



**HAL**  
open science

**Mise en oeuvre d'une solution de veille et de recherche  
de l'image par son contenu : cerner les limites, proposer  
des scénarios et anticiper les conséquences du passage  
des mots aux pixels**

Lydie Bennet

► **To cite this version:**

Lydie Bennet. Mise en oeuvre d'une solution de veille et de recherche de l'image par son contenu : cerner les limites, proposer des scénarios et anticiper les conséquences du passage des mots aux pixels. domain\_shs.info.docu. 2012. mem\_00945663

**HAL Id: mem\_00945663**

**[https://memsic.ccsd.cnrs.fr/mem\\_00945663v1](https://memsic.ccsd.cnrs.fr/mem_00945663v1)**

Submitted on 12 Feb 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CONSERVATOIRE NATIONAL DES ARTS ET METIERS

Ecole Management et Société

Département Culture Information Technique et Société (CITS)

INTD

MEMOIRE pour obtenir le

Titre professionnel "Chef de projet en ingénierie documentaire" INTD

RNCP niveau I

Présenté et soutenu par

*Lydie Bennet*

le 14 novembre 2012

Mise en œuvre d'une solution de veille et  
de recherche de l'image par son contenu :

cerner les limites, proposer des scénarios  
et anticiper les conséquences du passage  
des mots aux pixels

Jury : Sylvie Breton  
Françoise Profit

**Promotion 42**

*À ceux qui m'ont soutenue  
et à celui qui n'a pu le faire.*

## Remerciements

Je remercie l'équipe du service Achat d'art qui m'a accueillie et qui a toujours été disponible pour répondre à mes questions. Je remercie Sylvie Breton, ma responsable de stage, qui m'a accompagnée, au sens propre comme au sens figuré, et dont je garde les conseils et l'enthousiasme. Je remercie également Françoise Profit qui a levé mes incertitudes quant au mémoire et aux priorités à l'issue de cette formation.

Je voudrais aussi remercier l'équipe pédagogique qui m'a permis de découvrir un « nouveau monde » passionnant et de confirmer ainsi ma démarche de reconversion professionnelle. Je voudrais les saluer pour leur accompagnement.

Enfin, je remercie mes camarades de promotion, avec qui j'ai partagé les angoisses mais aussi des moments d'enthousiasme voire de fierté pour le travail accompli. Sans leur présence, cette année n'aurait pas été aussi enrichissante.

# Notice

BENNET Lydie. Mise en œuvre d'une solution de veille et de recherche de l'image par son contenu : cerner les limites, proposer des scénarios et anticiper les conséquences du passage des mots aux pixels. 2012. 1 volume, 105 p. Mémoire, Titre professionnel « Chef de projet en ingénierie documentaire » niveau I, INTD, 2012.

Ce mémoire rend compte de la mission de stage dont l'objectif était d'élaborer une solution de veille sur des images publicitaires produites dans le secteur du luxe. Cette demande a orienté la réflexion vers les systèmes de reconnaissance d'image par le contenu, dont les promesses sont grandes mais les résultats encore décevants, notamment pour une utilisation par le grand public ou dans le milieu professionnel.

Ainsi, afin de proposer des solutions de mise en œuvre de cette veille, le mémoire analyse, dans un premier temps, le contexte de l'entreprise et l'environnement technologique au sein desquels la veille se développe. Puis, dans un deuxième temps, il se concentre sur la problématique de l'image et sur les technologies afférentes. Enfin, dans un troisième temps, il expose deux scénarios de veille et des suggestions organisationnelles pour la mettre en œuvre.

Pour des raisons de confidentialité, le nom de l'entreprise dans laquelle j'ai effectué le stage a été retiré. Certaines informations ont été volontairement supprimées, notamment celles ayant trait à l'organisation et à la stratégie de l'entreprise, comme l'organigramme.

Descripteurs : agrégateur RSS, gestion de la connaissance, image, image fixe, image numérique, indexation automatique, indexation collaborative, indexation libre, indexation manuelle, intelligence économique, moteur de recherche, publicité, recherche d'images, recherche en texte intégral, recherche textuelle, reconnaissance de document, réseau social, RSS, similarité, veille, veille collaborative, veille concurrentielle, veille stratégique, Web 2.0, Web 3.0, Web sémantique, Web social.

# Table des matières

Remerciements .....	3
Notice.....	4
Liste des figures.....	10
Introduction.....	11
Première partie Contextualisation de la mission.....	14
1 La mission au cœur de la problématique.....	15
1.1 L'entreprise et le rôle des documentalistes.....	15
1.1.1 Une entreprise internationale .....	15
1.1.2 Mon positionnement dans l'entreprise.....	15
1.1.2.1 La Direction Artistique et le service Achat d'art.....	15
1.1.2.2 La place de la documentation.....	16
1.2 Le projet à réaliser .....	16
1.2.1 L'enjeu .....	16
1.2.2 Une consommation ostentatoire .....	17
1.2.3 Le projet.....	17
2 L'analyse de la situation.....	18
2.1 L'analyse de l'existant dans l'entreprise .....	18
2.1.1 Questionnaire et entretiens.....	18
2.1.2 Les entretiens avec les services Marketing et R&D.....	18
2.1.2.1 Le Marketing .....	18
2.1.2.2 Le service Recherche et Développement.....	19
2.2 L'analyse des besoins du demandeur.....	20
2.2.1 Trois besoins identifiés .....	20
2.2.2 De fortes contraintes détectées.....	21
3 La veille sur Internet .....	22
3.1 Définitions et circonscription de la veille idoine.....	22
3.1.1 La veille : définitions et démarches.....	22

3.1.1.1	La définition et les recommandations du rapport Martre .....	22
3.1.1.2	La norme AFNOR .....	22
3.1.1.3	La démarche d'une activité de veille .....	23
3.1.2	La circonscription de la veille .....	23
3.1.2.1	Des ajustements au cadre du projet .....	23
3.1.2.2	Une veille créative .....	24
3.1.2.3	Une veille web .....	24
3.2	Le Web 2.0 et 3.0 .....	24
3.2.1	Les définitions d'un environnement en mutation.....	24
3.2.1.1	Le Web 2.0 de Tim O'Reilly .....	24
3.2.1.2	Le Web social.....	25
3.2.1.3	Le Web de données.....	26
3.2.1.4	Le Web <sup>2</sup> .....	27
3.2.2	Les risques .....	27
3.2.2.1	L'infobésité : mythe ou réalité ? .....	27
3.2.2.2	La qualité de l'information .....	28
Deuxième partie Veille et recherche par l'image .....		29
4	Les particularités de l'image-document .....	30
4.1	L'analyse de l'image .....	30
4.1.1	La fiche technique de l'image numérique .....	30
4.1.2	Les concepts de connotation et de dénotation.....	30
4.1.2.1	Le contenu de l'image .....	30
4.1.2.2	La relation de l'image au texte .....	31
4.2	Nommer pour rechercher et retrouver .....	31
4.2.1	La problématique de l'indexation .....	31
4.2.1.1	L'indexation automatique.....	31
4.2.1.2	L'indexation manuelle .....	32
4.2.2	Les mots sont-ils inévitables ? .....	32
4.2.2.1	Le choix des mots .....	32
4.2.2.2	Le manque de pertinence .....	33
5	La reconnaissance par l'image.....	34
5.1	Le Content-Based Image Retrieval (CBIR).....	34
5.1.1	Définitions et explications .....	34
5.1.1.1	50 ans de recherche .....	34

5.1.1.2	Le principe de fonctionnement du CBIR .....	35
5.1.2	Exemples d'applications dans la recherche .....	35
5.1.2.1	Instituts spécialisés (INRIA, INA) .....	35
5.1.2.2	Les expériences et études menées avec un CBIR .....	36
5.2	Le CBIR au défi du monde réel.....	37
5.2.1	Les pratiques de recherche d'images .....	37
5.2.1.1	Les besoins des utilisateurs.....	37
5.2.1.2	La recherche textuelle incontournable .....	38
5.2.2	Les moteurs de recherche généralistes et les images.....	39
5.2.2.1	Le mode d'emploi d'un moteur de recherche.....	39
5.2.2.2	Un exemple caractéristique : Google Search Image.....	39
5.2.2.3	D'autres moteurs : Bing et Yahoo !.....	40
6	La reconnaissance faciale .....	41
6.1	Les technologies mises en œuvre .....	41
6.1.1	Reconnaître des visages .....	41
6.1.2	Les problèmes juridiques comme freins de son essor ? .....	41
6.2	Exemples concrets de reconnaissance faciale .....	43
6.2.1	Google et la reconnaissance des chats.....	43
6.2.2	Scene Tap à La Nouvelle Orléans .....	43
6.2.3	Face.com : la reconnaissance visuelle en réseau .....	44
6.3	Les limites technologiques .....	44
Troisième partie Scénarios pour une veille visuelle .....		46
7	Le rappel du contexte.....	47
7.1	Le cahier des charges fonctionnel.....	47
7.1.1	L'analyse de l'existant.....	47
7.1.2	L'analyse des besoins : les fonctions de l'outil idéal .....	47
7.1.3	Les spécifications applicatives .....	48
7.2	Le cahier des charges technique .....	50
7.2.1	Les spécifications techniques .....	50
7.2.2	L'identification des risques .....	51
8	Le scénario d'amorçage .....	53
8.1	Les solutions pour l'automatisation de la veille .....	53
8.1.1	Le potentiel de l'automatisation.....	53
8.1.2	Le sourcing .....	53



8.1.2.1	Définition.....	53
8.1.2.2	Les annuaires et les répertoires.....	54
8.1.2.3	Le tableau récapitulatif du sourcing .....	54
8.1.3	Les flux RSS et les agrégateurs .....	55
8.1.3.1	Définition et pratiques des flux RSS.....	55
8.1.3.2	La solution retenue .....	56
8.2	Les solutions pour la recherche d'images .....	57
8.2.1	Le moteur de recherche spécialisé.....	57
8.2.1.1	Sa technologie et ses fonctionnalités .....	57
8.2.1.2	Une comparaison de moteurs de recherche .....	57
8.2.2	La recherche par la couleur et la similarité .....	59
8.2.2.1	La recherche par la couleur.....	59
8.2.2.2	La recherche par la similarité .....	60
8.3	Les limites de ce scénario .....	60
9	Le scénario de pérennisation.....	62
9.1	Les logiciels et plateformes de veille.....	62
9.1.1	L'environnement technologique et juridique .....	62
9.1.1.1	Différents produits existants .....	62
9.1.1.2	L'activité de veille est-elle menacée ? .....	63
9.1.2	Les recommandations.....	65
9.1.2.1	Les critères de comparaison.....	65
9.1.2.2	Les résultats .....	66
9.2	La veille collaborative .....	67
9.2.1	Dans quel but ?.....	67
9.2.1.1	Une définition du travail collaboratif .....	67
9.2.1.2	Le Knowledge Management (KM).....	67
9.2.2	La place du documentaliste.....	68
9.2.2.1	Le réseau social d'entreprise : un réseau de veille utopique ?.....	68
9.2.2.2	Le « community manager » du réseau de veille .....	69
	Conclusion .....	70
	Bibliographie.....	74

Annexes .....	88
Annexe 1 Matrice du questionnaire.....	89
Annexe 2 Exemples de visuels du luxe en Parfums-Beauté.....	95
Annexe 3 Accès aux sites Internet en France .....	96
Annexe 4 Exemples d'images problématiques pour la reconnaissance faciale .....	97
Annexe 5 Schémas des processus de veille et recherche selon le diagramme d'activités d'UML (Unified Modeling Language) .....	99
Annexe 6 Sourcing .....	101
Annexe 7 Exemples d'agrégateurs de flux offrant une solution de diffusion.....	102
Annexe 8 Moteurs de recherche .....	104
Annexe 9 Musée de l'Hermitage .....	105

## Liste des figures

Figure 1 - Ce qu'un ordinateur "voit" .....	34
Figure 2 – Organisation du tableau des spécifications applicatives du projet .....	49
Figure 3 – Organisation du tableau des spécifications applicatives des outils différenciés .....	50
Figure 4 – Panorama du marché de la veille.....	62

# **Introduction**

Mon stage s'est déroulé dans l'entreprise X, grand nom du secteur du luxe, dans sa branche Parfums-Beauté. J'ai intégré l'équipe de documentation du service Achat d'art dédié à la gestion des images, de leur conception à leur archivage, au sein de la Direction Artistique.

Mon projet consistait à étudier la faisabilité de la mise en œuvre d'une veille et d'une recherche d'images photographiques produites à l'extérieur de l'entreprise. Cette solution est conçue comme un outil de décision stratégique. La difficulté majeure résidait dans le fait que la veille ne devait pas s'effectuer à partir de mots mais à partir du contenu de l'image.

Puisqu'un service au sein de l'entreprise suit l'actualité des publications visuelles dans la presse et les magazines, le périmètre d'application de la veille s'est donc circonscrit à Internet. En conséquence, je devais élaborer une (ou des) procédure(s) de veille et de recherche sur Internet permettant d'acquérir :

- les créations de photographes, de mannequins connus ;
- celles des nouveaux photographes et mannequins détectés ;
- les créations de la concurrence, avec une attention particulière sur les photographes, les mannequins ; les ambiances, les univers ;
- les tendances.

En premier lieu, la solution devait concerner à la fois les images fixes et animées. Néanmoins, en raison des difficultés techniques rencontrées en matière de reconnaissance par le contenu visuel, j'ai limité le mémoire aux images fixes, même si quelques allusions à la vidéo peuvent apparaître.

En effet, la recherche et la reconnaissance par le contenu des images est un sujet très étudié, par les universitaires et par les grands acteurs de l'Internet. Par reconnaissance visuelle ou reconnaissance par le contenu visuel j'entends l'aptitude d'un programme à identifier les éléments composant une image, à les qualifier, à les indexer. Cette fonctionnalité permet de rechercher une image depuis un visuel et non plus depuis des mots. Par le terme « visuel », je fais référence à l'expression visuelle sous forme d'image.

Lors de mes recherches bibliographiques, j'ai été étonnée de la quantité importante d'articles portant sur ce sujet. Par exemple, dans ScienceDirect<sup>1</sup>, une recherche avancée<sup>2</sup> sur les termes « Content-based image retrieval »<sup>3</sup> AND « Search engine » en « Arts and Humanities », « Computer science » et « Social sciences », depuis 2011 a remonté 489 articles. La recherche est donc dynamique et, puisqu'une découverte chasse l'autre, un algorithme rend obsolète le précédent, la date de publication a été un critère de sélection décisif.

En outre, je me suis abonnée à de nombreux flux RSS de journaux en ligne et de blogs spécialisés pour constituer une veille sur ce sujet. J'ai été aussi surprise du dynamisme de cette actualité que j'imaginai davantage limitée aux laboratoires de recherche. Ainsi, j'ai assez vite compris qu'une course était engagée sur la résolution du problème de la reconnaissance visuelle.

---

<sup>1</sup> ScienceDirect : plateforme d'accès aux revues scientifiques de l'éditeur Elsevier.

<sup>2</sup> Recherche effectuée le 14 juillet 2012.

<sup>3</sup> Content-based image retrieval : recherche fondée sur le contenu de l'image (cf. explication p.34 et suivantes).

Toutefois, comme je le développerai dans la deuxième partie du mémoire, l'algorithme qui permet une telle fonctionnalité n'a pas encore été découvert. Aussi, de nombreuses approches existent mais aucune n'est réellement pertinente ou accessible aux utilisateurs communs d'Internet. De ce fait, les procédures de veille et de recherche se fondent toujours sur une indexation textuelle des données, y compris pour les images. C'est la raison pour laquelle les recherches d'images, y compris par des moteurs de recherche spécialisés, occasionnent des résultats peu voire non pertinents.

En conséquence, comment mettre en place un système de veille des images (de leur contenu) - c'est-à-dire une automatisation de la collecte, de l'analyse et du traitement, de la diffusion et de l'archivage -, soutenu par un moteur de recherche pertinent pour des recherches ponctuelles urgentes, alors que la technologie développée actuellement ne permet pas de satisfaire une telle démarche ?

Ainsi, quelles solutions peuvent pallier cette absence de réponse technique ? Quels sont les risques induits par ces solutions ? S'il n'est pas possible de satisfaire la demande telle qu'elle a été formulée à l'origine de la mission, quels sont néanmoins les points qui peuvent être résolus et amorcer un système de veille, en attendant, peut-être, la découverte d'un algorithme adéquat ?

Pour y répondre, je m'intéresserai dans une première partie au contexte de la mission et du projet fondé d'une part, sur l'analyse de l'entreprise et d'autre part, sur la circonscription du périmètre de la veille demandée. Puis, dans une deuxième partie, j'étudierai le statut de l'image, et ses caractéristiques, et les technologies qui ont été développées, avec plus ou moins de succès, pour permettre la reconnaissance par son contenu. Enfin, dans une troisième partie, je présenterai les scénarios que j'ai proposés pour répondre à la mission du stage.

# **Première partie**

## **Contextualisation de la mission**

# 1 La mission au cœur de la problématique

---

## 1.1 L'entreprise et le rôle des documentalistes

### 1.1.1 Une entreprise internationale

L'entreprise X est une grande marque internationale, positionnée dans le secteur du luxe. Le siège social est installé dans la région parisienne et compte 900 employés. Le déploiement international représente près de 3 000 employés, répartis sur tous les continents.

J'ai effectué mon stage dans la branche Parfums-Beauté, qui s'articule en trois pôles : les parfums, les produits cosmétiques et les soins.

La culture d'entreprise est fondée sur la discrétion, la recherche de l'excellence et la création, trois valeurs issues de son histoire.

Le nom de l'entreprise, c'est-à-dire sa marque, nourrit la légende et engendre un imaginaire visuel alimenté par des campagnes publicitaires qui recherchent toujours l'avant-gardisme.

### 1.1.2 Mon positionnement dans l'entreprise

#### 1.1.2.1 La Direction Artistique et le service Achat d'art

Mon stage relevait de la Direction Artistique (DA). Celle-ci regroupe les créatifs de l'entreprise, c'est-à-dire les personnes qui inventent et rendent concrets les concepts pour :

- les photographies et les films publicitaires ;
- les packagings<sup>4</sup> ;
- l'aménagement intérieur et les vitrines des boutiques ;
- les supports de vente chez les détaillants ;
- le site Internet, etc.

En conséquence, la DA se compose de différents métiers : architectes, designers, graphistes, etc.

J'étais affectée au service Achat d'art dont la mission est de prendre en charge la chaîne de création de l'image publicitaire (fixe et animée) :

- la conception et le choix des campagnes visuelles : choix des thèmes, des photographes, des réalisateurs, des mannequins ;
- la création des visuels : suivi du processus créatif ;
- la mise à disposition des visuels pour les marchés internationaux ;
- la gestion des droits (d'auteur, d'image) ;
- la conservation patrimoniale des images.

---

<sup>4</sup> Packaging : conditionnement et emballage du produit.



En conséquence, ce service a une dimension stratégique fondamentale : il est au cœur des processus de réflexion, de décision et d'action de l'entreprise. Pour mener à bien ses missions, il travaille de manière transversale avec d'autres services comme le Marketing, la Communication, etc.

### **1.1.2.2 La place de la documentation**

Au sein du service Achat d'art, les personnes en charge de la documentation gèrent et conservent le fonds d'images par des outils de Digital Asset Management (DAM) afin de mettre ce fonds à la disposition des collaborateurs, les accès dépendant du profil de chacun. Le DAM est un outil de gestion du fonds photographique ; le logiciel retenu gère les différents formats de fichier, les droits d'auteur, le bilinguisme et se présente avec une double interface, le back et le front-office<sup>5</sup>, ce dernier étant l'intranet de la Direction Artistique [17, Haupais, p.60-61]. L'image occupe ainsi une place centrale, ce qui positionne les personnes qui gèrent son cycle de vie à la confluence de différents métiers.

Ce fonds d'images dématérialisé est très majoritairement constitué de photographies publicitaires, exclusivement celles des campagnes de l'entreprise. On dénombre plus de 20 000 images fixes et 100 images animées et, en moyenne, ce fonds s'accroît de plus de 1 000 photographies par an.

Le fonds s'enrichit depuis peu d'un nouveau genre de documents, celui des documents de travail internes des architectes, graphistes, designers, etc., constituant ainsi l'historique de création des différents produits de communication de l'entreprise.

## **1.2 Le projet à réaliser**

### **1.2.1 L'enjeu**

D'après une étude d'Élyette Roux, les parfums et les cosmétiques représentent le marché le plus important des marques de luxe. Le rythme de lancement des nouveaux produits connaît une accélération exponentielle depuis les années 1980 : en effet, à cette époque, une grande marque lançait un nouveau parfum féminin tous les sept ans environ contre moins de trois ans aujourd'hui, soit en moyenne un nouveau parfum tous les ans avec l'alternance féminin-masculin. Élyette Roux explique ainsi que le temps de lancement d'un nouveau parfum est de deux ans en moyenne : six mois pour le concept, douze mois pour l'élaboration, la fabrication et les tests (jus et flacon) et six mois pour la communication. Les grandes marques sont donc en lancement perpétuel [46, Roux, p.27-28].

Le marché est dès lors saturé de nouveaux produits qui s'ajoutent aux classiques. En conséquence, attirer l'attention du public sur un nouveau produit ou sur des produits mythiques, face à la concurrence, est un enjeu majeur en raison des investissements engagés. « Cosmétique Magazine » relevait en mars 2007 que le groupe LVMH<sup>6</sup> réalisait à lui seul la moitié des investissements publicitaires du secteur avec 49,7 millions d'euros pour ses marques Dior, Kenzo et Guerlain [46, Roux, p.29].

Pour rentabiliser ses investissements, la seule clientèle traditionnelle du luxe ne suffit plus. Ainsi, depuis plusieurs années, le secteur du luxe s'est démocratisé sur le marché des cosmétiques auprès des classes moyennes. Pour se déterminer, cette nouvelle cible se fonde sur un raisonnement valeur-prix, et non plus seulement qualité-prix, que cette valeur soit symbolique, affective ou émotionnelle [46, Roux, p.25].

---

<sup>5</sup> Back et front office : le back office est l'interface de l'administrateur d'un site, d'une base de données, etc. tandis que le front office est l'interface de l'utilisateur.

<sup>6</sup> LVMH : initiales de Louis Vuitton Moët Hennessy, groupe français de l'industrie du luxe.

## 1.2.2 Une consommation ostentatoire

La fonction sociale de prestige de la consommation a été mise en évidence dès la fin du XIX<sup>e</sup> siècle par Thornstein Veblen sous la notion de « conspicuous waste », traduite par prodigalité ostentatoire, dépense ou consommation de prestige<sup>7</sup>. Ainsi, René-Maurice Deremaux note : « Au sein du luxe, les consommateurs ne regardent pas tant l'utilité d'un produit, que sa valeur ajoutée en terme d'image, la démarcation sociale qu'il permet, sa qualité ou encore sa rareté » [47, Dereumaux, p.72]. En conséquence, la clientèle achète davantage un symbole que l'efficacité d'un produit.

Or, si l'efficacité peut se démontrer par des arguments, des chiffres (donc des informations textuelles), le statut projeté par une marque - et surtout du luxe -, ne se prouve pas, il relève du ressenti conféré par l'imaginaire afférent. Par exemple, René-Maurice Deremaux constate que « les jeunes dirigeants se reconnaissent plutôt dans le sobre et l'élégant de Gucci, Ralph Lauren ou encore Armani alors que les "bourgeois bohêmes" se sentent plus proches des univers de Christian Lacroix ou de Kenzo. » [47, Dereumaux, p.74].

De ce fait, les entreprises du luxe communiquent sur leur image de marque, dans laquelle les consommateurs actuels et potentiels se reconnaissent ou souhaiteraient être reconnus. Cette image de marque est leur bien le plus précieux et leur principal outil de promotion. Dès lors, un produit de luxe est le résultat d'un objet concret et d'une image subjective. Cette dernière est suggérée et concrétisée grâce à des images publicitaires fondées sur l'histoire, la tradition, la beauté [47, Dereumaux, p.71-74].

Le luxe doit aussi démontrer sa modernité et son avant-gardisme, qui le distinguent des marques grand public. L'art est ainsi un vecteur de communication du luxe. La création visuelle publicitaire du secteur du luxe fait donc appel à des photographes et des réalisateurs reconnus pour leurs talents créatifs. Dès lors, l'achat d'un produit de luxe devient « le passeport pour un univers d'exception, un univers idéal. » [47, Dereumaux, p.77].

## 1.2.3 Le projet

Ainsi, la création visuelle est au cœur de l'activité de l'entreprise, le message étant porté par l'univers imaginaire de la marque et l'atmosphère suscitée par l'image. Dans ce secteur concurrentiel, l'entreprise souhaite pouvoir répondre à plusieurs questions :

- Comment l'imaginaire du luxe et de l'image de marque est-il traduit visuellement dans les campagnes publicitaires de la concurrence ?
- Quels univers créatifs les concurrents utilisent-ils pour faire passer leurs messages ?
- Quels mannequins, quels photographes sont choisis pour donner vie à ces univers ?
- Quels sont les thèmes à la mode et ceux qui vont le devenir ?
- Quelles sont les bonnes idées, les mauvaises ?
- Quelles idées de l'entreprise les concurrents ont-ils reprises ? etc.

Ces questions supposent la mise sous surveillance du secteur. Toutefois, contrairement à la nature habituelle de la veille, le service Achat d'art n'a pas besoin d'informations chiffrées ou textuelles mais d'ordre visuel. Aussi le projet doit-il proposer une solution de recherche et de veille des images produites par le secteur du luxe - le marché Parfums-Beauté en particulier -, pour y détecter les signaux faibles<sup>8</sup> visuels, afin d'aider à la prise de décision créative et donc stratégique.

---

<sup>7</sup> VEBLEN Thornstein. The Theory of the leisure class. 1899 [48, Baudrillard, p.24].

<sup>8</sup> Signaux faibles : éléments de l'environnement devant faire l'objet d'une veille (cf. p.22).

## 2 L'analyse de la situation

---

### 2.1 L'analyse de l'existant dans l'entreprise

#### 2.1.1 Questionnaire et entretiens

Afin d'élaborer une solution de veille dédiée au service Achat d'art, la première étape consiste à vérifier qu'un tel dispositif n'est pas déjà déployé dans l'entreprise. En effet, les services impactés par la création visuelle sont nombreux, l'entreprise grande, aussi est-il envisageable qu'une personne ait mis en place "sa" réponse à "son" besoin, sans la faire partager à l'ensemble de ses collègues.

Si l'on sait déjà que dans le service Recherche et Développement (R&D) une documentaliste occupe la fonction de veilleur, aucune autre veille n'a été identifiée. Le service Marketing<sup>9</sup>, avec lequel travaille étroitement le service Achat d'art lors de la phase de conception créative, semble particulièrement susceptible d'en avoir développé, l'analyse du marché et de la concurrence étant au fondement de ses activités.

J'ai ainsi mené différents entretiens avec plusieurs services dont les deux susmentionnés. La matrice du questionnaire est présentée en Annexe 1. Il est organisé en deux parties :

- d'une part, l'analyse du système informationnel existant : le rapport à l'information, le Système d'information personnel (SIP) et le Système d'information collectif (SIC) ;
- et d'autre part, les attentes, les besoins et les freins dans le cadre d'une solution de veille visuelle.

Les entretiens se sont déroulés en face-à-face, pendant 45 minutes en moyenne, sur le poste de travail de l'interviewé, afin d'observer ses méthodes de travail.

#### 2.1.2 Les entretiens avec les services Marketing et R&D

##### 2.1.2.1 Le Marketing

Cinq personnes du Marketing ont accepté de me répondre, dans des domaines très divers : le Marketing produit<sup>10</sup>, le Marketing digital (sur Internet), le Merchandising<sup>11</sup> et l'Architecture pour les points de vente.

Ces entretiens ont permis de mettre en évidence plusieurs éléments communs :

- toutes les personnes interrogées ont adhéré au projet d'une veille visuelle sur les créations et les tendances du secteur du luxe ;
- les pratiques en matière informationnelle sont proches, orientées web : la création d'alertes Google, l'organisation des sites préférés en marque-pages, très peu de fils RSS,

---

<sup>9</sup> Marketing : ensemble des techniques permettant de concevoir l'offre d'un bien ou d'un service en fonction de l'analyse des attentes des consommateurs, des capacités de l'entreprise et des contraintes environnementales (concurrentielles, légales, culturelles, etc.).

<sup>10</sup> Marketing produit : service chargé de définir, d'adapter, de différencier les biens et services qui composent l'offre de l'entreprise, en fonction de la stratégie arrêtée en amont.

<sup>11</sup> Merchandising : service dédié aux points de vente (assortiment et présentation des produits).

l'inscription à des newsletters<sup>12</sup>, le recours à Google pour rechercher des informations ou des images et la réception de messages internes par la messagerie ;

- l'archivage des messages, des images, des données se fait principalement dans la messagerie, sur un serveur, dans un dossier partagé avec des collaborateurs ou dans un dossier personnel ;
- une culture écran prononcée : les personnes utilisent très peu le papier, uniquement dans le cas de réunion pour diffuser une information (souvent sous forme de power point).

Bien que satisfaites de leur organisation, qui répond *a priori* à leurs besoins, les personnes interrogées (sauf une) reconnaissent que leur démarche est chronophage (alors que la veille n'est pas leur cœur de métier), « artisanale » (pour reprendre un mot souvent employé) et qu'elles doivent probablement laisser passer des informations qui pourraient les intéresser.

Une veille marketing a été identifiée, sous forme de newsletter mensuelle, réalisée par le Marketing produit. Elle consiste à informer des lancements de produits et de boutiques de la concurrence. Elle est réalisée avec l'information collectée par l'abonnement à des sites professionnels, la consultation régulière de blogs<sup>13</sup>, sites, la réception d'alertes Google, la lecture de magazines papier, etc. Elle est diffusée par la messagerie aux personnes qui la demandent (aujourd'hui près de 150 personnes) sous la forme d'un power point de 200 diapositives environ. Ceci induit que :

- chacun des 150 destinataires est susceptible de l'enregistrer dans un dossier ou sur un serveur, alourdissant ainsi le réseau ;
- la solution adoptée n'offre aucune possibilité de recherche dans un ou plusieurs numéros ;
- aucun archivage n'est prévu.

Toutefois, l'exhaustivité de la newsletter ainsi que la qualité des informations incitent les collaborateurs à s'y abonner.

#### **2.1.2.2 Le service Recherche et Développement**

J'ai rencontré la responsable de la veille du service de Recherche et Développement. Contrairement au Marketing qui s'intéresse au lancement des produits, la R&D se concentre sur la fabrication des produits, donc sur des informations scientifiques et techniques. Son expertise est particulièrement intéressante puisqu'elle utilise un logiciel de veille et publie un bulletin d'information mensuel diffusé par messagerie à 150 destinataires environ, sous format Word.

Bien que la collecte soit professionnalisée par le logiciel de veille, des problèmes identiques à celle de la veille marketing sont soulevés :

- le poids de la diffusion (puisque le bulletin comprend des images d'illustration) ;
- le problème de recherche dans un ou plusieurs bulletins ;
- l'absence d'archivage autre que sous la forme d'un PDF sur le poste de travail de la documentaliste.

Elle a testé pendant quelques mois une plateforme collaborative de veille mais le résultat n'a pas été à la hauteur des attentes (et du coût d'un tel outil) : en effet, elle a constaté

---

<sup>12</sup> Newsletter : document périodique envoyé par messagerie électronique à une liste de diffusion de personnes qui s'y sont inscrites.

<sup>13</sup> Blog : espace personnel d'expression, dans lequel chacun fait partager ses intérêts, ses avis, etc.

que, si dans la vie privée le partage d'informations est courant dans les réseaux sociaux, les pratiques collaboratives en milieu professionnel sont difficiles et plus longues à mettre en place (ce problème est développé p.67).

## **2.2 L'analyse des besoins du demandeur**

La mise en place d'une veille sur les créations visuelles est une demande du service Achat d'art, auquel je suis affectée. En conséquence, si les précédents entretiens m'ont permis de me faire une idée des pratiques informationnelles de certains services, l'entretien avec l'assistante Achat d'art, qui pratiquera cette veille, est davantage orienté sur les problèmes qu'elle rencontre, ses attentes, ses besoins et les résultats qu'elle attend d'une telle veille.

### **2.2.1 Trois besoins identifiés**

L'assistante Achat d'art travaille exclusivement sur les images. Ses tâches demandent beaucoup de réactivité et sont difficilement planifiables à longue échéance. C'est le cas par exemple lorsqu'une demande de recherche de visuels lui parvient : les délais de réponse accordés sont brefs alors que la recherche est longue. Ses fonctions la mettent en relation avec différents métiers. Dans le cadre de ma mission, nos interlocuteurs communs sont :

- le service Marketing produit avec lequel elle travaille pour la réalisation du brief créatif<sup>14</sup> ;
- la responsable Achat d'art qui corrige, redéfinit les orientations créatives, valide (avec la Direction Artistique) le brief créatif ;
- la chef de projet en ingénierie documentaire, responsable du cycle de vie des images.

