

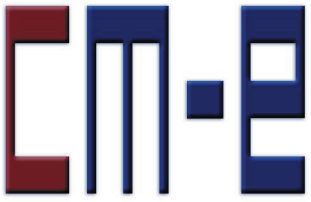
CLUSTERING
Meta-ENGINE

Julien AGULLO

Souâad AIDH

Ibrahima OUANE

Lamyae SELLAKE



Sommaire

I. Installation

1. Préambule
2. Installation en local
3. Installation sur un serveur distant
4. Net Draw

II. Contexte

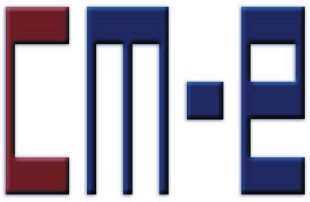
1. L'équipe projet
2. Présentation du projet
3. Définitions
4. Objectif du projet
5. Les moteurs de recherche

III. Principe d'utilisation

1. La recherche
2. La comparaison des résultats
3. Export sur Excel
4. Export sur Net Draw

IV. Conclusion

1. Contraintes
2. Perspectives d'avenir
3. Le rapport scientifique
4. Conclusion



Installation

1. Préambule

Le moteur de recherche **Clustering Meta-Engine** est un site Web dynamique. Vous pouvez soit l'exécuter directement en local, soit le placer sur un serveur Web afin de l'utiliser à partir de n'importe quel ordinateur relié à Internet.

Tout d'abords, décompressez le fichier **CME_IET_2007.zip** sur votre ordinateur. Le dossier qui apparaît se nomme « **Clustering Meta-Engine** ».

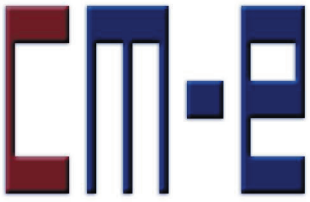
A l'intérieur, vous trouverez :

- un dossier « **CME** » contenant des fichiers PHP et des images qui composent le moteur de recherche.
- un dossier « **Logiciel** » contenant les logiciels à installer pour faire fonctionner le site.

Le code source se trouvant dans les fichiers PHP ne peut pas être vendu, échangé et diffusé. Cependant, vous pouvez le modifier, le corriger et l'adapter à vos besoins sans qu'il vous soit nécessaire de demander une autorisation.

Pour un bon fonctionnement du méta moteur, ne modifiez pas les fichiers PHP si vous n'avez pas de bonnes connaissances en programmation.

Un site Internet en PHP ne peut pas être lu sur un ordinateur qui ne dispose pas d'un serveur web installé. Pour utiliser le méta moteur CME, vous pouvez procéder de deux façons : soit installer un serveur PHP sur votre système, soit placer le site chez un hébergeur utilisant le PHP.



2. Installation en local

Si vous souhaitez utiliser le moteur de recherche CME directement sur votre système car vous ne possédez pas d'hébergeur, vous devez installer le logiciel Easy PHP. Le programme se trouve dans le dossier « **Logiciel** » mais vous pouvez aussi le télécharger sur le site Internet <http://www.easyphp.org/>.

Une fois le logiciel installé et avant son premier lancement, vous devez faire une modification dans ses fichiers d'initialisation pour que le méta moteur puisse fonctionner correctement. En effet, le script exécuté lors de la recherche prend un certain temps et si aucune modification n'est faite, il se coupera automatiquement au bout de 30 secondes.

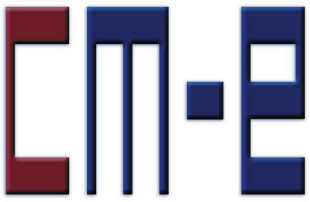
Pour remédier à cela, vous devez ouvrir le fichier php.ini se trouvant dans le dossier Apache du logiciel installé (C:\Program Files\EasyPHP1-8\apache). Une fois le fichier ouvert avec le bloc note, recherchez la ligne contenant « **max_execution_time** » et modifiez la valeur par défaut à 30 par 3600. Modifiez également la ligne « **max_input_time** » juste en dessous par 3600 (au lieu de 60) ainsi que la ligne « **memory_limit** » par 32M (au lieu de 8M).

Après avoir réalisé cette modification, vous pouvez lancer le logiciel. Réduisez la fenêtre qui apparaît et faites un clic droit sur l'icône **Easy PHP** qui s'est ouvert dans la barre des tâches de Windows (à côté de l'heure), puis cliquez sur **Administration**. Ceci ouvrira la page d'accueil du serveur local sur votre navigateur par défaut. Pour faire fonctionner le moteur de recherche, il suffit alors de créer un **alias**.

Pour cela, cliquez sur le bouton « **ajouter** » et remplissez le formulaire demandé. Donnez un nom à l'alias que vous créez, par exemple CME, puis remplissez le chemin où est situé le moteur de recherche (exemple : C:\Documents and Settings\All user\Clustering Meta-Engine\CME\). Une fois l'alias créé, il suffit de cliquer sur son nom pour lancer le site.

Une nouvelle fenêtre du navigateur devrait s'ouvrir avec une adresse du type : **http://127.0.0.1/CME/**

Cette adresse signifie que vous êtes bien en local, c'est-à-dire que le site Internet que vous visitez se trouve sur votre ordinateur. Si jamais vous ne voyez pas le moteur de recherche, actualisez la fenêtre du navigateur plusieurs fois.



