

Novembre 2008



# i-mag

Magazine de l'AAE ENSIMAG



>>> A l'heure de  
l'internet global

42



## Souplesse, rapidité, puissance... vous parlez d'Atos Origin

Partenaire informatique mondial des Jeux Olympiques pour 6 olympiades de 2002 à 2012, le contrat conclu entre Atos Origin et le Comité International Olympique (CIO) constitue à ce jour le plus important contrat informatique lié au monde du sport. Une épreuve de longue haleine pour nos collaborateurs qui ont la responsabilité principale de l'informatique, en matière de conseil, d'intégration de systèmes, de gestion des opérations, de sécurité de l'information et de développement des logiciels applicatifs. Parmi les exploits des athlètes, il y a aussi ceux de nos collaborateurs. Atos Origin exerce ses activités sous les noms d'Atos Origin, Atos Worldline et Atos Consulting.

### Étudiant(e) / Jeunes diplômé(e)s :

Vous êtes uniques, nous saurons vous proposer un parcours à la hauteur de vos aspirations. Pour en savoir plus et postuler, retrouvez-nous sur [atosoriginrecrute.fr](http://atosoriginrecrute.fr)



**ADVANCE YOUR CARRER >>**

[atosoriginrecrute.fr](http://atosoriginrecrute.fr) | PRENEZ DE L'AVANCE SUR VOTRE RÉUSSITE



# Ready for the challenge of a lifetime?



Fermat's integrated risk and performance management software is recognised by financial institutions worldwide.

Diverse team of 30 different nationalities are servicing more than 100 client sites in 30 countries.

Succeeding at Fermat means having the abilities that match the company's major characteristics: very rapid growth, strategic challenges, responsiveness and international scope.

[www.fermat.eu](http://www.fermat.eu)



**Fermat**

INTEGRATED RISK AND PERFORMANCE  
MANAGEMENT SOFTWARE

BEIJING . BRUSSELS . DUBAI . FRANKFURT . GRENOBLE . HONG KONG . NEW YORK . PARIS . SÃO PAULO . SEOUL . SINGAPORE . SYDNEY . WARSAW

## LA SÉCURITÉ, UN DÉFI PERMANENT

Leader mondial de la sécurité numérique, avec un chiffre d'affaires annuel de plus de 1,6 Md€, Gemalto propose des solutions de sécurité intégrées, du développement de logiciels jusqu'à la création et la fabrication d'outils de sécurité numérique comme les cartes à puce, cartes SIM, passeports électroniques...

Ses produits et services sont utilisés par plus d'un milliard de personnes à travers le monde pour de nombreuses applications, notamment dans les télécommunications, les services financiers, les administrations, la gestion des identités, le contenu multimédia, la gestion des droits numériques, la sécurité informatique et les transports en commun.

Philippe PROUST,  
(Ensimag 1988)  
Responsable  
de l'équipe  
Security Labs  
à La Clotat



### En quoi consiste la mission de votre équipe ?

Nous sommes chargés d'assurer l'adéquation des produits et des solutions Gemalto avec les exigences de sécurité des applications et des marchés visés : carte bancaire, solutions Pay-TV, carte d'identité... Pour cela, nous développons les algorithmes cryptographiques intégrés dans les produits et vérifions que les protocoles et les programmes offrent le niveau de robustesse requis. Toutes les composantes d'une solution sont considérées, depuis le Serveur jusqu'à l'objet portable sécurisé en passant par le middleware sur PC. En particulier, les objets portables – les cartes à puce par exemple – doivent résister à plusieurs types d'attaques, tant physiques que logiques. Les attaques physiques sont de trois ordres : les attaques intrusives, par modification du microcircuit (silicium) ; les attaques par perturbation, sur les conditions de fonctionnement (alimentation...) et visant à corrompre une partie sensible d'un programme ; les attaques par canaux cachés, basées sur l'exploitation d'informations (grandeurs physiques) obtenues par l'observation de l'objet en fonctionnement. D'autre part, il faut évidemment prémunir les objets sécurisés contre les attaques plus connues du monde de l'informatique (Buffer overflow...). Il faut en effet savoir que les dernières cartes à puce intègrent les technologies les plus avancées : VM Java, pile TCPIP, technologie webserviceur... - et que cette intégration nécessite une sécurisation préalable et méticuleuse.

### C'est une activité en perpétuelle évolution...

Notre objectif étant d'embarquer dans les objets portables une cryptographie faisant appel aux algorithmes les plus avancés, nous suivons au jour le jour les progrès de la communauté scientifique. Ainsi, nous remplaçons progressivement nos algorithmes DES par des AES, nous avons porté la taille de nos clés RSA à 2048 bits et offrons déjà de la cryptographie sur courbes elliptiques. La sécurité est une activité très variée où l'on apprend tous les jours quelque chose. Nous nous livrons à un duel intellectuel sans fin avec les hackers. C'est d'autant plus excitant que nous essayons de conserver toujours une longueur d'avance.

### De quels moyens disposez-vous pour cela ?

Mon équipe est composée d'une cinquantaine de personnes. Le labo en accueille une dizaine pour vérifier à travers des expériences pratiques que nos produits ne présentent plus de failles. L'implémentation des algorithmes de cryptage en employe également une dizaine, tout comme la conception de protocoles et de programmes sûrs. Le reste de l'équipe est réparti sur des activités de cryptographie avancée, de validation à travers des méthodes formelles et de certification de produits (critères communs, FIPS...). Et nous élargissons notre domaine de compétences et nos prérogatives, Gemalto proposant désormais des solutions complètes et intégrées, mettant en œuvre des serveurs, les PCs des usagers, les réseaux de télécommunication en plus des objets portables...

### Comment avez-vous intégré Gemalto ?

Tout simplement en répondant à une annonce relayée par l'Ensimag. En fait, j'ai été recruté en mars 2005 pour une mission de six mois en R&D dans le cadre de mon stage de fin d'études. À l'issue de celui-ci, Gemalto m'a proposé un CDI dans la section Identity and Security Services.

La présence de nombreux anciens de l'école m'a permis de me sentir très vite à l'aise dans cette entreprise où la population est jeune et cosmopolite. Mon intégration a également été facilitée par le séminaire Star One, organisé à Shanghai pour les nouveaux arrivants.

