

LIVRE BLANC

LES MESURES DE FILTRAGE DE CONTENUS

14 avril 2008



TABLE DES MATIERES

1. QUELLES RESPONSABILITES POUR LES PRESTATAIRES DE STOCKAGE ?	7
1.1 LA DIRECTIVE SUR LE COMMERCE ELECTRONIQUE DU 8 JUIN 2000	7
1.2 LA LOI POUR LA CONFIANCE DANS L'ECONOMIE NUMRIQUE DU 21 JUIN 2004	7
1.2.1 L'absence d'obligation générale de surveiller les informations stockées et de rechercher des faits ou des circonstances révélant des activités illicites	9
1.2.2 Le caractère « manifestement illicite » des informations dénoncées	10
1.2.3 Le mécanisme dit de présomption par notification	10
1.3 LA JURISPRUDENCE SUR LA RESPONSABILITE DES HEBERGEURS	11
1.3.1 La jurisprudence de principe	11
1.3.2 Une jurisprudence contraire : le cumul des qualités d'hébergeur et d'éditeur	12
1.3.3 Le revirement de cette jurisprudence contraire : une définition de l'hébergement fondée sur la fonction exercée, à savoir le stockage de données à la demande du destinataire du service	13
1.3.4 Les obligations prétoriennes des hébergeurs	14
2. L'ETAT DE L'ART EN MATIERE DE FILTRAGE DE CONTENUS	19
2.1 LES ACTEURS DU FILTRAGE DE CONTENUS	21
2.1.1 Les principaux utilisateurs de logiciels de filtrage de contenus	22
2.1.2 Les principaux éditeurs, producteurs et intégrateurs de logiciels de filtrage	26
2.1.3 Quelques titulaires de droits	36
2.1.4 Quelques acteurs de la recherche-développement identifiés en matière de filtrage de contenus	37
2.2 LES PRINCIPAUX OUTILS DE FILTRAGE	44
2.2.1 Les principaux outils commerciaux de filtrage de contenus	44
2.2.2 Les principaux brevets en matière de filtrage de contenus	56
2.3 LES PRINCIPAUX CONTRATS FAISANT REFERENCE AU FILTRAGE DE CONTENUS	57
2.3.1 Les conditions générales d'utilisation de Yahoo	57
2.3.2 Les conditions générales d'utilisation de Lycos	58
2.3.3 Les conditions générales d'utilisation de YouTube (Google)	59
2.3.4 Les conditions générales d'utilisation de Gmail (Google)	59
2.3.5 Les conditions générales d'utilisation de MySpace	59
3. LA PRATIQUE D'EBAY	61
3.1 PRESENTATION DES PROGRAMMES PROACTIFS DE RECHERCHE D'ANNONCES MANIFESTEMENT ILLICITES MIS EN PLACE PAR EBAY	61
3.1.1 L'outil de recherche par mots-clés	61
3.1.2 Les mesures de limitation	63
3.2 LE CADRE JURIDIQUE DANS LEQUEL LES MOYENS DE LUTTE CONTRE LA CONTREFAÇON SONT MIS EN ŒUVRE PAR EBAY	64
3.2.1 L'absence d'obligation générale de surveillance et de recherche de faits ou de circonstances révélant des activités illicites	64
3.2.2 La loi du Bon Samaritain	64
4. ANNEXES	67

AVANT-PROPOS

A l'initiative de Monsieur Alexandre Menais, directeur juridique d'eBay, il a été créé un groupe de travail sur le cadre juridique et technique de la mise en œuvre de mesures de filtrage de contenus par les prestataires de services sur internet agissant en qualité d'intermédiaire.

Le Livre blanc issu de ces travaux, se veut être une base de réflexion et d'échanges élargie, permettant d'aboutir à terme à la rédaction de documents de référence en matière de filtrage de contenus (code de bonne conduite, norme...).

Ont participé au groupe de travail : eBay, Alain Bensoussan, directeur du département Technologies émergentes du cabinet Alain Bensoussan-Avocats et Pierre Saurel, directeur du département Expertise technique, du cabinet Alain Bensoussan-Avocats.

La méthodologie suivie par le groupe de travail est la suivante :

- analyse du cadre juridique :
 - o cadre juridique actuel (cadres légal et jurisprudentiel) dans lequel les prestataires de services sur internet agissent en qualité d'intermédiaire en France ;
 - o cadre juridique dans lequel eBay met en œuvre des mesures de filtrage de contenus et présentation de ces mesures ;
- analyse du cadre technique :
 - o identification des principaux acteurs intervenant en matière de filtrage de contenus (utilisateurs de logiciels de filtrage de contenus, éditeurs, producteurs et intégrateurs de logiciels de filtrage de contenus, titulaires de droits, acteurs français de la recherche-développement en matière de filtrage de contenus) ;
 - o identification des principaux articles et travaux de recherche publiés en

- matière de filtrage (bibliographie et webographie commentées) ;
- identification des principaux outils de filtrage de contenus texte ou multimédia (images, vidéos, sons) en termes de logiciels et de brevets ;
 - identification des principaux contrats faisant référence au filtrage de contenus.

INTRODUCTION

Aux termes de l'article 6-I 7° de la loi pour la confiance dans l'économie numérique (dite LCEN), les hébergeurs ne sont pas tenus d'une obligation générale de surveiller les informations qu'ils stockent ni d'une obligation générale de rechercher des faits ou des circonstances révélant des activités illicites.

La jurisprudence tend, au vu des dernières décisions, à appliquer aux hébergeurs de services susceptibles de contenir de « manière régulière » des contenus illicites, une obligation particulière de surveillance des contenus qu'ils stockent dès lors qu'ils ont été notifiés, et ce sans prévoir l'étendue de cette obligation.

eBay, à l'instar d'autres plates-formes de l'internet, a volontairement mis en place des outils pour lutter contre un usage illicite de ses sites internet, et notamment des mesures de filtrage de contenus.

eBay considère agir en « Bon Samaritain » lorsqu'elle met en œuvre des moyens de lutte contre la fraude notamment contre la contrefaçon sur ses sites internet alors qu'elle n'a aucune obligation de le faire.

Des premiers débats autour des mesures à mettre en place pour lutter contre les comportements illicites ont déjà lieu, et vont continuer, en France et en Europe. Il a paru utile de porter cette réflexion et de poursuivre ce travail par la publication d'un Livre blanc sur le cadre juridique et technique de la mise en œuvre de mesures de filtrage de contenus par les hébergeurs.

Ce Livre blanc est une première version et ne se prétend pas exhaustif. Il a pour vocation d'être enrichi par tous les acteurs qui sont parties prenantes dans la lutte contre les comportements illicites sur l'internet. L'objectif est d'aboutir, avec le concours de tous, à l'écriture d'un document à visée pédagogique sous forme de conseils, de recommandations et d'un exposé des meilleures pratiques.

Cette première version aborde, dans une première partie, un rappel du droit positif s'agissant des responsabilités des prestataires de stockage en droit français avec, notamment, un récapitulatif des diverses jurisprudences en ce domaine.

En deuxième partie, nous avons essayé de proposer un état de l'art en matière de filtrage, en nous basant principalement sur des données publiques relatant les pratiques des différents acteurs de l'internet, tout en opérant un inventaire des différents outils qui peuvent être mis à disposition dans ce domaine. Cet état ne se veut pas exhaustif et peut être déjà caduque mais constitue une première photographie des pratiques qui se complètera au gré des versions. Le style des références est volontairement cursif pour ne pas alourdir le Livre blanc.

Enfin, une troisième partie fait part de la pratique d'eBay que nous exposons déjà depuis plusieurs mois et que nous souhaitons confirmer plus encore.

Des annexes figurent dans ce Livre blanc ; elles ont été utiles aux rédacteurs pour l'élaboration du Livre blanc.

Il est à noter que nous ne trouvons pas dans cette première version du Livre blanc, de recommandations sur le filtrage. Le groupe de travail qui pourrait se renforcer et se compléter autour de ce premier référentiel pourra être amené à faire des propositions.

Par ailleurs, l'Association française de normalisation (Afnor) a lancé un groupe de réflexion autour de la rédaction d'une norme de spécifications destinée, notamment à préciser l'état de l'art en matière de filtrage de contenus, à fournir les règles, lignes directrices et caractéristiques des mesures de filtrage de contenus et de leurs résultats, et à indiquer le seuil de tolérance.

Dans le même temps, des travaux ont été lancés en ce domaine autour du rapport Olivennes.

Il est probable que ces initiatives se retrouveront le moment venu et permettront ainsi de fixer, sans alourdir le poids des textes législatifs et réglementaires, une solution durable et pragmatique répondant aux spécificités de chacun.

Alexandre Menais
eBay France
Directeur Juridique

1. QUELLES RESPONSABILITES POUR LES PRESTATAIRES DE STOCKAGE ?

1.1 LA DIRECTIVE SUR LE COMMERCE ELECTRONIQUE DU 8 JUIN 2000

L'article 14 de la directive sur le commerce électronique¹ dispose que « les Etats membres veillent à ce que, en cas de fourniture d'un service de la société de l'information consistant à stocker des informations fournies par un destinataire du service, le prestataire ne soit pas responsable des informations stockées à la demande d'un destinataire du service à condition : a) que le prestataire n'ait pas effectivement connaissance de l'activité ou de l'information illicite (...) ou b) que le prestataire, dès le moment où il a de telles connaissances, agisse promptement pour retirer les informations ou rendre l'accès à celles-ci impossible ».

L'article 15 de la directive énonce, quant à lui, que « les Etats membres ne doivent pas imposer aux prestataires, pour la fourniture des services visés aux articles 12, 13 et 14, une obligation générale de surveiller les informations qu'ils transmettent ou stockent, ou une obligation générale de rechercher activement des faits ou des circonstances révélant des activités illicites ».

L'objectif du législateur européen est double : d'une part, le développement du commerce électronique dans la société de l'information² et, d'autre part, l'harmonisation des législations et des jurisprudences des Etats membres en matière de responsabilité des prestataires de services agissant en qualité d'intermédiaire³.

1.2 LA LOI POUR LA CONFIANCE DANS L'ECONOMIE NUMRIQUE DU 21 JUIN 2004

Les articles 14 et 15 de la directive sur le commerce électronique ont été transposés à l'article 6 de la loi pour la confiance dans l'économie numérique⁴.

¹ Directive 2000/31/CE du 8 juin 2000 relative à certains aspects juridiques de services de la société de l'information et notamment du commerce électronique, dans le marché intérieur, dite « directive sur le commerce électronique ».

² Considérant 2 de la directive sur le commerce électronique : « Le développement du commerce électronique dans la société de l'information offre des opportunités importantes pour l'emploi dans la Communauté, en particulier dans les petites et moyennes entreprises. Il facilitera la croissance économiques des entreprises européennes ainsi que leurs investissements dans l'innovation, et il peut également renforcer la compétitivité des entreprises européennes, pour autant que tout le monde puisse accéder à l'Internet ».

³ Considérant 40 de la directive sur le commerce électronique : « Les divergences existantes et émergentes entre les législations et les jurisprudences des Etats membres dans le domaine de la responsabilité des prestataires de services agissant en qualité d'intermédiaires empêchent le bon fonctionnement du marché intérieur, en particulier en gênant le développement des services transfrontaliers et en produisant des distorsions de concurrence (...) ».

⁴ Loi n°2004-575 pour la confiance dans l'économie numérique du 21 juin 2004.

La loi pour la confiance dans l'économie numérique prévoit, à son article 6-I 2°, que la responsabilité des hébergeurs ne peut être engagée¹ du fait des activités ou des informations stockées à la demande d'un destinataire de ces services :

- que s'ils avaient effectivement connaissance de leur caractère illicite ou de faits et circonstances faisant apparaître ce caractère ;
- ou si, dès le moment où ils en ont eu connaissance, ils n'ont pas agi promptement pour retirer ces informations ou en rendre l'accès impossible.

L'article 6-I 7° de la loi pour la confiance dans l'économie numérique précise en outre que les hébergeurs ne sont pas tenus d'une obligation générale de surveiller les informations qu'ils stockent, ni même d'une obligation générale de rechercher des faits ou des circonstances révélant des activités illicites.

Ce régime de responsabilité encadrée et limitée des prestataires de stockage a été élaboré par le législateur pour éviter que les hébergeurs ne soient condamnés pour des faits commis par les personnes dont ils hébergent les contenus.

En effet, avant que la loi du 30 septembre 1986 sur la liberté de communication ne soit modifiée par l'article L.43-8 de la loi du 1^{er} août 2000² et que la loi pour confiance dans l'économie numérique ne soit adoptée, plusieurs décisions de justice avaient condamné des hébergeurs du fait des contenus mis en ligne par des tiers.

Ainsi, dans l'affaire Estelle Hallyday, la Cour d'appel de Paris avait condamné l'hébergeur du site www.altern.org au motif qu'il hébergeait des photographies d'Estelle Hallyday partiellement ou complètement dénudées mises en ligne par un webmaster³.

Cette décision a fait l'objet de critiques de la part des défenseurs d'un régime de responsabilité encadrée et limitée propre aux hébergeurs, qui considéraient que de telles condamnations pouvaient, à terme, entraver la liberté d'expression sur internet et la création de sites internet et avoir pour conséquence la délocalisation des hébergeurs dans des pays où la loi leur serait plus favorable^{4,5}.

Ce nouveau régime se justifiait également par les contraintes techniques rencontrées par les prestataires de stockage.

La loi du 1^{er} août 2000 a ainsi mis fin à cette situation et cette position du législateur n'a pas été remise en cause depuis avec la loi pour la confiance dans l'économie numérique.

¹ Sur le plan pénal, les règles sont sensiblement identiques.

² Article L.43-8 de la loi du 1^{er} août 2000 : « Les personnes physiques ou morales qui assurent, à titre gratuit ou onéreux, le stockage direct et permanent pour mise à disposition du public de signaux, d'écrits, d'images, de sons ou de messages de toute nature accessibles par ces services, ne sont pénalement ou civilement responsables du fait du contenu de ces services que si, ayant été saisies par une autorité judiciaire, elles n'ont pas agi promptement pour empêcher l'accès à ce contenu ».

³ Cour d'appel de Paris, Estelle Hallyday c/ Valentin Lacambre, 10 février 1999.

⁴ « Quelle responsabilité pour les fournisseurs d'hébergement internet ? », revue Lamy Droit des affaires, mars 1999.

⁵ « Le web dans le collimateur de la justice », Libération, 27 décembre 1999.

1.2.1 L'absence d'obligation générale de surveiller les informations stockées et de rechercher des faits ou des circonstances révélant des activités illicites

L'article 6-I 7° de la loi pour la confiance dans l'économie numérique précise que l'hébergeur n'est pas tenu d'une obligation générale de surveiller les informations qu'il stocke ni même d'une obligation générale de rechercher des faits ou des circonstances révélant des activités illicites. Le juge peut toutefois imposer une telle mesure de surveillance ciblée et temporaire.

A cet égard, la Commission européenne a précisé que :

- « Il s'agit là d'un point important car la surveillance générale de millions de sites et de pages Web serait, en pratique, impossible et déboucherait sur une charge disproportionnée pour les prestataires intermédiaires et sur des coûts d'accès plus élevés aux services de base pour les utilisateurs »¹.

Les parlementaires français, quant à eux, ont souligné que :

- « Les hébergeurs ne sont pas soumis à une obligation générale de surveillance des contenus qu'ils stockent. Ils ne sont donc en aucun cas tenus de faire une recherche a priori des contenus illégaux »².

L'absence d'obligation de surveillance préalable tient également au fait que les litiges liés aux informations hébergées doivent normalement être réglés directement entre l'internaute et le tiers dont les droits seraient lésés.

C'est ainsi que l'article 6-I 5° de la loi pour la confiance dans l'économie numérique prévoit que la demande de retrait d'une information litigieuse adressée à l'hébergeur doit être accompagnée de la copie de la correspondance adressée à l'auteur de cette information demandant son retrait, ou la justification de ce que l'auteur n'a pu être contacté.

La loi du 5 mars 2007 relative à la prévention de la délinquance³ modifie l'article 6-I 7° de la loi pour la confiance dans l'économie numérique.

L'article 40 de la loi du 5 mars 2007 prévoit que les hébergeurs doivent participer outre à la lutte contre les crimes contre l'humanité, l'incitation à la haine raciale et la pornographie infantile, à la lutte contre l'incitation à la violence et les atteintes à la dignité humaine en mettant en place un dispositif permettant à toute personne de porter à leur connaissance ces infractions et en informant promptement les autorités publiques compétentes.

¹ Commission des Communautés européennes, Premier rapport sur l'application de la directive 2000/31/CE du Parlement européen et du Conseil du 8 juin 2000 relative à certains aspects juridiques des services de la société de l'information, et notamment du commerce électronique, dans le marché intérieur (« directive sur le commerce électronique »), COM (2003) 702 final du 21 novembre 2003, spéc. p.15.

² Avis sur le projet de loi pour la confiance dans l'économie numérique, document Assemblée nationale n°608 du 11 février 2003.

³ Loi n°2007-297 du 5 mars 2007 relative à la prévention de la délinquance.

La loi du 5 mars 2007 ne remet donc pas en cause le principe posé par l'article 6-I 7° de la loi pour la confiance dans l'économie numérique selon lequel les hébergeurs ne sont pas tenus d'une obligation générale de surveiller les informations stockées et de rechercher des faits ou des circonstances révélant des activités illicites.

1.2.2 Le caractère « manifestement illicite » des informations dénoncées

L'article 6-I 2° de la loi pour la confiance dans l'économie numérique subordonne l'engagement de la responsabilité civile des hébergeurs à la connaissance par les hébergeurs du « caractère illicite » des informations en cause.

Ce texte doit être interprété au regard de la décision du Conseil constitutionnel du 10 juin 2004 aux termes de laquelle l'hébergeur ne pourra voir sa responsabilité engagée pour ne pas avoir retiré une annonce dénoncée comme illicite par un tiers que si celle-ci présente « manifestement un tel caractère » ou si son retrait a été « ordonné par un juge »¹.

Les hébergeurs ne se voient en conséquence imposer d'obligation de retrait des contenus que lorsque les informations qui leur sont signalées présentent un caractère manifestement illicite.

1.2.3 Le mécanisme dit de présomption par notification

La loi pour la confiance dans l'économie numérique prévoit, à son article 6-I 5°, que la connaissance des faits litigieux est présumée acquise par l'hébergeur lorsque lui sont notifiés les différents éléments suivants :

- la date de notification ;
- les éléments permettant l'identification du notifiant ;
- les éléments d'identification du destinataire de la notification ;
- la description des faits litigieux et leur localisation précise ;
- les motifs pour lesquels le contenu doit être retiré comprenant la mention des dispositions légales et des justifications de fait ;
- la copie de la correspondance adressée à l'auteur ou à l'éditeur des informations ou activités litigieuses demandant leur interruption, leur retrait ou leur modification, ou la justification de ce que l'auteur ou l'éditeur n'a pu être contacté.

La mise en cause de la responsabilité de l'hébergeur suppose donc une démarche active de l'émetteur de la notification auprès de l'auteur des informations ou activités illicites, ou une justification de ce que ce dernier n'a pu être contacté.

¹ Conseil constitutionnel, décision n° 2004-496 DC du 10 juin 2004.

1.3 LA JURISPRUDENCE SUR LA RESPONSABILITE DES HEBERGEURS

1.3.1 La jurisprudence de principe

La jurisprudence considère qu'un prestataire qui assure l'hébergement de contenus de tiers, quels que soient ces contenus (propos, annonces...) doit être entendu comme un prestataire d'hébergement au sens de l'article 6-I 2° de la loi pour la confiance dans l'économie numérique et qu'il ne saurait être considéré comme responsable des contenus qu'il héberge, sauf à avoir été informé par un tiers de la présence d'un contenu manifestement illicite et n'avoir pas agi promptement pour empêcher l'accès à ce contenu.

- Tribunal de grande instance de Paris, ordonnance de référé, 8 février 2002, SA Télécom City, José Macia et Nicolas Bakar c/ SA Finance Net 8 février 2002, à propos d'un forum de discussion : « Dès lors que la société Finance Net ne peut être reconnue responsable du fait du contenu des messages concernés que si, ayant été saisie par une autorité judiciaire, elle n'a pas agi promptement pour empêcher l'accès à ce contenu, en application de l'article 43-8 de la même loi »¹.
- Tribunal de grande instance de Paris, 11 février 2003, Association Amicale des déportés d'Auschwitz et des camps de Haute Silésie à Monsieur Timothy Koogle², président de la société de droit américain Yahoo Inc., à propos d'un site de vente aux enchères : « S'agissant de communication sur le réseau internet, il convient de distinguer les apporteurs de contenus ou d'informations qu'ils souhaitent voir rendre accessibles sur le réseau, d'une part, et les prestataires de services ou intermédiaires techniques qui assurent cette accessibilité, d'autre part ».
- Tribunal de grande instance de Lyon, 21 juillet 2005, Groupe Mace c/ Gilbert D., à propos d'un forum de discussion : « Le responsable d'un forum non modéré ou modéré a posteriori doit être considéré comme un hébergeur au sein de la loi puisqu'il assure le stockage direct des messages diffusés sans porter de regard préalable sur ces derniers »³.
- Cour d'appel d'Aix-en-Provence, 13 mars 2006, SA Lucent Technologies c/ SA Escota, SA Lycos France, Nicolas B., à propos d'un hébergeur de pages personnelles : la responsabilité d'un hébergeur de pages personnelles ne peut être engagée « sur la base des moyens fournis pour la création d'un site puisque le préjudice n'est pas en relation directe avec cette assistance »⁴.

¹ Tribunal de grande instance de Paris, ordonnance de référé, 8 février 2002, SA Télécom City, José Macia et Nicolas Bakar c/ SA Finance Net 8 février 2002.

² Tribunal de grande instance de Paris, 11 février 2003, Association Amicale des déportés d'Auschwitz et des camps de Haute Silésie c/ Monsieur Timothy Koogle.

³ Tribunal de grande instance de Lyon, 21 juillet 2005, Groupe Mace c/ Gilbert D.

⁴ Cour d'appel d'Aix-en-Provence, 13 mars 2006, SA Lucent Technologies c/ SA Escota, SA Lycos France, Nicolas B.

1.3.2 Une jurisprudence contraire : le cumul des qualités d'hébergeur et d'éditeur

Certains tribunaux ont jugé qu'un prestataire de stockage pouvait cumuler les qualités d'hébergeur et d'éditeur en se fondant, d'une part, sur le fait qu'il proposait une structure de présentation de l'information et, d'autre part, qu'il percevait une rémunération pour son service d'hébergement.

- Cour d'appel de Paris, 7 juin 2006, Tiscali Media c/ Dargaud Lombard à propos d'un hébergeur de pages personnelles : « Si la société Tiscali Media a (...) exercé les fonctions techniques de fournisseur d'hébergement (...), son intervention ne saurait se limiter à cette simple prestation technique dès lors qu'elle propose aux internautes de créer leurs pages personnelles à partir de son site www.chez.tiscali.fr (...) de sorte que la société Tiscali Media doit être regardée comme ayant aussi la qualité d'éditeur dès lors qu'il est établi qu'elle exploite commercialement le site www.chez.tiscali.fr puisqu'elle propose aux annonceurs de mettre en place des espaces publicitaires payants directement sur les pages personnelles »¹.
- Cour d'appel de Paris, 7 mars 2007, Hôtels Méridiens c/ Sedo, Stéphane H., à propos d'un site de courtage aux enchères de noms de domaine : « La société intimée ne peut bénéficier de la qualité d'intermédiaire technique au sens de l'article 6 de la loi n°2004-575 du 21 juin 2004 relative à la confiance dans l'économie numérique dès lors qu'il résulte des éléments de la procédure que la société Sedo déploie une activité, qui en tout état de cause, ne se limite pas à celle d'hébergeur de sites internet ou de fournisseur d'accès à internet. (...) en effet, il résulte des éléments de la procédure que la société intimée, d'une part, édite un site internet consacré aux noms de domaine qu'elle propose à la vente et, d'autre part, réalise des liens hypertextes publicitaires de sorte qu'elle exploite commercialement le site www.sedo.fr »².
- Tribunal de grande instance de Paris, ordonnance de référé, 22 juin 2007, Jean-Yves L. dit Lafesse c/ MySpace, à propos d'un hébergeur de pages personnelles : « S'il est incontestable que la société défenderesse exerce les fonctions techniques de fournisseur d'hébergement, elle ne se limite pas à cette fonction technique (...) en effet, imposant une structure de présentation par cadres, qu'elle met manifestement à la disposition des hébergés et diffusant, à l'occasion de chaque consultation, des publicités dont elle tire manifestement profit, elle a le statut d'éditeur et doit en assumer les responsabilités »³.

Cette jurisprudence est critiquable.

La loi pour la confiance dans l'économie numérique ne limite pas l'activité d'hébergement à sa prestation technique mais retient une définition de l'hébergement fondée sur la fonction exercée, à savoir le stockage de données à la demande du destinataire du service.

¹ Cour d'appel de Paris, 7 juin 2006, Tiscali Media c/ Dargaud Lombard.

² Cour d'appel de Paris, 7 mars 2007, Hôtels Méridiens c/ Sedo, Stéphane H.

³ Tribunal de grande instance de Paris, ordonnance de référé, 22 juin 2007, Jean-Yves L. dit Lafesse c/ MySpace.

D'autre part, le fait que l'hébergeur perçoive une rémunération ne modifie en rien le régime juridique des prestataires de stockage qui lui est applicable, l'article 6-I 2° de la loi pour la confiance dans l'économie numérique visant toutes les prestations de stockage qu'elles soient réalisées à titre gratuit ou à titre onéreux.

1.3.3 Le revirement de cette jurisprudence contraire : une définition de l'hébergement fondée sur la fonction exercée, à savoir le stockage de données à la demande du destinataire du service

Cette jurisprudence tendant à requalifier les prestataires de stockage en éditeur, a été remise en cause.

- Tribunal de grande instance de Paris, 13 juillet 2007, Christian C., Nord Ouest Production c/ Dailymotion, UGC Images : « La commercialisation d'espaces publicitaires ne permet pas de qualifier la société Dailymotion d'éditeur de contenus dès lors que lesdits contenus sont fournis par les utilisateurs eux-mêmes, situation qui distingue fondamentalement le prestataire technique de l'éditeur, lequel, par essence même, est personnellement à l'origine de la diffusion, raison pour laquelle il engage sa responsabilité »¹.
- Tribunal de grande instance de Paris, 19 octobre 2007, SARL Zadig Productions, Monsieur JV, Monsieur MV c/ Google Inc., AFA : « Le fait pour la société défenderesse d'offrir aux utilisateurs de son service GOOGLE VIDEO une architecture et les moyens techniques, permettant une classification des contenus, au demeurant nécessaire à leur accessibilité par la public, ne permet pas de la qualifier d'éditeur de contenu dès lors qu'il est constant que lesdits contenus sont fournis par les utilisateurs eux-mêmes, situation qui distingue fondamentalement le prestataire technique de l'éditeur, lequel, par essence même, est personnellement à l'origine de la diffusion et engage à ce titre sa responsabilité »².
- Cour d'appel de Paris, 12 décembre 2007, Google Inc. c/ Benetton, Bencom : « le fait qu'elle [Google Inc.] offre aux créateurs de blogs, à travers la plate-forme Blogger, une fonctionnalité d'installation et de présentation ou un système de protection contre des commentaires indésirables ne démontre pas sa qualité d'éditeur du contenu de ces blogs »³.
- Tribunal de commerce de Paris, 20 février 2008, Flach Film et autres c/ Google France, Google Inc. : « Le fait pour [Google Inc. et Google France] d'organiser la présentation du site, d'offrir aux internautes les moyens de classer et de présenter leurs vidéos, de subordonner le stockage de vidéos à l'acceptation de conditions générales ne leur confère pas le contrôle des contenus et des internautes (...). En outre, Google Inc. et Google France ne prennent aucune initiative dans le choix et la présentation des œuvres (...). Il en résulte que Google Inc. et Google France n'ont pas la qualité d'éditeur, et qu'elles agissent donc, en exploitant le service Google Vidéo, en qualité d'hébergeur »⁴.

¹Tribunal de grande instance de Paris, 13 juillet 2007, Christian C., Nord Ouest Production c/ Dailymotion, UGC Images, 13 juillet 2007.

²Tribunal de grande instance de Paris, 19 octobre 2007, SARL Zadig Productions, Monsieur JV, Monsieur MV c/ Google Inc., AFA.

³Cour d'appel de Paris, 12 décembre 2007, Google Inc. c/ Benetton, Bencom.

⁴Tribunal de commerce de Paris, 20 février 2008, Flach Film et autres c/ Google France, Google Inc.

Les tribunaux retiennent une définition de l'hébergement fondée sur la fonction exercée, à savoir le stockage de données à la demande du destinataire du service qui fournit les contenus (image, texte, vidéo, son) mis à la disposition du public.

1.3.4 Les obligations prétoriennes des hébergeurs

Certaines décisions tendent à accroître les obligations légales des hébergeurs : obligation de retirer promptement les contenus qui leur sont signalés et qui présentent en outre un caractère manifestement illicite, obligation de lutter contre les infractions « graves », obligation de détenir et conserver les données d'identification, et absence d'obligation de surveillance a priori.

1.3.4.1 De l'absence d'obligation légale de surveillance a priori à une obligation prétorienne particulière de surveillance

- Tribunal de grande instance de Paris, 13 juillet 2007, Christian C., Nord Ouest Production c/ Dailymotion, UGC Images : le tribunal, tout en rappelant que selon l'article 6-I 7° de la loi pour la confiance dans l'économie numérique « les prestataires techniques ne sont pas soumis à une obligation générale de surveiller les informations qu'ils transmettent ou stockent, ni à une obligation générale de rechercher des faits ou circonstances révélant des activités illicites », a considéré qu'il incombait à la société Dailymotion « de procéder à un contrôle a priori » sur les vidéos. Le tribunal justifie sa décision par le fait que « la société Dailymotion doit être considérée comme ayant connaissance à tout le moins de faits et circonstances laissant penser que des vidéos illicites sont mises en ligne »¹.
- Tribunal de commerce de Paris, ordonnance de référé, 26 juillet 2007 Parfums Christian Dior, Kenzo Parfums, Parfums Givenchy et Guerlain c/ sociétés DMIS : le tribunal a imposé aux sociétés DMIS, hébergeurs du site de petites annonces en ligne www.vivastreet.fr « de mettre en place un système de surveillance pour une durée de six mois (6), afin de prévenir ou de retirer toutes annonces litigieuses ». En l'espèce, il s'agissait des annonces « proposant la vente hors du réseau de distribution sélective des demanderesses de parfums et produits cosmétiques dont le texte utilise les dénominations des demanderesses et/ou comportant un tableau de concordance ou d'équivalence avec ses dénominations et/ou offrant à la vente des parfums ou cosmétiques de grandes marques présentés comme génériques »².
- Tribunal de commerce de Paris, 20 février 2008, Flach Film et autres c/ Google France, Google Inc. : « Si l'hébergeur n'est pas tenu à une obligation de surveillance générale, il est tenu à une obligation de surveillance, en quelque sorte particulière, à partir du moment où il a eu connaissance du caractère illicite du contenu »³.

¹ Tribunal de grande instance de Paris, 13 juillet 2007, Christian C., Nord Ouest Production c/ Dailymotion, UGC Images.

² Tribunal de commerce de Paris, ordonnance de référé, 26 juillet 2007 Parfums Christian Dior, Kenzo Parfums, Parfums Givenchy et Guerlain c/ sociétés DMIS.

³ Tribunal de commerce de Paris, 20 février 2008, Flach Film et autres c/ Google France, Google Inc.

La jurisprudence tend, au vu des dernières décisions, à appliquer aux hébergeurs une obligation particulière de surveillance des contenus qu'ils stockent dès lors qu'ils ont été notifiés.

1.3.4.2 Vers un élargissement des obligations des hébergeurs

o Les contours de la notion de « caractère manifestement illicite »

Les hébergeurs ne se voient imposer d'obligation de retrait des contenus que lorsque les informations qui leur sont signalées présentent un caractère manifestement illicite¹.

En effet, les hébergeurs ne sauraient se voir ériger en juges ni voir leur responsabilité engagée par les auteurs des contenus à raison d'une intervention abusive, qui serait en outre susceptible de porter atteinte à la liberté d'expression.

Ainsi, Monsieur le député Jean Dionis du Séjour énonce dans son rapport que « l'ajout de l'adverbe « manifestement » a visé notamment à protéger les hébergeurs dans les cas où le litige portait sur des droits d'auteur ou des droits de propriété, car l'appréciation du bien fondé d'un droit est très délicate dans ces domaines »².

La jurisprudence n'est pas totalement fixée quant aux contours de la notion de « manifestement illicite » :

- la Cour d'appel de Paris, dans un arrêt du 24 novembre 2006, a jugé que, s'agissant d'un site mis en ligne par l'Association des anciens amateurs de récits de guerre et de l'Holocauste diffusant une compilation d'écrits et de propos antisémites et révisionnistes téléchargeables « ce site, dont le contenu est constitutif d'infractions pénales, est manifestement illicite »³ ;
- dans un jugement du 15 novembre 2004⁴, confirmé le 8 novembre 2006 par la Cour d'appel de Paris⁵, le Tribunal de grande instance de Paris a refusé de qualifier de « manifestement illicite » les contenus contestant l'existence du génocide arménien aux motifs que le caractère manifestement illicite des informations litigieuses « ne peut être la conséquence que d'un manquement délibéré à une disposition de droit positif explicite et dénuée d'ambiguïté ». Or, si la loi du 29 janvier 2001 relative à la reconnaissance du génocide arménien de 1915 « concerne explicitement le sujet du génocide arménien et ne pouvait, à ce titre, échapper à la connaissance du fournisseur d'hébergement du site internet, [elle] ne met cependant aucune obligation à la charge des particuliers et constitue seulement une prise de position officielle (...) du pouvoir législatif français sur cet événement historique » ;

¹ Conseil constitutionnel, décision n° 2004-496 DC du 10 juin 2004.

² Assemblée Nationale, Rapport n°612 fait au nom de la Commission des affaires économiques, de l'environnement et du territoire sur le projet de loi, pour la confiance dans l'économie numérique (extraits).

³ Cour d'appel de Paris, Tiscali Access et autres c/ Free, UEJF et autres, 24 novembre 2006.

⁴ Tribunal de grande instance de Paris, Comité de défense de la cause arménienne c/ M. Aydin et France Télécom, 15 novembre 2004.

⁵ Cour d'appel de Paris, Comité de défense de la cause arménienne c/ M. Aydin et France Télécom, 8 novembre 2006.

- Cour d'appel de Paris, 6 juin 2007, Lycos France c/ Abdelhadi S., SA Dounia et SAS iEurop : « Même s'il est reconnu à l'hébergeur une marge d'appréciation dans l'interprétation de la licéité des données qu'un particulier lui dénonce, des propos portant de façon évidente atteinte à l'intimité de la vie privée (...) sont manifestement illicites et l'hébergeur doit en conséquence les retirer ou en rendre l'accès impossible sans que cela n'ait à être requis par une décision de justice »¹ ;
- dans l'affaire opposant Google Inc. à Benetton, la Cour d'appel de Paris a jugé que que l'hébergeur doit « lorsqu'il se voit dénoncer des données dont le contenu est déclaré illicite, non s'en remettre à l'appréciation des juges, mais apprécier si un tel contenu a un caractère manifestement illicite et, dans cette hypothèse, supprimer ou rendre inaccessible de telles données »².

Ainsi donc, la jurisprudence la plus récente adopte une interprétation extensive de la notion de « manifestement illicite » alors même que les avis et rapports présentés au cours des débats parlementaires précisent, d'une part, que la notion de « manifeste » renvoie à la notion d'évidence, pour un profane ou un non-professionnel du droit et, d'autre part, que la question des droits de propriété intellectuelle est une question complexe ne pouvant entraîner une obligation d'agir de l'hébergeur.

○ **L'obligation de retirer « sans délai » les contenus notifiés**

Aux termes de l'article 6-I 2° de la loi pour la confiance dans l'économie numérique, les prestataires de stockage ne sont responsables du fait des activités ou des informations stockées à la demande d'un destinataire de ces services :

- que s'ils avaient effectivement connaissance de leur caractère illicite ou de faits et circonstances faisant apparaître ce caractère ;
- ou si, dès le moment où ils en ont eu connaissance, ils n'ont pas agi promptement pour retirer ces informations ou en rendre l'accès impossible.

La loi pour la confiance dans l'économie numérique ne définit pas le terme « promptement ».

Par ordonnance de référé du 13 mai 2008, le Tribunal de grande instance de Toulouse, tout en rappelant que selon l'article 6-I 2° de la loi pour la confiance dans l'économie numérique les prestataires de stockage « ne peuvent pas voir leur responsabilité civile engagée du fait des activités ou des informations stockées à la demande d'un destinataire de ces services si elles n'avaient pas effectivement connaissance de leur caractère illicite ou de faits et circonstances faisant apparaître ce caractère ou si, dès le moment où elles en ont eu cette connaissance, elles ont agi promptement pour retirer ces données ou en rendre l'accès impossible », a considéré que pour pouvoir être qualifié de prompt le retrait doit avoir eu lieu sans délai³.

¹ Cour d'appel de Paris, 6 juin 2007, Lycos France c/ Abdelhadi S., SA Dounia et SAS iEurop.

² Cour d'appel de Paris, Google Inc. c/ Benetton, 12 décembre 2007.

³ Tribunal de grande instance de Toulouse, ordonnance de référé, 13 mars 2008, Krim K. c/ Pierre G. Amen.

- **L'étendue de l'obligation de détenir et conserver les données d'identification**

La loi pour la confiance dans l'économie numérique n'impose pas aux hébergeurs de vérifier les données d'identification qu'ils doivent recueillir¹.

Au cours des travaux parlementaires relatifs à cette loi, les parlementaires avaient en effet refusé une telle obligation de vérification des données dans les termes suivants :

- « La réserve est d'ordre juridique et tient à la compatibilité d'une telle obligation au regard des dispositions de la directive communautaire du 8 juin 2000. Celle-ci ne prévoit en effet aucune obligation de ce type à la charge des intermédiaires techniques de la société de l'information. Elle n'ouvre pas, par ailleurs, expressément aux Etats membres la faculté d'exiger la vérification de contenus »².

Peu de décisions ont été rendues à ce jour sur l'obligation de vérification des hébergeurs :

- le Tribunal de grande instance de Paris, par ordonnance du 2 février 2004, a estimé qu'un hébergeur n'était pas tenu de vérifier les informations qui lui sont communiquées³ ;
- en 2006, la Cour d'appel de Paris, confirmant un jugement du Tribunal de grande instance de Paris du 16 février 2005, a estimé que la société Tiscali Média avait commis une négligence, au sens de l'article 1383 du Code civil, en se contentant des coordonnées fantaisistes d'identification fournies par le client⁴ lesquelles ne permettaient pas l'identification de la personne concernée⁵ ;
- dans l'affaire opposant Google Inc. à Benetton, la Cour d'appel de Paris vient de juger que Google Inc. « ne pouvait se contenter de fournir (...) une adresse IP en les renvoyant au fournisseur d'accès du blog litigieux pour obtenir l'identité de l'auteur (...), alors qu'en qualité d'hébergeur, elle devait disposer (...) des

¹ Le projet de décret portant application de l'article 6 de la loi pour la confiance dans l'économie numérique précise que les personnes fournissant des prestations d'hébergement au sens de l'article 6-I 2° doivent détenir et conserver les données suivantes :

- 1°) pour chaque opération de création, les données permettant d'identifier l'origine de la création des contenus ;
- 2°) les informations fournies lors de la souscription d'un contrat par un utilisateur ou lors de la création d'un compte ;
- 3°) lorsque la souscription du contrat ou du compte est payante, les informations relatives au paiement.

Le projet de décret précise que les données mentionnées aux 2°) et 3°) ne doivent être conservées que dans la mesure où l'hébergeur les collecte habituellement.

Le projet de décret prévoit par ailleurs que les données d'identification des créateurs de contenus doivent être conservées un an à compter du jour de la création des contenus, pour chaque opération contribuant à la création d'un contenu.

² Avis sur le projet de loi pour la confiance dans l'économie numérique, document Sénat, n°351, 11 juin 2003.

³ Tribunal de grande instance de Paris, ordonnance de référé, Métrobus c/Ouvaton, 2 février 2004.

⁴ nom : Bande ; prénom : Dessinée ; adresse : rue de la BD.

⁵ Cour d'appel de Paris, Tiscali Média c/ Dargaud Lombard, Lucky Comics, 7 juin 2006.