3. Installation chez un hébergeur

Si vous possédez un hébergeur Web tel que Free, vous pouvez y placer le moteur de recherche afin de l'utiliser de n'importe où. Il vous faudra pour cela un logiciel FTP afin de transférer les fichiers de votre machine sur le serveur.

Vous devez vérifier au préalable que le serveur Web exécute bien du PHP.

Si cette condition est remplie, installez le logiciel FileZilla se trouvant dans le dossier « **Logiciel** ». Une fois installé, entrez dans les champs adresse, utilisateur et mot de passe situés dans le menu du haut les informations concernant votre hébergement.

Il ne vous reste plus qu'à placer les fichiers sur le serveur distant en les glissant directement depuis le répertoire Windows où ils sont situés vers la fenêtre « **site distant** » du logiciel.

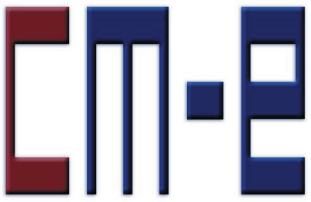
Attention à prendre tous les fichiers et dossiers du répertoire « CME ».

Vous pouvez maintenant utiliser le moteur de recherche en entrant l'adresse de votre hébergeur. Ceci vous permettra de l'utiliser de n'importe où et de le partager avec vos connaissances.

4. Net Draw

Le logiciel Net Draw, développé par Steve Borgatti, permet de dessiner des réseaux. Vous le trouverez aussi dans le répertoire « **Logiciel** ». Ce dernier n'a pas besoin de s'installer sur votre système d'exploitation, il suffit de le lancer directement à partir du fichier « netdraw.exe » se trouvant dans le dossier « **Net Draw** ».

Vous trouverez plus d'information sur l'utilisation du logiciel dans la partie « **Export sur Net Draw** ».



1. L'équipe projet

Le projet Clustering Meta-Engine fut initié par **Julien AGULLO** du Master Intelligence Economique et Territoriale de l'Université de Toulon. Il a été réalisé afin de subvenir à un besoin de plus en plus important en matière de collecte d'informations.

Entrepris dans le cadre d'un contrôle continu en Intelligence Economique coordonné par le professeur **Luc Quoniam**, le projet fut rejoint par d'autres élèves afin d'aider à le réaliser. Ainsi, **Souâad Aïdh**, **Ibrahima Ouane** et **Lamyae Sellak** ont rejoint l'équipe afin de tester le méta moteur mais aussi et surtout écrire ce document de synthèse et le rapport scientifique.

2. Présentation du projet

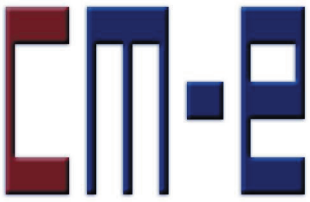
Clustering Meta-Engine est un méta moteur de recherche qui permet de connaître et de cartographier la composition du web. Ainsi à partir d'une requête plus ou moins complexe, nous pouvons connaître les mots clés associés et voir les liaisons existantes entre eux.

Le méta moteur Clustering Meta-Engine va rechercher ces mots clés dans différents moteurs de recherche spécialisés dans ce que l'on appelle les clusters. Lorsque les mots clés sont récupérés, le méta moteur fait une liste unique de tous ces clusters, puis les analyses afin de savoir quels sont les termes qui apparaissent le plus souvent.

Pendant la comparaison, il ne regarde pas si les clusters correspondent exactement entre eux mais vérifie si un cluster apparaît dans un autre et vice-versa.

Une fois la comparaison terminée, les données apparaissent triées et classées pour chaque moteur. L'utilisateur peut alors soit exporter la liste vers le logiciel **Excel** de Microsoft, soit l'exporter vers le logiciel **Net Draw**.

En exportant la liste vers le logiciel Excel, l'utilisateur pourra classer les mots clés, faire des recherches dans la liste, comparer les résultats obtenus mais surtout, pourra effectuer une sauvegarde de la liste obtenue.



L'export dans le logiciel Net Draw permet, quand a lui, d'avoir une cartographie de la liste des résultats obtenus. Ainsi l'utilisateur pourra classer les résultats par groupe, connaître la composition dans le domaine recherché sur le Web.

3. Définitions

Méta moteur : c'est un moteur de recherche qui va rechercher les informations dans d'autres moteurs de recherche. A l'inverse d'un moteur classique, le méta moteur ne possède pas de base de données et n'a pas besoin d'indexer le Web lors de la recherche de nouvelles pages ou alors de modifications. Cependant, il doit toujours vérifier les mises à jour des moteurs dans lesquels il effectue ses recherches afin de fonctionner correctement.

Cluster : c'est un terme ou un mot clé qui est associé à ce que l'on recherche. Cette association peut se faire de différentes façons. La manière la plus fréquente reste la répétition des mots dans les pages. Ainsi, lorsqu'un moteur de cluster va indexer le web, il enregistre dans sa base les mots de chaque page trouvée puis en dégage ceux qui apparaissent le plus souvent.

Clustering : le clustering représente l'action de rechercher des clusters sur le web.

4. Objectif du projet

Du fait que le Web connaît une croissance exponentielle, la recherche d'informations devient de plus en plus vaste. Pour pouvoir limiter les résultats, nous devons alors rendre notre recherche plus complexe et plus précise.

Connaître la composition du Web dans le domaine qui nous intéresse est essentiel afin de rentrer plus précisément au cœur du sujet.