### Qu'est-ce qui vous incite à poursuivre votre carrière chez Gemalto ?

Gemalto offre de réelles possibilités d'évolution. Ainsi, l'an dernier, j'ai eu l'opportunité d'intégrer une équipe Identity Custom Solution. Au sein de celle-ci, je participe à l'installation chez nos clients de solutions clé en main de production de cartes à puces, comme par exemple le centre de production de cartes de santé algériennes à Alger. De par son rôle d'interface client et la dimension internationale de ses projets, c'est un poste très enrichissant professionnellement. De plus, Gemalto permet à chacun de consolider ses connaissances et, surtout, de développer de nouvelles compétences. Le groupe dispose en effet d'un intranet très riche en documentation et propose un service d'apprentissage dans la plupart des sciences de l'ingénieur : électronique, réseaux, programmation... Sans oublier la possibilité d'effectuer jusqu'à trois formations tous les ans.

### Comment envisagez-vous votre avenir professionnel ?

Je veux continuer à évoluer en travaillant sur des projets internationaux, parfaire mon expertise et gagner en gestion de projet. La Sécurité informatique et les DSI offrent un large panel de métiers, j'espère devenir « Grand » dans l'un d'eux ;

Nicolas ANEAS,  
(Ensimag 2005)  
Ingénieur  
Projet  
à Meudon



Gemalto emploie plus de 10 000 salariés, dont 1 300 ingénieurs en R&D, répartis dans 85 bureaux implantés dans 40 pays. Gemalto favorise la diversité et compte plus de 40 nationalités différentes parmi ses collaborateurs à son siège social de Meudon. La politique RH privilégie la promotion interne, la mobilité fonctionnelle et géographique. Gemalto recrute actuellement des jeunes diplômés pour ses activités R&D, avec des compétences en langages (Java, C++, Assembleur) et des connaissances en architecture logicielle et en cryptographie. Les ingénieurs R&D acquièrent une expérience technique qui leur permettra d'évoluer vers du management, de l'expertise, du consulting ou du marketing. Les stages de fin d'études et les VIE sont la principale source de recrutement.

Contact Emploi/Stages/VIE:  
sylvie.mollet@gemalto.com

**gemalto**  
security to be free

# éditorial



**Jean-Marc Darrigol**  
Ensimag 1988  
Directeur Ingénierie  
au sein de Yahoo

## Jean-Marc Darrigol

Ensimag 1988, est Directeur Ingénierie au sein de Yahoo à Echirrolles. Il commence sa carrière chez Bull puis prend très tôt le virage de l'Internet. Il rejoint Kelkoo en mars 2000 comme responsable des sites de marque blanche, avant de prendre la direction de la R&D en 2004.

**L**es plus anciennes références aux interactions sociales qui seraient possibles avec un réseau d'ordinateurs datent de 1962, et la première démonstration publique du réseau ARPANET a eu lieu en 1972. Mais c'est le début des années 90, avec la création au CERN des standards à la base du web et l'apparition du navigateur Mosaic, qui marque le début d'Internet tel que le grand public le connaît. Web, courrier électronique et messagerie instantanée ont explosé dans les deux dernières décennies et Internet est maintenant universel, omniprésent, global.

Impossible d'être exhaustifs devant un domaine aussi large, donc nous essayons seulement dans cette revue de donner au lecteur quelques pistes de réflexion.

Réflexion sur la taille tout d'abord. Avec une croissance exponentielle du nombre de systèmes connectés, Internet oblige à inventer des moyens pour l'appréhender. C'est ce que Franck Ghitalla nous explique dans son article sur la cartographie du web. Les méthodes de recherche doivent également être revues, les mots clés ne permettant qu'un accès partiel à la masse d'information existante. Gilles Vandelle nous explique ainsi comment les réseaux sémantiques sont une piste pour améliorer la pertinence des recherches en ligne. Côté infrastructures, les besoins sans cesse croissants rendent nécessaires la mise en place de technologies innovantes, telles que le grid computing dont Rémy Amouroux nous explique les principes.

Réflexion sur les services ensuite. Internet a révolutionné notre façon de vivre en créant de nouveaux services pour les utilisateurs ou en changeant des services existants. Nous avons sélectionné trois exemples pour l'illustrer : le vote par Internet, qui entre maintenant dans une phase de maturité (Benoît Chenon), le crédit de particulier à particulier, service qui est maintenant rendu possible à grande échelle (Marc Duleba) et enfin l'accès aux informations multi-sources (Pierre Chappaz). Trois gouttes d'eau dans un océan de sites et de services évoluant sans cesse.

Pour compléter ces deux axes, Laurent Gatignol nous parle de l'Internet sur mobile, qui participe à cette universalité de l'accès au réseau. Enfin, le besoin d'évolution rapide des sites dans un environnement technique complexe et l'impact direct des modifications sur les utilisateurs obligent les équipes de développement à envisager de nouveaux processus. Dans ce contexte, les méthodes dites Agile sont pertinentes et très utilisées par les sociétés Internet, comme nous l'explique Alexandre Boutin.

Au final, cette revue effleure à peine la surface de ce domaine, mais il faudrait une encyclopédie pour en faire le tour, et celle-ci serait obsolète à peine écrite. Des domaines entiers de l'économie sont seulement à l'aube d'une révolution dont on a du mal à mesurer les impacts. Je pense aux loisirs numériques et à l'accès aux produits culturels par exemple. Nous sommes en présence d'un média et d'un pan d'économie qui croissent de façon exponentielle et évoluent sur des cycles d'une fréquence inconnue jusque là. Comment dans ce cas tirer des plans sur l'avenir ?

Je voudrais terminer en remerciant les auteurs qui m'ont fait l'amitié de participer à cette revue, et en vous souhaitant une bonne lecture. \*

INGÉNIERIE / FINANCE DE MARCHÉ / TELECOM / BIO-INFORMATIQUE / TEST & VALIDATION / E-CONSULTING

## Plus loin dans le futur, plus proche dans vos projets



Autour de la Finance de marché (38%) et des Télécoms (30%), Datavance partage la culture et les valeurs des élèves de l'INPG (ENSIMAG...). Comme vous, nos 850 collaborateurs aiment résoudre des problèmes complexes et en traiter tous les aspects, réaliser des projets dans leur intégralité, rester toujours ouverts aux autres.

Depuis la création de Datavance il y a 18 ans, notre développement appartient à leur esprit d'innovation, à leur audace, à leur professionnalisme.

