

ETUDE BIBLIOMETRIQUE POUR L'EVALUATION DES USAGES DES LICENCES DE LOGICIELS LIBRES ET OPEN SOURCE

Rachid ALHIANE, r.alhiane@marseille-innov.org

CHERPA – IEP d'Aix-en-Provence, Université Paul Cézanne Aix-Marseille III
25 rue Gaston de Saporta - 13625 Aix-en-Provence cedex 1 (France)

Mots clefs :

Open source, logiciels libres, Source-Forge, bibliométrie, licences de logiciels libres et open source, fouille d'information, Mathéo Analyzer

Keywords:

Open source, free software, Source-Forge, bibliometrics, free/open source software licenses, data mining, Mathéo Analyzer

Palabras clave :

Código abierto, software libre, Source-Forge, bibliometría, licencias, licencias de software libres/abiertas, minería de datos, Mathéo Analyzer,

Résumé

Les études bibliométriques ont pour but d'aider à la compréhension de grands volumes de données scientifiques et techniques.

Facilement accessibles sur les plates-formes d'hébergement de projets open source¹ comme Source-Forge², de grandes masses de données semi-structurées sur les projets de logiciels libres présentent un intéressant champ pour les chercheurs en sciences de l'information et de la communication pour réaliser des études bibliométriques visant à découvrir les propriétés cachées et les tendances de ce mode de développement de logiciels.

Cet article expose les réflexions, les traitements, la méthode et les outils mis en œuvre pour faciliter l'exploitation d'un corpus bibliographique contenant les données techniques des 1000 plus actifs projets open source sur Source-Forge à la date du 10 août 2007. L'objectif de ce travail est de comprendre les usages faits des licences de logiciels libres et open source.

¹ Dans ce travail, nous assimilons logiciel libre et open source, qui s'avèrent très proches.

² Source-Forge (<http://sourceforge.net/>) est une plateforme d'hébergement et de gestion des projets de développement de logiciels libres et open source.

1 Introduction et motivations du travail

Le mouvement des logiciels libres et open source enregistre ces dernières années une crédibilité et un essor incroyables aux yeux de beaucoup d'entreprises et de chercheurs en économie et gestion. Cependant, des questions demeurent encore floues concernant son mode de fonctionnement interne. En effet, au moment où certains projets open source arrivent à fédérer une communauté de plusieurs milliers de contributeurs volontaires et bénéficient d'une abondance de ressources en termes de capacité de développement, on constate sur les plates-formes d'hébergement de projets open source que beaucoup de ces projets restent inactifs, faute d'arriver à attirer les contributeurs volontaires. Se pose donc la question de savoir pourquoi certains projets open source sont plus attractifs que d'autres pour les contributeurs volontaires ? et quel rôle jouent les licences libres et open source dans cette attractivité ?

Pour répondre à ces questions, nous avons procédé à une étude bibliométrique sur les 1000 plus actifs projets open source sur Source-Forge. Cette activité est mesurée à partir d'un certain nombre d'indicateurs comme le nombre de bug rapportés / résolus, le nombre de téléchargements, le taux de fréquentation des forums de discussion, l'activité des mailing-lists et le nombre d'inscrits. Nous avons préféré cette approche à l'analyse de la totalité des projets de Source-Forge, car beaucoup d'entre eux sont inactifs.

Concernant les plates-formes d'hébergement des projets open source (Free Software Foundation³, Freshmeat⁴, Objectweb⁵, Rubyforge⁶, Source-Forge, SourceKibitzer⁷...), le choix de Source-Forge s'explique par le fait que cette plate-forme est de loin la plus connue et la plus utilisée par les créateurs et les développeurs de logiciels libres et open source. Le tableau suivant présente la liste des plates-formes open source les plus utilisées par les créateurs de projets open source:

Tableau 1 : Les plates-formes d'hébergement des projets open source (données collectées le 19/09/2008)

Nom Forge	URL	Nombre de projets hébergés	Nombre d'utilisateurs
Source-Forge	http://sourceforge.net/	180,000	1.900,000
Free Software Fondation	http://directory.fsf.org	5,000	NC
Freshmeat	http://www.freshmeat.net	45,010	406,255
Rubyforge	http://rubyforge.org	6,542	31,164
Object web	http://forge.objectweb.org	151	11,197

Le 19 septembre 2008, nous avons recensé sur Source-Forge plus de 180,000 projets open source et plus de 1.900,000 utilisateurs enregistrés. Chacun de ces projets est identifié par plusieurs champs comme le nom du projet (*Project Unix Name*), la licence utilisée (*License*), le(s) système(s) d'exploitation compatible(s)

³ Free Software Foundation <http://directory.fsf.org>

⁴ Freshmeat <http://www.freshmeat.net>

⁵ Objectweb <http://forge.objectweb.org>

⁶ Rubyforge <http://rubyforge.org>

⁷ SourceKibitzer <http://www.sourcekibitzer.org>

(*Operating System*), le nombre de développeurs qui participent au projet (*Developers*), etc. (plus de détails ci-dessous). Chacun de ces champs est susceptible de faire l'objet d'une analyse quantitative mais pour notre étude, nous focalisons notre attention sur le champ « *License* » afin d'étudier les usages faits des licences libres et open source. Ces dernières sont fortement dominées par la présence de la licence GPL (Général Public License) créée dans les années 80 par Richard Stallman, chercheur au M.I.T. (Massachusetts Institut of Technology) et fondateur de la Free Software Foundation⁸. Cette licence représente la majorité des licences libres actuellement utilisées.

Néanmoins, derrière la suprématie de la GPL, d'autres licences coexistent. Leur liste est longue⁹, mais on pourra citer comme exemple la LGPL (Lesser General Public License), une version dérivée de la GPL et moins contraignante que celle-ci, la MPL (Mozilla Public License) ou des licences plus anciennes comme la BSD (Berkeley Software Distribution) et ses dérivées.

Concernant leurs auteurs, les licences libres et open source proviennent d'associations, d'universités et d'entreprises. En effet, on constate que lorsqu'une entreprise développe un logiciel libre, il est fréquent de voir apparaître une licence attachée à son projet ou à elle-même. On peut d'ailleurs remarquer que ces licences comportent souvent le nom de l'entreprise dans leur propre nom. Quant aux universités, elles ont tendance à choisir une seule licence qui sera utilisée pour tous les projets développés.

D'une manière plus générale, on remarque que l'on peut d'ailleurs presque systématiquement associer une licence spécifique à un projet particulier, comme Perl avec la licence Artistic ou BSD avec le système Unix de l'université de Berkeley.