L'objectif du méta moteur est donc de dégager les termes correspondants à ce que l'on recherche directement sur des sites spécialisés en clustering. Le fait d'aller chercher dans plusieurs moteurs rend l'analyse encore plus précise.

Une fois les résultats analysés, nous pouvons affiner nos recherches pour mieux appréhender le domaine qui nous intéresse.

Le **méta moteur CME** peut aussi être utilisé dans un cadre professionnel que personnel. En effet, de nombreuses entreprises peuvent l'utiliser à des fins stratégiques lors d'une étude de marché, d'une veille concurrentielle... Le particulier peut, quand à lui, rechercher des informations sur une personnalité, un événement historique...

Les possibilités d'un tel méta moteur sont très variées mais elles restent encore peu exploitées par les internautes recherchant de l'information sur le Net.

5. Les moteurs de recherche

Le méta moteur **Clustering Meta-Engine** va rechercher les clusters dans des moteurs de recherche spécialisés dans le domaine du clustering.

Nous avons retenus les quatre moteurs suivants :



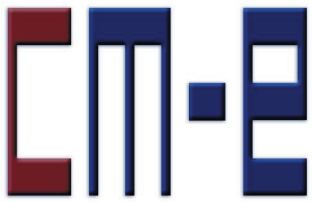
Nous avons choisi ces moteurs d'une part parce qu'ils sont gratuits, mais aussi pour la pertinence des résultats trouvés. Il existe en effet de nombreux moteurs de cluster sur le Web mais ceux là font partis des plus aboutis.

La sélection s'est faite aussi par rapport à une contrainte technique. En effet, certains moteurs sont conçus de telle manière qu'il est impossible d'aller rechercher par un script les différents mots clés qui en ressortent.

Le moteur de recherche **Clusty** appartient au groupe **Vivísimo**. Les résultats obtenus sont donc parfois identiques vu qu'ils travaillent à peu près sur la même base de données.

Voici d'autres moteurs qui n'ont pas été retenus :





Principe d'utilisation

1. La recherche

La recherche des mots clés est la première étape du méta moteur. Lorsque la requête est entrée, un script va chercher dans les moteurs de recherche sélectionnés les clusters qui ressortent et les affiche tels qu'ils apparaissent dans leur moteur respectif.

Il est possible de réaliser des requêtes complexes à la manière de Google. Ainsi nous pouvons mieux affiner nos recherches dans un certain domaine.

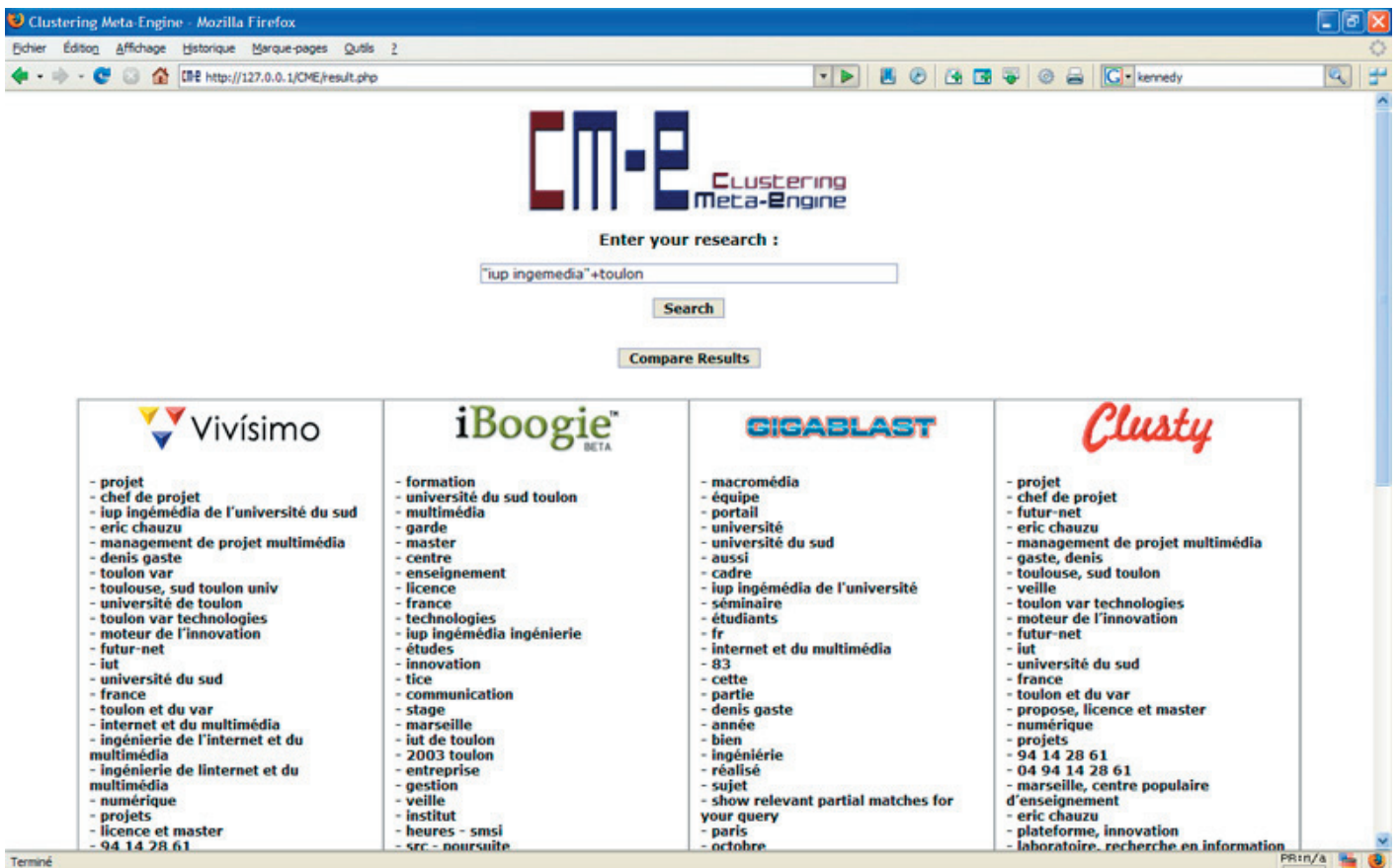
Voici la liste des outils disponibles :