éléments d'identité qui lui étaient demandés »¹. Dans cette affaire, l'hébergeur ne détenait aucune donnée d'identité (nom, prénom, adresse, n° de téléphone) mais seulement l'adresse IP et une adresse e-mail.

Il résulte de ces décisions que :

- la nature des données recueillies doit être conforme aux exigences de la loi pour la confiance dans l'économie numérique (affaire Google Inc. c/ Benetton) ;
- les vérifications de l'hébergeur doivent a minima porter sur l'absence de caractère fantaisiste des données fournies (affaire Tiscali Media) ;
- à ce jour, la jurisprudence ne semble pas exiger, en l'absence de caractère « manifestement fantaisiste » des données de vérification de la réalité des données fournies.

¹ Cour d'appel de Paris, Google Inc. c/ Benetton, 12 décembre 2007.

2. L'ÉTAT DE L'ART EN MATIÈRE DE FILTRAGE DE CONTENUS

Le filtrage de contenus est l'axe des technologies clés 2010 identifié par le ministère de l'Économie, des Finances et de l'Industrie dès novembre 2006¹.

L'état de l'art en matière de filtrage de contenus permet de constater une césure nette entre les technologies s'appuyant sur l'analyse lexicographique et celles constituées sur la base d'une analyse sémantique.

Les technologies s'appuyant sur la seule analyse lexicographique se sont développées depuis les années 60 et sont devenues des technologies matures pour analyser des bases de données de grande taille grâce à l'évolution de la puissance des calculateurs organisés ou non en réseau (grid computing). La maturité et l'usage de ces technologies ont évolué ces dernières années du fait notamment de l'amélioration de la puissance de calcul².

Les travaux concernant l'analyse sémantique sont moins aboutis et manifestement plus complexes.

L'analyse sémantique de contenus est technologiquement moins aboutie de sorte qu'elle est plus ou moins opérationnelle et intégrée selon les types de contenus et plus précisément les médias qui portent ces contenus.

Des techniques spécifiques ont été développées pour analyser les aspects sémantiques des textes, des images, sons et vidéos.

Les techniques d'analyse et de traitement développées pour chacun de ces médias sont hétérogènes mais s'appuient souvent sur des méthodes statistiques (bayésiennes, stochastiques, etc.) de sorte que le résultat du filtrage (analyse et identification) ne peut intrinsèquement pas être certain au sens des probabilités³.

Pour être performantes, les technologies de filtrage développées en matière de traitement du son, de l'image ou de la vidéo sont souvent intégrées dans des outils spécifiques à un métier ou un secteur d'activité.

En fonction des domaines d'application, les taux de réussite des filtrages sémantiques sectoriels peuvent parfois dépasser les 90 %⁴.

¹ Annexe 4 : Bibliographie commentée, n°3.43 et 3.44.

² Annexe 4 : Bibliographie commentée, n°3.18.

³ Ensemble de mesure nulle.

⁴ Annexe 4 : Bibliographie commentée, n°3.42.

Dans presque tous les cas, les technologies développées semblent s'appuyer sur une structuration en deux grandes fonctionnalités, la première ayant un rôle d'analyse des données et la seconde un rôle de comparaison avec une base de contenus de référence (base de filtres de référence).

Cette base de filtres de référence peut être constituée de mots-clés (dans le domaine de l'analyse lexicographique), d'images prototypes de référence (dans le domaine du traitement d'images), de sons de référence (dans le domaine de l'analyse de sons et/ou de la parole).

Ces bases de filtres de référence sont également indispensables pour mettre en place des mécanismes automatiques d'apprentissage sous la forme de généralisation et/ou de catégorisation.

La taxinomie retenue dans le cadre de la présente description de l'état de l'art a naturellement été constituée en distinguant d'une part les technologies d'analyse lexicographique et, d'autre part, les technologies d'analyse sémantique.

De plus, il a été nécessaire de distinguer au sein des technologies d'analyse sémantique, celles qui relèvent du traitement de textes, d'images, de sons ou de vidéos.

Cette classification se retrouve aussi bien lorsque l'on étudie :

- les acteurs qui utilisent ou offrent des technologies de filtrage de contenus ;
- les projets de recherche-développement en matière de filtrage de contenus ;
- les brevets déposés en matière de filtrage de contenus.

L'intégration dans des logiciels en production des technologies les plus récentes issues de la recherche-développement paraît être principalement le fait de sociétés américaines dont l'activité principale se situe à l'échelle mondiale et est déployée dans le cadre d'applications et de bases de données sur internet. Ceci ressort notamment des titulaires de droits sur les brevets concernant les technologies identifiées. Il est vraisemblable que d'autres acteurs stratégiques dans le domaine de la défense disposent et/ou maîtrisent des technologies équivalentes mais ne rendent pas cette information publique.

Dans ce contexte, on peut rappeler l'importance de l'exploration de données ou « data mining » qui a pour objet :

- l'extraction, au moyen de méthodes automatiques ou semi-automatiques, d'un savoir ou d'un ensemble d'informations à partir de grandes quantités de données, lesquelles sont généralement stockées dans un ou plusieurs entrepôts de données ou « datawarehouse » ;
- l'utilisation industrielle ou opérationnelle de ce savoir ou de cet ensemble d'informations.

Le « data mining » peut s'accompagner d'outils dits de « fouille de textes » ou « text mining » lesquels permettent d'associer aux principes du « data mining » l'analyse lexicographique ou linguistique multilingue de données non structurées, notamment des courriers électroniques.

Il est notamment utilisé dans le domaine de la lutte contre le spam ou tout autre courrier électronique indésirable ainsi que dans celui, plus général, de l'analyse de contenus.

Le « data mining » est ainsi l'un des socles recherche-développement sur lequel reposent les différentes technologies de filtrage de contenus.

Les technologies d'analyse de contenus se distinguent de par leur fonctionnement mais également de par leurs performances lesquelles varient très largement selon la nature et/ou le format du contenu analysé (texte, photo, vidéo ou son).

S'agissant d'un contenu texte, l'analyse peut être :

- syntaxique (analyse de chaînes de caractères, de mots-clés, le cas échéant, analyse multi-critères) ;
- sémantique (analyse du contexte).

S'agissant d'un contenu photo, vidéo ou son, l'analyse est, en revanche, exclusivement sémantique avec :

- pour l'image, des technologies de traitement d'images ou de reconnaissance de formes ;
- pour la vidéo, des technologies de reconnaissance de schèmes permettant de catégoriser les vidéos.

Les technologies d'analyse syntaxique s'avèrent relativement opérationnelles. En revanche, les technologies d'analyse sémantique ne le sont pas universellement, leur efficacité variant selon le caractère plus ou moins fermé de la zone d'intervention.

Dans cette perspective, ont été identifiés :

- les principaux acteurs du filtrage de contenus ;
- les principaux outils de filtrage de contenus ;
- les principaux contrats faisant référence au filtrage de contenus ;

2.1 LES ACTEURS DU FILTRAGE DE CONTENUS

Les principaux acteurs intervenant en matière de filtrage de contenus identifiés sont :

- les utilisateurs de technologies de filtrage de contenus ;
- les éditeurs, concepteurs et/ou intégrateurs de technologies ;
- les titulaires de droits ;
- les acteurs, notamment français, en matière de recherche-développement dans ce domaine.

Ils doivent être envisagés en distinguant :

- le filtrage de contenus texte et/ou multimédia (photos, vidéos, sons) ;
- le filtrage d'URL ;
- le filtrage de messages électroniques (anti-spam) ;
- le filtrage à but de contrôle parental.

2.1.1 Les principaux utilisateurs de logiciels de filtrage de contenus

2.1.1.1 L'utilisation de filtres par des moteurs de recherche

o **Google**

Google est le leader mondial dans le secteur des moteurs de recherche.

Son moteur de recherche est doté d'un double système de filtrage qui s'exerce tout à la fois sur la recherche elle-même et sur le contenu de la recherche.

Le filtrage de la recherche s'exerce selon les systèmes dits des « pages ranks »¹ et des « adwords »².

Quant au filtrage du contenu de la recherche, le moteur de recherche de Google propose à ses utilisateurs l'option de filtrage « SafeSearch »³.

Cette fonctionnalité se présente comme un filtrage de type « contrôle parental », filtrant les sites présentant un contenu à caractère sexuel et les excluant ainsi des résultats de la recherche.

Cette fonctionnalité se dédouble avec :

- un filtrage modéré ne s'appliquant qu'aux seules images à caractère sexuel ;
- un filtrage strict s'appliquant à tous les résultats de la recherche, aux recherches d'images mais aussi aux recherches standard sur le web.

o **Yahoo**

Yahoo est l'un des acteurs majeurs dans le domaine des moteurs de recherche.

De la même manière que Google avec son système des « adwords », le moteur de recherche Yahoo dispose de systèmes de filtrage de la recherche, notamment :

- le « Pay per Click »⁴ ;
- l'« option de ciblage »⁵, qui est un système de filtrage fonctionnant par listes de mots-clés et permettant, au moyen notamment des mots-clés choisis par

¹ Annexe 4 : Bibliographie commentée, n°1.36.

² Annexe 4 : Bibliographie commentée, n°1.32.

³ Annexe 4 : Bibliographie commentée, n°5.5.

⁴ Annexe 4 : Bibliographie commentée, n°1.32.

⁵ Annexe 4 : Bibliographie commentée, n°1.34.

l'utilisateur, d'orienter la recherche de ce dernier vers le site web effectivement ciblé.

○ **Exalead**

Exalead est une société française, éditrice de logiciels, et notamment, d'un logiciel de moteur de recherche.

Afin d'améliorer son service de recherche d'images sur internet, Exalead s'est dotée, récemment, d'un filtre de tri d'images¹.

Ce filtre, utilisant la technologie développée par LTU Technologies, est mis en œuvre au moyen de deux logiciels :

- l'Image-Seeker, un système complet de gestion des images avec indexation et recherche par le contenu ;
- l'Image-Filter, un logiciel de gestion de contenu visuel permettant l'analyse automatique des images, autrement dit une plate-forme de classification de contenus².

2.1.1.2 L'utilisation de filtres par des hébergeurs de contenus vidéo

○ **Dailymotion**

Dailymotion est une société française qui propose aux utilisateurs inscrits, d'héberger les contenus vidéo que ces derniers souhaitent mettre en ligne.

Dailymotion est l'un des quatre principaux acteurs qui, avec YouTube, Soapbox et MySpace, se partagent ce marché restreint.

En juillet 2007, Dailymotion a annoncé avoir opté pour la solution d'identification et de filtrage de contenus, proposée par Audible Magic, et intitulée « Content Identification Services »³.

Cette solution de filtrage de contenus vidéo met en œuvre une technologie de type « fingerprinting » et consiste à analyser un contenu vidéo en référence à une base de données d'« empreintes digitales » des versions originales protégées par le droit de la propriété intellectuelle. L'analyse du contenu est notamment réalisée grâce au spectre sonore de la vidéo de telle sorte que cette solution de filtrage de contenus vidéo s'avère principalement basée sur une technologie de filtrage de contenus audio.

En octobre 2007, Dailymotion a indiqué compléter son système de filtrage de contenus en recourant à la solution « Signature » proposée par l'Ina⁴.

¹ Annexe 4 : Bibliographie commentée, n°1.22.

² Annexe 4 : Bibliographie commentée, n°1.24.

³ Annexe 4 : Bibliographie commentée, n°1.6.

⁴ Annexe 4 : Bibliographie commentée, n°1.7 et 1.8.

En application de ce système, Dailymotion devrait scanner chaque vidéo avant leur mise en ligne, et rechercher la présence d'une empreinte invisible à l'œil nu, laquelle serait ajoutée par l'Ina. Si une telle signature est présente, Dailymotion agirait selon les ordres des ayants-droit qui ont marqué leur vidéo, ces derniers pouvant :

- décider d'interdire totalement la diffusion de la vidéo, auquel cas DailyMotion rejettera la vidéo ;
- choisir de bénéficier de la mise en ligne et de partager avec la plate-forme les revenus publicitaires générés par la vidéo copiée.

o **YouTube**

YouTube, filiale de la société Google Inc., est une société américaine dont l'objet est de proposer aux internautes d'héberger les contenus vidéo que ces derniers souhaiteraient mettre en ligne, afin que ceux-ci soient rendus accessibles à tous.

Créée en février 2005, YouTube a été rachetée par Google Inc. en novembre 2006.

YouTube est le leader mondial de la vidéo en ligne.

De la même manière que Dailymotion, YouTube a opté pour la solution de filtrage proposée par Audible Magic¹ et l'a complétée par une solution complémentaire qu'elle a, en revanche, elle même développée, la « Video Identification »².

2.1.1.3 L'utilisation de filtres par MySpace, hébergeur de contenus a la fois audio, vidéo et image

MySpace, société américaine, est une communauté privée qui vise à permettre aux internautes inscrits, de mettre en ligne des contenus audio, vidéo et image sur des pages personnalisables.

Comprenant plus de 200 000 membres, MySpace est le leader mondial dans ce secteur et devance notamment Facebook, son concurrent le plus direct.

MySpace a choisi la solution de logiciel de filtrage de contenus vidéo proposée par la société Audible Magic³.

Elle a également choisi la solution de filtrage de contenus audio proposée par Gracenote⁴, laquelle repose sur une technologie d'identification par empreintes acoustiques, autrement dit de type « audio fingerprinting ».

¹ Annexe 4 : Bibliographie commentée, n°1.18.

² Annexe 4 : Bibliographie commentée, n°1.3 et 1.35.

³ Annexe 4 : Bibliographie commentée, n°1.21.

⁴ Annexe 4 : Bibliographie commentée, n°1.20 et 1.28.

2.1.1.4 L'utilisation de filtres par des fournisseurs de messagerie électronique

○ **Yahoo**

Yahoo fournit des messageries électroniques.

En cette qualité, Yahoo a longtemps eu recours au système **DomaineKeys**, un système de filtrage tout à la fois antispamming, antiphishing, et antispoofing¹. Ce système permettait à Yahoo de sécuriser ses messageries électroniques en les dotant d'un mécanisme destiné à vérifier, pour tout message reçu, le domaine de l'expéditeur et l'intégrité du message.

Néanmoins, Yahoo et Cisco ont, récemment, mis au point, un nouveau système de filtrage anti-spam reposant sur un principe d'authentification, le **DomaineKeys Identified Mail (DKIM)**².

Ce système consiste à joindre une signature numérique chiffrée dans le courrier électronique d'un expéditeur, cette signature dans le courrier électronique sortant permet de s'assurer de l'identité de l'expéditeur, via notamment les serveurs de courriers électroniques.

L'Internet Engineering Task Force (IETF), organisation technique ouverte et autoorganisée, dont l'objet consiste à participer à l'élaboration des standards du web, a d'ores et déjà approuvé ce système, de même d'ailleurs que Microsoft alors pourtant que cette dernière œuvre déjà sur sa solution **Sender ID**.

○ **Microsoft**

Microsoft propose des services de courrier électronique au travers de ses plate-formes de messageries électroniques, MSN et Hotmail.

Dans ce cadre, et depuis 2005, Microsoft a généralisé l'utilisation de son filtre antispam baptisé « **Sender ID** ».

Ce filtre vise à éliminer tout spam et tout autre courrier électronique non désiré, des boîtes électroniques des utilisateurs.

Il est le produit de la combinaison des technologies « **Sender Policy Framework** » de Pobox.com et « **Caller ID** » de Microsoft³.

La technologie qu'il met en œuvre consiste à cacher chaque e-mail au moyen d'un identifiant unique permettant d'authentifier le serveur émetteur et de s'assurer ainsi qu'il ne sert pas à l'envoi de spam.

La première version de **Sender ID** avait été rejetée en septembre 2004 par l'IETF au motif qu'il ne permettait pas un accès facile à la concurrence.

¹ Annexe 4 : Bibliographie commentée, n°4.5.

² Annexe 4 : Bibliographie commentée, n°1.30.

³ Annexe 4 : Bibliographie commentée, n°4.1.

- **Lycos**

Lycos, au travers de sa plate-forme Caramail, est un fournisseur de messageries électroniques.

Elle dispose d'un système de filtrage antispam et antivirus, intitulé « Jubii ».

Le fonctionnement de ce système est brièvement décrit dans les conditions générales d'utilisation de Lycos. Dans ces conditions générales d'utilisation¹, il est notamment indiqué que :

- « Jubii organise les courriers électroniques « entrants » en fonction de leur fiabilité dans les différents dossiers et/ou identifie de manière particulière certains de ces courriels (...).Les courriers électroniques suspectés d'être des spams sont marqués et placés dans des dossiers désignés en tant que tels (...) ».

2.1.2 Les principaux éditeurs, producteurs et intégrateurs de logiciels de filtrage

2.1.2.1 Les principaux éditeurs, producteurs et intégrateurs de logiciels de filtrage de contenus texte

- **Ilog**

Fondée en 1987, Ilog² est une société française qui se livre à des activités d'édition de logiciels, principalement, d'optimisation et de visualisation système. Néanmoins, un grand nombre des solutions qu'elle propose, telles notamment les logiciels « Ilog Rules » et « Ilog Jrules », comportent un ou plusieurs modules de filtrage, ce qui fait d'elle l'un des acteurs majeurs du marché du filtrage.

Ces modules de filtrage relèvent du domaine du filtrage de contenus texte et mettent en œuvre des technologies d'analyse de contenus de type « syntaxique » (mots-clés, chaînes de caractères etc.) mais aussi de type « sémantique » (analyse du contexte).

Elle dispose de deux principaux établissements respectivement situés en France et aux Etats-Unis et de plusieurs filiales en Europe mais aussi au Japon.

En 2007, elle a réalisé un chiffre d'affaires de 161,5 millions de dollars.

Elle s'adresse à une clientèle internationale, principalement composée d'entreprises.

- **Websense**

Websense³ est une société de droit californien créée en 1994, dont le cœur de métier est l'édition de logiciels de filtrage, en particulier de logiciels de filtrage d'URL.

¹ Annexe 7 : Extraits de conditions générales d'utilisation, n°2.

² <http://www.ilog.fr/products/supplychain/>.

³ <http://www.websense.com/global/en/>.

Elle est implantée en Amérique du Nord, en Amérique latine, en Europe avec des bureaux en France, en Italie, en Grande Bretagne, aux Pays-Bas et en Espagne mais aussi dans le reste du monde, en Australie, en Israël, au Moyen-Orient ou en Asie.

En 2001, SurfControl et Websense se partageaient l'essentiel du marché mondial des solutions de filtrage web destinées aux entreprises, avec respectivement 21,9 % et 17,6 % de part de marché. En 2005, le chiffre d'affaires de Websense s'élevait à 196,2 millions de dollars.

Néanmoins, depuis le rachat, en 2007, de son principal concurrent, SurfControl, Websense est devenue le premier éditeur de solutions de filtrage web dans le monde.

Cette absorption a, par ailleurs, permis à Websense de se diversifier, en mettant à sa disposition un portefeuille de technologies complémentaires à celles dont elle disposait d'ores et déjà., telles que :

- Blackspider, un fournisseur de messagerie sécurisée ;
- Apero, un anti-spyware etc.

Néanmoins, le cœur de métier de Websense demeure le filtrage de contenus par URL, de type « syntaxique », par mots-clés et/ou par liste noire ou « blacklist ».

Dans ce domaine, Websense propose trois principaux produits :

- « Websense Enterprise » ;
- « Websense Express » ;
- « Websense Content Protection Suite ».

Websense propose, en outre, des logiciels antispam, tels que « Websense Email Security » ou encore « Websense Hosted Email Security ».

○ **ISS**

Créée en 1994, Internet Security Systems, Inc. (ISS)¹, filiale d'IBM, est une société américaine dont le siège social est situé à Atlanta, en Géorgie.

Elle dispose de filiales aux Etats-Unis, en Asie, en Australie, en Europe et au Moyen-Orient et s'adresse à une clientèle principalement composée d'entreprises et d'organisations publiques et répartie sur le monde entier.

Elle est l'un des leaders du marché des produits et services de protection dynamique contre les menaces liées à Internet.

Les produits et services proposés par ISS s'appuient sur des fonctions intelligentes de sécurité proactives, conçues par son équipe de recherche et développement, la X-Force. Elle édite notamment un logiciel de filtrage d'URL, le « Proventia Web Filter », et dispose, dans ce cadre, et selon ISS, de la plus grande base d'URL filtrées (60 millions) et fournit le plus grand nombre de rapports (100).

¹ <http://www.iss.net/issEn/delivery/prdetail.jsp?type=France&oid=25401>.

○ **Secure Computing**

Secure Computing¹ est une société américaine, leader mondial des passerelles de sécurisation des entreprises.

Elle propose une gamme de solutions mettant en œuvre la technologie « TrustedSource » et destinées à permettre aux entreprises de sécuriser leurs passerelles web, leur réseau et leur messagerie et de gérer les accès et les identités.

Les clients de Secure Computing interviennent dans les secteurs de la banque, des services financiers, de la santé, des télécommunications, de l'industrie, des services publics, ou sont des collectivités locales et administrations. Ils sont localisés aux Etats-Unis mais aussi en Europe, au Japon, en Chine, sur le littoral Pacifique et en Amérique latine.

En 2006, Secure Computing a réalisé un chiffre d'affaires de 176,7 millions de dollars.

Secure Computing entretient des relations étroites avec les grandes agences du gouvernement américain, dans le cadre notamment de plusieurs contrats de recherche sur la sécurité de pointe.

A l'exception de certains grands comptes, toutes les solutions que propose Secure Computing sont commercialisées par le biais de leurs partenaires, tels que Alternative Technology, Blue Coat Systems, Cisco, Hewlett-Packard, McAfee, Microsoft...

En octobre 2003, Secure Computing a racheté N2H2 et a lancé, en 2005, une nouvelle version des logiciels de filtrage « Smartfilter » et « Smartfilter Bess edition ».

Cette même année 2005, Secure Computing a racheté Cyberguard et a ainsi enrichi sa gamme de logiciels de filtrage avec la solution « Webwasher ».

Cette solution « Webwasher » met en œuvre une technique de filtrage par URL destinée à lutter contre le contenu Web inapproprié et comporte, par ailleurs, une fonction antispam conjuguant plusieurs méthodes de détection du courrier indésirable.

○ **Olféo**

Olféo² est une société française qui se livre à des activités d'édition de logiciels, notamment, de filtrage d'URL.

Elle est fortement implantée en France et vise une clientèle diversifiée allant de l'administration, dont le Conseil général et l'Eure et Loire, aux entreprises, telles qu'ING Direct ou Air France, en passant par les établissements scolaires et par les centres hospitaliers.

Selon Olféo, elle serait l'un des seuls éditeurs sur le marché à proposer une solution de filtrage fondée sur une analyse humaine des URL.

Cette technologie intitulée « filtrage dynamique » consiste à identifier chaque site qui ne serait pas reconnu la première fois par l'outil de filtrage afin de le renvoyer vers l'éditeur pour qu'il effectue un classement immédiat. Elle présente le principal avantage de

¹ <http://www.securecomputing.com/index.cfm?skey=233&lang=fr>.

² <http://www.olfeo.com/>.

personnaliser l'outil de filtrage à l'utilisation propre de l'entreprise utilisatrice ; si celle-ci utilise spécifiquement des sites dans un secteur d'activité particulier ou à une région du monde précise ou encore que certains utilisateurs ont des centres d'intérêt particuliers, l'outil de filtrage s'y adapte.

- **Optenet**

Implantée en Europe, en Amérique du Nord et en Amérique latine, la société Optenet¹ est spécialisée dans le filtrage de sites web, et propose notamment des solutions de filtrage par URL.

Elle compte parmi sa clientèle un certain nombre de fournisseurs d'accès à internet mais se présente avant tout comme l'un des principaux acteurs du marché du filtrage au titre du contrôle parental lequel filtrage s'adresse aux particuliers.

Dans ce cadre, elle propose une solution familiale et payante qui se nomme « Web Filter ».

Ce logiciel n'est pas uniquement destiné aux parents dans le cadre du contrôle parental. Dans ce cadre néanmoins, il permet aux parents de contrôler l'accès de leurs enfants à des contenus Internet qui ne leur seraient pas appropriés. Il leur permet, notamment, d'établir les horaires où la navigation est permise, d'établir un nombre maximum d'heures de navigation ou de limiter le téléchargement de fichiers.

Cette solution de filtrage n'est pas entièrement dépendante de listes d'URL préétablies, mais peut s'adapter à la nature dynamique d'internet. En effet, la fonction « reporting » supervise toute l'activité internet des internautes permettant ainsi aux administrateurs de réseau d'obtenir des informations en temps réel sur l'usage d'internet au sein de leur organisation et d'améliorer les performances du système de filtrage.

- **Microsoft**

Microsoft² est l'un des acteurs majeurs des secteurs du filtrage antihameçonnage ou « antiphishing » et antispam.

La solution de filtrage antihameçonnage qu'elle propose est une fonctionnalité d'Internet Explorer permettant de détecter les sites web d'hameçonnage.

Ce filtre antihameçonnage s'exécute en arrière-plan lorsque l'internaute navigue sur le web.

Il fonctionne sur un principe de filtrage par liste noires ou « blacklists » et opère comme suit :

- il compare l'adresse des sites web que vous visitez à une liste de sites présentés à Microsoft comme légitimes, cette liste étant stockée sur votre ordinateur ;

¹ <http://education.optenet.fr/3-1/index.htm>.

² <http://www.windowshelp.microsoft.com/Windows/fr-FR/Help/1a460290-632a-4fb1-b50b-4df7e40771c41036.mspx>

- il analyse les sites visités pour détecter toute caractéristique propre à un site web d'hameçonnage ;
- avec l'accord de l'utilisateur, il envoie certaines adresses de sites à Microsoft afin qu'elles soient confrontées plus minutieusement à la liste de sites d'hameçonnage d'ores et déjà signalés et ainsi régulièrement mise à jour.

Outre un logiciel de filtrage antihameçonnage, Microsoft propose des solutions antiphishing et antispam, notamment le logiciel « Outlook SP2 ».

Ce filtre « Courrier indésirable » analyse chaque message électronique, recherche et détecte le contenu suspect ou frauduleux et les caractéristiques associées aux spams et aux messages de phishing.

o **GFI Software**

Fondée en 1992, GFI Software¹ est l'un des acteurs du marché de la réalisation de logiciels et fournit, notamment, une seule source intégrée destinée à permettre aux administrateurs de résoudre les problèmes de sécurisation de réseaux, de contenus et de messageries.

GFI est une entreprise internationale qui possède des bureaux à Malte, Londres, Hong-Kong, Hambourg...

Elle est l'un des partenaires de Microsoft et est membre certifié du partenariat Microsoft Gold Certified Partner.

2.1.2.2 Les principaux éditeurs, producteurs et intégrateurs de logiciels de filtrage de contenus vidéo et/ou audio

o **Audible Magic**

Audible Magic² est une société de droit californien qui fournit des solutions de gestion de contenus numériques, des CMS (Content Management Services) et des services d'antipiratage pour les entreprises du secteur des media et du divertissement en ligne, ainsi que pour d'autres acteurs économiques, gouvernementaux ou encore éducatifs.

Elle édite, notamment, des logiciels de filtrage de contenus qu'elle propose dans le domaine très spécifique des contenus vidéo.

Elle propose ainsi un logiciel intitulé « Content Identification Services » qui met en œuvre une technologie de type « fingerprinting » consistant à analyser un contenu vidéo en référence à une base de données d'« empreintes digitales » des versions originales protégées par le droit de la propriété intellectuelle. L'analyse du contenu est notamment réalisée grâce au spectre sonore de la vidéo. Autrement dit, il s'agit d'une solution de filtrage de contenus vidéo basée essentiellement sur une technologie de filtrage de contenus audio.

¹ <http://www.gfsfrance.com>

² <http://www.audiblemagic.com/company/about.asp>

Cette technologie est celle qu'ont adoptés les principaux acteurs du marché de la fourniture d'espaces de stockage vidéo et/ou audio, tels que Dailymotion, MySpace, Microsoft avec Soapbox, Google avec You Tube, Break.com, Eyespot, GoFish, Grouper, étant précisé que certains d'entre eux, notamment Dailymotion et Google, ont complété leur système de filtrage en recourant à des solutions complémentaires développées par l'Ina s'agissant de Dailymotion et en interne en ce qui concerne Google.

Audible Magic propose, par ailleurs, des solutions de filtrage de contenus audio, le cas échéant appliquée à la vidéo, notamment la solution « CopySense Appliance ».

S'agissant du filtrage de contenus, « CopySense Appliance »¹ agit à différents niveaux et permet de filtrer le trafic peer-to-peer, en bloquant l'échange :

- soit de tous les fichiers peer-to-peer ;
- soit des fichiers dont le contenu serait protégé par le droit de la propriété intellectuelle ;
- soit des fichiers pédophiles ;
- soit des fichiers pornographiques

Concrètement, « CopySense Appliance » s'installe sur les routeurs d'un réseau ou sur les passerelles qui mènent à l'internet ; il crée une copie de l'ensemble du trafic, identifie les paquets de données qui utilisent le FTP ou la technologie peer-to-peer, et recrée les fichiers échangés pour les identifier.

Cette technologie peut concerner tout aussi bien les fichiers de type MP3 que les fichiers échangés via une messagerie instantanée.

Elle a été adoptée par Sony Music.

Audible Magic est représentée en France en exclusivité par Ayala Europe.

○ **KDDI**

KDDI, second opérateur de téléphonie mobile au Japon, a annoncé au mois d'octobre 2007 avoir mis au point une technologie qui permettrait d'analyser les contenus disponibles sur les sites de partage de vidéos afin d'identifier les vidéos filmées par des amateurs et les vidéos originales protégées par le droit de la propriété intellectuelle².

Selon KDDI, le taux de réussite de cette nouvelle technologie s'élèverait à 98 %³.

○ **Google**

Outre le logiciel de filtrage développé par Audible Magic, YouTube utilise, depuis le mois octobre 2007, un nouveau logiciel de filtrage de contenus vidéo, le « Video Identification », que Google a développé en interne⁴.

¹ <http://www.audiblemagic.com/products-services/copysense/>

² <http://www.audiblemagic.com/products-services/copysense/>.

³ <http://www.audiblemagic.com/products-services/copysense/>.

⁴ Annexe 4, Bibliographie commentée, n°1.3 et n°1.35.

Selon Google, ce nouveau logiciel permet :

- de bloquer le renvoi sur les sites des clips ayant fait l'objet d'une censure préalable ;
- de supprimer certains comptes ;
- de limiter la longueur maximale des vidéos postées à 10 minutes ;
- d'opérer une analyse préalable du programme vidéo visant à déterminer si son contenu est ou non protégé par le droit de la propriété intellectuelle.

Google considère que son logiciel n'est pas fiable à 100 % et espère un taux de réussite de l'ordre de 80 à 90 %¹.

○ **Advestigo**

Advestigo², société française, est l'un des acteurs majeurs de la « Protection d'Actifs Numériques ».

Elle propose notamment un logiciel « Advestigate »³ qui permet de détecter si un fichier, notamment vidéo, est ou non protégé par le droit de la propriété intellectuelle et ainsi de le distribuer légalement sur les sites de partage de contenus numériques.

La technologie mise en œuvre par ce logiciel ne s'appuie ni sur des métadonnées, ni sur des watermarks, ni sur des hashes de fichiers.

Il s'agit d'une technologie de calcul d'empreintes numériques, la « Théraographie », qui permet de reconnaître des copies exactes ou approchées, totales ou partielles d'un contenu original.

La technologie fonctionne pour tous types de formats, audio ou vidéo, et est relativement indifférente aux diverses transformations dont les fichiers vidéo peuvent faire l'objet : extraits, flou, altération de la couleur, accélération ou ralentissement, insertion de sous-titres.

○ **Thomson**

A l'origine exclusivement fabricant de matériel audiovisuel pour le grand public, le groupe Thomson⁴ est aujourd'hui devenu fournisseur de solutions professionnelles avec notamment les créateurs de contenu ou les diffuseurs numériques.

Il est d'ailleurs titulaire de brevets dans le domaine du MP3 et intervient dans le secteur des Digital Rights Management (DRM) dont l'objet est de sécuriser les fichiers numériques tant vidéo qu'audio.

¹ Annexe 4, Bibliographie commentée, n°1.3.

² <http://www.advestigo.com/>

³ Annexe 4 : Bibliographie commentée, n°1.27.

⁴ <http://www.thomson.net/GlobalEnglish/Products/content-tracking-and-security/nexguard/Pages/default.aspx>

C'est dans ce cadre que le groupe Thomson a lancé, pour Microsoft Windows Media Video 9, une solution de marquage numérique « Nexguard »¹.

Cette solution repose donc sur une technologie de type « watermarking ».

Cette technologie est mise en œuvre par les outils « Embedder » et « Investigator » :

- le logiciel « Embedder » ajoute un code invisible aux fichiers vidéo à la demande ou en streaming ;
- le logiciel « Investigator » permettant, lorsque la copie illégitime de fichier est découverte, d'extraire l'information codée pour remonter à son origine.

Selon Thomson, cette technologie résisterait aux altérations et transformations dont les fichiers peuvent faire l'objet.

o **L'Institut national de l'audiovisuel (Ina)**

Mis en place le 6 janvier 1975, l'Ina est un organisme français ayant le statut d'Etablissement Public de l'Etat à caractère Industriel et Commercial ou Epic.

L'Ina a développé une technologie de type « fingerprinting », intitulée « Signature » et visant à filtrer le contenu vidéo².

Cette technologie, par opposition aux technologies de tatouage de type « watermarking », repose sur la prise d'empreintes numériques ou signatures, chacune correspondant à une séquence d'images.

Le partenaire agréé par l'Ina pour prendre les empreintes des contenus numérisés est la société MPO eMedia.

Selon l'Ina, "le contenu qui a été protégé par un producteur ou un diffuseur en utilisant cette technologie sera automatiquement détecté et sera soit rejeté soit géré conformément aux accords avant d'être mis en ligne".

En octobre 2007, l'Ina et Dailymotion ont signé un accord de collaboration permettant à Dailymotion de mettre en œuvre cette nouvelle technologie sur l'ensemble de sa plate-forme.

Dailymotion va donc scanner chaque vidéo avant leur mise en ligne, et rechercher la présence d'une empreinte invisible à l'œil nu, laquelle sera ajoutée par l'Ina. Si une telle signature est présente, elle agira selon les ordres des ayant droits qui ont marqué leur vidéo, ces derniers pouvant :

- soit décider d'interdire totalement la diffusion de la vidéo, auquel cas Dailymotion rejettera la vidéo ;
- soit choisir de bénéficier de la mise en ligne et de partager avec la plate-forme les revenus publicitaires générés par la vidéo copiée.

¹ Annexe 4 : Bibliographie commentée, n°1.5.

² <http://www.ina.fr/entreprise/activites/recherche-audiovisuelle/signature.html>.

○ **Vivacode**

Vivacode¹ a mis au point une technologie dite de « dissémination protégée » pour les contenus numériques, notamment e-book, presse digitale, e-learning, vidéos, musique.

Elle propose une solution professionnelle de « Digital Dissemination Management » qui permet, selon Vivacode :

- « à l'utilisateur final, de recopier ses achats sur ses différentes machines et de disséminer une version « découverte » à son réseau relationnel ;
- à ce réseau de découvrir ce contenu et de le relier au site du vendeur de contenu à l'origine de la transaction, sans que la chaîne de protection ne soit rompue ».

○ **Gracernote**

Gracernote est une société américaine qui, notamment, se livre à des activités d'édition de solutions de filtrage de contenus audio.

Elle propose un filtre antipiratage².

Ce filtre met en œuvre une technologie d'identification par empreintes acoustiques.

Cette technologie d'identification par empreintes acoustiques de type « audio fingerprinting » repose, par ailleurs, sur une base de données discographiques fonctionnant par mots-clés, que le logiciel antipiratage interroge une fois l'empreinte acoustique du morceau reconnue.

Cette solution a été adoptée par des acteurs majeurs du marché de la musique, tels qu'Universal et par certains sites spécialisés dans le partage de contenus audio-vidéo, parmi lesquels MySpace³.

Cette clientèle fait de Gracernote l'un des principaux éditeurs de logiciels de filtrage de contenus audio.

○ **TDF**

TDF⁴ est une entreprise française qui propose des services de diffusion de contenus audiovisuels et réalise des prestations pour les opérateurs télécoms. Elle se développe également dans le domaine du Multimédia avec, en particulier, de la numérisation de contenus, de l'encodage, des archivages, de la diffusion de contenus sur tous supports, notamment, sur internet.

¹ <http://www.vivacode.eu/cms/index.php?lang=fr>.

² http://www.gracernote.com/gn_korea/gn_products/music_id.html.

³ Annexe 4 : Bibliographie commentée, n°1.28.

⁴ <http://www.tdf.fr/>.

TDF a développé un logiciel de filtrage de contenus audio, « Wavessence »¹, qui est, par ailleurs, commercialisé par l'une de ses filiales, TV-Radio.com.

Ce logiciel met en œuvre une technologie qui repose sur le traitement numérique de signaux au moyen duquel il est possible d'extraire une empreinte numérique représentative de temporelle de n'importe quelle forme d'onde, puis de repérer des similitudes entre les différents signaux, ce par simple comparaison entre leurs empreintes respectives.

En novembre 2007, l'Ina et TDF ont signé un contrat de partenariat visant à coordonner leurs technologies en matière de traçage de contenus audiovisuels, « Signature » pour l'Ina et « Wavessence » pour TDF².

2.1.2.3 LTU Technologies, le principal éditeur de logiciels de filtrage d'images fixes

LTU Technologies³ est une société française fondée en 1999 par des chercheurs issus du MIT Media Lab, de l'Université d'Oxford et de l'Inria (Institut national de recherche en informatique et automatique) et qui se livre à des activités d'édition de logiciels.

Elle dispose d'une assise mondiale grâce notamment à sa filiale américaine, basée à Washington, aux Etats-Unis.

Elle est l'une des premières entreprises à avoir développé des logiciels de recherche et de classification d'images dans des domaines et marchés relativement variés tels les investigations liées à la pédo-pornographie, l'intelligence économique et militaire, la protection de la propriété industrielle, les médias, les sciences de la vie et les nouvelles technologies.

Les logiciels développés par LTU Technologies sont :

- « Image-Seeker », un système complet de gestion des images avec indexation et recherche par le contenu ;
- « Image-Filter », un logiciel de gestion de contenu visuel permettant l'analyse automatique des images, autrement dit une plate-forme de classification de contenu.

Le filtre de tri dont Exalead, le moteur de recherche français, s'est doté, ce afin d'améliorer son service de recherche d'images sur internet, utilise précisément la technologie développée par LTU Technologies⁴.

¹ <http://www.tdf.fr/search/?format=builtinlong&method=boolean&sort=score&config=htdig&restrict=&exclude=&words=wavessence>.

² Annexe 4 : Bibliographie commentée, n°1.23.

³ <http://www.ltutech.com/fr/>.

⁴ Annexe 4 : Bibliographie commentée, n°1.22.

2.1.2.4 Criteo, éditeur de logiciels de filtrage de type « filtrage collaboratif »

Criteo¹ est une société française qui se livre à des activités de conseil en système informatique et édite certaines solutions logicielles, telles notamment le logiciel « Criteo ».

Elle a un bureau en France et un bureau aux Etats-Unis.

Le logiciel « Criteo » est un moteur prédictif temps réel reposant sur une technologie de type « filtrage collaboratif », développée en interne en collaboration avec des chercheurs de l'Inria.

Le principe mis en œuvre consiste à comparer les goûts des internautes entre eux, et d'en extrapoler des recommandations.

Les recommandations sont donc faites sur les goûts relatifs des utilisateurs et non sur les caractéristiques intrinsèques des produits.

2.1.3 Quelques titulaires de droits

2.1.3.1 Universal Music Group (UMG) et Warner Music Group

En 2006, UMG et Warner Music Group ont conclu avec YouTube, des accords aux termes desquels :

- YouTube s'est engagée à intégrer des technologies de filtrage de contenu non autorisé par les ayants droit ;
- UMG et Warner Music Group se sont engagées à autoriser les usagers de YouTube à accéder à des titres issus de leur catalogue respectif, le contenu pouvant, par ailleurs, être intégré gratuitement dans les vidéos générées par les usagers

2.1.3.2 Motion Picture Association of America (MPAA)

Le syndicat MPAA a conclu un accord avec le site de téléchargement et d'échange de vidéos Guba.com.

Cet accord s'est vu concrétisé par l'utilisation, par Guba.com, d'une technologie dite « propriétaire de filtrage », nommée « Johnny ».