Pour les années à venir, c'est à vous qu'il appartiendra. C'est pourquoi nous privilégierons, toujours, vos idées et votre technicité. Favoriserons votre trajectoire professionnelle pour qu'elle réponde à vos attentes. Et vous offrirons un environnement de travail de qualité.

Retrouvez toutes nos offres, notre culture, nos actualités sur [www.datavanceRH.com](http://www.datavanceRH.com)

18 ans  
d'expérience



**DATAVANCE**  
avec VOUS

Conseil & Expertise en Technologies innovantes

# sommaire

## « A l'heure de l'internet global »

### EDITORIAL

Par Jean-Marc Darrigol 3

### La Naissance d'une géographie du Web ?

Par Franck Ghitalla 6

### Utilisation des réseaux sémantiques dans la recherche Web

Par Gilles Vandelle 12

### Web Services et Grid Computing : la nouvelle ère des systèmes distribués

Par Rémy Amouroux 14

### Le vote par internet : quel avenir ?

Par Benoît Chenon 18

### Crédit de particulier à particulier. L'argent sans la banque ?

Par Marc Duleba 20

### Wikio, un portail d'information 2.0

Par Pierre Chappaz 18

### L'internet mobile

Par Laurent Gatignol 26

### L'agilité au Service du Web

Par Alexandre Boutin 28

### Vie de l'association

Par Mouna Beyk 31

Photographie de couverture : © Nmedia - Fotolia.com

Edition : i-mag est publié par l'Association des anciens élèves de l'ENSIMAG, 68, Bld de Port Royal 75005 Paris - Tél.06 28 50 08 85; site : www.aae-ensimag.com ; e-mail : contact@aae-ensimag.com ; rédacteur en chef : Jean-Marc Darrigol. PUBLICITE : Editions 50, 87, route de Grigny, 91130 Ris Orangis. CONCEPTION GRAPHIQUE ET MAQUETTE : tél. 04 42 99 38 31. IMPRESSION : Imprimerie de Montligeon, 61400 La Chapelle Montligeon. Dépôt légal : mai 2008 ; ISSN 1774-7104.



**Marquez de votre empreinte  
tous les domaines de la Sécurité  
Rejoignez-nous!**

digital - Photo Getty Images, 387 945 925 - 7 photos et vidéos dans le monde entier



**you  
and us,** worldwide™

Acteur majeur dans les domaines de la protection des citoyens, des transports, des entreprises et des Etats, Sagem Sécurité applique ses solutions dans plus de 70 pays. Grâce à son expertise unique au monde, elle maîtrise les technologies utilisant notamment la biométrie où elle est leader mondial,

les terminaux sécurisés et les cartes à puce. Vous aussi, contribuez à rendre le monde plus sûr en rejoignant ses équipes : de nombreuses opportunités vous y sont proposées. [www.sagem-securite.com](http://www.sagem-securite.com)





## La naissance d'une géographie du web ?

Par Franck Ghitalla

Maître de Conférences en Sciences de l'Information - Président de l'association de recherche WebAtlas

*"Nous avons créé le Web, et nous avons pour devoir de le comprendre". Ainsi s'est exprimé Tim Berners-Lee le 10 novembre 2006 en fondant le W.S.R.I. (Web Science Research Initiative). Cette initiative récente du M.I.T. et de l'université de Southampton de promouvoir l'émergence d'une "science du web" montre, si besoin était, tout le chemin qu'il reste à parcourir pour comprendre ce réseau dont la raison nous échappe encore. Le web représente en effet un « e-cosystem » (1) documentaire relativement inédit. Certains principes de son organisation sont encore mal connus, même si depuis quelques années on commence à percevoir certaines propriétés génériques de cet espace hypertexte ouvert et dynamique.*

### Franck Ghitalla

Maître de Conférences en Sciences de l'Information à l'Université de Technologie de Compiègne depuis 1999.

Président de l'association de recherche loi-1901 WebAtlas - **Consultant en management de l'information**

#### Thèmes de recherche :

technologies web, cartographies des systèmes d'information et des réseaux, théorie et expérimentation autour des Network-Sciences.

Mène une réflexion théorique sur la « géographie de l'information » à l'heure des réseaux distribués comme le web et des systèmes complexes. En particulier, il s'agit d'essayer de comprendre les principes de l'architecture documentaire du web. Ce travail théorique s'accompagne du développement d'une ingénierie scientifique et donc différents projets technologiques comme la société R.T.G.I. lauréate du concours OSEO-ANVAR en 2006, le portail dédié à l'ingénierie web [web-mining.fr](http://web-mining.fr) ou les projets expérimentaux dirigés par l'association de recherche WebAtlas dont il est le président.

Si l'on quitte quelques instants notre navigateur-web et que l'on oublie les listes à plat que nous donnent les moteurs de recherche, alors s'ouvre un univers où les documents web et leurs liens sont inextricablement mêlés. Ils forment un système complexe qui ne semble plus rien devoir à l'organisation physique des réseaux, se développant de lui-même au niveau le plus abstrait de toute cette série de couches de machines et de protocoles qui forment l'Internet. L'activité de nombreux chercheurs et ingénieurs, dans le secteur de la recherche académique comme industrielle, est aujourd'hui concentrée sur l'étude du web comme architecture dont il s'agit de construire les modèles possibles en s'appuyant sur des données expérimentales ou des données indexées dans les grandes bases d'indexation des moteurs de recherche.

A parcourir les publications issues du champ des computer sciences et du web-mining ces dernières années, on se rend compte combien les chercheurs et les ingénieurs contribuent petit à petit à le dessiner comme un système topologique complexe, esquissant à différentes échelles et par touches successives son modèle à la fois graphique et logique. C'est là, dans

(1) *The Laws of the Web, Patterns in the Ecology of Information*, B.A. Huberman, M.I.T.-Press, 2001.

The logo for Grenoble INP ENSIMAG is centered in a white circle. It features the text 'Grenoble INP' in a dark green font above 'ENSIMAG' in a lighter green font. To the right of the text are several vertical bars of varying colors (orange, blue, purple, red, green, blue) and a stylized network diagram of nodes and lines in the background.