- **Le plus (+)** placé entre 2 mots, permet de rechercher des clusters liés aux 2 mots et non à l'un ou à l'autre comme le fait l'espace.
- **L'espace ()**, contrairement au plus, permet de chercher des clusters associés soit au premier mot, soit au second mot, soit au deux. Il ressort donc plus de résultats mais au détriment de la précision.
- **Les guillemets (« ou ')** permettent, quand ils sont placés autour de plusieurs mots, de rechercher le terme exactement tel qu'il est écrit.

A partir de tous ces outils, nous pouvons former des requêtes complexes et obtenir des résultats beaucoup plus précis.

Il est possible de cliquer sur un résultat dans la liste afin de relancer une recherche correspondant au terme cliqué.

Voici une capture d'écran de ce que l'on obtient d'une recherche :



2. La comparaison des résultats

En cliquant sur le bouton « Compare Results » lorsque la recherche est terminée, un script de comparaison est exécuté. Le principe de ce script est assez simple mais sa réalisation fut complexe.

Une liste est créée regroupant tous les clusters de chaque moteur. Puis le premier élément de la liste est comparé à tous les autres. Lorsqu'un terme contient cet élément où lorsqu'un terme est contenu dans cet élément alors on incrémente le compteur lié à ce dernier. Ensuite le script continue avec le second élément du tableau et ainsi de suite en repartant à chaque fois du début.

Nous pouvons ainsi en déduire les termes qui apparaissent le plus souvent dans la liste et la recomposer entre les quatre moteurs. Les résultats obtenus apparaissent entre parenthèses à côté de chaque cluster.

Il est intéressant de regarder les résultats les mieux classés car ils représentent les domaines dans lesquels la recherche se rapproche le plus. Les résultats du bas sont aussi intéressants puisqu'ils représentent des termes qui ne sont pas trop associés donc pouvant être innovant dans le domaine.

Voici une capture d'écran de ce que l'on obtient lors d'une comparaison :



Clustering Meta-Engine - Mozilla Firefox

Enter your research :

Generate XLS Data Format Results 208 for "iup ingemedia"+toulon Generate VNA Data Format

Vivísimo	iBoogie ^{BETA}	GIGABLAST	Clusty
<ul style="list-style-type: none"> - projet (8) - université de toulon (8) - université du sud (8) - iup ingemédia de l'université de toulon (8) - internet et du multimédia (7) - iup ingemédia de l'université du sud (6) - france (6) - ingénierie de l'internet et du multimédia (6) - compétences déployée dans le cadre (6) - chef de projet (5) - management de projet multimédia (5) - ingénierie de l'internet et du multimédia (5) - 04 94 14 28 61 (5) - centre de formation flash (5) - eric chazu (4) - toulon var (4) - futur-net (4) - projets (4) - 94 14 28 61 (4) 	<ul style="list-style-type: none"> - agence edufrance (7) - université du sud toulon (6) - france (6) - formation (5) - centre (5) - multimédia (4) - iut de toulon (4) - licence (3) - tice (3) - numérique (3) - konoyaro.net - book et divers (3) - garde (2) - iup ingemédia ingénierie (2) - études (2) - veille (2) - dossiers - secrétariat de l iup (2) - r-pur - 1ère centrale de réservation d appn (2) - diplôme (2) - studio - bee (2) - master (1) - enseignement (1) - technologies (1) - innovation (1) - communication (1) 	<ul style="list-style-type: none"> - d (72) - université (14) - net (14) - université du sud (8) - internet et du multimédia (7) - fr (6) - france (6) - cadre (4) - iup ingemédia de l'université (3) - denis gaste (3) - sujet (3) - recherche (3) - 04 (3) - provence (3) - macromédia (2) - portail (2) - étudiants (2) - 83 (2) - port (2) - équipe (1) - aussi (1) - séminaire (1) - cette (1) - partie (1) - année (1) 	<ul style="list-style-type: none"> - projet (8) - université du sud (8) - université de toulon (8) - iup ingemédia de l'université de toulon (8) - france (6) - ingénierie de l'internet et du multimédia (6) - université du sud toulon-var (6) - compétences déployée dans le cadre (6) - chef de projet (5) - management de projet multimédia (5) - 04 94 14 28 61 (5) - centre de formation flash (5) - futur-net (4) - eric chazu (4) - projets (4) - 94 14 28 61 (4) - marseille, centre populaire d'enseignement (4) - laboratoire, recherche en information et communication (4) - compétences (4)

Entre les 2 boutons servant aux différents exports, nous pouvons voir combien de résultats ont été obtenus lors de la recherche. Il n'apparaît que maintenant car pendant la comparaison, toutes les répétitions dans chaque moteur ont été supprimées.