2 Contexte de la recherche

Au début des années 80, les grands éditeurs de logiciels, tel le géant Microsoft, ont vu naître sur leur marché un phénomène nouveau : le « mouvement libre ». Poussé par des associations comme la Free Software Foundation (FSF) ou l'Open Source Initiative¹⁰ (OSI), il tente de promouvoir un nouveau mode de production mais également de distribution des logiciels et se fonde sur une utilisation différente du droit de propriété intellectuelle.

Ce mouvement trouve ses origines dans l'évolution qu'avait connue l'industrie informatique à partir des années 70. En effet, la mutation marquante durant cette période consiste en un phénomène d'ouverture faisant de l'industrie informatique une activité produisant des systèmes compatibles et concurrents. Cette mutation s'est accentuée dans les années 80 avec l'avènement du micro-ordinateur, la démocratisation de l'informatique et la domination de l'industrie par le couple Intel-Microsoft.

La montée en puissance des logiciels libres à amener de nombreux éditeurs de logiciels propriétaires à prendre en considération le danger que représente ce mouvement sur leurs activités. Dans le même temps, les logiciels libres et open source offrent des opportunités énormes aux éditeurs de logiciels de pouvoir toucher un large marché constitué essentiellement de PME et TPE dont le niveau d'équipement en TIC est encore très faible dans de nombreux pays.

⁸ Free Software Foundation <http://www.fsf.org>

⁹ Nous verrons en détails ces différents types de licences.

¹⁰ Open Source Initiative <http://www.opensource.org>

Aujourd'hui, les logiciels libres s'imposent comme un modèle nouveau et original dans le monde de l'édition de logiciels, modèle avec lequel il faut désormais compter et qui, en réduisant la domination d'acteurs jusque-là, en situation de quasi-monopole et en imposant un nouveau mode de production de logiciels, a modifié en profondeur le paysage informatique mondial.

Les expressions open source et logiciel libre trouvent leurs racines dans la définition donnée par M. Richard Stallman. Elle met en avant quatre libertés fondamentales pour toute personne ayant acquis un logiciel :

- Celle d'exécuter un programme pour tous les usages.
- Celle d'étudier le fonctionnement du programme pour l'adapter à ses besoins. Pour cela, l'accès au code source est indispensable.
- Celle de redistribuer des copies pour aider d'autres utilisateurs.
- Celle d'améliorer le programme pour aider d'autres utilisateurs en publiant ces améliorations. Pour cela, l'accès au code source est indispensable.

Ces logiciels sont régis par une catégorie de licences dites libres ou open source. Le trait distinctif de ces licences, pierres angulaires du nouveau modèle, est qu'elles donnent naissance à des contrats stipulant que tout utilisateur a le droit d'utiliser, modifier et distribuer le logiciel, dès lors qu'il préserve son caractère ouvert en laissant son code source disponible et transmet ces mêmes droits à tout utilisateur. Ainsi définis, les logiciels open source s'opposent aux logiciels propriétaires dont les licences, elles, n'accordent à l'utilisateur qu'un simple droit d'utilisation, soumis à une condition de redevance.

3 Procédure d'exploitation des données de Source-Forge

3.1 Etape 1

La première étape dans notre étude consiste à s'assurer que les données définissant les projets à analyser sont structurées de la même manière (voir la FIG. 1). Une analyse manuelle de quelques projets sur Source-Forge permet d'identifier les données que nous souhaitons aspirer automatiquement via un logiciel développé sur mesure :

Tableau 2 : Description des champs définissant un projet open source sur Source-Forge

Registered	Date d'enregistrement du projet sur Source-Forge.
Developers	Les personnes qui participent au développement d'un projet open source. L'administrateur du projet a la possibilité d'attribuer des droits limités ou larges aux participants du projet : développeurs, testeurs, débogueurs...
Project Unix name	Le nom donné par l'administrateur à son projet.
Summary of the project	Description du projet.
License	Le(s) licence(s) choisie(s) pour le projet en question.
Intended audience	Le public cible du projet : Advanced End Users, Developers, Telecommunications, Financial and Insurance Industry, etc.
Programming language	Le langage de programmation utilisé pour écrire le code source d'un projet open source : C++, C, Java, Perl, etc.
Topic	Cartographie des domaines dans lesquels un projet open source peut être classé par son administrateur : Development software, System, Multimedia, Internet, Office/Business, etc.
Operating system	Le(s) système(s) d'exploitation avec le(s)quel(s) un projet open source fonctionne : Linux, Windows, Solaris, etc.
Development status	Statut de développement d'un projet open source : 1-Planning, 2-Pre-Alpha, 3-Alpha, 4-Beta, 5-Production/Stable, 6-Mature, 7-Inactive.
User interface	Détermine l'outil logiciel utilisé pour la réalisation de l'interface graphique d'un projet open source.

Il existe d'autres champs définissant un projet open source sur Source-Forge comme par exemple, *Database Environment*, *Project Admins*, *Activity Percentile*, *Public Forums*, *Mailing Lists*, *Bugs*, etc.

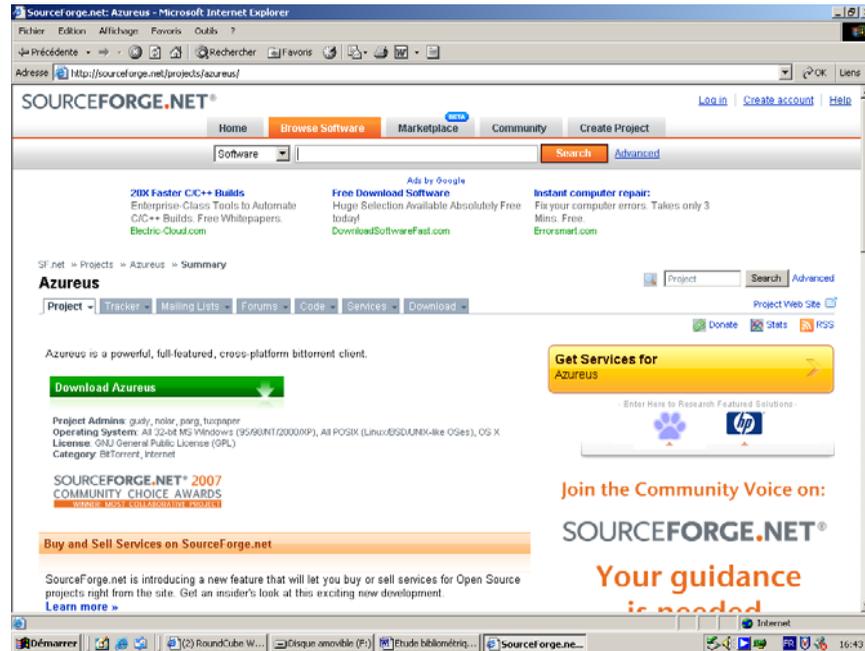


Figure 1 : Page d'accueil d'un projet de logiciel libre et open source hébergé sur Source-Forge

3.2 Etape 2

La deuxième étape dans notre étude, consiste à identifier la liste des projets à aspirer sur Source-Forge. Le lien suivant <http://sourceforge.net/top/mostactive.php> présente la liste des projets open source les plus actifs sur Source-Forge. Pour notre étude, nous avons sélectionné les 1000 premiers.

