n'était pas intéressé par le jeu"	11 974	36	3,0
Top Arrêts de la Coupe du Monde 2010 (16/06/2010)	4 210	12	2,9
Blatter : "Les ingérences politiques ne sont pas tolérées"	704	2	2,8
Domenech : "c'est plus que décevant"	10 660	30	2,8
But refusé : Allemagne - Angleterre, 39ème	30 954	72	2,3
En route pour le Brésil	2 336	5	2,1
Retour en catimini pour les Bleus	7 436	14	1,9
Patrice Evra : "Un traître à éliminer du groupe !"	80 117	150	1,9
Thuram : "j'ai proposé qu'Evra ne mette plus les pieds en équipe de France"	55 667	103	1,9
Escalettes : "ce n'est pas ma nature de démissionner"	11 949	22	1,8
La polémique sur les vuvuzelas	6 137	11	1,8
Sondage : les français jugent les bleus	561	1	1,8

### Annexe 2. Annotation automatique des vidéos par YouTube et création de Snippets Google et Facebook

Apport personnel



**Chaton trop mignon**

Un chaton qui écarte les pattes quand on arrête de lui caresser le ventre. Trop mignon :) Et pour celles et ceux qui se poseraient la question, non il ne s'a...

1 sur 1 Choisissez une miniature

Pas de vignette

**Chaton trop mignon - YouTube**

[www.youtube.com/watch?v=eddYjWSZvp8](http://www.youtube.com/watch?v=eddYjWSZvp8)  
 Bloquer tous les résultats de [www.youtube.com](http://www.youtube.com)  
 29 nov. 2009 - 17 s - Ajouté par virginviola  
 Un **chaton** qui écarte les pattes quand on arrête de lui caresser le ventre. Trop mignon :) Et pour celles et ceux ...

```

<meta property="og:url"
content="http://www.youtube.com/watch?v=eddYjWSZvp8">
<meta property="og:title" content="Chaton trop mignon">
<meta property="og:description" content="Un chaton qui écarte
les pattes quand on arrête de lui caresser le ventre. Trop mignon
;) Et pour celles et ceux qui se poseraient la question, non il ne
ss#39;a...">
<meta property="og:type" content="video">
<meta property="og:image"
content="http://i2.ytimg.com/vi/eddYjWSZvp8/mqdefault.jpg">
<meta property="og:video"
content="http://www.youtube.com/v/eddYjWSZvp8?
version=3&autohide=1">
<meta property="og:video:type" content="application/x-
shockwave-flash">
<meta property="og:video:width" content="480">
<meta property="og:video:height" content="360">
<meta property="og:site_name" content="YouTube">
<meta property="fb:app_id" content="87741124305">
<meta name="twitter:card" value="player">
<meta name="twitter:site" value="@youtube">
<meta name="twitter:player"
value="https://www.youtube.com/embed/eddYjWSZvp8">
<meta property="twitter:player:width" content="480">
<meta property="twitter:player:height" content="360">
<meta name="attribution" content="youtube none/>

```

```

<div id="watch-container" itemscope
itemtype="http://schema.org/VideoObject">
<link itemprop="url" href="http://www.youtube.com/watch?
v=eddYjWSZvp8">
<meta itemprop="name" content="Chaton trop mignon">
<meta itemprop="description" content="Un chaton qui écarte les
pattes quand on arrête de lui caresser le ventre. Trop mignon :) Et
pour celles et ceux qui se poseraient la question, non il ne
ss#39;a...">
<meta itemprop="duration" content="PT0M17S">
<meta itemprop="unlisted" content="False">
<meta itemprop="paid" content="False">
<span itemprop="author" itemscope
itemtype="http://schema.org/Person">
<link itemprop="url"
href="http://www.youtube.com/user/virginviola">
</span>
<link itemprop="thumbnailUrl"
href="http://i2.ytimg.com/vi/eddYjWSZvp8/hqdefault.jpg">
<span itemprop="thumbnail" itemscope
itemtype="http://schema.org/ImageObject">
<link itemprop="url"
href="http://i2.ytimg.com/vi/eddYjWSZvp8/mqdefault.jpg">
<meta itemprop="width" content="320">
<meta itemprop="height" content="180">
</span>
<link itemprop="embedURL"
href="http://www.youtube.com/v/eddYjWSZvp8?
version=3&autohide=1">
<meta itemprop="playerType" content="Flash">
<meta itemprop="width" content="480">
<meta itemprop="height" content="360">
<meta itemprop="isFamilyFriendly" content="True">

```

### Annexe 3 : popularité des outils Facebook Open Graph, Schema.org, Vidéo Sitemap sur Google.

Tiré de Google tendance de recherche, sur la période 1<sup>er</sup> janvier 2010 / 1<sup>er</sup> aout 2012