3. Export dans Excel

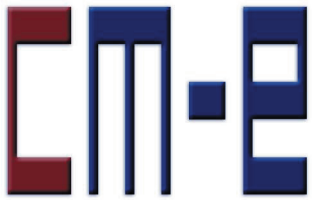
Le logiciel **Excel de Microsoft** est un tableur qui permet de créer des tableaux puis de les analyser. En cliquant sur le bouton « **Generate XLS Data Format** », un script est exécuté servant à créer un fichier Excel contenant un tableau des différentes listes.

Le fichier qui s'ouvre automatiquement peut soit être enregistré sur le disque dur de l'utilisateur, soit être ouvert directement avec le logiciel **Excel**. L'enregistrer permet d'avoir une sauvegarde de la recherche afin d'être réutilisée ultérieurement.

Le principe de cet export est de trier les résultats de différentes manières à partir des fonctionnalités d'**Excel**. Nous pouvons alors réaliser des comparaisons, faire des statistiques, des graphiques...

Voici une capture d'écran d'un export dans Excel :

A	B	C	D
1 VIVISIMO	IBOOGIE	GIGABLAST	CLUSTY
2 internet et du multimédia	agence edufrance	d	internet et du multimédia
3 projet	université du sud toulon	net	projet
4 université du sud	france	université	université du sud
5 iup ingénierie de l'université du sud	formation	internet et du multimédia	iup ingénierie de l'université du sud
6 ingénierie de l'internet et du multimédia	centre	université du sud	ingénierie de l'internet et du multimédia
7 ingénierie de l'internet et du multimédia	multimédia	fr	ingénierie de l'internet et du multimédia
8 université de toulon	iut de toulon	france	france
9	licence	cadre	université de toulon
10 france	tice	denis gaste	compétences déployée dans le cadre
11 iup ingénierie de l'université de toulon	numérique	provence	chef de projet
12 compétences déployée dans le cadre	konoyaro.net - book et divers	iup ingénierie de l'université	management de projet multimédia
13 chef de projet	garde	sujet	04 94 14 28 61
14 management de projet multimédia	iup ingénierie ingénierie	recherche	provence À aix, iup ingénierie
15 04 94 14 28 61	études		4 centre de formation flash
16 centre de formation flash	veille	macromédia	eric chauzu
17 eric chauzu	dossiers - secrétariat de l iup	portail	
18	r-pur - 1ère centrale de réservation d appn	étudiants	denis gaste
19 denis gaste	diplome		83 toulon var
20 toulon var	studio - bee		toulon var technologies
21 toulon var technologies	master	port	projets
22 projets	enseignement	équipe	94 14 28 61
23 94 14 28 61	technologies	aussi	marseille, centre populaire d'enseignement
24 laboratoire, recherche en information et communication	innovation	séminaire	laboratoire, recherche en information et communication
25 marseille, centre populaire d'enseignement	communication	cette	provence
26 provence	stage	partie	compétences
27 compétences	marseille	année	colloque, tice méditerranée
28 colloque, tice méditerranée	2003 toulon	bien	desu services multimédia en ligne
29 desu services multimédia en ligne	entreprise	ingénierie	toulouse, sud toulon univ
30 toulouse, sud toulon univ	gestion	réalisé	moteur de l'innovation
31 moteur de l'innovation	institut	show relevant partial matches for your	
32 futur-net	heures - smsi	query	futur-net
33 iut	src - poursuite	paris	iut
		octobre	toulon et du var



Principe d'utilisation

A première vue ce fichier n'est pas du tout exploitable vu qu'il possède beaucoup trop d'information. On remarque d'ailleurs qu'il n'y a aucune cohésion dans la position de chaque cluster car le script les a situés aléatoirement dans une zone de la fenêtre. En effet, ils ont été positionnés de façon à être regroupés par moteur.

Voici la correspondance des formes et des couleurs :

- **Triangle jaune** : Vivisimo
- **Losange vert** : IBoogie
- **Carré bleu** : Gigablast
- **Rond rouge** : Clusty

Dans le menu situé à droite on peut voir tous les clusters qui ont été analysés. Ils sont classés par « **ID** » et possèdent des noms différents afin qu'il n'y ait pas de conflit entre les liens.

Vous pouvez cliquer sur le menu déroulant ou est écrit « **ID** » afin de voir les autres catégories de classement.