Grenoble INP  
ENSIMAG

## L'Ensimag remercie les entreprises qui lui ont renouvelé leur confiance en lui versant la taxe d'apprentissage en 2008

4D, ACCI FRANCE MSP, ACCENTURE, ADCO, ADPCCO IT, ADI ALTERNATIVE INVESTMENTS, ADR, AGIR INFORMATIQUE, AIR FRANCE, AIRBUS, AICATEI LUCCENT, ALCION GROUP, ALP'EMPLOI, ALTEN, ALTI, ALUMINIUM PÉCHINEY, ALYOTECH, AMADDEUS, AMT GIE, ANCO, ANSALDO STS, ANTIDOT, AQUILA CONSULTING SARL, ARCELORMITTAL, AREVA, ARJOWIGGINS, ARMONYS TECHNOLOGIES, ASSET TECHNOLOGY, ASTER, ASTIRUM SAS, AIGS ORIGIN, AUBAY, AUBAY PROUPE, AXA, AXAITO, BANQUE DE FRANCE, BERTIN TECHNOLOGIES, BIOMPRIFLUX, BND PARIBAS, BUII, BUSINESS ET DECISION, BUSINESS OBJECT, CADEXTAN, CAI GRAPH SA, CALISTEL, CALYON, CAP GEMINI, CASINO, CILME INFORMATIQUE, COFACE, COH-RAMI, COMPLETEL SAS, CREDIT AGRICOLE ASSET MANAGEMENT, CREDIT AGRICOLE F SA, CRII TECHNOLOGY, CSC COMPUTER SCIENCE SA, CSSI, DANAF, DASSAULT DATA SERVICES, DASSAULT SYSTEMES, DECAI OG BENIF INFORMATIQUE, DELCOURT OLIVIER, DERIVEDPETS, DOCCA POWER, DREAMSOFT, DXO LADS, DXO SILICON GROUP, EADS, EDF, EUROCOPTER, FERMAT, FININFO, FORCLUM, FOST, FRANCE TELECOM, GDF, GEA, HEWLETT PACKARD, HIGH TECH INTELLIGENCE, HITACHI DATA SYSTEMS, HITV, IBM FRANCE, IMRA EUROPE SAS, INFORMATIQUE CDC, INOVEN, IXIS ASSET MANAGEMENT, KELKOO, KOELIS, LADINAL Groupe SAFRAN, MACDONALD'S France SERVICES, MACGLAN INGENIERIE, MARDEN PRODUCTS, MEDIA FRANCE, MEDIA INTERNATIONAL, MENTOR GRAPHICS LTD, MICROSOFT, MISES BANKING CONSULTING, MISES INTERNATIONAL BANKING SYSTEMS, MURFX, NATIXIS, NFRA RHONE ALPES, NFUII LY OPSTION, NFXFO CONSULTING, NIRVA SYSTEMS, NORTEL NETWORK, O2 INFORMATIQUE, ORPHEUR TECHNOLOGIES, OPEN WDC SA, OPTALAN, ORACLE FRANCE, OTO RESEARCH, PGDI, PHOTOMATON, PROCALI, QUARTZ DATA SYSTEMS, RANDSTAD, RENAULT, RENAULT TRUCK, REPUTERS FINANCIAL SOFTWARE, RORFO GESTIONS, SABEM DEFENSE SECURITE, SABEM TELECOMMUNICATIONS, SCAI ABENT DISTRIBUTED TECHNOLOGIES, SCHINDLER ELECTRIC, SEGA, SIGEC RA, SII, SILICOMP INGENIERIE AQL, SMART UP, SNECMA SA, SNECMA SERVICES, SOBT, SOCIETE DES HONDIERES D'USSEL, SOCIETE GENERALE, SOGETI, SOPHIS TECHNOLOGY, SOPRA, SI MICROELECTRONICS, STAEELIS, STERIA, STEVA, SULLY GROUP, SUMMIT SYSTEMS, SUN MICROSYSTEMS FRANCE, SUPPORT TECHNOELEMENT PROGRAM, SWORD SA, SYNCHROTRON SOLEIL, SYSTEMEL, TEAM TRADE, TEAMLOC, TELECOM, TELINDUS SA, TEUCHOS EXPLOITATION, THALES, TOTAL FRANCE, TRADING PLACES, TRANGO VIRTUAL PROCESSORS, TRILOGIE SA, UNILOG ITS, UNION DE CREDIT POUR LE BATIMENT, VISIORI DRF, VOLURII, WIND RIVER, XPROX, XII INX INTERNATIONAL INC.

<http://ensimag.grenoble-inp.fr>

ÉCOLE NATIONALE SUPÉRIEURE  
D'INFORMATIQUE ET DE MATHÉMATIQUES APPLIQUÉES

# dossier

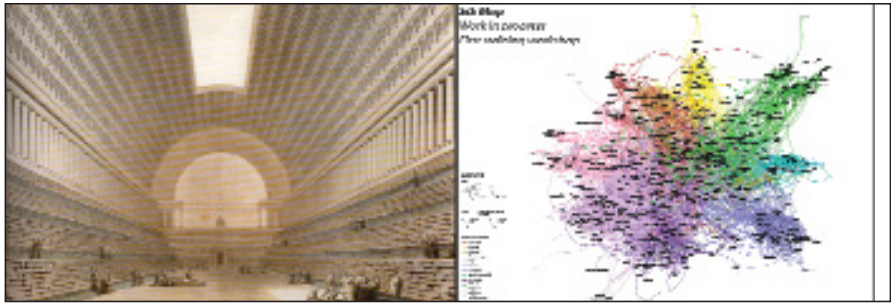
## A L'HEURE DE L'INTERNET GLOBAL

cet exercice patient et minutieux de « traduction » des données statistiques, techniques et formelles en propriétés spatiales puis en modèle général que le web commence à « prendre corps », comme (enfin) rendu visible et appréhendable. Dans ces tâches d'exploration de grandes BDD associées à des systèmes d'information, la construction de modèles théoriques du web peut s'appuyer sur des bases mathématiques et statistiques mais aussi sur la théorie des graphes qui représentent un puissant moyen de réduction de la complexité de certains systèmes. Si les graphes peuvent être mobilisés dans des tâches de calcul, ils peuvent aussi être visualisés, offrant alors des types particuliers de « vues cartographiques » représentatives de la distribution de l'information sur le web, à grande échelle comme localement.