Pour collecter et présenter les images en vue de la réalisation du brief créatif ou pour répondre à une demande ponctuelle sur un visuel, l'assistante Achat d'art recourt à différentes sources et outils, principalement de nature dématérialisée :

- Internet : elle y effectue ses recherches ponctuelles et continues d'images par l'abonnement à des newsletters, des alertes Google, la consultation de sites et blogs qu'elle a identifiés comme des références, l'utilisation de moteurs de recherche ;
- intranet :
  - elle utilise le DAM pour trouver des visuels des précédentes campagnes de l'entreprise pour s'inspirer ou faire évoluer les concepts artistiques ;
  - si besoin, elle accède également à l'intranet du service Marketing pour trouver des informations plus précises sur les briefs qui lui sont transmis.

Les images qu'elle collecte lors de ses campagnes de recherche sont classées dans un dossier personnel, organisé en sous-dossiers. Cette organisation est peu pratique car elle ne peut pas lancer des requêtes de recherche or son fonds est composé de près de 6 000 images, ce qui rend une consultation image par image impossible.

De ce fait, trois besoins ont été identifiés et ont donné naissance au projet de veille :

- répondre de façon rapide et pertinente à une demande de visuels sur la concurrence, sur un mannequin, un photographe ;
- trouver une information depuis un visuel : identifier tel mannequin, tel photographe de telle campagne ;

---

<sup>14</sup> Brief créatif : document qui présente les stratégies créatives retenues.

- capter les nouveautés : en matière de créations, de tendances, mais aussi de mannequins, de photographes, etc.

Une solution de recherche et de veille pourrait impacter son poste de la manière suivante :

- se recentrer sur son cœur de métier ;
- accélérer le processus de décision stratégique ;
- prendre un avantage sur la concurrence en repérant de futurs artistes, égéries, concepts qui assureraient une position avant-gardiste ;
- contribuer à la réactivité de l'entreprise sur la concurrence.

### **2.2.2 De fortes contraintes détectées**

La première, qui revient dans tous les entretiens, est l'impérieuse confidentialité : nul ne doit savoir, en externe comme en interne (sauf les personnes autorisées), les sujets mis sous surveillance - les thèmes, produits mais aussi les photographes, mannequins, concurrents, etc. -, afin que les pistes créatives demeurent secrètes. La solution devra garantir la confidentialité des recherches et des acquisitions et donc s'adapter à la sécurité du système existant. Ce qui induit la mise à l'écart de solutions type Software as a Service (SaaS)<sup>15</sup> ; l'utilisation de Google est même à débattre puisque le moteur enregistre toutes les requêtes, à moins de trouver des outils permettant d'effacer les traces (cf. p.60).

Une autre contrainte est le fait que l'assistante Achat d'art n'est pas documentaliste. Certes, elle a suivi une formation pour se familiariser aux bases de la recherche, comme les opérateurs booléens. Toutefois, comme me l'a précisé la responsable de la veille R&D, les logiciels professionnels ne sont pas intuitifs pour des non spécialistes et plus leurs fonctionnalités sont importantes (en quantité ou en qualité) plus leur prise en main et leur utilisation sont complexes.

Or, la contrainte principale de ce projet est la nature même de l'objet de la recherche et de la veille : l'image. Lorsque nous recherchons une information, lorsque nous voulons créer des alertes, nous avons l'habitude d'utiliser des mots-clefs. Dans ce projet, l'utilisation de mots-clefs semble bien insuffisante pour traduire et faire émerger d'Internet des univers imaginaires, des ambiances, des émotions, des sentiments, etc.

L'enjeu est donc de trouver une ou des solutions permettant d'automatiser la collecte et de faciliter la recherche d'images connotant des émotions très subjectives donc difficilement interprétables en mots. Dès lors, dans quelle mesure est-il possible de se fonder, non pas sur des critères textuels pour mettre en place cette surveillance, mais sur le contenu de l'image lui-même ?

Pour y répondre je vais étudier les spécificités d'une veille menée dans l'environnement demandé, Internet.

---

<sup>15</sup> SaaS : abonnement à un logiciel en ligne grâce à un programme client-serveur (en opposition à l'achat d'une licence).

## 3 La veille sur Internet

---

### 3.1 Définitions et circonscription de la veille idoine

#### 3.1.1 La veille : définitions et démarches

Plusieurs mémoires relatifs à une mission de veille ont déjà présenté de manière détaillée la veille<sup>16</sup>. Aussi, je ne m'attacherai qu'aux éléments pertinents dans le cadre de mon projet.

##### 3.1.1.1 La définition et les recommandations du rapport Martre

Le rapport Martre [16, Martre et al.] est certes ancien (1994) mais il est cité par toutes les sources lues ayant trait à la veille et à l'intelligence économique. Ce rapport définit les caractéristiques de la veille, fait un constat sur son utilisation et une série de recommandations. Je reprends ici ses définitions, les préconisations étant évoquées plus loin (cf. p. 68).

La veille à mettre en place dans le cadre de ma mission doit délivrer une « information utile » c'est-à-dire « Celle dont les différents niveaux de l'entreprise ont besoin pour élaborer et mettre en œuvre [...] la stratégie et les tactiques nécessaires à l'atteinte des objectifs définis par l'entreprise dans le but d'acquérir et/ou de préserver, voire d'améliorer sa situation dans son environnement concurrentiel. » [16, Martre et al., p.102]. Dans mon projet, cette information utile sera à interpréter depuis les visuels.

De plus, « Ces actions, au sein de l'entreprise, s'ordonnent en un cycle ininterrompu [...] » [16, Martre et al., p.11] : l'information utile collectée va certes satisfaire le besoin du demandeur mais aussi susciter chez lui de nouveaux besoins, relançant un nouveau cycle de recherche et de collecte. Cette information est recueillie par un « acquéreur » c'est-à-dire la personne qui recueille l'information utile [16, Martre et al., p.102].

Le rapport parle davantage d'intelligence économique que de veille. Cette distinction est couramment faite dans les sources traitant de ce sujet. La veille est un outil de l'intelligence économique : elle délivre l'information utile qui, analysée, traitée et diffusée aux bons acteurs, permettra d'anticiper l'avenir et d'agir - l'action relevant de l'intelligence économique [16, Martre et al., p.11 et 19]. Or, l'objectif assigné à la veille visuelle est d'être un outil de décision stratégique, s'inscrivant ainsi dans les définitions du rapport Martre.

Ainsi, la veille a deux approches : d'une part, elle recherche les informations à caractère anticipatif, que l'on appelle des « signaux faibles » ; d'autre part, elle acquiert des informations contribuant à une meilleure connaissance de l'environnement à partir de données peu accessibles [9, Balmisse et al., p.24].

##### 3.1.1.2 La norme AFNOR

La veille est également définie par la norme de l'Association française de normalisation (AFNOR)<sup>17</sup> : « [...] politique d'intelligence économique, laquelle englobe la mise en place d'une fonction d'observation et de surveillance en vue de détecter, analyser et suivre tous les signaux susceptibles de conforter, d'infléchir ou de remettre en cause sa stratégie ou

---

<sup>16</sup> [5, Miaux], [8, Toupin Lemaire], [11, Ha], [13, Perbost] et [14, Ajalbert].

<sup>17</sup> Norme XP X 50-053 « Prestations de veille et prestations de mise en place d'un système de veille », 1998. [15, AFNOR].

les décisions prises, sans omettre la protection du capital connaissance (protection du patrimoine). » [15, AFNOR, p.4].

Si la veille est définie comme une « Activité continue et en grande partie itérative visant à une surveillance active de l'environnement technologique, commercial, etc., pour en anticiper les évolutions », la demande qui a été formulée relève de la mise en place d'un système de veille, c'est-à-dire d'un « Ensemble structuré réunissant les compétences répondant à des besoins de veille. » [15, AFNOR, p.6]. La prestation de veille peut être réalisée en interne ou bien être externalisée auprès d'un prestataire.

### **3.1.1.3 La démarche d'une activité de veille**

Afin de mettre en place un système de veille, plusieurs étapes sont à respecter. Celles décrites ci-dessous sont reprises de la norme AFNOR [15, AFNOR, p.4, 8, 9] :

1. la définition ou le recueil des axes de surveillance et des finalités : décrire les sujets d'intérêt et identifier les finalités ;
2. la détermination des types d'informations utiles : dans le cadre du projet, il s'agit d'images ;
3. l'identification et la sélection des sources d'information, appelées couramment le *sourcing*<sup>18</sup> ;
4. la collecte et la sélection des informations selon les axes de surveillance et les finalités assignés à l'étape 1 ;
5. le traitement et l'analyse des données collectées ;
6. la synthèse et la mise en perspective : dégager le sens, les aspects stratégiques de l'information collectée pour aider à la décision ;
7. la communication des résultats ;
8. la validation et le réajustement : la veille est un processus itératif, elle évolue, se réoriente sans cesse vers de nouveaux axes de surveillance.

## **3.1.2 La circonscription de la veille**

### **3.1.2.1 Des ajustements au cadre du projet**

Par rapport à cette démarche, j'ajouterais une autre étape essentielle, celle de la capitalisation des données acquises : en effet, les images collectées doivent être archivées pour alimenter un fonds dans lequel l'assistante Achat d'art fera des recherches rapides pour répondre à une demande ponctuelle ou pour mettre au point un brief créatif.

De plus, le modèle conçu par François Jakobiak<sup>19</sup>, met en évidence trois catégories d'acteurs, généralement reprises par la littérature : les observateurs, les experts et les décideurs. Or, dans le projet, l'experte, c'est-à-dire l'assistante Achat d'art, est aussi l'observatrice, celle qui devra acquérir l'information utile.

---

<sup>18</sup> Sourcing : (cf. p.53).

<sup>19</sup> JAKOBIAK François. Pratique de la veille technologique. In CASTANET Caroline. Mise en place d'une veille mutualisée par le biais d'une plateforme de veille. 2010. 1 volume, 132 p. Mémoire, Titre professionnel « Chef de projet en ingénierie documentaire » niveau I, INTD, 2010. p.23.



### **3.1.2.2 Une veille créative**

La littérature dédiée à la veille distingue plusieurs types de veille. Dans le cadre du stage, il s'agit d'une veille créative. Stéphane Gorla en propose différentes approches, parmi lesquelles j'en retiendrai deux, adaptées au contexte du projet [3, Gorla, p.60] :

- « Elle peut simplement correspondre à une veille de l'univers créatif » ;
- « Elle peut être comprise comme une veille qui est destinée aux personnels créatifs ».

Le système de veille à mettre en place devra en effet repérer les signaux faibles de la création visuelle - relatifs aux nouveaux codes graphiques, aux atmosphères, aux artistes et mannequins -, qu'ils soient issus de la concurrence ou des cabinets de tendance. Outre la dimension anticipative, il devra également identifier les items déjà utilisés avec succès ou non par la concurrence et, plus largement, par le secteur du luxe, soit pour s'en inspirer soit pour s'en démarquer.

### **3.1.2.3 Une veille web**

Les images publicitaires dans le secteur du luxe sont diffusées par l'affichage dans la rue, sur les présentoirs (publicité sur le lieu de vente) et surtout dans la presse. Plusieurs services créatifs de l'entreprise sont abonnés à des magazines pour étudier les réalisations concurrentes.

Le système de veille à élaborer est circonscrit à Internet. En effet, celui-ci diffuse la plupart des images publicitaires magazines, avec néanmoins des adaptations. Le choix de se concentrer sur Internet tient au fait qu'il est « à la fois une source d'informations [...] et un accès à un ensemble de technologies qui [...] facilitent la collecte d'informations. » [4, Khénissi et al., p.138].

De plus, le Web invisible, c'est-à-dire toutes les pages web non indexées par un moteur de recherche, regorge d'images, notamment celles des blogs, qui confèrent un véritable avantage concurrentiel puisque ces images ne sont pas mises à la portée de tous.

Enfin, la veille portera uniquement sur les images fixes en ligne. En effet, la complexité du sujet m'a obligée à limiter la nature des images. Toutefois, en fonction de l'efficacité du système mis en place, il sera envisageable, dans un second temps, d'élargir la veille aux images animées.

Pour résumer, le projet est une veille créative des images fixes diffusées sur Internet dans le secteur du luxe, en Parfums-Beauté. Dès lors, après avoir défini les contours de la veille, je vais m'intéresser aux caractéristiques de l'environnement de recherche et de collecte.

## **3.2 Le Web 2.0 et 3.0**

### **3.2.1 Les définitions d'un environnement en mutation**

#### **3.2.1.1 Le Web 2.0 de Tim O'Reilly**

Le « Web 2.0 » est une expression qui marque le deuxième âge du Web, selon les nouvelles fonctionnalités qu'il offre aux usagers. Elle a été inventée en 2003 et consolidée par Tim O'Reilly<sup>20</sup> dans un article paru en 2005 où il définit les sept piliers de ce nouveau Web

---

<sup>20</sup> Fondateur de la société O'Reilly Media, supporteur du logiciel libre et du mouvement open source (expression qui s'applique aux logiciels dont la licence respecte les critères de l'Open Source Initiative - OSI -, soit la possibilité de libre redistribution, d'accès au code source et de travaux dérivés).

[30, O'Reilly]. Véronique Mesguich les résume en ces mots : « Simplicité d'accès, partage, mutualisation, individualisation, réutilisabilité des contenus » [12, Mesguich, p.58].

L'évolution technologique et ergonomique des applications a permis aux internautes de changer de statut et de prendre leur autonomie. Ainsi, les utilisateurs ne sont plus seulement des consommateurs de données que des professionnels diffusent, ils sont désormais producteurs potentiels d'informations [2, Foenix-Riou, p.6] : ils créent leur propre support (par exemple un blog), partagent leurs données sur des plateformes communes (images ou vidéos), mettent en place des systèmes automatiques de collecte, de veille, etc.

### 3.2.1.2 Le Web social

Ainsi, la caractéristique principale du Web 2.0, est le développement de réseaux, de communautés de partage, dans lesquels l'internaute se reconnaît et à l'intérieur desquels il développe des relations avec d'autres internautes soit qu'il connaît réellement soit avec lesquels il partage des intérêts, des avis communs. On parle alors de « Web social ».

Pour quasiment chaque type de contenu, des plateformes ou des réseaux dédiés se sont créés ; en matière d'images fixes :

- Flickr<sup>21</sup> : site de partage gratuit de photos et de vidéos ;
- Pinterest<sup>22</sup> : réseau social qui permet de punaiser (*to pin*) sur un tableau en liège virtuel des images ou des vidéos (plus de 10 millions d'utilisateurs)<sup>23</sup> ;
- Tumblr<sup>24</sup> : service entre le blog et l'agrégateur de contenu (réception des flux dans son tableau de bord personnel) ; ses utilisateurs privilégient la diffusion d'images de qualité (artistiques, créatives), exprimant un style de vie ;
- Instagram<sup>25</sup> : service de partage de photos qui met l'accent sur l'aspect esthétique et qui permet le partage social et les commentaires sur Twitter et Facebook (qui l'a racheté)<sup>26</sup>.

Ce Web social se manifeste également sous la forme d'indexation collaborative. Les internautes associent des mots-clefs choisis librement aux images afin de caractériser leur contenu : ces mots-clefs sont appelés tags et l'action tagging. Des communautés d'échanges thématiques se constituent à partir de ces tags, qui servent aussi de points d'entrée ou de recherche, y compris pour une image [36, Ménard, p.6].

De ce Web 2.0, Véronique Mesguich en tire les conclusions professionnelles suivantes : « Blogs, [...] réseaux sociaux constituent ainsi autant de nouvelles cordes à l'arc des professionnels de la veille, permettant de traquer le plus en amont possible les tendances ou informations émergentes. » [12, Mesguich, p.58].

---

<sup>21</sup> Flickr : <<http://www.flickr.com/>>

<sup>22</sup> Pinterest : <<http://pinterest.com/>>

<sup>23</sup> Kevin. Pinterest, le réseau social qui monte, qui monte... In Presse-Citron [blog en ligne]. 09.02.2012 [consulté le 11.07.2012] <<http://www.presse-citron.net/pinterest-le-reseau-social-qui-monte-qui-monte>>

<sup>24</sup> Tumblr : <<https://www.tumblr.com/>> : Tumblr revendique 20 milliards de publications et 50 millions de blogs d'après BLADIER Cyril. Les clés du Picture Marketing : analyse d'une nouvelle tendance du Web. In Presse-Citron [blog en ligne]. 25.04.2012 [consulté le 11.07.2012]. <<http://www.presse-citron.net/les-cles-du-picture-marketing-analyse-dune-nouvelle-tendance-du-web>>

<sup>25</sup> Instagram : <<http://instagram.com/>>

<sup>26</sup> BOULESTAN Rénaud. Instagrille 2.0 : nouvelle version du client Instagram pour PC. In Presse-Citron [blog en ligne]. 10.07.2012 [consulté le 11.07.2012]. <<http://www.presse-citron.net/instagrille-2-0-nouvelle-version-du-client-instagram-pour-pc?>>

### 3.2.1.3 Le Web de données

#### *Définition*

Le Web a abordé un nouvel âge : le Web 3.0 que Tim O'Reilly appelle « Web sémantique » ou « Web de données ». Cette expression fait référence à la vision de la W3C des données liées (ou linked data)<sup>27</sup>.

Le Web est constitué de multiples dépôts de données. Pour que ces données interagissent et permettent d'accéder à des informations de valeur ajoutée, elles doivent être disponibles et mises en relations (par des inférences) grâce à un format standard, accessible et gérable par des outils de Web sémantique (ce standard est défini par la W3C). Les vocabulaires définissent des concepts (et réduisent les ambiguïtés), classent des termes, caractérisent des relations (on parle d'ontologies).

L'objectif de ce Web de données n'est donc pas l'accumulation de données mais la capacité d'y extraire du sens depuis cette masse devenue ingérable à échelle d'homme [26, Mangematin]. Ainsi, Véronique Mesguich constate que ce Web de données répond aux besoins des professionnels de la veille, avec :

- « d'une part, un accès formalisé à la connaissance et aux données, via des ontologies spécialisées, dans un métalangage permettant de s'affranchir de la barrière des langues et des ambiguïtés linguistiques ;
- d'autre part, des possibilités d'analyse de l'information accrues par des traitements linguistiques sophistiqués. » [12, Mesguich, p.59].

#### *Google et « The Knowledge Graph »*

La recherche sémantique devient ainsi un enjeu, que Google ne peut laisser passer. Ainsi, le célèbre moteur de recherche a lancé en mai dernier le « Knowledge Graph » (pour l'instant disponible uniquement pour la version anglo-saxonne) afin d'effectuer des recherches complémentaires et d'obtenir de nouveaux renseignements sur les entités nommées (noms propres) à partir de « l'intelligence collective du Web ». Google promet trois avancées :

1. « Trouver la chose correcte » : désambiguïsation de la requête par la compréhension des mots utilisés dans leur contexte de recherche ;
2. « Avoir le meilleur résumé » : affichage d'un résumé et d'informations connexes sous la forme d'un encadré ;
3. « Aller plus profond et plus loin » : proposition de liens sur des sujets proches<sup>28</sup>.

La question est de savoir si, dans le cadre d'une recherche d'images, ce Web sémantique - et bientôt la recherche par le « Knowledge Graph » sur des noms de marque,

---

<sup>27</sup> W3C : World Wide Web Consortium : communauté internationale développant les standards (ensemble de technologies et de protocoles) ouverts pour assurer la croissance à long terme du Web. [25, W3C.org].

<sup>28</sup> ANDRIEU Olivier. Knowledge Graph : Google officialise son moteur sémantique. In Abondance Actualité [liste de diffusion]. 21.05.2012 [consulté le 16.08.2012]. <<http://www.abondance.com/actualites/20120521-11478-knowledge-graph-google-officialise-son-moteur-semantique.html>>

SINGHAL Amit. Introducing the Knowledge Graph : things, not strings. In Google™ Official Blog [blog en ligne]. 16.05.2012 [consulté le 17.08.2012]. <<http://googleblog.blogspot.fr/2012/05/introducing-knowledge-graph-things-not.html>>

de photographes, de mannequins, etc. plutôt que sur le contenu de l'image -, sera pertinent. J'aborderai cette question dans la deuxième partie du mémoire (cf. p.32 et suivantes).

#### 3.2.1.4 Le Web<sup>2</sup>

La dernière étape du Web est appelée le « Web Squared » (Web<sup>2</sup>), expression lancée par Tim O'Reilly lors du cinquième sommet Web 2.0 en octobre 2009. Ce concept, symbolisé par Twitter<sup>29</sup>, est apparu en 2009 et se caractérise selon Béatrice Foenix-Riou [2, Foenix-Riou, p.240-241] par :

- le Web temps réel : les informations sont publiées sur le Web de façon instantanée et publique c'est-à-dire qu'elles sont à la fois envoyées à un groupe de destinataires et analysées par des logiciels de traitement de l'information, en temps réel ;
- l'« Information Shadow » : chaque information est rattachée à d'autres données cachées que des logiciels vont rendre visibles et utilisables, comme les applications de réalité augmentée<sup>30</sup> ;
- les métadonnées implicites : l'exemple le plus connu est celui des appareils photos numériques qui indiquent la date de prise de vue, le modèle de l'appareil, etc. Dans Flickr, par exemple, ces données sont intégrées dans une base et sont réutilisées par le service. Aujourd'hui, des systèmes auto-apprenants génèrent automatiquement des métadonnées en recoupant des informations, comme dans les services de reconnaissance faciale (cf. p. 41 et suivantes) ;
- les bases de données interconnectées : l'interconnexion des bases de données permet d'accéder à tous les domaines relatifs à une donnée (principes du Web sémantique).

Les notions d'instantanéité et de métadonnées implicites sont au cœur du travail de veilleur et de documentaliste. Comme pour le Web sémantique, le problème de mon projet est d'arriver à concilier ces nouvelles fonctionnalités avec la recherche d'images.

### 3.2.2 Les risques

#### 3.2.2.1 L'infobésité : mythe ou réalité ?

Les implications du Web social peuvent être également négatives. Comme le constate Julie Mangematin, « En deux ans, l'Humanité a généré davantage de données que depuis le début de son Histoire. » [26, Mangematin]. Ce phénomène est appelé « infobésité » et est devenu un sujet qui serait caractéristique de notre époque web.

Pourtant, comme le souligne Aref Jdey, cette surabondance de données a toujours existé. Ainsi, « Leibniz qualifiait la croissance massive des ouvrages disponibles à son époque par "*horrible*" [...]. En 1550, Anton Francesco Doni, l'homme de lettres italien, affirmait ouvertement qu'il y avait, à son époque, "tellement de livres qu'il n'était même pas possible d'avoir le temps pour lire les titres." »<sup>31</sup>. Robert Burton, dans son ouvrage *The Anatomy of Melancholy*, publié en 1621, « détaille dans tous les sens les symptômes et les difficultés

---

<sup>29</sup> Twitter : réseau social créé en 2006 qui permet d'envoyer instantanément et gratuitement de brefs messages (limités à 140 caractères) par messagerie instantanée ou SMS.

<sup>30</sup> Réalité augmentée : elle ajoute en temps réel des éléments virtuels (2D ou 3D) au monde environnant (cf. p.36, exemple d'IBM).

<sup>31</sup> JDEY Aref. L'infobésité, ce non-problème qui accompagne l'Humanité depuis des siècles. In Demain la veille [blog en ligne]. 01.08.2012 [consulté le 15.08.2012].

<<http://www.demainlaveille.fr/2012/08/01/linfobesite-ce-non-probleme-qui-accompagne-lhumanite-depuis-des-siecles/>>

de ce malaise général, de sur-information. »<sup>32</sup>. Ainsi, pour Aref Jdey, cette surabondance d'informations a toujours existé, à toutes les époques.

Dans le cadre du Web, l'article d'Alex Hudson comptabilise ce qui se passe en soixante secondes : par exemple, on dénombre 168 millions de courriels envoyés, 694 445 recherches sur Google, 20 000 nouvelles publications sur Tumblr, 6 600 nouvelles photos sur Flickr, etc.<sup>33</sup>.

Dans le cadre d'une veille web, le nombre de données collectées est ainsi considérable, sans avoir de garantie quant à la qualité du contenu.

### **3.2.2.2 La qualité de l'information**

En effet, avant le Web, la recherche d'informations fournissait un nombre moins important de données qui étaient généralement validées, structurées, indexées car issues de bases de données en ligne. Elles étaient donc fiables [12, Mesguich, p.59]. Aujourd'hui, on privilégie la fraîcheur de l'information délivrée, qui est bien souvent reprise sans vérification, y compris par les journalistes. Or, chaque internaute est potentiellement producteur de données. Une enquête sur l'émetteur est ainsi conseillée afin de qualifier la source avant de réutiliser l'information.

De même, si l'indexation collaborative offre des opportunités nouvelles de recherche, les tags utilisés sont choisis par les internautes et dépendent donc des connaissances, des capacités et de la subjectivité de chaque indexeur plus ou moins aguerri à l'exercice.

Ainsi, le projet consiste en une veille créative du secteur du luxe mise en place sur le Web, devenu social, devenant sémantique, avec de nouvelles fonctionnalités qui augmente ses capacités, comme le temps réel ou l'indexation implicite.

Ce cadre posé, je vais étudier la question de l'image, plusieurs fois évoquée, et surtout les techniques permettant de rechercher et de retrouver une image par son contenu.

---

<sup>32</sup> JDEY Aref. Infobésité : arrêtez avec ce non-sens. In Demain la veille [blog en ligne]. 29.08.2010 [consulté le 15.08.2012]. <<http://www.demainlaveille.fr/2010/08/29/infobesite-arretez-avec-ce-non-sens/>>

<sup>33</sup> HUDSON Alex. The age of information overload. In BBC News [site en ligne]. 14.08.2012 [consulté le 25.08.2012]. <[http://news.bbc.co.uk/2/hi/programmes/click\\_online/9742180.stm](http://news.bbc.co.uk/2/hi/programmes/click_online/9742180.stm)>

# **Deuxième partie**

## **Veille et recherche par l'image**

## 4 Les particularités de l'image-document

---

### 4.1 L'analyse de l'image

#### 4.1.1 La fiche technique de l'image numérique

L'image photographique consiste à « capter, [...] enregistrer et [...] révéler la lumière sur différents supports natifs via un procédé chimique (pour la photo argentique) ou électronique (pour la photo numérique). » [19, Bayle, p.14]. Avec le numérique, l'image fixe est transformée en fichiers informatiques.

La photo numérique utilise des capteurs CCD (*charged coupled services*) qui, lorsque la lumière les atteint, transmettent les informations saisies sous forme numérique. Chaque capteur comprend trois capteurs qui mesurent la luminosité des lumières rouge (R), vert (V) et bleu (B) [44, GFII, p.10-11 et 24].

Une image est constituée de points, appelés pixels, qui sont les plus petits éléments constitutifs d'une image numérique (ses atomes), exprimés en bit. La définition d'une image renvoie au nombre de pixels ; la résolution est le nombre de pixels par unité de surface, exprimé en points par pouce.

Les données internes, dites techniques, sont générées automatiquement selon le format Exif<sup>34</sup> ; elles comprennent une trentaine de rubriques comme le modèle de l'appareil, la date et l'heure de la prise de vue, les réglages de l'appareil, etc. Les données internes textuelles, encapsulées dans le fichier numérique, décrivent le contenu. Elles permettent de saisir manuellement, selon le modèle IPTC-NAA IIM<sup>35</sup>, des données de type administratif, juridique et sémantique.

En 2001, le standard XMP<sup>36</sup> est développé. Il permet de traiter les accents et le multilinguisme pour les légendes. Il se fonde sur le Dublin Core<sup>37</sup>, organisé en quinze éléments qui renseignent le contenu, la propriété intellectuelle et la version.

#### 4.1.2 Les concepts de connotation et de dénotation

##### 4.1.2.1 Le contenu de l'image

Une fois définie la partie technique de l'image photographique, il faut s'intéresser à son contenu, au sens qu'elle porte. Selon Roland Barthes dans « La Chambre claire », c'est par le *studium* (ou champ d'intérêt culturel), que le *Spectator* (celui qui regarde l'image) s'intéresse à une photographie, c'est-à-dire aux intentions de l'*Operator* (le photographe), dans lesquels il y reconnaît ses propres goûts. Ce *studium* est dérangé par le *punctum* qui touche, point le *Spectator*. Or, pour Roland Barthes, « Ce que je peux nommer ne peut réellement me poindre. L'impuissance à nommer est un bon symptôme de trouble. » [22, Barthes, p.84]. Il soulève ainsi un problème qui est la capacité à nommer ce qui est susceptible de troubler, de retenir l'attention, de marquer.

---

<sup>34</sup> Exchangeable Image File.

<sup>35</sup> Information Interchange Model (IMM), créé en 1991 par l'international Press & Telecommunications Council (IPTC) avec la Newspaper Association of America (NAA).

<sup>36</sup> Extensible Metadata Platform.

<sup>37</sup> Norme de description pour l'Internet.

En outre, Roland Barthes définit la photographie comme étant « une émanation du référent » [22, Barthes, p.126]. Dans « Le message photographique », il écrit : « [...] c'est précisément cette perfection analogique qui [...] définit la photographie » [24, Barthes, p.128]. Dès lors, il en extrait deux messages : « un message dénoté, qui est l'*analogon* lui-même, et un message connoté, qui est la façon dont la société donne à lire, ce qu'elle en pense. » [24, Barthes, p.129]. En conséquence, la lecture d'une photo dépend du savoir de son lecteur, de sa capacité à y identifier les signes, les codes. Ainsi, l'experte assistante Achat d'art n'interprète pas une photographie comme un consommateur cherchant à se reconnaître dans le visuel ou comme la clientèle du luxe s'identifiant aux codes véhiculés.

Dans la publicité, afin que les messages atteignent leur(s) cible(s), ces signifiés sont transmis le plus clairement possible. Roland Barthes distingue trois types de messages :

- un message linguistique : le code de la langue ;
- un message iconique non-codé : l'image littérale dénotée ;
- un message iconique codé : l'image littérale connotée [23, Barthes, p.42-43].

Afin d'éviter tout égarement en raison de la polysémie de l'image (notamment due à la subjectivité de chacun), le message linguistique vient délimiter les interprétations possibles des signes (le message iconique codé).

#### **4.1.2.2 La relation de l'image au texte**

Toutefois, je l'ai expliqué précédemment (cf. p.17), le message linguistique est restreint dans la photographie publicitaire du secteur du luxe : elle se contente souvent du nom de la marque, du produit, parfois d'un bref slogan. Le message dénoté montre le produit ; mais comme le dit Coline Haupais, « la photographie publicitaire de luxe revêt un aspect artistique très fort privilégiant l'esthétique et la forme au contenu et au réalisme ». Ainsi, les couleurs, les jeux de lumières, de formes, les choix et poses des mannequins (s'il y en a) participent à la création d'un univers totalement connoté, dont l'interprétation est donc d'autant plus large que le texte sensé la canaliser est réduit voire absent. En effet, la photographie créative « travaille ses effets visuels et porte plus sur l'affect que produira la photo » [17, Haupais, p.60] (cf. des exemples en Annexe 2).

En conséquence, si la veille fondée sur le message dénoté semble réalisable (exemple : parfum féminin, rouge à lèvres, etc.), celle prospectant les tendances, les univers créatifs paraît beaucoup plus aléatoire : d'une part, le choix des mots-clefs des axes de surveillance et de recherche dépend de la subjectivité de l'acquéreur et d'autre part, les images acquises relèvent de la subjectivité de l'indexation (ou du tagging) de celui qui les met en ligne, en fonction du message qu'il veut véhiculer.

## **4.2 Nommer pour rechercher et retrouver**

### **4.2.1 La problématique de l'indexation**

#### **4.2.1.1 L'indexation automatique**

Frédéric Bayle explique que l'objectif de l'indexation de l'image est de : « [...] pallier, par la description et la caractérisation, et dans le cadre d'une gestion rationnelle des documents, la polysémie inhérente à l'image et ainsi d'essayer de circonscrire son contenu en contrôlant les éléments formels et sémantiques. Cette indexation a donc pour finalité de représenter rationnellement l'image afin de pouvoir la rechercher et la retrouver dans un fonds iconographique. » [19, Bayle, p.24-25].



Dans mon projet, le « fonds iconographique » est très vaste puisqu'il s'agit d'Internet : banques d'images en ligne, sites des marques concurrentes, des magazines en ligne, sites de partage de photographies, blogs personnels, etc. En conséquence, l'indexation permettant de rechercher une image à partir de concepts clefs est très aléatoire et amatrice.