3.3 Etape 3

À l'aide d'un script développé sous le langage Perl, nous avons procédé dans cette troisième étape à l'extraction automatique des données techniques définissant les 1000 projets open sources sélectionnés. La fin de cette opération a permis de disposer d'un corpus bibliographique prêt à être exploité.

Rappelons que pour notre étude sur les usages des licences libres et open source, nous focalisons notre analyse sur le champ « *License* ».

3.4 Etape 4

La dernière étape dans ce processus consistait à charger le corpus bibliographique obtenu dans l'étape 3 dans le logiciel Mathéo-Analyser afin de procéder à son analyse. Ce logiciel développé par l'entreprise Mathéo Software¹¹, permet d'analyser des corpus bibliographiques de brevets et d'articles scientifiques (voir figure 2).

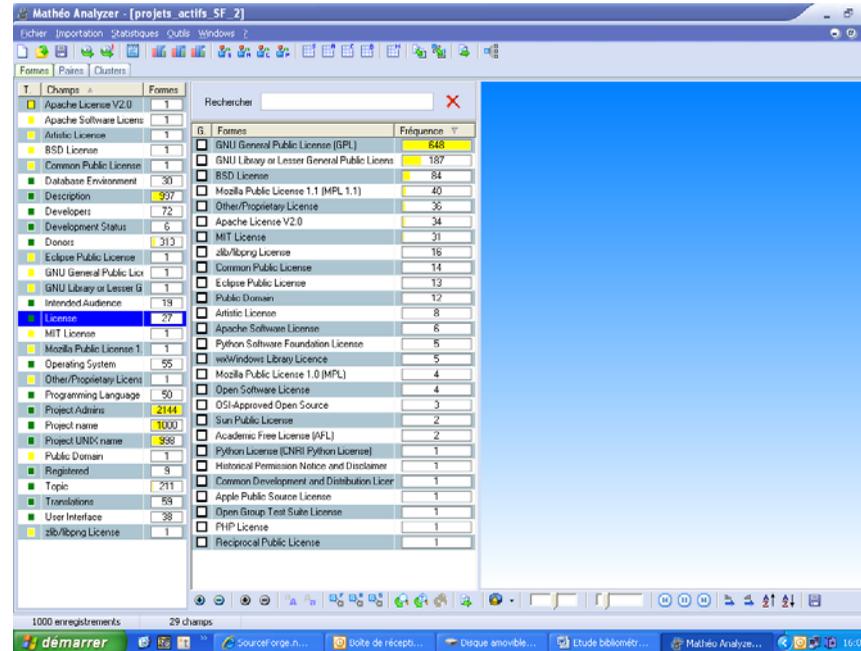


Figure 2 : Copie d'écran de Matheo-Analyser

4 Résultats de l'analyse

Les statistiques présentées ici sur les projets (répartition des licences, répartition thématique des licences), concernent seulement les licences les plus représentatives des 1000 projets open source analysés.

Source- Forge répartit ses projets dans différents thèmes et sous-thèmes. Pour des raisons de clarté, les sous-thèmes de Source-Forge ont été regroupés dans cette étude en thèmes plus généraux

¹¹ Mathéo Software <http://www.matheo-software.com/>

4.1 Provenance géographique des licences

La provenance géographique des licences a été obtenue en recherchant le lieu du créateur de la licence. L'encyclopédie libre «Wikipedia»¹² a été d'une grande utilité dans ce travail. Cependant, quelques difficultés sont à évoquer :

- Certaines licences proviennent d'une collaboration impliquant des équipes de différentes nationalités et zones géographiques ce qui pose problème au niveau de l'affectation de la licence en question à une zone géographique particulière,
- Dans certaines licences, il est difficile de trouver les informations sur l'auteur de la licence car très souvent, seul le texte de la licence et éventuellement sa version sont disponibles,
- Certains projets analysés utilisent des licences imprécises du type « *Other/proprietary license* » ou « *OSI approved open source* »,
- Certains projets utilisent des licences du type « *Public Domain* » ce qui signifie que le projet est tombé dans le domaine public,

Sur les 1000 projets open source analysés, 27 licences ont été identifiées. La grande majorité provient des Etats-Unis, avec près de 87% des licences. Le reste vient de l'Europe du nord, avec près de 13%. Une seule licence vient d'une collaboration impliquant une équipe américaine et française. Certaines licences ne sont pas comptabilisées dans ces statistiques car, comme l'on a déjà évoqué ci-dessus, il est difficile de trouver la vraie origine de ces licences.

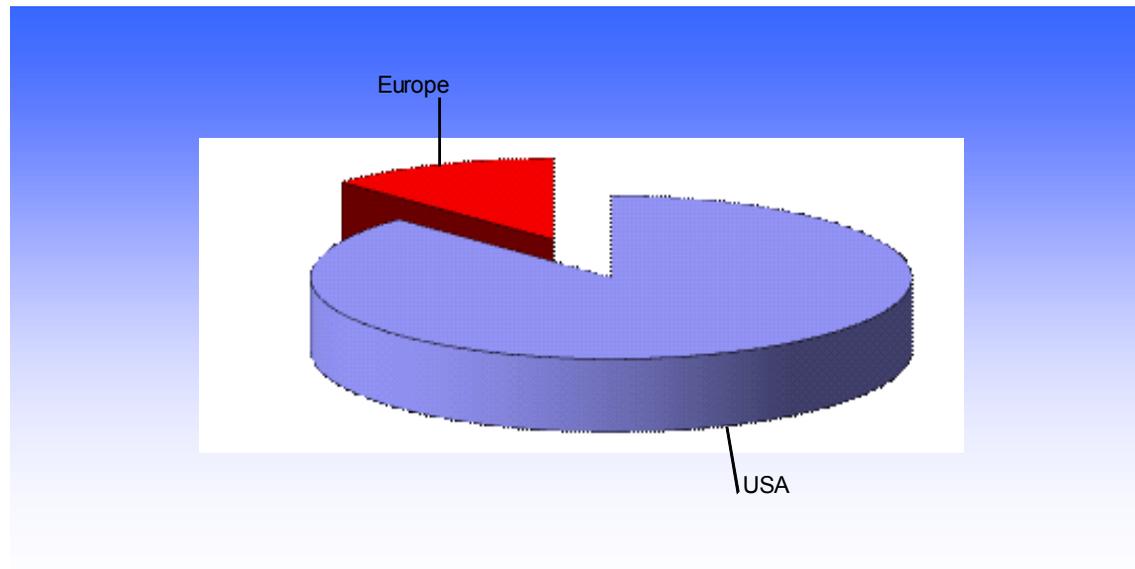


Figure 3 : Répartition géographique des Licences.