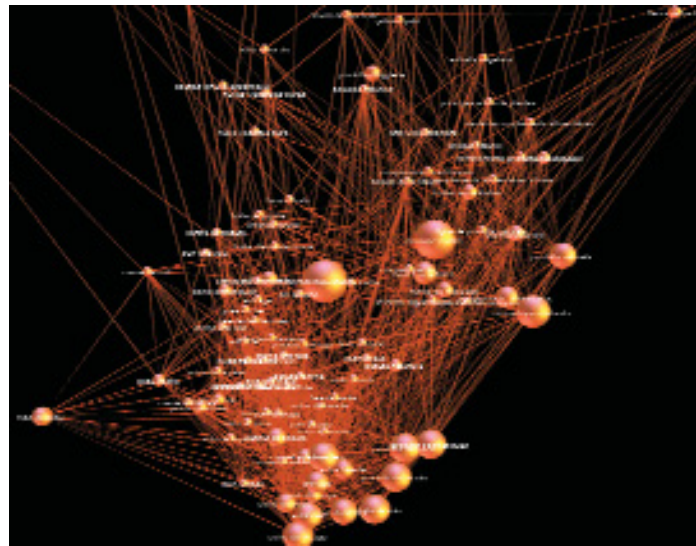
### A la recherche de patterns

Les milliards de pages qui constituent le web peuvent être indexées (décrites) et comparées par leurs contenus ou les liens hypertextes qu'elles distribuent entre elles, formant de vastes systèmes que seule la théorie des graphes peut aider à réduire. C'est en projetant les pages (ou leur « contenu ») dans un espace théorique (à titre de noeuds) ainsi que leurs liens réciproques (à titre d'arcs ou de vecteurs) que l'on peut alors comprendre les propriétés topo-logiques du web comme système d'information. Les données extraites du web sont traitées sous forme de matrice de graphes, laissant ainsi apparaître aux chercheurs certaines propriétés de cet espace documentaire inédit, propriétés que l'on peut aussi visualiser.

Techniquement, les données traitées dans cette ingénierie d'exploration des structures du web sont constituées surtout par des indicateurs de contenu associés à des URL (quels sites parlent de quoi ?) et par la distribution entre ces dernières des liens hypertextes (les sites qui parlent de la « même chose » sont-ils reliés entre eux ?). Dans ce type d'approche, un effort particulier porte donc sur l'étude conjointe des degrés de similarité de contenu entre plusieurs (dizaines) milliers de documents et leur



*A gauche, Etienne Louis Boullée, vue de la nouvelle salle projetée pour l'agrandissement de la Bibliothèque Nationale, 1791. A droite, près de 2.000 sites publiés sur le web européen et traitant des problématiques « science dans la société » (WebAtlas). Le web ne ressemble pas à une bibliothèque et la logique de distribution de l'information n'épouse pas la rigueur des arborescences ou de hiérarchies encyclopédiques. Cette architecture distribuée n'est pourtant pas non plus gouvernée par le hasard et présente des « strong regularities » qu'il faut expérimentalement arriver à décrire. L'étude des propriétés natives de ces territoires numériques peut permettre d'adapter les technologies de search existantes ou de comprendre la nature de certains phénomènes comme la propagation d'une information ou la constitution de communautés sociales sur le réseau.*



*Dans certaines conditions, on peut apercevoir le regroupement de certains mots ou expressions en projetant sous forme de graphe leurs relations statistiques d'occurrence (tel mot apparaît sur le web plus souvent avec tel autre). Ici, ont été visualisées en trois dimensions les relations de hiérarchie (grosseur du nœud suivant l'importance du mot dans le corpus) et de proximité (ou de « complémentarité » thématique) des expressions principales de près de 27.000 pages consacrées aux cosmétiques : radicaux libres, principes actifs, cosmétiques naturels, produits cosmétiques, produits de soins, peaux sèches, produits de beauté, produits de soin, soins du visage... On peut tirer de là des indicateurs puissants sur la façon dont les internautes s'emparent de certaines thématiques, et comment ils le font.*

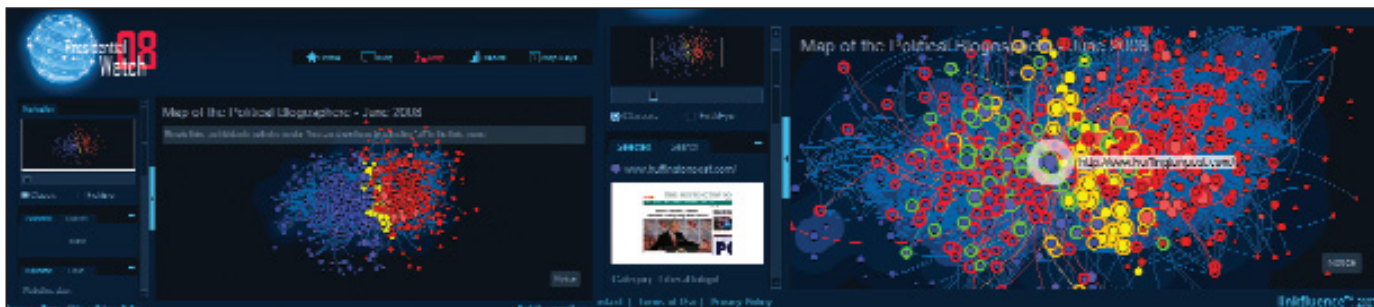
proximité topologique en termes de liens hypertextes. La production de « cartographies du web » suppose donc la fusion de plusieurs traditions scientifiques : celle de l'indexation des documents et des moteurs de recherche, celle de la théorie des graphes (qui a commencé dès le XVIII<sup>e</sup> siècle !) et celle,

enfin, des techniques d'Information Visualisation (InfoViz) et des interfaces destinées à mapper de grands ensembles complexes de données.