Néanmoins, dès le niveau de la collecte d'images, il est possible de bénéficier des informations embarquées grâce à l'indexation automatique. En effet : « La description automatique du contenu visuel consiste à extraire automatiquement d'une image des descripteurs numériques significatifs, compacts et structurés qui seront utilisés et comparés au moment de la recherche de similarité visuelle. » [45, p.37].

L'automatisation a pour but l'exploitation de larges collections d'images, qui, à échelle humaine, ne serait pas réalisable. Dès lors, l'exploitation du fonds tentaculaire d'Internet pourrait trouver une solution en recourant à l'indexation déjà existante des images. Cependant, dans quelle mesure « des descripteurs numériques significatifs » sont-ils pertinents dans une recherche visuelle fondée sur des émotions, des impressions ?

#### **4.2.1.2 L'indexation manuelle**

En effet, si ces techniques fonctionnent sur des images dites simples, l'enjeu sur des images plus complexes est de parvenir à mêler les caractéristiques du contenu des images aux mots-clefs et aux légendes (quand il y en a) [44, GFII, p.55-56]. Les systèmes d'extraction des caractéristiques du contenu se concentrent principalement sur les couleurs et les textures. Pour aller plus profondément, il faut passer par une phase d'apprentissage, encore limitée aujourd'hui. De même, après extraction, les métadonnées de transcription requièrent une étape d'apprentissage du système [45, p.31 et 59-60].

Toutefois, même si les perspectives offertes par l'automatisation sont prometteuses, notamment quant au volume de données traitées, elles demeurent pour le moment encore imprécises. Ainsi, l'indexation automatique doit être vérifiée voire complétée avant d'intégrer une nouvelle image au fonds : des critères, tels que les couleurs ou les textures, sont restreints au regard des connotations exprimées dans l'imaginaire du luxe.

Enfin, l'objectif est aussi de constituer, grâce à la veille, une banque d'images, soit le DAM actuellement utilisé pour les visuels de l'entreprise, soit une nouvelle base spécialement dédiée aux opérations de veille. L'assistante Achat d'art, ses responsables, peut-être ensuite les créatifs, pourront y rechercher l'historique créatif d'un objet, d'une marque ou une idée par exemple. L'indexation après collecte à usage interne pourrait se fonder sur une liste d'autorité, un thésaurus métier (en fonction des outils disponibles) afin de dépasser ou de réduire les risques de subjectivité qui conduiraient à l'occultation de certaines images.

### **4.2.2 Les mots sont-ils inévitables ?**

#### **4.2.2.1 Le choix des mots**

L'indexation automatique peut certes choisir à notre place les mots adéquats, même si cette opération n'est pas encore performante. Cette difficulté à qualifier l'image par des mots-clefs se manifeste particulièrement lorsque l'on recherche une image. Aujourd'hui, cette recherche s'effectue majoritairement de manière textuelle par des moteurs, soit sur Internet, soit dans des banques d'images.

Ainsi, il relève de la personne qui recherche, collecte, archive de « comprendre comment et par quoi son regard est guidé » [20, Regimbeau, p.132] c'est-à-dire d'être capable de déterminer dans l'image le message dénoté mais surtout, dans le cas du luxe, de traduire en mots les éléments, - le *punctum* de Roland Barthes -, qui déclenchent le trouble du *Spectator* et en conséquence l'intérêt de la veille. Comme le dit Michel Melot,

« L'indexation n'est pas celle de l'image, mais celle des questions que l'on est susceptible de lui poser. » [21, Melot, p.361]. De ce fait, les mots retenus pour procéder à la recherche d'images puis à leur capitalisation dépendent des questions que l'on souhaite poser aux images et en conséquence sont liés aux objectifs assignés à la veille.

#### **4.2.2.2 Le manque de pertinence**

Les images diffusées sur Internet sont rarement indexées avec des mots-clefs. En conséquence, les recherches se fondent principalement sur le texte environnant les images. Aussi, bien que les difficultés d'interpréter l'image et de choisir les mots adéquats soient connues (et donc combattues), la recherche d'images offre souvent des résultats décevants, notamment lorsque la requête est un peu complexe.

Par exemple, Éleine Ménard relate une requête formulée avec le mot anglais « car » (voiture) qui a remonté dans ses résultats une photo de Sarah Palin (femme politique américaine républicaine) parce que cette photo avait été indexée par le nom de fichier « palin-in-the-car.jpg » [36, Ménard, p.7].

L'interprétation d'une image procède ainsi de la capacité de chacun à décoder les signes connotés, cette capacité dépendant de son milieu culturel d'appartenance. Elle relève ainsi de la subjectivité de celui qui met à disposition et de celui qui recherche, collecte, archive. Elle est aussi impactée par le contexte de gestion et de diffusion (comme le recours à des vocabulaires contrôlés ou implicites). La traduction de l'image en mots pose ainsi des problèmes importants. Cette difficulté n'est pas nouvelle pour les documentalistes, qui avant Internet travaillaient avec des banques d'images. Michel Melot témoigne : « Longtemps les documentalistes ont rêvé de saisir l'image en lui accolant une liste de descripteurs plus ou moins longue, de la réduire en mots, autant dire la réduire en miettes. L'image n'est ni un concept, ni une association d'idées successives. Elle ne parle pas. Elle n'a pas de grammaire. Elle montre, reproduit, informe, imagine et l'essentiel de ce qu'elle exprime est irréductible au langage. » [21, Melot, p.361].

Ainsi, selon Laurent Gervereau, « le seul équivalent de l'image est l'image elle-même. »<sup>38</sup>. Cette affirmation suggère de trouver une autre forme de veille, toujours créative et web, mais qui ne s'effectuerait non plus par des mots-clefs mais par son contenu, par les caractéristiques visuelles qui composent l'image.

---

<sup>38</sup> GERVEREAU, Laurent. Voir, comprendre, analyser les images. In [19, Bayle, p.15].

## 5 La reconnaissance par l'image

---

### 5.1 Le Content-Based Image Retrieval (CBIR)

#### 5.1.1 Définitions et explications

##### 5.1.1.1 50 ans de recherche

L'automatisation de la compréhension des images, ou vision par ordinateur, est une branche de l'intelligence artificielle [38, Harzallah, p.1]. Plus précisément, l'identification d'objets et d'individus, c'est-à-dire l'interprétation par la machine de chaque élément d'une image, relève de l'analyse de scènes [42, Ponce, p.1]. Différents problèmes se posent dans cette démarche :

- la variabilité :
  - des points de vue (dans le sens d'angles de vue) ;
  - de la disposition des sources lumineuses ;
  - des couleurs, textures, formes, etc.
- l'arrière-plan de la scène qui noie l'objet ou l'individu [42, Ponce, p.1].

Les premiers travaux sur l'interprétation des scènes visuelles remontent aux années 1960 ; ils consistaient à décrire les images et à en extraire des formes (des cylindres par exemple) [42, Ponce, p.2].

À la fin des années 1990, les équipes de Cordelia Schmid de l'INRIA<sup>39</sup> et de David Lowe de l'université de Colombie britannique (Canada) ont réorienté les recherches vers le pixel ou un ensemble parcimonieux de points d'intérêt, c'est-à-dire une zone de l'image saillante [42, Ponce, p.2]. De cette réorientation naissent des stratégies fondées sur différents algorithmes : en effet, si l'œil humain voit dans l'image des couleurs, des textures, identifie des objets, des personnes, la machine ne reconnaît dans cette même image qu'une série de chiffres.



Figure 1 - Ce qu'un ordinateur "voit"

Photographie de Kenneth Willardt [téléchargée le 28.09.2012]

<[http://www.kwpcf.com/uploads/images/3408/2006\\_056\\_D\\_051F2\\_4dc.h576.jpg](http://www.kwpcf.com/uploads/images/3408/2006_056_D_051F2_4dc.h576.jpg)>

---

<sup>39</sup> INRIA : établissement public consacré aux sciences numériques <<http://www.inria.fr/>>

### 5.1.1.2 Le principe de fonctionnement du CBIR<sup>40</sup>

Rémy Vieux définit le Content-Based Image Retrieval (CBIR) comme l'amélioration de « la recherche sur [...] des ensembles de données d'images fondée sur le contenu de l'image. » [41, Vieux, p.6]. Antoine Raulin<sup>41</sup> a décrit les principes de la reconnaissance automatique des photos :

- l'analyse pixellique de l'image : si un texte est composé de parties, sous-parties, paragraphes, l'image comprend des pixels [17, Haupais, p.43]. Toutefois, l'information visuelle n'est pas dans chaque pixel mais plutôt dans la manière dont ils sont arrangés [18, Rebai, p.12] ;
- la prise d'empreinte de l'image avec encodage des couleurs, des textures et des formes ;
- la création d'une base de connaissances d'images ;
- la détection d'images similaires ;
- le couplage éventuel avec des filtres d'analyse.

Les systèmes de CBIR se fondent sur les caractéristiques physiques de l'image, appelées de bas niveau (en opposition aux caractéristiques de haut niveau issues de l'indexation humaine) et extraites automatiquement [36, Ménard, p.2-3] :

- la couleur : il s'agit de la caractéristique la plus importante ;
- la texture : elle propose une description de la structure locale de l'image ;
- la forme : ses descripteurs offrent des indices importants quant à l'identité des objets présents dans l'image et permettent de décrire la structure géométrique du contenu [45, p.93, 96 et 98].

Ainsi, dans ces systèmes, l'utilisateur peut rechercher des images à partir de ces caractéristiques ou bien par la soumission d'un exemple d'image afin de trouver des images disposant des mêmes caractéristiques, c'est-à-dire similaires [36, Ménard, p.7]. En effet, les systèmes de CBIR sont capables d'extraire eux-mêmes les caractéristiques à rechercher. Ils impliquent des processus d'apprentissage importants à partir de fonds constitués d'un minimum de 50 000 images [44, GFII, p.53].

Les systèmes de CBIR qui utilisent les métadonnées de bas niveau offrent de bons résultats notamment dans les recherches extrêmement précises telles que les reconnaissances d'empreintes digitales, de visages, de plagats, etc. [45, p.56].

## 5.1.2 Exemples d'applications dans la recherche

### 5.1.2.1 Instituts spécialisés (INRIA, INA)

Le projet LEAR<sup>42</sup>, issu d'un partenariat entre l'INRIA et le CNRS, a pour principal sujet d'étude la reconnaissance visuelle d'objet et l'analyse de scène. Les objectifs sont la détection de catégorie d'objet, la recherche d'image, l'indexation vidéo et l'analyse des mouvements humains [40, Schmid]. L'équipe est saisie par des entreprises et met en place des solutions adaptées à leurs besoins.

L'Institut national de l'audiovisuel (INA) a participé au projet « Quaero », au sein d'une initiative collaborative de cinq ans comprenant vingt-quatre partenaires industriels

---

<sup>40</sup> L'objet du mémoire n'est pas un état de l'art des systèmes de CBIR. Celui-ci est présenté de manière très simple. Pour en savoir plus sur ces technologies, on peut lire [29], [38], [42], [45].

<sup>41</sup> RAULIN Antoine. Image numérique et nouvelles technologies. In [44, GFII, p.52].

<sup>42</sup> Learning and Recognition in Vision.

et universitaires. Ce programme étudie le traitement automatique des contenus multimédias et multilingues, l'objectif étant de développer des technologies d'accès au contenu des documents textuels, audios et images. Les domaines d'application sont larges, aussi bien grand public que professionnel, comme la télévision, le téléphone mobile, la gestion de contenus, la numérisation des bibliothèques, la diffusion du patrimoine audiovisuel<sup>43</sup>.

### 5.1.2.2 Les expériences et études menées avec un CBIR

Le système de CBIR a été choisi par une équipe universitaire chinoise pour identifier les espèces de papillons depuis les images présentes dans leur base de données, avec une efficacité atteignant jusqu'à 84 % de réussite pour certaines espèces. Outre l'efficacité, le choix du CBIR a été justifié par une série d'avantages :

- l'obtention automatique des caractéristiques d'une image ;
- l'affichage rapide des résultats pour les utilisateurs, grâce à la similarité entre les caractéristiques saisies et celles présentes dans la base de données ;
- l'interaction de l'utilisateur sous forme de feed-back<sup>44</sup> afin d'ajuster le système [37, Wang et al., p.25].

Dans un autre exemple d'application, une équipe de chercheurs du Massachusetts Institute Technology (MIT) a mis au point une bague prototype, dont le but est de permettre aux personnes mal ou non-voyantes d'identifier certains éléments de leur environnement. Des tests ont été effectués dans une boutique. La bague, connectée au smartphone par Bluetooth<sup>45</sup>, envoie les clichés réalisés par un clic de l'utilisateur pour les soumettre à une application de traitement de l'image, traduisant en mots ce qu'elle a reconnu. En fonction de la requête, elle transmet vocalement, par l'oreillette Bluetooth, les informations demandées (pour l'instant limitées à la distance, la couleur, au prix ou à la valeur d'un billet)<sup>46</sup>.

Enfin, IBM s'est très tôt intéressé à la recherche d'images et notamment par le contenu visuel. Je présente un de ces travaux pour le Musée de l'Hermitage (cf. p.59). Aujourd'hui, IBM travaille sur la réalité augmentée et développe des applications, en phase de tests, dans les points de vente américains. Ainsi, un consommateur pourra faire apparaître des informations telles que les ingrédients, le prix, etc. en passant l'appareil photo intégré à son portable devant un produit. La nouveauté d'IBM est que son application sera en mesure de reconnaître les produits de manière autonome. Cette technique de reconnaissance par l'image peut aider au quotidien les consommateurs ; elle procure surtout à IBM une masse de données (monnayable) sur les habitudes de consommation<sup>47</sup>.

Ces technologies ne sont pas accessibles à tout internaute, contrairement à la recherche textuelle. De plus, le fossé sémantique, c'est-à-dire l'écart entre l'information extraite de

---

<sup>43</sup> INA EXPERT. Quaero. In Ina Expert [site en ligne]. [s.d.] [consulté le 09.09.2012].

<<http://www.ina-sup.com/recherche/quaero>>

<sup>44</sup> Feed-back : ou rétroaction : mécanisme permettant à un dispositif automatisé de modifier son fonctionnement d'après les résultats obtenus par ce fonctionnement.

<sup>45</sup> Bluetooth : norme de communication sans fil, utilisée pour relier différents appareils par ondes hertziennes.

<sup>46</sup> DURAND-PARENTI Chloé. Une bague-caméra pour faciliter la vie des non-voyants. In Le Point.fr [site en ligne]. 30.08.2012 [consulté le 31.08.2012].

<[http://www.lepoint.fr/science/une-bague-avec-camera-embarquee-pour-faciliter-la-vie-des-non-voyants-30-08-2012-1500603\\_25.php](http://www.lepoint.fr/science/une-bague-avec-camera-embarquee-pour-faciliter-la-vie-des-non-voyants-30-08-2012-1500603_25.php)>

<sup>47</sup> SEDOURAMANE Hugo. IBM teste une application de réalité augmentée en points de vente. In Le Journal du Net [site en ligne]. 03.07.2012 [consulté le 03.07.2012].

<<http://www.journaldunet.com/ebusiness/internet-mobile/ibm-application-realite-augmentee-0712.shtml>>

la donnée visuelle et l'interprétation de cette donnée par l'utilisateur dans un contexte, demeure un problème quant à la pertinence de ces technologies. Elles sont au cœur des travaux de recherche et donc un objet davantage orienté système qu'utilisateur. Aussi, les technologies de CBIR peuvent-elles convenir pour une recherche d'image menée dans un environnement professionnel non-scientifique ?

## 5.2 Le CBIR au défi du monde réel

### 5.2.1 Les pratiques de recherche d'images

#### 5.2.1.1 Les besoins des utilisateurs

Selon Pascal Frion<sup>48</sup>, le chercheur d'information néophyte sur Internet est caractérisé par les attributs suivants :

- il utilise entre un et deux mots-clefs par requête ;
- il ne fait pas de distinction entre bandeau d'adresse et bandeau de recherche ;
- il fait peu de distinction entre les types d'outils ;
- il utilise très peu les outils payants ;
- il recourt aux outils gratuits ;
- il a peu de perception du Web profond (invisible)<sup>49</sup> ;
- il préfère la rapidité de l'outil à ses qualités ;
- il utilise les outils sans préoccupation des options de recherche avancée ;
- il utilise deux opérateurs logiques au maximum ;
- il manque d'organisation dans la recherche ;
- il a peu d'intérêt pour la recherche multisupports ;
- il a peu ou manque de formation en informatique, sur Internet, en anglais, en veille stratégique.

Trois profils sont alors identifiés :

- le navigateur (« browser ») : sans but, l'internaute effectue des recherches aléatoires, par sérendipité, occasionnant de longues sessions. Il accorde de l'importance à la facilité de l'utilisation et de la manipulation du système ;
- le surfeur (« surfer ») : l'internaute progresse avec une certaine clarté vers une finalité déterminée, mais encore de façon exploratoire au début de sa session de recherche ;
- le chercheur (« searcher ») : l'internaute sait très bien ce qu'il recherche et comment rechercher, à l'aide de requêtes cohérentes aboutissant à des sessions courtes. Aussi privilégiera-t-il des systèmes offrant des résultats complets et clairs [29, Datta et al., p.9].

Les utilisateurs manifestent comme premiers besoins l'ergonomie du système de recherche et la présentation des résultats. Suite à une étude auprès de ses étudiants, JungWon Yoon a mis en évidence que les fonctionnalités importantes étaient la quantité de réponses pertinentes, la qualité et la taille des images [28, Yoon, p.271]. Il a aussi relevé la difficulté de trouver les mots-clefs appropriés à la recherche. Il a également constaté que les

---

<sup>48</sup> FRION Pascal. Dans la recherche d'information. Observations sur le manque de méthode sur Internet. In [4, Khénissi et al., p.147].

<sup>49</sup> Web profond ou invisible : partie non indexée par les moteurs de recherche généralistes.

internautes, plus ou moins amateurs, souhaitent pouvoir effectuer des requêtes sur des images avec des métadonnées « comme la date de l'image, le type, la taille, la qualité, le copyright et la source des images » afin de réaliser une première sélection [28, Yoon, p.273]. Or nous avons vu précédemment que ces métadonnées pouvaient être relativement faciles à extraire par les systèmes d'indexation automatique.

Néanmoins, il est difficile pour l'utilisateur de traduire sa requête en caractéristiques de bas niveau de CBIR. Aussi, une modalité de recherche, permettant de contourner les problèmes de mots-clés et ces caractéristiques, est l'interface de requêtes par l'exemple, appelée plus couramment recherche par similarité (cf. p.35) : l'utilisateur choisit une image, la transmet au système en lui demandant de trouver des images ayant les mêmes caractéristiques. Certains systèmes proposent de rechercher par région d'image et non sur l'image entière [45, p.28]. Encore faut-il au départ trouver l'image pertinente à soumettre.

### **5.2.1.2 La recherche textuelle incontournable**

En conséquence, les outils de recherche par le contenu visuel ont pour le moment une utilisation limitée [43, Nobis, p.5]. De ce fait, la recherche d'image repose sur l'exploitation commune de la recherche textuelle et de l'interrogation du contenu visuel. Toutefois, les outils grand public, et spécifiquement les moteurs de recherche généralistes qui sont les plus populaires pour effectuer des sessions de recherche, privilégient les métadonnées et le texte descriptif autour de l'image, par un système de pondération. Dès lors, la recherche d'images pâtit de la dépendance vis-à-vis du contenu linguistique associé à l'image et de l'ignorance du contenu de cette même image [43, Nobis, p.2].

De plus, l'utilisateur, en raison de son expérience de recherche par les moteurs auxquels il est familiarisé (au moins pour la recherche simple), préfère interagir avec le système par des mots-clés issus du langage naturel, donc à un niveau sémantique [45, p.56].

De ce fait, les tendances actuelles des travaux en matière de recherche d'image s'orientent vers la fusion des deux modalités, le texte et le contenu visuel. En effet, d'une part, la recherche par mots-clés peut poser des problèmes d'ambiguïté d'interprétation de la requête et d'autre part, les seules caractéristiques de bas niveau des systèmes de CBIR perdent en qualité sémantique [39, He et al., p.223].

Dès lors, la combinaison des deux systèmes permettrait de pallier les problèmes respectifs. En effet, « Le couplage de solutions CBIR avec l'analyse sémantique et linguistique du contenu textuel associé à l'image permet d'affiner les résultats en optimisant la pertinence des réponses et la minimisation du bruit. » [44, GFII, p.53]. Comme le démontrent H. Bannour et C. Hudelot dans leurs travaux, l'approche sémantique fondée uniquement sur du texte est limitée dans la recherche d'images. Ainsi, dans WordNet<sup>50</sup>, les concepts de « Baleine » et « Humain » sont plus proches que ceux de « Baleine » et « Requin », en raison du critère de classe, celle des mammifères. Cependant, dans le cadre d'une recherche d'images, la « Baleine » et le « Requin » sont des concepts plus proches car ces deux animaux évoluent dans le même environnement et partagent des caractéristiques visuelles. En conséquence, l'approche sémantique, donc par des requêtes textuelles, doit être capable d'intégrer des similarités conceptuelles et contextuelles, la pertinence de la recherche d'images ayant des exigences de résultats différentes de la simple recherche textuelle [35, Bannour et al., p.2-4].

---

<sup>50</sup> WordNet : base de données lexicale développée par les linguistes du laboratoire des sciences cognitives de l'université de Princeton (New Jersey, États-Unis).

## 5.2.2 Les moteurs de recherche généralistes et les images

### 5.2.2.1 Le mode d'emploi d'un moteur de recherche

Lors des enquêtes sur les besoins des utilisateurs présentées précédemment, l'utilisation des moteurs de recherche dits généralistes pour effectuer des recherches d'images est quasi générale. Mes entretiens menés auprès des membres des services Marketing et Achat d'art ont donné les mêmes résultats<sup>51</sup>.

Un moteur de recherche est constitué de trois éléments [2, Foenix-Riou, p.16-17] :

- un robot (ou « bot », « crawler », « spider ») qui parcourt le Web de liens en liens et rapatrie dans sa base une copie des pages visitées. Cette démarche n'est pas exhaustive : un contenu d'une page web n'est jamais intégralement indexé, ni toutes les pages d'un site web, ni toutes les pages web ;
- un index qui comprend l'ensemble des mots de toutes les pages rapatriées par le robot et qui relie ainsi les mots à l'URL<sup>52</sup> des pages dont ils proviennent. En général, le texte intégral et les balises méta (titre, mots-clefs) sont indexés ;
- une interface web qui permet à l'internaute d'effectuer une requête par mots sur l'index du moteur.

Ainsi, faire une recherche sur un moteur revient à effectuer une requête par mots sur le texte intégral de toutes les pages web. En raison de la quantité d'informations disponibles, la recherche génère du bruit, encore plus important avec les images puisque les robots éprouvent des difficultés à les indexer [19, Bayle, p.42]. En effet, la recherche de l'image par les moteurs de recherche en langage naturel est effectuée sur « les titres, les descriptions, les résumés, ou même encore la balise ALT<sup>53</sup> [...] et certaines balises du code source de la page HTML où figure l'image » [19, Bayle, p.41]. Toutefois, ces éléments sont parfois mal ou non renseignés et ainsi le référencement et donc la visibilité de ces images par les moteurs de recherche sont affectés.

Afin de réduire le bruit, l'internaute peut utiliser les fonctionnalités de recherche avancée. Néanmoins, comme présenté en page 37, il accuse un déficit de maîtrise des outils informatiques d'Internet. La simplicité et la rapidité de la recherche simple prévalent.

### 5.2.2.2 Un exemple caractéristique : Google Search Image

En 2011, Google Search Image annonçait indexer plus de 10 milliards d'images [2, Foenix-Riou, p.178]. Il est l'outil le plus utilisé dans l'entreprise X pour effectuer des recherches d'images externes. Il est également le moteur le plus populaire d'une manière générale<sup>54</sup>. Par exemple, l'affichage des résultats sous forme de vignettes, bien qu'ayant posé quelques problèmes juridiques<sup>55</sup>, a été repris par ses concurrents. De plus, il a développé des fonctionnalités qui facilitent les requêtes des internautes. Ainsi, depuis juin 2011, Google

---

<sup>51</sup> Plus d'une visite sur deux sur Internet est issue d'un moteur de recherche (cf. Annexe 3) [27, Médiamétrie eStat, p.1].

<sup>52</sup> URL : de l'anglais Uniform Resource Locator, soit en français localisateur uniforme de ressource, souvent appelée « adresse web ». L'URL désigne une chaîne de caractères utilisée pour adresser les ressources (document HTML, image, son, etc.) du Web.

<sup>53</sup> Balise ALT : balise HTML d'une image permettant de la décrire avec des mots.

<sup>54</sup> Google représente plus de 93 % des moteurs de recherche en France (cf. Annexe 3) [27, Médiamétrie eStat, p.1].

<sup>55</sup> BATTISTI Michèle. Google Images ou le droit d'auteur au défi du droit à l'information. In Paralipomènes [blog en ligne]. 22.06.2011 [consulté le 28.08.2012].

<<http://paralipomenes.net/wordpress/archives/4983>>. ISSN 2260-3336



Search Image propose la recherche inversée, depuis un fichier image ou depuis l'URL d'une image, afin d'afficher les images similaires<sup>56</sup>.

De plus, pour les internautes anglo-saxons, la recherche par image utilise désormais le « Knowledge Graph », lancé en mai 2012 (cf. p.26)<sup>57</sup>, alliant des capacités sémantiques à l'algorithme de vision artificielle utilisé seul jusqu'à cette date. Sean O' Malley, ingénieur à Mountain View (firme de Google), explique ainsi sur le blog dédié à la recherche que dorénavant le moteur reconnaîtra avec plus de précision les images : « [...] désormais vous verrez que l'image ci-dessus n'est pas seulement une "fleur", mais plus spécifiquement un oiseau de paradis, qui est populaire à Hawaii. »<sup>58</sup>.

Outre ces innovations technologiques, l'ergonomie de Google Search Image contribue beaucoup à son succès :

- le recours traditionnel aux mots-clefs, modalité rassurante pour l'internaute car connue ;
- le glisser-déposer des images pour une recherche inversée ;
- l'utilisation de filtres pour affiner les résultats : par date, taille, couleur, type (exemples : visage, photos, clipart, dessins au trait).

Ainsi, le réflexe Google en matière de recherche, qui incite les internautes à utiliser cet outil y compris pour des images, est renforcé par la simplicité de son utilisation. Je présenterai par ailleurs (cf. p.60) la performance de ses résultats. Google est ainsi le seul outil gratuit et accessible à tous à rassembler, la vision artificielle, la sémantique et l'ergonomie.

### **5.2.2.3 D'autres moteurs : Bing et Yahoo !**

Bing, développé par Microsoft, investit également la recherche d'images. Elle constitue un marché important puisqu'aux États-Unis elle représente 7 % des recherches effectuées sur Bing<sup>59</sup>. Comme Google Search Image, Bing affiche ses résultats sous forme de vignettes. Il propose un historique des recherches et surtout un système ergonomique de filtres pour affiner les résultats : par taille, couleur, style (tous, photo, clipart, dessin), disposition (toutes, carrées, larges, hautes), personne (tous, visages uniquement, têtes et épaules).

Yahoo ! présente ses résultats de recherche, fournis par Bing<sup>60</sup>, sous forme de vignettes, avec quelques filtres mais moins développés que ceux de Google et de Bing : par date, taille, couleur (noir et blanc ou couleur). Ainsi, à moins de constater une meilleure pertinence de résultats, les fonctionnalités proposées par Yahoo ! n'incitent pas à le préférer.

---

<sup>56</sup> ANDRIEU Olivier. Google améliore (à son tour) la recherche d'images. In Abondance actualités [liste de diffusion]. 04.07.2012 [consulté le 07.07.2012].

<<http://www.abondance.com/actualites/20120704-11638-google-ameliore-a-son-tour-la-recherche-dimages.html>>

<sup>57</sup> Ibidem.

<sup>58</sup> O'MALLEY Sean. Find smarter, more comprehensive Search by Image results. In Google Inside Search. The Official Google Search Blog [blog en ligne]. 02.07.2012. [consulté le 17.08.2012]. <<http://insidesearch.blogspot.fr/2012/07/find-smarter-more-comprehensive-search.html>>

<sup>59</sup> G. Jérôme. Bing et Yahoo ! : recherche d'images revue. In Génération Nouvelles technologies [site en ligne]. 22.06.2012 [consulté le 02.07.2012]. <<http://www.generation-nt.com/bing-yahoo-recherche-images-actualite-1595331.html>>

<sup>60</sup> Ibidem.

## 6 La reconnaissance faciale

---

### 6.1 Les technologies mises en œuvre

#### 6.1.1 Reconnaître des visages

J'ai annoncé que l'une des demandes était de pouvoir identifier le nom d'un mannequin (ou d'une égérie) ou les photos qu'il/elle a pu faire à partir d'une image repérée (cf. p.20). Aussi, la technologie de reconnaissance faciale semble indiquée pour répondre à ce besoin de recherche.

Cette technologie a été développée pour identifier les personnes suspectes ou indésirables dans les casinos, les gares, les stades de sport, tous les espaces publics en général. Elle s'est également imposée au sein des aéroports pour des raisons de sécurité suite aux attentats et tentatives d'attentats aériens. Les essais en laboratoire ont donné d'excellents résultats mais ils ne sont pas représentatifs des conditions du monde réel. Ainsi, le logiciel de reconnaissance perd de sa pertinence face à la variabilité des sources lumineuses ou d'angles de prises de vue, soit les mêmes failles que celles éprouvées par le CBIR (cf. p.34), mais aussi en fonction de la position des caméras et de la distance avec les sujets. L'autre difficulté du système est que la reconnaissance faciale se fonde sur la confrontation du visage capté avec celui enregistré dans une base afin de comparer les deux images ; or, l'image de référence peut être ancienne, la plupart des passeports étant valides dix ans. Malgré cela, on considère que le logiciel de reconnaissance faciale obtient de meilleurs résultats qu'un garde-frontière en situation depuis plus de deux heures devant vérifier visuellement de multiples types raciaux de passagers [31, Gohringer, p.5-7].

L'équipe IMEDIA de l'INRIA explore cette question. Des techniques de détection (localisation du visage, voire de ses composants) et de segmentation (délimitation spatiale de la surface du visage dans une image) sont développées : elles n'identifient pas une personne détectée, elles « apportent [...] une information spatiale de présence et de localisation du visage dans une image [...] étape indispensable et préalable avant de procéder à la recherche de personnes dans les documents multimédias. » [45, p.109].

Les méthodes de détection et de reconnaissance de visages 2D<sup>61</sup> sont sensibles à l'éclairage et à l'orientation de prise de vue du visage. D'autres facteurs, tels que l'expression faciale, le maquillage, les lunettes, etc. accentuent les difficultés de reconnaissance [45, p.110]. Or, dans mon projet, ces facteurs sont particulièrement prégnants. En effet, les photographies du luxe, de nature artistique, mettent en scène des mannequins, des égéries dans des atmosphères jouant sur la lumière, les poses suggestives, les accessoires, l'expression des visages, puisque peu voire pas de textes sont utilisés pour véhiculer le message. Et la demande est encore plus impactée par le maquillage, branche de l'entreprise dans laquelle est positionné le service (cf. les exemples d'images publicitaires en Annexe 4). En conséquence, les besoins d'identification de mannequins et de recherche utilisant les technologies de reconnaissance faciale s'annoncent délicats à satisfaire.

#### 6.1.2 Les problèmes juridiques comme freins de son essor ?

Si la reconnaissance faciale est une technologie beaucoup exploitée, les problèmes juridiques qu'elle engendre ne vont-ils pas inciter les entreprises travaillant dessus à s'orienter vers d'autres sujets d'études, la reconnaissance faciale devenant trop délicate ?

---

<sup>61</sup> 2D : objet qui ne comporte que les dimensions de la longueur et de la largeur.

Facebook a lancé en juin 2011 un outil de reconnaissance faciale (« Tag suggest ») qui consiste à reconnaître automatiquement les visages des membres du réseau social et à associer les images au profil des personnes concernées. Les utilisateurs, qui peuvent choisir de désactiver cette fonction<sup>62</sup>, ont ainsi créé la plus grande base de données au monde répertoriant 300 millions de photos<sup>63</sup>. En août 2011, le land de Hambourg en Allemagne s'est alarmé du stockage des informations privées des utilisateurs de Facebook, en contradiction avec la loi européenne. Cet été, la Norvège a lancé une enquête sur le traitement des photos stockées et des données ainsi disponibles. Finalement, Facebook a annoncé fin septembre suspendre son service dans l'Union européenne et effacer les données collectées au plus tard le 15 octobre. Le réseau social travaille néanmoins pour trouver une solution permettant d'utiliser son outil en respect des règles européennes<sup>64</sup>.