¹² Wikipedia <http://www.wikipedia.org>

4.2 Les auteurs des licences libres et open source

Il s'agit ici de classer les auteurs des licences étudiées dans les catégories suivantes :

- **Université** : les universités, et en général leurs laboratoires de recherche, ont été à l'origine de plusieurs projets de logiciels libres et open source. Ceci a eu pour conséquence la création d'un bon nombre de licences pour diffuser ces logiciels.
- **Entreprise** : cette catégorie regroupe les entreprises qui ont choisi une politique d'ouverture en libérant le code source de certains de leurs logiciels propriétaires ou en créant des projets de logiciels libres. En conséquence, ces entreprises ont créé des licences particulières pour diffuser ces logiciels.
- **Association** : cette catégorie regroupe les individus, les fondations ou les associations créant des licences libres, comme la FSF, LATEX ou des associations d'entreprises.

Une fois encore, les 27 licences ont été étudiées pour obtenir ces chiffres : Les licences libres des 1000 projets proviennent des associations (18 licences, soit 67%). Ensuite, on trouve les entreprises avec 7 licences (26%) et pour finir les universités avec 2 licences (7%)

L'explication de cette répartition est assez simple à comprendre. En effet, les associations ont tendance à créer des licences assez génériques et très utilisées (comme la GPL créée par la FSF), alors que les entreprises ont tendance à créer une licence par type de projets. On trouvera souvent une ou plusieurs licences associées à une entreprise. A contrario, chaque université crée sa propre licence, et tous les projets de l'université utiliseront la même licence.

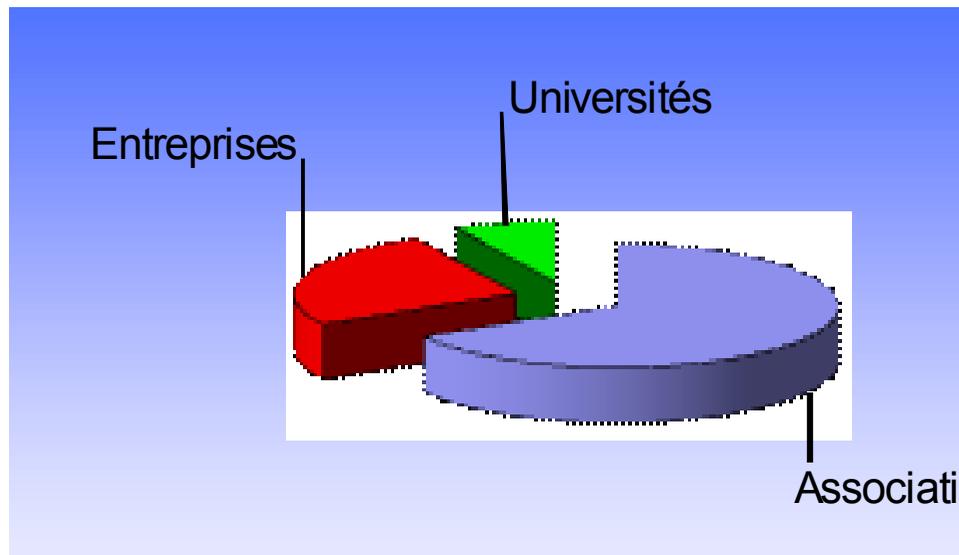


Figure 4 : Répartition des origines des licences

4.3 Répartition des licences

La licence GPL est de loin la plus représentée, puisque sur les 1000 projets open source étudiés on trouve près de 648 projets utilisant la GPL, soit plus de 64%. La licence LGPL, arrive en deuxième position avec plus de 18 % des projets, soit près de 187 projets. Puis en troisième position, on trouve la licence de type BSD, qui représente environ 8% des projets, soit 84 projets.

On trouve ensuite, la MPL, la licence du projet Mozilla avec une représentation de 4% suivit respectivement par other/proprietary license, la licence Apache et la licence MIT avec environ 3% chacune.

Comme on le voit, la répartition est assez inégale, les licences de la FSF (GPL et LGPL) représentant plus de 83% des projets étudiés.