La visualisation de grands graphes révèle ainsi quelques propriétés structurelles du web fascinantes. Certains patterns ou schèmes organisationnels de ce sys-

# dossier

## A L'HEURE DE L'INTERNET GLOBAL



Des dispositifs expérimentaux d'exploration dynamique de grandes masses de données existent aujourd'hui. Une interface de type cartographique permet de jouer sur les niveaux de zoom, de visualiser différents niveaux et type de liens entre les documents et, aussi, de tracer l'évolution temporelle du système mappé (ajout/suppression de liens hypertextes parmi ces quelques 2.000 blogs, propagation d'une information via l'indexation des posts et commentaires...) sans oublier l'accès direct aux URL. Ici, le site de veille des blogs dédiés aux élections américaines développé par la SAS compiègnaise R.T.G.I. (<http://presidentialwatch08.com/>)

tème distribué et ouvert sont aujourd'hui connus et testés sur des jeux de données variés, d'autres restent à découvrir. Les questions de « couches », de « hauteur », voire de « profondeur » y tiennent une place déterminante mais aussi « d'orientation » et de « concentration », comme dans une physique des phénomènes magnétiques. Les principes qui permettent de décrire les lois de cette architecture documentaire constituée de dizaines de milliards de pages comme ceux qui président à la construction de « cartographies du web » reposent sur un ensemble d'hypothèses issues de la communauté scientifique (essentiellement américaine) dès les années 1996-1997. Les notions de « distance », « centralité », « éloignement », « proximité », « voisinage », « concentration » ou de « couches » s'enracinent dans une série de travaux orientés vers l'exploration semi-automatique de l'architecture documentaire du web. Dans ces contributions récentes se mêlent aussi bien des questions d'algorithmes de traitement des données que des méthodes d'extraction de l'information via des crawling-systems ou bien encore d'optimisation de process déjà existants comme les grands systèmes YAHOO!, Google ou le projet IBM-Webfountain. Cette lignée de contributions théoriques et de comptes-rendus d'expérimentations esquisse une possible géographie générale du système. C'est le cas, en particulier, des travaux de J. Kleinberg sur l'organisation générale des corpus de pages web distribués en agrégats dotés de « cœurs » où sont concentrés les meilleurs scores de

Hubs/Authorities. HITS (Hypertext Induced Topic Search) a été conçu pour détecter ces cœurs de ressources à partir de l'analyse de leurs liens réciproques à des échelles réduites ou moyennes du web (2). On parle ainsi maintenant couramment « d'agrégats de documents web » pour qualifier ce processus de concentration thématique et hypertextuelle à l'œuvre sur le web. Davison (IBM) développe le concept de « topical localities » dès 2000 (3) tandis que l'on commence à explorer la morphologie générale du système en 1998 avec la publication du fameux modèle du web « en nœud papillon » par A. Tomkins et al. (4). Ces premiers travaux pionniers ont donné naissance à une famille d'algorithmes et de méthodes qui portent

(2) Jon M. Kleinberg, "Authoritative sources in a hyperlinked environment", *Journal of the ACM (JACM)*, Volume 46, Issue 5 (September 1999)

(3) Brian D. Davison, "Topical locality in the Web", *Proceedings of the 23rd Annual International Conference on Research and Development in Information Retrieval, SIGIR 2000*, ACM-Press.

(4) Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, D. Sivakumar, Andrew Tomkins, Eli Upfal, "Graph structure in the web: experiments and models", <http://www.almaden.ibm.com/webfountain/publications/./resources/TheWebas-a-Graph.pdf>

(5) <http://www.nd.edu/~alb/>

(6) Duncan Watts, *Six Degrees : The Science of a Connected Age*, W. W. Norton & Company; 1st edition (February 2003)

aussi bien sur les techniques d'extraction des données web (crawling systems) que sur leur description (indexation) ou bien même leur visualisation (web information visualization).

Les questions de topologie documentaire sur le web (ou dans les grands systèmes comme certains intranets d'entreprise) occupent donc une place centrale depuis lors en termes de recherche et développement. La communauté des « explorateurs du web » est actuellement concentrée sur deux problématiques essentielles. Celle de la temporalité ou l'étude des phénomènes dynamiques sur le web, à laquelle a grandement contribué V.-L. Barabasi et son équipe de l'Université de Notre-Dame en proposant des modèles de scénarios évolutifs (5). De son côté, D. Watts à l'Université de Columbia développe une série d'hypothèses sur les scénarios de propagation de l'information dans des grands systèmes comme le web (peut-on modéliser la propagation des virus informatiques ? Celle d'une rumeur sur un réseau de blogs ? Quand et comment un système comme le web vit-il certaines phases critiques de mutation ?...) (6).

### Enjeux d'une « géographie de l'information »

Les méthodes et les outils de web-mapping concentrent des enjeux importants en termes de recherche et d'innovation. On peut les ranger parmi les technologies de knowledge mapping et de knowledge discovery et incarnent les efforts réalisés par les ingénieurs et les chercheurs pour comprendre les règles qui

# dossier

## A L'HEURE DE L'INTERNET GLOBAL



La construction d'une cartographie réclame plusieurs étapes.

1. Choix des URL et indexation de leurs contenus et de leurs liens hypertextes réciproques

2. Les données sont traitées sous forme de matrices de graphes (nœuds = sites, arcs = liens hypertextes)

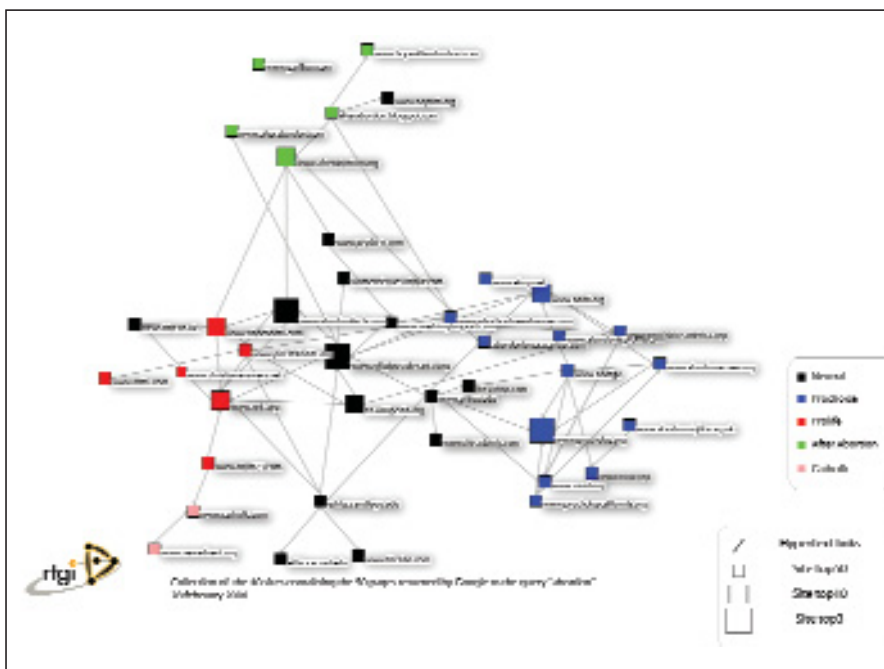
3. Le graphe est spatialisé dans le logiciel Graphiltre. On applique une force aux nœuds-sites pour qu'ils se repoussent tandis que les arcs-liens les retiennent : le graphe prend forme.