Cette technologie consiste à analyser les vidéos, à numériser et à générer une empreinte pour chacune d'elle et aboutit à neutraliser la distribution ou l'échange de toute donnée que Johnny viendrait à reconnaître comme interdite.

2.1.3.3 Sony Music

De la même manière qu'UMG, Sony Music a conclu des accords avec YouTube et CBS en vertu desquels :

¹ <http://www.criteo.com/fr/home.aspx>

- YouTube et CBS se sont engagées à recourir à des techniques de filtrage de contenus ;
- Sony Music s'est engagée à leur fournir du contenu.

2.1.3.4 La Sacem

La Sacem a participé activement à la négociation de la charte entre les fournisseurs d'accès à internet, les professionnels de la musique et les pouvoirs publics français.

La Sacem, en la personne de son Président de directoire, Bernard Miyet, réclame qu'il soit procédé à des tests sur les différentes solutions de filtrage de contenus.

Le 6 décembre 2007, ce-dernier déclarait :

- « Le problème n'est pas de déterminer la source ou le destinataire, mais de faire bloquer tout fichier qui circulerait sans autorisation. Ce qui éviterait le problème des libertés publiques. Dès l'an 2000, nous avons lancé un dispositif de recherche *[des oeuvres illicites, NDLR]*, avec la Sacem allemande, les producteurs de disques anglais, allemands et belges, et qui nous avait permis, sur une cinquantaine d'oeuvres, de repérer facilement sur quels protocoles elles circulaient et de trouver les adresses IP. Quand j'ai vu ce système, je me suis dit qu'il y avait une solution, mais il fallait être attentif au respect des libertés publiques.

La Sacem a donc sollicité la Cnil. Cette dernière nous a dit que cela nécessitait des modifications législatives. Nous avons cessé toute opération de cette nature, sachant néanmoins que la solution existait »¹

2.1.4 Quelques acteurs de la recherche-développement identifiés en matière de filtrage de contenus

Les acteurs de la recherche-développement sont extrêmement nombreux et variés aussi bien en France, en Europe que dans le monde.

Cette dispersion est liée à la diversité des technologies fondamentales qui sont utilisées et intégrées dans les technologies « mères » du filtrage de contenu (analyse de textes, traitement de l'image, analyse du son et de la parole, analyse de la vidéo).

2.1.4.1 Quelques acteurs français

- o **L'Institut national de recherche en informatique et automatique (Inria)**

L'Inria² a l'ambition d'être au plan mondial, un institut de recherche au cœur de la société de l'information.

Placé sous la double tutelle des ministères de la Recherche et de l'Industrie, il a pour vocation d'entreprendre des recherches fondamentales et appliquées dans les domaines des sciences et technologies de l'information et de la communication (Stic). L'institut assure

¹ Annexe 4 : Bibliographie commentée, n°1.10.

² <http://www.inria.fr/>.

également un fort transfert technologique en accordant une grande attention à la formation par la recherche, à la diffusion de l'information scientifique et technique, à la valorisation, à l'expertise et à la participation à des programmes internationaux.

Jouant un rôle fédérateur au sein de la communauté scientifique de son domaine et au contact des acteurs industriels, l'Inria est un acteur majeur dans le développement des Stic en France.

L'Inria dispose de 8 centres de recherche situés à Rocquencourt, Rennes, Sophia Antipolis, Grenoble, Nancy, Bordeaux, Lille et Saclay et sur d'autres sites à Paris, Marseille, Lyon et Metz.

L'Inria développe de nombreux partenariats avec le monde industriel et favorise le transfert et la création d'entreprises (83) dans le domaine des STIC, notamment au travers de sa filiale Inria-Transfert, promoteur de 4 fonds d'amorçage: I-Source 1 et 2 (technologies de l'information et de la communication), C-Source (multimédia) et T-Source (télécommunications).

L'Inria est actif au sein d'instances de normalisation comme l'IETF, l'Iso ou le W3C dont il a été le pilote européen de 1995 à fin 2002. Enfin l'institut entretient d'importantes relations internationales : en Europe, l'Inria s'implique fortement dans le 6e PCRDT où il participe à plus de 100 actions ainsi que dans le consortium ERCIM, qui regroupe 17 organismes de recherche.

Son budget est de 162 M€ht, dont 20% proviennent de contrats de recherche et de produits de valorisation.

L'Inria est à l'origine d'un certain nombre de travaux qui se révèlent exploitables dans le domaine du filtrage de contenus, tels :

Sur le thème « Images et vidéo : perception, indexation, communication » :

- « Imedia»¹ : l'objectif est de développer des méthodes d'indexation par le contenu, de recherche interactive, et de navigation dans des bases d'images, dans un contexte multimédia. Les axes de recherche sont l'indexation d'images par le contenu, la recherche interactive dans des grandes bases d'images, la navigation et l'indexation multimédia.
- « Lagadic»² : l'objectif est de modéliser et d'élaborer des stratégies de perception et d'action autour des techniques d'asservissement visuel pour des applications dans tous les secteurs de la robotique, en vision par ordinateur, réalité augmentée, animation virtuelle et cogniscience. Les axes de recherche sont :
 - o la modélisation d'informations visuelles optimales pour les différents capteurs de vision ;
 - o la spécification et la réalisation de tâches de haut niveau en environnement complexe ;
 - o la conception d'algorithmes de traitements d'images temps réel.

¹ <http://www.inria.fr/recherche/equipes/imedia.fr.html>.

² <http://www.inria.fr/recherche/equipes/lagadic.fr.html>.

- « Temics »¹ : l'objectif est de développer les concepts et les outils d'analyse, de modélisation, de codage, et de tatouage d'images, et plus généralement des informations vidéo manipulées en communication multimédia. Les axes de recherche sont :
 - l'analyse et modélisation de séquences vidéo ;
 - le codage conjoint source-canal ;
 - le tatouage avec la prise en compte, notamment, de l'impact mutuel du tatouage, de la représentation et de l'indexation des données au sein d'une base, notamment pour le traçage de copies illicites.

- « Vista »² : les travaux portent sur :
 - l'analyse de scènes ou de phénomènes physiques dynamiques, pour des objectifs de détection, d'interprétation et de décision sur des événements temporels, ainsi que pour des besoins de mesures ;
 - le couplage perception-commande dans des systèmes automatisés ou robotiques, pour des tâches de surveillance, de guidage et de manipulation, de navigation et d'exploration.

- « Perception »³ : l'objectif est d'interpréter des images et des vidéos en termes de représentations visuelles tridimensionnelles et de descriptions symboliques. Les principaux axes de recherche sont :
 - la modélisation d'objets et de scènes à partir de plusieurs images ;
 - la représentation de scènes visuelles dynamiques ;
 - les modèles calculatoires et la vision biologique ;
 - les réseaux de caméras, la réalité augmentée et les systèmes interactifs

- « Willow »⁴ : l'objectif est de développer des modèles géométriques, physiques, et statistiques, appropriés de toutes les composantes du processus d'interprétation des images, y compris l'illumination, les matériaux, les objets, les scènes, et les activités humaines. Les axes de recherche sont :
 - la modélisation, l'analyse, et la reconnaissance d'objets et de scènes tridimensionnels ;
 - la capture et la classification des activités humaines ;
 - la reconnaissance de catégories d'objets et de scènes.

Sur le thème « Images et vidéo : perception, indexation, communication » :

- « Metiss »⁵ : les axes de recherche sont :
 - la caractérisation, identification et vérification du locuteur ;

¹ <http://www.inria.fr/recherche/equipes/temics.fr.html>.

² <http://www.inria.fr/recherche/equipes/vista.fr.html>.

³ <http://www.inria.fr/recherche/equipes/perception.fr.html>.

⁴ <http://www.inria.fr/recherche/equipes/willow.fr.html>

⁵ <http://www.inria.fr/recherche/equipes/metiss.fr.html>.

- la modélisation, détection d'informations et indexation d'enregistrements audio ;
 - la séparation de sources et traitement avancé du son.
- « Aviz »¹ : l'objectif est d'améliorer les méthodes d'analyse et de visualisation de grandes quantités de données en intégrant profondément le processus d'analyse et celui de visualisation d'information pour permettre de comprendre plus facilement et rapidement ces données. Les axes de recherche sont :
- les méthodes de visualisation et de navigation dans des masses de données ;
 - les méthodes d'analyse et de réduction des masses de données afin de les rendre visualisables ;
 - les méthodes d'évaluation pour mesurer l'efficacité et l'utilisabilité des visualisations, navigations et analyses ;
 - les outils logiciels pour réaliser et déployer des systèmes d'analyse visuelle pouvant gérer, chercher, visualiser et analyser des masses de données avec des temps de réponse interactifs.

○ **Le laboratoire de recherche informatique de l'Ecole Télécom Lille 1 (Enic)**

Le laboratoire de recherche informatique de l'Enic² réalise des recherches dans le domaine, notamment, de la fouille de données multimédias ou « data mining ».

Dans ce cadre, l'objectif est d'étudier des solutions d'indexation et de structuration des contenus multimédia et audiovisuels (images, vidéo, 3D) pour en faciliter l'accès, voire l'exploitation des contenus. L'objectif de la recherche est aussi d'analyser les usages des contenus, afin de faciliter leur interprétation et leur exploitation dans des contextes d'usages variés.

Jean-Philippe Vandeborre³, notamment, se livre à des recherches dans le domaine de l'indexation de modèles tridimensionnels ou « modèles 3D » utilisés dans les environnements virtuels, les simulations, les jeux vidéos, etc., l'indexation de telles données consistant alors à trouver des descripteurs mathématiques, de forme notamment, invariants aux transformations géométriques comme la translation, la rotation et la mise à l'échelle, mais également indépendants du niveau de facétisation des modèles 3D.

○ **Le Laboratoire d'informatique en image et systèmes d'information (Liris)**

Liris, créé en 2003 à la suite du regroupement de plusieurs laboratoires de recherche lyonnais (Ligim, Lisi, RFV) et d'individualités du domaine des sciences et techniques de l'information et de la communication.

Il est associé au CNRS avec le label UMR 5205.

¹ <http://www.inria.fr/recherche/equipes/aviz.fr.html>.

² http://www.enic.fr/recherche/informatique_reseaux.php?id=75&&niv=2.

³ <http://www.telecom-lille1.eu/people/vandeborre/>.

Le laboratoire a deux thèmes principaux de recherche, l'image numérique et les systèmes d'information, chaque thème étant décliné selon quatre axes scientifiques :

- les connaissances et les systèmes complexes ;
 - les images et vidéos, leur segmentation et l'extraction d'informations ;
 - la modélisation et la réalité augmentée ;
 - les systèmes d'information communicants.
- **Le laboratoire Signal image communications (Sic)**

Le laboratoire Sic¹ est rattaché à l'Institut de Recherche XLIM UMR CNRS 6172 depuis le 1er janvier 2008 et est ainsi devenu le « département Sic » du Laboratoire XLIM.

Les activités du laboratoire relèvent des Sciences et Technologies de l'Information et de la Communication, et s'inscrivent en Image et en Communications sans fil.

Ses principaux axes de recherche sont :

- l'informatique géométrique et graphique avec la modélisation et l'animation d'objets géométriques;
- le traitement et l'analyse (de séquences) d'images couleur et/ou texturées avec l'évaluation de la qualité visuelle des traitements et des supports de restitution d'images ou de vidéos couleur ;
- les systèmes de communication sans fil avec l'optimisation de la qualité et/ou de la robustesse du lien radioélectrique, en particulier par l'étude de la propagation d'ondes radioélectriques.

L'équipe-projet « Icones »², notamment, concentre ses travaux de recherche autour « du traitement, de la caractérisation et de l'analyse de signaux et images multi sources et multi composantes avec une spécificité concernant les images couleur texturées statiques et dynamiques ».

Ses travaux concernent l'introduction, dans toute la chaîne des traitements de paramètres issus, d'une part, de modèles du système visuel humain, d'autre part, de modèles physiques liés aux aspects des surfaces analysées.

Ils visent également à permettre d'évaluer la qualité des chaînes de traitements, des supports de reproduction d'images ou de vidéos couleur (sur écran ou sur papier).

- **Le Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur (Limsi) du Centre national de la recherche scientifique (CNRS)**

Le Centre national de la recherche scientifique (CNRS)³ est un organisme public de recherche, un établissement public à caractère scientifique et technologique, placé sous la tutelle du ministère de l'Enseignement supérieur et de la Recherche.

¹ <http://www.sic.sp2mi.univ-poitiers.fr/liens/seminaires.php>

² <http://www.sic.sp2mi.univ-poitiers.fr/themes/icones/index.php>

³ <http://www.cnrs.fr/fr/organisme/presentation.htm>

Le CNRS comporte plus de 1 250 laboratoires (propres, mixtes ou associés), dont le Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur (Limsi).

Les thèmes de recherche de ce laboratoire¹ couvrent un large spectre disciplinaire, allant du « thermodynamique au cognitif », en passant par la mécanique des fluides, l'énergétique, l'acoustique, l'analyse et la synthèse vocale, le traitement de la langue parlée et du texte, la vision et la perception, la réalité virtuelle et augmenté etc.

Son rapport d'activités pour l'année 2007² fait notamment état de divers groupes de recherche dont les travaux intéressent la matière du filtrage, tels :

- « Groupe Langues, Information et Représentations (LIR) : les activités de recherche de ce groupe sont essentiellement consacrées au traitement des données écrites, à leur analyse, leur compréhension ou leur reproduction ainsi qu'à l'acquisition de connaissances nécessaires, principalement morphologiques et sémantiques. La quantité impressionnante de données écrites aujourd'hui disponibles électroniquement est une mine d'informations et la fouille de données dans les textes est un des enjeux majeurs de la société de l'information. Les recherches développées dans le groupe LIR s'inscrivent dans cette dynamique, avec une implication croissante dans des projets nationaux et internationaux. Les compétences variées et complémentaires des membres du groupe LIR permettent de combiner approches symboliques et statistiques, et constituent un des atouts majeurs du groupe qui participe ainsi pleinement à l'évolution du traitement des langues ;
- « Groupe Traitement du Langage Parlé (TLP) » : les recherches de ce groupe portent sur la modélisation de la parole et son traitement automatique. Pour extraire et structurer l'information présente dans un document audio, le groupe de recherche développe des modèles et des algorithmes fondés sur la prise en compte conjointe des diverses sources d'information visant à un processus global de décodage du signal. Ces recherches sur les modélisations acoustique, lexicale, et linguistique, sont réalisées dans un contexte multilingue et s'appuient sur de grands corpus oraux représentatifs de nombreux domaines applicatifs.
- « Action Transversale COPTE : Corpus Parole Texte Evaluation » : l'action transversale COPTE fait le lien entre deux domaines du Traitement Automatique du Langage Naturel : la reconnaissance de la parole et l'analyse de l'écrit. L'objectif est de fusionner les approches propres aux deux domaines sur des problèmes ouverts situés à l'interface des deux disciplines. Le domaine de la reconnaissance de la parole aborde l'analyse du langage par l'étude du signal sonore et doit donc nécessairement prendre en compte les aspects propres à la parole : temporalité et spontanéité. De son côté l'analyse de l'écrit, aborde l'analyse du langage par l'étude des signes, où les aspects qui priment sont plutôt la nature statique et préparée du support d'information étudié.
- « Action Thématique Sémantique et Mémoire Episodique » : les recherches ont été centrées sur deux thèmes : (1) l'exploration du contenu de la mémoire sémantique et la relation entre mémoire sémantique et mémoire épisodique ; (2) les conditions

¹ <http://www.limsi.fr/Scientifique/Domaines>

² Annexe 4 : Bibliographie commentée, n°3.23.

d'élaboration des inférences causales. Ces deux séries de recherches ont permis de mettre en évidence le rôle des situations stockées dans la mémoire sémantique.

○ **Le Laboratoire d'intégration des systèmes et des technologies (List) du Cea**

Situé en île de France sud (Saclay et Fontenay aux Roses), le List¹ est un centre de recherche technologique s'intéressant aux systèmes à logiciel prépondérant et organisé selon trois thématiques présentant de forts enjeux sociétaux et économiques :

- les systèmes embarqués ;
- les systèmes interactifs : ingénierie de la connaissance, robotique et réalité virtuelle et les interfaces sensorielles² ;
- les capteurs et le traitement des signaux.

Ce laboratoire a notamment participé, avec son analyseur syntaxique « Lima », à la campagne d'évaluation « Easy ».

Cet analyseur « Lima » repose sur l'implémentation d'une « grammaire de dépendance ».

Selon Romaric Besançon et Gaël Chalendar³ :

- les résultats obtenus sont encourageants ;
- néanmoins, le traitement de corpus plus généraux couvrant des phénomènes syntaxiques plus variés nécessiteront très probablement des ressources financières supplémentaires ou la mise en place de traitements particuliers.

2.1.4.2 Quelques acteurs internationaux

Les acteurs de la recherche-développement sont très nombreux au niveau mondial et on y retrouve naturellement les grandes institutions.

Dans le présent document nous n'avons cité que le MIT et l'université de Harvard mais la liste des centres de recherche travaillant sur les technologies permettant de réaliser du filtrage est extrêmement vaste et ne se réduit en aucune manière à ces deux acteurs.

○ **Le Massachusetts Institute of Technology (MIT)**

Le MIT⁴ est une université ainsi qu'un important centre de recherche, situé à Boston, aux Etats-Unis.

Le MIT comporte de nombreux laboratoires, parmi lesquels, notamment :

- « the Media Laboratory » ;
- « the Computer Science and Artificial Intelligence Laboratory (CSAIL) ».

¹ http://www-list.cea.fr/fr/presentation/presentation_list.htm

² Annexe 4 : Bibliographie commentée, n°3.16.

³ Annexe 4 : Bibliographie commentée, n°3.6.

⁴ <http://www.mit.edu/>

Ces laboratoires comportent eux mêmes des groupes de recherche, dédiés à un domaine très spécifique, tel que, par exemple, le « Networks and Mobile Systems » consacré aux questions portant sur les télécommunications sans fil.

○ **L'université de Harvard**

Le « Berkman Center for Internet & Society »¹ est un programme de recherche piloté par l'université d'Harvard.

Cet organisme s'est intéressé à la question du filtrage de contenus, au travers, notamment, d'articles plus généraux² portant sur :

- l'interopérabilité de l'internet ;
- le second forum de la gouvernance d'internet ou « Internet Governance Forum », qui s'est tenu à Rio de Janeiro courant l'année 2007 ;
- l'influence des technologies sur le débat politique.

2.2 LES PRINCIPAUX OUTILS DE FILTRAGE

2.2.1 Les principaux outils commerciaux de filtrage de contenus

2.2.1.1 Les systèmes assimilables à des systèmes de filtrage : les systèmes des « adwords » et des « pages' ranking » utilisés par le moteur de recherche Google

« Adwords »³ est le nom du système publicitaire de Google, qui vise à afficher des annonces-texte ciblées. Les annonceurs paient lorsque l'internaute clique sur la publicité selon un système d'enchère et de qualité : plus le prix au clic est élevé et plus l'annonce est pertinente pour l'utilisateur, plus l'annonce est en évidence.

Sans être un système de filtrage de contenus en tant que tel, le système des « adwords » s'en rapproche puisqu'il consiste en un filtrage par mots-clés permettant un ciblage de la requête afin de pouvoir associer à cette requête les publicités qui seraient en lien avec elle.

Le système des « pages' ranking »⁴ est le système au moyen duquel Google attribue à chaque page web, le score qui détermine, par la suite, son positionnement.

De la même manière que le système des « adwords », ce système des « pages' ranking », sans constituer un véritable système de filtrage de contenus, repose sur des principes assimilables à ceux qui fondent les technologies ou systèmes de filtrage de contenus.

Dans ce système, en effet, le score attribué à chaque page est notamment fonction :

¹ <http://cyber.law.harvard.edu/> ;

<http://cyber.law.harvard.edu/about>

² Annexe 4 : Bibliographie commentée, n°3.5.

³ Annexe 4 : Bibliographie commentée, n°1.32.

⁴ Annexe 4 : Bibliographie commentée, n°1.36.

- de la date du document ;
- de la fréquence avec laquelle le contenu du document est changé ;
- de la manière dont un document est choisi parmi les résultats de la requête ;
- de la mise à jour du document, repérable notamment au travers de la modification du texte des ancrés ;
- du sujet du document, etc.

2.2.1.2 Les principaux outils de filtrage antimalware (antispam, antihameçonnage ou « antiphishing » et antispyware)

o Les logiciels de filtrage antispam de type « bayésien »

Les techniques utilisées par la plupart des logiciels antispam sont statiques. La protection que ces logiciels offrent risque d'être contournée relativement facilement.

Pour pallier ce risque, certains logiciels antispam mettent en œuvre une technologie dynamique, fondée sur le théorème mathématique élaboré par Thomas Bayes. On parle alors de filtrage bayésien¹.

Le filtrage bayésien repose sur le principe selon lequel la plupart des événements dépendent les uns des autres de sorte que la probabilité qu'un événement se répète dans le futur peut être déduite de la survenance, dans le passé, de ce même événement.

La mise en œuvre d'un filtre bayésien suppose que l'utilisateur du filtre élabore, au préalable, une base de données de mots-clés et de signes, notamment d'adresses URL, qu'il tire généralement de l'analyse d'échantillons composés, respectivement, de spams, courriers indésirables, et de hams, courriers valables.

Une valeur de probabilité doit ensuite être accordée à chaque mot ou signe contenu dans la matrice. Cette valeur est le produit de calculs prenant en compte le nombre de fois où un mot est apparu dans un spam par différence avec le nombre de fois où ce même mot est apparu dans un ham.

A l'arrivée de chaque message, le filtre bayésien sépare les mots les plus importants du message aux fins de l'identification du message en spam ou en ham, de ceux qui le sont moins, cette appréciation se réalisant au moyen d'une interrogation de la base de données spam et ham qui aura été préalablement élaborée. A partir des mots reconnus comme les plus importants, le filtre bayésien est, dès lors, en mesure d'identifier si le message analysé est ou non un spam.

Les points forts de la méthode bayésienne résident dans le fait :

- qu'elle prend en compte l'intégralité du message analysé, reconnaissant outre les mots-clé identifiant un spam, les mots-clé qui dénotent que le message analysé n'en est pas un ;

¹ Annexe 4 : Bibliographie commentée, n°4.2, 4.3 et 4.4.

- qu'elle permet aux filtres antispam qui la mettent en œuvre, de s'adapter aux éventuelles tentatives de contournement de certains spammeurs consistant, notamment, en la dénaturation de certains mots afin de neutraliser la base de données ;
- qu'elle est sensible à l'utilisateur avec, notamment, une prise des habitudes de ce dernier ;
- qu'elle est multilingue et internationale.

Selon un article de la BBC paru en mai 2003, cette approche bayésienne serait d'ailleurs efficace à près de 99,7 %.

Nombreux sont donc les logiciels de filtrage anti-spam qui mettent en œuvre cette méthode, tels, par exemple, le logiciel d'Outlook ou l'Internet Message Filter du serveur Exchange.

Outre ces logiciels de filtrage, d'autres logiciels de filtrage auxquels l'essentiel des moteurs de recherche, notamment Google, ont recours, visent à filtrer les sites spammés.

Dans ce cadre, la méthode utilisée reste sémantique puisqu'elle consiste à associer un « taux de spam » aux mots-clés de la requête de l'internaute, puis à comparer ce taux à un « taux de probabilité de spam » associé au contenu des sites indexés dans la base de données du moteur de recherche utilisateur du filtre.

o **Les logiciels de filtrage antimalware proposés par Websense**

« Websense ThreatSeeker » est une technologie qui vise à offrir une protection préventive contre les menaces de sécurité en ligne.

« Websense ThreatSeeker » intervient en amont, recherche les menaces sur internet avant que les clients ne soient victimes d'une attaque et protège les clients avant la création de correctifs et de signatures.

Websense indique, sur son site web, que cette technologie :

- « - utilise plus de 100 processus et systèmes propriétaires pour déchiffrer les menaces émergentes complexes ;
- utilise une combinaison d'algorithmes mathématiques, de profilage du comportement, d'analyse de codes ainsi qu'un vaste réseau de machines d'exploration de données ou « data mining » ;
- fournit aux logiciels de sécurité Websense des informations continues concernant les menaces protégeant les clients en quelques minutes »¹.

¹ <http://www.websense.com/global/en/ProductsServices/ThreatSeeker/>

« Websense Email Security »¹, « Websense Hosted Email Security »² et « Websense Web Security »³ sont des logiciels de filtrage destinés à protéger l'utilisateur d'une boîte de messagerie électronique contre les menaces, en entrée et sortie, de type « spam », « virus », « pertes de données » etc.

Ces logiciels s'adressent, essentiellement, à une clientèle d'entreprises.

« Websense Express »⁴ est, quant à elle, une solution globale de filtrage de contenus qui, néanmoins, comporte une fonctionnalité antimalware.

- **Les logiciels et fonctionnalités de filtrage antispam et antihameçonnage proposés par Microsoft**

Internet Explorer est doté d'une fonctionnalité permettant de détecter les sites Web d'hameçonnage ou de « phishing ».

Ce filtre antihameçonnage ou « antiphishing » s'exécute en arrière-plan lorsque l'internaute navigue sur le Web.

Il fonctionne sur un principe de filtrage par listes noires ou « blacklists » et opère comme suit :

- il compare l'adresse des sites Web que vous visitez à une liste de sites présentés à Microsoft comme légitimes, cette liste étant stockée sur votre ordinateur ;
- il analyse les sites visités pour détecter toute caractéristique propre à un site Web d'hameçonnage ;
- avec l'accord de l'utilisateur, il envoie certaines adresses de sites à Microsoft afin qu'elles soient confrontées plus minutieusement à la liste de sites d'hameçonnage d'ores et déjà signalés et ainsi régulièrement mise à jour.

Le filtre antihameçonnage de Microsoft :

- bloque uniquement les sites que des réviseurs de chez Microsoft, ou des employés de fournisseurs de données tiers, ont vérifiés être des sites web d'hameçonnage ;
- et offre un système de commentaires basé sur le web, permettant aux utilisateurs et aux propriétaires de sites web de signaler toute erreur le plus rapidement possible, les rapports étant ensuite vérifiés par Microsoft afin, le cas échéant, de corriger et compléter la liste noire.

Par ailleurs, le logiciel « Outlook SP2 »⁵ de Microsoft bénéficie d'une fonctionnalité de protection contre le phishing qui s'associe au filtre « Courrier Indésirable » amélioré lequel vise à protéger les internautes contre les menaces de type « spam ».

¹ <http://www.websense.com/global/fr/ProductsServices/EmailSecurity/index.php>

² <http://www.websense.com/global/fr/ProductsServices/HostedEmailSecurity/index.php>

³ <http://www.websense.com/global/fr/ProductsServices/HostedWebSecurity/>

⁴ <http://www.websense.com/global/fr/ProductsServices/Express/index.php>

⁵ http://www.microsoft.com/switzerland/athome/fr/security/email/outlook_sp2_filters.mspx.

Comme indiqué ci-avant, ce logiciel « Outlook SP2 » met en œuvre la méthode bayésienne dont les principes ont été précédemment décrits.

S'agissant du filtre « Courrier Indésirable », celui-ci analyse chaque message électronique entrant, à la recherche de contenu suspect ou frauduleux et de caractéristiques qui sont associées aux spams et aux messages de phishing, et procède de la manière suivante :

- soit les messages ont été falsifiés de telle manière à ce qu'ils apparaissent provenir d'un expéditeur légitime, auquel cas les messages frauduleux sont envoyés dans le dossier « courrier indésirable » qui les convertit automatiquement au format texte brut, inoffensif, et qui désactive tous les liens contenus dans ces messages ;
- soit les messages contiennent des liens vers des sites web soupçonnés de phishing conçus pour à des sites légitimes, auquel cas les messages suspects restent dans la boîte de réception avec toutefois désactivation des liens qu'ils comportent.

○ **Le logiciel « Proventia Mail Filter » proposé par ISS**

« Proventia Network Mail Filter », proposée par IBM via ISS, est une solution globale antispam et de sécurité de l'e-mail qui surveille le contenu du trafic e-mail afin d'éliminer le spam et de bloquer le contenu indésirable ou illégal.

IBM indique, sur son site web, que :

- « - Proventia Network Mail Filter allie des techniques d'analyse perfectionnées à une base de données qui répertorie plus de 200 000 exemples de spam et plus de 20 millions de sites Web ;
- les courriers électroniques inoffensifs sont identifiés et transférés instantanément tandis que le courrier au contenu indésirable est automatiquement bloqué »¹.

« Proventia Mail Filter » est disponible sous la forme d'un logiciel autonome ou d'un module en option de « Proventia Network Multi-Functional Security ».

○ **Les principaux logiciels de filtrage antimalware proposés par Secure Computing**

« Smartfilter »² et « Smartfilter Bess edition » sont des solutions de filtrage antispam et anti-phishing, de manière plus générale, anti-malware.

Ces solutions reposent sur la technologie « TrustedSource »³ qui est une technologie de filtrage d'URL, autrement dit une technologie de filtrage texte.

Le filtrage est réalisé au moyen d'une liste noire ou « blacklist ».

« Webwasher »⁴ est également une solution de filtrage antimalware, visant à offrir une sécurité tant en entrée qu'en sortie.

¹ <http://www-935.ibm.com/services/fr/index.wss/offering/gts/f1027049>.

² <http://www.securecomputing.com/index.cfm?skey=85>.

³ <http://www.securecomputing.com/index.cfm?skey=233&lang=fr>

⁴ <http://www.securecomputing.com/index.cfm?skey=22&lang=en>

2.2.1.3 Les principaux outils de filtrage de contenus texte

○ Le logiciel « Ilog JRules » proposé par Ilog

Le logiciel « Ilog JRules »¹ est l'un des principaux logiciels proposés par Ilog.

Il s'agit d'un système de gestion de règles métier (en anglais, business rule management system ou BRMS) pour environnement Java.

Ce logiciel « Ilog JRules » comporte des modules et fonctionnalités de filtrage de contenus texte dont Ilog décrit le fonctionnement dans le guide d'utilisation du logiciel².

Deux types de technologies d'analyse de contenus sont ainsi mises en œuvre :

- une technologie de type « syntaxique » (analyse de chaînes de caractères et de mots-clés ;
- une technologie de type « sémantique ».

La technologie de type « syntaxique » permet de filtrer les éléments de vocabulaire et les types d'expressions issus d'un ensemble syntaxique abstrait et reconnu par le système de filtrage comme valides.

La technologie de type « sémantique », optionnelle, permet d'assurer la complétude du système de filtrage au moyen d'une analyse de contenus au travers d'une analyse du contexte entourant les éléments de vocabulaires ou expressions analysés par le filtre.

○ Les principaux logiciels proposés par Websense

« Websense Enterprise » est une solution de filtrage de contenus.

Websense indique, sur son site web, que :

- ce logiciel « fonctionne à 100% de ses capacités de 50 à 250 000 utilisateurs et est conçu pour des réseaux de pratiquement n'importe quelle configuration ».

Trois options de déploiement sont disponibles:

- déploiement intégré sur un serveur distinct étroitement intégré à la plateforme de la passerelle réseau afin d'offrir un filtrage "pass-through" qui optimise la stabilité, l'extensibilité et la performance ;
- déploiement incorporé à une appliance ou un produit de la passerelle afin de réduire les dépenses matérielles et d'améliorer la facilité d'utilisation, notamment à distance ;

¹ <http://www.ilog.fr/products/jrules/>

² <http://www.ilog.com/products/jrules/documentation/jrules67/globalfiles/data.html>

- déploiement autonome utilisant un agent réseau afin de délivrer des capacités de filtrage "pass-by" dans n'importe quel environnement réseau »¹.

« Websense Express » est une solution de filtrage qui s'adresse principalement aux entreprises de moyenne taille lesquelles entendent protéger leurs employés contre les risques liés à l'usage d'internet.

Elle vise à interdire l'accès aux contenus inappropriés, et comme indiqué ci-avant, comporte une fonctionnalité antimalware qui vise à bloquer les menaces de sécurité avant qu'elles ne puissent infecter le système d'information de l'utilisateur du logiciel.

« Websense Content Protection Suite » est une solution de filtrage de contenus reposant sur la technologie « PreciseID » de Websense.

Il s'agit d'une solution d'identification et de classification des informations, destinée à offrir une protection contre les fuites d'informations.

Websense indique, sur son site web, que :

- « la technologie Precise ID, brevetée, exploite de multiples méthodes de détection afin d'aider les entreprises à découvrir, contrôler et protéger des données, incluant :
 - la technologie de l'empreinte ;
 - des règles ;
 - des lexiques ;
 - des dictionnaires ;
 - un appariement exact et partiel ;
 - une analyse statistique »².

La technologie d'empreinte ou de type « fingerprinting », mise en œuvre par « Precise ID » :

- « génère une « empreinte des informations », représentation mathématique d'un groupe de caractères, de mots, de phrases ou de champs de données d'un document, d'un message ou d'une base de données et identifie avec précision les données sensibles et leurs métadonnées »³.

Elle permet ainsi au logiciel « Websense Content Protection Suite » de protéger tout type de données qu'elles soient désactivées, utilisées ou encore dynamiques.

Selon Websense :

- « Contrairement aux méthodes de simple hachage ou d'appariement exact ou partiel, les algorithmes avancés de la technologie PreciseID utilisent des techniques d'appariement des données granulaires. Ne requiert aucune modification ni aucun repérage initial des données originales. La technologie PreciseID est optimisée

¹ <http://www.websense.com/global/fr/ProductsServices/WebsenseEnterprise/index.php>

² <http://www.websense.com/global/fr/ProductsServices/PreciseID/>

³ <http://www.websense.com/global/fr/ProductsServices/PreciseID/>

pour les applications en temps réel et protégée contre la rétro-ingénierie du contenu sensible et la manipulation des données »¹.

- **Le logiciel « Proventia Web Filter » proposé par ISS**

« Proventia Web Filter » est une solution de filtrage d'URL, fonctionnant, selon la fiche produit fournie par ISS², sur la base d'un grand nombre d'URL filtrées (60 millions) et fournissant un nombre conséquent de rapports.

- **Le logiciel « Olféo » proposé par Olféo**

« Olféo » est une solution de filtrage d'URL qui repose sur une technologie qu'Olféo a intitulée « le filtrage dynamique ».

Cette technologie consiste à identifier chaque site qui ne serait pas reconnu la première fois par l'outil de filtrage afin de le renvoyer vers l'éditeur pour qu'il effectue un classement immédiat.

Olféo indique, dans la fiche produit³, que cette technologie présente le principal avantage de personnaliser l'outil de filtrage à l'utilisation propre de l'entreprise utilisatrice ; si celle-ci utilise spécifiquement des sites dans un secteur d'activité particulier ou à une région du monde précise ou encore que certains utilisateurs ont des centres d'intérêt particuliers, l'outil de filtrage s'y adapte.

- **Le logiciel « Web Filter » proposé par Optenet**

« Web Filter »⁴ est une solution de filtrage d'URL.

Ce logiciel n'est pas uniquement destiné aux parents dans le cadre du contrôle parental.

Dans ce cadre néanmoins, il permet aux parents de contrôler l'accès de leurs enfants à des contenus internet qui seraient jugés inappropriés. Il leur permet, notamment, d'établir les horaires où la navigation est permise, d'établir un nombre maximum d'heures de navigation ou de limiter le téléchargement de fichiers.

Cette solution dispose d'une fonction « reporting » qui supervise toute l'activité internet des internautes et qui, par là-même, permet aux administrateurs de réseau d'obtenir des informations en temps réel sur l'usage d'internet au sein de leur organisation et d'améliorer les performances du système de filtrage.

2.2.1.4 Les principaux logiciels de filtrage de contenus vidéo

- **Le logiciel « Content Identification Services » proposé par Audible Magic**

« Content Identification Services »⁵ est un logiciel de filtrage de contenus audio, mais avant tout, de contenus vidéo.

¹ <http://www.websense.com/global/fr/ProductsServices/PreciseID/>

² http://www.idepro.fr/upload/Fournisseur/ISS/ISS_Proventia_WebFilter.pdf

³ <http://www.olfeo.com/pdf/offre.pdf>

⁴ <http://www.optenet.com/fr/webfilter.asp?c=1>

⁵ <http://www.audiblemagic.com/products-services/contentsvcs/>

Ce logiciel repose sur une technologie de type « fingerprinting » laquelle consiste à analyser un contenu vidéo en référence à une base de données d'« empreintes digitales » des versions originales protégées par le droit de la propriété intellectuelle.

Il convient de relever que l'analyse des contenus est notamment réalisée au moyen d'une analyse du spectre sonore de la vidéo.

En définitive, cette solution de filtrage de contenus vidéo met en œuvre une technologie de filtrage de contenus audio.

- **Le logiciel « Signature » de l'Ina**

« Signature »¹ est une solution de filtrage de contenus vidéo.

Elle met en œuvre une technologie de type « fingerprinting » : des empreintes numériques ou signatures sont réalisées, chaque empreinte correspondant à une séquence d'images.

2.2.1.5 Les principaux logiciels de filtrage de contenus audio

- **Le logiciel « Gracenote Music ID » proposé par Gracenote**

« Gracenote Music ID »² est une solution de filtrage de contenus audio.

Cette solution met en œuvre une technologie de type « audio fingerprinting », autrement dit d'identification par empreintes acoustiques.

Cette technologie s'exécute au moyen d'une base de données discographiques laquelle fonctionne par mots-clés, que le logiciel antipiratage interroge une fois l'empreinte acoustique du morceau reconnue.

- **Le logiciel « Wavessence » proposé par TDF**

« Wavessence »³ est une solution de filtrage de contenus audio.

Ce logiciel met en œuvre une technologie qui repose sur le traitement numérique de signaux au moyen duquel il est possible d'extraire une empreinte numérique représentative de n'importe quelle forme d'onde, puis de repérer des similitudes entre les différents signaux, ce par simple comparaison entre leurs empreintes respectives.

2.2.1.6 Les principaux logiciels de filtrage de contenus audio ou vidéo

- **Le logiciel « CopySense Appliance » proposé par Audible Magic**

« CopySense Appliance »⁴ est une solution de filtrage de contenus audio ou vidéo.

¹ <http://www.ina.fr/entreprise/activites/recherche-audiovisuelle/signature.html>

² http://www.gracenote.com/gn_korea/gn_products/music_id.html

³ <http://www.tdf.fr/multimedia/les-offres-tdf/tvradiocom-moteur-de-recherche-audiovisuel/>

⁴ <http://www.audiblemagic.com/products-services/copysense/>

Elle agit à différents niveaux et permet de filtrer le trafic peer-to-peer, en bloquant l'échange :

- soit de tous les fichiers peer-to-peer ;
- soit des fichiers dont le contenu serait protégé par le droit de la propriété intellectuelle ;
- soit des fichiers pédophiles ;
- soit des fichiers pornographiques

Concrètement, « CopySense Appliance » s'installe sur les routeurs d'un réseau ou sur les passerelles qui mènent à l'internet ; il crée une copie de l'ensemble du trafic, identifie les paquets de données qui utilisent le FTP ou la technologie peer-to-peer, et recrée les fichiers échangés pour les identifier.

○ **Le logiciel « Advestigate » proposé par Advestigo**

« Advestigate »¹ est une solution de filtrage de contenus audio ou vidéo qui permet de détecter si un fichier, notamment vidéo, est ou non protégé par le droit de la propriété intellectuelle et ainsi de le distribuer légalement sur les sites de partage de contenus numériques.

La technologie mise en œuvre par ce logiciel ne s'appuie ni sur des métadonnées, ni sur des watermarks, ni sur des hashes de fichiers.

Il s'agit d'une technologie de calcul d'empreintes numériques, de type « fingerprinting » donc, la « Théraographie », qui, selon Advestigo, permet de reconnaître des copies exactes ou approchées, totales ou partielles d'un contenu original.

Cette technologie, selon Advestigo, serait relativement indifférente aux diverses transformations dont les fichiers vidéo peuvent faire l'objet : extraits, flou, altération de la couleur, accélération ou ralentissement, insertion de sous-titres.

○ **Les logiciels « Nexguard », « Embedder » et « Investigator » de Thomson**

« Nexguard »² est une solution de filtrage de contenus audio ou vidéo.

Cette solution repose sur une technologie de type « watermarking ».

Cette technologie est mise en œuvre par les outils « Embedder » et « Investigator » :

- le logiciel « Embedder » ajoute un code invisible aux fichiers vidéo à la demande ou en streaming ;
- le logiciel « Investigator » permet, lorsque la copie illégitime de fichier est découverte, d'extraire l'information codée pour remonter à son origine.