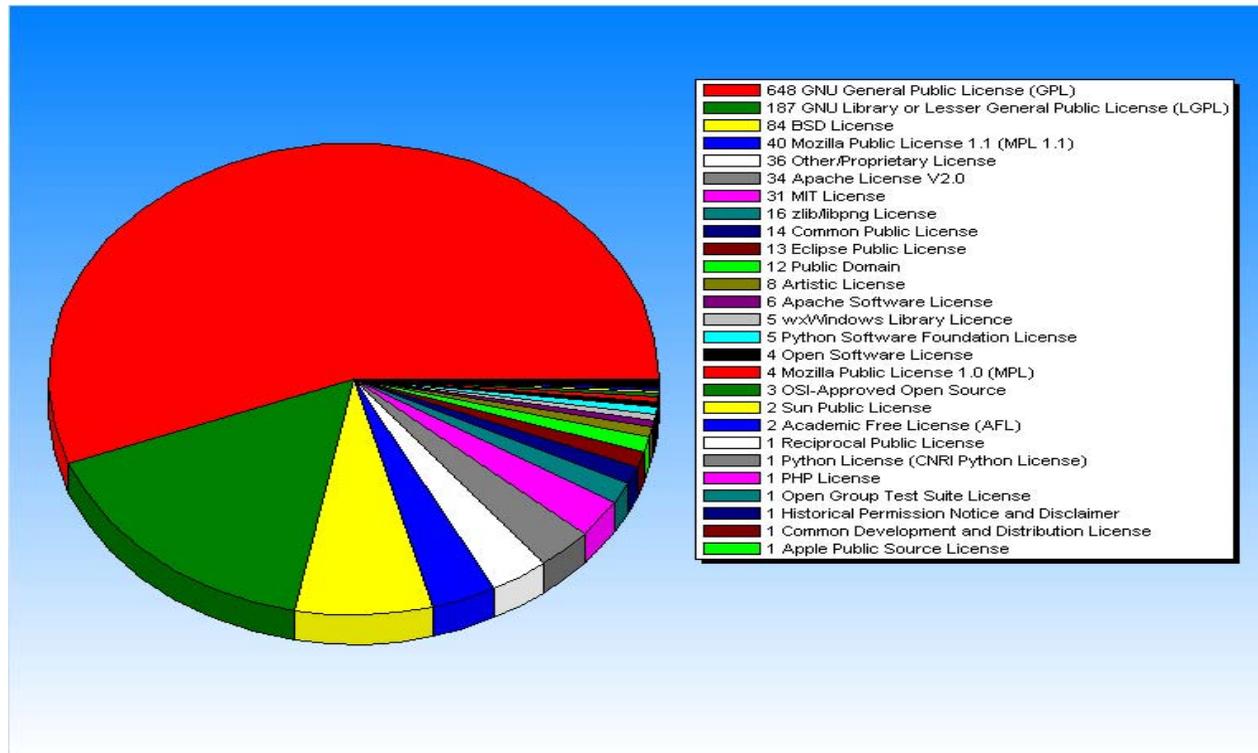


Figure 5 : Répartition des projets par licences

4.4 Répartition thématique des licences

Pour les 27 licences identifiées, on cherche ici à savoir dans quel(s) domaine(s) ces licences sont utilisées, mais également, pour chaque licence, quelles étaient ses thématiques dominantes.

4.4.1 Répartition générale

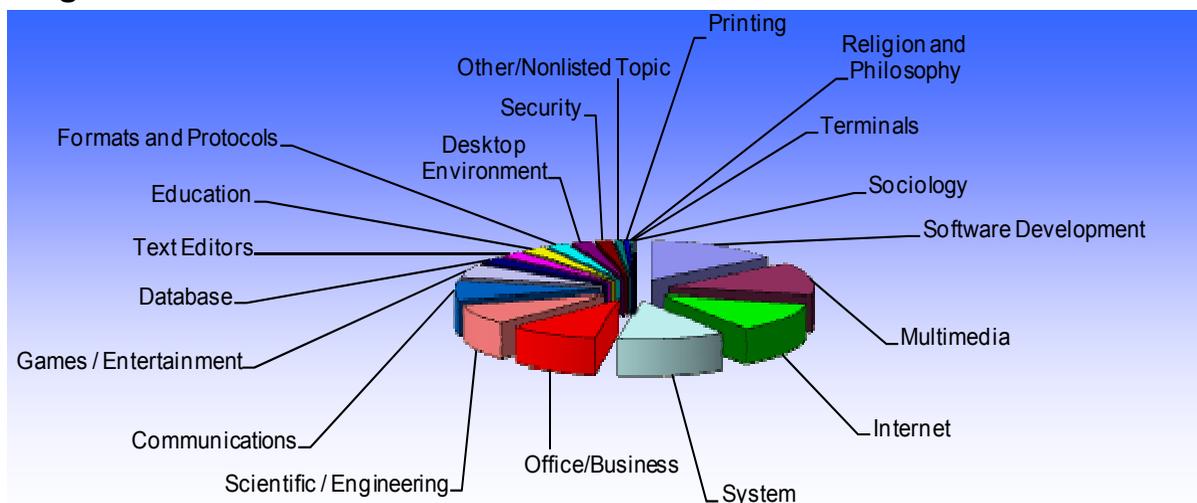


Figure 6 : Répartition générale des projets par thèmes

TAB. 2 - Répartition générale des projets en pourcentage

Software Development	15 %	Database	3 %
Multimedia	13 %	Text Editors	3 %
Internet	12 %	Education	3 %
System	12 %	Formats and Protocols	2 %
Office/Business	10 %	Desktop Environment	2 %
Scientific/Engineering	9 %	Security	2 %
Communications	7 %	Other/Non listed Topic	1 %
Games/Entertainment	5 %	Printing	1 %

La répartition générale des projets permet de constater que 15% des projets libres étudiés sont liés au développement, 13% au multimédia, 12% chacun pour Internet et système et 10% à la bureautique.

L'importante représentation des projets liés au développement est significative de l'utilisation habituelle de systèmes libres comme GNU/Linux.

La présence des projets liés à la multimédia en 2^{ème} position montre la bonne implantation des projets libres dans les applications liées au son et à l'image. Ces applications se sont accentuées avec l'explosion de l'Internet.

La présence des projets liés à Internet et au système montre également la bonne implantation des projets libres dans les applications d'infrastructure.

La présence de projets liés à la bureautique est plutôt récente. En effet, on a longtemps reproché l'absence de projets de bureautiques libres.

4.4.2 La GPL

La Licence GPL (General Public License), rédigée par Free Software Foundation (FSF), est une licence qui fixe les conditions légales de distribution des logiciels libres du projet GNU. L'objectif de cette licence est de garantir à l'utilisateur les droits suivants (appelés libertés) sur un programme informatique :

- La liberté d'exécuter le logiciel, pour n'importe quel usage ;
- La liberté d'étudier le fonctionnement d'un programme et de l'adapter à ses besoins, ce qui passe par l'accès aux codes sources ;
- La liberté de redistribuer des copies ;
- La liberté d'améliorer le programme et de rendre publiques les modifications afin que l'ensemble de la communauté en bénéficie.

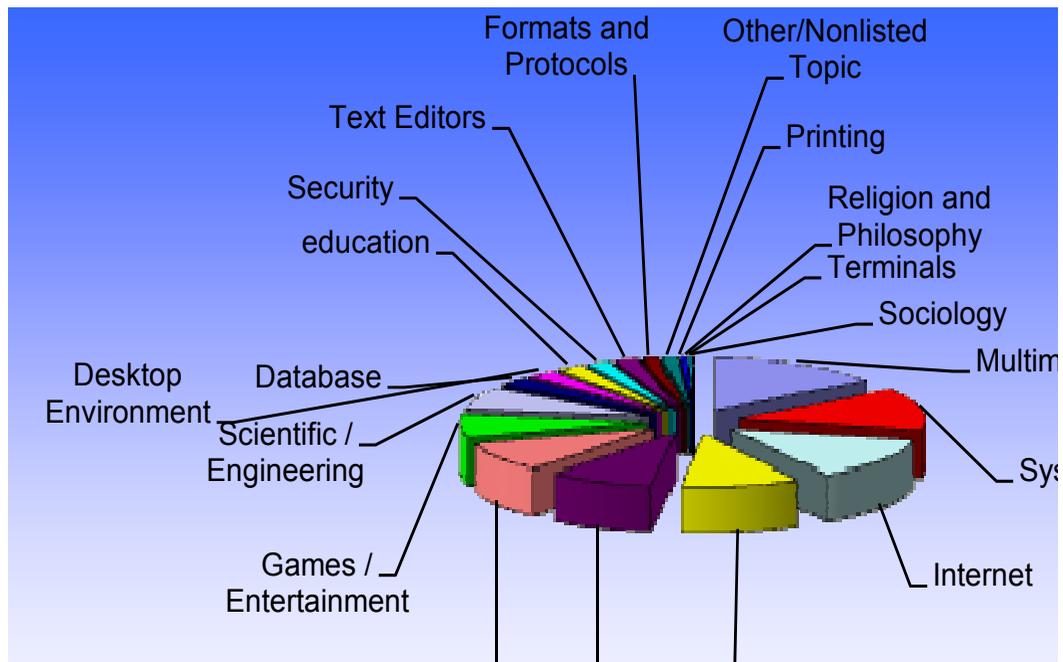


Figure 7 : Répartition de la GPL par thèmes

Tableau 3 : Répartition de la GPL en pourcentage

Multimedia	15 %	Database	3 %
System	14 %	Desktop Environment	2 %
Internet	13 %	Education	2 %
Office/Business	10 %	Security	2 %
Software Development	9 %	Text Editors	2 %
Communications	9 %	Formats and Protocols	1 %
Games/Entertainment	7 %	Other/Nonlisted Topic	1 %
Scientific/Engineering	7 %		