Sur demande des utilisateurs et pour éviter tout procès, YouTube, site d'hébergement et de partage en ligne de vidéos (propriété de Google), a mis au point un système de floutage, qui permet de masquer les visages présents dans une vidéo. Néanmoins, la technologie étant nouvelle, le site prévient que le système a parfois des difficultés « à détecter les visages en fonction de l'angle, l'éclairage, les obstructions et de la qualité de la vidéo. »<sup>65</sup>. La reconnaissance faciale avait mis en danger les activistes protestataires durant les révolutions arabes de 2011. YouTube a donc dû proposer une solution, qui est peut-être une première étape vers l'anonymat visuel.

Des associations de défense des libertés individuelles s'intéressent à la question du fichage généralisé rendu possible grâce aux applications de reconnaissance faciale. Par exemple, une équipe de l'université Carnegie Mellon de Pittsburg en Pennsylvanie a photographié au hasard des étudiants sur le campus et est parvenue à identifier près d'un tiers d'entre eux en comparant leurs photos à celles publiées et tagguées sur Facebook et indexées par Google (donc en libre-accès), grâce à un système de reconnaissance faciale. Son rapport final se conclut ainsi : « Désormais, on peut partir d'un visage anonyme pris dans la rue et aboutir à une masse d'informations sensibles sur cette personne. » [34, Eudes].

Or, d'après Emily Turettini<sup>66</sup>, 665 millions de smartphones et tablettes seront équipés de la reconnaissance faciale fin 2012. Ainsi, les technologies permettant d'améliorer cet axe de recherche pourraient être ralenties voire suspendues en raison des conséquences de leur utilisation pour le grand public. Il faut donc être vigilant sur l'évolution de ces systèmes, qui néanmoins ont déjà engendré des réalisations impressionnantes.

---

<sup>62</sup> LE MONDE.FR, BLOOMBERG. La Norvège se penche sur la reconnaissance faciale de Facebook. In Le Monde.fr [site en ligne]. 06.08.2012 [consulté le 15.08.2012]. <[http://www.lemonde.fr/technologies/article/2012/08/06/la-norvege-se-penche-sur-la-reconnaissance-faciale-de-facebook\\_1742878\\_651865.html](http://www.lemonde.fr/technologies/article/2012/08/06/la-norvege-se-penche-sur-la-reconnaissance-faciale-de-facebook_1742878_651865.html)>

<sup>63</sup> TURRETTINI Emily. Comment se défendre contre la reconnaissance faciale ? In Huffington Post [site en ligne]. 29.08.2012 [consulté le 29.08.2012]. <[http://www.huffingtonpost.fr/emily-turettini/reconnaissance-faciale-logiciels-facebook\\_b\\_1831555.html?utm\\_hp\\_ref=france](http://www.huffingtonpost.fr/emily-turettini/reconnaissance-faciale-logiciels-facebook_b_1831555.html?utm_hp_ref=france)>

<sup>64</sup> AFP. Facebook suspend son outil de reconnaissance faciale dans l'UE. In Le Point.fr [site en ligne]. 21.09.2012 [consulté le 24.09.2012] <[http://www.lepoint.fr/high-tech-internet/facebook-suspend-son-outil-de-reconnaissance-faciale-dans-l-ue-21-09-2012-1508879\\_47.php](http://www.lepoint.fr/high-tech-internet/facebook-suspend-son-outil-de-reconnaissance-faciale-dans-l-ue-21-09-2012-1508879_47.php)>

LE MONDE.FR, AFP. Facebook suspend sa reconnaissance faciale en Europe. In Le Monde.fr [site en ligne]. 22.09.2012 [consulté le 24.09.2012] <[http://www.lemonde.fr/technologies/article/2012/09/22/facebook-suspend-sa-reconnaissance-faciale-en-europe\\_1764136\\_651865.html](http://www.lemonde.fr/technologies/article/2012/09/22/facebook-suspend-sa-reconnaissance-faciale-en-europe_1764136_651865.html)>

<sup>65</sup> CONWAY Amanda. Face blurring : when footage requires anonymity. In Broadcasting Yourself. The Official YouTube blog [blog en ligne]. 18.07.2012 [consulté le 26.07.2012]. <<http://youtube-global.blogspot.de/2012/07/face-blurring-when-footage-requires.html>>

<sup>66</sup> Idem.

## 6.2 Exemples concrets de reconnaissance faciale

### 6.2.1 Google et la reconnaissance des chats<sup>67</sup>

Google a mis en place, dans son laboratoire secret, une équipe de recherche sur la simulation du cerveau humain. Les scientifiques ont ainsi créé le plus grand réseau neural pour apprentissage automatique. Les 16 000 processeurs informatiques de ce cerveau virtuel dénombrent plus d'un milliard de connexions et tournent sur Internet afin d'apprendre par eux-mêmes. Les chercheurs leur ont soumis des vignettes d'images, choisies de façon aléatoire, chacune extraite des 10 millions de vidéos en ligne sur YouTube, mais ils n'ont pas dit aux ordinateurs ce qu'ils devaient rechercher ni ne leur ont appris au préalable à reconnaître des items.

Le résultat est que ce réseau neural a fait ce que les humains font sur YouTube : s'intéresser aux chats ! En effet, à force de traiter des images de chats, le réseau neural a fini par inventer lui-même "son" concept de chat.

Comme nous l'avons vu, quand un ordinateur n'ayant aucun apprentissage regarde une image, il n'y voit que des milliers de pixels de différentes couleurs. Avec une période d'apprentissage, il peut reconnaître certaines caractéristiques, comme celles attestant la présence d'un visage humain, et les identifier. Dans cette étude, les ordinateurs n'avaient pas d'apprentissage préalable et ne savaient pas ce qu'ils devaient regarder : ils devaient apprendre par eux-mêmes à déterminer les caractéristiques pertinentes de chaque photo. C'est l'arrangement des pixels, image après image, qui a suggéré que les caractéristiques d'une tête de chat étaient pertinentes et donc devaient être identifiées.

Les résultats encourageants ont fait sortir le projet du laboratoire pour le continuer dans une division de recherche. En effet, les applications potentielles concernent la recherche d'image, la reconnaissance vocale et la traduction linguistique automatique. Ainsi, il faut demeurer vigilant sur l'actualité des recherches en matière de vision et d'intelligence artificielles car l'algorithme "miracle" sera peut-être bientôt découvert.

En attendant, Google est sur le point de racheter Viewdle<sup>68</sup>, une firme ukrainienne de technologie de reconnaissance faciale, qui permet de tagger automatiquement les photos. Elle serait utilisée, entre autres, pour le réseau social Google+, Android<sup>69</sup> et Picasa<sup>70</sup>.

### 6.2.2 Scene Tap à La Nouvelle Orléans<sup>71</sup>

La reconnaissance faciale dans l'espace public n'a pas que des applications de surveillance ou de protection. Ainsi, la start-up<sup>72</sup> Scene Tap a mis au point une application disponible sur smartphone afin de fournir des informations en temps réel sur la fréquentation des vingt-cinq bars de la ville. Par exemple, pour The Library : « Remplissage : 20 %. Ambiance : calme. Âge moyen des clients : 29 ans. Hommes : 40 %, femmes : 60 %. ».

---

<sup>67</sup> [32, Oresmus], [33, Markoff].

<sup>68</sup> MILLS Elinor. Google buying face recognition firm Viewdle. In CENT News [site en ligne]. 01.10.2012 [consulté le 03.10.2012]. <[http://news.cnet.com/8301-1023\\_3-57523863-93/google-buying-face-recognition-firm-viewdle/](http://news.cnet.com/8301-1023_3-57523863-93/google-buying-face-recognition-firm-viewdle/)>

<sup>69</sup> Android : système d'exploitation open source pour les terminaux mobiles élaboré par l'entreprise Android rachetée par Google.

<sup>70</sup> Picasa : logiciel gratuit de gestion et de partage des photos proposé par Google.

<sup>71</sup> [34, Eudes].

<sup>72</sup> Start-up : jeune entreprise innovante avec un potentiel de développement fort et ayant besoin, de ce fait, d'un investissement important.

Ces données sont rendues disponibles grâce au dispositif installé par la start-up. Deux caméras de surveillance sont reliées à un ordinateur : l'une au-dessus de la porte compte les entrées et les sorties, calculant ainsi le taux de remplissage ; l'autre filme les clients tandis que l'ordinateur traite les caractéristiques des visages. Les images sont transférées à un serveur couplé à une base de données, catalogue d'environ 500 000 visages de volontaires anonymes, créé à l'origine par Intel et revendu à Scene Tap. Comme dans tout processus de reconnaissance faciale, le serveur informatique compare les visages filmés à ceux présents dans la base de données puis établit les statistiques de fréquentation.

Pour le fondateur de la start-up, Marc Doering, les logiciels sont quasi infaillibles, puisqu'ils étudient la texture de la peau, la forme et la taille de chaque partie de la tête (c'est-à-dire les caractéristiques de bas niveau considérées comme pertinentes pour les visages). Finalement, la seule limite pourrait venir de la gestion de la base de données qui doit sans cesse trouver de nouveaux moyens, légaux, pour s'enrichir.

### 6.2.3 Face.com : la reconnaissance visuelle en réseau<sup>73</sup>

J'ai évoqué précédemment le service de reconnaissance faciale de Facebook (cf. p.42). Pour se renforcer dans ce domaine, Facebook a racheté cet été la start-up israélienne Face.com<sup>74</sup>. Cette dernière avait déjà créé pour le réseau social deux applications :

- « Photo Finder », pour retrouver des photos de soi ou de ses amis publiées par des tiers ;
- « Photo Tagger », qui prend les albums de chaque membre, les trie, taggue automatiquement les visages connus et aide à identifier les autres.

La société dispose d'une autre application, « Klik », qui identifie avant même de prendre la photo : lorsque l'utilisateur regarde une personne sur l'écran de son smartphone, son nom s'inscrit au-dessus de sa tête. Ses logiciels sont également efficaces pour les visages en arrière-plan.

Face.com ne s'arrête pas à ses propres recherches. Ainsi, elle a ouvert sa plateforme à des développeurs indépendants étrangers, créateurs d'applications inédites. Par exemple, la société Farelock de Baltimore (Maryland, États-Unis) a créé une « pointeuse faciale » pour les usines et les chantiers, qui se substitue à la carte de pointage ou à la badgeuse : pour signaler leurs arrivées et départs, les employés présentent leur visage à une tablette.

## 6.3 Les limites technologiques

J'ai exposé dans cette deuxième partie les technologies associées à la recherche par le contenu de l'image, leurs potentialités et leurs limites.

Si, à l'aube du développement de ces technologies, les chercheurs étaient optimistes quant à la découverte rapide d'un algorithme pertinent en la matière, force est de constater

---

<sup>73</sup> ZDNET FRANCE. Reconnaissance faciale : Facebook s'offre l'éditeur Face.com.

In Challenges.fr [site en ligne]. 20.06.2012 [consulté le 02.07.2012].

<<http://www.challenges.fr/high-tech/20120620.ZDN7259/reconnaissance-faciale-facebook-s-offre-l-editeur-face-com.html>>

ORESKOVIC Alexei, HUET Nathalie. Facebook s'offre Face.com et sa reconnaissance faciale. In Challenges.fr [site en ligne]. 18.06.2012 [consulté le 02.07.2012].

<<http://www.challenges.fr/high-tech/20120618.REU9189/facebook-s-offre-face-com-et-sa-reconnaissance-faciale.html>>

[34, Eudes].

<sup>74</sup> Face.com : <<http://face.com/>>

que, vingt ans après les premières études sur la recherche pixellique, le graal algorithmique demeure encore inconnu et fait l'objet d'une véritable course entre équipes de chercheurs.

Aussi, malgré ses résultats très encourageants sur l'apprentissage automatique des chats par leur grand réseau neural, Google est très circonspect et s'explique : « Il pourrait sembler surprenant que ce type de modèle de reconnaissance soit si difficile. Après tout, un enfant de trois ans peut le faire. Mais les réseaux neuraux d'un cerveau de trois ans contiennent bien plus de connexions qu'une configuration massive de Google. ». Jeff Dean, membre de l'équipe de recherche de Google, raconte aussi que les ordinateurs apprennent des quantités de concepts qui ont peu de sens pour les humains : « Par exemple, ils sont intrigués par des objets semblables à des outils orientés à 30°, incluant les spatules et les pince-nez. » [32, Oresmus].

D'une manière générale, Éline Ménard relève trois faiblesses principales dans le système de CBIR [36, Ménard, p.4] :

- la pertinence des résultats obtenus ;
- la vitesse d'exécution ;
- la complexité de mise en œuvre et d'utilisation.

Ainsi, pour la plupart, ces systèmes demeurent à un stade expérimental.

Dès lors, la recherche textuelle semble incontournable, à condition que les images pertinentes correspondantes aient bien été indexées d'une part, et de trouver les bons mots-clefs d'autre part.

En effet, comme le résume Coline Haupais, « Le problème de la recherche d'image via métadonnées réside dans ce que le texte fonctionne par signes (le langage) là où l'image fonctionne par analogie. Enfin, bien souvent, les utilisateurs se souviendront de l'ambiance d'une image, de ce qu'elle leur a évoqué plutôt que ce qu'elle représentait en réalité. Il est donc extrêmement difficile de retranscrire en requête informatique une demande d'images d'utilisateurs. » [17, Haupais, p.40].

En conséquence, en me fondant sur l'identification des besoins et la définition de la veille d'une part, et sur l'étude de l'image et des technologies annexes d'autre part, je vais tenter dans une troisième et dernière partie de proposer des scénarios pouvant satisfaire la demande, malgré les limites qui ont été mises en évidence.

# **Troisième partie**

## **Scénarios pour une veille visuelle**

## 7 Le rappel du contexte

---

### 7.1 Le cahier des charges fonctionnel

#### 7.1.1 L'analyse de l'existant

L'assistante Achat d'art de l'entreprise X, grand acteur du secteur du luxe, réalise différentes tâches dont le point commun est la nature des informations qu'elle utilise : l'image.

Sa mission principale est de réaliser des briefs créatifs pour les campagnes publicitaires. Elle recherche des visuels qui répondent aux concepts demandés dans le brief. Elle doit aussi pouvoir satisfaire des demandes ponctuelles sur des images publicitaires dans le domaine du luxe.

Pour ce faire, elle recourt principalement à des outils dématérialisés et surtout à Internet. Une sélection de sites favoris, de newsletters, de moteurs de recherche a été réalisée au fur et à mesure des pratiques mais l'assistante Achat d'art se demande si d'autres sources ou d'autres outils, qu'elle ne connaît pas, ne seraient pas meilleurs pour l'aider dans ses démarches. Elle n'a pas de formation de documentaliste et recourt donc à des outils simples et intuitifs.

En outre, les images qu'elle a collectées depuis 2007 (près de 6 000) sont regroupées au sein d'un dossier personnel, organisé en sous-dossiers : cette organisation ne lui permet pas d'effectuer des recherches au sein des visuels, qu'elle doit consulter l'un après l'autre pour éventuellement trouver des pistes créatives intéressantes à soumettre à la responsable du service.

Ces démarches lui prennent beaucoup de temps or les recherches ne constituent pas son cœur de métier et l'empêchent de se concentrer sur ses missions initiales.

Des veilles existent au sein de l'entreprise toutefois elles ne répondent pas à ses besoins en termes de collecte d'images.

#### 7.1.2 L'analyse des besoins : les fonctions de l'outil idéal

Trois besoins principaux ont été identifiés :

- répondre de façon rapide et pertinente à une demande d'images publicitaires d'après le brief créatif ;
- trouver une information depuis un visuel : identifier tel mannequin, tel photographe de telle campagne ;
- capter les nouveautés : en matière de créations, de tendances, mais aussi de mannequins, de photographes, etc.

Le dernier besoin, qui peut faire l'objet d'une veille, semble être le premier à satisfaire. En effet, c'est au sein de la collection d'images ainsi constituée par la veille que l'assistante Achat d'art pourra effectuer ses premières recherches pour répondre aux briefs. De plus, les images livrées par la veille pourraient être accompagnées des informations relatives aux noms des photographes, mannequins, marques. La satisfaction de ce besoin pourrait ainsi accélérer ou du moins faciliter celle des deux autres.

Ainsi, on en déduit que l'outil devra :



- être capable de faire de la veille et de la recherche ;
- réduire et optimiser le temps de recherche ;
- être interopérable avec les systèmes existants, notamment le DAM, pour y classer les images collectées ;
- être accessible à d'autres membres du service ;
- être simple ;
- assurer la confidentialité et la sécurité de la veille et de la recherche.

Ces premières conclusions servent de point de départ à la réflexion du cahier des charges. En effet, le marché des outils de veille est riche : « les annuaires électroniques, les moteurs de recherche, les métamoteurs, les listes de diffusion, les outils de surveillance, le Web visible et invisible, les blogs, les flux ou fils RSS, le Web 2.0, l'intranet, l'extranet, les logiciels de veille propriétaire et open source etc. » [4, Khénissi et al., p.140].

La norme AFNOR XP X 50-053 définit ainsi une prestation de veille : « Réalisation d'un ensemble de fonctions selon un cahier des charges spécifique à l'entité cliente, et faisant appel aux outils, sources et expertise proposés par un prestataire interne ou externe. » [15, AFNOR, p.6]. En conséquence, pour déterminer l'outil ou les outils les mieux adaptés, j'ai réalisé un cahier des charges, fonctionnel et technique, dans lequel j'ai traduit les besoins en fonctionnalités et en solutions pratiques.

Pour des raisons de confidentialité, je n'évoquerai dans les prochains paragraphes que les grands axes de réflexion - les spécifications applicatives pour le cahier des charges fonctionnel et les spécifications techniques et les risques pour le cahier des charges technique -, sans entrer dans les détails.

### **7.1.3 Les spécifications applicatives**

J'ai élaboré un tableau en trois parties afin de préciser les spécifications applicatives (c'est-à-dire les services et qualités attendus) correspondant à :

- l'ensemble de la demande (le projet initial : veille et recherche) ;
- la prestation de veille ;
- un moteur de recherche.

Chaque partie est composée des fonctionnalités que l'outil ou les outils retenus devront disposer.

Pour la première partie, j'ai déterminé six fonctionnalités fondées sur le processus de veille et intégrant la fonction de recherche dont le besoin avait été manifesté. Ces fonctionnalités sont traduites en objectifs, pour lesquels j'ai défini les leviers envisageables, les éléments et les interactions prévisibles, un ordre de classement (selon les besoins exprimés) et le niveau de flexibilité (obtention imposée, souhaitée, révisable, négociable). Je pourrais résumer ma démarche par le schéma ci-après :

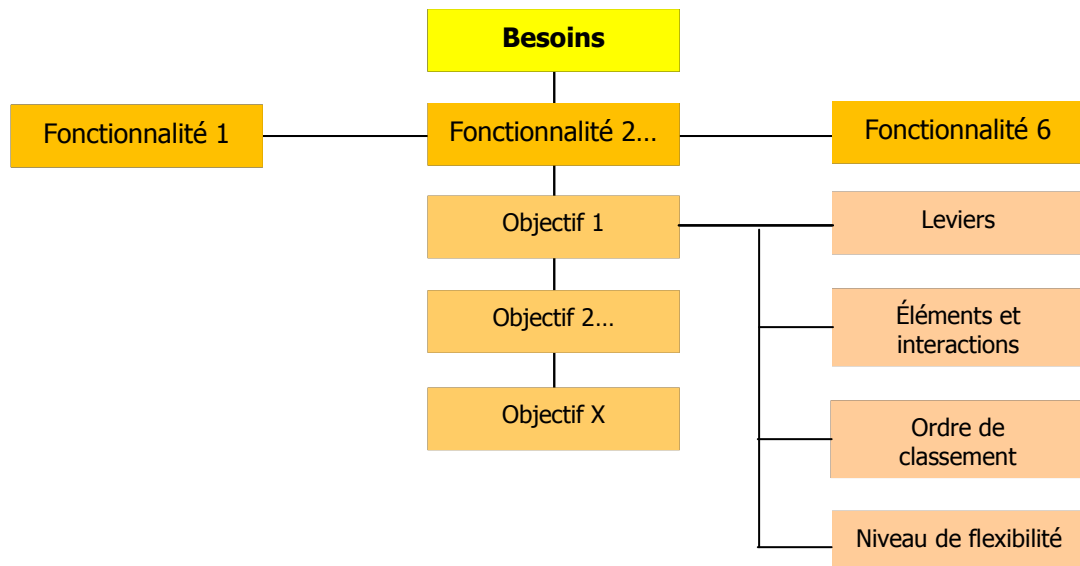


Figure 2 – Organisation du tableau des spécifications applicatives du projet

Les fonctionnalités identifiées et leurs objectifs se présentent ainsi :

- fonctionnalité de « Recherche » pour :
  - trouver des visuels de manière rapide et pertinente ;
  - trouver les informations relatives à un visuel ;
  - rechercher sur Internet et dans le DAM ;
  - centraliser la recherche en un même lieu ;
  - conserver les formules de recherche pour les réutiliser ultérieurement.
- fonctionnalité d'« Alerte » pour :
  - identifier les nouveaux mannequins, nouvelles égéries, nouveaux photographes ; nouvelles tendances avant la concurrence ;
  - connaître les dernières réalisations des mannequins, égéries et photographes ;
  - connaître les dernières créations des concurrents ;
  - connaître l'actualité du secteur ;
- fonctionnalité d'« Acquisition/collecte » pour :
  - détenir l'information adéquate ;
  - centraliser les documents dans un même lieu ;
  - faciliter l'enregistrement des nouveaux documents ;
  - caractériser les documents acquis ;
  - éviter les doublons ou la perte de documents ;
- fonctionnalité de « Consultation » pour :
  - être autonome dans la recherche ;
  - accéder à l'outil (aux outils) depuis différents postes ;
  - obtenir un document sous forme dématérialisée ou papier ;
  - recourir à des modes de visualisation différents ;
- fonctionnalité de « Diffusion » pour :

- centraliser la diffusion ;
- adapter la diffusion à la cible ;
- fonctionnalité de « Sécurité » pour :
  - garantir la confidentialité lors de la recherche et de l'acquisition ;
  - sécuriser les accès à l'outil (aux outils) ;
  - sécuriser l'accès au fonds.

Pour les deux autres parties, relatives aux outils de veille et de recherche, j'ai organisé différemment ma réflexion, le degré de précision étant plus poussé. Ainsi, pour chacune, j'ai préféré déterminer les fonctionnalités dont j'ai ensuite spécifié les caractéristiques et pour lesquelles j'ai défini les réponses attendues. De manière schématique :

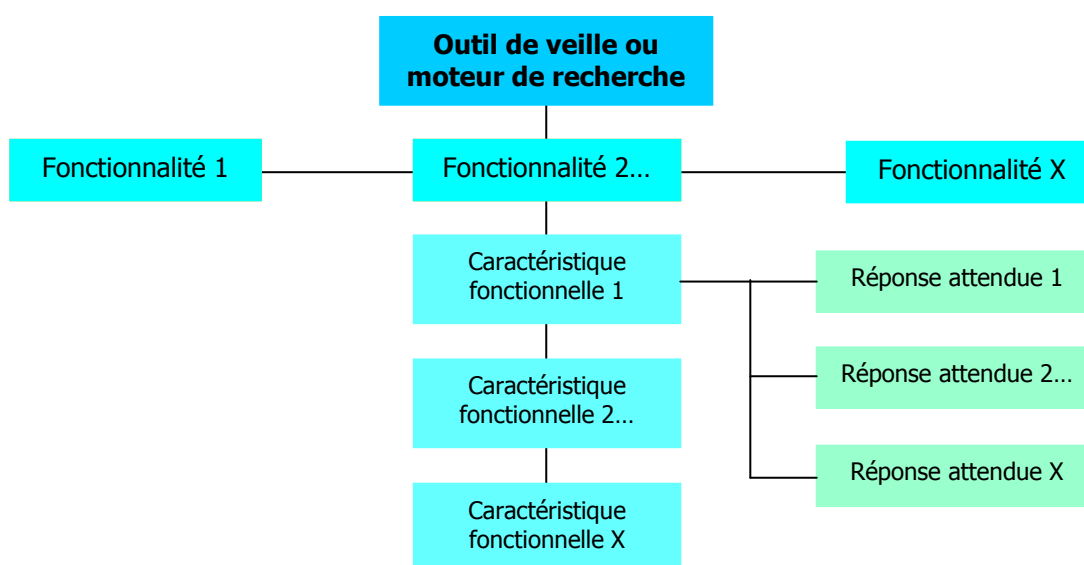


Figure 3 – Organisation du tableau des spécifications applicatives des outils différenciés

## 7.2 Le cahier des charges technique

La réalisation du cahier des charges fonctionnel, dont je ne peux reproduire les éléments dans ce mémoire, m'a permis d'exprimer les besoins en termes de fonctions de services, de résultats et de contraintes. J'en ai déduit deux processus correspondant aux besoins identifiés (cf. Annexe 5). Dans les schémas représentés, le terme « plateforme de veille » est utilisé pour faire référence à un unique outil de prestation de veille, un agrégat d'outils de prestation de veille ou bien à une plateforme de veille, la solution n'étant pas encore retenue à ce niveau de la réflexion.

### 7.2.1 Les spécifications techniques

J'ai choisi d'étudier séparément les spécifications techniques et les performances attendues de la solution de prestation de veille et celle du moteur de recherche. Néanmoins, pour chacune, j'ai organisé ma réflexion en quatre parties : les fonctionnalités demandées, leurs spécifications techniques, les acteurs impliqués et le niveau de flexibilité.

Pour la solution de prestation de veille, j'ai ainsi relevé les fonctionnalités suivantes :

- « Acquisition » : les sources, les requêtes, la méthodologie (automatisation, embarquement des métadonnées, centralisation, etc.) ;

- « Livraison des résultats » : le fonctionnement des alertes (périodicité, lieu d'arrivée), leur niveau de qualification, leur format et les fonctionnalités associées (comme l'historique), etc. ;
- « Consultation » : les différentes possibilités d'accéder aux données acquises, aux modes d'affichage, la pérennité des hyperliens, etc. ;
- « Analyse/traitement » : l'indexation (automatique, manuelle), le tagging, le téléchargement vers le DAM, l'historique, le panier, etc. ;
- « Diffusion » : les modalités de transmission des données acquises (format, accès, mode d'affichage, etc.) ;
- « Archivage dans la solution de veille » : la nature des données et des documents archivables, la structuration, l'accès, etc.
- « Interrogation » : la capacité à rechercher une information sur Internet et dans le DAM, l'emplacement dans l'interface, le format d'affichage des résultats, etc.
- « Validation et réajustement » : le processus itératif avec le prestataire et l'autonomie de l'utilisateur ;
- « Contraintes techniques » : les problèmes éventuels d'interopérabilité, les impératifs de sécurité et de confidentialité, la disponibilité de la solution aux heures de travail (maintenance), etc.

Pour le moteur de recherche, les fonctionnalités sont les suivantes :

- « Accès » : depuis le poste ou Internet ou le DAM ou la solution de veille ;
- « Saisie de la requête » : la recherche simple, avancée (opérateurs booléens, filtres), autocomplétion, sauvegarde des requêtes ;
- « Recherche » : l'environnement (Internet, DAM), dans tous les formats, dans le Web invisible et les réseaux sociaux, la recherche d'images similaires, à partir d'images fournies dans les résultats (rebonds), intégrant les technologies du CBIR, dans les données en français et en anglais, etc.
- « Affichage des résultats » : dans une interface unique quelle que soit l'origine, la détermination de la forme, le dédoublonnage, l'affinage par filtres, la clusterisation<sup>75</sup>, la possibilité de générer un flux RSS du résultat de la requête, le panier, etc.
- « Ergonomie » : l'enchaînement des actions utilisateurs, l'enchaînement des fenêtres ;
- « Contraintes techniques » : l'interopérabilité, les impératifs de sécurité et de confidentialité, la disponibilité de la solution aux heures de travail (maintenance), etc.

## 7.2.2 L'identification des risques

J'ai identifié et réparti les risques liés au développement de ce projet en trois familles<sup>76</sup> :

- « Risques techniques » :
  - le manque de pertinence des résultats en raison notamment des technologies disponibles ;
  - l'interopérabilité ;
  - la confidentialité et la sécurité ;

---

<sup>75</sup> Clusterisation : organisation automatique d'un ensemble de documents en sous-groupes (clusters), fondée sur un calcul de proximité entre documents.

<sup>76</sup> MESSAGER Véronique. Gestion de projet [PDF], 24.10.2011, INTD, CNAM. 17.12.2011. volume 2/2, p.18

- « Risques externes » :
  - le respect de la confidentialité par le prestataire ;
  - la réglementation (cf. p.63-64) ;
- « Risques organisationnels » :
  - l'interfaçage avec les autres outils rendant ces derniers inopérants (par exemple pour les utilisateurs du DAM) ;
  - la disponibilité du service informatique ;
  - la disponibilité des futurs utilisateurs, notamment pour les former.

Pour chaque risque ainsi relevé, j'ai appliqué la matrice de probabilité et d'impact<sup>77</sup> suivante :

	Impacts		
Occurrence	Forts	Moyens	Faibles
Très probable	Critique	Critique	Majeur
Probable	Critique	Majeur	Mineur
Peu probable	Majeur	Majeur	Mineur

J'ai alors proposé pour chaque risque des actions afin de réduire leur potentialité de probabilité et de nuisance.

À l'issue du cahier des charges technique, deux scénarios se sont alors dégagés, que I. Boulifa Tamboura et Z. Ben Ammar Mamlouk identifient sous forme de phases : « [...] une phase de lancement ou de démarrage désignée "amorçage" et une phase de systématisation et de formalisation désignée "pérennisation". » [6, Boulifa et al., p.188]. J'ai repris ces deux expressions pour présenter mes recommandations.

---

<sup>77</sup> MESSAGER Véronique. Gestion de projet [PDF], 24.10.2011, INTD, CNAM. 17.12.2011. volume 2/2, p.20. [consulté le 10.09.2012].

## 8 Le scénario d'amorçage

---

### 8.1 Les solutions pour l'automatisation de la veille

#### 8.1.1 Le potentiel de l'automatisation

D'après C. Deschamps et N. Moinet, « [...] on ne peut parler de veille sur Internet qu'à partir du moment où l'on a mis en place un dispositif permettant d'automatiser la surveillance d'une ou plusieurs des page(s) retenue(s) comme sources d'information. » [1, Deschamps et al., p.65]. En effet, si le besoin de recherche d'un visuel, mannequin, photographe ou d'information peut se manifester à n'importe quel moment, l'acquisition d'images sur les tendances, la concurrence, un artiste ou une égérie, qui permettra de créer un fonds en mesure de répondre à ces demandes, nécessite un système automatisé.

Toutefois, cette automatisation a ses limites : on ne peut pas mécaniser toutes les activités de traitement de l'information. Il est intéressant d'automatiser les phases qui sont les plus chronophages, à savoir la collecte et la diffusion. L'analyse réalisée par un expert, dans le cas l'assistante Achat d'art, qui comprend et connaît les attentes de sa hiérarchie, constituera « une valeur ajoutée au processus de veille » [14, Ajalbert, p.26-27].

Néanmoins, même si la phase d'acquisition est automatisée, sa mise en place sera longue en raison du travail intellectuel à réaliser : la définition des axes de surveillance, la sélection des sources, la saisie des requêtes à automatiser, les tests, etc.

C. Deschamps et N. Moinet précisent que « Cette étape [celle de l'automatisation] se positionne donc logiquement après celle du sourcing. » [1, Deschamps et al., p.65].

#### 8.1.2 Le sourcing

##### 8.1.2.1 Définition

Le sourcing est l'opération qui consiste à identifier les types d'information, les organisations et les sources pertinentes d'après les axes de surveillance définis et les données recherchées [11, Ha, p.21]. Selon Jean-François Miaux, cinq critères permettent d'évaluer la qualité d'une source :

- sa richesse : une grande quantité et variété d'informations utiles ;
- sa performance dans le temps : des informations fraîches délivrées de manière systématique et continue ;
- sa fiabilité : des informations correspondant fidèlement à l'information d'origine ;
- sa discrétion : idée que plus une information est connue, moins elle a de valeur ;
- sa vulnérabilité : la susceptibilité à ne plus émettre d'information [5, Miaux, p.33].

Pour Jean-François Miaux, une veille web nécessite une évaluation préalable de la qualité des informations délivrées car :

- la qualité des informations sur Internet est très variable ;
- n'importe qui est désormais producteur et éditeur d'opinions, d'expériences ;
- les sources sont parfois difficiles à identifier ;
- l'objectif des sites à surveiller est parfois vague voire ambigu ;

- des enjeux économiques peuvent influencer les données délivrées [5, Miaux, p.34].