Selon Thomson³, cette technologie résisterait aux altérations et transformations dont les fichiers peuvent faire l'objet.

¹ <http://www.advestigo.com/protection.php?men=2&rub=2&LANG=FRA>

² <http://www.thomson.net/GlobalEnglish/Products/content-tracking-and-security/nexguard/Pages/default.aspx>

³ <http://www.thomson.net/GlobalEnglish/Products/content-tracking-and-security/nexguard/Pages/default.aspx>

2.2.1.7 Les principaux logiciels de filtrage de contenus image fixe, proposés par LTU Technologies

LTU Technologies indique, sur son site web¹ :

- avoir « développé un système temps-réel qui permet à la machine à voir, comprendre et expliciter le contenu visuel » (photographie, dessin, illustration, document visuel numérique), lequel reposerait sur un analyseur à forte sensibilité capable d'indexer, de reconnaître et de comparer des images à partir de leurs composantes visuelles ;
- et que sa technologie permet de distinguer les images dupliquées, les images clones et les images similaires.

Le processus d'analyse de l'image s'effectue en trois étapes :

- la segmentation de l'image ;
- l'indexation laquelle consiste à extraire, pour chaque image préalablement segmentée, un identifiant, également appelé signature numérique ou ADN de contenu ;
- le déchiffrement du contenu de l'ADN numérique au moyen de différents modules experts lesquels interrogent la base de données afin d'opérer la comparaison indispensable à l'identification du contenu.

Cette technologie est mise en œuvre par deux logiciels :

- « Image-Seeker » ;
- « Image-Filter ».

« Image-Seeker »² est un système complet de gestion des images avec indexation et recherche par le contenu et qui se caractérise par une interface web entièrement configurable, par une recherche par le texte et par l'image et par une architecture-serveur supportant de grands volumes.

« Image-Filter »³ est un logiciel de gestion de contenu visuel permettant l'analyse automatique des images, autrement dit une plate-forme de classification de contenu

2.2.1.8 Le logiciel « Criteo », un exemple de logiciel de filtrage de type « filtrage collaboratif »

« Criteo » est un moteur d'analyse comportementale qui vise à permettre à ses utilisateurs, principalement des sites marchands, de recommander à leurs clients, visiteurs de leurs sites, les produits susceptibles de leur plaire, en fonction du comportement que ces derniers auront adopté sur le site.

¹ <http://www.ltutech.com/fr/technologie-et-produits.technologie.html>

² <http://www.ltutech.com/fr/technologie-et-produits.image-seeker.html>

³ <http://www.ltutech.com/fr/technologie-et-produits.image-filter.html>

A cette fin, le moteur analyse les différentes affinités des visiteurs que ces derniers ont exprimé soit de manière explicite (notation des produits ou articles), soit de manière implicite (clics, pages vues, produits achetés etc.). Le moteur calcule, ensuite, au moyen d'un algorithme prédictif, les personnes les plus en affinités afin de pouvoir déduire de ce qu'une personne est susceptible d'apprécier ce produit puisque la personne avec laquelle elle est en affinités l'a d'ores et déjà apprécié et acheté¹.

La technologie ainsi mise en œuvre par ce logiciel « Criteo » est une technologie de filtrage de type « filtrage collaboratif »².

Cette technologie est le résultat de 3 années de développement en partenariat avec l'INRIA.

Criteo indique, sur son site web³ :

- « Qu'en plus de sa facilité d'installation et d'utilisation; la solution « Criteo » offre les fonctionnalités suivantes :
 - le processus de recommandation est entièrement automatisé, il suit naturellement les tendances des consommateurs et il est complètement intuitif pour l'utilisateur final.
 - peu intrusif, le moteur a seulement besoin d'identifiants d'utilisateurs et de produits et de votes pour commencer à recommander des produits aux consommateurs. Il n'a pas besoin d'autres informations telles que le nom de l'utilisateur, son e-mail, des informations sur le produit, etc.
 - il fonctionne en temps réel. Afin d'optimiser la précision des recommandations tout en permettant un accès très rapide, notre processus de recommandation se déroule en deux étapes. La première étape est un processus offline utilisé pour analyser les votes et générer un modèle. La seconde étape utilise le modèle pour répondre aux requêtes de recommandations en temps réel ».

¹ <http://www.criteo.com/fr/comment-ca-marche.aspx>

² Le filtrage collaboratif, traduction française d'un terme anglais introduit en 1994, le « collaborative filtering », désigne l'ensemble des méthodes qui, au travers des opinions et évaluations d'un groupe, a pour finalité d'élaborer des systèmes de recommandations. Les techniques consistant à comparer les utilisateurs entre eux relèvent du filtrage collaboratif dit « utilisateurs ». En revanche, les techniques consistant à comparer, entre eux, les articles devant être notés ou recommandés, sont du ressort du filtrage collaboratif dit « objets ou par article ».

Il existe au moins deux grands types de filtrage collaboratif dit « par article » : les modèles binaires lesquels se basent exclusivement sur le fait qu'un utilisateur a, ou non, acheté et/ou sélectionné un bien donné, et les modèles avec évaluations lesquels consistent à inviter les utilisateurs à noter eux mêmes les différents produits.

³ <http://www.criteo.com/fr/a-propos/technologie.aspx>

2.2.2 Les principaux brevets en matière de filtrage de contenus

Afin de mettre en évidence une partie des technologies de filtrage de contenus brevetées, nous avons procédé à une recherche parmi les bases de données rassemblant les demandes de brevets et brevets français, européens, internationaux¹ et américains². Cette recherche a été réalisée notamment à partir des noms des acteurs intervenant dans ce secteur et que nous avons précédemment identifiés.

Les brevets relevés sont répertoriés dans le tableau ci-annexé³. Les abstracts des brevets listés, qui consistent en une présentation succincte de l'invention, apparaissent également en annexe⁴.

Certains acteurs, tels que Google, ont déposé de nombreux brevets qui peuvent se révéler connexes à la technologie de filtrage de contenus.

Lorsque l'on analyse sommairement ces brevets, on constate qu'il est possible de les regrouper en onze catégories :

- le filtrage lexicographique ;
- l'analyse de base de données et extractions ;
- le prototypage de données ;
- le taggage de contenus et l'affectation de metadonnées ;
- l'analyse et/ou traitement d'images ;
- l'analyse de sons et/ou de la parole ;
- l'analyse de vidéos ;
- le filtrage web et/ou filtrage réseau ;
- le filtrage de flux ;
- le filtrage de courriers électroniques et de messages ;
- autres types de filtres de contenus.

Cent cinquante et un brevets ont été identifiés comme relevant de technologies pertinentes pour être intégrées dans des outils de filtrage.

Les brevets retenus ne portent pas sur les technologies élémentaires (analyse de Fourier, détection de cibles, etc.) qui sont déjà bien maîtrisées et exploitées dans des industries autres que le secteur des Technologies de l'information et de la communication.

Après regroupement de ces brevets, on constate la répartition suivante :

- 15 brevets concernant le filtrage lexicographique ;
- 23 brevets concernant l'analyse de base de données et extractions ;
- 29 brevets concernant le prototypage de données ;
- 19 brevets concernant le taggage de contenus et l'affectation de metadonnées ;
- 3 brevets concernant l'analyse et/ou traitement d'images ;
- 7 brevets concernant l'analyse de sons et/ou de la parole ;
- 9 brevets concernant l'analyse de vidéos ;

¹ <http://fr.espacenet.com/>

² <http://www.uspto.gov/patft/index.html>

³ Annexe 5 : Tableau des brevets.

⁴ Annexe 6 : Abstract des brevets.

- 17 brevets concernant le filtrage web et/ou filtrage réseau ;
- 14 brevets concernant le filtrage de flux ;
- 11 brevets concernant le filtrage de courriers électroniques et de messages ;
- 4 brevets concernant autres types de filtrages de contenus.

L'analyse de cette répartition permet à nouveau de confirmer qu'il convient de distinguer l'analyse lexicographique de l'analyse sémantique.

Au sein de l'analyse sémantique, différentes technologies ressortent qui se distinguent en fonction de la nature du média (son/parole, image, vidéo).

Quelques technologies transversales ressortent également en la matière. Elles portent notamment sur le taggage de contenus et l'affectation de meta-données ainsi que sur le prototypage de données, ces technologies permettant notamment d'affecter un sens à des données stockées pour lesquelles le sens n'est pas immédiatement identifiables (vidéo notamment).

2.3 LES PRINCIPAUX CONTRATS FAISANT REFERENCE AU FILTRAGE DE CONTENUS

L'état de l'art en matière de filtrage de contenus ne peut exclure une recherche de quelques contrats, généralement d'adhésion, prévoyant la possibilité pour le prestataire de services sur internet agissant en qualité d'intermédiaire, de mettre en place des technologies de filtrage, ceci étant accepté par l'utilisateur.

2.3.1 Les conditions générales d'utilisation de Yahoo

Les conditions d'utilisation du service Yahoo ! France, dans leur version en date du 19 mars 2008 (<http://fr.docs.yahoo.com/info/utos.html>), disposent à l'article 6 « Comportement des utilisateurs et protection des mineurs » :

- « Des outils visant à la protection des mineurs sont disponibles
Sur le service Yahoo ! Search, vous pouvez si vous le souhaitez, intervenir dans la configuration des résultats de recherche en vous rendant dans la section « Préférences » de yahoo ! Search et en paramétrant le filtre des contenus adultes. L'efficacité de ces filtres est estimée à 95 %. Il demeure que les adultes ayant la garde de mineurs ont l'obligation de surveiller leur utilisation d'Internet. Il est ainsi de leur responsabilité de déterminer les services et les utilisations qu'ils jugent adaptés à ces mineurs. Pour ce faire, ils pourront s'inspirer des conseils prodigués par l'AFA, en collaboration avec les pouvoirs publics à l'adresse <http://www.pointdecontact.net/protectiondelenfance.html>
Lorsque vous mettez en ligne des contenus qui ne sont pas tous publics sur Yahoo ! Groupes, Yahoo ! 360° ou Flickr, veuillez les placer dans les catégories « Adulte » ou utiliser la fonctionnalité équivalente mise à votre disposition dans le service. Cela vise à empêcher l'accès à des internautes mineurs à des contenus qui ne leur sont pas destinés ».

2.3.2 Les conditions générales d'utilisation de Lycos

Les conditions générales d'utilisation pour les services gratuits de Lycos Europe GmbH, dans leur version en date du 16 mai 2007 (http://login.lycos.fr/lsu/print_agb.php), prévoient à l'article XVIII 1., que :

- « 1. Jubii

Jubii est une plateforme de communication qui offre à ses utilisateurs :

- a) Une large gamme de moyens de communication, tels que notamment : messagerie électronique, SMS, et messagerie instantanée ;
- b) Stockage en ligne et partage de fichiers tels que notamment : photos, vidéos et autres documents ;
- c) Gestion et administration des contacts.

Jubii organise les courriers électroniques « entrants » en fonction de leur fiabilité dans différents dossiers, et/ou identifie de manière particulière certains de ces courriels. Les critères pour qu'un courrier électronique soit considéré comme fiable sont déterminés en fonction de votre comportement passé (i.e. est-ce que l'expéditeur est déjà dans vos contacts ; à quelle fréquence répondez-vous aux courriels de l'expéditeur ; effacez-vous immédiatement les courriels de l'expéditeur ?).

Jubii fournit des filtres anti-spam et anti-virus pour protéger les systèmes de traitement de l'information de Lycos, de la même manière que votre boîte de messagerie électronique. Les courriers électroniques suspectés d'être des spams sont marqués et placés dans des dossiers désignés en tant que tels où ils peuvent être consultés par vous.

Lycos décline expressément toute responsabilité pour le cas où lesdits filtres échoueraient à filtrer tous les spams, virus, ou autres logiciels dangereux.

A chaque courrier électronique envoyé par l'intermédiaire de Jubii, Lycos pourra ajouter une mention indiquant l'origine du message (par exemple « powered by Lycos »), ou de la publicité.

L'utilisation de Jubii pour envoyer massivement des courriers électroniques (« Spamming »), faire du « mail-bombing » ou envoyer toute autre forme de message publicitaire ou de marketing vous est interdite et toute personne utilisant Jubii à de telles fins sera tenue pour seule et entière responsable des conséquences d'un tel acte . Il vous est interdit de déguiser ou de masquer votre identité lorsque vous envoyez un courrier électronique par l'intermédiaire de Jubii ».

Il est ainsi expressément indiqué que Jubii comporte effectivement des filtres antisпам et antivirus dont la parfaite efficacité n'est toutefois pas garantie par Lycos, cette dernière déclinant toute responsabilité dans le cas où les filtres viendraient à échouer à filtrer tous les spams.

2.3.3 Les conditions générales d'utilisation de YouTube (Google)

Les conditions générales d'utilisation de YouTube, dans leur version en date du 20 mars 2008 (<http://fr.youtube.com/t/terms>), et dans un article 7 intitulé « Politique de protection des droits d'auteur », renvoient à l'adresse http://fr.youtube.com/t/copyright_notice portant « Notification d'infraction aux droits d'auteur », laquelle adresse renvoie elle-même au programme de vérification du contenu mis en place par YouTube (http://fr.youtube.com/t/copyright_program), programme en vertu duquel :

- « YouTube s'engage à aider les détenteurs de droits d'auteur à trouver et supprimer du site les contenus présumés en infraction aux droits de propriété. C'est pourquoi, nous avons créé un outil de vérification des droits d'auteur qui permet aux titulaires de ces droits de rechercher les contenus qu'ils estiment être en infraction et de fournir à YouTube les informations requises pour localiser ces contenus.

Cet outil est spécialement conçu pour que les entreprises détentrices de droits d'auteur puissent soumettre de multiples demandes de suppression. Il est possible d'envoyer des notifications individuelles en suivant ces instructions.

Si vous avez déjà un compte YouTube, vous pouvez demander l'accès à cet outil en remplissant une demande de participation au programme de vérification du contenu de YouTube. Imprimez-la, puis envoyez-la par fax au numéro indiqué sur le document. Ce formulaire identifie vos représentants et atteste légalement que vous êtes le détenteur des droits d'auteur du document au sujet duquel vous voulez contacter YouTube. Si vous n'avez pas de compte, veuillez en créer un. Vous pourrez ensuite accéder à la demande de participation au programme de vérification du contenu ».

2.3.4 Les conditions générales d'utilisation de Gmail (Google)

Les conditions d'utilisation de Gmail, telles que rédigées au 20 mars 2008 (http://mail.google.com/mail/help/intl/fr/terms_of_use.html), comportent un article 4 « Contenu du service », dont le point c) aux termes duquel :

- « [Google se réserve le droit de] détecter, prévenir, ou lutter contre des problèmes de fraude, de sécurité ou des problèmes d'ordre techniques (y compris, notamment, le filtrage des e-mails non sollicités ou "spam") ».

2.3.5 Les conditions générales d'utilisation de MySpace

Les conditions générales d'utilisation de myspace.com, dans leur version au 28 février 2008 (<http://www.myspace.com/index.cfm?fuseaction=misc.terms>), stipulent à l'article 9 que :

- « Protéger les droits d'auteur et autres droits de propriété intellectuelle MySpace respecte la propriété intellectuelle des autres et exige que ses Utilisateurs fassent de même. Vous n'êtes pas autorisé à mettre en ligne, télécharger, intégrer, publier, envoyer par courrier électronique, transmettre ou autrement rendre disponible tout contenu qui enfreint tout droit d'auteur, brevet, marque, secret de fabrication ou autre droit de propriété de toute personne ou entité. MySpace se réserve le droit d'annuler l'Adhésion des contrefacteurs.

Si vous pensez que votre travail a été copié et publié sur ou par le biais des Services MySpace d'une manière qui constitue une atteinte à vos droits d'auteur, veuillez faire parvenir à l'Agent des Droits d'Auteur de MySpace une notification décrivant la violation revendiquée avec toutes les informations qui suivent : (a) identification de l'œuvre protégée par les droits d'auteur dont la contrefaçon est alléguée, ou, dans le cas où plusieurs œuvres protégées par les droits d'auteur sont concernées, une liste représentative de telles œuvres ; (b) identification raisonnablement suffisante de l'œuvre dont la contrefaçon est alléguée et les informations nous permettant de situer le contenu concerné sur les Services MySpace (l'indication de l'URL ou des URLs menant au contenu litigieux est suffisante) ; (c) informations raisonnablement suffisantes pour nous permettre de vous contacter, comme une adresse, un numéro de téléphone et, le cas échéant, une adresse électronique ; (d) une déclaration de votre part indiquant que vous estimez en toute bonne foi que l'usage contesté n'est pas autorisé par le titulaire des droits d'auteur, son agent ou la loi ; (e) une déclaration de votre part qui atteste, sous peine de parjure, que les informations ci-dessus figurant dans votre notification sont exactes et que vous êtes le titulaire des droits d'auteur ou autorisé à agir au nom du titulaire des droits d'auteur ; et (f) votre signature physique ou électronique. L'Agent des droits d'auteur de MySpace chargé de recevoir les notifications des infractions alléguées peut être contacté à l'adresse suivante : Copyright Agent, MySpace, Inc., 8391 Beverly Blvd., n° 349, Los Angeles, CA 90048, Etats-Unis ; Télécopie : (310) 388-0892 ; Attn : Copyright Agent. Il peut également être contacté par courrier électronique en cliquant ici : <http://collect.myspace.com/index.cfm?fuseaction=misc.contactInput&primarySubject=2&secondarySubject=32>. MySpace fournit certains outils et technologies qui aident les titulaires de droits d'auteur à contrôler leurs œuvres protégées par les droits d'auteur ».

3. LA PRATIQUE D'EBAY

eBay a toujours cherché à lutter contre un usage illicite de ses services par des tiers mal intentionnés et a mis en place un ensemble de mesures à cet effet.

Allant au-delà de ses obligations légales, eBay a volontairement mis en œuvre des moyens de lutte contre la contrefaçon au stade de la notification avec le formulaire général de notification et le programme VeRO¹, mais également en dehors de toute notification grâce aux recherches quotidiennes qu'elle effectue sur ses sites internet pour identifier les annonces manifestement contrefaisantes, à l'adoption de règlements et de messages d'avertissement destinés à prévenir la mise en ligne d'annonces potentiellement contrefaisantes et à une politique d'information des utilisateurs. En outre, elle suspend les utilisateurs contrevenants et a mis en place des mesures pour empêcher les utilisateurs suspendus de réutiliser le site.

3.1 PRESENTATION DES PROGRAMMES PROACTIFS DE RECHERCHE D'ANNONCES MANIFESTEMENT ILLICITES MIS EN PLACE PAR EBAY

Pour identifier les annonces manifestement illicites mises en ligne sur ses sites internet, eBay a volontairement mis en place des outils de recherche proactive²

3.1.1 L'outil de recherche par mots-clés

eBay effectue quotidiennement des recherches proactives sur l'ensemble de ses sites internet pour identifier les annonces manifestement illicites, c'est-à-dire celles qui proposent à la vente des produits illicites tels que armes, objets nazis ou qui comportent des termes indiquant qu'il s'agit manifestement d'un produit contrefait.

¹ eBay a créé un programme de coopération avec les titulaires de droits de propriété intellectuelle, le programme VeRO, qui leur permet de signaler facilement et rapidement les annonces qui portent atteinte à leur droit de propriété intellectuelle. Le programme VeRO réunit aujourd'hui plus de 18 000 entreprises, associations et particuliers titulaires de droits de propriété intellectuelle dans le monde. Le programme VeRO fonctionne selon un régime de « retrait sur notification ». Le titulaire de droits de propriété intellectuelle adresse à eBay un formulaire de notification VeRO en précisant les numéros d'annonces contrefaisantes. La notification est transmise aux équipes dédiées d'eBay qui traitent promptement l'annonce ainsi notifiée. eBay envoie un courrier électronique au vendeur lui notifiant le retrait de son annonce et le renvoyant aux règlements relatifs à la mise en ligne d'annonces contrevenantes. La participation au programme VeRO est gratuite, simple et rapide. En effet, pour participer au programme VeRO, le titulaire de droits de propriété intellectuelle doit simplement remplir, puis faxer, un formulaire de notification VeRO, téléchargeable en ligne, en indiquant ses coordonnées et les numéros d'annonces potentiellement contrefaisantes ainsi que les droits potentiellement contrefaits.

² « Flag filters » (filtres de signalement), « delay filters » (filtres retard) et « block filters » (mesures de limitation).

Concernant l'identification d'annonces manifestement contrefaisantes, eBay mène ces recherches en utilisant des mots-clés, tels que des marques ou des mots-clés révélateurs d'une activité contrefaisante. Il en est ainsi, par exemple, des termes « copie », « imitation », « faux », « réplique », « reproduction », « contrefaçon ».

Les annonces ainsi détectées sont examinées manuellement afin de déterminer si elles proposent effectivement manifestement à la vente un produit contrefait.

S'il est avéré que le vendeur propose à la vente des objets manifestement contrefaisants, eBay procède, dans les meilleurs délais :

- à l'annulation de l'annonce ainsi qu'à son retrait du site ;
- à la notification au vendeur du retrait de l'annonce ;
- au remboursement de tous les frais relatifs à la mise en ligne de l'annonce ;
- à l'examen de son compte en vue d'une éventuelle suspension.

Ces recherches ne concernent toutefois que les cas dans lesquels le caractère manifestement contrefaisant apparaît à partir du seul contenu de l'annonce dans la mesure où eBay n'est pas experte des produits proposés à la vente sur ses sites et ne peut donc intervenir dans les autres cas d'atteinte aux droits de propriété intellectuelle.

eBay demande aux titulaires de droits de coopérer avec elle pour perfectionner ses outils d'identification des annonces manifestement contrefaisantes.

Toutefois, certaines annonces peuvent ne pas être détectées parce que, par exemple, le vendeur a intentionnellement mal orthographié le nom de la marque.

D'autres annonces peuvent par contre être à tort détectées comme illicites et retirées du site, le mot-clé faisant référence à un tout autre objet que celui pour lequel l'outil de recherche par mots-clés a été paramétré.

Aussi, eBay adapte continuellement ses outils de recherche proactifs d'annonces manifestement illicites pour s'adapter à l'ingéniosité des contrefacteurs.

3.1.2 Les mesures de limitation


Parallèlement à l'outil de recherche par mots-clés, eBay a mis en place des mesures de limitation.

Les annonces qui proposent dans certaines conditions des objets particulièrement exposés à la contrefaçon sont bloquées. Les annonces ainsi bloquées ne sont pas ensuite examinées manuellement.

Il s'agit par exemple :

- des annonces contenant une mise aux enchères sur un ou trois jours privilégiées par les contrefacteurs ;

Vous ne pouvez pas soumettre votre annonce à cause des problèmes suivants

 **Cher vendeur,**


A des fins de sécurité, des restrictions peuvent être placées sur certaines mises en vente. Cet objet doit être mis en vente pour une durée de 5, 7 ou 10 jours. Veuillez revenir à la page précédente pour modifier la durée de l'annonce.

Nous vous remercions pour votre coopération et vous souhaitons beaucoup de succès sur eBay.

[En savoir plus](#) sur le règlement eBay sur les objets interdits, contestables ou contrevenants.

(#12341-2147.18) [Revenir à la page suivante et modifier mon annonce](#)

- des annonces proposant à la livraison ce type d'objets en dehors de la zone de langue concernée ;

 **Attention !**

Cher vendeur,

A des fins de sûreté eBay peut imposer des restrictions sur les comptes. En l'occurrence, vous avez atteint temporairement une limite de mise en vente l'international.

Pour qu'eBay puisse lever cette limite, nous vous demandons de bien vouloir prendre contact avec notre service consommateur, l'adresse suivante : cbtquestions@ebay.com, en donnant une liste de vos pseudos eBay. Nous pourrions alors effectuer les vérifications nécessaires pour lever la limite de mise en vente l'international.

Nous vous remercions pour votre coopération et vous souhaitons beaucoup de succès sur eBay

3.2 LE CADRE JURIDIQUE DANS LEQUEL LES MOYENS DE LUTTE CONTRE LA CONTREFAÇON SONT MIS EN ŒUVRE PAR EBAY

3.2.1 L'absence d'obligation générale de surveillance et de recherche de faits ou de circonstances révélant des activités illicites

Aux termes de l'article 6-I 7° de la loi pour la confiance dans l'économie numérique les hébergeurs ne sont pas soumis à une obligation générale de surveiller les informations qu'ils stockent, ni à une obligation générale de rechercher des faits ou des circonstances révélant des activités illicites.

L'article 14 2° de la directive sur le commerce électronique énonce que le régime de responsabilité dérogatoire des prestataires de stockage « ne s'applique pas lorsque le destinataire du service agit sous le contrôle ou l'autorité du prestataire ». Cet article a été transposé à l'article 6-I 2°, alinéa 2, de la loi pour la confiance dans l'économie numérique¹.

A cet égard, la Commission européenne a précisé que le terme « contrôle » faisait référence au « contrôle des activités [du destinataire du service] et non à celui des informations elles-mêmes [mises en ligne par le destinataire du service] »².

Le régime de responsabilité limitée des prestataires de stockage ne s'applique donc pas lorsque le destinataire du service agit sur ordre, sous la direction ou sous le lien hiérarchique de l'hébergeur.

Ainsi donc, le régime de responsabilité de l'hébergeur prévu par l'article 6-I 2° de la loi pour la confiance dans l'économie numérique s'applique à eBay, le fait qu'eBay ait mis en place des outils de recherche par mots-clés afin d'identifier les annonces manifestement illicites qui seraient mises en ligne sur ses sites internet étant sans incidence.

eBay, en sa qualité d'hébergeur, n'a aucun devoir général de surveillance et de rechercher des faits ou des circonstances révélant des activités illicites sur ses sites internet. Toutefois, elle procède spontanément à la régulation des contenus diffusés sur ses sites internet et agit en « Bon Samaritain ».

3.2.2 La loi du Bon Samaritain

Bien qu'elle n'ait aucun devoir général de surveillance et de contrôle, eBay a volontairement mis en place un ensemble de mesures destinées à lutter contre un usage illicite de ses services par des tiers mal intentionnés.

eBay agit en « Bon Samaritain » lorsqu'elle régule volontairement le contenu de ses sites internet.

La loi du Bon Samaritain est aux Etats-Unis un ensemble de règles destinées à protéger tout citoyen portant assistance, notamment les secouristes bénévoles, contre toute poursuite judiciaire.

¹ A cet égard, il convient de relever que le terme « contrôle » est la traduction du terme anglais « to control » qui signifie « avoir le pouvoir sur » et se distingue donc du terme « to monitor » qui signifie « surveiller ».

² Bulletin d'information Communications commerciales, janvier 1999.

Plus généralement, la loi du Bon Samaritain désigne les dispositions qui exonèrent les personnes des dommages qu'ils pourraient causer lorsqu'ils réalisent des bonnes actions qu'ils n'ont aucune obligation légale de réaliser.

Des clauses du Bon Samaritain sont contenues dans diverses lois, telles que :

- le « Communications Decency Act » de 1996 selon lequel « aucun fournisseur ou utilisateur d'un service informatique interactif ne peut voir sa responsabilité engagée du fait de toute action prise de bonne foi pour restreindre l'accès ou la disponibilité d'un contenu que le fournisseur ou l'utilisateur considère obscène, lubrique, dégoûtant, extrêmement violent, harcelant ou critiquable pour toute autre raison, que ce contenu soit constitutionnellement protégé ou non, ou de toute action prise pour permettre ou rendre disponible au fournisseur de contenus d'informations ou à d'autres parties les moyens techniques pour restreindre l'accès aux contenus décrits au paragraphe ci-dessus »¹ ;
- le « Digital Millennium Copyright Act » de 1998 qui a inspiré la directive sur le commerce électronique et selon lequel les fournisseurs d'un service informatique interactif sont, sous certaines conditions, exonérés de toute responsabilité, directe ou indirecte, des contenus qu'ils auraient soit accueilli sur leurs serveurs, en ce qui concerne les hébergeurs et les chat room, soit véhiculé, en ce qui concerne les fournisseurs d'accès à l'internet².

Les tribunaux américains ont appliqué la clause du Bon Samaritain à eBay.

Ainsi, dans l'affaire Gentry de 2002³, la Cour d'appel de Californie a rejeté la plainte formée contre eBay sur le fondement de la clause du Bon Samaritain contenue dans le Communications Decency Act de 1996⁴.

¹ Section 230 du Communications Decency Act de 1996.

² Section 512 du Digital Millennium Copyright Act de 1998.

³ Cour d'appel de Californie, Lars Gentry c/ eBay Inc., 26 juin 2002.

⁴ Des personnes avaient vendu sur eBay des objets sportifs dédicacés. Les acheteurs de ces objets ont décidé de poursuivre eBay pour négligence et violation de l'article 1739.7 du Code civil californien qui réglemente la vente d'objets sportifs dédicacés en obligeant les vendeurs à fournir un certificat d'authenticité. Appliquant la jurisprudence Zeran de 1997 selon laquelle les actions en justice visant à mettre en cause la responsabilité d'un fournisseur de services informatiques interactifs en raison de l'exercice des fonctions éditoriales traditionnelles d'un éditeur sont interdites en application de la section 230 du CDA, la Cour d'appel de Californie a jugé que « rendre responsable eBay pour ne pas avoir fourni une garantie conformément à la section 1739.7 serait contraire [...] à la section 230 du Communication Decency Act de 1996 ». En effet, la section 230 du CDA met à l'abri les fournisseurs de services informatiques interactifs contre les actions en justice initiées par des personnes soutenant avoir subi un préjudice causé par un contenu fourni par un tiers. La Cour d'appel de Californie a considéré que si elle imposait une responsabilité au titre de l'article 1739.7 du Code civil californien à eBay, elle reconnaîtrait alors eBay comme étant responsable d'un contenu provenant d'une tierce partie et la traiterai ainsi comme un éditeur, c'est-à-dire comme l'émetteur originel du contenu. Au contraire, elle a considéré qu'eBay avait simplement rendu disponible les descriptions des faux produits aux autres utilisateurs de son site internet, que ce sont les vendeurs des faux objets sportifs dédicacés et non pas eBay qui ont choisi la catégorie sous laquelle leurs produits devaient être identifiés et que se sont donc eux qui ont faussement décrit leurs produits comme comportant des dédicaces authentiques. Enfin, s'agissant de l'argument des appelants selon lequel eBay connaissait ou aurait dû connaître le comportement illégal ou frauduleux des intimés et n'a pas pris de mesures pour assurer le respect de la loi, elle a rappelé que ceci est classiquement le type de plainte que l'arrêt Zeran a considéré comme étant empêché par la section 230 du CDA. Dans ces conditions, elle a rejeté la plainte formée contre eBay.

Par ailleurs, dans l'affaire Hendrickson de 2001, le Tribunal de première instance de Californie reprenant les travaux parlementaires relatifs au Digital Millennium Copyright Act de 1998 ci-après, a rejeté la plainte formée contre eBay sur le fondement de la clause du Bon Samaritain :

- « L'implication volontaire d'eBay dans une surveillance limitée de son site internet concernant les contrefaçons « apparentes » conformément au programme VeRO, ne peut, en soi, mener le tribunal à conclure qu'eBay a le droit et la capacité de contrôler toute activité illicite selon les termes du DMCA. Les précédents législatifs montrent que le Congrès n'a pas cherché à ce que des sociétés telles qu'eBay soient pénalisées lorsqu'elles réalisent des efforts volontaires pour combattre le piratage sur internet : cette législation n'a pas pour but de décourager les fournisseurs de services de surveiller leur service en vue de contrôler les contenus illicites. Les tribunaux ne doivent pas conclure que le fournisseur de services perd son droit à la limite de responsabilité en vertu de l'article 512 uniquement parce qu'il s'implique dans un programme de surveillance ».

eBay agit donc en « Bon Samaritain » lorsqu'elle met en œuvre des moyens de lutte contre la contrefaçon sur ses sites internet alors qu'elle n'a aucune obligation de le faire.

4. ANNEXES

Le présent Livre blanc comprend sept annexes :

- annexe 1 : référentiel légal
- annexe 2 : schémas sur les types de filtrage
- annexe 3 : glossaire
- annexe 4 : bibliographie commentée
- annexe 5 : tableau des brevets déposés ayant trait ou susceptibles d'avoir un lien avec le filtrage de contenus
- annexe 6 : abstract des brevets
- annexe 7 : extraits de conditions générales d'utilisation

ANNEXE 1

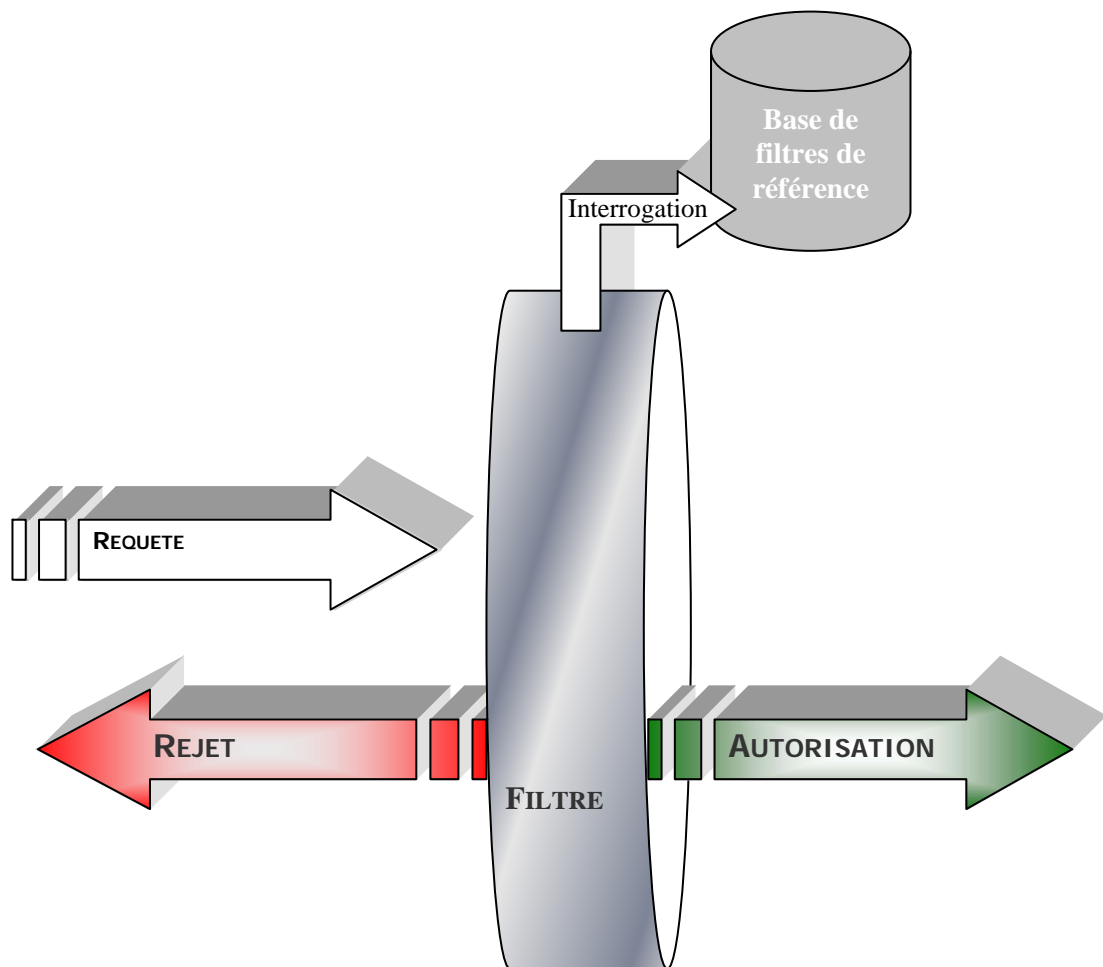
REFERENTIEL LEGAL

- Directive 2000/31/CE du Parlement européen et du Conseil du 8 juin 2000 relative à certains aspects juridiques des services de la société de l'information, et notamment du commerce électronique, dans le marché intérieur (« Directive sur le commerce électronique »)
- Loi n°2004-575 pour la confiance dans l'économie numérique du 21 juin 2004
- Loi n°2007-297 du 5 mars 2007 relative à la prévention de la délinquance

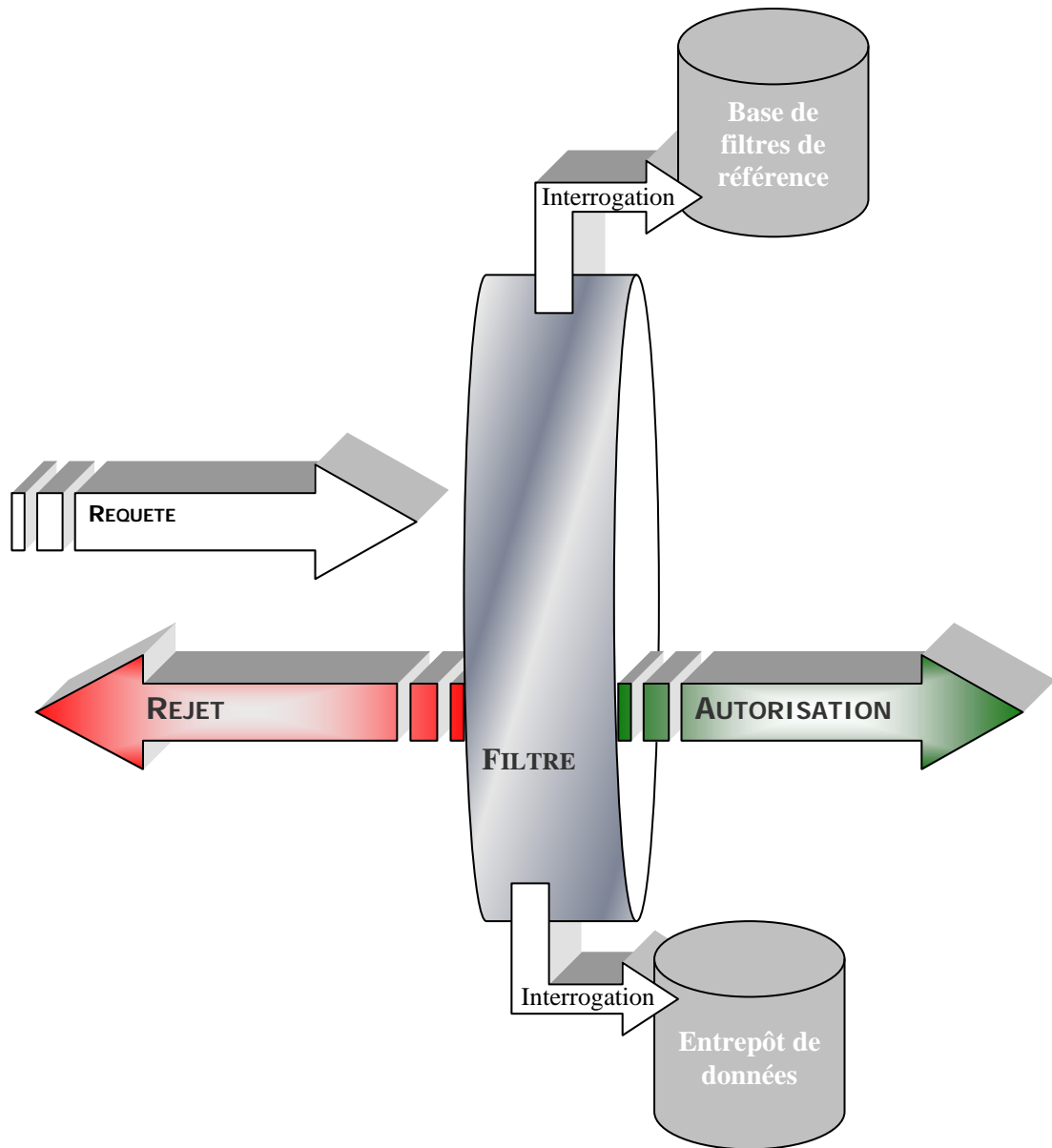
ANNEXE 2

SCHEMAS SUR LES TYPES DE FILTRAGE DE CONTENUS

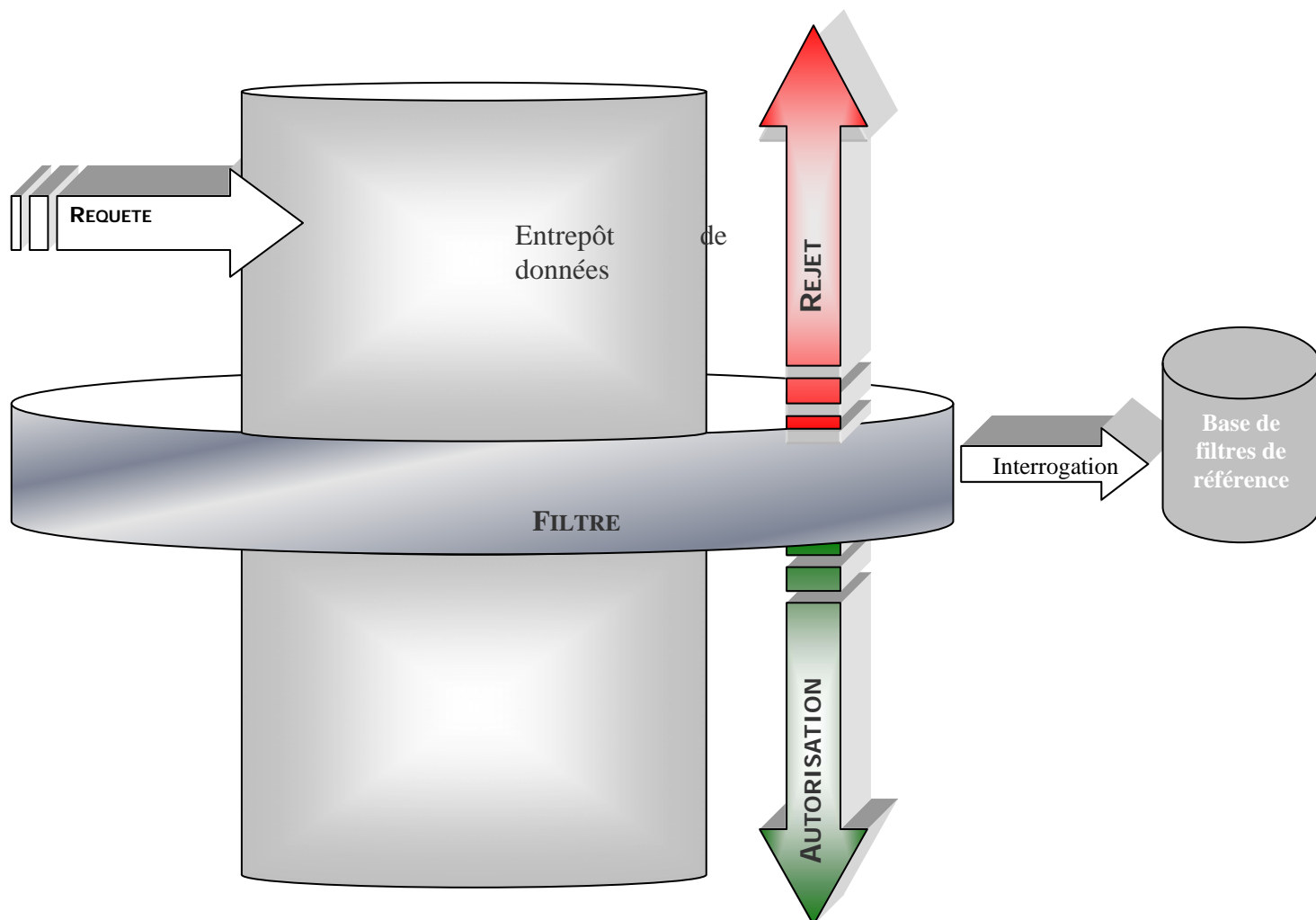
1. LE FILTRAGE DE CONTENUS DE FLUX



2. LE FILTRAGE DE CONTENUS AVEC BASE DE DONNEES OPERE A PRIORI



3. LE FILTRAGE DE CONTENUS AVEC BASE DE DONNEES OPERE A POSTERIORI OU « DATA MINING »



ANNEXE 3

GLOSSAIRE

Data mining : processus d'extraction de données réalisé au moyen de méthodes automatiques ou semi-automatiques et en vue d'une utilisation industrielle ou opérationnelle.