À l'exception des projets liés au développement, on retrouve ici une répartition assez similaire à la répartition générale : la licence GPL est présente en effet dans plus de 64% des projets étudiés.

4.4.3 LGPL

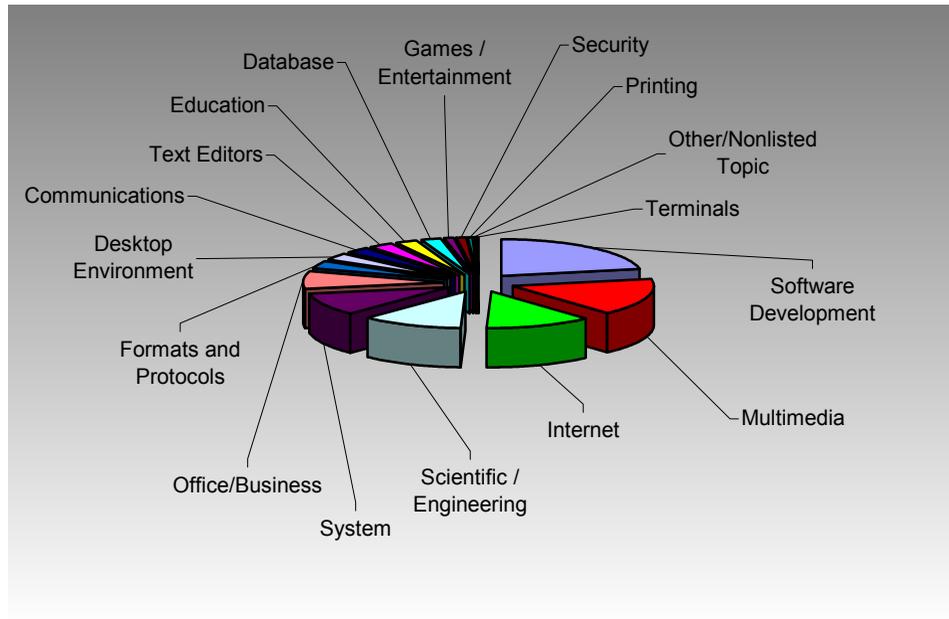


Figure 8 : Répartition de la LGPL par thèmes

Tableau 4 : Répartition de la LGPL en pourcentage

Software Development	22 %	Communications	3 %
Multimedia	16 %	Text Editors	3 %
Internet	12 %	Education	3 %
Scientific/Engineering	12 %	Database	2 %
System	10 %	Games/Entertainment	1 %
Office/Business	8 %	Security	1 %
Formats and Protocols	3 %	Printing	1 %
Desktop Environment	3 %		

La licence LGPL a été écrite pour les bibliothèques logicielles de programmation afin de permettre leur utilisation avec des logiciels dont la licence n'est pas compatible avec la GPL. En effet, la LGPL est jugée moins contraignante que la GPL car elle autorise l'incorporation du programme libre à des programmes « propriétaires ». C'est la raison pour laquelle on constate une forte implication des projets dans les activités de développement (22%).

4.4.4 BSD

La licence BSD (Berkeley software distribution License) est une licence libre utilisée pour la distribution de logiciels de l'université Berkeley. Elle permet de réutiliser tout ou partie du logiciel sans restriction, qu'il soit intégré dans un logiciel libre ou propriétaire.

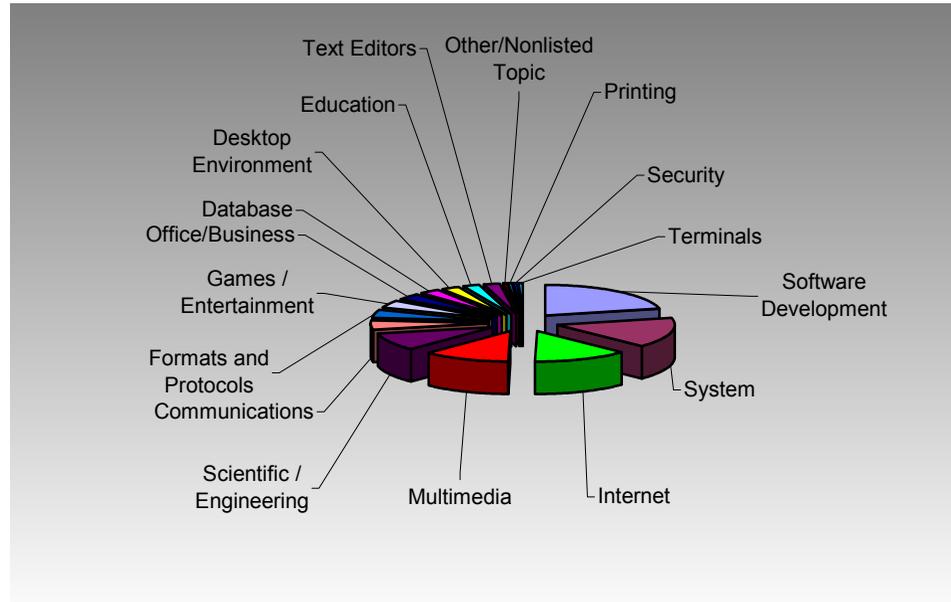


Figure 9 : Répartition de la BSD par thèmes

Tableau 5 : Répartition de la BDS en pourcentage

Software Development	21 %	Database	3 %
System	16 %	Desktop Environment	3 %
Internet	13 %	Education	2 %
Multimedia	12 %	Text Editors	2 %
Scientific/Engineering	9 %	Other/Nonlisted Topic	1 %
Communication	5 %	Printing	1 %
Formats and Protocols	4 %	Security	1 %
Games/Entertainment	4 %	Terminals	1 %
Office/Business	3 %		

La licence BSD a une importante représentation dans les projets liés au développement. Cela peut être expliqué par la souplesse de cette licence : elle autorise la réutilisation de tout ou partie du logiciel sans restriction, qu'il soit intégré dans un logiciel libre ou propriétaire. Viennent ensuite les thèmes liés au système et à Internet ce qui montre une importante implication des projets utilisant la licence BSD dans les applications d'infrastructure.