Une fois alerté, il est possible de s'intéresser aux éléments pouvant composer le sourcing. Ainsi, selon Frédéric Bayle, les principaux acteurs de la diffusion d'images photographiques sur le Web sont :

- les photothèques publiques et privées (exemples : BNF, INA, Agence de la RMN<sup>78</sup>) ;
- les banques d'images commerciales (exemples : Corbis, Getty Images, Fotolia) ;
- les agences de presse (exemples : Agence France Presse, Reuters) ;
- les collectifs de photographes (exemples : L'Œil public, Agence Vu, Tendence floue) ;
- les plateformes collaboratives, qui sont des plateformes de partage et d'échanges (exemples : Flickr, Picasa, Istockphoto) [19, Bayle, p.52].

Des plateformes qui proposent, par abonnement, l'accès à des bouquets d'agences, à des fonds spécialisés, etc. se développent. De plus, dans le secteur du luxe, les magazines féminins - qui ont également leur version web -, sont aussi des sources intéressantes.

### **8.1.2.2 Les annuaires et les répertoires**

D'autres outils de sourcing populaires, notamment auprès des professionnels de l'information au début d'Internet, sont en perte de vitesse. C'est le cas, en premier lieu, des annuaires généralistes dont l'objectif est « d'offrir un recensement de sites web, classés dans des catégories thématiques. Tous les domaines sont couverts et tous les types de sites - grand public et professionnels - sont répertoriés. » [2, Foenix-Riou, p.108]. Les annuaires thématiques se concentrent sur un domaine, un sujet particulier. Généraliste ou thématique, le contenu éditorial des annuaires est géré par une équipe qui veille, répertorie et évalue les sites. Leur usage est recommandé pour défricher un sujet.

Les répertoires sélectifs, élaborés selon le même principe que les annuaires généralistes, proposent un référencement plus poussé dans chaque domaine. Ils ont une dimension qualitative plus forte et sont recommandés pour l'identification de sites de référence. Ainsi, ils ne contiennent pas de sites commerciaux ou grand public.

Pour la réalisation du sourcing, je n'ai pas eu recours à ces outils peu pertinents pour mon projet. Les premières références de sources à surveiller m'ont été transmises par l'assistante Achat d'art.

### **8.1.2.3 Le tableau récapitulatif du sourcing**

En effet, lors de notre premier entretien en février, l'assistante Achat d'art m'a fourni une liste de sites et blogs qu'elle avait identifiés et qu'elle consultait régulièrement. Ils ont ainsi été mes premières sources de référence, complétés par d'autres lors de mes divers entretiens. Je me suis abonnée aux flux RSS (cf. p.55) et/ou aux newsletters lorsqu'ils existaient afin de les évaluer.

Ces sources proposent également d'autres références, notamment dans les blogrolls, c'est-à-dire la colonne dans laquelle le rédacteur d'un blog propose sa sélection de blogs favoris. Dans l'univers du luxe, fermé, ces cooptations sont très informatives. Pour les sources identifiées, j'ai mis en place un même type de suivi pour évaluation.

---

<sup>78</sup> RMN-Grand Palais : établissement public sous la tutelle du ministère de la Culture, issu de la Réunion des musées nationaux et du Grand Palais et dont l'objet est de gérer et valoriser trente-quatre musées nationaux et leurs collections. Son catalogue en ligne comprend plus de 200 000 images photographiques des œuvres d'art conservées dans les musées nationaux et régionaux français. <<http://www.photo.rmn.fr/>>

J'ai enfin élaboré le sourcing à partir des dossiers de photographies réalisés par l'assistante Achat d'art, classés par les noms des mannequins et des photographes. Créés et alimentés depuis 2007, ils comprennent plus de 4 000 photos de mannequins et près de 2 000 créations de photographes. Pourvus d'un titre et de quelques métadonnées embarquées (par indexation automatique, cf. p.31), ces fichiers m'ont permis de retrouver les sites d'agences, des sites personnels ou encore ceux sur lesquels sont diffusées ces photos.

À partir des sources repérées et de leur évaluation, j'ai pu élaborer un tableau de sourcing organisé en colonnes : Catégorie (nature de la source) ; Nom ; URL ; Remarques (informations sur le contenu, les modalités de recherche, les accès, etc.) ; Fonctionnalités (flux RSS, flux thématiques RSS, Twitter, Tumblr, YouTube, Vimeo, newsletter).

Plus de 200 références ont ainsi été répertoriées, chacune pouvant avoir différentes URL : par exemple, pour le site du magazine *Madame Figaro*, il y a l'URL de la page d'accueil, celle de la page Beauté, celle de la page Tendances. L'Annexe 6 reproduit les catégories des références retenues pour réaliser l'acquisition des données.

### 8.1.3 Les flux RSS et les agrégateurs

Puisque nous sommes dans un "scénario d'amorçage", il n'est pas encore question de recourir à un outil trop professionnel. En effet, l'assistante Achat d'art doit d'abord tester ce sourcing, l'ajuster et se familiariser avec le processus de veille.

Pour ce faire, j'ai recherché des outils push<sup>79</sup> relativement simples, lui permettant toutefois de simuler une approche plus professionnelle de la veille. Ainsi, je lui ai proposé de travailler avec des flux RSS et un agrégateur.

#### 8.1.3.1 Définition et pratiques des flux RSS

La quasi-totalité des personnes que j'ai rencontrées lors des entretiens ne connaissaient pas l'expression de « flux (ou fil) RSS ». L'abréviation RSS signifie « Really Simple Syndication », littéralement « format simple pour syndiquer du contenu » c'est-à-dire un format universel que les sites et blogs peuvent utiliser pour publier leurs articles<sup>80</sup>. Le RSS est donc « un protocole de diffusion de contenu, qui permet d'être alerté automatiquement dès que de nouvelles informations sont publiées sur des blogs, sites d'actualités, moteurs de recherche... » [2, Foenix-Riou, p.197]. En effet, des moteurs de recherche proposent de transformer les résultats d'une requête en flux RSS (comme Google Actualités, mais aussi ScienceDirect, etc.).

Béatrice Foenix-Riou précise : « Pour exploiter un fil RSS et le lire, il faut disposer d'un outil spécifique, appelé lecteur ou agrégateur de flux RSS. » [2, Foenix-Riou, p.198]. Elle en présente plusieurs :

- hors connexion : à télécharger préalablement sur son ordinateur (FeedDemon ou Newzle) ;
- en ligne :
  - intégrés aux navigateurs Internet (Mozilla, Internet Explorer, etc.) ;
  - ressemblant à des portails personnalisés (iGoogle, Netvibes, etc.) ;
  - se présentant comme des outils hors connexion (Google Reader, etc.).

---

<sup>79</sup> Push : l'information est "poussée" jusqu'à l'utilisateur, grâce à des systèmes mis en place en amont.

<sup>80</sup> Vincent. Une introduction aux flux RSS. In Veille & Recherche [blog en ligne]. 13.06.2012 [consulté le 04.07.2012]. <<http://veille-et-recherche.fr/une-introduction-aux-flux-rss/>>



En outre, j'ai relevé des agrégateurs développant des fonctionnalités particulières :

- pour fusionner plusieurs flux en un seul (Simple Feed Sorter, RSS Mix, FeedMyinBox)<sup>81</sup> ;
- pour visualiser ses flux :
  - Feedly<sup>82</sup> : il permet d'afficher une sélection de flux RSS, stockée sous Google Reader, dans un format magazine, au sein de son navigateur (extensions Chrome ou Firefox), sa tablette ou son smartphone ;
  - sous forme de PDF : exemples : Feed2PDF et FiveFilters : il faut choisir cinq articles maximum d'un même flux que l'on veut voir apparaître dans son PDF<sup>83</sup>. L'Annexe 7 présente des essais réalisés à partir du blog Armocromia.

### 8.1.3.2 La solution retenue

Pour amorcer l'utilisation des flux RSS, j'en ai sélectionné six, réguliers sans être envahissants et comportant beaucoup de visuels. Je précise que les sites ou blogs dont ils sont issus sont parfois moins intéressants que leurs flux, qui vont directement à l'essentiel, l'image :

- <http://www.armocromia.com/> : maquillage ;
- <http://www.latoutepetiteagence.com/leblog/> : maquillage, parfum, joaillerie, mode, etc. ;
- <http://www.soblacktie.com/> : photos, mannequins, parfum, joaillerie, etc. ;
- <http://thefashionisto.com/> : photos de mode masculine ;
- <http://trendland.com/> : photos, mode, design, etc. ;
- <http://www.webandluxe.com/> : marques, parfum, joaillerie, voitures.

J'ai retenu l'agrégateur Google Reader. Il présente en effet plusieurs atouts :

- s'abonner à des flux RSS de sites et de requêtes lancées sur les moteurs de recherche (seulement si ces derniers proposent cette fonctionnalité) ;
- organiser ses flux selon une arborescence par dossiers, sous-dossiers, maîtrisée par les utilisateurs potentiels ;
- tagger les articles, ce qui permet de les retrouver plus facilement puisque les tags apparaissent sous l'arborescence des dossiers. Ainsi, si l'on a tagué plusieurs articles avec « Helmut Newton », une étiquette apparaîtra sous l'arborescence et en cliquant dessus, tous les articles tagués avec cette expression s'afficheront ;
- lancer une recherche par mots sur l'ensemble des flux, dans tous les articles archivés ou ceux d'un flux particulier ;
- envoyer un article par courriel ;
- étudier des statistiques : on peut ainsi évaluer l'activité d'un flux (nombre de nouveaux articles) et observer quels sont ceux que l'on consulte le plus, ce qui permet d'ajuster le sourcing.

Toutefois, l'un des inconvénients est que toutes les sources ne proposent pas de flux RSS ou alors pas dans les domaines thématiques d'intérêt. Aussi, la veille mise en place

---

<sup>81</sup> Simple Feed Sorter : <<http://wordpress.org/extend/plugins/simple-feed-sorter/>> ;

RSS Mix : <<http://www.rssmix.com/>> ; FeedMyinBox : <<http://www.feedmyinbox.com/>>

<sup>82</sup> Feedly : <<http://www.feedly.com/>>

<sup>83</sup> Feed2PDF : <<http://feed2pdf.appspot.com/>> ; FiveFilters : <<http://fivefilters.org/pdf-newspaper/>>

ne couvre qu'une partie du sourcing. Un tel scénario de prestation de veille est dès lors très limité.

## **8.2 Les solutions pour la recherche d'images**

### **8.2.1 Le moteur de recherche spécialisé**

#### **8.2.1.1 Sa technologie et ses fonctionnalités**

La solution la plus simple de la recherche d'images est la sérendipité c'est-à-dire la découverte aléatoire par une navigation de liens en liens. Ainsi, le responsable du studio maquillage, que j'ai interviewé en juillet, préfère ne pas élaborer de stratégie de recherche par requêtes, mots-clefs, sites et blogs favoris, flux, etc. et rebondir de lien en lien, de site en site. Même si cette démarche requiert beaucoup de temps, il la préfère car, selon lui, elle lui permet de trouver des images qu'il ne voit pas ailleurs, donc rares et ainsi à forte valeur ajoutée.

Cependant, la demande initiale souhaitait la mise en place de processus automatisés afin de faire gagner du temps à l'assistante Achat d'art. Aussi, je me suis intéressée aux moteurs de recherche qui, à côté de la prestation de veille, pourraient l'aider à répondre aux demandes ponctuelles émises par sa hiérarchie.

En matière de recherche d'images, Frédéric Bayle distingue deux types de moteurs de recherche, les généralistes et les spécialisés. Certains moteurs de recherche généralistes comme Google ont développé une fonction avancée de l'image (cf. p.39) [19, Bayle, p.42-43]. Leur volume de résultats est souvent plus important, ces acteurs ayant indexé des milliards d'images sur le Web. Toutefois, le fait que les requêtes soient lancées sur les éléments textuels périphériques et les métadonnées du fichier génère du bruit et un manque de pertinence.

Les moteurs spécialisés proposent moins de fonctionnalités et de filtres mais utilisent des technologies d'indexation et de recherche d'images [19, Bayle, p.42-43]. Ainsi, les algorithmes de ces moteurs de recherche d'images :

- parcourent des bases de données ;
- analysent les images et définissent leur signature numérique (fondée sur les couleurs, les formes et les textures) ;
- comparent les images entre elles et sélectionnent les plus ressemblantes ;
- recherchent par similarité [44, GFII, p.56].

#### **8.2.1.2 Une comparaison de moteurs de recherche**

J'ai réalisé un tableau comparatif de trente-deux moteurs de recherche d'images<sup>84</sup>, dont les critères relèvent du cahier des charges technique. Google Search Image, étudié au préalable, a donc servi de mesure « étalon ». En effet, l'assistante Achat d'art, ainsi que de nombreuses autres personnes rencontrées, utilisent principalement voire exclusivement ce moteur. J'ai effectué les recherches à partir des requêtes suivantes :

- deux noms de mannequins ;
- trois noms de photographes ;

---

<sup>84</sup> Beaucoup de ces moteurs sont en mode Bêta donc en phase de tests et d'évolution continue, ainsi certains ont depuis disparu ou ont été modifiés.

- deux concepts visuels : lèvres bleues (facile), femme tête renversée (complexe).

Le tableau est organisé en deux temps :

- les fonctionnalités de recherche avec ces six critères d'évaluation :
  - l'accès au moteur : en ligne, logiciel installé ou en SaaS ;
  - les modalités de la recherche simple ;
  - les modalités de la recherche avancée : opérateurs booléens, filtres, clusters, etc. ;
  - les aides à la recherche : autocomplétion, vocabulaire contrôlé, propositions ;
  - un historique et la sauvegarde des requêtes ;
  - la méthode de recherche du moteur ;
- les fonctionnalités de consultation des résultats avec sept critères d'évaluation :
  - le format d'affichage ;
  - le dédoublement des réponses ;
  - la clusterisation automatique des images ;
  - les filtres ;
  - les fonctionnalités caractéristiques du moteur (exemple : générer des flux RSS depuis les résultats) ;
  - l'ergonomie en termes de visualisation et de manipulation ;
  - la pertinence des résultats.

Ne pouvant m'attarder sur chaque moteur, je vais m'intéresser à certains d'entre eux, connus ou méritant de l'être. Cependant, je précise au préalable qu'à l'issue des différents tests, Google Search Image est celui qui a présenté les résultats les plus pertinents.

Parmi les moteurs de recherche en ligne, je présenterais Ixquick<sup>85</sup>. Ce moteur, en français, m'a intéressée en raison des contraintes du demandeur. En effet, il revendique une politique de confidentialité, puisqu'il n'enregistre pas l'adresse IP<sup>86</sup>. On peut effectuer des recherches par langage naturel ou avec des opérateurs booléens. L'ergonomie est confortable et les résultats sont pertinents et ont proposé des images que je n'avais pas trouvées avec d'autres moteurs.

Je me suis également intéressée aux métamoteurs. Ils présentent l'avantage « de lancer une même requête sur plusieurs moteurs » [2, Foenix-Riou, p.100]. Cependant, il faut être prudent car d'une part, certains moteurs refusent d'être interrogés par des métamoteurs et d'autre part, la même requête étant envoyée à plusieurs, elle sera comprise différemment selon chacun et les résultats seront donc réduits « au plus petit dénominateur commun » [2, Foenix-Riou, p.102].

Parmi les métamoteurs, j'ai retenu ZapMeta<sup>87</sup>. La liste des moteurs de recherche qu'il consulte est présentée sur la page d'accueil. Il est interrogeable par langage naturel et opérateurs booléens. Il propose des fonctionnalités d'autocomplétion, de filtre, de surbrillance des mots, de suggestions d'autres termes, et conserve un historique de recherche. Son originalité réside dans la page des résultats sous la forme d'un petit bloc au sein duquel les résultats ont été classés automatiquement sous la forme d'une arborescence, selon trois critères : « Dimensions » (taille des images), « Sujet »

<sup>85</sup> Ixquick : <<https://www.ixquick.com/fra/>>

<sup>86</sup> Adresse IP : IP pour Internet Protocol : numéro d'identification pour chaque ordinateur connecté à Internet permettant de localiser n'importe quel ordinateur sur le réseau Internet.

<sup>87</sup> ZapMeta : <<http://www.zapmeta.fr/?stname=image>>

et « Domaine » (domaine de l'URL du site : exemples : .com, .eu, etc.). C'est ce qui est couramment appelé la « clusterisation des résultats ». Sur la capture d'écran présentée dans l'Annexe 8, les chiffres entre parenthèses témoignent qu'une petite partie seulement des résultats est traitée mais le développement de l'indexation automatique devrait accroître la performance de ce genre d'outil.

Enfin, je présenterais Exalead<sup>88</sup>, acteur bien connu du marché. L'internaute peut utiliser la recherche simple en langage naturel néanmoins, contrairement à beaucoup de recherches avancées complexes pour les non initiés, celle proposée par Exalead est aisée car elle présente, par des exemples concrets, l'utilisation des opérateurs booléens. De même, sur la page de résultats, le moteur propose des filtres de manière ergonomique. Cependant, la pertinence des résultats n'est pas très bonne. Exalead s'est investi également sur des chantiers technologiques approfondis en matière de recherche d'images.

## 8.2.2 La recherche par la couleur et la similarité

### 8.2.2.1 La recherche par la couleur

Exalead a développé « Chromatik »<sup>89</sup>, un moteur de recherche fondé sur la couleur<sup>90</sup>. Dans la palette de couleurs proposées, il faut en choisir une ou plusieurs, les résultats de la recherche étant les photos composées par ces couleurs. On peut également déterminer la luminosité, la saturation et surtout y adjoindre une recherche textuelle. Dans l'Annexe 8, j'ai fait un test avec divers bleus. Les résultats contiennent en effet les couleurs. Toutefois, quand j'ajoute le terme « lips » (lèvres), le moteur ne trouve aucun résultat. Si la technologie est très intéressante, le fonds Flickr utilisé par Chromatik n'est pas suffisamment riche pour faire des recherches dans le domaine du projet.

Cette technologie utilise le système appelé QBIC (« Query By Image Content »). Le site du Musée de l'Hermitage de Saint-Petersbourg l'utilise depuis 2003 pour son catalogue en ligne<sup>91</sup>. Le système QBIC est proposé en tant que recherche rapide pour trouver une œuvre. Deux modalités sont offertes, illustrées en Annexe 9 :

- par la couleur : sélectionner les couleurs que l'on souhaite voir apparaître dans l'œuvre ;
- par la mise en page : choisir des formes (cercle, carré, etc.), leur couleur et leur emplacement sur la toile virtuelle.

Les résultats sont très impressionnants, on peut remarquer que le moteur a su interpréter ma demande. Néanmoins, le système s'appuie sur un fonds solide et réduit, celui du musée (comparé aux millions de pages d'Internet). Le moteur a été mis au point par IBM<sup>92</sup> qui précise qu'il permet de rechercher dans le contenu visuel des images appartenant à de larges bases de données avec comme critères le pourcentage de couleur, la répartition des couleurs et les textures, sans utiliser de mots pour décrire l'image recherchée.

---

<sup>88</sup> Exalead : <<http://www.exalead.com/search/image/>>

<sup>89</sup> Chromatik : <<http://chromatik.labs.exalead.com/>>

<sup>90</sup> EXALABS. What is Chromatik ? In Exalead [site en ligne]. [s.d.] [consulté le 12.09.2012]. <<http://labs.exalead.com/project/chromatik>>

<sup>91</sup> Digital Collection. In The State Hermitage Museum [site en ligne]. [s.d.] [consulté le 12.09.2012]. <<http://www.hermitagemuseum.org/fcgi-bin/db2www/browse.mac/category?sellLang=English>>

<sup>92</sup> FINN Robert. Querying by image content. In IBM Research [site Internet en ligne]. [s.d.] [consulté le 29.09.2012]. <[http://domino.research.ibm.com/comm/wwwr\\_thinkresearch.nsf/pages/image396.html](http://domino.research.ibm.com/comm/wwwr_thinkresearch.nsf/pages/image396.html)>

Un autre moteur de recherche d'images utilise cette technologie, TinEye<sup>93</sup>, néanmoins ses résultats sont peu pertinents par rapport à la qualité artistique des photos recherchées car son fonds est celui des images de Flickr (cf. Annexe 8 pour voir un exemple de recherche).

### 8.2.2.2 La recherche par la similarité

TinEye propose un autre service, celui de la recherche par similarité. Il propose ainsi soit de télécharger une image, soit d'entrer une URL d'image. J'avais présenté ce type de recherche avec Google Search Image (cf. p.39).

Ce mode de requête est proposé par la plupart des moteurs de recherche d'images par le contenu (appelé aussi « Query By Exemple » ou QBE ou recherche par l'exemple) [45, p.99]. Selon Pierre Nobis, trois types de requête peuvent être identifiés [43, Nobis, p.3] :

- la requête globale : sur toute l'image ;
- la requête partielle : sur une zone précise (ou région d'intérêt) ou sur un ou plusieurs paramètres ;
- la requête esquissée : fondée sur la reconnaissance par la forme.

Du point de vue utilisateur, ce mode de requête est plus simple à utiliser que d'avoir à définir les caractéristiques de bas niveau des couleurs et des textures [45, p.38].

J'ai testé plusieurs moteurs de recherche par similarité et une fois encore Google Search Image a présenté les résultats les plus pertinents, avec l'ergonomie la plus confortable. Les nombreux filtres proposés ainsi que son fonds expliquent, à mon avis, cette conclusion concernant une recherche d'images particulières relatives à l'univers de mon projet. En effet, de nombreux moteurs fondent leurs recherches sur Flickr, qui dispose d'une immense base, or, ces images trop banales ne correspondent pas au genre d'images recherchées par l'assistante Achat d'art.

## 8.3 Les limites de ce scénario

Pour conclure, ce "scénario d'amorçage", composé d'une part, d'un agrégateur recevant des flux RSS d'alertes, de newsletters et d'autre part, d'un moteur de recherche d'images, peut aider à la mise en place des pratiques et des réflexes et permet de tester des sites et les flux RSS.

Néanmoins, ces outils seront vite dépassés car :

- la couverture est trop faible : tous les sites ne disposent pas de flux ;
- les informations ne sont pas analysées, traitées ;
- l'utilisatrice sera vite débordée par le nombre d'articles collectés dans l'agrégateur ;
- il faut assurer une maintenance des flux (problème de la pérennité des hyperliens).

En outre, même si l'on peut essayer de préserver la confidentialité par des solutions telles que les serveurs proxy (qui fait le « tampon » entre le poste utilisateur et le Web), le chiffrement ou encore le recours au réseau Tor<sup>94</sup>, les exigences de sécurité ne sont pas remplies.

---

<sup>93</sup> TinEye : <<http://labs.tineye.com/multicolor>>

<sup>94</sup> Réseau Tor : système construit sur une multitude de nœuds : la demande transite par un chemin aléatoire, ne permettant pas de connaître la provenance de la connexion.

Enfin, la combinaison d'outils présente plusieurs inconvénients voire risques. En effet, il faut toujours un travail humain pour activer chaque outil donc le gain de temps promis par l'automatisation se réduit. De plus, les erreurs risquent de se produire à chaque étape de manipulation ; de même, il peut y avoir autant de pannes techniques que d'outils utilisés [14, Ajalbert, p.26].

En conséquence, la mise en œuvre d'une veille de portée stratégique nécessite la professionnalisation de la fonction. En effet, la veille est le métier d'un professionnel de l'information et non une polyvalence supplémentaire.

## 9 Le scénario de pérennisation

### 9.1 Les logiciels et plateformes de veille

#### 9.1.1 L'environnement technologique et juridique

##### 9.1.1.1 Différents produits existants

Le marché de la veille est foisonnant, proposant des solutions de différentes natures. La figure 4 schématise sous la forme de trois ellipses les offres du marché de la veille selon trois fonctions : la collecte, l'analyse et la diffusion. Il me semble que les fonctions de collecte et d'analyse sont incontournables pour répondre au besoin, la diffusion étant pour le moment à la discrétion de la hiérarchie directe de l'assistante Achat d'art. Ainsi, selon ce schéma, il faudrait ainsi étudier ce que G. Balmisse et D. Meingan appellent les « outils de classification automatique ».

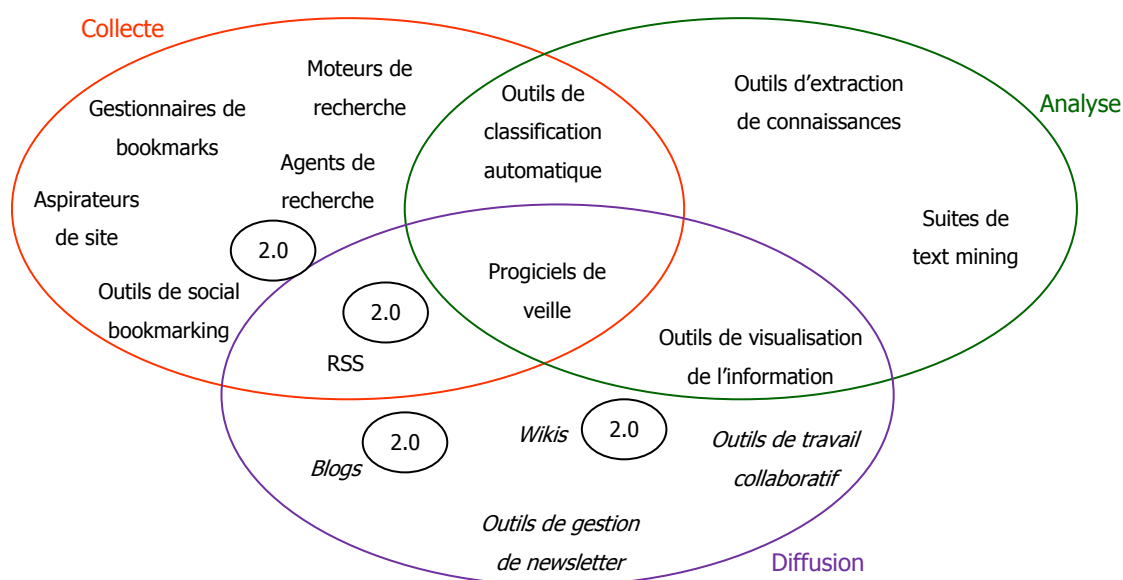


Figure 4 – Panorama du marché de la veille, d'après [9, Balmisse et al., p.110]

En italique, les outils qui ne sont pas strictement réservés à la veille.

Toutefois, il est envisageable qu'une partie des données extraites par l'outil retenu puissent intéresser les créatifs, en dehors du cycle de décision stratégique. Aussi, à moyen ou long terme, une diffusion aux équipes créatives de l'entreprise serait à prévoir. En conséquence, il me paraît judicieux d'orienter mes recherches sur ce que G. Balmisse et D. Meingan appellent des « progiciels de veille », c'est-à-dire « des outils intégrés qui rassemblent des fonctionnalités couvrant le champ complet des besoins en matière de veille [...] ». Plus précisément, ils offrent, à des degrés divers, des fonctionnalités couvrant l'ensemble du périmètre du trèfle fonctionnel (collecte, analyse et diffusion). » [9, Balmisse et al., p.118].

D'une manière générale, la littérature dédiée parle de « plateforme de veille ». Pour C. Deschamps et N. Moinet, celle-ci permet d'« Automatiser, rationaliser et professionnaliser la surveillance de son environnement » mais aussi d'« Optimiser la couverture de sources électroniques multiples (flux RSS, sites web, blogs, forums, bases de données) à partir d'un point d'entrée unique. » [1, Deschamps et al., p.75]. Le sourcing, utilisé et testé dans le « scénario d'amorçage », pourra alors être repris dans son intégralité (ou dans une version

corrigée) et des requêtes lancées avec des moteurs de recherche, comme « femme tête renversée », pourront également être intégrées.

Cependant, le travail humain ne peut pas être évacué : il faut définir les axes de veille, paramétrer la plateforme, suivre les résultats de l'outil pour le « rendre intelligent » et pertinent par rapport aux besoins. Il faut ainsi compter plusieurs mois d'apprentissage pour obtenir une plateforme efficace.

#### **9.1.1.2 L'activité de veille est-elle menacée ?<sup>95</sup>**

L'actualité judiciaire réclame la prudence. En effet, au Royaume-Uni, la cour d'appel du Pays de Galles a condamné le 27 juillet 2011 la société Meltwater, prestataire de veille, à « payer des droits pour les références des articles accompagnés des liens permettant à ses clients de consulter les articles en consultation libre sur Internet. »<sup>96</sup>, ce qui induit que désormais « au Royaume-Uni, les clients d'un prestataire de veille [...] vont payer pour les liens qui leur sont proposés. », soit « la même somme que ceux qui recevront également des extraits des articles répondant à leurs profils. »<sup>97</sup>.

Au Canada, les universités du Western Ontario et de Toronto ont accepté « de payer pour les liens proposés dans les courriers électroniques de leurs enseignants »<sup>98</sup>.

En France, « Par principe, le lien hypertexte est libre. »<sup>99</sup>. Le jugement pris au Royaume-Uni ne nous concerne donc pas (encore ?). Néanmoins, il faut avoir à l'esprit le jugement du tribunal de grande instance de Nanterre du 25 juin 2009, selon lequel : « La diffusion d'un assemblage de flux RSS sur un site est considérée comme une agrégation d'informations relevant du régime applicable aux hébergeurs. », suggérant ainsi une responsabilité *a posteriori*. Michèle Battisti explique que « si l'information devait être traitée c'est-à-dire reformulée, précisée, vous devenez éditeur, responsable *a priori* des propos diffusés. »<sup>100</sup>.

Enfin, cet été, l'Allemagne a pris une décision, surnommée la « Lex Google », qui pourrait inspirer d'autres pays. En effet, le gouvernement allemand a adopté le 29 août « un projet de loi qui vise à mieux protéger la propriété intellectuelle des éditeurs de presse sur le Web, obligeant les moteurs de recherche à leur reverser des commissions. »<sup>101</sup>. Michèle Battisti précise que « Cette taxe toucherait tous les portails de liens et non uniquement

---

<sup>95</sup> Pour aborder les risques juridiques courants de la veille, on peut lire l'article suivant : BATTISTI Michèle. La veille sous une pluie de droits. In Paralipomènes [blog en ligne]. 06.10.2012 [consulté le 06.10.2012]. <<http://paralipomenes.net/wordpress/archives/8786>>. ISSN 2260-3336

<sup>96</sup> BATTISTI Michèle. Web 2.0 : nouveaux usages, nouveaux droits ? In Paralipomènes [blog en ligne]. 01.06.2012 [consulté le 28.08.2012]. <<http://paralipomenes.net/wordpress/archives/8172>>. ISSN 2260-3336

<sup>97</sup> Ibidem.

<sup>98</sup> Ibidem.

<sup>99</sup> BATTISTI Michèle. Droit d'auteur, responsabilité et documentation en pratique. In Paralipomènes [en ligne]. 23.05.2012 [consulté le 28.08.2012]. <<http://paralipomenes.net/wordpress/archives/8145>>. ISSN 2260-3336

<sup>100</sup> Ibidem.

<sup>101</sup> LE MONDE.FR, AFP. Polémique sur la « Lex Google » en Allemagne. In Le Monde.fr [site en ligne]. 30.08.2012 [consulté le 31.08.2012]. <[http://www.lemonde.fr/technologies/article/2012/08/30/polemique-sur-la-lex-google-en-allemande\\_1753099\\_651865.html](http://www.lemonde.fr/technologies/article/2012/08/30/polemique-sur-la-lex-google-en-allemande_1753099_651865.html)>



un mastodonte comme Google »<sup>102</sup>. Aurélie Filipetti, ministre de la Culture, n'est pas défavorable au projet de loi allemand<sup>103</sup>. En effet, dans un entretien au journal « Le Monde », elle déclare : « Il serait normal que ces moteurs de recherche contribuent à financer la vraie valeur ajoutée produite par les journaux. Le travail éditorial est fait par la presse, pas par ces plateformes qui pourtant en tirent un profit commercial évident. »<sup>104</sup>. Les éditeurs de presse français se réjouissent de la décision allemande : en 2009, le GESTE<sup>105</sup> voulait « obtenir de Google qu'il partage avec les journaux indexés une partie des revenus publicitaires générés par le service. »<sup>106</sup>. L'association de la presse d'information politique et générale (IPG) proposent ainsi la création d'un système de droits voisins sur l'indexation des contenus<sup>107</sup> qui constituerait un « droit d'utilisation de leurs contenus par les portails Internet ou les moteurs de recherche : ces derniers devraient les rémunérer pour exploiter ces contenus »<sup>108</sup>. Ainsi, l'association IPG demande de modifier l'article L.355-4 du code de la Propriété industrielle relatif à la contrefaçon. Outre une sanction pénale pour la reproduction des contenus sur Internet, la presse veut également punir les liens hypertextes sans rémunération vers les articles. L'alinéa 3 de l'article deviendrait (les ajouts sont en italique) : « Est puni de la peine d'amende prévue au premier alinéa le défaut de versement de la rémunération due à l'auteur, [...] *ainsi qu'à l'organisme de presse* au titre de la copie privée [...] *mais aussi de l'utilisation de liens hypertextes ou de toute technique équivalente permettant d'accéder à des contenus de presse.* »<sup>109</sup>.