Data warehouse : base de données particulière organisée de façon à faciliter l'analyse de toutes les données produites par une entreprise.

Filtrage collaboratif : traduction du terme anglais « collaborative filtering », le filtrage collaboratif regroupe l'ensemble des méthodes qui visent à construire des systèmes de recommandation utilisant les opinions et évaluations d'un groupe pour aider l'individu. Il existe trois principaux types de filtrage collaboratif :

- le filtrage collaboratif passif qui repose sur l'analyse des comportements ;
- le filtrage collaboratif actif qui repose sur du déclaratif (notes, commentaires...) de la part des utilisateurs, lequel peut être soit utilisateurs, soit objets ;
- le « Content Based » qui repose sur une classification « objective » des éléments filtrés.

Filtre : système servant à séparer des éléments dans un flux, ce flux pouvant être un flux de matières, un flux électronique, un flux d'informations, et dont l'action du filtre consiste à retenir, supprimer, rediriger ou modifier les éléments indésirables du flux et, à en laisser passer librement les éléments utiles.

Filtre bayésien : filtre utilisant une méthode probabiliste fondée sur le théorème de Thomas Bayes et visant à prédire si un courrier électronique est légitime ou s'il s'agit d'un spam.

Filtre informatique : programme capable de traiter un ensemble d'informations pour en extraire un sous-ensemble d'informations pertinentes.

Fonction de hashage : méthode permettant de caractériser une information ou donnée consistant à faire subir une suite de traitements reproductibles à la donnée fournie en entrée afin de générer une empreinte servant à identifier la donnée initiale.

Hameçonnage ou phishing : action consistant à envoyer un courrier électronique à un destinataire qui ne l'a pas sollicité et à l'occasion duquel l'expéditeur a maquillé son identité de manière à obtenir des informations personnelles sur ce dernier à des fins frauduleuses et/ou malveillantes.

Malware : tout type de programme nocif introduit sur un ordinateur à l'insu de l'utilisateur et regroupant les virus, vers, spywares, keyloggers, chevaux de Troie, backdoors.

Hypertext Transfer Protocol Secured (HTTPS) : version sécurisée du protocole HTTP, littéralement « protocole de transfert hypertexte », protocole de communication client-serveur développé pour le World Wide Web.

Pourriel ou spam : communication électronique, notamment courrier électronique, non sollicitée par les destinataires et expédiée en masse à des fins publicitaires ou malhonnêtes.

Serveur mandataire ou proxy : serveur informatique ayant pour fonction de relayer des requêtes entre un poste client et un serveur, principalement utilisé pour assurer la journalisation des requêtes ou « logging », la sécurité du réseau local, le filtrage et l'anonymat.

Spoofing : technique consistant à falsifier l'adresse électronique d'une personne ou d'une société afin que le destinataire du message envoyé fasse confiance et ouvre le message.

Tatouage numérique ou « watermarking » : technique permettant d'ajouter des informations de copyright ou d'autres messages de vérification à un fichier ou signal audio, vidéo, une image ou un autre document numérique et consistant à introduire un message caché dans le signal hôte, généralement appelée marque ou bien simplement message, et consistant en un ensemble de bits, dont le contenu dépend de l'application.

Technique de type « fingerprinting » : technique permettant d'ajouter des informations de copyright ou d'autres messages de vérification à un fichier ou signal audio, vidéo, une image ou un autre document numérique et consistant en la prise d'empreintes numériques destinées à marquer le fichier ou signal.

Uniform Resource Locator (URL) : format de nommage universel désignant une ressource sur internet et consistant en une chaîne de caractères ASCII imprimables, se décomposant en cinq parties : le nom du protocole, le langage utilisé pour communiquer sur le réseau (le protocole http, le format html...), un identifiant et un mot de passe, le nom du serveur, le numéro de port et le chemin d'accès à la ressource.

ANNEXE 4

BIBLIOGRAPHIE ET WEBOGRAPHIE COMMENTEES

1. ARTICLES ET RAPPORTS SUR LE MARCHÉ DU FILTRAGE

1.1 Aguilla Nicolas, "*Google ressert les mailles de ses filtres*", 15 février 2007, <http://www.infos-du-net.com/actualite/9898-filtres-google.html>

Mise en place par Google d'un filtre avertissant l'internaute des sites présentant un risque de contamination virale ou de spyware.

1.2 Anselin Elisabeth, "*Signature de la charte FAI*", 28 juillet 2004, <http://www.sacem.fr/portailSacem/jsp/ep/contentView.do;jsessionid=HIK3K2wOC4bZDcuIIRU7DnCDRIDkpeFumprM6r1VM85GBC7yRJeU!879839327?channelId=-536879932&contentId=536883980&programId=536882534&programPage=%2Fep%2Fprogram%2Feditorial.jsp&pageTypeId=8586&contentType=EDITORIAL>

Participation active de la société des auteurs, compositeurs et éditeurs de musique (Sacem) à la négociation de la charte entre fournisseurs d'accès à internet, professionnels de la musique et pouvoirs publics, et engagement de la Sacem à offrir son répertoire aux services de musique en ligne légaux.

1.3 Bellenger Cédric, "*Google lance son système de filtrage pour Youtube*", 17 octobre 2007, <http://www.loadings.fr/blog/2007/10/17/google-lance-son-systeme-de-filtrage-pour-youtube>

Lancement du programme « video identification » de YouTube. Ce système permettra de bloquer le renvoi sur le site des clips déjà censurés.

1.4 Borland John, "*Sony Music choisit Audible Magic pour se protéger du piratage*", 2 juin 2004, <http://www.zdnet.fr/actualites/internet/0,39020774,39155146,00.htm>

Sony Music utilise le logiciel CopySense qui permet à un administrateur d'identifier les fichiers qui transitent sur le réseau de son organisation (entreprise ou université) et de bloquer ceux qui figurent dans la liste des œuvres protégées par le copyright.

1.5 Chêne Olivier, "*Watermarking : Thomson va adapter nexguard à Windows Media Video 9*", 21 mars 2007, <http://www.vnunet.fr/fr/news/2007/03/21/thomson-point-de-d-voiler>

Annnonce de la création d'une solution de filtrage de type « watermarking » pour le format vidéo WMV9. Cette solution est mise en œuvre au moyen des outils Embedder et Investigator. Ce marquage devrait résister aux changements de format des données.

1.6 Dailymotion, communiqué de presse, "Dailymotion choisit la solution de fingerprinting d'Audible Magic pour détecter les vidéos protégées par des droits", 13 juillet 2007,
<http://www.dailymotion.com/press/fr>

Audible Magic et son service d'identification de contenus fournira à Dailymotion des solutions d'identification et de filtrage des contenus protégés par le droit d'auteur.

1.7 Dailymotion, communiqué de presse, "Dailymotion renforce son dispositif de détection des vidéos protégées avec la technologie « signature » de l'INA", 8 octobre 2007,
<http://www.dailymotion.com/press/fr>

Signature d'un accord de partenariat entre Dailymotion et l'Ina afin d'installer le système « signature » basé sur une technique d'empreinte digitale d'une séquence vidéo.

1.8 Dailymotion, communiqué de presse, "Dailymotion announces full implementation of INA technology for detection of copyright video", 25 février 2008,
<http://www.dailymotion.com/press/fr>

Annonce de la fin de l'implémentation du système « signature » de l'Ina sur l'ensemble de Dailymotion, l'ensemble des producteurs étant invité à enregistrer leurs vidéos dans la base de données.

1.9 Devillard Arnaud, "Les FAI, déchargés de l'obligation de filtrer les échanges de musique", 10 mars 2005, <http://www.01net.com/editorial/270396/les-fai-decharges-de-l-obligation-de-filtrer-les-echanges-de-musique>

Analyse du rapport d'étude de Messieurs Gilles Kahn et Antoine Brudigou recommandant que les fournisseurs d'accès à internet ne filtrent pas les contenus musicaux sur internet. Ils préconisent le filtrage opéré sur l'ordinateur de l'internaute uniquement à sa demande.

1.10 Devillard Arnaud, "Interview de Bernard Miyet (SACEM) : Nous souhaitons des tests sur le filtrage des contenus", 6 décembre 2007,
[http://www.01net.com/editorial/365559/bernard-miyet-\(sacem\)-nous-souhaitons-des-tests-sur-le-filtrage-des-contenus](http://www.01net.com/editorial/365559/bernard-miyet-(sacem)-nous-souhaitons-des-tests-sur-le-filtrage-des-contenus)

Interview de Monsieur Bernard Miyet, président du directoire de la Sacem, sur le rapport Olivennes. La Sacem a mis en place, au cours de l'année 2000, un système de filtrage qu'elle a finalement retiré, ce dernier ayant été considéré comme illicite par la Commission nationale de l'informatique et des libertés. Depuis, aucun système de filtrage n'a été mis en œuvre par la Sacem.

1.11 Dimitri T., "Une nouvelle technologie de filtrage des contenus en ligne", 31 octobre 2007,
<http://www.generation-nt.com/piratage-video-plates-formes-kddi-technologie-filtrage-kddi-actualite-44834.html>

KDDI, opérateur japonais de téléphonie mobile, annonce la mise au point d'une technologie permettant d'analyser les contenus disponibles sur les sites de partage en ligne afin de distinguer les vidéos filmées par des amateurs de celles filmées par des professionnels.

1.12 Dumout Estelle, "MySpace et Soapbox trahis par leur technologie de filtrage vidéo", 14 juin 2007, <http://www.zdnet.fr/actualites/internet/0,39020774,39370275,00.htm>

Les défaillances constatées des systèmes de filtrage dont la finalité est de bloquer les contenus sous copyright.

1.13 Dutheil Christophe, "Filtrage de contenu : Websense adapte ses solutions aux PME", 4 juillet 2007, <http://www.vnunet.fr/fr/news/2007/07/04/filtrage-de-contenus-websense>

Annnonce de la solution Websense Express, solution de filtrage d'URL dédiée aux PME.

1.14 Filippone Dominique, "Panorama des offres de filtrage de contenus web", <http://www.journaldunet.com/solutions/0606/060607-panorama-filtrage-contenus-web-1.shtml>

Panorama des critères de choix entre les différentes solutions de filtrage et comparaison des différentes offres des éditeurs.

1.15 Google, "Comment les filtres fonctionnent-ils? Quelle est la meilleure façon de les utiliser", www.adwords.google.com/support/bin/answer.py?hl=fr&answer=29173

Manière dont il conviendrait d'utiliser des filtres afin de fournir un niveau supplémentaire de personnalisation dans la recherche de l'internaute.

1.16 Grenier Frantz, "Résultats du filtrage des séries et des films", 20 juillet 2007, www.journaldunet.com/ebusiness/internet/dossier/070720-test-audible-magic-dailymotion-youtube/3-test-series-et-films.shtml

La base de données d'Audible Magic dispose de trop peu d'empreintes de films et de séries, notamment françaises, pour que son système de filtrage puisse être efficace.

1.17 La rédaction du site Clubic, "L'INA protège Canal+ des services de piratage vidéo", 25 juillet 2007, <http://www.clubic.com/actualite-77338-ina-protege-canal-services-partage-video.html>

Adoption par le groupe Canal+ de la technologie « signature » de l'Institut national de l'Audiovisuel (Ina). Ce dispositif permet le filtrage des contenus vidéo par un marquage numérique du fichier vidéo.

1.18 La rédaction JDN et JDN solution, "Le filtrage des contenus sur Dailymotion et You Tube est-il efficace ?", 9 février 2008, <http://www.journaldunet.com/ebusiness/internet/dossier/080211-test-filtrage-dailymotion-youtube>

Tests réalisés par le Journal du Net sur l'efficacité des systèmes de filtrage vidéo de DailyMotion et de YouTube.

1.19 La rédaction du JDN, "Marché des éditeurs de solution de filtrage web dans le monde", 5 septembre 2002, www.journaldunet.com

Tableau de répartition des parts de marché des éditeurs de solutions de filtrage web dans le monde pour l'année 2001.

1.20 La rédaction JDN, "MySpace filtre les contenus protégés", 31 octobre 2006, <http://www.journaldunet.com/breve/6120/myspace-filtre-les-contenus-proteges.shtml>

Annonce par MySpace de l'adoption d'un système de filtrage développé par Gracenote afin d'identifier les contenus vidéo.

1.21 La rédaction du site numerama, "Microsoft confirme son choix pour Audible Magic", 27 mars 2007, <http://www.numerama.com/magazine/4322-Filtrage-Microsoft-confirme-son-choix-pour-Audible-Magic.html>

Avec AOL, Yahoo et MySpace, Microsoft a décidé de créer un vaste consortium pour apporter une alternative à YouTube et au piratage.

1.22 La rédaction de 01net, "Le moteur français Exalead cherche les visages", 20 avril 2007, <http://www.01net.com/editorial/346640/le-moteur-francais-exalead-cherche-les-visages>

Le moteur de recherche Exalead a introduit un filtre de tri intitulé « visage » afin de pouvoir, lors d'une requête, ne rechercher que des images.

1.23 La rédaction Zdnet, "L'INA et TDF rapprochent leurs technologies de filtrage de contenus audio et vidéo", 23 novembre 2007, <http://www.zdnet.fr/actualites/internet/0,39020774,39375808,00.htm>

Signature d'une alliance pour le développement dans le domaine du traçage de contenus audiovisuels. Il s'agit d'une combinaison entre la technologie de filtrage vidéo « signature » et de la technologie de filtrage audio « wavessence ».

1.24 LTU Technologies, "Image Filter", <http://www.ltutech.com>

Présentation du logiciel de contrôle parental image-filter fonctionnant grâce à l'analyse automatique des images par la forme, la couleur, la texture pour en extraire une signature numérique.

1.25 Olivennes Denis, "Rapport au ministre de la culture et de la communication sur le développement et la protection des œuvres culturelles sur les nouveaux réseaux", novembre 2007, <http://lesrapports.ladocumentationfrancaise.fr/BRP/074000726/0000.pdf>

Le recours massif et diversifié au téléchargement illégal observé en France et dont les effets économiques négatifs impactent la création et les industries culturelles, s'explique par l'utilisation de technologies adaptées à tous les types d'œuvres et en constante évolution : téléchargement par un réseau de « peer to peer », mise à disposition de contenus illégaux sur des sites hébergeant des contenus. Inversement, le téléchargement légal peine à se développer, compte tenu des conditions dans lesquelles il s'exerce : verrouillage de l'œuvre achetée par des mesures techniques de protection qui limitent la libre utilisation et la conservation de celle-ci, disponibilité tardive de l'œuvre. Tel est le constat du présent rapport dont l'objectif est d'inverser cette tendance, de trouver les moyens de développer l'offre légale d'œuvres sur internet et de désinciter l'offre illégale, en complément des outils déjà existants, tant juridiques que techniques, qui peuvent être mis en œuvre.

1.26 Puel Hélène, "Techniquement il n'y a aucune contrainte de filtrage de contenu pour les FAI", 23 août 2007, <http://www.01net.com/editorial/356835/-techniquement-il-n-y-aucune-contrainte-au-filtrage-de-contenu-par-les-fai/>

Interview du représentant France d'Audible Magic qui explique la manière dont une solution de filtrage de contenus est mise au point.

1.27 Rees Marc, "Advestigo : une solution de filtrage pour les sites de vidéo", 25 janvier 2007, <http://www.pcimpact.com/actu/print.php?id=34266&c=1>

Annonce de la sortie du logiciel Advestigate : système automatisé de filtrage de contenus sous copyrights fonctionnant grâce à une comparaison, pour chaque fichier vidéo, de leur empreinte numérique.

1.28 Rees Marc, "MySpace adopte le filtre anti-piratage de Gracenote", 31 octobre 2006, <http://www.pcimpact.com/actu/news/32428-myspace-gracenote-musique.htm>

MySpace filtre le contenu importé par les usagers en utilisant la technologie de tatouage numérique de Gracenote.

1.29 Rees Marc, "Filtrage : Viacom insatisfait des efforts de Google sur Youtube", 22 octobre 2007, <http://www.pcimpact.com/actu/news/39584-viacom-charte-google-youtube-filtrage.htm>

Description du fonctionnement du filtre utilisé par YouTube : les titulaires de droits fournissent à Google une copie de leurs vidéos protégées par le droit d'auteur afin que Google élabore une matrice de comparaison des données qui sont ensuite importées sur YouTube.

1.30 Rees Marc, "Le système anti-spam DKIM sur la voie de la standardisation", 25 mai 2007, <http://www.pcimpact.com/actu/news/36578-DKIM-standardisation-IETF-RFC-DomainKeys-Ide.htm>

Alliance entre le système « Internet Identified Mail » de Cisco Systems et du système « DomainKeys » de Yahoo. Le système DKIM consiste à joindre une signature numérique chiffrée dans le courrier électronique pour s'assurer de l'identité de l'expéditeur afin d'éliminer le spam.

1.31 Saiz Jérôme, "Websense avale SurfControl", 27 avril 2007, <http://www.lesnouvelles.net/articles/business/websense-achete-surfcontrol>

Rachat par Websense de Surfcontrol. Websense, principalement connue pour le filtrage d'URL, peut désormais compter sur des technologies de filtrage spam, spyware...

1.32 Seobook, "Google Adwords & Yahoo ! PPC Tips", 1^{er} décembre 2006, <http://www.seobook.com/overture-adwords.pdf>

Description du fonctionnement du système des « adwords » de Google, au travers notamment d'une comparaison avec le système « Pay Per Click » de Yahoo.

1.33 Team criteo, "Allo Ciné choisit la technologie Criteo", 11 octobre 2006, <http://blog-fr.criteo.com/?p=102>

La technologie Criteo permet un filtrage collaboratif très fiable.

1.34 Yahoo, "Option de ciblage",

http://help.yahoo.com/l/fr/yahoo/vsm/sps/start/overview_matchtypes.html#standard

L'option de ciblage de Yahoo est un filtre par liste de mots-clés permettant grâce aux mots-clé choisis par l'utilisateur d'orienter la recherche vers le site internet ciblé.

1.35 Youtube, "Version bêta du service d'identification video Youtube",

http://fr.youtube.com/t/video_id_about/

Description du service d'identification vidéo mis en place par YouTube, à l'occasion de laquelle il est proposé aux titulaires de droits de participer au programme.

1.36 "Comment Google attribue un score à une page Web",

www.scriptol.fr/seo/brevet-google-pagerank.php

Explication sur la manière dont Google attribue à chaque page un score, score qui détermine ensuite la position de la page dans les résultats proposés par le moteur de recherche. A cette occasion, les causes de l'effet dit « sandbox » sont dévoilées.

1.37 "Gracenote acquiers cutting edge audio fingerprinting technology from philips electronics and announces long-term research agreement with Philips research",

<http://sev.prnewswire.com/computerelectronics/20050830/SFTU06730082005-1.html>

Accord signé entre Gracenote et Philips afin d'améliorer le système de filtrage audio basé sur un système d'empreinte digitale et déjà développé par Gracenote.

1.38 "MPO Emedia est choisie par l'INA comme partenaire technique pour son filtrage signature", 23 novembre 2007,

<http://www.secteurpublic.fr/public/article.tpl?id=11544&rub=8270&t=MPO+EMEDIA+est+choisi+par+l%27INA+comme+partenaire+technique+pour+sa+technologie+de+filtrage+%22signature%22>

La technologie utilisée par le système « signature » est une technologie de marquage de type « watermarking ».

2. WEBOGRAPHIE GENERALE

2.1 Drothier Yves JDN solution, "Filtrage web face aux nouveaux usages de l'internet en entreprise", 11 mai 2006,

<http://www.journaldunet.com/solutions/0605/060511-filtrage-contenu-evolution.shtml>

Réflexions visant à rendre le filtrage web dans les entreprises le plus optimum possible, notamment au regard de toutes les évolutions que ce domaine a pu et continue à connaître.

2.2 Lemire Daniel, "Le filtrage collaboratif par article",

http://benhur.telug.ugam.ca/SPIP/inf6460/article.php3?id_article=104&id_rubrique=17&sem=Semaine%2014

Description des différents algorithmes de recommandation qui permettent d'effectuer un filtrage collaboratif par article.

2.3 Ormes Sarah UKOLN the library association and UKOLN, "Introduction to filtering", <http://www.ukol.ac.uk/public/earl/issuepapers/filtering.html>

La meilleure utilisation et les inconvénients de trois logiciels de filtrage : le Keyword Blocking, le Site Blocking et le Web Rating Systems.

2.4 Villeneuve Nart, "Choisir sa technique pour contourner la censure", www.rsf.org/article.php3?id_article=14981

Analyse des différentes techniques de filtrage des contenus internet et des différentes technologies permettant de contourner ce filtrage.

2.5 Wikipédia, "Censure de l'internet", http://fr.wikipedia.org/wiki/Censure_de_l%27Internet

Les quinze principaux Etats « ennemis d'internet », les différentes techniques de censure étatique, ainsi que les principaux opposant à cette censure.

2.6 Wikipédia, "Exploration de données", http://fr.wikipedia.org/wiki/Exploration_de_donn%C3%A9es

Processus d'extraction, par des méthodes automatiques ou semi-automatiques, d'un savoir ou d'une connaissance à partir de grandes quantités de données. Ce processus peut s'accompagner d'outils d'analyse lexicographique dits de « text mining » et ainsi être utilisé dans la lutte contre le spam, et de manière plus générale, dans le domaine de l'analyse de contenus.

2.7 "FAQ web mining", <http://www.web-datamining.net/forum/faq.asp>

Foire aux questions expliquant ce qu'est le Web Mining, son utilité et ses limites, le différenciant du Web Content Mining, du Web Structure Mining, du Web Usage Mining, du data Webhouse, du filtrage collaboratif et d'un simple fichier « log ».

2.8 Wikipédia, "Filtrage collaboratif", http://fr.wikipedia.org/wiki/Filtrage_collaboratif

Systèmes de recommandation utilisant les opinions et évaluations d'un groupe pour aider l'individu. Il existe des systèmes de filtrage « actif », « passif », « utilisateurs » et « objets ». Ces systèmes de filtrage collaboratif peuvent être exploités à des fins commerciales.

2.9 Wikipédia, "Filtrage d'internet", [http://fr.wikipedia.org/wiki/Filtrage-d'internet](http://fr.wikipedia.org/wiki/Filtrage-d%27internet)

Description des objectifs poursuivis par les différents systèmes de filtrage mis en place sur internet, de leurs aspects techniques, et des perspectives dessinées par le projet PRINCIP.

2.10 Wikipédia, "Filtrage usenet", http://wikipedia.org/wiki/Filtrage_de_Usenet

Décision de ne pas diffuser des forums ou hiérarchies dont le contenu serait jugé illégal ou inintéressant ou qui seraient destinés à accueillir des contenus binaires, par là-même, très volumineux.

2.11 Wikipédia, "Tatouage numérique",

http://fr.wikipedia.org/wiki/Tatouage_num%C3%A9rique

Définition du tatouage numérique comme une technique permettant d'ajouter sur un fichier audio, vidéo ou image, de façon visible ou invisible, des informations, notamment de copyright.

3. ARTICLES ET RAPPORTS CONCERNANT LES PROJETS SCIENTIFIQUES ET DE RECHERCHE SUR LES TECHNOLOGIES PERMETTANT LE FILTRAGE DE CONTENUS

3.1 Alata Olivier, Augereau Bertrand, Carre Philippe et Tremblais Benoit, SIC, « *ICONES, Images Couleur, mouvemeNt, rElief et Surfaces* », « *Optimisation de modèles pour les signaux et Images Multicomposantes (MIM)* »,

<http://www.sic.sp2mi.univ-poitiers.fr/themes/icones/icones-001.php>

Construction et optimisation de modèles multi-échelle et multi-résolution ainsi que probabilistes, sur tout la chaîne de traitements d'images et de vidéos couleur, en prenant en compte la nature vectorielle des informations et éventuellement des spécificités du système visuel humain, en liaison avec le thème PERLE. Les modèles mathématiques essentiellement étudiés se rapportent aux processus différentiels itératifs, aux techniques Espace-Echelle-Fréquence et aux processus stochastiques.

3.2 ANR, "Edition 2006 du Programme « Masse de Données et Connaissances Ambiantes »", 2006,

<http://www.agence-nationale-recherche.fr/documents/aap/2006/selection/mdca.pdf>

Liste des projets sélectionnés pour le programme « Masse de Données et Connaissances Ambiantes ».

3.3 Augereau Bertrand, Helbert David et Tremblais Benoit, Sic, « *ICONES, Images Couleur, mouvemeNt, rElief et Surfaces* », « *Contenu Implicite des Séquences d'Images (CISI)* »,

<http://www.sic.sp2mi.univ-poitiers.fr/themes/icones/icones-004.php>

Divers problèmes spécifiques aux séquences d'images scalaires ou vectorielles, notamment à l'objet implicite contenu dans ces séquences, le mouvement.

3.4 Balakrishnan Hari, Shenker Scott, Welfish Michael, MIT Laboratory for Computer Science, « *Semantic-Free Referencing in Linked Distibued Systems* »,

<http://nms.lcs.mit.edu/papers/sfr-iptps03.pdf>

Description des méthodes de référencement et des différents composants des liens de référencement.

3.5 Berkman Center for Internet & Society, Harvard, « *The Filter, november 2007* »,

<http://cyber.law.harvard.edu/node/485>

Rassemblement de plusieurs articles relatifs au filtrage :

- l'interopérabilité de l'internet ;
- second forum de la gouvernance d'internet ou « Internet Governance Forum », Rio de Janeiro, 2007 ;
- influence des technologies sur le débat politique.

3.6 Besançon Romaric et de Chalendar Gaël, "L'analyseur syntaxique de LIMA dans la campagne d'évaluation EASY", juin 2005,

http://www.list.cea.fr/fr/publications/docs/si/ingenierie_connaissance/fr/TALN2005_besancon_chalendar_easy.pdf

Le LIC2M, laboratoire du CEA/LIST, a participé à la campagne d'évaluation « Easy » avec l'analyseur syntaxique de son système « Lima », un analyseur syntaxique robuste qui implémente une grammaire de dépendance. Les résultats obtenus sur le corpus d'exemples sont encourageants et permettent de valider les techniques utilisées. En revanche, le traitement de corpus plus généraux couvrant des phénomènes syntaxiques plus variés nécessiteront sûrement le développement de ressources supplémentaires ou la mise en place de traitements particuliers.

3.7 Bimbot Frédéric, "Thème de recherche Systèmes Cognitifs, METISS : Modélisation et expérimentation pour le traitement des informations et des signaux sonores (équipe-projet)", 2007, www.inria.fr/recherche/equipes/metiss.fr.html

Rapport d'activité de l'équipe projet « Metiss » comportant trois volets : la caractérisation du locuteur (notamment pour la vérification vocale d'identité), le suivi de locuteur et des classes de sons pour l'indexation d'enregistrements sonores et le traitement « avancé » de signaux sonores (par exemple, la séparation de sources dans le cas sous-déterminé). Les fondements scientifiques s'inscrivent dans le cadre des mathématiques appliquées, du traitement du signal, de la modélisation probabiliste, de l'estimation statistique et de la théorie de la décision. L'équipe projet s'appuie sur les outils de traitement de signal au niveau de la représentation du signal (représentations adaptives), de sa paramétrisation (analyse spectrale) et de sa décomposition (séparation de sources). Les approches probabilistes interviennent au niveau de la modélisation acoustique (modèle de distribution) et de la classification (tests d'hypothèses et reconnaissance). Les travaux font également place à des algorithmes de décodage et de poursuite tels que, par exemple, l'algorithme de Viterbi et la Matching Pursuit. Les principaux secteurs industriels concernés sont le secteur des télécommunications, de l'internet et du multimédia, et sont susceptibles de s'étendre aux domaines de la production musicale et audiovisuelle et des logiciels éducatifs et des jeux.

3.8 Bletsas Aggelos, Khisti Ashish, Reed David P., Lippman Andrew, IEEE, "A simple Cooperative Diversity Method Based on Network Path Selection", mars 2006,

<http://pubs.media.mit.edu/pubs/papers/IEEE,March2006.pdf>

Description du « diversité coopérative » ou "Cooperative diversity" qui est un système permettant de générer des gains au regard de l'affaiblissement de la puissance des transmissions sans fil.

3.9 Bouali Fatma, Mongy Sylvain et Djeraba Chabane, "Analyzing User's Behavior On A Video Database", 2005, Sixth International Workshop on Multimedia Data Mining "Mining Integrated Media and Complex Data", Chicago, USA,

<http://perso.numericable.fr/~mondanis/publis/MDMBookChapter23.pdf>

L'analyse du comportement des utilisateurs dans les grosses banques de données est un problème émergent. La croissance importante de la vidéo dans la vie de tous les jours est en lien direct avec

l'usage de la vidéo. Les utilisateurs de ces vidéos ont besoin de systèmes logiciels intelligents qui mettent à profit la source d'informations cachées dans le comportement de l'utilisateur sur de grandes bases de données vidéo pour récupérer et parcourir les vidéos. Les auteurs proposent deux niveaux de modèle pour l'approche fondée sur la modélisation des comportements des utilisateurs sur moteur de recherche vidéo. Le premier modèle niveau vise à modéliser le comportement de l'utilisateur et le regroupement sur une seule séquence vidéo. Le second modèle vise à modéliser le comportement des utilisateurs et de regroupement sur un ensemble de séquences vidéo.

3.10 Bouali Fatma, Mongy Sylvain, Djeraba Chabane, "Video Usage Mining", 2006, Encyclopedia of Multimedia, pages 928-934.. Edited by Borko Furht, <http://perso.numericable.fr/~mondenis/publis/videousagemining.pdf>,

Utilisation du data mining vidéo pour générer des profils d'utilisateurs sur un moteur de recherche vidéo dans le contexte de la production cinématographique.

3.11 Boujemaa Nozha, "Thème de recherche Systèmes Cognitifs, IMEDIA : Images et multimédia : indexation, navigation et recherche (équipe-projet)", 2007, www.inria.fr/recherche/equipes/imedia.fr.html

Rapport d'activité de l'équipe projet « Imedia » ayant pour objectif de développer des méthodes d'indexation par le contenu, de recherche interactive et de navigation dans des bases d'images, dans un contexte multimédia. Le rapport traite des bases d'images « génériques » (web) et des bases d'images « spécifiques » à un domaine d'application ciblé (visages, images médicales, ...). Ces deux catégories relèvent respectivement de la recherche d'images et de la reconnaissance d'objets. En réalité, la recherche d'images est une problématique plus vaste qui englobe la reconnaissance d'objets et intègre les interactions avec l'utilisateur. Plus généralement, ce rapport a pour objet de répondre au problème complexe de l'accès intelligent aux données multimédia dans leur globalité.

3.12 Business Development Research Consultants BDRC, "Intermediate Evaluation of the Safer Internet Action Plan conducted for The European Commission volume 1 Final Report", 31 mai 2001, www.bdrc.co.uk

Le premier volume présente la méthodologie ainsi que les opinions de l'ensemble des participants. Ce volume décrit les conclusions et recommandations en matière d'actions à mener, notamment en ce qui concerne la classification et le filtrage, celles-ci étant des technologies émergentes.

3.13 Business Development Research Consultants BDRC, "Intermediate Evaluation of the Safer Internet Action Plan conducted for The European Commission volume 2 Context and Appendices", 31 mai 2001, www.bdrc.co.uk

Le second volume présente le contexte légal entourant les contenus illicites et les nouvelles technologies justifiant une surveillance.

3.14 Carre Philippe, Sic, « *ICONES, Images Couleur, mouvemeNt, rElief et Surfaces* », <http://www.sic.sp2mi.univ-poitiers.fr/themes/icones/index.php>

Les activités de recherche de l'équipe ICONES se sont organisées autour du traitement, de la caractérisation et de l'analyse de signaux et images multisources et multicomposantes avec une spécificité concernant les images couleur texturées statistiques et dynamiques. Les objectifs sont de développer des modèles pour traiter et analyser les images numériques à travers les échelles et à travers le temps, et de gérer leur reproduction sur différents supports, et ce en intégrant éventuellement des connaissances sur les contenus déterminées a priori.

3.15 Caromel Denis, "Thème de recherche *Systèmes communicants, OASIS : Objets actifs, sémantique, Internet et sécurité (équipe-projet)*", 2007, www.inria.fr/recherche/equipes/oasis.fr.html

Rapport d'activité de l'équipe projet Oasis ayant pour objectif de proposer des principes fondamentaux, des techniques et des outils pour la construction, l'analyse, la validation, la vérification et la maintenance des systèmes fiables, dans le cadre des applications réparties (réseaux internet et intranet, cartes à puce et terminaux).

3.16 CEA LIST, « *Les Systèmes Interactifs* », http://www-list.cea.fr/fr/programmes/systemes_interactifs/systemes_interactifs.htm

Après avoir relevé que l'interaction Homme-Système, au cœur des systèmes intelligents de demain, est omniprésente dans la vie quotidienne des citoyens et dans l'entreprise, expliquant que l'interactivité, située au cœur d'un triptyque Homme – Information – Environnement, est au centre de tous les grands programmes de recherche dans le monde, la Recherche Développement du LIST s'est structurée selon trois axes :

- ingénierie de la connaissance ;
- robotique ;
- réalité virtuelle et interfaces sensorielles.

3.17 CEA LIST, « *Multimedia Multilingual Knowledge Engineering Laboratory(LIC2M)* », http://wwwlist.cea.fr/fr/programmes/systemes_interactifs/docs/gb/poster_lic2m_images_gb.pdf, http://wwwlist.cea.fr/fr/programmes/systemes_interactifs/docs/gb/poster_lic2m_texte_gb.pdf

Présentation de l'activité de traitement de l'information visuelle et textuelle de l'équipe de recherche « Réalité virtuelle, cognitive et interfaces », qui traite l'axe de l'ingénierie de la connaissance.

3.18 CEA et Bull, communiqué de presse, « *Le CEA et BULL annoncent une performance record pour la recherche d'images dans les très grandes bases de données* », 4 février 2008, http://wwwlist.cea.fr/fr/actualites/actu_2008/CP_CEA_Bull_RecordRechercheImages.pdf, <http://www.bull.com/fr/bulldirect/N23/hot.html>

Le Centre d'énergie atomique (CEA) et Bull annoncent avoir atteint une performance record dans la recherche d'images dans les très grandes bases de données. Ainsi, le nouveau moteur permet d'effectuer une recherche de 3,7 millions d'images par seconde, ce qui est cinq fois plus rapide que précédemment. Cette performance record a été obtenue sur un supercalculateur conçu et fourni par Bull, en utilisant le logiciel de recherche multimédia spécialement développé par le CEA LIST dans le cadre du projet FAME2. Il ouvre la voie à un vaste champ applicatif allant de

la veille stratégique à la comparaison d'images médicales, des « fouilles » de données sur internet au commerce électronique ou à la gestion de contenu.

3.41 CEA, « DETECT : un suivi et une détection de thèmes multilingues », [http://www-list.cea.fr/fr/programmes/systemes_interactifs/docs/ingenierie_connaissance/projet DETECT.pdf](http://www-list.cea.fr/fr/programmes/systemes_interactifs/docs/ingenierie_connaissance/projet_DETECT.pdf)

Le projet DETECT a pour objet le suivi et la détection des sujets d'intérêt qui peuvent être exprimés en différentes langues. Ce projet a abouti à un démonstrateur d'un système capable d'identifier des thèmes dans des documents multilingues.

3.19 Chaar Sana-Leila, Ferret Olivier et Fluhr Christian, « Filtrage multi-document orienté par un profil utilisateur », octobre 2002, http://wwwlist.cea.fr/fr/publications/docs/si/ingenierie_connaissance/fr/chaar_cide_2002.pdf

Présentation d'une méthode de filtrage permettant de sélectionner à partir d'un ensemble de documents, les extraits de textes les plus significatifs relativement à un profil défini par un utilisateur. Pour ce faire, l'accent a été mis sur l'utilisation conjointe de profils structurés et d'une analyse thématique des documents. Cette analyse permet en particulier d'étendre le vocabulaire définissant un profil en fonction du document traité en sélectionnant les termes de ce dernier les plus étroitement liés aux termes du profil. Cette capacité ouvre ainsi la voie à une plus grande finesse du filtrage en permettant la sélection d'extraits de documents ayant un lien plus ténu avec les profils mais davantage susceptibles d'apporter des informations nouvelles et donc intéressantes.