4.4.5 MPL 1.1

La MPL ou Mozilla Public License en anglais (Licence publique de Mozilla) est une licence open source créée par Netscape lors de la libération du code source de ce qui devait devenir Netscape Communicator 5 en mars 1998. Ce dernier formera la base du projet Mozilla.

Les limitations de cette licence se sont assez vite apparues, et une version 1.1 est créée. Cependant, certains projets open source continuent d'être publiés sous la l'ancienne version (MPL 1.0) et non sous la version 1.1 : dans notre étude, 4 projets open source sont publiés avec la version 1.0 de la MPL (voir Fig 5 ci-dessus)

Le projet Mozilla produit et propose le célèbre navigateur Web Firefox et le client de messagerie Thunderbird. Il offre aussi d'autres produits et technologies, comme Bugzilla d'où une importante représentation des projets liés à Internet (23%)

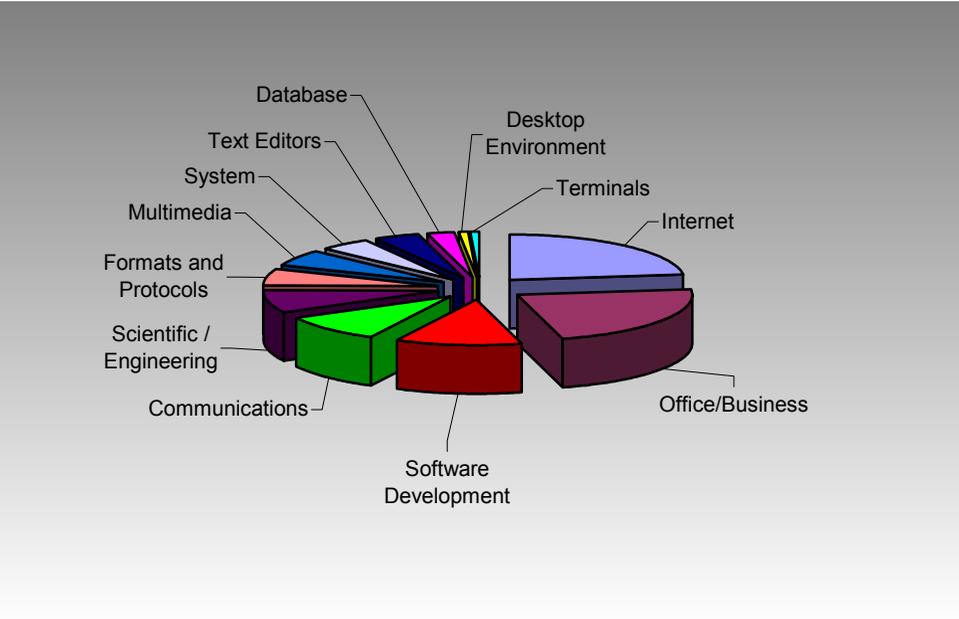


Figure 10 : Répartition de la MPL 1.1 par thèmes

Tableau 6 : Répartition de la MPL 1.1 par thèmes en pourcentage

Internet	23 %	Formats and Protocols	6 %
Office/Business	23 %	System	5 %
Software Development	11 %	Text Editors	4 %
Communications	10 %	Database	2 %
Scientific/Engineering	8 %	Desktop Environment	1 %
Multimedia	6 %	Terminals	1 %

4.4.6 Apache V2.0

La licence Apache est une licence de logiciel libre et open source. Elle est écrite par l'Apache Software Foundation pour le logiciel *Apache HTTP Server*. Ce dernier est le logiciel libre le plus populaire du Web. C'est donc logiquement normal qu'on trouve que les projets l'utilisant concernent les activités de développement et d'Internet.

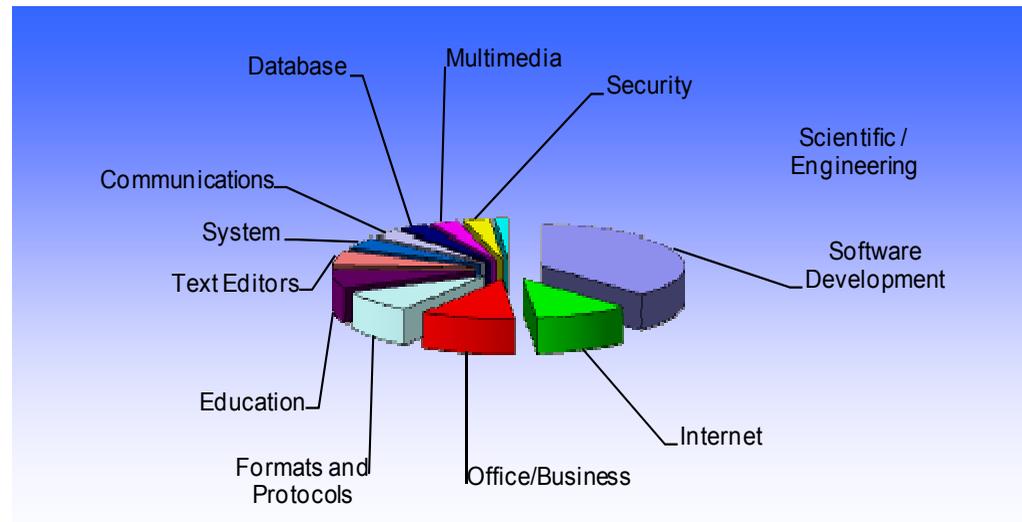


Figure 12 : Répartition de la licence Apache par thèmes

Tableau 6 : Répartition de la licence Apache en pourcentage

Software Development	37 %	System	5 %
Internet	10 %	Communications	3 %
Office/Business	10 %	Database	3 %
Formats and Protocols	9 %	Multimedia	3 %
Education	8 %	Security	3 %
Text Editors	6 %	Scientific/Engineering	2 %

4.4.7 MIT License

La Licence MIT, aussi appelée plus précisément la licence X11¹³, est une licence de logiciel du Massachusetts Institute of Technology (MIT) pour la diffusion de logiciels, en particulier le X Window System. Cette licence de logiciel libre et open source donne à toute personne recevant le logiciel le droit illimité de l'utiliser, le copier, le modifier, le fusionner, le publier, le distribuer, le vendre et de changer sa licence. La seule obligation est de mettre le nom des auteurs avec la notice de copyright.