Il est encore impossible d'anticiper la décision qui sera prise en Allemagne et peut-être en France (ou bien au niveau de l'Union européenne) à ce sujet. Pourtant, cette évolution législative pourrait impacter le choix d'une solution de veille et d'un prestataire. Dans les pages suivantes, je compare plusieurs outils de veille afin de cerner celui ou ceux qui pourraient répondre au mieux à la demande initiale toutefois je n'ai pas pu intégrer cette dimension. Aussi, une attention particulière sur ce point est conseillée si un prestataire extérieur est retenu : il faudra l'interroger sur les solutions qu'il compte mettre en œuvre si une telle décision sur les hyperliens de veille devenait applicable en France.

---

<sup>102</sup> BATTISTI Michèle. Taxer ou ne pas taxer ? Bataille autour du paiement du référencement de la presse. In Paralipomènes [blog en ligne]. 26.08.2012 [consulté le 03.09.2012].

<<http://paralipomenes.net/wordpress/archives/8514>>. ISSN 2260-3336

<sup>103</sup> Pour connaître l'historique de cette « taxe Google », on peut lire POUCHARD Alexandre. Votée, retardée, supprimée... Petite histoire de la « taxe Google ». In Le Monde.fr [site en ligne]. 18.09.2012 [consulté le 18.09.2012].

<[http://www.lemonde.fr/politique/article/2012/09/18/votee-retardee-supprimee-petite-histoire-de-la-taxe-google\\_1761318\\_823448.html](http://www.lemonde.fr/politique/article/2012/09/18/votee-retardee-supprimee-petite-histoire-de-la-taxe-google_1761318_823448.html)>

<sup>104</sup> TERNISIEN Xavier. « Sur le Net, rien n'est éditorialisé » : Aurélie Filipetti s'explique. In Le Monde.fr [site en ligne]. 04.09.2012 [consulté le 08.09.2012].

<[http://www.lemonde.fr/economie/article/2012/09/04/la-ministre-de-la-culture-s-explique-sur-la-presse-et-internet\\_1755524\\_3234.html](http://www.lemonde.fr/economie/article/2012/09/04/la-ministre-de-la-culture-s-explique-sur-la-presse-et-internet_1755524_3234.html)>

<sup>105</sup> GESTE : Groupement des éditeurs en ligne.

<sup>106</sup> CHAMPEAU Guillaume. En France aussi, la presse veut faire payer les liens hypertextes. In Numerama [site en ligne]. 04.09.2012 [consulté le 08.09.2012].

<<http://www.numerama.com/magazine/23607-filipetti-ouverte-a-une-taxe-google-financant-la-presse.html>>

<sup>107</sup> TESQUEST Olivier. Taxe Google : « Télérama » dévoile le projet des éditeurs de presse. In Télérama.fr [site en ligne]. 21.09.2012 [consulté le 04.10.2012].

<<http://www.telerama.fr/medias/taxe-google-telerama-devoile-le-projet-des-editeurs-de-presse.87027.php>>

<sup>108</sup> FEITZ Anne. Les éditeurs veulent leur « Lex Google ». In Les Échos.fr [site en ligne]. 05.09.2012 [consulté le 08.09.2012]. <<http://www.lesechos.fr/entreprises-secteurs/tech-medias/actu/0202246753016-les-editeurs-veulent-leur-lex-google-358789.php>>

<sup>109</sup> Olivier Tesquest (publié par). Projet de proposition de loi de l'association IPG. Droits voisins pour les organismes de presse. In Scribd [site en ligne]. 21.09.2012 [consulté le 06.10.2012]. 8 p. <<http://www.scribd.com/doc/106565372/Projet-de-proposition-de-loi-sur-les-droits-voisins-pour-les-organismes-de-presse>>

## 9.1.2 Les recommandations

### 9.1.2.1 Les critères de comparaison

Je n'ai pas trouvé de logiciels de veille spécialisés dans l'image. Ainsi, la capacité de reconnaissance par le contenu visuel n'a pas pu faire l'objet d'un critère. J'ai réalisé un tableau comparatif de seize solutions, fondé sur le cahier des charges technique et organisé selon les différentes étapes du processus de veille, avec des critères d'évaluation :

- la collecte :
  - la méthode : sourcing, requêtes, logiciel spécifique, etc. ;
  - la nature des sources : texte, multimédias, sites Internet, sources internes à l'entreprise, flux RSS, blogs, base de données en ligne, Web invisible, etc. ;
  - la valeur ajoutée : dédoublonnage, recherche sémantique, extraction des entités nommées, multilinguisme, etc. ;
  - l'automatisation : critères paramétrables ;
  - l'historique des requêtes et des collectes ;
  - les clients, notamment du secteur du luxe ;
- la livraison des résultats :
  - par alertes mail, RSS ;
  - sous forme de livrables ;
  - les fonctionnalités prévues pour ces alertes (exemple : sont-elles transférables ?) ;
- la consultation de ces résultats :
  - l'accès pour différents profils d'utilisateurs ;
  - l'organisation des résultats : consultation par axes de veille, par thèmes, etc. ;
  - le mode d'affichage ;
  - la pérennité des hyperliens ;
  - un panier de résultats dans lequel l'utilisateur place les articles qu'il souhaite consulter plus tard ;
- l'analyse (traitement des données acquises) :
  - l'indexation automatique des métadonnées embarquées par l'outil ;
  - le tagging des données collectées ;
  - l'indexation des données par l'utilisateur dans l'outil de veille ;
  - l'indexation des données par l'utilisateur depuis le DAM ;
  - le téléchargement des images dans le DAM avec les métadonnées natives ajoutées et le tagging ;
  - le téléchargement des images dans le DAM par glisser-déposer ;
  - l'historique des actions : qui a fait quoi et quand ;
- la diffusion :
  - l'impression ;
  - le transfert et l'export des mails, données, newsletters traités ;
  - la réalisation de livrables personnalisés ;
- l'archivage avec la création d'une base de connaissances :

- la nature des éléments archivés ;
- la structuration des archives : par axes de veille, thèmes, formats, etc. ;
- les métadonnées : sont-elles toutes archivées et sous quelle forme ? ;
- le mode de navigation dans les archives : par arborescence, moteur de recherche.

J'ai ajouté trois autres sections :

- l'interrogation car la solution devra permettre d'interroger Internet et le DAM, avec des capacités de recherche sémantique :
  - moteur de recherche intégré à l'outil ;
  - mode d'affichage des réponses ;
  - interrogation des archives de l'outil ;
  - interrogation d'Internet ;
  - interrogation du DAM ;
  - historique et conservation des requêtes ;
- les contraintes techniques, en lien avec les impératifs informatiques :
  - la confidentialité ;
  - la compatibilité Mac et PC ;
  - l'interopérabilité avec l'existant et les installations à venir (évolutivité) ;
  - le respect de la procédure de sécurité informatique ;
  - le mode d'installation et hébergement : licence, SaaS, sur le serveur du client ou du prestataire ;
- la validation et le réajustement pour évaluer le mode de travail avec le prestataire : les interactions avec le prestataire pour perfectionner l'outil et l'autonomie de l'utilisateur.

### **9.1.2.2 Les résultats**

Aucun outil ne répond à l'ensemble des critères exposés, auxquels il faut ajouter l'ergonomie et la facilité de prise en main puisque l'utilisatrice principale n'est pas une professionnelle de l'information. En conséquence, les solutions proposées par les éditeurs AMI Software, Arisem, Digimind, KB Crawl me paraissent trop complexes. En raison des impératifs de confidentialité et d'anonymat, l'entreprise ne souhaite pas de solution en SaaS ni d'un hébergement sur des serveurs extérieurs aux siens (cf. p.21). À l'issue du comparatif, j'ai ainsi relevé trois outils en capacité de répondre aux besoins : ceux des sociétés Antidot, Iscope et Qwam.

Les deux premiers prestataires ont développé des outils très avancés en matière de recherche sémantique, ce qui peut être pertinent en raison du problème soulevé par la recherche d'images par le texte. De plus, leurs solutions se construisent par brique, obtenant de ce fait un outil adapté aux besoins. Le logiciel KeyWatch d'Iscope semble aisé à prendre en main, ce qu'a confirmé une utilisatrice du ministère des Affaires étrangères et européennes. Antidot propose trois logiciels de base et est particulièrement avancé sur les questions sémantiques (notamment au niveau de son moteur de recherche). Comme Qwam, il offre des prestations avancées en matière de veille collaborative et de Knowledge Management (KM), deux développements ultérieurs à ne pas négliger.

## 9.2 La veille collaborative

### 9.2.1 Dans quel but ?

#### 9.2.1.1 Une définition du travail collaboratif

Selon A. Diallo et A. Jdey, le collaboratif concerne toutes les fonctions d'une entreprise et s'impose progressivement avec la diffusion, aussi bien dans la vie privée que professionnelle, des pratiques issues du Web 2.0, sous forme d'un principe de partage [10, Diallo et al., p.60-61]. Il est difficile voire impossible de concilier au sein d'un seul service toutes les expertises : aussi le collaboratif permet l'interaction d'acteurs aux compétences complémentaires qui sont censés enrichir le travail produit par le partage de leurs savoirs et savoir-faire. Ainsi, pour A. Diallo et A. Jdey, cela se traduit d'un point de vue technique, au niveau de l'outil de veille, par :

- l'ajout de commentaires et la possibilité d'annoter le contenu ;
- au moins trois niveaux de droits d'utilisateurs sur les applications en réseau : administrateur, contributeur et lecteur ;
- un système d'alerte ou de diffusion sélective en fonction du profil de l'utilisateur ;
- la personnalisation des contenus et la possibilité d'éditer à plusieurs [10, Diallo et al., p.61].

Cependant, Stéphanie Toupin Lemaire souligne que ce mode collaboratif doit être contrôlé [8, Toupin Lemaire, p.11] : en effet, le schéma traditionnel de l'émetteur au(x) récepteur(s) est complexifié en « un émetteur-récepteur et un, ou des récepteurs-émetteurs, diffusant à un ou plusieurs récepteurs. » [8, Toupin Lemaire, p.18]. De plus, les contributeurs doivent être motivés pour participer au partage de leur expertise (cf. p.19-20 le retour d'expérience de la responsable de la veille du service R&D). Le partage de l'information est une démarche difficile car sa possession est souvent interprétée comme conférant un pouvoir à celui qui la détient. Il faut en outre veiller les données transmises - vérifier si elles ne sont pas redondantes ou erronées -, et ne pas engendrer de la surinformation.

Laura Phirmis cite les exemples de grands groupes engagés dans cette démarche (Bouygues, Schlumberger, Danone) en mentionnant les risques liés, à savoir la remise en question de « certains principes du management tel que le contrôle vertical de l'information » [7, Phirmis, p.90].

#### 9.2.1.2 Le Knowledge Management (KM)

Le partage de l'information collectée et capitalisée ouvre la voie au Knowledge Management ou gestion de la connaissance. Laura Phirmis définit la notion de connaissance ainsi : « Une information est une donnée brute, sortie d'un contexte, qui remise dans un contexte et comprise par les individus devient la connaissance. » [7, Phirmis, p.15].

L'idée est que les salariés d'une même entreprise ont acquis, plus ou moins consciemment, au cours de leurs activités et expériences, beaucoup d'informations tacites donc inaccessibles par informatique : aucun outil de veille ni moteur de recherche ne peut retrouver ces informations. Elles peuvent donc être perdues ou inutilisées alors qu'associées à d'autres informations et contextualisées, elles représentent une grande valeur ajoutée.

Laura Phirmis repère deux générations du KM : la première, du début des années 1990, se limite à un « outil de "capitalisation de contenus et de gestion de la mémoire

d'entreprise<sup>110</sup> » [7, Phirmis, p.18]. Elle induit la création de bases de connaissances. La seconde génération « est caractérisée par l'aspect des relations sociales et informelles de la connaissance », qui se traduit d'un point de vue technologique par « des applications, personnalisables afin de mettre en relation les salariés pour échanger, partager et approfondir leur expertise, en d'autres termes : créer des réseaux de connaissances. » [7, Phirmis, p.18]. Les outils permettent alors « d'identifier l'auteur d'un document ou de localiser des profils d'expertises, de partager les résultats de ses recherches » [7, Phirmis, p.33].

Ainsi, Stéphanie Toupin Lemaire constate dans son mémoire que « les outils du KM [sont] de plus en plus utilisés en tant qu'outil de diffusion des résultats » [8, Toupin Lemaire, p.19]. L'objet traité par la veille dans le cadre de mon stage, l'image publicitaire, est tellement subjectif, faisant appel à une culture personnelle, notamment en matière d'art, et concerne des services très différents par nature (le Marketing et la Création) qu'une démarche mêlant le collaboratif et le KM peut parfaitement faire sens. Toutefois, les problèmes et les risques relevés nécessitent l'intervention d'une personne dédiée pour l'organiser.

## 9.2.2 La place du documentaliste

### 9.2.2.1 Le réseau social d'entreprise : un réseau de veille utopique ?

Laura Phirmis parlait de réseaux de connaissances. Selon le rapport Martre, un réseau est un « Ensemble de personnes partageant un intérêt pour une technique ou pour un objectif commun et échangeant sur ce sujet des informations et, éventuellement, des renseignements, indépendamment ou non des liens professionnels ou hiérarchiques pouvant les réunir. » [16, Martre et al., p.102]. Dans la dynamique du Web 2.0, du collaboratif, se développe depuis quelques années le réseau social d'entreprise, sorte de réseau social mais interne à l'entreprise. Il met en place des communautés thématiques de membres selon les projets ou objectifs à accomplir, contribuant à la mutualisation de services, dont la veille.

Rédigé en 1994, bien avant l'émergence de ces concepts, le rapport Martre déclare que « l'intelligence économique est indissociable de la notion de réseau. [...] Il faut donc favoriser la création de réseaux denses au sein desquels les savoirs professionnels particuliers et les logiques spécifiques de chaque fonction puissent se confronter et se combiner. » [16, Martre et al., p.68].

Ainsi, on pourrait imaginer la mise en place d'une communauté de salariés impactés par l'image publicitaire, en amont au niveau du processus de décision stratégique et en aval au niveau de la création visuelle. La définition du profil et du rôle de chacun (lecteur, contributeur, administrateur, etc.) conditionnerait l'accès et la diffusion des données collectées par l'outil de veille mais aussi le partage de connaissances et les retours d'expériences. Que ce soit ou non dans le cadre d'un réseau social d'entreprise, cette communauté permettrait de trouver les images, l'expertise, les décisions, les interlocuteurs, etc. propres à répondre aux questions qui se posent quotidiennement dans son travail. Néanmoins, ces nouvelles pratiques du « travailler ensemble », de partage, etc. sont difficiles à mettre en place car elles vont à rebours des modes de travail exercés depuis des décennies, des organisations hiérarchiques et des méthodes managériales.

---

<sup>110</sup> PRAX Jean-Yves. Le Manuel du Knowledge Management. Paris, Dunod, 2<sup>e</sup> édition, 2007. p.103. ISBN 978-2-10-051191-4

### **9.2.2.2 Le « community manager » du réseau de veille**

Le documentaliste pourrait trouver ainsi un rôle de médiation transversale. Le rapport Martre dresse le portrait de cet « animateur » : « Véritable interface entre l'acquéreur et l'exploitant, il transforme les demandes de celui-ci en orientations à l'usage du premier, il conduit le traitement de l'information acquise, la distribue à qui en a besoin et dirige l'ensemble de la fonction intelligence économique en animant les réseaux qui lui sont adaptés. » [16, Martre et al., p.102].

Le rapport identifie très clairement la notion de réseaux internes à l'entreprise dès 1994 : « Il apparaît [...] que la coordination des activités d'intelligence économique par "un fonctionnel type", en prise sur tous les secteurs de l'entreprise répond plus efficacement au développement d'un système d'ingénierie stratégique de l'information dans l'entreprise. [...] Il guide l'ensemble des actions d'intelligence économique selon une triple fonction. Celle-ci consiste à animer, voire créer des réseaux, à assurer le traitement de l'information, ainsi qu'à mémoriser et transmettre le savoir-faire relatif à la gestion de l'information utile. » [16, Martre et al., p.69].

Le documentaliste, devenu « community manager », manager de communauté(s), piloterait le dispositif de veille et l'animerait. Il s'inscrirait de la sorte au cœur des processus stratégiques de l'entreprise. À l'heure où l'accès facilité à l'information grâce à Internet donne l'illusion à chacun d'être suffisamment informé et de trouver les données dont il a besoin, ce nouveau rôle adapté au documentaliste-veilleur lui permettrait de se positionner au centre de l'entreprise comme vecteur des synergies des différents collaborateurs. En définitive, l'absence de solutions en matière de recherche et de veille par le contenu de l'image lui confère un avantage de crédibilité et de légitimité de recours, qu'il faut exploiter avant que des algorithmes ne donnent l'illusion de pouvoir se substituer à ses compétences.

# Conclusion

La demande initiale d'une veille des images par l'image n'a pu être satisfaite. En effet, si les recherches en matière de reconnaissance visuelle ont débuté il y a plus de cinquante ans, les résultats sont jusqu'à présent très insuffisants pour généraliser une solution au grand public. Les systèmes demeurent complexes à utiliser. Toutefois, les grands acteurs de l'Internet s'intéressent de très près à ces questions. Ainsi, le 4 septembre dernier, Google a obtenu un brevet intitulé « Reconnaissance automatique et à grande échelle d'objets de vidéo »<sup>111</sup>, qui avait été déposé auprès de l'USPTO<sup>112</sup> le 9 février 2009, complétant un autre brevet sur la reconnaissance faciale dans la vidéo<sup>113</sup>. En conséquence, on peut imaginer que la reconnaissance du contenu de l'image, devenue un enjeu, trouvera bientôt son algorithme. En attendant, il faut travailler sur les mots qui traduisent le contenu de l'image (sa dénotation) et son interprétation (ou connotation).

L'assistante Achat d'art compte beaucoup sur l'automatisation de la veille pour la soulager dans son travail, néanmoins, cette automatisation a aussi ses limites. Dans un premier temps, un travail de réflexion en amont doit être mené, ce que j'ai tenté durant mon stage. La pratique amènera à un réajustement constant du sourcing et des éventuelles requêtes. Toutefois, les solutions mises en place pour le "scénario d'amorçage", combinaison d'outils, peuvent être sources d'erreurs, ce qui demande ainsi une attention particulière. Elles risquent également de générer automatiquement beaucoup de flux, ce qui deviendra ingérable pour l'assistante Achat d'art. Ainsi, la solution devrait montrer rapidement ses limites, surtout pour une personne dont l'activité principale est orientée vers d'autres tâches. Il faudra alors envisager le passage à une solution plus pérenne, avec un outil de plus grande envergure, type plateforme, et surtout géré par une personne dont l'information est le cœur de métier.

Pour répondre à la demande qui avait été formulée, j'ai donc mis en place un agrégateur de flux RSS avec Google Reader, alimenté de six flux et d'une requête transformée en flux, pour se familiariser avec le processus de la veille. Cette solution a été mise en œuvre au cours d'une formation donnée à l'assistante Achat d'art et à la responsable de la documentation du service qui pourrait suppléer l'assistante Achat d'art. Je leur ai fourni un sourcing comprenant plus de 200 références à tester, corriger, enrichir. Ce document a été transposé en marque-pages, respectant la même organisation, et exporté sur les postes de travail de ces deux personnes. Ainsi, selon une approche pull<sup>114</sup>, elles peuvent aller consulter et évaluer les sources que j'ai relevées.

J'ai élaboré un cahier des charges fonctionnel et technique au sein duquel la démarche est décrite, de l'analyse de l'existant à la proposition d'outils, en passant par la schématisation des processus, les tableaux comparatifs des différents outils de veille et de recherche. Ainsi, si le service souhaite dépasser le premier scénario mis en place pour installer une solution pérenne, il disposera d'éléments (qui seront certes à réactualiser) sur lesquels se fonder.

En outre, dans le cadre du traitement des dossiers de photographies de l'assistante Achat d'art, j'ai réalisé des tableaux de suivi des photographes et des mannequins, aussi bien sur

---

<sup>111</sup> GOOGLE. Login to a computing device based on facial recognition. *Yoshimichi Matsuoka*. États-Unis, N° 13/247,652. 2012-09-04 [consulté le 06.10.2012].

<<http://tinyurl.com/8dwq9ol>>

<sup>112</sup> USPTO : United States Patent and Trademark Office, soit le Bureau américain des brevets et des marques de commerce.

<sup>113</sup> CHÉRON Antoine. Reconnaissance vidéo : Google pousse encore plus loin son intrusion dans la vie privée. In Atlantico.fr [site en ligne]. 13.09.2012 [consulté le 14.09.2012].

<<http://www.atlantico.fr/decryptage/reconnaissance-video-google-pousse-encore-plus-loin-intrusion-dans-vie-privee-antoine-cheron-480800.html>>

<sup>114</sup> Pull : l'utilisateur va chercher lui-même l'information.



le nombre de photos collectées par année et leur progression que sur leur contenu. Ces tableaux qui synthétisent cinq années de recherche et de collecte, peuvent être complétés, si l'assistante Achat d'art dispose de temps pour le faire.

Je n'ai pas eu le temps d'élaborer des indicateurs de performance pour la solution mise en place. Toutefois, si le service souhaite se fonder sur des éléments objectifs quant à l'utilité et la pertinence de l'outil, il pourra, dans un premier temps, recourir aux statistiques fournies par l'agrégateur.

Je suggère également de travailler sur le vocabulaire pour élaborer des requêtes. Comme je l'ai souvent répété et surtout développé dans la deuxième partie, la recherche et donc la veille par le contenu de l'image ne sont pas encore d'actualité, d'autant plus dans l'univers si subjectif et émotionnel de la création du secteur du luxe. Aussi le travail des mots est-il indispensable. Si l'assistante Achat d'art doit avoir en tête, de manière plus ou moins consciente, une liste de mots-clefs, je pense néanmoins qu'il faut étudier les requêtes lancées par les personnes en charge de la création lorsqu'elles utilisent le moteur de recherche de l'intranet. Cela permettrait d'identifier leurs besoins, leurs centres d'intérêt. Même si la veille ne leur est pas destinée dans un premier temps, cette démarche permettrait de déterminer leurs sources d'inspiration pour mettre en image les décisions stratégiques visuelles.

Pendant ces mois de vigilance sur ces sujets, dans le cadre de mon stage et de la rédaction de mon mémoire, j'ai constaté que l'actualité en matière de reconnaissance de l'image, de recherche par le contenu, de Web sémantique était extrêmement riche. Aussi, puisque mon stage sert de point de départ à la mise en place d'un système de veille, je conseille aux personnes qui vont approfondir le sujet de rester en alerte quant aux innovations et aux orientations juridiques et politiques. En outre, si elles doivent choisir un outil de veille professionnel, je leur recommande la plus grande vigilance quant à l'évolutivité de la solution retenue, qui doit être en mesure de s'adapter aux formats du Web et surtout ceux du Web sémantique.

En effet, l'implication d'acteurs tels que Google sur ce créneau suppose que le linked data est probablement la technologie qui rendra ou non accessibles les données sur le Web. Ne pas disposer d'outils adaptés à cette évolution pourrait empêcher l'entreprise d'acquérir les données dont elle a besoin. La substitution d'un outil par un autre est de nos jours très rapide et il est quasiment impossible de suivre ce rythme effréné. C'est pourquoi il me semble pertinent de se concentrer sur les solutions qui disposent de fortes capacités d'adaptation et qui sont composées de modules ou briques actualisables, sans avoir à changer l'ensemble du système à chaque progrès majeur.

Enfin, pour les prochaines années, j'engagerais une réflexion sur la place de la veille au sein de l'entreprise, sous forme de veille collaborative, selon une démarche de Knowledge Management ou dans le cadre d'un réseau social d'entreprise.

En effet, plusieurs veilles ont été détectées lors du stage. Il pourrait être intéressant de regrouper leur accès en un seul point d'entrée, au sein, par exemple, d'un intranet dédié. Cette démarche a besoin d'un fort appui de la hiérarchie car elle affronte des pratiques installées depuis longtemps au cœur des entreprises. Un travail d'accompagnement est ainsi à prévoir :

- d'une part, rassurer les personnes qui réalisent ces veilles pour leur expliquer qu'elles ne seront pas dépossédées de leur travail mais qu'au contraire celui-ci sera valorisé ;
- d'autre part, garantir aux services la confidentialité des données qu'ils ne souhaitent pas mutualiser mais qui seront accessibles selon les profils.

Cette solution aurait comme avantages :

- de faciliter l'accès et donc d'augmenter l'audience des veilles réalisées ;
- de valoriser ces veilles et de mettre en exergue les experts par domaine ;
- d'archiver les éditions qui circulent sous forme de PDF, avec des métadonnées, et de permettre d'effectuer des recherches ;
- d'alléger les serveurs, puisque les PDF des veilles ne seraient plus stockées sur chaque poste mais sur un environnement commun, l'intranet, et les mails de diffusion ne contiendraient plus que l'hyperlien permettant d'y accéder.

Puisque l'entreprise X privilégie la création, la gestion de cet intranet reviendrait à un documentaliste de la Direction Artistique. Pour les mêmes raisons, dans le cas où un réseau social d'entreprise serait développé à la place d'un intranet, je pense qu'il serait judicieux de créer une communauté de veille animée par cette même personne.

De ce fait, le service Achat d'art serait renforcé au cœur de la stratégie de l'entreprise et ses documentalistes seraient positionnées comme ses animatrices.

# **Bibliographie**

La bibliographie a été arrêtée le 9 octobre 2012. Elle est articulée en trois parties, une bibliographie principale, une bibliographie complémentaire et un index des auteurs. Les références bibliographiques des deux premières parties sont ordonnées par ordre antéchronologique, de la plus récente à la plus ancienne, puis par ordre alphabétique du nom des auteurs. Ce mode de classement a été choisi en raison de la forte actualité du secteur et de la rapide obsolescence des recherches.

La bibliographie principale est divisée en cinq thématiques. Chaque référence bibliographique fait l'objet d'une courte présentation. Les chiffres qui les précèdent correspondent aux appels bibliographiques insérés dans le développement du mémoire.

## **Bibliographie principale thématique**

### **Veille**

1-DESCHAMPS Christophe, MOINET Nicolas. La boîte à outils de l'intelligence économique. Paris, Dunod, 2011, 185 p. ISBN 978-2-10-055112-5

Ce livre propose des outils très divers, répartis en six dimensions, à choisir selon ses besoins.

2-FOENIX-RIOU Béatrice. Recherche éveillée sur Internet : mode d'emploi : outils et méthodes pour explorer le Web : Web visible, Web invisible, Web social, Web temps réel. Paris, Lavoisier, Bases Publications, 2011. 368 p. ISBN 978-2-7430-1342-4

Livre qui détaille avec précision et simplicité les outils du Web pour effectuer des recherches avec de nombreux conseils.

3-GORIA Stéphane. Éléments de veille créative pour contribuer à l'innovation produit : la mise en œuvre de cartes d'affrontement de produits. Revue internationale d'intelligence économique [en ligne], Paris, Lavoisier. 2011, numéro 1. [consulté le 20.08.2012], volume 3, p.57-72.

<<http://www.cairn.info/revue-internationale-d-intelligence-economique-2011-1-page-57.htm>>. ISSN 2101-647X

Cet article est intéressant pour les multiples définitions détaillées que l'auteur propose de la veille.

4-KHÉNISSI Mohamed Ghazi, GHARBI Jamel-Eddine. "La veille stratégique". Bilan de la culture numérique la veille du 2.0. Les Cahiers du numérique [en ligne], Paris, Lavoisier. 2010, numéro 1. [consulté le 19.08.2012], volume 6, p.135-156.

<<http://www.cairn.info/revue-les-cahiers-du-numerique-2010-1-page-135.htm>>. ISSN 1622-1494

Cet article propose une évaluation de la maîtrise de l'environnement numérique par des salariés en Tunisie, notamment dans la perspective d'une veille sur Internet.

5-MIAUX Jean-François. Mise en œuvre d'une activité de veille. Le cas Réseau Ferré de France. 2010. 1 volume, 104 p. Mémoire, Titre professionnel « Chef de projet en ingénierie documentaire » niveau I, INTD, 2010

Ce mémoire délivre des conseils pratiques sur la mise en œuvre d'un système de veille avec un cas concret d'application.

6-BOULIFA TAMBOURA Inès, BEN AMMAR MAMLOUK Zeineb. Identification et validation des facteurs critiques de succès pour la mise en place d'un dispositif de veille stratégique. La Revue des Sciences de Gestion [en ligne], Paris, Direction et Gestion (la RSG). 2009/3. [consulté le 20.08.2012], numéro 237-238, p.187-193. <<http://www.cairn.info/revue-des-sciences-de-gestion-2009-3-page-187.htm>>. ISSN 1160-7742

Cet article propose des facteurs de réussite pour un projet de veille.

7-PHIRMIS Laura. Capitaliser et partager les connaissances pour assurer une meilleure gouvernance associative. Le cas de la mise en place d'un système de veille au Service Actions Diabète de l'Association Française des Diabétiques. 2009. 1 volume, 122 p. Mémoire, Titre professionnel « Chef de projet en ingénierie documentaire » niveau I, INTD, 2009

Ce mémoire s'intéresse au Knowledge Management, le définit, propose des outils et expose sa pertinence dans le cadre d'un projet de veille.

8-TOUPIN LEMAIRE Stéphanie. Outils et méthodes de diffusion des résultats de la veille. Le cas du Centre International d'Études pédagogiques (CIEP). 2009. 1 volume, 100 p. Mémoire, Titre professionnel « Chef de projet en ingénierie documentaire » niveau I, INTD, 2009

Ce mémoire s'attache à la notion de travail collaboratif dans le cadre du Web 2.0 et au Knowledge Management, appliqués à la diffusion de la veille.

9-BALMISSE Gilles, MEINGAN Denis. La veille 2.0 et ses outils. Paris, Lavoisier, Hermès Science publications, 2008. 231 p. Collection Management et informatique (dirigée par Nicolas Manson). ISBN 978-2-7462-1929-8. ISSN 1365-7361

Ce livre définit la veille 2.0, explique les technologies qui la supportent et propose de très nombreux outils.

10-DIALLO Alpha, JDEY Aref. La veille : collaborative, forcément collaborative... Documentaliste – Sciences de l'information [en ligne], Paris, ADBS. 2008, numéro 4. [consulté le 19.08.2012], volume 45, p.60-62. <<http://www.cairn.info/revue-documentaliste-sciences-de-l-information-2008-4-page-58.htm>>. ISSN 0012-4508

Cet article présente le travail collaboratif et expose ainsi sa pertinence pour l'activité de veille.

11-HA Seolin. Étude de faisabilité pour l'automatisation de veille réglementaire à moyens constants : le cas du Centre de documentation de l'expertise du Muséum national d'histoire naturelle. 2008. 1 volume, 108 p. Mémoire, Titre professionnel « Chef de projet en ingénierie documentaire » niveau I, INTD, 2008

Ce mémoire donne une définition développée de la veille ainsi qu'un focus sur son automatisation.

12-MESGUICH Véronique. Web 2.0, web 3.0, les nouveaux habits de la veille. Documentaliste – Sciences de l'information [en ligne], Paris, ADBS. 2008, numéro 4. [consulté le 19.08.2012], volume 45, p.58-59. <<http://www.cairn.info/revue-documentaliste-sciences-de-l-information-2008-4-page-58.htm>>. ISSN 0012-4508

Cet article présente les nouvelles fonctionnalités du Web 2.0 et 3.0 et leurs conséquences sur le travail de veille.

13-PERBOST Olivia. Organisation des informations et partage des connaissances dans un système de veille : l'exemple d'un partenariat entre le centre de documentation de la direction de l'architecture et du patrimoine et une sous-direction au ministère de la culture et de la communication. 2007. 1 volume, 93 p. Mémoire, Titre professionnel « Chef de projet en ingénierie documentaire » niveau I, INTD, 2007

Le mémoire propose plusieurs définitions de la veille ainsi que ses typologies et fournit des solutions de diffusion.