3.20 Chaar Sana-Leila, « Extraction de segments thématiques pour la construction de résumé multi-document orienté par un profil utilisateur », juin 2003, http://wwwlist.cea.fr/fr/publications/docs/si/ingenierie_connaissance/fr/chaar_recital_2003.pdf

Cet article présente une méthode qui vise à donner à un utilisateur la possibilité de parcourir rapidement un ensemble de documents par le biais d'un profil utilisateur. Un profil est un ensemble de termes structuré en sous-ensembles thématiquement homogènes. L'analyse des documents se fonde sur l'extraction des passages les plus étroitement en relation avec ce profil. Cette analyse permet en particulier d'étendre le vocabulaire définissant un profil en fonction du document traité en sélectionnant les termes de ce dernier les plus étroitement liés aux termes du profil. Cette capacité ouvre ainsi la voie à plus grande finesse du filtrage en permettant la sélection d'extraits de documents ayant un lien plus ténu avec les profils mais davantage susceptibles d'apporter des informations nouvelles et donc intéressantes. La production du résumé résulte de l'appariement entre les segments délimités lors de l'analyse des documents et les thèmes du profil.

3.21 Chaar Sana-Leila, Ferret Olivier et Fluhr Christian, « Génération de résumé multi-document guidé par un profil utilisateur », mars 2004, http://wwwlist.cea.fr/fr/publications/docs/si/ingenierie_connaissance/fr/setit_2004_chaar.pdf

L'article présente une méthode qui permet la construction de résumés multidocuments orientés par un profil utilisateur. Cette méthode vise à donner à un utilisateur la possibilité de parcourir rapidement un ensemble de documents selon un point de vue particulier. Ce point de vue est représenté par le biais d'un profil utilisateur. L'article met l'accent sur l'utilisation conjointe de profils structurés et d'une analyse discursive des documents pour extraire des passages de textes en rapport avec les attentes de l'utilisateur. Cette capacité ouvre ainsi la voie à une plus grande

finesse du filtrage et permet de sélectionner les extraits des documents ayant un lien plus ténu avec les profils. La production du résumé résulte de l'appariement entre les segments délimités lors de l'analyse des documents et les thèmes du profil.

3.22 Chaumette François, "Thème de recherche Systèmes Cognitifs, LAGADIC : Asservissement visuel en robotique, vision et animation (équipe-projet)", 2007, www.inria.fr/recherche/equipes/lagadic.fr.html

Rapport d'activité de l'équipe projet « Lagadic » ayant pour objectif de modéliser et d'élaborer des stratégies de perception et d'action autour des techniques d'asservissement visuel pour des applications dans les secteurs de la robotique, de la vision par ordinateur, de la réalité augmentée, de l'animation virtuelle et de la cognoscience. Lagadic dispose d'un parc robotique constitué d'un robot manipulateur, d'une cellule de vidéosurveillance, d'un robot mobile et en commun avec l'équipe-projet « Visages », d'un robot médical.

3.23 CNRS Limsi, "Rapport d'activité 2007", 2007, <http://rs2007.limsi.fr/RS2007FFdef.pdf>

Ce rapport d'activité couvre la période 2005-2007, et reprend l'ensemble des recherches effectuées par l'Unité propre de recherche du CNRS (Limsi), dont quatre concernent plus particulièrement la présente bibliographie :

- « Groupe Langues, Information et Représentations (LIR) : les activités de recherche de ce groupe sont essentiellement consacrées au traitement des données écrites, à leur analyse, leur compréhension ou leur reproduction ainsi qu'à l'acquisition de connaissances nécessaires, principalement morphologiques et sémantiques. La quantité impressionnante de données écrites aujourd'hui disponibles électroniquement est une mine d'informations et la fouille de données dans les textes est un des enjeux majeurs de la société de l'information. Les recherches développées dans le groupe LIR s'inscrivent dans cette dynamique, avec une implication croissante dans des projets nationaux et internationaux. Les compétences variées et complémentaires des membres du groupe LIR permettent de combiner approches symboliques et statistiques, et constituent un des atouts majeurs du groupe qui participe ainsi pleinement à l'évolution du traitement des langues ».
- « Groupe Traitement du Langage Parlé (TLP) » : les recherches de ce groupe portent sur la modélisation de la parole et son traitement automatique. Pour extraire et structurer l'information présente dans un document audio, le groupe de recherche développe des modèles et des algorithmes fondés sur la prise en compte conjointe des diverses sources d'information visant à un processus global de décodage du signal. Ces recherches sur les modélisations acoustique, lexicale, et linguistique, sont réalisées dans un contexte multilingue et s'appuient sur de grands corpus oraux représentatifs de nombreux domaines applicatifs ».
- « Action Transversale COPTE : Corpus Parole Texte Evaluation » : l'action transversale COPTE fait le lien entre deux domaines du Traitement Automatique du Langage Naturel : la reconnaissance de la parole et l'analyse de l'écrit. L'objectif est de fusionner les approches propres aux deux domaines sur des problèmes ouverts situés à l'interface des deux disciplines. Le domaine de la reconnaissance de la parole aborde l'analyse du langage par l'étude du signal sonore et doit donc nécessairement prendre en compte les aspects propres à la parole : temporalité et spontanéité. De son côté l'analyse de l'écrit, aborde l'analyse du langage par l'étude des signes, où les aspects qui priment sont plutôt la nature statique et préparée du support d'information étudié ».

- « Action Thématique Sémantique et Mémoire Episodique » : les recherches ont été centrées sur deux thèmes : (1) l'exploration du contenu de la mémoire sémantique et la relation entre mémoire sémantique et mémoire épisodique ; (2) les conditions d'élaboration des inférences causales. Ces deux séries de recherches ont permis de mettre en évidence le rôle des situations stockées dans la mémoire sémantique ».

3.24 Commission des Communautés Européennes, "Résumé de l'analyse d'impact et l'évaluation ex ante, document de travail des services de la Commission", 27 février 2008, http://ec.europa.eu/information_society/activities/sip/programme/index_en.htm

Rapport synthétique sur l'accès sur internet à un contenu illégal ou dangereux, et plus particulièrement, sur l'accès des enfants et adolescents au contenu pornographique diffusé sur internet et mise en place d'un plan d'actions avec une évaluation coûts-avantages de chaque option stratégique.

3.25 Communication de la Commission au Conseil, au Parlement Européen, au Comité Economique et Social Européen et au Comité des régions, "Evaluation finale de la mise en œuvre du plan d'action communautaire pluriannuel visant à promouvoir une utilisation plus sûr d'internet par la lutte contre les messages à contenu illicite et préjudiciable diffusés sur les réseaux mondiaux", 6 novembre 2006, http://ec.europa.eu/information_society/activities/sip/programme/decision/index_en.htm

Evaluation finale du plan d'actions pour un internet plus sûr de 2003 à 2004, réalisée par trois experts indépendants. Le programme initial comportait quatre lignes d'actions : ligne directe et sensibilisation, filtrage, classement et autorégulation. Au titre de l'année 2003-2004, le financement de l'Europe ne concernait que la ligne directe et la sensibilisation. Au titre du filtrage, il est signalé que les entreprises ont réalisé de grands progrès techniques dans ce domaine et offrent différentes solutions aux utilisateurs. La labellisation et le classement restent encore fondamentaux pour garantir un internet plus sûr.

3.26 Daoudi Mohamed, Filali Ansary Tarik, Vandeborre Jean-Philippe, Equipe FOX-MIIRE, “A framework for 3d CAD models retrieval from 2D images”, <http://www.telecom-lille1.eu/people/vandeborre/papers/filaliAnnalsTelecom2005.pdf>

La gestion de grandes bases de données de modèles tridimensionnels (utilisés dans des applications de CAO, de visualisation, des jeux ...) est un domaine très important. La capacité de caractériser et rechercher facilement des modèles est une question clé pour les concepteurs et les utilisateurs finaux. Deux approches principales existent : la recherche par l'exemple d'un modèle tridimensionnel, et la recherche par une vue 2D ou des photos. L'article présente un cadre pour la caractérisation d'un modèle 3D par un ensemble de vue (appelées vue caractéristiques), et un processus d'indexation de ces modèles avec une approche probabiliste bayésienne en utilisant les vues caractéristiques. Le système est indépendant du descripteur utilisé pour l'indexation. L'article illustre les résultats en utilisant différents descripteurs sur une collection de modèles tridimensionnels fournis par le constructeur automobile Renault. Il présente également les résultats sur l'indexation de modèles 3D à partir de photos.

3.27 Daouadi Mohamed Tierny Julien et Vandeborre Jean-Philippe, "Graphes de Reeb de Haut Niveau de Maillages Polygonaux 3D", 14 juin 2006, http://www.telecom-lille1.eu/people/tierny/stuff/papers/tierny_3dpvt06_pre.pdf

Cet article présente une méthode originale pour la construction de graphes de Reeb invariants de haut niveau-entités topologiques qui offrent une bonne vue d'ensemble de la structuration d'un objet 3D. Dans ce but, nous proposons un algorithme d'extraction de sommets caractéristiques simple et précis. Ces sommets sont utilisés pour le calcul d'une fonction d'application invariante, visuellement intéressante. De plus, nous proposons un nouvel algorithme de construction de graphe de Reeb, basé sur l'analyse de connexité de lignes de niveau discrètes. Cet algorithme apporte une solution pratique au problème de suppression de points critiques non significatifs, produisant en sortie des graphes bénéficiant de bonnes propriétés descriptives. L'invariance géométrique de ces graphes et leur forte tolérance à la variation de pose du modèle et à la variation d'échantillonnage du maillage en font de bons descripteurs, exploitables dans diverses applications, comme la déformation de maillage, la compression, l'indexation 3D, la métamorphose, etc.

3.28 Djeraba Chabane, Daoudi Mohamed, Tombelle Christophe et Zeng Huicheng, “Adult Image Filtering For Internet Safety”, http://www-rech.enic.fr/MIIRE/publis/Multimedia_security.pdf

Pour faire face au filtrage de contenus, trois notions sont nécessaires : la recherche d'information d'image, la catégorisation de texte et la catégorisation d'image. Même si ces trois notions sont nécessaires à l'efficacité du filtrage de contenus, les auteurs se concentrent sur le champ d'application de la communication sur l'image et du filtrage de contenus.

3.29 Djeraba Chabane, Lew Stanislas, Simovici Dan, Mongy Sylvain et Ihaddadene Nacim, *Eye/gaze Tracking in web , image and video documents*", octobre 2006, http://perso.numericable.fr/~mondenis/publis/lew_eyegaze.pdf

La démonstration se concentre sur l'eye tracking sur le Web, d'images et de données vidéo. Elle utilise l'état de l'art des mesures, tel que, par exemple, le scan path, afin de déterminer comment l'utilisateur voit les documents Web, les images et les vidéos.

3.30 Djeraba Chabane et Bouali Fatma, "Recherche textuelle et visuelle-Indexation par concepts", mai 2004, <http://www-rech.enic.fr/coresa2004/articles/p193-djeraba.pdf>

Les concepts utilisateurs permettent :

- l'exploitation des descriptions visuelles et textuelles ;
- une intégration simple et naturelle des connaissances liées au domaine d'application et leur réutilisation dans la définition d'autres concepts ou dans d'autres requêtes, ce qui évite d'avoir à gérer des connaissances a priori.

3.31 Dieng-Kuntz Rose, "Thème de recherche Systèmes Cognitifs, EDELWEISS : Echanges, Documents, Extraction, Langages, Web, Ergonomie, Interactions, Sémantique, Serveurs (équipe-projet)", 2007, www.inria.fr/recherche/equipes/edelweiss.fr.html

Rapport d'activité de l'équipe projet « Edelweiss » qui vise à proposer des modèles, des méthodes et des outils pour aider des communautés virtuelles de pratique et/ou d'intérêt à gérer leurs connaissances de manière collaborative via le Web, en interagissant avec des ressources d'information et des personnes « annotées sémantiquement », c'est-à-dire indexées par des « ontologies ». Les communautés de pratique et/ou d'intérêt sont des groupes de personnes ayant un intérêt commun ou une passion commune pour un sujet ou un problème, et qui partagent leurs idées et leurs expériences et recherchent en commun des solutions.

3.32 Enic, Eurecom, INT, Liris, Renault et TGS, "SEMANTIC-3D : Compression, indexation et tatouage de données 3D", 2004, <http://www-rech.enic.fr/coresa2004/articles/p045-lavoue2.pdf>

Les avancées récentes en imagerie 3D rendent possible la création, le stockage, mais aussi la transmission de modèles 3D. Le projet SEMANTIC-3D vise à développer de nouvelles techniques d'indexation, de compression et de tatouage d'objets 3D dans le cadre d'une application industrielle pour un système d'information et de communication sécurisée dans lequel la variété des terminaux et des réseaux utilisés doit être prise en compte.

3.33 Feamster Nick, Balazinska Magdalena, Harfst Greg, Balakrishnan Hari, Karger, David, MIT Laboratory for Computer Science, "Infranet : Circumventing Web Censorship and Surveillance", <http://wind.lcs.mit.edu/papers/usenixsec2002.pdf>

De plus en plus de pays et d'entreprises bloquent l'accès à internet. « Infranet » est une alternative à ces mesures. Il s'agit d'un système qui fournit aux clients l'accès aux sites censurés tout en leur permettant de continuer à héberger des contenus non sensibles et non censurés. Ce système consiste, notamment, à fournir du contenu censuré en transformant les données censurées en des images non censurées.

3.34 Fekete Jean-Daniel, « Thème de recherche Systèmes Cognitifs, AVIZ : Analyse et Visualisation », 2007, www.inria.fr/recherche/equipes/aviz.fr.html

Rapport d'activité de l'équipe projet « Aviz » dont le projet, pluridisciplinaire, vise à améliorer les méthodes d'analyse et de visualisation de grandes quantités de données en intégrant profondément le processus d'analyse et celui de visualisation d'informations pour permettre de comprendre plus facilement et rapidement ces données. L'équipe projet s'est essentiellement concentrée sur la visualisation de très gros réseaux (de l'ordre de millions de sommets et d'arrêtes) et les séries

temporelles (plusieurs milliards d'enregistrements capturés en continu et en temps réel). Les domaines d'application incluent l'analyse de grands réseaux sociaux (Wikipédia, les développeurs de logiciels libres), les réseaux biologiques, l'intelligence économique, les bibliothèques numériques et les séries temporelles issues des activités de chercheurs.

3.35 Ferret Olivier, « Segmenter et structurer thématiquement des textes par l'utilisation conjointe de collocations et de la récurrence lexicale », juin 2002, http://www-list.cea.fr/fr/publications/docs/si/ingenierie_connaissance/fr/chaar_cide_2002.pdf

L'article expose une méthode réalisant de façon intégrée deux tâches de l'analyse thématique : la segmentation et la détection de liens thématiques. Cette méthode exploite conjointement la récurrence des mots dans les textes et les liens issus d'un réseau de collocations afin de compenser les faiblesses respectives des deux approches. L'article présente son évaluation concernant la segmentation sur un corpus en français et un corpus en anglais, et propose une mesure d'évaluation spécifiquement adaptées à ce type de systèmes.

3.36 Ferret Olivier, "Filtrage thématique d'un réseau de collocations", juin 2003, http://wwwlist.cea.fr/fr/publications/docs/si/ingenierie_connaissance/fr/taln2003_ferret_filtre_thematique.pdf

Les réseaux lexicaux de type « WordNet » présentent une absence de relations de nature thématiques, relations pourtant très utiles dans des tâches telles que le résumé automatique ou l'extraction d'information. Cet article propose une méthode visant à construire automatiquement à partir d'un large corpus un réseau lexical dont les relations sont préférentiellement thématiques. En l'absence d'utilisation de ressources de type dictionnaire, cette méthode se fonde sur un principe d'auto-amorçage : un réseau de collocations est d'abord construit à partir d'un corpus puis filtré sur la base des mots du corpus que le réseau initial a permis de sélectionner. L'article montre au travers d'une évaluation portant sur la segmentation thématique que le réseau final, bien que de taille bien inférieure au réseau initial, permet d'obtenir les mêmes performances que celui-ci pour cette tâche.

3.37 Guillemot Christine, "Thème de recherche Systèmes Cognitifs, TEMICS : Traitement, modélisation et communication d'images numériques (équipe-projet)", 2007, www.inria.fr/recherche/equipes/temics.fr.html

Rapport d'activité de l'équipe projet « Temics » ayant pour objectif de développer les concepts et les outils d'analyse, de modélisation, de codage, et de tatouage d'images, et plus généralement, des informations vidéo manipulées en communication multimédia. Les travaux de l'équipe projet portent plus particulièrement sur les problèmes suivants :

- l'interaction avec le contenu et la navigation dans les scènes vidéo 3D ;
- la représentation compacte et robuste aux bruits de transmission des images et des signaux vidéo ;
- le marquage ou « tatouage » des images et des signaux vidéo à des fins de protection contre les copies illicites, et à des fins d'authentification.

3.38 Gros Patrick, "Thème de recherche Systèmes symboliques, TEXMEX : Techniques d'exploitation des données multimédia (équipe-projet)", 2007, www.inria.fr/recherche/equipes/textmex.fr.html

Rapport d'activité de l'équipe projet « Texmex » dont les deux axes de travail sont :

- la définition et l'évaluation de nouveaux descripteurs de documents pour les images fixes, les vidéos et les textes et de descripteurs faisant intervenir plusieurs média et de méta-données associées aux documents ;
- l'élaboration de statistiques pour l'exploration des grands volumes de données, la gestion et la mise en place de stratégies de calcul des méta-données et des descripteurs associés aux documents, l'analyse de la qualité des données, l'étude de stratégies économes d'exploitation (navigation, indexation, recherche) et la définition de supports systèmes et matériels pour un accès rapide à ces données.

3.39 Horaud Radu Patrice, "Thème de recherche Systèmes Cognitifs, PERCEPTION : Interprétation et Modélisation d'Images et de Vidéos", 2007, www.inria.fr/recherche/equipes/perception.fr.html

Rapport d'activité de l'équipe projet « Perception » ayant pour objectif d'interpréter des images et des vidéos en termes de représentations visuelles tri-dimensionnelles et de descriptions symboliques. L'approche utilise la théorie de l'information et des modèles cognitifs. D'une part, l'équipe projet met au point des modèles mathématiques nécessaires pour extraire des informations géométriques et physiques du stimulus (les images) ainsi que les méthodes, algorithmes et logiciels réalisant une implémentation optimale de ces modèles. D'autre part, l'équipe projet étudie les mécanismes biologiques et cognitifs qui gouvernent les interactions entre les connaissances a priori et le stimulus sensoriel : l'équipe projet essaie de décrire ces mécanismes en tant que modèles mathématiques et calculatoires en vue de les réaliser sous la forme de logiciels.

3.40 Huet Stéphane, Sébillot Pascale et Gravier Guillaume, "Thème de recherche Systèmes Cognitifs, Utilisation de la linguistique en reconnaissance de la parole : un état de l'art", mai 2006, http://hal.inria.fr/view_by_stamp.php?&halsid=bdn23muiekk1b2f0ufpu6s96b6&label=INRIA-RRRT&langue=fr&action_todo=view&id=inria-00077386&version=2

Rapport de recherche n°5917 de l'équipe projet « Edelweiss » partant du postulat que pour transcrire des documents sonores, les systèmes de reconnaissance de la parole font appel à des méthodes statistiques, notamment aux chaînes de Markov cachées et aux modèles N-grammes. Selon les auteurs, même si ces techniques se sont révélées performantes, elles approchent du maximum de leurs possibilités avec la mise à disposition de corpus de taille suffisante et il semble nécessaire, pour tenter d'aller au-delà des résultats actuels, d'utiliser des informations supplémentaires, en particulier liées au langage. Intégrer de telles connaissances linguistiques doit toutefois se faire en tenant compte des spécificités de l'oral (présence d'hésitations par exemple) et en étant robuste à d'éventuelles erreurs de reconnaissance de certains mots. Ce document présente un état de l'art des recherches de ce type, en évaluant l'impact de l'insertion des informations linguistiques sur la qualité de la transcription.

3.41 Jedynek Bruno, Zheng huicheng, Daoudi Mahamed et Barret Didier, "Maximum Entropy Models for Skin Detection", décembre 2002, <http://www.ee.iitb.ac.in/~icvgip/PAPERS/250.pdf>

Les auteurs considèrent comme une séquence de 3 modèles la « skin detection » construite à partir d'une grande collection d'images labellisées. Chaque modèle est un modèle d'entropie maximale à l'égard de contraintes concernant les distributions marginales. Les modèles sont emboîtés. Le premier modèle, appelé le modèle de base est bien connue des praticiens. Les pixels sont considérés comme indépendants. La performance, mesurée par la courbe ROC sur la base de données Compaq est impressionnante pour un modèle simple. Toutefois, seule un examen de l'image révèle des résultats très irréguliers. Le second modèle est le modèle caché de Markov qui comprend des contraintes plus douce pour la solution. La courbe de ROC obtenue montre de meilleures performances que le modèle de base. Enfin, le dégradé de couleur est inclus. Grâce au rapprochement de l'arbre de Bethe, on obtient une expression analytique simple pour les coefficients de l'entropie maximale associée au modèle. La performance, par rapport aux précédents modèles est une fois de plus améliorée.

3.42 Li Yiping, "Un système de segmentation du chinois basé sur des triplets", juin 2003, http://www.list.cea.fr/fr/publications/docs/si/ingenierie_connaissance/fr/taln2003_li_recital_li_segchinois.pdf

Un des problèmes rencontrés lors de l'analyse de textes en chinois est qu'il n'existe pas de séparateur entre les mots dans cette langue. Le mot étant une unité linguistique fondamentale en traitement automatique de la langue, il est nécessaire d'identifier les mots dans un texte chinois afin que des analyses de plus haut niveau puissent être réalisées. Le but de cet article est de présenter un système d'identification des mots basé sur un algorithme utilisant des triplets de catégories grammaticales et des fréquences de mots. Ce système comprend deux dictionnaires : l'un dédié aux mots et à leurs fréquences, l'autre aux triplets des catégories correspondantes. Les tests qui ont été effectués révèlent que 98,5% des phrases sont découpées correctement. Certaines erreurs sont dues à la taille limitée du dictionnaire utilisé. Une réflexion sur la création de nouvelles catégories et des études proposant des règles grammaticales sont en cours de réalisation afin d'augmenter la performance du système.

3.43 Minefi, "Technologies clés 2010 : Sécurisation des transactions électroniques et des contenus", novembre 2006, http://www.dgemp.minefi.gouv.fr/techno_cles_2010/html/tech_10.html

Présentation et description des technologies utilisées pour sécuriser les transactions électroniques et à prévenir, détecter et limiter les attaques malveillantes à l'encontre des systèmes mais aussi des contenus.

3.43 Minefi, "Technologies clés 2010 : Technologies du web sémantique", novembre 2006, http://www.dgemp.minefi.gouv.fr/techno_cles_2010/html/tech_13.html

Description et portée des technologies du « web sémantique » lesquelles se rapportent à un ensemble de modèles et d'outils visant à permettre aux contenus numériques d'être partagés.

3.44 Minefi, "Technologies clés", http://www.industrie.gouv.fr/techno_cles_2010/pdf/technocles2010-1.pdf

Ce rapport traite des technologies de l'information et de la communication, dont deux concernent plus particulièrement la présente bibliographie :

- la « Gestion et diffusion des contenus numériques » (textes, photos, documents composites, fichiers audio ou vidéo, logiciels...) font référence à un ensemble de technologies permettant d'organiser, d'accéder et d'acheminer les contenus tout en garantissant leur intégrité et en gérant les contraintes liées aux droits de diffusion.
- les « Technologies du Web sémantique » sont à la base des moteurs de recherche, des interfaces de navigation et des plates-formes collaboratives de demain. Cette description est plus large que la notion de Semantic Web telle que développée dans le cadre du W3C.

3.45 MIT, "The Media Laboratory United States Patents", décembre 2006

Rapport d'activité du laboratoire media du MIT, de décembre 2006, portant, notamment, sur les techniques de tatouage audio ou « Radio Frequency Identification (RFID) tags ».

3.46 MIT, "Media Laboratory : projects, october 2007", <http://www.media.mit.edu/research/ml-projects.pdf>

Rapport d'activités du laboratoire media Lab US Patents du MIT, d'octobre 2007, portant, notamment sur les modes de télécommunication sans fil et sur les techniques de tatouage audio ou « Radio Frequency Identification (RFID) tags ».

3.47 MIT, "Infranet", 21 mars 2008, <http://nms.lcs.mit.edu/projects/infranet/>

Description d'« Infranet », un système qui permettant de contourner la censure d'internet.

3.48 MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), "ELIZA Creator Remembered", 21 mars 2008, <http://www.csail.mit.edu/index.php>

Décès, le 5 mars 2008, de Joseph Weizenbaum, professeur émérite en sciences informatiques et l'un des pères de l'intelligence artificiel avec, notamment, son célèbre langage de programmation, Eliza.

3.49 MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), "Networks and Mobile Systems", <http://nms.csail.mit.edu/>

Le « Networks and Mobile Systems », un des groupes de recherche du MIT CSAIL, conduit des recherches dans des domaines variés relevant des architectures réseaux et de l'informatique mobile.

3.50 MIT Computer Science and Artificial Intelligence Laboratory (CSAIL), "Research activities", <http://www.csail.mit.edu/research/activities/activities.php>

Page web comprenant les différents liens vers les sites des différents groupes de recherche que comporte le CSAIL

3.51 MIT Laboratory for Computer Science, "Semantic-Free Referencing", <http://nms.lcs.mit.edu/projects/sfr/>

Article portant sur le référencement sur internet et proposant l'adoption d'une nouvelle méthode de référencement.

3.52 Mongy Sylvain et Djeraba Chabane, " Extraction de comportements liés à la lecture de bandes-annonces cinématographique" ,
<http://perso.numericable.fr/~mondennis/publis/mongybda2006>

Cet article est organisé comme suite. La section 2 présente l'état de l'art dans le domaine de la fouille des usages de la vidéo et en spécifie les particularités. La section 3 décrit le contexte applicatif de notre approche qui est le montage de films. Elle présente ensuite notre modélisation à deux niveaux des comportements des utilisateurs exploitant un moteur de recherche de bandes-annonces. La section 5 présente l'outil de démonstration. Enfin la section 6 regroupe les conclusions et introduit les futures lignes directrices de notre travail.

3.53 Moreau Fabienne et Sébillot Pascale, "Thème de recherche Systèmes symboliques, Contributions des techniques du traitement automatique des langues à la recherche d'information", février 2005,
http://hal.inria.fr/view_by_stamp.php?&halsid=snpk7oeoq3ligg5c8vg988ger0&label=INRIA-RRRT&langue=fr&action_todo=view&id=inria-00070523&version=1

Rapport de recherche n°5484 de l'équipe projet « Texmex » qui traite des techniques issues du traitement automatique des langues (TAL) permettant de mettre à jour des informations morphologiques, syntaxiques et sémantiques sur les unités lexicales composant des textes. Ces divers types de connaissance ont été partiellement exploités par de nombreux travaux s'intéressant à l'interrogation de bases documentaires. Ce document tente d'évaluer l'impact des différentes sortes d'informations linguistiques pouvant être acquises par des techniques de TAL, sur les systèmes de recherche d'informations et d'en évaluer les performances.

3.54 Nain Philippe, "Thème de recherche Systèmes communicants, MAETRO : Modèles pour l'analyse des performances et le contrôle des réseaux (équipe-projet)", 2007,
www.inria.fr/recherche/equipes/maestro.fr.html

Rapport d'activité de l'équipe projet « Maestro » s'intéressant à la modélisation, à l'évaluation des performances, à l'optimisation et au contrôle des systèmes à événements discrets (SED). Les contributions scientifiques sont de nature théorique, avec le développement de nouveaux formalismes de modélisation, et de nature pragmatique avec la réalisation d'outils logiciel pour l'évaluation des performances des SED.

3.55 NewScientist, "Web « camouflage » aims to beat censors", 22 juillet 2002
<http://www.newscientist.com/article.ns?id=dn2577>

Description d'un système permettant de contourner l'accès limité à internet et consistant, notamment, à cacher le contenu censuré au travers d'images digitales non censurées.

3.56 Parlement européen et Conseil européen, "Décision n° 854/2005/CE: instituant un programme pluriannuel visant à promouvoir une utilisation plus sûre de l'internet et des nouvelles technologies en ligne", 11 mai 2005,

http://ec.europa.eu/information_society/activities/sip/programme/decision/index_en.htm

Institution d'un programme communautaire pour la période 2005-2008 visant à promouvoir une utilisation plus sûre de l'internet et des nouvelles technologies en ligne, notamment pour les enfants, et à lutter contre les contenus illicites et les contenus non désirés par l'utilisateur final. Ce programme est dénommé « Safer Internet plus ».

3.57 Parlement européen et Conseil européen, "Décision n° 1151/2003/CE: modifiant la décision n° 276/1999/CE : adoptant un plan d'action communautaire pluriannuel visant à promouvoir une utilisation plus sûre d'Internet par la lutte contre les messages à contenu illicite et préjudiciable diffusés sur les réseaux mondiaux", 16 juin 2003,

http://ec.europa.eu/information_society/activities/sip/programme/decision/index_en.htm

Tout en offrant une multitude de possibilités nouvelles, les nouvelles technologies en ligne, les nouveaux utilisateurs et les nouveaux types d'utilisation font courir de nouveaux dangers et accroissent la dangerosité actuelle d'internet. Afin d'obtenir une utilisation plus sûre d'internet un besoin de coordination se fait sentir. Dès lors, la participation de tous les acteurs concernés et en particulier d'un grand nombre de fournisseurs de contenus dans les différents secteurs devrait être encouragée. Par conséquent, le plan d'action issu de la décision n°276/1999/CE est prolongé de 2 ans.

3.58 Parlement européen et Conseil européen, "Décision n° 276/1999/CE : Adoptant un plan d'action communautaire pluriannuel visant à promouvoir une utilisation plus sûre d'Internet par la lutte contre les messages à contenu illicite et préjudiciable diffusés sur les réseaux mondiaux", 25 janvier 1999

http://ec.europa.eu/information_society/activities/sip/programme/decision/index_en.htm

Il est essentiel pour de créer un environnement internet plus sûr en luttant contre l'utilisation illicite des possibilités techniques d'internet, notamment les infractions contre les enfants, le commerce d'être humains, diffuser des idées racistes et xénophobes. Pour cette raison, la communauté doit contribuer à la réalisation de cet objectif par une action spécifique en complément des politiques menées par les Etats membres. Par conséquent, la Commission considère la promotion de l'autoréglementation de l'industrie et des systèmes de suivi du contenu, le développement des outils de filtrage et des systèmes de classement fournis par l'industrie et une sensibilisation accrue portant sur les services offerts par l'industrie jouera un rôle crucial dans la consolidation de cet environnement. Il convient d'encourager, au niveau européen, la mise à disposition des consommateurs d'outils de filtrage et d'initier la création de systèmes de classement tels que la norme « platform for internet content selection » (PICS) lancée par le consortium World Wide Web.

3.59 Perez Patrick, "Thème de recherche Systèmes Cognitifs, VISTA : Vision spatio-temporelle et apprentissage (équipe-projet)", 2007, www.inria.fr/recherche/equipes/vista.fr.html

Rapport d'activité de l'équipe projet « Vista » dont les travaux portent sur deux grandes catégories de problèmes pouvant interagir :

- l'analyse de scènes ou de phénomènes physiques dynamiques, pour des objectifs de détection, d'interprétation et de décision sur des événements temporels, ainsi que pour des besoins de mesures ;
- le couplage perception-commande dans des systèmes automatisés ou robotiques, pour des tâches de surveillance, de guidage et de manipulation, de navigation et d'exploration.

L'équipe projet s'intéresse à plusieurs types d'imageries spatio-temporelles, relevant principalement de l'imagerie optique (vidéo, infra-rouge), mais aussi acoustique (sonar, échographie). Trois secteurs d'applications ont principalement motivés les études :

- métrologie du mouvement et des déformations (imagerie météorologique, imagerie médicale, visualisation expérimentale en mécanique des fluides) ;
- vision robotique et systèmes de surveillance (sonar, transports, endoscopie) ;
- indexation de vidéos par le contenu.

3.60 Ponce Jean, "Thème de recherche Systèmes Cognitifs, Willow : Modèles de la reconnaissance visuelle d'objets et de scènes (équipe-projet)", 2007, www.inria.fr/recherche/equipes/willow.fr.html

Rapport d'activité de l'équipe projet « Willow » portant sur les problèmes de représentation dans le domaine de la reconnaissance visuelle. Le but est de développer des modèles géométriques, physiques et statistiques appropriés de toutes les composantes du processus d'interprétation des images, y compris l'illumination, les matériaux, les objets, les scènes, et les activités humaines. Cela permettra de faire face à des défis scientifiques tels que la modélisation, l'analyse et la recherche d'objets tridimensionnels, la description et la classification des activités humaines, la reconnaissance de catégories d'objets et l'interprétation de scènes complexes. Les modèles seront également utilisés dans des applications telles que l'analyse quantitative de données visuelles dans des domaines tels que l'archéologie, l'anthropologie, et la conservation du patrimoine culturel, la « post-production » de films et les effets spéciaux, l'annotation, l'interprétation, ou encore la recherche de segments vidéo dans des bases de données audiovisuelles.

3.61 "Proposition de décision du parlement européen et du conseil instituant un programme communautaire pluriannuel visant à protéger les enfants lors de l'utilisation de l'internet et d'autres technologies de communication", 27 février 2008, http://ec.europa.eu/information_society/activities/sip/programme/index_en.htm

Depuis le lancement du plan d'action pour un internet plus sûr, les technologies et leur utilisation ont fortement évolué. Dès lors, protéger les enfants des contenus et des comportements préjudiciables en ligne et restreindre la diffusion de contenus illégaux sont devenus des priorités pour l'ensemble des acteurs d'internet. Les réalisations à ce jour, sont :

- un réseau européen de lignes directes : point de contact où le public peut signaler des contenus illicites ;
- un réseau de sensibilisation européen et une journée pour un internet plus sûr coordonnée par le réseau ;
- des informations sur l'efficacité des logiciels de filtrage par des essais indépendants ;

- appui aux initiatives d'autorégulation du domaine du classement des contenus et de la téléphonie mobile.

Ce nouveau programme est conçu pour tenir compte des évolutions prochaines de l'environnement en ligne et des menaces qui en résulteront. A cet effet, il sera nécessaire de concevoir des actions appropriées pour protéger les enfants dans l'environnement en ligne au cours de la période 2009-2013.

3.62 Schmid Cordelia, "Thème de recherche Systèmes Cognitifs, LEAR : Apprentissage et reconnaissance en vision par ordinateur (équipe-projet)", 2007,
www.inria.fr/recherche/equipes/lear.fr.html

Rapport d'activité de l'équipe projet « Lear » qui aborde le problème de la reconnaissance d'objets et de l'interprétation de scènes pour des images statiques et des séquences d'images vidéo. C'est le problème de la vision par ordinateur : il est aujourd'hui impossible de déterminer de façon automatique le contenu d'une image ou d'une séquence vidéo. La piste suivie par l'équipe-projet repose sur l'ajout de l'apprentissage aux techniques de vision par ordinateur que sont la description d'images et la géométrie. Une solution, même partielle, au problème de la reconnaissance d'objets et de l'interprétation de scènes a de nombreuses applications. Dans ce cadre, l'équipe-projet s'est intéressée en particuliers aux applications en recherche d'images et en indexation vidéo.

3.33 Shepherd Michael et Watters Carolyn, "Technologies de filtrage de contenu et fournisseurs de services Internet", 22 mars 2000,
<http://www.ic.gc.ca/epic/site/smt-gst.nsf/fr/sf05251f.html>

Ce rapport, commissionné par Industrie Canada, traite des mécanismes que les fournisseurs de services internet (FSI) ont le choix d'offrir et que les utilisateurs peuvent eux-mêmes choisir d'utiliser pour filtrer le contenu mis à leur disposition sur internet, en autorisant par ailleurs leur accès. Il s'agit d'une simple description des mécanismes de filtrage sans qu'il ne soit prodigué de conseils ou de recommandations d'ordre juridique.

3.64 Tabbone Salvatore-Antoine, "Thème de recherche Systèmes Cognitifs, QGAR : Recherche d'information graphique par l'analyse et la reconnaissance (équipe-projet)", 2007,
www.inria.fr/recherche/equipes/qgar.fr.html

Rapport d'activité de l'équipe projet « Qgar » dont le thème scientifique majeur est la reconnaissance de graphiques. Les objectifs sont l'indexation et la recherche d'informations, dans le contexte de la documentation technique. Le problème de base est celui de la conversion d'une information faiblement structurée, telle que l'image d'un document papier ou un fichier PDF, en une information enrichie des structures qui la rendent exploitable directement au sein d'un système d'informations.

3.65 The 2020 Science Group, "About the Report", juillet 2005
http://research.microsoft.com/towards2020science/downloads/T2020S_ReportA4.pdf

Rapport contenant les postulats de départ et les conclusions d'un groupe composé de scientifiques internationaux et se réunissant pour mesurer l'impact de l'informatique et des sciences informatiques sur la science en 2020.

3.66 Urruty Thierry, "KpyrRec : a Recursive Multidimensional Indexing Structure", janvier 2005, International Journal of parallel, Emergent and Dtributed Systems

L'émergence des technologies numériques dans le secteur du multimédia a mis en valeur l'importance des problèmes d'indexation multidimensionnelle et de recherché par le contenu dans la recherche informatique. Ce travail fait partie d'un projet visant à mettre en place un outil de recherche vidéo, destiné à des utilisateurs professionnels de grandes bases de films d'entreprise. Nous présentons dans cet article la synthèse d'une étude analytique des performances des structures d'indexation multidimensionnelles concluant sur le fort impact du facteur du nombre de données analysées par la requête sur le temps de réponse. Nous proposons en conséquence une nouvelle structure d'indexation KpyrREc et montrons ses performances comparées à d'autres techniques récentes de la littérature.

3.67 Urruty Thierry, Belkouch Fatima et Djeraba Chabane, " Kpyr, une structure efficace d'indexation de documents vidéo", 8 juillet 2005,

<http://134.214.81.35/articles/a542c1FRgOUaXZ9LA.pdf>

Motivés par les récents besoins de structures d'indexation efficaces et adaptés à de réelles applications sur des bases de données vidéo, nous présentons, dans cet article Kpyr une nouvelle structure d'indexation multidimensionnelle. L'idée générale de Kpyr est d'utiliser un algorithme de classification pour diviser l'espace en sous espaces sur lesquels nous appliquerons la technique Pyramidale. Nous réduisons ainsi l'espace de recherche concerné par une requête et améliorons par conséquence la rapidité des recherches. Nous avons montré que notre approche procure des résultats expérimentaux intéressants et performants pour les requêtes par « fenêtrage » (Windows Query) et par les plus proches voisins (KNN Query).

3.68 Valette Mathieu, "Détection et interprétation automatique de contenus illicites et préjudiciables sur internet", projet PRINCIP, <http://faculty.arts.ubc.ca>

PRINCIP est un système de détection automatique des pages web ayant un contenu illégal ou préjudiciable, développé par des laboratoires de recherche européens. Cet article critique le système de filtrage par mots-clés en ce qu'il ne tient pas compte de l'intertextualité

3.69 Vandeborre Jean-Philippe, Samir Chafik et Daoudi Mohamed, "Automatic 3D face recognition topological techniques", juillet 2005,

<http://www.enic.fr/people/vandeborre/papers/samirICME2005.pdf>

Dans le présent document, les informations en trois dimensions de la forme topologique du visage humain pour l'identification sont utilisées. Il est proposé une nouvelle méthode pour représenter en 3D les visages comme un graphe topologique. L'enregistrement des surfaces se fait en trouvant automatiquement la topologie des composantes connexes, et ensuite on bâtit sa topologie graphique en représentant les modifications topologiques importantes sur le visage. Le calcul de la similitude entre les visages 3D traité à l'aide d'une stratégie grossière, tout en préservant la cohérence de la représentation graphique des structures, qui se traduit par l'établissement d'une correspondance entre les parties du visage. Les expériences faites avec les 144 visages en données 3D montrent l'efficacité de notre approche.

3.70 Walfish Mickael, MIT Laboratory for Computer Science, "Using DHTs to Untangle the web DNS", <http://nms.lcs.mit.edu/papers/sfr-isw03.pdf>

Article relatif au référencement sur internet portant proposition de l'adoption d'un nouveau système de référencement.

3.71 Zitnick Lawrence, Takeo Kanade, The Robotics Institute de la Carnegie Mellon University de Pittsburgh, "Content-Free Image Retrieval", mai 2003, <http://research.microsoft.com/users/larryz/ZitnickCFIR03.pdf>

Présentation d'une méthode de localisation des images, la « Content-Free Image Retrieval ». Il s'agirait d'un système capable de repérer et localiser les images pertinentes afin, le cas échéant, d'en analyser le contenu. Cette méthode repose sur un algorithme mettant en œuvre les principes du filtrage de type « collaboratif », destiné à prédire les préférences des individus en référence aux préférences passées d'autres individus.