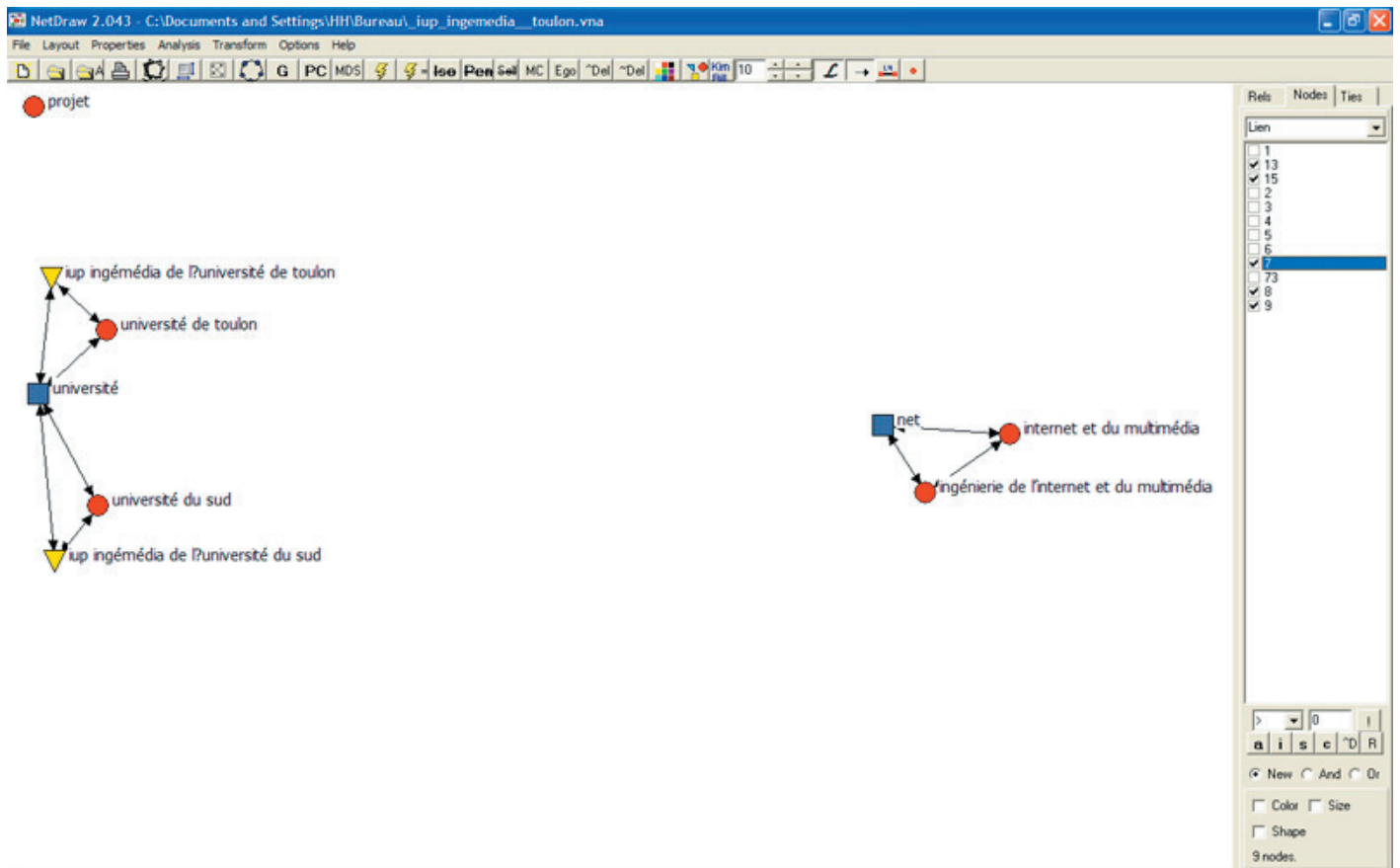
Vous trouverez donc deux autres catégories qui sont :

- **Moteur**
- **Lien**

En sélectionnant « **Moteur** », vous affichez la liste des moteurs utilisés. Vous n'avez qu'à tout décocher pour enlever les clusters, puis sélectionner « **Lien** » dans le menu .

Maintenant nous allons reclasser les clusters en fonction des liens. Il suffit pour cela de cocher, une après l'autre, les cases correspondantes à la note qui a été attribuée lors de la comparaison.

Voici ce que l'on obtient après avoir affiché les 5 plus grands liens :