4. Mise en forme graphique : les sites les plus connectés sont mis en hauteur et rendus plus gros. Des couleurs sont choisies pour les différents types de sites. Un angle de vue est choisi pour la « scène 3D ».

5. Rendu : La scène contenant le graphe en 3D est transformée en image avec Persistence of Vision. Les nœuds sont peints comme des boules transparentes colorées et les liens comme des rubans pointus.

6. En indexant régulièrement les données, on peut ainsi mapper le réseau des adhérents à l'APRIL et comparer à différentes étapes l'organisation des données.

gouvernement ce vaste système que l'on décrivait jusqu'à présent comme composé de données « en grandes masses, dynamiques et peu structurées ». Leur première valeur est d'ordre expérimentale pour les chercheurs : les « vues » du web constituent des modèles de synthèse théorique et des instruments de concentration de l'expertise. Mais les cartographies ne sont jamais stables, non seulement parce qu'elles évoluent dans le temps, mais aussi parce qu'elles sont numériquement manipulables. C'est en faisant varier, à partir d'un modèle de départ, les traitements sur les données et les types de visualisation qu'apparaissent les



Résultats de la requête « abortion » sur YAHOO ! spatialisés sous forme de graphes. Alors que dans l'interface du moteur les résultats sont affichés par défaut sous forme de listes à plat, un graphe des liens hypertextes montre lui des affinités communautaires évidentes : en bleu les « pro-choice » défenseurs du droit à l'avortement aux Etats-Unis et, d'un autre côté, les « pro-life » ardents défenseurs du droit « à la vie ». Cette cartographie montre à quel point les documents web sont organisés en agrégats, parfois très compétitifs.

patterns les plus stables. Pour les scientifiques, produire des « formes » visuelles du web c'est donc entrer dans un dialogue intense entre modèles théoriques possibles et sessions expérimentales de relevé et de traitement des données : projections graphiques des connaissances acquises elles sont aussi comme un concentré des pistes et des hypothèses encore à parcourir. Les cartographies de documents web peuvent aussi préfigurer de nouvelles méthodes ou de nouveaux algorithmes qui viendront améliorer les systèmes d'information web comme les moteurs de recherche. La manifestation visuelle de propriétés statistiques de grands corpus de pages ou de sites permet de mettre à jour des contraintes auxquelles seront confrontées les techno-

logies et les méthodes d'extraction comme d'indexation de documents qui se révèlent regroupés en agrégats thématiques (donc éventuellement d'opinion, de communautés sociales ou d'intérêts utilisant chacune leur propre vocabulaire ou folksonomies...). Au-delà de ces « cœurs » thématiques peuplés de mots et de liens hypertextes, à moyenne échelle, les cartographies manifestent des lignes de fracture entre corpus (boundaries), des niveaux d'inclusion, de complémentarité ou de hiérarchie, laissant apercevoir les linéaments d'une géographie générale de l'information sur le web. En termes d'outils de recherche avancés pour l'utilisateur ou de services pour des experts, elles laissent aussi préfigurer de nouvelles interfaces pour la recherche d'information. En retour, la recherche de patterns dans des données indexées contribue souvent à la valorisation du patrimoine informationnel des entreprises puisqu'elles révèlent souvent des principes d'organisation jusque là implicites, pour tout dire invisibles.

Enfin, il paraît évident que les cartographies dynamiques de grandes masses de données web contribuent à l'amélioration des systèmes de veille stratégique et constituent donc aussi de puissants outils d'aide à la décision



## Utilisation des réseaux sémantiques dans la recherche Web

Par Gilles VANDELLE

Dirigeant équipe Search Backend Platform dans le centre de R&D de Yahoo

### Les bases du Web Search

Cela fait plus de 10 ans que les grands moteurs de recherche sont apparus sur le Web. Une grande partie des efforts a été consacrée à conserver une grande pertinence et fraîcheur des données malgré une croissance exponentielle des volumes et l'explosion de la fraude. Mais, en fait, la recherche se fait toujours selon le même principe. On extrait tous les documents qui contiennent les mots-clés de la requête (ou un grand nombre quand ceux-ci sont trop nombreux) et on applique une fonction de pertinence ou « ranking » à chacun d'eux de façon à pouvoir les ordonner. Pour ce faire, on

**Gilles Madelle** dirige l'équipe Search Backend Platform dans le centre de R&D de Yahoo basé à Grenoble. Diplômé de l'Ecole Centrale de Paris en 82 il a participé à la création de startups en France et en Californie. Il travaille dans le domaine de l'internet depuis 1998 où il s'est spécialisé sur les techniques de machine learning.

associe à chaque document un vecteur formé par ses termes pondérés par une valeur. La formule la plus classique est appelée  $tf \cdot idf$  pour « term frequency x inverse document frequency »

$$w_{t,d} = tf_t \cdot \log \frac{|D|}{|\{t \in d\}|}$$

On considère que plus un terme est fréquent dans l'index moins il est important. A noter que cette pondération a été largement révisée car les spammers créaient des documents avec un nombre artificiellement élevé de termes de façon à augmenter le «  $tf$  » de leur document. Pour mesurer la similarité entre la requête et un document, on calcule le cosinus de l'angle formé par ces deux vecteurs.

$$\text{Sim}(Q, D_i) = \frac{\sum_j w_{Q,j} w_{i,j}}{\sqrt{\sum_j w_{Q,j}^2} \sqrt{\sum_j w_{i,j}^2}}$$

En fait, cette approche pose un problème car deux mots différents sont considérés comme des dimensions différentes du modèle et donc orthogonales. Par exemple, un document contenant UEFA mais pas football ne sera donc pas considéré comme un bon résultat pour la requête « football ».

### Introduction de la sémantique

Depuis plusieurs années les laboratoires travaillent sur l'introduction d'un peu de sémantique dans les moteurs de recherche. On peut distinguer deux types d'approches : ascendantes et descendantes. La première part du document d'origine et pousse à le compléter avec des informations qui permettent de mieux l'analyser et donc de mieux l'indexer. L'autre, à l'inverse, part de la requête en essayant de mieux la comprendre afin de promouvoir des résultats plus pertinents.