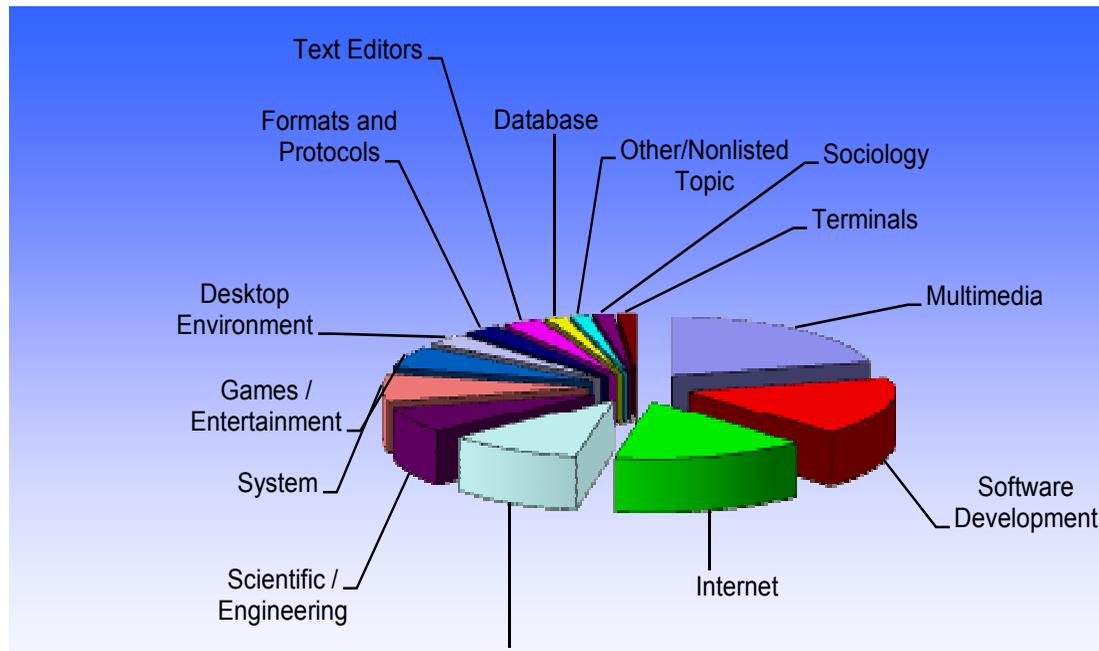


Figure 13 : Répartition de MIT License par thèmes

¹³ X11 est un système qui gère l'interaction homme-machine par l'écran, la souris et le clavier de certains ordinateurs en réseau

Tableau 7 : Répartition de MIT License en pourcentage

Multimedia	20 %	Desktop Environment	3 %
Software Development	16 %	Formats and protocols	3 %
Internet	14 %	Text Editors	3 %
Office/Business	11 %	Database	2 %
Scientific/Engineering	8 %	Other/Nonlisted Topic	2 %
System	8 %	Sociology	2 %
Games/Entertainment	6 %	Terminals	2 %

La licence MIT a une forte représentation dans les projets liés à la multimédia avec près de 20% des projets.

4.4.8 Les autres licences

- Zlib/libpng License : la licence zlib est une licence libre qui détermine comment peuvent être distribuées les bibliothèques zlib et libpng. Les grandes thématiques rencontrées concernent le multimédia (27%) et le développement (19%).
- Common Public License (CPL) : la CPL est une licence Open Source d'IBM. Elle autorise la modification du code source et garantit le respect des brevets logiciels. Pour cette raison, la CPL n'est pas compatible Licence publique générale GNU (GPL). Cette licence a une forte représentation dans les projets liés au développement avec près de 40% des projets.
- Eclipse Public License (EPL) : l'EPL est une licence libre utilisée à l'origine par le logiciel Eclipse d'IBM. Eclipse IDE est un environnement de développement intégré libre (le terme Eclipse désigne également le projet correspondant, lancé par IBM) extensible, universel et polyvalent, permettant potentiellement de créer des projets de développement mettant en œuvre n'importe quel langage de programmation. Eclipse IDE est principalement écrit en Java (à l'aide de la bibliothèque graphique SWT, d'IBM), et ce langage, grâce à des bibliothèques spécifiques, est également utilisé pour écrire des extensions. C'est donc logiquement que les activités qui y sont liées concernent principalement le développement avec près de 48% des projets.

5 Conclusion

Dans cet article, nous avons présenté les méthodes et les outils utilisés pour réaliser une étude bibliométrique sur les usages des licences de logiciels libres et open source. Cette étude a porté sur les 1000 plus actifs projets open source de Source-Forge.

Sur les 27 licences libres et open source identifiées, nous avons procédé à une répartition thématique de quelques unes (les plus importantes). Cela nous a permis de connaître les domaines dont lesquels ces licences sont utilisées, ainsi que pour chaque licence, sa thématique dominante. Résultat : on constate que pour les licences GPL, MIT License, Zlib/libpng License, la thématique dominante est le « **Multimedia** ». Pour les licences LGPL, BSD, Apache V2.0, CPL, la thématique dominante est le « **Software Development** ». Enfin, pour la licence MPL 1.1, sa thématique dominante est « **Internet** ».

Cela montre que si l'attractivité d'un projet open source dépend de plusieurs paramètres comme, par exemple, le statut et la notoriété du mainteneur du projet, le nombre de contributeurs participant déjà au projet, le stade de développement, la capacité du projet à répondre aux besoins des utilisateurs, etc. il ne faut pas négliger le rôle que joue le choix de la licence libre et open source dans cette attractivité. Les résultats ci-dessus indiquent que certaines licences open source s'avèrent plus adaptées que d'autres pour certains domaines d'activité. Il en découle que la pertinence du choix de la licence permettra d'assurer au logiciel en développement une forte attractivité de contributeurs volontaires et une large diffusion auprès des utilisateurs.

En prenant certaines précautions, une étude bibliométrique similaire sur l'ensemble des projets de Source-Forge, permettra de fournir des résultats beaucoup plus intéressants sur les licences libres et leurs thématiques dominantes. Une telle étude pourrait, par exemple, permettre à des porteurs de projets de développement de logiciels libres de disposer d'informations pertinentes sur la /les licence(s) le(s) plus adaptée(s) à leur projet.

6 Bibliographie

- [1] **KLINCEWICZ K.**, *Innovativeness of Open Source Software Projects*. School of Innovation Management, Tokyo Institute of Technology, 2005
- [2] **SOUFRON J.B., SALLANTIN J.**, *The Evolution of Free/Libre and Open Source software Licenses: A Dynamic Model*. CERSA, 2005
- [3] **WEISS D.**, *Quantitative Analysis of Open Source Projects on Source-Forge*. Poznan University of Technology, Poland, 2005