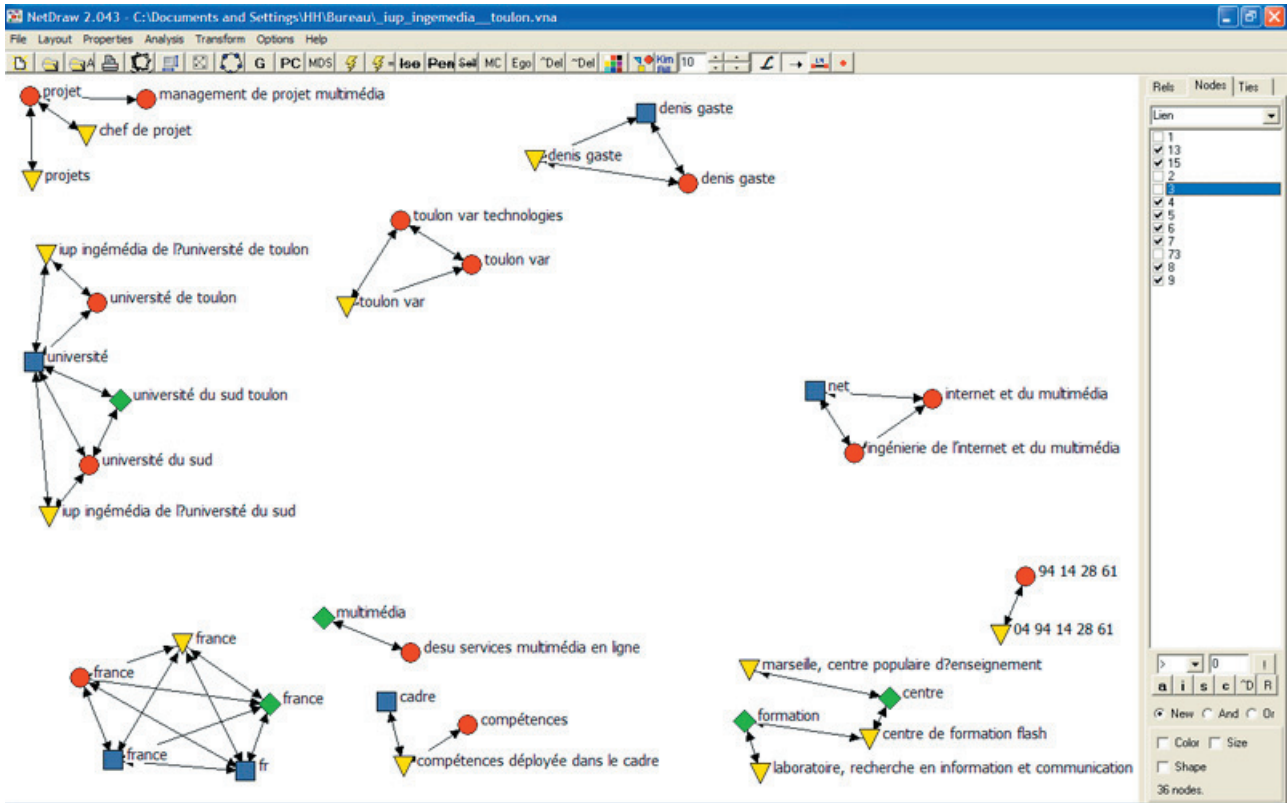


A chaque fois qu'un cluster apparaît vous pouvez le positionner en cliquant sur lui puis en le déplaçant. Ceci déplacera automatiquement les flèches entre les clusters avec qui il est associé.

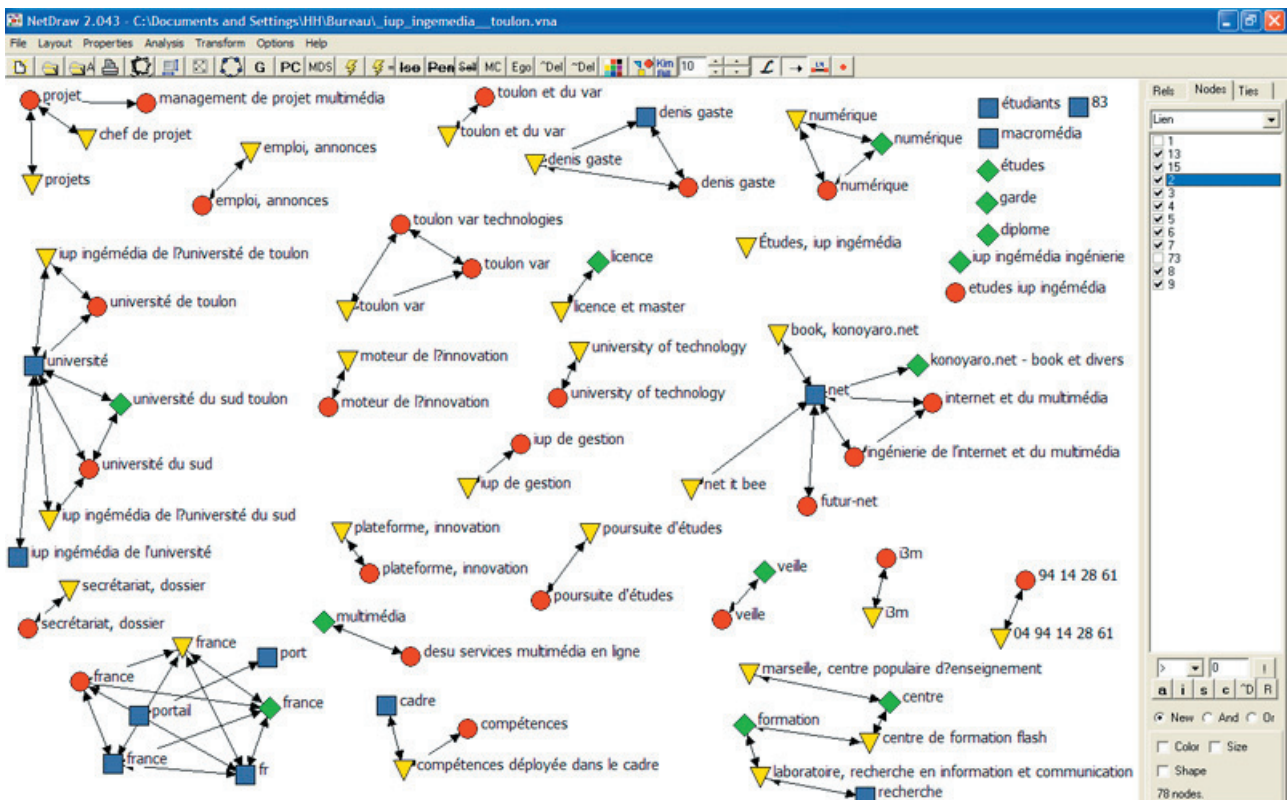
Après avoir tout positionné, continuez d'afficher les clusters par ordre de décroissance. Au bout de quelques liens et de repositionnements vous obtiendrez une cartographie beaucoup plus claire et vous commencerez à en dégager des groupes.

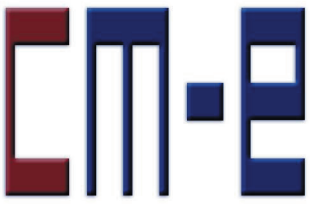
Vous pouvez également supprimer des clusters qui ne vous paraissent pas pertinents. Pour cela, cliquez sur ces derniers et appuyez sur **SUPPR**. Ceci effacera aussi les liens avec les clusters auxquels ils sont rattachés.