14-AJALBERT Marc. Les limites de l'automatisation de la veille : Le cas d'une société de travail temporaire. 2006. 1 volume, 48 p. Mémoire, Titre professionnel « Chef de projet en ingénierie documentaire » niveau I, INTD, 2006

Ce mémoire expose les limites de l'automatisation de la veille et présente des solutions pour y remédier.

15-ASSOCIATION FRANÇAISE DE NORMALISATION (AFNOR). Prestations de veille et prestations de mise en place d'un système de veille. XP X 50-053. Paris, AFNOR, Avril 1998. 23 p. ISSN 0335-3931

Norme AFNOR relative aux activités de veille qui présente les bonnes pratiques et propose des outils pour aider à la mise en place d'un système de veille.

16-MARTRE Henri, CLERC Philippe, HARBULOT Christian, COMMISSARIAT GÉNÉRAL AU PLAN. Intelligence économique et stratégie des entreprises. [en ligne]. Paris, La Documentation française, 1994. [consulté le 20.08.2012], 167 p.

<<http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/074000410/0000.pdf>>

Rapport sur l'intelligence économique et la veille : présentation historique, définitions, comparaisons internationales, etc. dont les conclusions, lexicales et stratégiques, sont toujours d'actualité.

## **Image**

17-HAUPAIS Coline. De l'image au langage. Enjeux de la gestion et de l'organisation des connaissances audiovisuelles dans un service de documentation d'une grande entreprise de luxe. 2011. 1 volume, 100 pages. Mémoire, Titre professionnel « Chef de projet en ingénierie documentaire » niveau I, INTD, 2011

Ce mémoire a été réalisé dans la même entreprise que la mienne et traite donc de problématiques auxquelles j'ai été confrontée, notamment sur les images.

18-REBAI Ahmed. Interactive Object Retrieval using Interpretable Visual Models [en ligne]. Thèse de doctorat, Science de l'informatique, Université Paris Sud – Paris XI, 2011. [consulté le 31.08.2012]. 198 p.

<<http://tel.archives-ouvertes.fr/docs/00/59/69/16/PDF/Thesis.pdf>>

Cette thèse scientifique (avec beaucoup d'algorithmes, graphiques, tableaux) fait le point, dans un premier temps, sur les solutions et technologies développées pour la recherche d'images.

19-BAYLE Frédéric. Conception d'un thésaurus iconographique dans le cadre du développement d'une agence de photographies créatives sur Internet : Le cas de l'agence photos Voz'Image. 2008. 1 volume, 141 p. Mémoire, Titre professionnel « Chef de projet en ingénierie documentaire » niveau I, INTD, 2008

Ce mémoire s'intéresse à l'image, et notamment la photographie d'art, à la recherche d'images et ainsi aux métadonnées de la photographie, au langage documentaire, aux normes et aux implications du Web 2.0 et 3.0.

20-REGIMBEAU Gérard. L'image d'art entre analyse critique et analyse documentaire. Documentaliste – Sciences de l'information [en ligne], Paris, ADBS. 2007, numéro 2. [consulté le 20.08.2012.2012], volume 44, p.130-137. <<http://www.cairn.info/revue-documentaliste-sciences-de-l-information-2007-2-page-130.htm>>. ISSN 0012-4508

Cet article met en lumière le problème de traitement documentaire de l'image (analyse et indexation).

21-MELOT Michel. L'image n'est plus ce qu'elle était. Documentaliste – Sciences de l'Information [en ligne], Paris, ADBS. 2005, numéro 6. [consulté le 31.08.2012], volume 42, p.361-365.

<<http://www.cairn.info/revue-documentaliste-sciences-de-l-information-2005-6-page-361.htm>> ISSN 0012-4508.

Cet article s'intéresse au rapport entre l'image et le texte et soulève les difficultés de nommer et d'indexer une image.

22-BARTHES Roland. La Chambre claire. Note sur la photographie. Paris, Gallimard, 1980. 192 p. Collection Cahiers du cinéma. ISBN 207020541X

Ce livre explique les concepts clefs de l'image qui éclairent sur la difficulté de déterminer des termes de recherche pour les images.

23-BARTHES Roland. Rhétorique de l'image. Communications [en ligne], Paris, Éditions du Seuil. 1964, numéro 4. [consulté le 20.08.2012], p.40-51.

<[http://www.persee.fr/web/revues/home/prescript/article/comm\\_0588-8018\\_1964\\_num\\_4\\_1\\_1027](http://www.persee.fr/web/revues/home/prescript/article/comm_0588-8018_1964_num_4_1_1027)>. ISSN 0588-8018

Cet article est considéré comme fondateur sur l'analyse sémiologique de l'image photographique, fondée sur un exemple extrait de la publicité.

24-BARTHES Roland. Le message photographique. Communications [en ligne], Paris, Éditions du Seuil. 1961, numéro 1. [consulté le 31.08.2012], p.127-138.

<[http://www.persee.fr/web/revues/home/prescript/article/comm\\_0588-8018\\_1961\\_num\\_1\\_1\\_921](http://www.persee.fr/web/revues/home/prescript/article/comm_0588-8018_1961_num_1_1_921)>. ISSN 0588-8018

L'auteur analyse l'image photographique en dégagant deux messages, l'un dénoté et l'autre connoté, et s'intéresse au rapport qu'elle entretient avec le texte.

## **Internet**

25-W3C.ORG. Semantic Web. In W3C [site en ligne]. mis à jour le 25.08.2012 [consulté le 29.09.2012]. <[http://www.w3.org/standards/semanticweb/#w3c\\_overview](http://www.w3.org/standards/semanticweb/#w3c_overview)>

Cette page web de la W3C définit ce que l'on appelle le Web sémantique ou Web 3.0.

26-MANGEMATIN Julie. Web 3.0 : la révolution des datas annonciatrice de l'Internet du futur. In Atlantico.fr [site en ligne]. 11.08.2012 [consulté le 15.08.2012].

<<http://www.atlantico.fr/decryptage/web-30-revolution-datas-annonciatrice-internet-futur-424010.html?page=0,0>>

Cet article en ligne traite de la multiplication exponentielle des données et des conséquences sur le traitement de l'information.



27-MÉDIAMÉTRIE ESTAT. La fréquentation des sites Internet français. Classement CybereStat [en ligne]. 07.08.2012 [consulté le 15.09.2012], 5 p.

<<http://www.mediametrie.fr/internet/communiqués/telecharger.php?f=1ecfb463472ec9115b10c292ef8bc986>>

Enquête de Médiamétrie sur les habitudes des Français sur Internet, qui met en évidence le quasi monopole de Google dans les moteurs de recherche.

28-YOON JungWon. Searching images in daily life. Library & Information Science Research [en ligne], Elsevier. Octobre 2011. [consulté le 17.08.2012], volume 33, p.269-275.

<<https://cassiope.cnam.fr/http/www.sciencedirect.com/science/article/pii/S0740818811000612>>. ISSN 0740-8188

Cet article expose les résultats d'une enquête menée dans un milieu universitaire sur les besoins des personnes qui recherchent des images. Il en déduit des implications technologiques et méthodologiques pour les systèmes de recherche.

29-DATTA Rittendra, JOSHI Dhiraj, LI Jia, WANG James Z. Image Retrieval : Ideas, Influences, and Trends of the New Age. ACM Computer Surveys [en ligne], New York, ACM Digital Library. Avril 2008, numéro 2. [consulté le 25.08.2012], volume 40, article 5, 60 p.

<[http://dl.acm.org/ft\\_gateway.cfm?id=1348248&type=pdf&coll=DL&dl=GUIDE&CFID=121023225&CFTOKEN=74117176](http://dl.acm.org/ft_gateway.cfm?id=1348248&type=pdf&coll=DL&dl=GUIDE&CFID=121023225&CFTOKEN=74117176)>. ISSN 0360-0300

Cet article est souvent cité par des publications plus récentes. En effet, il expose clairement les fondements et principes du CBIR.

30-O'REILLY Tim. What is web 2.0, design, patterns and business models for the next generation software [blog en ligne]. 30.09.2005 [consulté le 31.08.2012].

<<http://oreilly.com/web2/archive/what-is-web-20.html>>

Tim O'Reilly reproduit sur son blog l'intervention au cours de laquelle il a défini le Web 2.0. Une version traduite est proposée ici : BOISSEAU Jean-Baptiste (traduction), KAPLAN Daniel (révisions). O'REILLY Tim. Qu'est ce que le web 2.0 : Modèles de conception et d'affaires pour la prochaine génération de logiciels. InternetActu [site en ligne]. 21 avril 2006. [consulté le 31.08.2012]. <<http://www.internetactu.net/?p=6421>>

## **CBIR**

31-GOHRINGER Carl. The application of face recognition in airports. Biometric Technology Today [en ligne], Elsevier. Juillet-août 2012, numéro 7. [consulté le 24.08.2012], volume 2012, p.5-9.

<<https://cassiope.cnam.fr/http/www.sciencedirect.com/science/article/pii/S0969476512701480>>. ISSN 096964765

Cet article expose les fonctionnalités offertes par la reconnaissance faciale dans les aéroports.

32-ORESMUS Will. In Artificial Intelligence Breakthrough, Google Computers Teach Themselves To Spot Cat on YouTube. In Slate.fr [site en ligne]. 27.06.2012 [consulté le 26.07.2012].

<[http://www.slate.com/blogs/future\\_tense/2012/06/27/google\\_computers\\_learn\\_to\\_identify\\_cats\\_on\\_youtube\\_in\\_artificial\\_intelligence\\_study.html](http://www.slate.com/blogs/future_tense/2012/06/27/google_computers_learn_to_identify_cats_on_youtube_in_artificial_intelligence_study.html)>

Cet article en ligne expose l'expérience réalisée par le réseau neural de Google qui a abouti à l'identification des chats, sans apprentissage préalable.

33-MARKOFF John. How Many Computers to Identify a Cat ? 16,000. The New York Times.com [site en ligne]. 25.06.2012 [consulté le 26.07.2012].

<[http://www.nytimes.com/2012/06/26/technology/in-a-big-network-of-computers-evidence-of-machine-learning.html?\\_r=1](http://www.nytimes.com/2012/06/26/technology/in-a-big-network-of-computers-evidence-of-machine-learning.html?_r=1)>

Le journaliste expose, dans cet article accessible en ligne, l'expérience réussie par le réseau neural de Google de reconnaissance de l'image des chats, sans phase préalable d'apprentissage.

34-EUDES Yves. Souriez, vous êtes identifiés. M le magazine du Monde [site en ligne]. 06.05.2012 [consulté le 02.07.2012].

<[http://www.lemonde.fr/technologies/article/2012/05/04/souriez-vous-etes-identifies\\_1694639\\_651865.html](http://www.lemonde.fr/technologies/article/2012/05/04/souriez-vous-etes-identifies_1694639_651865.html)>

Article qui expose une utilisation commerciale de la reconnaissance faciale et présente ainsi son fonctionnement.

35-BANNOUR Hichem, HUDELLOT Céline. Combinaison d'information visuelle, conceptuelle, et contextuelle pour la construction automatique de hiérarchies sémantiques adaptées à l'annotation d'images. In Reconnaissance des Formes et Intelligence Artificielle (RFIA) [en ligne]. 24-27.01.2012, Lyon. 8p. Mis en ligne dans Hal le 17.01.2012 [consulté le 13.08.2012]. <<http://hal.archives-ouvertes.fr/hal-00656489/>>

Cet article scientifique expose une expérience menée sur la sémantique afin d'améliorer les résultats de classification et d'annotation des images.

36-MÉNARD Éline. Méthodes et défis du repérage d'images sur le Web : Jean et John cherchent-ils de la même manière ? The Partnership : The Canadian Journal of Library and Information Practice and Research [en ligne], Partnership. 2012, numéro 1. [consulté le 14.08.2012], volume 7, 14 p.

<<http://journal.lib.uoguelph.ca/index.php/perj/article/view/1863/2525>>. ISSN : 1911-9593  
Article de vulgarisation sur le CBIR, l'indexation des images, le collaboratif, la recherche d'information multilingue et les moteurs de recherche et d'actualité.

37-WANG Jiangning, JI Liqiang, LIANG Aiping, YUAN Decheng. The identification of butterfly families using content-based image. Biosystems Engineering [en ligne], Elsevier. Janvier 2012, numéro 1. [consulté le 24.08.2012], volume 111, p.24-32.

<<https://cassiope.cnam.fr/http/www.sciencedirect.com/science/article/pii/S1537511011001784>>. ISSN 1537-5110

Cet article scientifique présente des explications intéressantes sur le CBIR, notamment avec un cas pratique sur la recherche des familles de papillons grâce au contenu de l'image.

38-HARZALLAH Hedi. Contribution à la localisation et à la reconnaissance d'objets dans les images [en ligne]. 2011. 1 volume, 120 p. Thèse de doctorat, Mathématiques et Informatique, Université de Grenoble, 2011. [consulté le 25.08.2012]. <[http://tel.archives-ouvertes.fr/docs/00/62/80/27/PDF/19220\\_HARZALLAH\\_2011\\_archivage\\_1\\_.pdf](http://tel.archives-ouvertes.fr/docs/00/62/80/27/PDF/19220_HARZALLAH_2011_archivage_1_.pdf)>

Thèse difficile d'accès pour des non-mathématiciens (nombreux algorithmes) mais qui, dans ses premières pages, explique les méthodes et techniques de description d'images rencontrées dans d'autres références bibliographiques.

39-HE Ruhan, XIONG Naixue, YANG Laurence T., PARK Jong Hyuk. Using Multi-Modal Semantic Association Rules to fuse keywords and visual features automatically for Web image retrieval. Information Fusion [en ligne], Elsevier. Juillet 2011. [consulté le 31.08.2012], volume 12, p.223-230.

<<https://cassiope.cnam.fr/http/www.sciencedirect.com/science/article/pii/S1566253510000230>>. ISSN 1566-2535

Cet article scientifique expérimente une fusion de mots-clés et de caractéristiques visuelles de l'image pour effectuer des recherches d'images sur le Web.

40-SCHMID Cordelia. Rapport d'activité. Project/Team LEAR : Learning and Recognition in Vision [en ligne]. INRIA et CNRS. 2011 [consulté le 17.08.2012], 23 p.

<<http://raweb.inria.fr/rapportsactivite/RA2011/lear/lear.pdf>>

Ce rapport d'activité de l'équipe LEAR de l'INRIA présente aux professionnels leurs axes de recherche et leurs réalisations.

41-VIEUX Rémy. Extraction de descripteurs pertinents et classification pour le problème de recherche des images par le contenu [en ligne]. 2011. 1 volume, 102 pages. Thèse de

doctorat, Informatique, Université de Bordeaux 1, 2011. [consulté le 25.08.2012].  
<<http://www.theses.fr/2011BOR14244>>

Cette thèse, difficile d'accès pour les non-initiés à l'informatique, délivre cependant les connaissances de base pour aborder la recherche d'images par le contenu.

42-PONCE Jean, Comment donner un sens à l'image numérique ? Les Cahiers de l'INRIA [en ligne], INRIA. Novembre 2009, numéro 435. [consulté le 13.08.2012], p.1-3.  
<<http://hal.inria.fr/inria-00511918/PDF/inria-n435-nov09.pdf>>

Article de vulgarisation scientifique sur la recherche d'image, la vision artificielle et les méthodes élaborées depuis les années 1960.

43-NOBIS Pierre. Chercher des images par l'image. In Service de documentation de l'Académie de Rouen [site en ligne]. 15.11.2008, mis à jour le 03.03.2012 [consulté le 14.08.2012]. <<http://documentation.spip.ac-rouen.fr/spip.php?article205>>

Article de vulgarisation sur les outils et les méthodes de recherche par le contenu de l'image.

44-GROUPEMENT FRANÇAIS DE L'INDUSTRIE DE L'INFORMATION (GFII). Des clichés et des clics. Le poids de la photographie dans la société de l'information. Ruth Martinez (coordinatrice). Paris, ADBS, 2007. 113 pages. ISBN 978-2-84365-094-9. ISSN 1762-8288  
Ce livre s'intéresse aux différents aspects de la photographie dans l'environnement numérique, de la recherche d'images à leur diffusion.

45-L'indexation multimédia. Description et recherche automatique (sous la direction de Patrick Gros). Paris, Lavoisier, Hermès Science, 2007. 330 pages. XI pages. Série Signal et image (dirigée par Francis Castanié et Henri Maître). ISBN 978-2-7462-1492-7

Ce livre scientifique couvre les thèmes de la recherche d'images par le contenu, des métadonnées afférentes à l'image, des méthodes d'interrogation et de recherche, de la description de l'image et de son indexation, de la reconnaissance faciale, entre autres.

## **Luxe**

46-ROUX Élyette. Le luxe au temps des marques. Géoéconomie [en ligne], Paris, Éditions Choiseul. 2009/2. [consulté le 03.09.2012], numéro 49, p.19-36.

<<http://www.cairn.info/revue-geoéconomie-2009-2-page-19.htm>>. ISSN 1284-9340

Cet article présente les stratégies élaborées dans le secteur du luxe.

47-DEREUMAUX René-Maurice. Le luxe et l'image de marque. Market Management [en ligne], Paris, Eska. 2007, numéro 1. [consulté le 03.09.2012], volume 7, p.70-78.  
<<http://www.cairn.info/revue-market-management-2007-1-page-70.htm>>. ISSN 1179-3572

L'auteur expose dans cet article l'importance de la notion d'image de marque dans le secteur du luxe.

48-BAUDRILLARD Jean. La morale des objets. Communications [en ligne], Paris, Éditions du Seuil. 1969, numéro 13. [consulté le 03.09.2012], p.23-50.

<[http://www.persee.fr/web/revues/home/prescript/article/comm\\_0588-8018\\_1969\\_num\\_13\\_1\\_1184](http://www.persee.fr/web/revues/home/prescript/article/comm_0588-8018_1969_num_13_1_1184)>. ISSN 0588-8018

Dans cet article, le philosophe développe le concept d'ostentation dans les modes de consommation.

## **Bibliographie complémentaire**

DELENGAINE, Xavier. Organiser sa veille sur Internet. Au-delà de Google... Outils et astuces pour le professionnel. Paris, Éditions Eyrolles, 2012. 320 p. ISBN 978-2-212-13295-3

GUILLAUMIN Matthieu, MENSINK Thomas, VERBEEK Jakob, SCHMID Cordelia. Face recognition from caption-based supervision. International Journal of Computer Vision [en ligne], Springer. 2012, numéro 1. [consulté le 24.08.2012], volume 96, p.64-82. <<http://hal.inria.fr/inria-00585834/PDF/paper.pdf>>. ISSN 0920-5691

LI Billy Y.L., LIU Wanquan, AN Senjian, KRISHNA Aneesh, XU Tianwei. Face recognition using scales of discriminant color space transform. Neurocomputing [en ligne], Elsevier. 01.10.2012. [consulté le 13.08.2012], volume 94, p.68-76.

<<https://cassiope.cnam.fr/http/www.sciencedirect.com/science/article/pii/S0925231212003360>>. ISSN 092562312

GUOYONG Duan, YANG Jing, YANG Yilong. Content-Based Image Retrieval Research. Physics Procedia [en ligne], Elsevier. 2011. [consulté le 17.08.2012], volume 22, p.471-77.

<<https://cassiope.cnam.fr/http/www.sciencedirect.com/science/article/pii/S1875389211007279>>. ISSN 1875-3892

LIU Chengjun. Extracting discriminative color features for face recognition. Pattern Recognition Letters [en ligne], Elsevier. 15.10.2011, numéro 14. [consulté le 13.08.2012], volume 32, p.1796-1804.

<https://cassiope.cnam.fr/http/www.sciencedirect.com/science/article/pii/S0167865511002455>. ISSN 0167-8655

CARON-FASAN Marie-Laurence, LESCA Humbert, BUITRAGO Alex, CASAGRANDE Annette. Comment collecter des données numériques et textuelles, utiles à la phase d'exploitation

d'un dispositif de veille anticipative : problématique et proposition d'un outil. The Journal of Competitive Intelligence and Management [en ligne], Society of Competitive Intelligence Professionals. 2010 [consulté le 22.08.2012], volume 1, numéro 2, p.5-18. <[http://atlas.irit.fr/PIE/VSSST/Actes\\_VSSST\\_2010\\_Toulouse/Session%2010%20-%20Signaux%20faibles%201/CARON-%20LESCA-BUITRAGO-CASAGRANDE.pdf](http://atlas.irit.fr/PIE/VSSST/Actes_VSSST_2010_Toulouse/Session%2010%20-%20Signaux%20faibles%201/CARON-%20LESCA-BUITRAGO-CASAGRANDE.pdf)>.

ISSN 1540-4242

LEITZELMAN Mylène. « La veille 2.0. » Outiller les interactions sociales au sein du processus de veille. Les Cahiers du numérique [en ligne]. Paris, Lavoisier. 2010, numéro 1. [consulté le 19.08.2012], volume 6, p.119-133. <<http://www.cairn.info/revue-les-cahiers-du-numerique-2010-1-page-119.htm>>. ISSN 1622-1494

LESCA Humbert et al. La surinformation causée par l'Internet : un facteur d'échec paradoxal largement avéré : veille stratégique – cas concrets, retours d'expérience et piste de solution. La Revue des Sciences de Gestion [en ligne], Paris, Direction et Gestion (la RSG). Septembre-décembre 2010/5. [consulté le 19.08.2012], numéro 245-246, p.35-42. <<http://www.cairn.info/revue-des-sciences-de-gestion-2010-5-page-35.htm>>. ISSN 1160-7742

TOLLARI Sabrina et al. Exploitation du contenu visuel pour améliorer la recherche textuelle d'images en ligne. Document numérique [en ligne] Paris, Lavoisier. 2010, numéro 1. [consulté le 14.08.2012], volume 13, p.187-209. <<http://www.cairn.info/revue-document-numerique-2010-1-page-187.htm>> ISSN 1279-5127

COURRIER Serge. Produire des fils RSS et Atom. Paris, ADBS, 2009. 63 p. Collection L'essentiel sur... ISBN 978-2-84365-101-4

GERVEREAU, Laurent. Voir, comprendre, analyser les images. Paris, La Découverte, 2004. 198 p. Collection «Guides Repères ». ISBN 2707142999

## **Index des auteurs**

AFNOR : p.22, 23, 48

AFP : p.42, 63

AJALBERT Marc : p.22, 53, 61

ANDRIEU Olivier : p.26, 40

BALMISSE Gilles, MEINGAN Denis : p.22, 62

BANNOUR Hichem, HUDELLOT Céline : p.38

BARTHES Roland : p.30, 31, 32

BATTISTI Michèle : p.39, 63, 64  
BAUDRILLARD Jean : p.17  
BAYLE Frédéric : p.30, 31, 33, 39, 54, 57  
BLADIER Cyril : p.25  
BLOOMBERG : p.42  
BOULESTAN Rénaud : p.25,  
BOULIFA TAMBOURA Inès, BEN AMMAR MAMLOUK Zeineb : p.52  
CASTANET Caroline : p.23  
CHAMPEAU Guillaume : p.64  
CHÉRON Antoine : p.71  
CONWAY Amanda : p.42  
DATTA Rittendra, JOSHI Dhiraj, LI Jia, WANG James Z. : p.37  
DEREUMAUX René-Maurice : p.17  
DESCHAMPS Christophe, MOINET Nicolas : p.53, 62  
DIALLO Alpha, JDEY Aref : p.67  
DURAND-PARENTI Chloé : p.36  
EUDES Yves : p.42, 43, 44  
FEITZ Anne : p.64  
FINN Robert : p.59  
FOENIX-RIOU : p.25, 27, 39, 54, 55, 58  
FRION Pascal : p.37  
GERVEREAU Laurent : p.33  
GFII : p.30, 32, 35, 38, 57  
GOHRINGER Carl : p.41  
GORIA Stéphane : p.24  
GROS Patrick : p.32, 35, 38, 41, 60  
HA Seolin : p.53  
HARZALLAH Hedi : p.34  
HAUPAIS Coline : p.16, 31, 35, 45  
HE Ruhan, XIONG Naixue, YANG Laurence T., PARK Jong Hyuk : p.38  
HUDSON Alex : p.28  
JAKOBIAK François : p.23  
JDEY Aref : p.27, 28, 67  
KHÉNISSI Mohamed Ghazi, GHARBI Jamel-Eddine : p.24, 37, 48  
LE MONDE.FR : p.42, 63, 64  
MANGEMATIN Julie : p.26, 27  
MARKOFF John : p.43

MARTRE Henri, CLERC Philippe, HARBULOT Christian, COMMISSARIAT GÉNÉRAL AU PLAN :  
p.22, 68, 69  
MÉDIAMÉTRIE ESTAT : p.39, 96  
MELOT Michel : p.32, 33  
MÉNARD Éline : p.25, 33, 35, 45  
MESGUICH Véronique : p.25, 26, 28  
MESSAGER Véronique : p.51, 52  
MIAUX Jean-François : p.22, 53, 54  
MILLS Elinor : p.43  
NOBIS Pierre : p.38, 60  
O'MALLEY Sean : p.40  
O'REILLY Tim : p.24, 25, 26, 27  
ORESKOVIC Alexei, HUET Nathalie : p.44  
ORESMUS Will : p.43, 45  
PERBOST Olivia : p.22  
PHIRMIS Laura : p.67, 68  
PONCE Jean : p.34  
POUCHARD Alexandre : p.64  
RAULIN Antoine : p.35  
REBAI Ahmed : p.35  
REGIMBEAU Gérard : p.32  
ROUX Élyette : p.16  
SCHMID Cordelia : p.34, 35  
SEDOURAMANE Hugo : p.36  
SINGHAL Amit : p.26  
TERNISIEN Xavier : p.64  
TESQUEST Olivier : p.64, 65  
TOUPIN LEMAIRE Stéphanie : p.22, 67, 68  
TURRETTINI Emily : p.42  
VIEUX Rémy : p.35  
W3C.ORG : p.26  
WANG Jiangning, JI Liqiang, LIANG Aiping, YUAN Decheng : p.36  
YOON JungWon : p.37, 38  
ZDNET FRANCE : p.44



# **Annexes**

## Annexe 1 Matrice du questionnaire

### ***Comprendre l'activité de l'interlocuteur***

- > Pouvez-vous donner votre intitulé de poste ?
  
- > Pouvez-vous décrire votre fonction ?
  
- > Quels sont vos rôles dans la chaîne d'activités ? Quel est votre rôle spécifique dans votre service ?
  
- > Quels sont les liens de votre service avec les autres ? Liens verticaux (hiérarchiques) et horizontaux ?

### **Analyser l'existant**

#### ***Cerner le rapport à l'information***

Selon vous, dans votre activité quotidienne, il vous arrive le plus souvent :

○ 1. de chercher de l'information ? ou de vous informer ?	a. L'information vous arrive, elle vous est fournie automatiquement : revues, rapports, notes, fils RSS...
	b. Vous faites appel à des circuits formels ou informels : les collègues à la machine à café, les collègues dans le service, dans les autres services ; en externe ; des outils comme Internet, l'intranet, etc.
○ 2. de produire de l'information ? ou d'informer les autres	a. Dans quel but ?
	b. Sous quelle forme ? (papier, dématérialisée)
	c. Qui en est destinataire ?
	d. Êtes-vous prêt à partager ce que vous avez fait ? La diffusion peut-elle être élargie ?
	e. Savez-vous qui serait intéressé ? Savez-vous comment vous pourriez le savoir ? (quelle est leur conscience de leur environnement)
○ 3. de gérer de l'information ? ou de classer, ranger, archiver...	a. Classement, rangement, organisation du bureau électronique, nommage des dossiers ?
	b. Est-ce nécessaire pour réaliser vos tâches ou complémentaire à l'organisation du travail ?

#### ***Définir le Système d'information personnel (SIP)***

Les moyens utilisés

○ 1. Avez-vous recours à des abonnements ?	a. à la presse magazine
	b. à la presse professionnelle, spécialisée
	c. papier et/ou électronique
	d. à des newsletters, blogs, sites

○ 2. Avez-vous recours à la pige presse ?	a. Si oui, qui, comment ?
	b. Sinon, pourquoi ?
○ 3a. Connaissez-vous les fils RSS ?	Non ou Si oui 3b
○ 3b. Êtes-vous abonné à des fils RSS ?	a. Si oui, lesquels ?
	b. Si oui, avec quel outil (Netvibes, Google, Yahoo, etc.) ?
	c. Sinon, pourquoi ?
	d. Avez-vous vu la fonctionnalité de fils RSS dans la messagerie ?
○ 4. Quelles sont vos sources principales d'informations ?	a. Humaines
	b. Matérielles (documents)
	c. Immatérielles : messagerie, intranets métiers ; autres

« Une pige presse est un dispositif de recueil permanent des investissements publicitaires réalisés sur le média presse. Les piges publicitaires presse sont généralement réalisées manuellement et quotidiennement. Certaines piges donnent accès à chaque publicité qui est recensée et scannée puis accessible pour l'abonné sur Internet. Dans le contexte français près de 1000 titres de presse sont "épluchés" par la pige TNS » <<http://www.definitions-marketing.com/Definition-Pige-presse>>

« Une pige plurimédia est une pige publicitaire qui couvre l'ensemble des média publicitaires (TV, Presse, Affichage, Cinéma, Internet, Radio) » <<http://www.definitions-marketing.com/Definition-Pige-plurimedia>>

Définitions prévues en cas de demande

### La quantité d'informations

○ 1. Le flux d'informations	a. qui vous parvient est-il volumineux ?
	b. La part à conserver est-elle volumineuse ou pouvez-vous la trier ?
○ 2. La quantité d'informations	a. dont vous avez besoin quotidiennement est-elle volumineuse ?
	b. La part à conserver est-elle volumineuse ou pouvez-vous la trier ?

### La production d'informations

○ 1. Produisez-vous de l'information ? (cf. réponses précédentes)	a. Dans quel but ?
	b. Sous quelle forme ? (papier, dématérialisée)
	c. Qui en est destinataire ?
	d. Êtes-vous prêt à partager ce que vous avez fait ? La diffusion peut-elle être élargie ?
	e. Savez-vous qui serait intéressé ? Savez-vous comment vous pourriez le savoir ? (quelle est la conscience de leur environnement)
○ 2. Transformez-vous l'information qui vous parvient ou que vous avez trouvée ?	a. Réalisation de notes, de synthèses, de dossiers... ?
	b. Partage avec des collègues intéressés par les mêmes sujets et si oui comment ?

## Le classement, l'archivage

○ 1. Classez-vous l'information qui vous parvient ou que vous trouvez ?	a. Si oui : Pourquoi ? Comment ? : dans la messagerie ; en dossier, sous-dossiers ; recours à des outils collectifs ; par ordre alpha, chrono...
	b. Sinon : Pourquoi ?
	c. Comment retrouvez-vous l'information dont vous avez besoin ? (recherche par mots-clefs, moteur de recherche...)
○ 2. Classez-vous vos mails ?	a. Si oui : Comment ?
	b. Sinon : Pourquoi ?
	c. Comment faites-vous pour retrouver un mail, une information ou un document contenu dans un mail (quelle recherche) ?
○ 3. Quand vous décidez de conserver une information, ce choix s'effectue en fonction :	a. de l'utilité de l'information ?
	b. du potentiel de l'information ? (non pertinente aujourd'hui mais peut-être plus tard) ?
	c. de la qualité de la source ?
○ 4. Quand vous conservez une information, déterminez-vous un délai ?	a. Oui : le temps de son utilisation et vous jetez ?
	b. Oui : le temps de son utilisation et vous classez ?
	c. Non : vous trie quand vous avez le temps, la mémoire est pleine... ?
○ 5. Comment retrouvez-vous les informations que vous détenez ?	a. Par arborescence : tout est classé en dossier, sous-dossiers... ?
	b. Avec un moteur de recherche, d'une fonction "Rechercher"... ?
	c. Tout est sur papier donc dans des dossiers, sous-chemises, dans les armoires ?