4. ARTICLES ET DOCUMENTS SUR LE FILTRAGE ANTI-SPAM ET ANTI-HAMECONNAGE

4.1 Dumout Estelle, ZDNet France, « Anti-spam : Microsoft relance Sender ID », <http://www.zdnet.fr/actualites/telecoms/0,39040748,39179631,00.htm>

Nouvelle version du filtre anti-spam « Sender ID », présentée par Microsoft.

4.2 Grandmontagne Yves, "Google victime de son filtre Bayésien anti-spam", 3 décembre 2003, <http://www.silicon.fr/fr/silicon/special-report/2003/12/03/google-victime-filtre-bayésien-anti-spam/>

Les raisons qui ont conduit Google à installer un filtre bayésien sur son moteur de recherche et définition technique des filtres bayésiens dont l'efficacité peut être remise en cause.

4.3 Lagadec Philippe, "Filtrage de messagerie et analyse de contenu", http://actes.sstic.org/SSTIC04/Filtrage_messagerie/SSTIC04-article-Lagadec-Filtrage_messagerie.pdf

Présentation des objectifs d'un filtrage de messagerie électronique, des différentes techniques utilisées et de leurs limites.

4.4 Livre blanc GFI, "Pourquoi le filtrage Bayésien est la technologie anti-spam la plus efficace", <http://www.gfsfrance.com/fr/whitepapers/why-bayesian-filtering.pdf>

Critique des systèmes dits de « filtrage statique des spams » et description du fonctionnement d'un filtre bayésien.

4.5 "Déjouer les arnaques en ligne à l'aide du filtre anti-hameçonnage de Microsoft", 28 octobre 2006, http://www.microsoft.com/france/athome/security/online/phishing_filter.msp

Les trois techniques de protection mises en œuvre par un filtre antihameçonnage.

4.6 "Filtre anti-spam", http://fr.docs.yahoo.com/mail/spamguard_domainkeys.html

Fiche technique en cinq étapes sur la technologie « Domainkeys » permettant le filtrage antisпам des adresses de messagerie électronique Yahoo.

5. ARTICLES ET DOCUMENTS SUR LE FILTRAGE AU TITRE DU CONTROLE PARENTAL

5.1 Béranger Anne-laure, "*Protection des enfants : AOL, MSN, Yahoo passent au filtre*", 12 juillet 2002, <http://www.journaldunet.com/0207/020712label.shtml>

Trois acteurs américains de l'internet ont eu recours aux critères et à la technologie de filtrage de l'Internet Content Rating Association (ICRA), qui permet aux parents, en entrant leurs propres références, et ce au moyen d'un petit logiciel, de filtrer les contenus disponibles sur internet.

5.2 Boutier Frédéric, "*Comment fonctionne le filtrage ?*", Micro Hebdo, 5 décembre 2005, <http://www.01net.com/article/297389.html>

Les différentes méthodes pour contrôler la navigation des enfants sur internet : les logiciels qui s'appuient sur des listes noires ou sur une analyse lexicographique, les filtres en général, et la recherche de marqueurs.

5.3 C. Ange-Gabriel, "*Un nouvel outil de filtrage des contenus pour adulte vient de voir le jour*", 21 mars 2006, <http://www.generation-nt.com/filtre-porno-ishield-guardware-actualite-12172.html>

L'éditeur américain, Guardware, propose un nouveau logiciel, baptisé iShield, proposant trois niveaux de protection dans la consultation d'internet.

5.4 "*Contrôle parental*", http://fr.wikipedia.org/wiki/Contrôle_parental

Présentation très générale sur les logiciels de contrôle ou de filtre parental sur internet.

5.5 Google, "*Filtrage Safesearch*", <http://google.fr/support/bin/answer.py?answer=35892&print=1>

Safesearch est une fonctionnalité du moteur de recherche Google permettant de filtrer les sites présentant un caractère sexuel et de les exclure des résultats de la recherche.

5.6 La rédaction du journal du net, "*Protection des enfants : un filtre, un guide pratique et des recommandations*", 12 février 2004, <http://www.journaldunet.com/0402/040212filtrage.shtml>

Installation d'un logiciel de filtrage des sites internet dans les établissements scolaires et diffusion, auprès des parents et des enfants, de guides pratiques d'utilisation du Web.

5.7 "*Projet Filtra : protocole de test et conditions de participation*", février 2006, <http://www.filtra.info/doc/ProtocoleConditionsParticipation.pdf>

Filtra permet d'apprécier l'efficacité des outils de filtrage des contenus illégaux ou immoraux, destinés aux familles. Ce protocole explique la méthode de travail adoptée par Filtra.

5.8 Rego Karine, "*Fonctionnement et limites des logiciels de contrôle parental*", 3 octobre 2006, http://www.linternaute.com/hightech/internet/contrôle_parental/fonctionnement.shtml

Etat des lieux des différents logiciels utilisés au titre du contrôle parental.

5.9 "*Test des logiciels de filtrage*", <http://delegation.internet.gouv.fr/mineurs/enquete.htm>

Résultats d'un test destiné à mesurer l'efficacité du filtrage exercé au titre du contrôle parental et tableaux comparatifs des différentes possibilités de filtrage

ANNEXE 5

TABLEAU DES BREVETS DEPOSES AYANT TRAIT OU SUSCEPTIBLES D'AVOIR UN LIEN AVEC LE FILTRAGE DE CONTENUS

Recherches effectuées sur les bases de données de brevets françaises (FR), européennes (EP), internationales (WO)¹ et américaines (US)².

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
ADVESTIGO	27 11 2003	EP1704695 SYSTEME D'INTERCEPTION DE DOCUMENTS MULTIMEDIAS	WO2003FR03502 WO2005064885 CA2547344 AU2003294095 US2007110089	Taggage de contenus et affectation de métadonnées
ADVESTIGO	27 11 2003	FR2863080 PROCEDE D'INDEXATION ET D'IDENTIFICATION DE DOCUMENTS MULTIMEDIAS	EP1697862 WO2005055086	Analyse de bases de données et extractions
ADVESTIGO SA	15 06 2005	FR2887385 PROCEDE ET SYSTEME DE REPERAGE ET DE FILTRAGE D'INFORMATIONS MULTIMEDIA SUR UN RESEAU	WO2006134310	Filtrage de flux
AUDIBLE MAGIC CORP	15 02 2001	WO0162004 PROCEDE ET APPAREIL D'IDENTIFICATION DE CONTENU MEDIA PRESENTE SUR UN DISPOSITIF DE DIFFUSION DE MEDIAS	US20000511632	Taggage de contenus et affectation de métadonnées

¹ <http://fr.espacenet.com/>

² <http://www.uspto.gov/patft/index.html>

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
BASCOM GLOBAL INTERNET SERVICE, CIRASOLE PETER, DEROSA ROBERT, FOX ROBERT	18 03 1998	WO9841913 PROCEDE ET SYSTEME DE FILTRAGE DU CONTENU D'INFORMATIONS RECUPEREES SUR UN RESEAU INFORMATIQUE INTERNET	US19970820955	Analyse de bases de données et extractions
BEEP SCIENCE AS, BREIVIK OEYVIND	08 11 2002	WO03040898 DISPOSITIF ET PROCEDE POUR LE CONTROLE DE LA POLICE DE CONTENU AU MOYEN D'UN ENVIRONNEMENT DE CONFIANCE DANS UN SYSTEME DE MESSAGERIE MULTIMEDIA	NO20010005471 NO316737	Filtrage de courriers électroniques et de messages
BIAP SYSTEMS INC, SLOTHOUBER LOUIS	21 08 2006	WO2007024736 SYSTEME ET METHODE POUR RECOMMANDER DES ARTICLES INTERESSANTS A UN UTILISATEUR	US20050709420P	Prototypage de données
BLACK PETER M, WATERS ANTHONY BRYAN	14 09 2001	US2002035573 METATAG-BASED DATAMINING	US20000630227 US20000703006 US20000738471	Analyse de bases de données et extractions
BLACKSPIDER TECHNOLOGIES, KAY JAMES	22 06 2006	WO2006136605 PROCEDE ET SYSTEME PERMETTANT DE FILTRER DES MESSAGES ELECTRONIQUES	EP1894369 GB20050012744	Filtrage de courriers électroniques et de messages
BUSINESS OBJECTS S A, WEBSTER RICHARD DAVID, CAMERON RICHARD BRUCE	22 08 2007	WO2008027765 APPARATUS AND METHOD FOR PROCESSING QUERIES AGAINST COMBINATIONS OF DATA SOURCES	US20060515441	Analyse de bases de données et extractions

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
BUSINESS OBJECTS SA	17 12 2004	EP1695235 APPARATUS AND METHOD FOR USING DATA FILTERS TO DELIVER PERSONALIZED DATA FROM A SHARED DOCUMENT	WO2004US42553 US20030531509P WO2005062807 CA2550113	Prototypage de données
BUSINESS OBJECTS SA	02 12 2005	EP1839200 SUPPORT LISIBLE PAR ORDINATEUR, PROCEDE ET APPAREIL DE CONSERVATION DES CONDITIONS DE FILTRAGE POUR INTERROGER DES SOURCES DE DONNEES MULTILINGUES SELON DIVERSES LOCALISATIONS LORS DE LA REGENERATION D'UN COMPTE RENDU	WO2006073633 WO2005US43563 US20040640469P US20050271702	Filtrage lexicographique
BUSINESS OBJECTS SA	30 03 2006	US2007074176 APPARATUS AND METHOD FOR PARALLEL PROCESSING OF DATA PROFILING INFORMATION	US20050720277P	Prototypage de données
BUSINESS OBJECTS SA	22 09 2006	US2007074155 APPARATUS AND METHOD FOR DATA PROFILE BASED CONSTRUCTION OF AN EXTRACTION, TRANSFORM, LOAD (ETL) TASK	US20050719958P	Analyse de bases de données et extractions
BUSINESS OBJECTS SA, KRINSKY ANTHONY SETH, HASSENFORDER MARCEL, CHEVRIER MARC, CRAS JEAN- YVES	08 02 2007	WO2007098320 APPAREIL ET PROCÉDÉ D'INTERROGATION FÉDÉRÉE DE DONNÉES NON STRUCTURÉES	US20060364564	Analyse de bases de données et extractions

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
BUSINESS OBJECTS SA, NAIBO ALEXIS- JEAN LAURENT, BOLF DIDIER MARC JEAN, MEINIEL PHILIPPE, REYNOLDS RICHARD THOMAS JR	25 05 2007	WO2007140290 APPARATUS AND METHOD FOR QUERYING DATABASES VIA A WEB SERVICE	US2007276815 US20060808860P	Analyse de bases de données et extractions
CANON RES CT FRANCE, INRIA INST NAT DE RECH	22 06 2005	EP1612727 DETECTION ET PREUVE AVEC UN PROTOCOLE "ZERO KNOWLEDGE" DE TATOUAGES NUMERIQUES DANS DES ENTITES MULTIMEDIA	FR20040007083	Analyse de vidéo
CHECKPOINT SYSTEMS INC, ROBERTS PAUL A, SIMON PAUL	20 07 2005	WO2006012380 SYSTEME ET PROCEDE D'AUTO- VERIFICATION PERMETTANT DE PROTEGER DES CONTENUS MULTIMEDIA NUMERIQUES	US20040590548P US20050154252 EP1771830 US2006016885 KR20070060083 CA2575051 AU2005267167	Taggage de contenus et affectation de métadonnées
CISCO TECH INC	31 08 2000	US7120931 SYSTEM AND METHOD FOR GENERATING FILTERS BASED ON ANALYZED FLOW DATA		Filtrage de flux
CISCO TECH IND	30 05 2000	US6829336 SYSTEM AND METHOD FOR ACTIVE FILTERING IN A TELECOMMUNICATIONS NETWORK		Filtrage de flux

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATÉGORIE DE FILTRAGE
CLEARPLAY INC	27 09 2001	WO0237853 FILTRAGE DE CONTENU MULTIMEDIA INDESIRABLE	EP1344399 WO2001US30481 US20000694873	Prototypage de données
CLICKSAFE COM LLC	21 02 2001	WO0163835 SYSTEME ET PROCEDE PERMETTANT D'IDENTIFIER ET D'EMPECHER L'ACCES A UN CONTENU D'INTERNET PORNOGRAPHIQUE ET ANALOGUE		Prototypage de données
COHESIA CORP	05 07 2000	WO0104812 REPRESENTATION, GESTION, FILTRAGE ET SYNTHESE DE CONTENU TECHNIQUE	US19990349753 EP1196877 US6658428 US6405211	Prototypage de données
CORETRUST INC, WOO JE-HAK, LEE HWAN-CHUL, CHO SANG-YOUNG, JEONG SEONG-HO, HA YOUNG-SOO, SHIN SEOG-KYOON, KIM SEONG-IL	10 01 2003	WO03058485 PROCEDE ET SYSTEME POUR LA PROTECTION DES INFORMATIONS D'UN CONTENU NUMERIQUE	KR20020001916 KR20020073773 EP1470497 US2005086501 AU2003202815	Taggage de contenus et affectation de métadonnées
CORVIGO	03 10 2003	US2005076084 DYNAMIC MESSAGE FILTERING	WO2005036341 EP1714201 US7257564 US2007239639 CN101069175 CA2540571 AU2004281052	Filtrage de courriers électroniques et de messages

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
FRAISSE THOMAS, DUTHEIL PIERRE	21 10 2003	FR2861195 PROCEDE ET DISPOSITIF DE FILTRAGE DE CONTENUS EN LIGNE	EP1676218 WO2005038670	Filtrage de flux
FRANCE TELECOM	08 08 2003	FR2852416 SYSTEME DE FILTRAGE PROGRESSIF DE CONTENUS		Analyse de bases de données et extractions
FUJITSU LTD	16 04 2002	EP1311100 METHODE, APPAREIL ET LOGICIEL DE FILTRAGE DE CONTENU	JP20010346835 US7203749 US2003093518 KR2003003995 JP2003150482	Filtrage web et/ou filtrage réseau
GESTWEB S P A	13 06 2001	EP1197878 PROCEDE ET SYSTEME DE FILTRAGE DU CONTENU D'INFORMATIONS RECUPERE DANS UN RESEAU DE COMMUNICATION DE DONNEES	IT2000MI02189	Analyse de bases de données et extractions
GOOGLE INC	24 07 2003	EP1547118 PROCEDE ET SYSTEME DE FOURNITURE DE PUBLICITES FILTRES ET/OU MASQUEES SUR L'INTERNET	WO2003US23010 US20020398101P US20030388166	Filtrage de flux
GOOGLE INC	27 02 2004	EP1604302 IDENTIFICATION D'INFORMATIONS CONNEXES EN FONCTION D'UN CONTENU ET/OU PRESENTATION D'INFORMATIONS CONNEXES EN ASSOCIATION AVEC DES ANNONCES PUBLICITAIRES LIEES AU CONTENU	WO2004US05979 US20030450775P US20030748870	Prototypage de données

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
GOOGLE INC	23 07 2004	EP1649395 PROCEDES ET SYSTEMES PERMETTANT DE COMPRENDRE UN SENS D'UN ITEM DE CONNAISSANCE AU MOYEN D'INFORMATIONS ASSOCIEES L'ITEM DE CONNAISSANCE	WO2004US23826 US20030491422P US20030690328	Prototypage de données
GOOGLE INC	23 07 2004	EP1649396 PROCEDES ET SYSTEMES DE DETERMINATION DU SENS D'UN DOCUMENT AFIN DE FAIRE CORRESPONDRE LE DOCUMENT AU CONTENU	WO2004US23827 US20030491422P US20030689903	Prototypage de données
GOOGLE INC	13 09 2004	EP1676211 SYSTEMES ET PROCEDES POUR FAIRE DES RECHERCHES AU MOYEN DE DEMANDES ECRITES DANS UN ENSEMBLE DE CARACTERES ET/OU LANGAGE DIFFERENT A PARTIR DE PAGES CIBLES	WO2004US29772 US20030676724	Filtrage lexicographique
GOOGLE INC	15 09 2004	EP1775665 MARQUAGE DE DOCUMENT BASE SUR DES CRITERES BASES SUR DES LIENS	EP20040784004 US20030507617P US20030748664	Taggage de contenus et affectation de métadonnées
GOOGLE INC	17 11 2004	EP1695232 PROCEDES ET SYSTEMES DESTINES A L'EXTRACTION D'INFORMATIONS	WO2004US38559 US20030731916	Prototypage de données
GOOGLE INC	29 03 2005	EP1735725 REECRITURE DE REQUETE AVEC DETECTION D'ENTITE	WO2005US10701 US20040813359	Prototypage de données
GOOGLE INC	10 05 2005	EP1787258 SYSTEME ET PROCEDE POUR EVALUER DES DOCUMENTS COMPRENANT UNE IMAGE	WO2005US15963 US20040841834	Analyse et ou traitement d'images
GOOGLE INC	26 07 2005	EP1622052	US20040900075	Filtrage lexicographique

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
		GENERATION DE DESCRIPTIONS DE DOCUMENTS A BASE DE PHRASES		
GOOGLE INC	26 07 2005	EP1622053 IDENTIFICATION DE LOCUTIONS DANS UN SYSTEME DE RECHERCHE D'INFORMATIONS	US20040900021	Filtrage lexicographique
GOOGLE INC	26 07 2005	EP1622054 RECHERCHE A BASE DE PHRASES DANS UN SYSTEME DE RECHERCHE D'INFORMATIONS	US20040900041	Filtrage lexicographique
GOOGLE INC	26 07 2005	EP1622055 INDEXAGE BASE SUR DES LOCUTIONS DANS UN SYSTEME POUR RECHERCHE D'INFORMATIONS	US20040900055	Filtrage lexicographique
GOOGLE INC	21 12 2005	EP1839203 ASSOCIATION DE CARACTERISTIQUES AVEC DES ENTITES, NOTAMMENT DES CATEGORIES OU DES DOCUMENTS DE PAGE WEB, ET/OU PONDERATION DE TELLES CARACTERISTIQUES	WO2005US46194 US20040026497	Prototypage de données
GOOGLE INC	25 01 2006	EP1844391 SYSTEME D'EXTRACTION D'INFORMATIONS BASE SUR DES INDEX MULTIPLES	WO2006US02709 US20050043695	Filtrage lexicographique
GOOGLE INC	29 11 2006	WO2007064656 FORMATAGE D'UN SITE DE RESEAU UTILISATEUR SUR LA BASE DES PREFERENCES UTILISATEUR ET DES DONNEES DE RENDEMENT DES FORMATS	US20050288431 US2007300152	Filtrage web et/ou filtrage réseau
GOOGLE INC	14 12 2006	WO2007070622 DETECTION ET REJET DE DOCUMENTS AGAÇANTS	US20050302495	Prototypage de données
GOOGLE INC, ACHARYA ANURAG,	15 09 2004	WO2005033978 RECUPERATION D'INFORMATION BASEE	US20030507617P US20030748664	Filtrage lexicographique

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
CUTTS MATT, DEAN JEFFREY, HAAHR PAUL, HENZINGER MONIKA, HOELZLE URS, LAWRENCE STEVE, PFLEGER KARL, SERCINOGLU OLCAN, TONG SIMON		SUR DES DONNEES HISTORIQUES	EP1668551 US2007100817 US2007094255 US2007094254 US2007088693 US2007088692 US2005071741 JP2007128547 CA2540573 AU2004277678	
GOOGLE INC, AGARWAL SUMIT, AXE BRIAN, GEHRKING DAVID, LAW CHING, MAXWELL ANDREW, RAJARAM GOKUL; WISEMAN LEORA	24 06 2005	WO2006115508 SUGGESTION D'INFORMATIONS DE CIBLAGE DESTINEES A DES PUBLICITES, TELLES QUE DES SITES WEB ET/OU DES CATEGORIES DE SITES WEB, PAR EXEMPLE	US20050112732 EP1897044 US2006242013 CA2605536 AU2005331031	Prototypage de données
GOOGLE INC, ANGELO MICHAEL, BRAGINSKY DAVID, GINSBERG JEREMY, TONG SIMON	27 06 2006	WO2007005431 DETERMINATION D'UN REFERENTIEL DE DONNEES DESIRE	US20050169285 US2007005568 AU2006266103	Prototypage de données
GOOGLE INC, BALUJA SHUMEET, COVELL MICHELE, FINK MICHAEL	27 11 2006	WO2007064640 DETECTION DE CONTENUS REPETITIFS DANS DES MEDIA DIFFUSES	US20050740760P US20060823881P WO2007064641	Filtrage de flux
GOOGLE INC, BRIN SERGEY, GOMES BENEDICT, TONG	27 07 2004	WO2005013153 MISE A DISPOSITION D'UNE INTERFACE UTILISATEUR AVEC ELARGISSEMENT DE LA	US20030629479 EP1654681 US2005027691	Analyse de bases de données et extractions

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
SIMON		REQUETE DE RECHERCHE	CN1849603 CA2533605 AU2004262352	
GOOGLE INC, BUNESCU RAZVAN CONSTANTIN, PASCA ALEXANDRU MARIUS	02 04 2007	WO2007115266 DESAMBIGUISATION D'ENTITES NOMMEES	US20060744091P US20060427678 US2007233656	Prototypage de données
GOOGLE INC, CAROBUS ALEXANDER PAUL, ROETTER ALEX, DAVENPORT BEN	05 10 2004	WO2005038575 SERVICE D'ANNONCES A CONTENU CIBLE DANS UN MESSAGE ELECTRONIQUE, TELS QUE DES BULLETINS DE MESSAGES ELECTRONIQUES	US20030509164P US20030699607 EP1671215 US7203684 US2005076051 KR20060086374 CA2541932 BRPI0415122 AU2004282878	Filtrage de courriers électroniques et de messages
GOOGLE INC, DJABAROV GUEORGUI	07 12 2005	WO2006065583 FOURNITURE D'INFORMATION UTILE ASSOCIEE A UN ARTICLE DANS UN DOCUMENT	US20040010316 EP1831836 US2006129910 CA2591686 AU2005316808	Taggage de contenus et affectation de métadonnées

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
GOOGLE INC, EGNOR DANIEL	30 12 2005	WO2006074054 INDEXAGE DE DOCUMENTS CONFORMEMENT A LA PERTINENCE GEOGRAPHIQUE	US20040024790 EP1839212 US2006149774 KR20070092758 CN101128823 CA2593420	Taggage de contenus et affectation de métadonnées
GOOGLE INC, EGNOR DANIEL, CHAUDHRY GEETA	30 12 2005	WO2006074056 IDENTIFICATION DE DOCUMENTS FAISANT AUTORITE	US20040024967 EP1859367 US2006149800 KR20070094941 EP1859367 CN101128822	Prototypage de données
GOOGLE INC, EGNOR DANIEL, HAAHR PAUL, LACKER KEVIN, LAMPING JOHN, SINGHAL AMITABH K, YANG KE	19 03 2007	WO2007126628 DISSEMINATION D'INFORMATIONS UTILES SUR DES PAGES INTERNET CONNEXES, TELLES LES PAGES D'UN SITE INTERNET	US20060396301 US2007233808	Filtrage web et/ou filtrage réseau
GOOGLE INC, FRANZ ALEXANDER M, HENZINGER MONIKA	29 12 2004	WO2005066847 SYSTEMES ET PROCEDES POUR AMELIORER LA QUALITE D'UNE RECHERCHE	US20030749730 EP1704495 US2005149499 BRPI0418230	Analyse de bases de données et extractions
GOOGLE INC, GEHRKING DAVID, LAW CHING, MAXWELL ANDREW	24 04 2006	WO2006116273 CATEGORISATION D'OBJETS, DE TYPE DOCUMENTS ET/OU GROUPES, PAR RAPPORT A UNE TAXINOMIE ET A DES STRUCTURES DE DONNEES DERIVEES DE LADITE CATEGORISATION	US20050112716	Prototypage de données
GOOGLE INC, GUHA	12 07 2006	WO2007021417	US20050202423	Analyse de bases de

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
RAMANATHAN V		MOTEUR DE RECHERCHE PROGRAMMABLE	US2007038616	données et extractions
GOOGLE INC, HAAHR PAUL, BAKER STEVEN	01 09 2004	WO2005026989 OBTENTION DE RAFFINEMENTS DE DEMANDE DE RECHERCHE	US20030500539P US20030668721 EP1665090 US2005055341	Analyse de bases de données et extractions
GOOGLE INC, HAGAN ROSE ANNE, RANA KULPREET SINGH	30 03 2006	WO2006105492 CIBLAGE AUTOMATIQUE DU CONTENU BASE SUR LES DROITS DE PROPRIETE INTELLECTUELLE	US20050094793 US2006230457	Prototypage de données
GOOGLE INC, KHALIQ SIRAJ, BROUGHER WILLIAM C	29 08 2005	WO2006039025 FOURNITURE D'INFORMATIONS ASSOCIEES A UN DOCUMENT	US20040953112 EP1797511 US2006074868 EP1797511 CN101061478 CA2583042	Taggage de contenus et affectation de métadonnées
GOOGLE INC, KONINGSTEIN ROS, SPITKOVSKY VALENTIN, HARIK GEORGES R, SHAZEER NOAM	30 12 2004	WO2005065401 SUGGESTION ET/OU FOURNITURE DE CRITERES DE CIBLAGE POUR ANNONCES PUBLICITAIRES	US20030750451 EP1709551 US2005228797 KR20060130155 CN1922604 CA2552236 BRPI0418256 AU2004311451	Prototypage de données
GOOGLE INC, LAWRENCE STEPHEN, KAHN OMAR, BHATLA NIKHIL	17 11 2004	WO2005066841 PROCEDES ET SYSTEMES POUR AMELIORER UN CLASSEMENT DE RECHERCHE AU MOYEN D'INFORMATIONS CONCERNANT L'ARTICLE	US20030749434 EP1700235 US2005149498	Prototypage de données

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
GOOGLE INC, PISCITELLO JOHN, WANG XUEFU, TONG SIMON, HAGAN BREEN	13 07 2005	WO2006019852 ENTRÉE URL DESIGNEE PAR DES TERMES	US20040953497 US20040587548P EP1769404	Filtrage web et/ou filtrage réseau
GOOGLE INC, REDDY BINDU, BRUNSMAN JONATHAN, MOSBERGER NING, BHAYA GAURAV RAVINDRA, SIRAJUDDIN SARAH, KALE DAVID, KOZENSKI JENNIFER L, SUNDARARAJAN ARVIND, AGARWAL PUNEET	13 12 2005	WO2007046830 RECHERCHE DANS DES DONNEES STRUCTUREES	US20050257282	Analyse de bases de données et extractions
GOOGLE INC, REDDY BINDU, SPIGHT MARSHALL, MOSBERGER NING	13 12 2005	WO2007046829 AJOUT D'ATTRIBUTS ET D'ETIQUETTES A DES DONNEES STRUCTUREES	US20050256883 US2007100862	Taggage de contenus et affectation de métadonnées
GOOGLE INC, RUHL JAN M, DATAR MAYUR D; LEE JESSICA Y	30 11 2005	WO2006065546 PROCEDE, SYSTEME ET INTERFACE UTILISATEUR GRAPHIQUE PERMETTANT DE FOURNIR DES AVIS CONCERNANT UN PRODUIT	US20040012500 US20040012846 EP1834249 CA2591441	Taggage de contenus et affectation de métadonnées

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
GOOGLE INC, SCHIRRIPA STEVEN R, HARADA MASANORI	15 06 2006	WO2006138473 CLASSIFICATION D'UN ELECTRONIQUE CONTENU	US20050153123 EP1899798 US2006288015	Prototypage de données
GOOGLE INC, SHELLEN JASON H, PARPARITA MIHAI	06 10 2006	WO2007044722 PAGE DE SUGGESTIONS PERSONNALISEES DE DONNEES DE CONTENU	US20050246596 US2007083520	Prototypage de données
GOOGLE INC, STEELBERG RYAN, STEELBERG CHAD	16 03 2006	WO2006127097 SYSTEME ET PROCEDE D'ETIQUETAGE DE DIFFUSIONS ET PUBLICITE CIBLEE	US20050135860 US20050135859 EP1891593 AU2006249685	Prototypage de données
GOOGLE INC, STOPPELMAN MICHAEL	29 06 2006	WO2007002828 RECOMMANDATIONS DE PRODUIT FONDEES SUR LE FILTRAGE COLLABORATIF DE DONNEES UTILISATEUR	US20050168561 US2007005437	Prototypage de données
GOOGLE INC, VINCENT LUC, ULGES ADRIAN	29 06 2007	WO2008003095 RECONNAISSANCE DE TEXTE DANS DES IMAGES	US20060479115 US20060479155 US20060479957	Analyse et ou traitement d'images
GOOGLE INC, WEISSMAN ADAM J	30 12 2003	WO2005069199 PROCEDES ET SYSTEMES POUR LA SEGMENTATION DE TEXTE	AU2003300437	Filtrage lexicographique
GOOGLE INC, YAGNIK JAY N	09 04 2007	WO2007120716 PROCEDE ET APPAREIL PERMETTANT DE RESUMER AUTOMATIQUEMENT UNE VIDEO	US20060791869P US20060454386 US2007245242	Analyse de vidéo
GOOGLE INC, YAN WEIPENG, TOKUSEI KENTARO	18 10 2006	WO2008008087 IDENTIFICATION DE REQUETES FRAUDULEUSES D'INFORMATIONS	US20050253004 US2007084915	Analyse de bases de données et extractions

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
GOOGLE INC, YU HUA, MORENO PEDRO	28 09 2006	WO2007041370 UTILISATION DE LA RECONNAISSANCE DE LA PAROLE POUR L'OBTENTION D'ANNONCES PUBLICITAIRES PERTINENTES PAR RAPPORT A UN CONTENU AUDIO ET / OU DE CONTENU AUDIO PERTINENT PAR RAPPORT A DES ANNONCES PUBLICITAIRES	US20050241834 US2007078708	Analyse du son et/ou de la parole
GOOGLE INC, ZHOU JIE, KHOPKAR CHIRAG, WALKOVER ASHER, KAPPLER PETER, LU CHARITY YUEH-CHWEN	17 11 2006	WO2007061877 METHODE DE DETECTION DE FRAUDE DANS DES PUBLICITES INTERNET	US20050282971 US2007129999	Filtrage web et/ou filtrage réseau
GRACENOTE INC	21 07 2000	US7228280 FINDING DATABASE MATCH FOR FILE BASED ON FILE CHARACTERISTICS	US19990354164 US19970838082	Analyse de bases de données et extractions
GRACENOTE INC	22 07 2002	EP1410380 IDENTIFICATION AUTOMATIQUE D'ENREGISTREMENTS SONORES	WO03009277 WO2002US23101 US20010306911P	Analyse du son et/ou de la parole
GRACENOTE INC	31 07 2002	EP1421521 IDENTIFICATION D'ENREGISTREMENTS COMPORTANT DES ETAPES MULTIPLES	WO03012695 WO2002US24054 US20010308594P	Taggage de contenus et affectation de métadonnées
GRACENOTE INC	21 01 2003	NO20030319 METHOD AND SYSTEM FOR FINDING MATCH IN DATABASE RELATED TO WAVEFORMS	WO0208943 EP1303817 US20000621619 WO2001US22891	Analyse du son et/ou de la parole

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
GRACENOTE INC	21 08 2006	US2007106405 METHOD AND SYSTEM TO PROVIDE REFERENCE DATA FOR IDENTIFICATION OF DIGITAL CONTENT	US20050709543P	Analyse de bases de données et extractions
GRUPE ROBERT	15 06 2001	US2002194487 SCANNING COMPUTER FILES FOR SPECIFIED CONTENT	US7043758	Filtrage de courriers électroniques et de messages
HITACHI LTD	25 08 1999	EP1063833 SYSTEME DE FILTRAGE DES DONNEES UTILISANT UN FILIGRANE ELECTRONIQUE	JP19990176289 JP2001005757 EP1063833	Taggage de contenus et affectation de métadonnées
HOME BOX OFFICE INC, GABRIEL MICHAEL, PROBST BRUCE E, DIBARTOLOMEO JEFFREY	23 07 2004	WO2005013056 CONTRÔLE D'ACCÈS À UN CONTENU	EP1654617 US20030627002	Taggage de contenus et affectation de métadonnées
HUAWEI TECH CO LTD, ZHAO QIN, ZHU YONGSHENG	28 05 2007	WO2008009224 SYSTÈME, DISPOSITIF ET PROCÉDÉ DE FILTRAGE DE CONTENU	CN20061100874 CN20061138277	Filtrage de flux
HUBEY HACI-MURAT	30 05 2002	US2003065632 SCALABLE, PARALLELIZABLE, FUZZY LOGIC, BOOLEAN ALGEBRA, AND MULTIPLICATIVE NEURAL NETWORK BASED CLASSIFIER, DATAMINING, ASSOCIATION RULE FINDER AND VISUALIZATION SOFTWARE TOOL	US20010294314P	Analyse de bases de données et extractions

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATÉGORIE DE FILTRAGE
HUNTINGTON STEPHEN G, COVINGTON STANLEY P	17 07 2002	US2003131098 NETWORK DATA RETRIEVAL AND FILTER SYSTEMS AND METHODS	US20020199451 US20010306107P US20010306056P US20010306106P US20010306792P US20010311142P US7149189	Filtrage web et/ou filtrage réseau
HUNTINGTON STEPHEN G, COVINGTON STANLEY P, MAJOR JOHN D, ROWLEY BEVAN S	14 09 2006	US2007011321 NETWORK DATA RETRIEVAL AND FILTER SYSTEMS AND METHODS	US20060531990; US20020199451 US20010306056P US20010306106P US20010306792P US20010311142P US7315894	Filtrage web et/ou filtrage réseau
IBM	16 08 2000	EP1079315 SYSTEME ET METHODE POUR PRENDRE EN COMPTE DES CARACTERISTIQUES SEMANTIQUES DANS LE CADRE DU TRANSCODAGE SYNTACTIQUE ET REGI PAR LE FORMAT DE DOCUMENTS	US19990383742 US6993476 JP2001109743 CA2313558 CN1142512C	Taggage de contenus et affectation de métadonnées
IBM	20 02 2001	US2002120369 SYSTEM AND METHOD TO MONITOR DATAMINING POWER USAGE	US6631309	Analyse de bases de données et extractions
IBM, BLUMRICH MATTHIAS, GARA ALAN, SALAPURA VALENTINA	17 03 2006	WO2006104747 PROCEDE ET APPAREIL SERVANT A FILTRER DES DEMANDES DE FURETAGE AU MOYEN DE REGISTRES DE SEQUENCES CONTENUES	US20050093130 EP1864224 US2006224836	Filtrage de flux

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
INSTITUT NATIONAL DE L'AUDIOVISUEL - INA	20 04 2006	US20060083429 SEARCH OF SIMILAR FEATURES REPRESENTING OBJECTS IN A LARGE REFERENCE DATABASE		Analyse de bases de données et extractions
INTERNATIONAL BUSINESS MACHINES CORPORATION	22 07 1998	EP0893920A2 SYSTEME POUR LA MODIFICATION DYNAMIQUE DU CONTENU D'UN FLUX DE DONNEES MULTIMEDIA	US898220	Filtrage de flux
IP CO LLC	02 07 2003	WO2004006063 SYSTEME, PROCEDE ET PRODUIT LOGICIEL POUR LE FILTRAGE SELECTIF DE CONTENU INCONVENANT DANS UNE EMISSION	WO2004006063 US20020187540	Analyse de vidéo
ISSARD BERTRAND	24 11 2005	FR2893731 METHODE ET SYSTEME DE CONTROLE D'ACCES AUX CONTENUS MULTIMEDIA ET DE FILTRAGE DES MESSAGES DE SOLLICITATION		Filtrage de courriers électroniques et de messages
JARMAN MATTHEW, VENN CHRISTOPHER, IVERSON BRENT	30 01 2007	WO2007120963 SYNCHRONIZING FILTER METADATA WITH A MULTIMEDIA PRESENTATION	US20060763525P US20060785547P US20070669138 US2007186235	Taggage de contenus et affectation de métadonnées
JUSTSYSTEM PITTSBURGH RESEARCH	18 10 1999	WO0026795 PROCEDE DE FILTRAGE DE MESSAGES SUR LA BASE DU CONTENU, PAR ANALYSE DES CARACTERISTIQUES DES TERMES A L'INTERIEUR DU MESSAGE	US19980183871	Filtrage lexicographique

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
KENT RIDGE DIGITAL LABS, LI HAIZHOU, WU JIANKANG, NARASIMHALU A DESAI	27 01 1999	WO0045375 PROCEDE ET APPAREIL DESTINES A L'ANNOTATION ET LA RECUPERATION VOCALES DE DONNEES MULTIMEDIA	US6397181 GB2361339	Analyse du son et/ou de la parole
KONINKILJKE PHILIPS ELECTRONICS N.V., YULE DAVID, BELL DAVID	12 12 2002	WO2003061216 SYSTEME DE TRANSFERT ET DE FILTRAGE DE DONNEES DE CONTENU VIDEO	EP02250350 US2004263914 CN1615615 AU2002356362 DE60219079T	Analyse de vidéo
KONINKL PHILIPS ELECTRONICS NV	06 11 2001	EP1336297 PROCEDE ET SYSTEME PERMETTANT DE LIMITER DES REPRESENTATIONS REPETITIVES PAR UN FILTRAGE DE CONTENU	WO2001EP13001 US20000709266 WO0239730 US6829778 CN1416644	Analyse de vidéo
KONINKL PHILIPS ELECTRONICS NV	15 03 2002	WO02080530 SYSTEME DE CONTROLE PARENTAL DANS DES PROGRAMMES VIDEO EN FONCTION D'INFORMATIONS RELATIVES AU CONTENU MULTIMEDIA	EP1378121 US20010822436	Taggage de contenus et affectation de métadonnées
KONINKL PHILIPS ELECTRONICS NV , PHILIPS CORP, HOLLEMANS GERRIT, BUIL VINCENT P	28 11 2006	WO2007063497 SYSTEME ET PROCEDE POUR PRESENTER UN CONTENU A UN UTILISATEUR	US20050741297P	Prototypage de données

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
KONINKL PHILIPS ELECTRONICS NV, PHILIPS SVENSKA AB	25 05 1998	WO9855942 VISUAL INDEXING SYSTEM	EP0916120 US6185363 US6125229 US19970867145	Taggage de contenus et affectation de métadonnées
LEBRAT FRANCOIS, TDF	18 03 2005	EP1741047 PROCÉDÉ DE RECHERCHE DE CONTENU, NOTAMMENT D'EXTRAITS COMMUNS ENTRE DEUX FICHIERS INFORMATIQUES	WO2005FR00673 FR20040003556	Filtrage lexicographique
LEGEND BEIJING LTD, WANG JIANG, GAO JIANZHONG, WANG NAN, ZHU GUANG, XIAO HANG	23 05 2002	WO03038667 FILTRE A CONTENU FONCTIONNANT PAR COMPARAISON ENTRE LA SIMILARITE DE CARACTERES DE CONTENU ET LA CORRELATION DE LA MATIERE	CN20011031420	Filtrage lexicographique
LUCENT TECHNOLOGIES INC	16 07 2004	EP1505603 SYSTEME D'IDENTIFICATION DE CONTENU	US20030629486 US2005027766 JP2005049878 CN1604081	Analyse de vidéo
MATRA COMMUNICATION (FR)	22 01 1996	EP0752181 FREQUENCY-DOMAIN ADAPTIVE-FILTER ACOUSTIC ECHO CANCELLER	WO1996FR00100 FR19950000777 WO9623384 US5848151 FR2729804 FI963769 CN1145710 CA2186281 BR9603890 EP0752181 ES2158276T AU708388B	Analyse de son et/ou de la parole
MCAFFEE INC	10 05 2002	US7237008		Filtrage de courriers

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
		DETECTING MALWARE CARRIED BY AN E-MAIL MESSAGE		électroniques et de messages
MCAFFEE INC, DIXON CHRISTOPHER J, PINCKNEY THOMAS	02 05 2006	WO2006119479 DETERMINATION DE REPUTATIONS DE SITES WEB FAISANT APPEL A UN TEST AUTOMATIQUE	US20050677786P US20050691349P US20060342297 US2006253458	Filtrage web et/ou filtrage réseau
MCAFFEE INC, DIXON CHRISTOPHER J, PINCKNEY THOMAS	02 05 2006	WO2006119481 INDICATION DE LA REPUTATION DE SITES WEB DANS DES RESULTATS DE RECHERCHE	US20050677786P US20050691349P US20060342322 US2006253582	Filtrage web et/ou filtrage réseau
MICROSOFT CORP	01 07 1999	EP1060597 PROCEDE ET DISPOSITIF DE FILTRAGE DE CONTENU	EP1058874 EP1051823 WO9935802 EP1051681 WO9935778 EP1051824 EP1053525 EP1840698 WO9935801 WO9935557 WO9935593 WO9935591 WO1999US00337 US19980070720P US19980075123P US19980107724 US19980107666 US19980189591	Autres types de filtrages de contenus
MICROSOFT CORP	23 05 2003	EP1385097 FILTRAGE DE CONTENU POUR LA	US20020183657 US2004006621	Filtrage web et/ou filtrage réseau