Voici ce que l'on obtient après avoir affiché les 8 plus grands liens :



Voici le résultat final de la cartographie où l'on peut remarquer les groupes obtenus :





Conclusion

1. Contraintes

Bien que Clustering Meta-Engine possède de nombreuses fonctionnalités afin d'aider à la recherche d'information, il reste néanmoins contraint à plusieurs conditions.

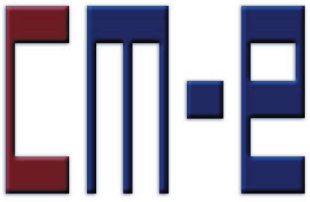
Tout d'abord l'installation sur certains serveurs Web distants peut poser quelques problèmes. En effet, étant un site Web demandant beaucoup de ressources, il peut dans certaines conditions, n'avoir pas le temps d'exécuter ses fonctions. Pour cela il faudrait avoir accès à la configuration du serveur Apache de votre hébergeur et suivre la même étape que lors de la modification du fichier **php.ini** vu dans la phase « **Installation en local** ». Si on n'a pas accès à cette configuration, il se peut que le script permettant de rechercher les clusters ne possède pas le temps qu'il lui faut afin de bien s'exécuter.

Un autre problème concerne les mises à jour des moteurs sélectionnés. Il peut arriver que les sociétés qui ont développés ces outils changent complètement ou en partie leurs scripts, le méta moteur ne pourra plus alors aller chercher les clusters. Cependant, si un tel cas se produit, cela n'empêchera pas la bonne utilisation des autres moteurs de recherche.

Le serveur Web **Easy PHP** ainsi que le logiciel **Net Draw** sont prévus pour le système d'exploitation Windows XP. Si vous installez ces logiciels sur un autre système, il se peut que vous rencontriez des problèmes de compatibilité. Il en est de même pour le site Internet Clustering Meta-Engine qui est optimisé pour les navigateurs **Firefox** de la société Mozilla et **Internet Explorer** de Microsoft, tous les 2 sous Windows XP.

Vous êtes libre d'utiliser le site Clustering Meta-Engine pour n'importe quelle finalité tant que cela reste dans le domaine de la recherche d'information. Toute modification du code source, pour que le méta moteur soit adapté à vos besoins, est autorisé sans même avoir besoin de demander une autorisation. Cependant si cela ce produit, il serait redevable de faire parvenir les modifications à l'adresse julien@wanadoo.fr si celles-ci apportent quelque chose de constructif pour le projet.

Il est strictement interdit de vendre le projet, de diffuser son code source ou encore de l'échanger par respect envers son développeur.



Si vous rencontrez un problème lors de l'utilisation du méta moteur, veuillez faire remonter l'information par mail à agullo.julien@wanadoo.fr afin de le corriger.

2. Perceptives d'avenir

Le moteur Clustering Meta-Engine pourrait faire l'objet d'un site Web à part entière et s'enrichir de fonctionnalités supplémentaires. Nous devons pour cela vérifier les droits concernant l'utilisation des informations prises dans les moteurs sélectionnés.

Si un tel projet se réalisait, alors d'autres moteurs de cluster viendraient compléter ceux déjà présent. Il serait également envisageable de réaliser la cartographie en Flash ce qui permettrait de ne pas avoir besoin d'un logiciel pour la visualiser.

La création d'une page concernant la recherche avancée, composée d'une manière similaire à celle de Google, aiderai les personnes peu expérimentées à réaliser des requêtes complexes dans le méta moteur.

Il serait possible d'utiliser la technologie AJAX afin de réaliser des requêtes dans un seul des moteurs sans avoir besoin de recharger la page complètement. En exploitant un peu mieux le CSS et le Javascript, nous pourrions rendre les colonnes de chaque moteur plus dynamique afin de les déplacer à coté d'une autre pour comparer les résultats.

Le référencement du site sera une étape importante pour sa réussite. En effet, si un site Web n'est pas référencé il ne peut pas être connu du grand public et ne sera pas exploité.

3. Le rapport scientifique

Pour compléter le projet Clustering Meta-Engine, un rapport scientifique devra être rédigé d'ici la fin du mois de février. La version finale de Clustering Meta-Engine datant du 20 décembre 2006, la rédaction du rapport n'a pas pu se faire avant les vacances de Noël 2006. En effet, le rapport devra comporter des analyses complètes de différentes requêtes, des statistiques et des analyses poussées des cartographies obtenues.

Une fois le rapport rédigé, nous le soumettrons à une revue de publication qui n'a pas encore été choisie définitivement. Le but de cet article sera de démontrer les atouts du clustering et l'importance que possède cette approche pour compléter le processus de recherche d'information.

L'article sera publié en anglais et en français mais cela dépendra de la revue de publication qui sera choisie.

4. Conclusion

Le projet Clustering Meta-Engine n'est qu'au début de sa conception. Les résultats obtenus sont déjà fort convaincants dans le processus de recherche d'information. Si le projet continu à se développer, le site Clustering Meta-Engine peut devenir une référence dans le domaine de la recherche de cluster.

Bien entendu, quelques informations du type juridique doivent être vérifiées et des autorisations doivent être demandées afin que le projet se réalise en toute légalité.

Le site Web **www.clustering-me.com** devra être disponible dans le courant du mois de janvier 2007. En attendant vous pouvez procéder selon les 2 manières expliquées dans le dossier ou encore l'utiliser à partir du site :

<http://www.virtual-it.fr/cme/>

Nous remercions toutes les personnes qui ont participé à la conception du méta moteur Clustering Meta-Engine.