### ***Définir le Système d'information collectif (SIC)***

○ 1. Partagez-vous l'information collectée ?	a. Si oui : Avec qui ? Comment ? (quels outils, quelles méthodes)
	b. Sinon : Pourquoi ?
○ 2. Quels outils collectifs utilisez-vous pour collecter, produire et/ou classer l'information ? (à proposer si la personne n'a pas d'idée)	a. Outils de travail :
	aa. bdd : si oui, la(es)quelle(s) (= compris comme dossier sur serveur ou fichier Excel)
	ab. Internet
	ac. intranet
	ad. workflow (informatise un processus)
	ae. ERP (système de gestion intégré)
	af. Ged (service commun avec ou non des espaces privés)
	b. Outils d'information :
	ba. mémoire / stockage, preuve
	bb. archives, records management, gestion des connaissances
bc. filtrage, exploitation	
bd. documentation, veille	

	c. Mixte :
	ca. travail coopératif
	cb. travail à distance
	cc. intranet

Workflow : suite de tâches ou opérations effectuées par une personne, un groupe de personnes, un organisme ; il suppose des circuits et modes de validation, la répartition des tâches, la définition de délais. <<http://fr.wikipedia.org/wiki/Workflow>>

ERP (Enterprise resource planning) : progiciel de gestion intégré : intègre les principales composantes fonctionnelles de l'entreprise : gestion de production, gestion commerciale, logistique, RH, comptabilité, contrôle de gestion, paie ; induit un système unifié où les utilisateurs de différents métiers travaillent dans un environnement applicatif identique qui repose sur une bdd unique (intégrité des données, non redondance de l'information, réduction des temps de traitement). <[http://fr.wikipedia.org/wiki/Progiciel\\_de\\_gestion\\_int%C3%A9gr%C3%A9](http://fr.wikipedia.org/wiki/Progiciel_de_gestion_int%C3%A9gr%C3%A9)>

Ged : procédé informatisé pour organiser et gérer des informations et des documents électroniques dans une organisation (système d'acquisition, de classement, stockage, indexation, consultation) ; elle participe aux processus de travail collaboratif, de capitalisation et d'échanges d'informations. <[http://fr.wikipedia.org/wiki/Gestion\\_%C3%A9lectronique\\_des\\_documents](http://fr.wikipedia.org/wiki/Gestion_%C3%A9lectronique_des_documents)>

Définitions prévues en cas de demande

### **Pour conclure**

- > Ce dispositif d'informations est-il satisfaisant ?
  - . Avez-vous les informations dont vous avez besoin pour votre activité ?
  - . Si vous ne les avez pas, les trouvez-vous facilement par vous-même ?
  - . Si oui, comment ? Sinon, pourquoi ?
- > Qu'est-ce qui vous bloque, vous pose problème, vous ralentit ?

### **Analyser les besoins pour la mise en place d'une veille concurrentielle de la communication visuelle dans le secteur du luxe**

#### ***Des solutions déjà trouvées***

o 1. Avez-vous déjà mis en place des "solutions" pour réaliser une veille concurrentielle d'images ?	a. Oui Non
	b. Si oui : Lesquelles ?
o 2. Savez-vous si un collègue a mené cette réflexion ?	a. Oui Non
	b. Si oui : Qui ?
o 3. Savez-vous qui ou	a. Oui Non

quel(s) service(s) seraient intéressés par cette veille d'images ?	b. Si oui : Qui ?
--	-------------------

### **Les attentes**

Idéalement,

o 1. Comment aimeriez-vous obtenir les informations nécessaires à votre activité ?	
o 2. Ces informations (en matière de veille d'images) seraient :	a. personnalisées à vos besoins propres
	b. généralistes
o 3. Comment souhaiteriez-vous utiliser ces informations ?	a. Les lire à l'écran, sur papier ?
	b. Les classer dans votre ordinateur ? Si oui, où ?
	c. Les avoir à disposition rapidement ?
	d. Les partager avec d'autres collègues, interagir avec eux sur le contenu, échanger des idées en temps réel ?

### **Les besoins**

Au quotidien, dans le cadre d'une veille concurrentielle,

o 1. Quelle est la nature des informations dont vous avez besoin ?	a. Des informations visuelles : photos, dessins, vidéos ?
	b. Des informations stratégiques : rapports, analyses, études, etc. sur les campagnes des concurrents, sur les tendances à venir ?
	c. Le contexte : qui (concurrent) communique quoi (produit) quand (date, période) où (dans / sur quel support) ?
o 2. Vous préférez :	a. recevoir une information définie selon vos besoins : d'après des critères enregistrés, des paramètres donc pouvant laisser passer certaines informations ?
	b. chercher par vous-même ce dont vous avez besoin ?
o 3. Quelle serait la manière la plus simple de consulter les informations délivrées ?	a. Les concentrer en une seule application :
	aa. via une plateforme connue : messagerie, intranets métiers...
	ab. ou une autre de votre service
	ac. ou une nouvelle différenciée
	b. Proposer des accès différenciés selon la nature de l'information : des fils RSS, des pages médias en ligne, une base de données pour les images

○ 4. Pour savoir quand consulter cette veille, le plus simple serait :	a. de définir une périodicité de mise à jour (un jour sur deux, une fois par semaine, etc.) ?
	b. de recevoir un message d'alerte quand une nouvelle information est ajoutée à la veille ?
	c. aucun dispositif, vous vous y rendez quand vous aurez le temps, quand vous en aurez besoin, au fur et à mesure
○ 5. Pensez-vous qu'il est plus simple :	a. qu'une seule personne soit chargée d'alimenter cette veille ?
	b. que chacun puisse y ajouter ses recherches, ses remarques, etc. ?
○ 6. Pouvez-vous donner des exemples concrets d'informations dont vous avez besoin, le type de requêtes que vous exécutez, les questions que vous posez ?	

### ***Les freins éventuels***

○ 1. Pour des raisons de confidentialité, est-il envisageable de travailler avec :	a. un prestataire extérieur ?
	b. un service en ligne (Saas : Software as a Service) ?
	c. en cloud computing ?
○ 2. Une solution de veille concurrentielle, nécessitant une courte formation, est-elle envisageable ?	
○ 3. S'il existe des outils dédiés à la veille et à la recherche d'informations, seriez-vous intéressé pour qu'on vous les présente ?	
○ 4. Pensez-vous qu'il est possible d'ouvrir cette veille concurrentielle ?	a. À tout le monde en interne ?
	b. À certaines personnes avec des accès restreints de lecture, de modifications...

Cloud computing : déporter sur des serveurs distants des stockages et des traitements informatiques traditionnellement localisés sur des serveurs locaux ou sur le poste de l'utilisateur : les utilisateurs ou les entreprises ne sont plus gérants de leurs serveurs informatiques mais peuvent accéder de manière évolutive à de nombreux services en ligne sans gérer l'infrastructure sous-jacente : les applications et les données ne sont plus sur l'ordinateur local mais dans un « nuage » composé de serveurs distants interconnectés. <[http://fr.wikipedia.org/wiki/Cloud\\_computing](http://fr.wikipedia.org/wiki/Cloud_computing)>

} Définition prévue en cas de demande

### **Ouverture**

> Avez-vous des remarques, des suggestions, des propositions... ?

## Annexe 2 Exemples de visuels du luxe en Parfums-Beauté



Campagne Homme, Guerlain 2008

Renne Castrucci (mannequin)

600 x 400 px, 27 ko, jpg

Téléchargée le 07.10.2012

<<http://tinyurl.com/9b7akdv>>



Campagne Rouge Dior, Dior 2010

Monica Bellucci (égérie)

510 x 382 px, 27,2 ko, jpg

Téléchargée le 07.10.2012

<<http://tinyurl.com/9ceauo2>>



Campagne Sublimage la crème Chanel 2011

700 x 360 px, 146 ko, jpg

Téléchargée le 07.10.2012

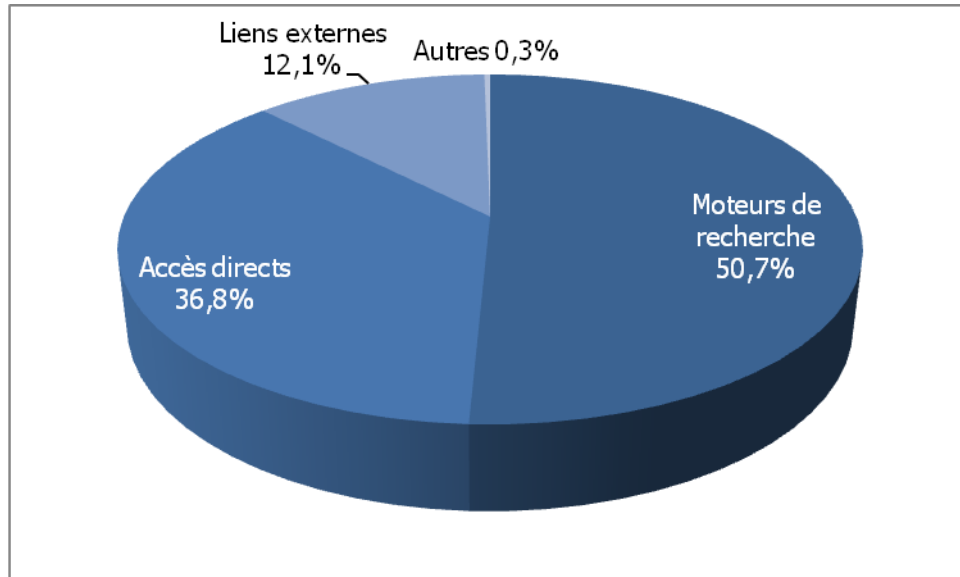
<<http://tinyurl.com/9jjoemq>>

En raison de la confidentialité demandée par l'entreprise X, les noms des photographes ne sont pas mentionnés pour les campagnes publicitaires.



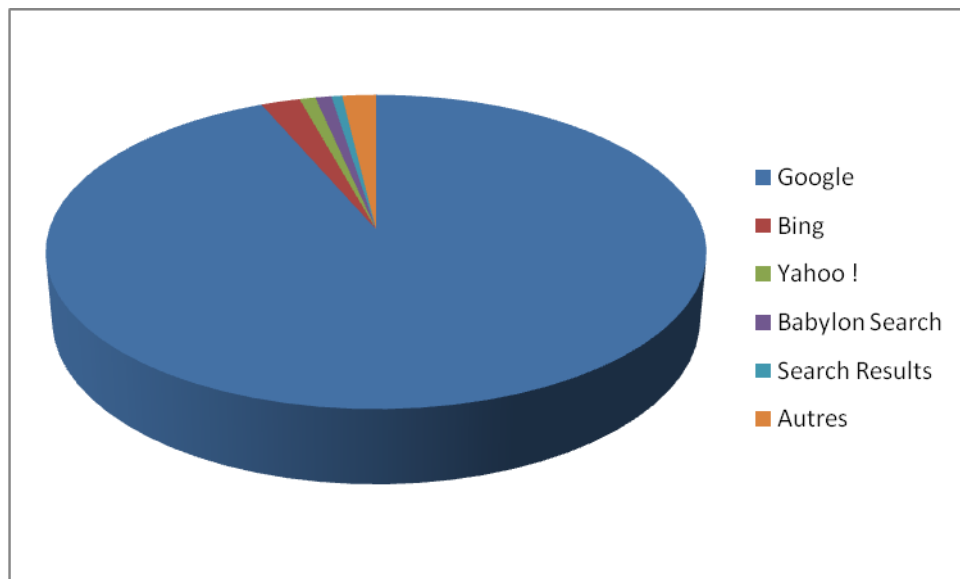
## Annexe 3 Accès aux sites Internet en France

### Annexe 3a – Répartition selon le type d'accès



[27, Médiamétrie eStat, p.1]

### Annexe 3b – Répartition selon le moteur de recherche



[27, Médiamétrie eStat, p.1]

## Annexe 4 Exemples d'images problématiques pour la reconnaissance faciale



Photo d'Inez van Lamsweerde et Vinoodh Matadin  
Campagne Louis Vuitton 1997

437 x 276 px, 26,3 ko, jpg

Téléchargée le 30.09.2012

<<http://tinyurl.com/9v9uxd9>>



Photo de Richard Burbridge  
Gisèle Bündchen, mars 2000

520 x 662 px, 107 ko, jpg

Téléchargée le 30.09.2012

<<http://photodonuts.com/richard-burbridge>>



Photo de Craig McDean

Coco Rocha, octobre 2006

800 x 1026 px, 117 ko, jpg

Téléchargée le 30.09.2012

<<http://tinyurl.com/9adurft>>





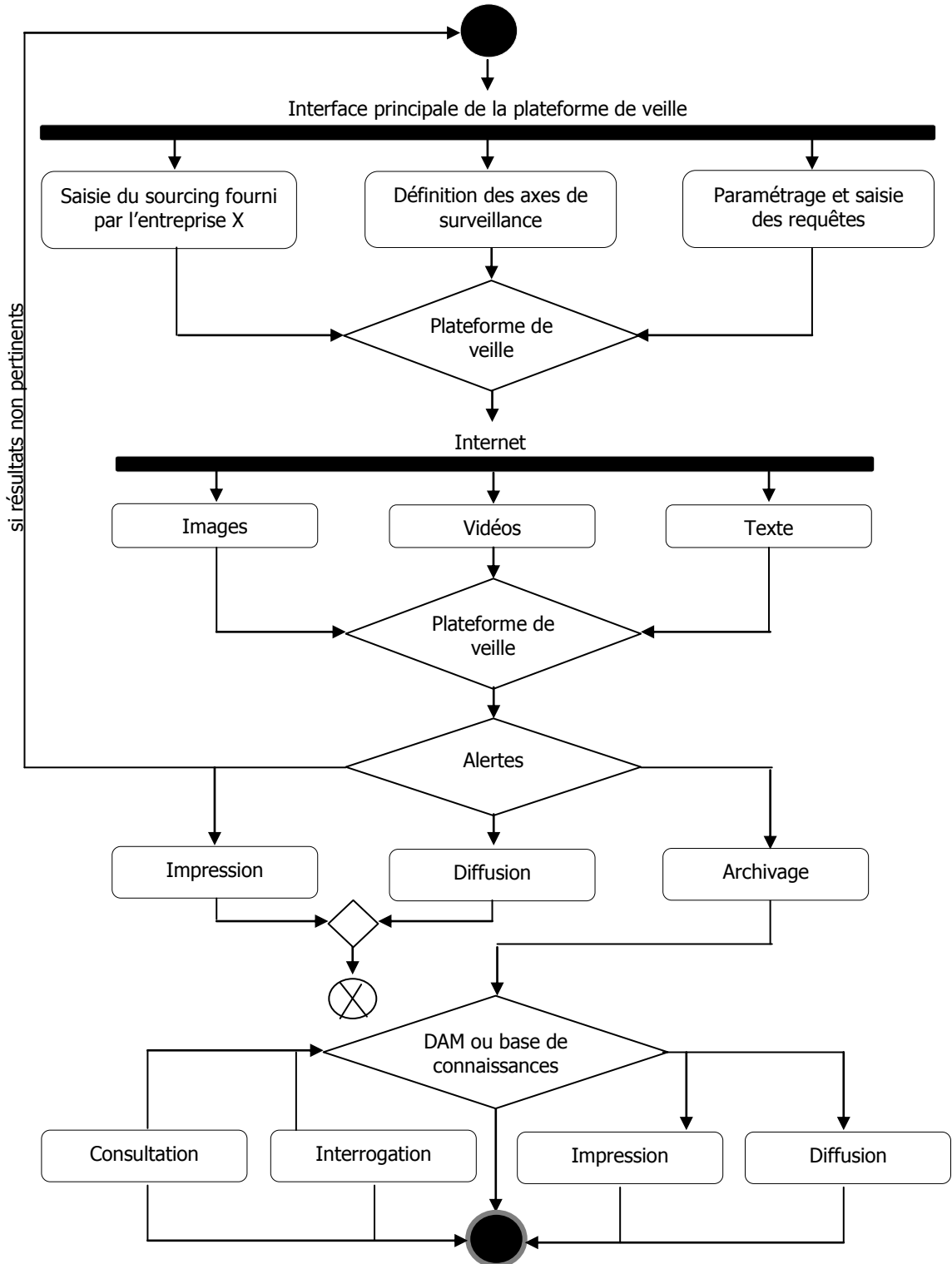
Photo de Sølve Sundsbø  
Edita Vilkeviciute, mai 2008  
480 x 639 px, 57 ko, jpg  
Téléchargée le 30.09.2012  
<<http://tinyurl.com/9skh3h7>>



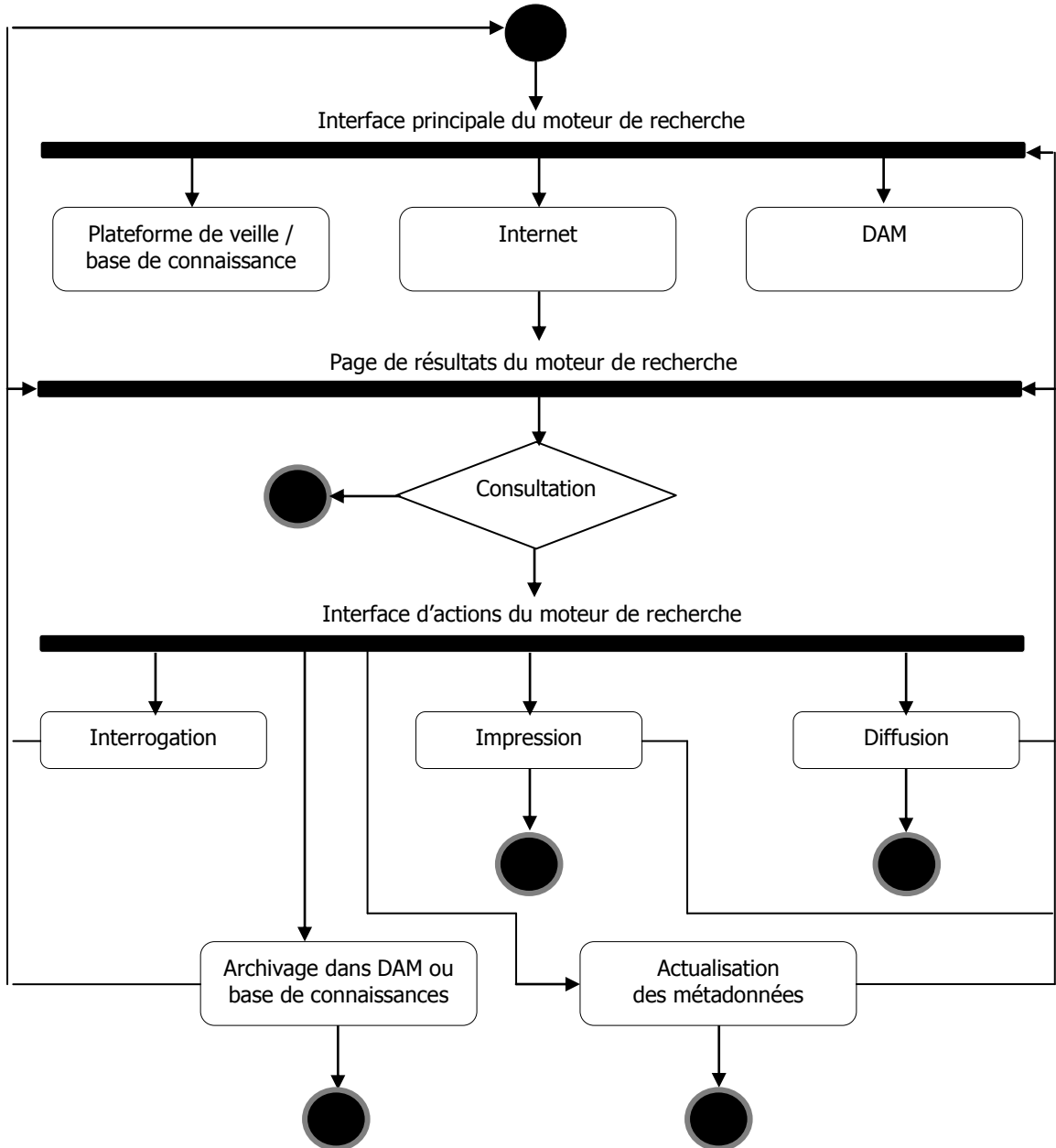
Photo d'Alexander Khokhlov  
Icône Wifi, 2012  
600 x 903 px, 94,5 ko, jpg  
Téléchargée le 30.09.2012  
<<http://tinyurl.com/9pchz2z>>

# Annexe 5 Schémas des processus de veille et recherche selon le diagramme d'activités d'UML (Unified Modeling Language)

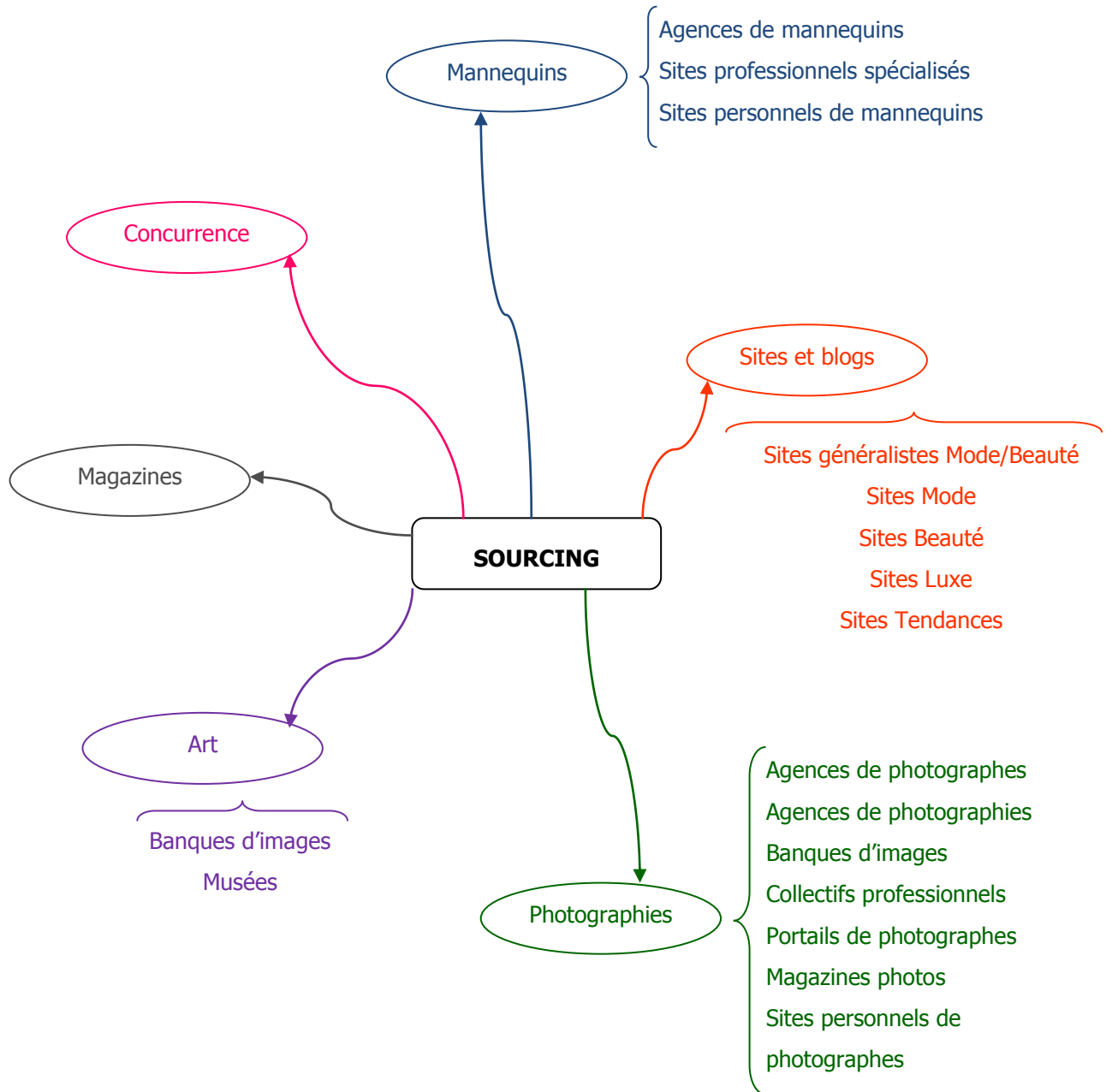
## Annexe 5a – Processus de veille



## Annexe 5b – Processus de recherche



## Annexe 6 Sourcing



# Annexe 7 Exemples d'agrégateurs de flux offrant une solution de diffusion

## Annexe 7a – Feed2PDF

<<http://feed2pdf.appspot.com/>>

Réalisé le 07.08.2012 à partir des flux du blog Armocromia <<http://www.armocromia.com/>>

collectés dans l'agrégateur Google Reader les 30 et 31.07.2012

Lancôme (3 pages) ; Estée Lauder (1 page) ; Louis Vuitton (1 page).

armocromia.com 31 July 2012

### Lancôme Midnight Roses Fall 2012



**Blush Highlighter** - 001 Moonlight Rose - Limited Edition - crafted with a new technology that combines a high-concentrated emulsion with an illuminating effect, this extremely fine, light-infused powder is adorned with an embossed rose.

Inspired by the atmosphere of Parisian nights and the mystery and magic of the Parisian cabaret scene, it's a both sophisticated and feminine collection.

**Le Regard Pro** - Limited Edition - with three universal shades of brows for shaping and defining the brows as well as a highlighter to apply just below.


Love this [joliprint PDF?](#) [Add it to your Reading List!](#) → [joliprint.com/mag](#)

Page 1

armocromia.com 31 July 2012

### Lancôme Midnight Roses Fall 2012

the arch for added depth. Also included is a brush and a pair of tweezers for perfect brow grooming.



**L'Absolu Rouge Mat Lipstick** - 100 Rouge Figueur and 901 Rose Sorbonne - Limited Edition - in a new matt velvet finish.

**Hypnose Brown Shaper** - Limited Edition - a transparent gel to structure the brows that can be used alone for a natural effect or combined with Le Regard Pro to ensure makeup lasts.

**Le Ceylan Khol** - 014 Prune Crépuscule - bewitching purple - Limited Edition

Love this [joliprint PDF?](#) [Add it to your Reading List!](#) → [joliprint.com/mag](#)

Page 2

armocromia.com 31 July 2012

### Lancôme Midnight Roses Fall 2012



**L'Absolu Creme de Mat Lip Gloss** - 100 Pouton D'Amour, 901 Magie Pourpre and 901 Pure Eclair - Limited Edition - the smoothness of a gloss, the coverage of a lipstick and the comfort of a cream texture with a matte finish.

**Versus In Love** - 433 Infusion de Prune - intense purple and 030 Givre Magique - transforming top coat that turns matte velvet matte - Limited Edition - Fantastic and very a fall trend, do you love it?!

**Color Design 5 Shadow and Liner Palette** - Violet Magnétique - Limited Edition -

Love this [joliprint PDF?](#) [Add it to your Reading List!](#) → [joliprint.com/mag](#)

Page 3

armocromia.com 30 July 2012

### NEW PURE COLOR VIVID SHINE Estee Lauder



Color as vivid, it's luminous. Light-reflecting prism pearls create multidimensional effects. For amplified color with unprecedented staying power. Clusters of up to eight different shaped pearls create diamond-like refractions. Silver and gold sparkle pearls enliven lip tones. All shades look great on every skin tone. Feel so good. Lightweight, creamy, conditioning.

From fall 2012 (September) a new line of 16 vivid lipsticks, featuring fresh colors, crystalline and gold particles inspired by the city of Paris. You can find more info visiting my others post about the new collection: [click here](#).

Love this [joliprint PDF?](#) [Add it to your Reading List!](#) → [joliprint.com/mag](#)

Page 1

armocromia.com 29 July 2012

### LOUIS VUITTON HAUTE COUTURE 2012-13

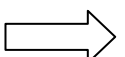


Amazing this fashion collection for fall 2012. It's a big one, big beautiful bags and coats, amazing 21 century style.

Do you like Louis Vuitton and its chic bags? For me these are the best!

Love this [joliprint PDF?](#) [Add it to your Reading List!](#) → [joliprint.com/mag](#)

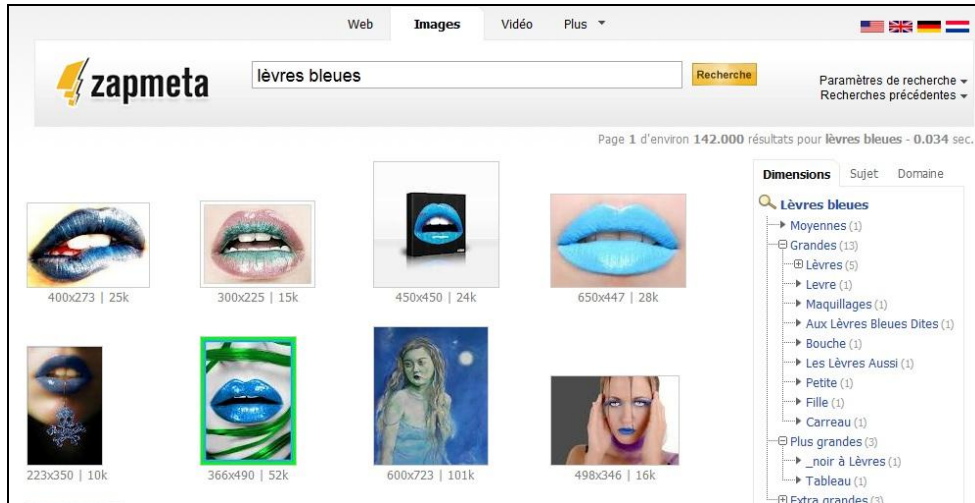
Page 1







## Annexe 8 Moteurs de recherche



ZapMeta [consulté le 12.09.2012] : recherche sur l'expression « lèvres bleues », avec une présentation des résultats sous forme de clusterisation à droite.  
 <<http://www.zapmeta.fr/?sttname=image>>

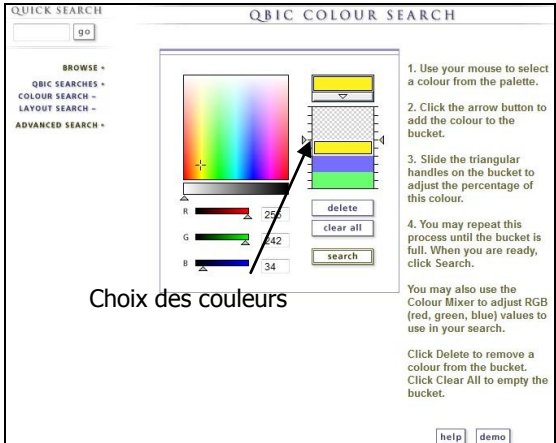


Chromatik [consulté le 12.09.2012] : recherche avec les couleurs sélectionnées dans la palette à gauche.  
 <<http://chromatik.labs.example.com/>>



TinEye [consulté le 12.09.2012] : recherche avec les couleurs sélectionnées dans la palette à droite.  
 <<http://labs.tineye.com/multicolor>>

## Annexe 9 Musée de l'Hermitage



**QBIC COLOUR SEARCH**

1. Use your mouse to select a colour from the palette.  
 2. Click the arrow button to add the colour to the bucket.  
 3. Slide the triangular handles on the bucket to adjust the percentage of this colour.  
 4. You may repeat this process until the bucket is full. When you are ready, click Search.

You may also use the Colour Mixer to adjust RGB (red, green, blue) values to use in your search.

Click Delete to remove a colour from the bucket. Click Clear All to empty the bucket.

Choix des couleurs

1) Stained Glass Panel: Execution a Righteous Man (11c)  
 UNKNOWN  
 Late 14th century (?)

2) View of Murnau  
 Kandinsky, Wassily 1908

3) Boats at Saintes-Maries  
 Van Gogh, Vincent 1888

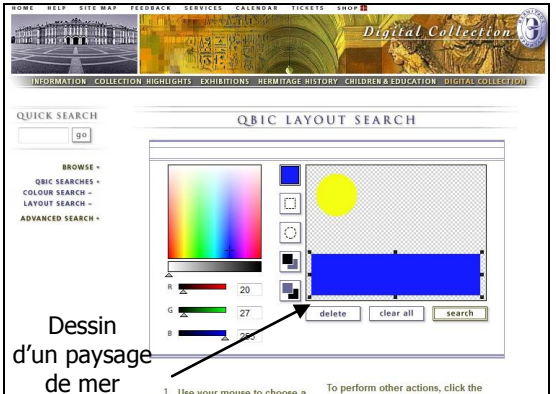
4) Stained Glass Panel: Antichrist Turning Stones into Bread (7b)  
 UNKNOWN  
 Late 14th century (?)

5) Portrait of Lydia Delectorskaya  
 Matisse, Henri 1947

6) Stained Glass Panel: Noah's Ark (7a)  
 UNKNOWN

Recherche par la couleur [consulté le 12.09.2012] :

<<http://www.hermitagemuseum.org/cgi-bin/db2www/qbicColor.mac/qbic?selLang=English>>



**QBIC LAYOUT SEARCH**

Dessin d'un paysage de mer

1) Flood  
 Guillaou, Charles 1893

2) Waterloo Bridge. Effect of Fog  
 Monet, Claude 1903

3) Panel 7. Cupid Carrying Psyche Up to Heaven  
 Denis, Maurice 1908

4) Poppy Field  
 Monet, Claude Circa 1890

5) View of the Versailles Palace from the Park  
 Perelle, Gabriel Mid-17th century

6) Island of Fishermen on Lake Maggiore  
 Premazzi, Luigi Second half of the 19th century

Recherche par la mise en forme [consulté le 12.09.2012] :

<<http://www.hermitagemuseum.org/cgi-bin/db2www/qbicLayout.mac/qbic?selLang=English>>