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
		RECHERCHE WEB	RU2003119093 MXPA03005801 KR20040002656 JP2004030678	
MICROSOFT CORP	19 05 2004	EP1489799 DISSIMULATION D'UN FILTRE DE MESSAGES DE COURRIER ELECTRONIQUE NON SOLLICITE (SPAM)	US20030601034 US2005015454 KR20040110086 JP2005011326 CN1573780	Filtrage de courriers électroniques et de messages
MICROSOFT CORP	19 08 2004	EP1513349 FILTRAGE D'UN FICHIER VIDEO EN POST- PRODUCTION A PARTIR D'INFORMATIONS DE CONTROLE CONTENUES DANS LE FICHIER LUI-MEME	US20030501081P US20030680072 US2005053288 KR20050025928 JP2005086830 CN1627824	Analyse de vidéo
MOORE ROBERT EDWARD (GB), HOWARD FRASER PETER	26 12 2001	US2003120947 IDENTIFYING MALWARE CONTAINING COMPUTER FILES USING EMBEDDED TEXT	US7114185	Filtrage lexicographique
NEUSTAR INC, FRIDMAN SHARON, VOLACH BEN	24 08 2007	WO2008025008 SYSTEME ET PROCEDE POUR FILTRER UN CONTENU D'INFORMATIONS CHOQUANTES DANS DES SYSTEMES DE COMMUNICATION	US20060839705P US20060839703P	Filtrage de flux

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
NIPPON ELECTRIC CO	25 06 2007	WO2007148817 SYSTEME DE RECOMMANDATION DE CONTENU, PROCEDE DE RECOMMANDATION DE CONTENU, ET PROGRAMME DE RECOMMANDATION DE CONTENU	JP20060173805 JP2004343321 JP2005327028	Prototypage de données
NOKIA CORPORATION, NOKIA INC	17 12 2001	WO02054302 FILTRAGE DE CONTENU PAR L'INTERMEDIAIRE DE METADONNEES D'IDENTIFICATION PIXEL PAR PIXEL	US09753844	Taggage de contenus et affectation de métadonnées
NOKIA CORP	04 02 2002	EP1364497 FILTRAGE DE CONTENU PUBLICITAIRE FORCE	WO02069585 WO2002IB00435 US20010794373	Filtrage lexicographique
PCSAFE INC	10 12 2004	EP1638016 METHODES ET APPAREILS ADAPTES A FILTRER DES URL, PAGES WEB, ET CONTENU	US20040609843P WO200603617	Filtrage web et/ou filtrage réseau
PREVUE INTERNATIONAL INC	01 05 1998	WO9852357 SYSTEME DE FILTRAGE DU CONTENU DE VIDEO	US19970857977 19970516	Analyse de vidéo
SAP AG	29 08 2003	EP1510935 MISE EN CORRESPONDANCE DE DONNES DU DATAWAREHOUSE A UN DATAMART	WO2005022406 US2005114248	Analyse de bases de données et extractions
SCREAMINGMEDIA INC	09 11 2000	WO0135281 MOTEUR DE CONTENU	US19990438004 19991110	Analyse de bases de données et extractions
SECURE COMPUTING CORP	18 09 1997	DE19741238 SYSTEM AND METHOD OF ELECTRONIC MAIL FILTERING	GB2317793 US19960715336 US19960715333	Filtrage de courriers électroniques et de messages

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATÉGORIE DE FILTRAGE
SECURE COMPUTING CORP, JUDGE PAUL, ALPEROVITCH DMITRI, KRASSER SVEN, SCHNECK PHYLLIS A, ZDZIARSKI JONATHAN A	06 06 2007	WO2007146696 SYSTEMES ET PROCÉDÉS PERMETTANT D'IDENTIFIER DES MESSAGES POTENTIELLEMENT MALVEILLANTS	US20060423313	Filtrage de courriers électroniques et de messages
SECURE COMPUTING CORP, JUDGE PAUL, SCHNECK PHYLLIS ADELE, YANG WEILAI, ZDZIARSKI JONATHAN ALEXANDER	10 11 2006	WO2007059428 SYSTEMES ET PROCEDES POUR FAIRE OBSERVER DES REGLES FONDEES SUR UN CONTENU	US20050736121P US20060383347	Filtrage de flux
SONY COMP ENTERTAINMENT	02 06 2004	EP1649407 PROCEDES ET SYSTEMES DE FORMATION DE FILTRES DE CONTENU ET DE RESOLUTION D'INCERTITUDE DANS DES OPERATIONS DE FILTRAGE DE CONTENU	EP1636968 WO2004109588 EP1636698 WO2004110022 WO2004110024 WO2004110020 WO2004110019 WO2004110023 WO2004109514 WO2004110018 WO2004109490	Analyse de bases de données et extractions

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
SONY CORP	24 11 2004	EP1538838 PROCESSEUR D'INFORMATION, PROCEDE DE TRAITEMENT D'INFORMATION ET PROGRAMME INFORMATIQUE	JP20030403728	Filtrage de flux
SONY CORPORATION	04 05 2007	EP1857963 DISPOSITIF D'ENREGISTREMENT, DISPOSITIF DE COLLECTE, PROCEDE ET PROGRAMME D'EXTRACTION	JP2006137824 JP2006256498	Filtrage de flux
SONY ELECTRONICS INC	04 04 2003	EP1495411 FILTRAGE D'UN CONTENU PAR UN MECANISME D'APPRENTISSAGE	WO03088061 WO2003US10223 US20020371111P US20030400018 US2003191753 GB2404060 CN1659531	Autres types de filtrages de contenus
SOPHOS PLC	16 08 2004	EP1509014 PROCEDE ET DISPOSITIF DE FILTRAGE DE COURRIER ELECTRONIQUE	US2005041789 GB2405229	Filtrage de courriers électroniques et de messages
SOPHOS PLC	19 06 2007	GB2439806 CLASSIFYING SOFTWARE AS MALWARE USING CHARACTERISTICS (OR "GENES")	US2008005796	Prototypage de données
SOURCELABS INC, PUGH WILLIAM, SWEET RYAN, JACOBSON STEVE, HANSSON CHRISTIAN, JEKEL ROSS ARDEN, RUAN YONGSHAO	26 06 2007	WO2008002937 PROCEDE DESTINE A AMELIORER L'EFFICACITE D'UN DISAGNOSTIC LOGICIEL EN EXPLOITANT LE CONTENU EXISTANT, LE FILTRAGE HUMAIN ET LES OUTILS DE DIAGNOSTIC AUTOMATISES	US20060816797P US2008034351	Autres types de filtrages de contenus

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
ST MICROELECTRONIC S SA	22 12 2000	FR2818859 PROCEDE ET DISPOSITIF DE FILTRAGE DES DONNEES RELATIVES AU GUIDE ELECTRONIQUE DE PROGRAMMES D'UN TELEVISEUR	EP1217831	Analyse de vidéo
SURFMONKEY COM INC	18 10 2000	WO0133371 SYSTEME ET PROCEDE DE FILTRAGE DU CONTENU POUR ADULTES SUR INTERNET	US19990435142	Filtrage web et/ou filtrage réseau
SYMANTEC CORP	30 11 1999	US6539430 SYSTEM AND METHOD FOR FILTERING DATA RECEIVED BY A COMPUTER SYSTEM	US5996011 US19970823123	Filtrage de flux
TELEFUNKEN SYSTEMTECHNIK	08 09 1990	DE4028602 IDENTIFYING HOPPING RADIO TRANSMISSIONS - TESTING ON BASIS OF PREVIOUSLY USED FREQUENCIES IN SEARCH RECEIVER BY TIME FILTER FUNCTION FOR COINCIDENCE		Analyse de son et/ou de la parole
THALES SA	15 07 2003	FR2857766 PROCEDE D'EXTRACTION D'INFORMATIONS PERTINENTES AVEC PRISE EN COMPTE DE L'OBJECTIF ET DE LA CIBLE	WO2005017771 EP1644852 US2006173802	Analyse et ou traitement d'images

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATÉGORIE DE FILTRAGE
THOMSON LICENSING, KENDALL SCOTT ALLAN, STUART ANTHONY EDWARD, SAHASRABUDHE RAJEEV MADHUKAR	14 04 2005	WO2006112822 REPLACEMENT AUTOMATIQUE DE CONTENU REPREHENSIBLE DE SIGNAUX AUDIO		Analyse du son et/ou de la parole
THOMSON LICENSING, LIN SHU, VANDERSCHAAR AUKE SJOERD	11 04 2006	WO2007117240 TECHNIQUE DE FILTRAGE ADAPTEE AU CONTENU / CONTENT-ADAPTIVE FILTER TECHNIQUE		Autres types de filtrages de contenus
TONEGUZZO GROUP PTY LTD	09 07 2001	EP1301890 FILTRAGE ET GESTION DE CONTENU	WO2001AU00823 AU2000PQ08657 WO0205148 US2003182573	Prototypage de données
TTP COMMUNICATIONS LTD, TINDALL PAUL GEOFFREY	06 10 2003	WO2004059959 PROCEDE DE FILTRAGE DE MESSAGES TEXTUELS DANS UN DISPOSITIF DE COMMUNICATION	EP1576792 GB20020030219	Filtrage lexicographique
VIGIL FRANK	06 10 2006	WO2007050368 SYTEME ET PROCEDE MIS EN OEUVRE PAR ORDINATEUR PERMETTANT D'OBTENIR DES INFORMATIONS SUR MESURE LIEES A UN CONTENU MULTIMEDIA	US2007136247 US2007094245	Taggage de contenus et affectation de métadonnées
WEBSense INC	28 01 2000	WO0155905 AUTOMATED CATEGORIZATION OF INTERNET DATA		Filtrage web et/ou filtrage réseau

DEPOSANT / TITULAIRE	DATE DE DEPOT	NUMERO ET TITRE	FAMILLE DE BREVETS	CATEGORIE DE FILTRAGE
WEBSense INC	06 12 2002	EP1318468 SYSTEME ET PROCEDE POUR UN FILTRE INTERNET	US20010017750	Filtrage web et/ou filtrage réseau
WEBSense INC, BADDOUR VICTOR L, CHENETTE STEPHAN, HUBBARD DAN, VERENINI NICHOLAS J, MESDAQ ALI A	09 07 2007	WO2008008339 SYSTÈME ET PROCÉDÉ D'ANALYSE DE CONTENU WEB		Filtrage web et/ou filtrage réseau
WEBSense INC, HUBBARD DAN, VERENINI NICHOLAS J	09 07 2007	WO2008008219 SYSTÈME ET PROCÉDÉ D'ANALYSE D'UN CONTENU INTERNET	US20060484240	Filtrage web et/ou filtrage réseau
WEBSense, INC.	28 01 2000	US6606659 SYSTEM AND METHOD FOR CONTROLLING ACCESS TO INTERNET SITES		Filtrage web et/ou filtrage réseau

ANNEXE 6

ABSTRACT DES BREVETS

ANNEXE 7

EXTRAITS DE CONDITIONS GENERALES D'UTILISATION

1. Conditions générales d'utilisation de Yahoo au 19 mars 2008

1.1.1.6.2 Des outils visant à la protection des mineurs sont disponibles

Sur le service Yahoo! Search, vous pouvez, si vous le souhaitez, intervenir dans la configuration des résultats de recherche en vous rendant dans la section « Préférences » de Yahoo! Search et en paramétrant le filtre des contenus adultes. L'efficacité de ces filtres est estimée à 95%. **Il demeure que les adultes ayant la garde de mineurs ont l'obligation de surveiller leur utilisation d'Internet. Il est ainsi de leur responsabilité de déterminer les services et les utilisations qu'ils jugent adaptés à ces mineurs.** Pour ce faire, ils pourront s'inspirer des conseils prodigués par l'AFA, en collaboration avec les pouvoirs publics à l'adresse :

<http://www.pointdecontact.net/protectiondelenfance.html>

Lorsque vous mettez en ligne des contenus qui ne sont pas « tous publics » sur Yahoo! Groupes, Yahoo! 360° ou Flickr, veuillez les placer dans les catégories « Adulte » ou utiliser la fonctionnalité équivalente mise à votre disposition dans le service. Cela vise à empêcher l'accès à des internautes mineurs à des contenus qui ne leur sont pas destinés.

Vous vous interdisez dans le cadre de l'utilisation des Services de vous livrer à des actes, de quelque nature que ce soit (notamment à des actes de consultation, téléchargement, envoi, diffusion, édition, émission, mise en ligne, publication ou de toute autre manière), qui seraient contraires à la loi française, porteraient atteinte à l'ordre public français, ou aux droits d'un tiers. En particulier, et sans que cette liste soit limitative, vous vous interdisez de :

1. vous livrer à des actes constitutifs d'apologie des crimes contre l'humanité, de négation de génocides, d'incitation à la violence, à la haine raciale ou à la pornographie infantile ;
2. vous livrer à des actes de diffamation, d'injure, de menaces, de chantage, de harcèlement ou à des actes attentatoires à la vie privée ou à la dignité humaine ;
3. porter atteinte d'une quelconque manière aux utilisateurs mineurs, de les inciter à se mettre en danger d'une quelconque manière ;
4. en particulier, transmettre, diffuser, éditer, publier ou rendre accessible tout Contenu qui pourrait être constitutif, sans que cette liste soit limitative, d'incitation à la réalisation de crimes et délits ; de propagation de fausses nouvelles ou d'informations financières couvertes par le secret, de même que tout Contenu destiné à représenter ou proposer à la vente des objets et/ou des ouvrages, des logiciels, des Contenus interdits par la loi ou portant atteinte aux droits de tiers ; d'atteinte à l'autorité de la justice ; d'atteinte à la vie privée, à la protection des données personnelles ou au secret des correspondances ; de divulgation d'informations couvertes par un secret relatives, notamment à l'adoption plénière, à un procès en cours, au suicide, ou à la santé d'un tiers, ou à une situation patrimoniale ou financière individuelle couverte par le secret ou par le droit à l'intimité de la vie privée ; ou encore d'acte mettant en péril des mineurs notamment par la fabrication, la transmission, la diffusion ou l'accessibilité de messages à caractère violent ou pornographique, de nature à porter atteinte à la dignité humaine ou de nature à permettre la fabrication d'explosifs ;
5. tenter d'induire en erreur d'autres utilisateurs en usurpant l'identité ou une dénomination sociale ou en portant atteinte à l'image ou à la réputation d'autres personnes et/ou en vous faisant passer pour un tiers ou pour un employé, un service habilité ou un affilié de Yahoo! ;
6. falsifier des données, messages ou documents, des en-têtes de messages ou de données d'identification ou de connexion à un Service ou manipuler de toute autre manière un identifiant de manière à dissimuler l'origine de la transmission d'un Contenu via le Service ;
7. vous livrer à une violation des droits de propriété intellectuelle (notamment en matière de musique, vidéo, animations, jeux, logiciels, bases de données, images, sons et textes) ou tout autre droit de propriété (ci-après dénommés collectivement les « Droits ») appartenant à autrui ;
8. expédier ou faire expédier des courriers électroniques ou des messages instantanés à des personnes qui ne les ont pas sollicités ou sans avoir respecté leurs droits reconnus par la loi, tels que des publicités, du matériel promotionnel, des chaînes de lettres ou toute autre forme de prospection directe non sollicitée ; mettre en ligne des messages à caractère promotionnel sur les Services Yahoo! ;
9. télécharger sciemment, afficher, émettre, diffuser, transmettre ou rendre accessible de toute autre manière tout Contenu comprenant ou constituant des virus informatiques ou tout autre code ou programme informatique conçus pour interrompre, détruire, détourner ou limiter les fonctionnalités ou les performances de tout logiciel, ordinateur, service ou outil de communications électroniques sans que cette énumération ne soit limitative ;
10. perturber, ralentir, bloquer ou altérer le flux normal des données échangées dans le cadre du Service, accélérer le rythme de défilement des Contenus du Service de telle manière que le fonctionnement du Service soit modifié ou altéré ou commettre toute autre action ayant un effet perturbateur équivalent sur les fonctionnalités du Service ;
11. accéder frauduleusement, se maintenir, entraver ou perturber les systèmes d'information de Yahoo ! et notamment des Services, les serveurs, les réseaux connectés au Service, ou refuser de se conformer aux conditions requises, aux procédures, aux règles générales ou aux dispositions réglementaires applicables aux réseaux connectés au Service.

2. Conditions générales d'utilisation de Lycos au 16 mai 2008

XVIII. Conditions complémentaires relatives à certains services

Lycos propose différents services sur Internet. Selon leur objectif et leur nature, ces services sont équipés de différentes fonctions et comportent des caractéristiques propres dont les principales sont succinctement décrites ci-après. Pour certains services, l'exécution d'une nouvelle procédure d'inscription est nécessaire. En complément, les conditions suivantes s'appliquent à l'utilisation des différents services :

1 Jubii

Jubii est une plateforme de communication qui offre à ses utilisateurs :

- Une large gamme de moyens de communication, tels que notamment : messagerie électronique, SMS, et messagerie instantanée;
- Stockage en ligne et partage de fichiers tels que notamment : photos, vidéos et autres documents ;
- Gestion et administration des contacts.

Jubii organise les courriers électroniques "entrants" en fonction de leur fiabilité dans différents dossiers, et/ou identifie de manière particulière certains de ces courriels. Les critères pour qu'un courrier électronique soit considéré comme fiable sont déterminés en fonction de votre comportement passé (i.e. est-ce que l'expéditeur est déjà dans vos contacts ; à quelle fréquence répondez-vous aux courriels de l'expéditeur ; effacez-vous immédiatement les courriels de l'expéditeur ?).

Jubii fournit des filtres anti-spam et anti-virus pour protéger les systèmes de traitement de l'information de Lycos, de la même manière que votre boîte de messagerie électronique. Les courriers électroniques suspectés d'être des spams sont marqués et placés dans des dossiers désignés en tant que tels où ils peuvent être consultés par vous.

Lycos décline expressément toute responsabilité pour le cas où lesdits filtres échoueraient à filtrer tous les spams, virus, ou autres logiciels dangereux.

A chaque courrier électronique envoyé par l'intermédiaire de Jubii, Lycos pourra ajouter une mention indiquant l'origine du message (par exemple "powered by lycos"), ou de la publicité.

L'utilisation de Jubii pour envoyer massivement des courriers électroniques (« Spamming »), faire du « mail-bombing » ou envoyer toute autre forme de message publicitaire ou de marketing vous est interdite et toute personne utilisant Jubii à de telles fins sera tenue pour seule et entière responsable des conséquences d'un tel acte. Il vous est interdit de déguiser ou de masquer votre identité lorsque vous envoyez un courrier électronique par l'intermédiaire de Jubii.

2 Espace de stockage gratuit

Dans la mesure où les Services Lycos que vous avez choisis contiennent un service d'espace de stockage en ligne pour la publication de Contenus sur le World Wide Web (i.e. hébergement Internet gratuit, weblogs, listes de liens, réseaux d'experts, services de rencontres), les conditions suivantes s'appliquent :

Lycos ne saurait être tenu pour responsable des difficultés ou des problèmes techniques que vous pourriez rencontrer avec le stockage en ligne ou avec l'outil de stockage.

Vous vous engagez à ne pas mettre l'espace de stockage à la disposition de tiers pour stocker leurs données et à ne pas utiliser l'espace mis à votre disposition par Lycos à d'autres fins que celles découlant des Services Lycos.

Dans l'hypothèse où vous inséreriez le contenu hébergé par Lycos au sein de sites Web extérieurs au réseau Lycos, Lycos se réserve le droit de désactiver ladite insertion ou de l'accepter en ajoutant à ce contenu la mention que celui-ci est hébergé par Lycos (i.e. "hébergé par Lycos").

Lycos se réserve le droit de limiter l'espace d'hébergement en ligne qu'il vous fournit.

3 Salons de conversation / Forum

Il est interdit de faire du commerce dans ou par l'intermédiaire des salons de conversation et des forums.

4 Informations financières

En collaboration avec ses partenaires, Lycos met à disposition, entre autres, des services procurant des informations sur des investissements. Les cours boursiers et les cours de fonds sont mis à disposition pour votre usage strictement privé exclusivement. Toutes les informations concernant des cours sont représentées différemment selon la bourse et le type de valeur sélectionnés, en général avec une temporisation de quinze (15) minutes. Malgré une collecte et une mise à disposition minutieuses, Lycos et ses partenaires n'engagent pas leur responsabilité du fait de la conformité, l'intégralité ou l'exactitude des informations boursières et économiques, cours, indices, prix, messages, informations générales sur le marché et autres contenus accessibles ("Informations financières"), mis à disposition pour consultation, affichés ou envoyés. Les Informations financières ainsi que les résultats créés et affichés avec les outils de représentation proposés sont exclusivement destinés à l'information des utilisateurs et ne peuvent être considérés comme un conseil en placement ou une autre recommandation obligatoire.

Dans la mesure où des informations financières sont envoyées à un numéro de téléphone mobile sous forme de messages courts (SMS) à la demande de l'utilisateur, ce dernier n'est autorisé à demander l'envoi qu'à son propre numéro de téléphone mobile exclusivement. En outre, les règles ci-dessus s'appliquent à l'envoi d'Informations financières par messages courts (SMS), dans la mesure où elle sont par nature applicables à l'envoi d'informations financières par l'intermédiaire des Services Lycos.

L'envoi d'informations financières sous forme de message court sera interrompu après quatre semaines si le membre ne fait pas appel au service du domaine "Informations financières" dans ce laps de temps.

5 Service WLAN

Dans le cadre de l'offre WLAN de Lycos, vous avez la possibilité de signaler la position d'un point d'accès WLAN ("Hotspot"). Toutefois, ceci ne vous est permis que si vous êtes le propriétaire de ce Hotspot ou si vous êtes habilité à cette fin par le propriétaire.

6 Sonique Media Player

Le Sonique Media Player ainsi que le logiciel pour réaliser des Plug-Ins et Skins vous sont proposés sous forme de téléchargement gratuit aux conditions du contrat de licence affiché dans le programme d'installation desdits logiciels. Lycos n'est pas responsable des éventuelles erreurs du logiciel Sonique et des dommages éventuels en résultant, quelle qu'en soit la nature (ni pour des supports de données défectueux, des erreurs dans l'infrastructure du réseau, la stabilité du système, etc.). Seuls les utilisateurs majeurs sont autorisés à proposer des Skins.

3. Les conditions générales d'utilisation de YouTube (Google) au 20 mars 2008

[HTTP://FR.YOUTUBE.COM/T/COPYRIGHT_NOTICE](http://fr.youtube.com/t/copyright_notice)

1. NOTIFICATION D'INFRACTION AUX DROITS D'AUTEUR

Pour envoyer une notification d'infraction, vous devez d'abord nous faire parvenir un communiqué écrit contenant les informations indiquées, en respectant le format suivant :

1. Envoyez-nous une déclaration stipulant que vous avez trouvé une vidéo sur YouTube qui, selon vous, constitue une infraction à vos droits d'auteur. Exemple : "Je confirme par la présente que j'estime que la vidéo identifiée ci-dessous constitue une infraction à mes droits d'auteur".
2. Indiquez-nous le pays dans lequel vos droits d'auteur s'appliquent.
3. Indiquez-nous le titre de la vidéo et l'URL complète de sa page de lecture.
4. Expliquez-nous pourquoi la vidéo constitue une infraction à vos droits d'auteur (le son a été copié, l'ensemble de la vidéo est une copie d'une de vos œuvres originales, etc.).
5. Identifiez la nature (film, morceau de musique, livre, etc.) et les informations (titre, éditeur, dates, etc.) de la vidéo dont vous êtes le titulaire des droits et qui vous semblent faire l'objet d'une infraction. Si ces informations sont disponibles sur Internet, nous vous conseillons de nous envoyer le lien.
6. N'oubliez pas de nous fournir vos coordonnées pour que nous puissions vous répondre (adresse électronique de préférence).
7. Donnez-nous vos coordonnées que nous pourrions faire suivre à l'expéditeur de la vidéo concernée. Ainsi, il pourra prendre contact directement avec vous pour résoudre le litige (adresse e-mail de préférence).
8. Ajoutez la déclaration suivante : "Je pense en toute bonne foi que l'utilisation du travail protégé par droits d'auteur décrite ci-dessus n'est pas autorisée par le titulaire des droits d'auteur (ou par un tiers le représentant) et n'est pas non plus autorisée par la loi."
9. Signez la déclaration. Si vous envoyez la déclaration par e-mail, vous pouvez y joindre votre signature scannée ou électronique.
10. Envoyez la déclaration par écrit à l'adresse suivante :

Attn	:	YouTube	Copyright	Infringement	Notification
1st		and		2nd	Floor
Gordon					House
Barrow					Street
Dublin					4
Irlande					
Adresse		e-mail	:		frcopyright@youtube.com
Fax : +353 1 437 0741					

Vous devez être sûr que le contenu que vous avez vu sur YouTube constitue une infraction à vos droits d'auteur. Si vous n'êtes pas sûr de bien connaître vos droits ou si vous n'êtes pas sûr qu'il y a atteinte à vos droits d'auteur, adressez-vous à un conseiller juridique avant de poursuivre. Attention : Dans votre pays, vous risquez peut-être des poursuites en cas d'allégation de mauvaise foi concernant des infractions aux droits d'auteur dans le cadre de ce processus.

Si vous avez beaucoup de vidéos à supprimer ou si vous pensez avoir souvent besoin de supprimer du contenu potentiellement illicite de YouTube, nous vous suggérons de vous inscrire à notre [Programme de vérification de contenu](#). Ce système nous permet d'être prévenus automatiquement, ne laissant aucune place à l'erreur, et augmente fortement la vitesse à laquelle nous pouvons supprimer tout contenu illicite.

1.1 NOTIFICATION DE CONTESTATION


Si vous décidez de nous envoyer un avis de contestation, veuillez suivre les instructions de notre [Centre d'aide](#). Toute personne indiquant sciemment et de façon erronée qu'un contenu ou une activité a été supprimée ou désactivée par erreur, pourra voir sa responsabilité engagée. Notez également que nous appliquons une politique qui offre la possibilité de résilier, le cas échéant, le compte ou l'abonnement des utilisateurs ayant violé le règlement en vigueur à plusieurs reprises.

Haut du formulaire

Nom d'utilisateur :

Mot de passe :

Bas du formulaire

[Inscription](#) [Compte](#) [Historique](#) [Aide](#) [Connexion](#) Site : 

Haut du formulaire

Bas du formulaire

Accueil

Vidéos

Chaînes

Communauté

[Envoyer une vidéo](#)

Haut du formulaire

Bas du formulaire

[paramètres](#)

Haut du formulaire

Bas du formulaire

Programme de vérification du contenu

YouTube s'engage à aider les détenteurs de droits d'auteur à trouver et supprimer du site les contenus présumés en infraction aux droits de propriété. C'est pourquoi, nous avons créé un outil de vérification des droits d'auteur qui permet aux titulaires de ces droits de rechercher les contenus qu'ils estiment être en infraction et de fournir à YouTube les informations requises pour localiser ces contenus.

Cet outil est spécialement conçu pour que les entreprises détentrices de droits d'auteur puissent soumettre de multiples demandes de suppression. Il est possible d'envoyer des notifications individuelles en suivant ces instructions.

Si vous avez déjà un compte YouTube, vous pouvez demander l'accès à cet outil en remplissant une demande de participation au programme de vérification du contenu de YouTube. Imprimez-la, puis envoyez-la par fax au numéro indiqué sur le document. Ce formulaire identifie vos représentants et atteste légalement que vous êtes le détenteur des droits d'auteur du document au sujet duquel vous voulez contacter YouTube. Si vous n'avez pas de compte, veuillez en créer un. Vous pourrez ensuite accéder à la demande de participation au programme de vérification du contenu.

Haut du formulaire

Bas du formulaire

Mon compte

[Vidéos](#) [Boîte de réception](#)
[Favoris](#) [Abonnements](#)
[Playlists suite...](#)

Aide et informations

[Centre d'aide](#) [Conseils de sécurité](#)
[Boîte à outils vidéo](#) [Droits d'auteur](#)
[API de développement](#) [Règlement de la communauté](#)
[Publicité](#)

YouTube

[Qui sommes-nous ?](#) [Presse](#)
[TestTube](#) [Contact](#)
[Conditions d'utilisation](#) [Blog](#)
[Confidentialité](#)

1. PROGRAMME DE VERIFICATION DU CONTENU

YouTube s'engage à aider les détenteurs de droits d'auteur à trouver et supprimer du site les contenus présumés en infraction aux droits de propriété. C'est pourquoi, nous avons créé un outil de vérification des droits d'auteur qui permet aux titulaires de ces droits de rechercher les contenus qu'ils estiment être en infraction et de fournir à YouTube les informations requises pour localiser ces contenus.

Cet outil est spécialement conçu pour que les entreprises détentrices de droits d'auteur puissent soumettre de multiples demandes de suppression. Il est possible d'envoyer des notifications individuelles en suivant [ces instructions](#).

Si vous avez déjà un compte YouTube, vous pouvez demander l'accès à cet outil en remplissant une [demande de participation au programme de vérification du contenu de YouTube](#). Imprimez-la, puis envoyez-la par fax au numéro indiqué sur le document. Ce formulaire identifie vos représentants et atteste légalement que vous êtes le détenteur des droits d'auteur du document au sujet duquel vous voulez contacter YouTube. Si vous n'avez pas de compte, veuillez [en créer un](#). Vous pourrez ensuite accéder à la demande de participation au programme de vérification du contenu.

4. **Contenu du service.** Google décline toute responsabilité quant aux contenus de tiers (notamment, tout virus ou autre programme malveillant) et Google n'a aucune obligation de surveillance de tels contenus de tiers. Google se réserve le droit à tout moment de retirer ou de refuser de diffuser tout contenu sur le service, tels que les contenus qui contreviennent aux stipulations du présent Accord. Google se réserve également le droit d'accéder, de lire, de conserver, et de communiquer toute information dont elle pense raisonnablement qu'elle est nécessaire aux fins de (a) respecter toute obligation légale ou réglementaire applicable, ainsi que toute demande d'une autorité judiciaire ou toute autre autorité publique; (b) faire valoir ses droits au titre du présent Accord, y compris les vérifications relatives à de potentielles violations de celui-ci ; (c) détecter, prévenir, ou lutter contre des problèmes de fraude, de sécurité ou des problèmes d'ordre techniques (y compris notamment le filtrage des e-mails non sollicités ou "spam"), (d) répondre aux demandes d'assistance technique des utilisateurs ou (e) protéger les droits, les biens ou la sécurité de Google, de ses utilisateurs et du public. Google décline toute responsabilité quant à l'exercice ou le non exercice de ses droits dans le cadre du présent Accord.

5. Conditions générales d'utilisation de MySpace au 28 février 2008

- 9 Protéger les droits d'auteur et autres droits de propriété intellectuelle MySpace respecte la propriété intellectuelle des autres et exige que ses Utilisateurs fassent de même. Vous n'êtes pas autorisé à mettre en ligne, télécharger, intégrer, publier, envoyer par courrier électronique, transmettre ou autrement rendre disponible tout contenu qui enfreint tout droit d'auteur, brevet, marque, secret de fabrique ou autre droit de propriété de toute personne ou entité. MySpace se réserve le droit d'annuler l'Adhésion des contrefacteurs.

Si vous pensez que votre travail a été copié et publié sur ou par le biais des Services MySpace d'une manière qui constitue une atteinte à vos droits d'auteur, veuillez faire parvenir à l'Agent des Droits d'Auteur de MySpace une notification décrivant la violation revendiquée avec toutes les informations qui suivent : (a) identification de l'œuvre protégée par les droits d'auteur dont la contrefaçon est alléguée, ou, dans le cas où plusieurs œuvres protégées par les droits d'auteur sont concernées, une liste représentative de telles œuvres ; (b) identification raisonnablement suffisante de l'œuvre dont la contrefaçon est alléguée et les informations nous permettant de situer le contenu concerné sur les Services MySpace (l'indication de l'URL ou des URLs menant au contenu litigieux est suffisante) ; (c) informations raisonnablement suffisantes pour nous permettre de vous contacter, comme une adresse, un numéro de téléphone et, le cas échéant, une adresse électronique ; (d) une déclaration de votre part indiquant que vous estimez en toute bonne foi que l'usage contesté n'est pas autorisé par le titulaire des droits d'auteur, son agent ou la loi ; (e) une déclaration de votre part qui atteste, sous peine de parjure, que les informations ci-dessus figurant dans votre notification sont exactes et que vous êtes le titulaire des droits d'auteur ou autorisé à agir au nom du titulaire des droits d'auteur ; et (f) votre signature physique ou électronique. L'Agent des droits d'auteur de MySpace chargé de recevoir les notifications des infractions alléguées peut être contacté à l'adresse suivante : Copyright Agent, MySpace, Inc., 8391 Beverly Blvd., n° 349, Los Angeles, CA 90048, Etats-Unis ; Télécopie : (310) 388-0892 ; Attn : Copyright Agent. Il peut également être contacté par courrier électronique en cliquant ici :

<http://collect.myspace.com/index.cfm?fuseaction=misc.contactInput&primarySubject=2&secondarySubject=32>. MySpace fournit certains outils et technologies qui aident les titulaires de droits d'auteur à contrôler leurs

œuvres protégées par les droits d'auteur.

10. Différends entre les Membres. Vous êtes seul responsable de vos interactions avec les autres Membres de MySpace. MySpace se réserve le droit, mais n'est soumise à aucune obligation, de s'impliquer de quelque façon que ce soit dans les différends entre vous et d'autres Membres.
11. Confidentialité L'utilisation des Services MySpace est également régie par notre Politique de Confidentialité, qui est incorporée au présent Contrat par référence: <http://www.myspace.com/index.cfm?fuseaction=misc.privacy>.
12. Exclusions MySpace ne saurait être tenue responsable et ne donne aucune garantie, expresse ou implicite, concernant le Contenu publié par les Utilisateurs ou l'exactitude et la fiabilité du Contenu publié par les Utilisateurs sur ou par le biais des Services MySpace, qu'il ait été fourni par des Utilisateurs des Services de MySpace ou qu'il ait pour origine les équipements ou la programmation associée utilisés dans les Services MySpace ; un tel Contenu publié par les Utilisateurs ne reflète pas nécessairement les opinions ou la politique de MySpace. Les profils et les applications tiers créés et publiés par des Membres sur le Site MySpace peuvent contenir des liens vers d'autres sites internet. MySpace n'est pas responsable du contenu, de l'exactitude ou des opinions exprimées sur de tels sites internet, et de tels sites internet ne sont pas nécessairement examinés, contrôlés ou vérifiés en termes de précision ou d'exhaustivité par MySpace. L'inclusion de tout lien vers un site internet dans les Services MySpace ne signifie pas que MySpace approuve ou soutienne le site accessible par ce lien. Lorsque vous accédez à des sites tiers, vous le faites à vos risques et périls. MySpace ne saurait en aucun cas être tenue responsable des publicités tiers ou des applications tiers, qui sont publiées sur ou par le biais des Services MySpace, ni des biens ou services fournis par ses annonceurs. MySpace ne saurait être tenue responsable de la conduite, en ligne ou hors ligne, de tout Utilisateur des Services MySpace. MySpace ne saurait être tenue responsable de toute erreur, omission, interruption, suppression, de tout défaut, retard d'opération ou de transmission, échec de la ligne de communication, vol ou destruction ou accès non autorisé, ou encore de toute altération de la communication de tout Utilisateur ou Membre. MySpace ne saurait être tenue responsable de tout problème ou mauvais fonctionnement technique de tout réseau ou de toute ligne téléphonique, systèmes informatiques en ligne, serveurs ou fournisseurs, matériel informatique, logiciels, échec de tout courrier électronique ou lecteur en raison de problèmes techniques ou d'encombrement du trafic sur l'Internet ou sur l'un quelconque des Services MySpace ou toute combinaison des éléments ci-dessus, notamment toute blessure ou tout dommage subi par les Utilisateurs ou tout dommage causé à l'ordinateur d'une quelconque personne relatif à ou résultant de la participation ou du téléchargement de contenu en relation avec les Services MySpace. En aucun cas, MySpace ne saurait être tenue responsable de toute perte ou de tout dommage, notamment les blessures personnelles et fatales, résultant de l'utilisation des Services

MySpace, de la participation à des manifestations organisées par MySpace, de tout Contenu publié par un Utilisateur sur ou par le biais des Services MySpace, ou de la conduite de tout Utilisateur des Services MySpace, que ce soit en ligne ou hors ligne. Les Services MySpace sont fournis « TELS QUELS » et selon leur disponibilité, et MySpace écarte expressément toute garantie d'adéquation à un usage particulier ou d'absence de contrefaçon. MySpace ne peut aucunement garantir et ne promet aucun résultat spécifique résultant de l'utilisation des Services MySpace.

- i. Limitation de responsabilité EN AUCUN CAS, MYSPACE NE SAURAIT ENCOURIR DE RESPONSABILITÉ ENVERS VOUS OU TOUT TIERS EN CAS DE DOMMAGES INDIRECTS, CONSECUTIFS, EXEMPLAIRES, ACCESSOIRES, SPECIAUX OU PUNITIFS, Y COMPRIS LES DOMMAGES POUR GAIN MANQUÉ RÉSULTANT DE VOTRE UTILISATION DES SERVICES MYSPACE, MÊME SI MYSPACE AVAIT CONNAISSANCE DE LA POSSIBILITÉ DE TELS DOMMAGES. NONOBTANT TOUTE DISPOSITION CONTRAIRE DES PRÉSENTES, LA RESPONSABILITÉ DE MYSPACE ENVERS VOUS POUR QUELQUE CAUSE QUE CE SOIT ET QUELLE QUE SOIT LA FORME DE L'ACTION, SERA À TOUT MOMENT LIMITÉE AU MONTANT PAYÉ, LE CAS ÉCHÉANT, PAR VOUS À MYSPACE POUR LES SERVICES MYSPACE DURANT LA DURÉE DE L'ADHÉSION.
- ii. Contrôles des exportations des États-Unis. Les logiciels disponibles relatifs aux Services MySpace (les « Logiciels ») sont soumis au contrôle des exportations des États-Unis. Aucun Logiciel ne peut être téléchargé à partir des Services MySpace ou autrement exporté ou réexporté au risque d'enfreindre les lois américaines sur l'exportation. Vous téléchargez ou utilisez tout Logiciel à vos risques et périls.
- iii. Différends. Les présentes Conditions d'Utilisation sont régies par, et interprétées conformément aux lois de l'État de New York, sans tenir compte des règles de conflit des lois. Vous et MySpace acceptez de vous soumettre à la juridiction exclusive des tribunaux situés dans l'État de New York pour résoudre tout différend en relation avec les présentes Conditions d'Utilisation ou les Services MySpace. CHACUNE DES PARTIES AUX PRÉSENTES RENONCE À TOUT DROIT QUI POURRAIT ÊTRE LE SIEN À UN PROCÈS DEVANT JURY PAR RAPPORT À TOUT LITIGE (INCLUANT, MAIS SANS TOUTEFOIS S'Y LIMITER TOUTE DEMANDE, DEMANDE RECONVENTIONNELLE, DEMANDE ENTRE DÉFENDEURS OU DEMANDE DE TIERCE PERSONNE) RESULTANT DE OU EN RELATION AVEC LES PRÉSENTES CONDITIONS D'UTILISATION. PAR AILLEURS, CHAQUE PARTIE AUX PRÉSENTES CERTIFIE QU'AUCUN REPRÉSENTANT OU AGENT DE L'UNE OU L'AUTRE DES PARTIES N'A DECLARE, EXPRESSÉMENT OU AUTREMENT, QU'UNE TELLE PARTIE NE CHERCHERAIT PAS DANS LE CAS D'UN TEL LITIGE, À RENONCER À SON DROIT À UN PROCÈS DEVANT JURY. CHACUNE DES PARTIES RECONNAÎT QUE CETTE SECTION REPRÉSENTE UNE CONDITION ESSENTIELLE POUR L'AUTRE PARTIE CONTRACTANTE À LA CONCLUSION DU PRESENT CONTRAT.