L'enrichissement des documents paraît en effet une bonne idée. Une proposition appelée micro-formats propose au créateurs de document HTML d'ajouter des balises qui ne serviront pas cette fois-ci à modifier l'affichage mais à mieux qualifier le texte qu'elles délimitent. Génial ! Seulement voilà, avant d'avoir une couverture suffisante du web, cela va prendre longtemps, très longtemps. Des approches qui s'appuient sur un enrichissement automatisé du contenu semblent donc nécessaires si l'on veut arriver rapidement à une grande couverture du Web.

L'approche descendante qui essaie de comprendre la requête est complémentaire de la précédente. Premièrement, on essaie de segmenter la requête. Par exemple la requête « restaurant chinois grenoble » pourra être découpée suivant le schéma « quoi / où », une fois identifié

que grenoble est bien une ville. Attention, toutefois aux faux amis comme « crème chantilly ».

## Représentation des connaissances

Pour effectuer toutes ces analyses, il est donc nécessaire de disposer de vastes bases de connaissances sous forme d'ontologies. Plusieurs projets existent dans ce domaine comme, par exemple, le langage OWL du W3C ou le projet dbpedia. Mais, la création d'ontologies pose certains problèmes. Premièrement, le web est un media vivant et il faut constamment faire des mises à jour. Deuxièmement, une approche trop théorique doit se plier à ce qui est communément admis dans la langue courante. Par exemple, une « machine à laver » est un « lave linge » et non un « lave vaisselle ». Quand on dispose d'une telle base, on peut améliorer notre système de recherche de documents. Pour chaque document, on sélectionne les concepts qui le représentent le mieux. On fait, de même, avec la requête. Notre fonction de similarité se voit donc enrichie de dimensions qui tiennent, cette fois-ci, compte de la sémantique.

## Web mining ou comment apprendre à partir du web

Comme le coût de création et de mise à jour des bases de connaissance devient vite prohibitif, on a pensé à automatiser ce processus. De nouvelles approches permettent de les fabriquer automatiquement à partir de connaissances présentes sur le web. On appelle cela web-mining. C'est un domaine passionnant qui permet d'extraire de cette immense source non structurée qu'est le Web des données utilisables pour l'analyse sémantique. L'enjeu est immense car il permettra d'appréhender non seulement la dimension multilingue et multiculturelle du web mais aussi son évolution rapide. De quoi s'agit-il ? Derrière la grande idée « extraire des connaissances du web », on peut différencier

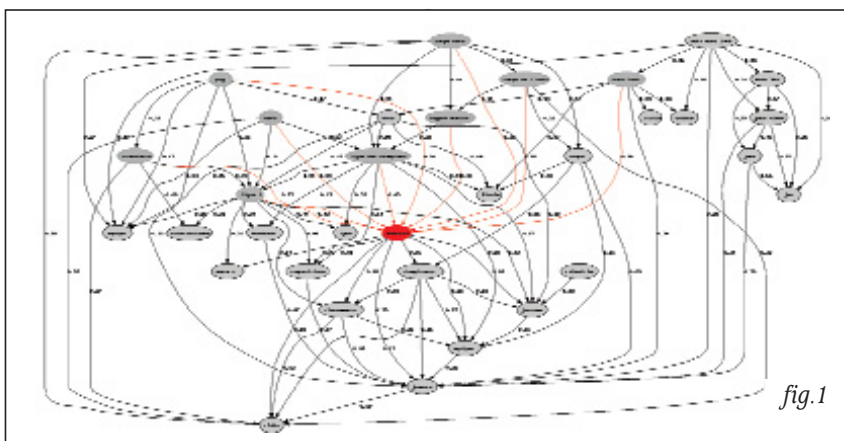


fig.1

deux familles. La première, extrait des informations de la structure même de « la toile » en explorant les liens entre les pages et les sites, la seconde s'intéressant plutôt aux concepts contenus dans ces pages.

Les liens entre les pages d'un même site et encore plus entre les pages de sites différents nous apportent beaucoup d'informations. Le texte (href) porté par ces liens est largement utilisé par les moteurs de recherche pour promouvoir une page lors de l'indexation. Certains vont bien au delà en proposant des cartographies thématiques du web. Cela permet d'identifier les courants de pensée, les vecteurs d'influence et leurs évolutions. Il n'est pas étonnant de voir les spécialistes du marketing se pencher de plus en plus sur ces études pour analyser, par exemple, l'image portée par une marque. L'autre approche consiste à bâtir des réseaux de concepts. En premier, il faut associer à chaque page un certain nombre de concepts représentatifs. Chaque concept est associé à une ou plusieurs listes de mots. Une fois ceci réalisé, on utilise les co-citations ou les coréférences entre ces concepts pour construire le graphe. Le graphe ainsi obtenu porte souvent un grand nombre de défauts dus entre autre à l'ambiguïté de la langue. Par exemple, « Paris Hilton » et « Ville de Paris » risquent d'être reliés par le terme « Paris ». Des analyses en composantes connexes et le clustering permettent d'éliminer un bon nombre de ces défauts.

Les graphes ainsi construits ne sont pas aussi rigoureux que les ontologies

construites par des experts mais représentent bien les relations entre les termes. La Fig1 donne le graphe fabriqué pour analyser la requête « UEFA ».

## les réseaux sociaux

Certains prévoient déjà le coup suivant. A la base, un constat simple : ce qui est un bon résultat pour moi ne l'est peut être pas pour quelqu'un d'autre. Il faut donc personnaliser les résultats. L'exploitation du profil des utilisateurs semble prometteur seulement voilà, la plupart d'entre nous rechigne à répondre à des questionnaires plus ou moins longs sur leurs centres d'intérêts. Il faut donc trouver des sources moins explicites. Une approche plus sournoise consiste à stocker les « clics » et conserver la trace laissée par chacun d'entre nous. Ce type d'approche est contestable d'un point de vue éthique sauf si l'on agrège les données individuelles au niveau d'un groupe. Nous voyons donc poindre de plus en plus de solutions basées sur des réseaux sociaux. Des solutions comme delicious.com permettent de partager ses tags avec la communauté et de mettre en valeur les pages taguées par les personnes de son réseau.

Dans les années qui viennent, nous allons voir encore progresser ces techniques de recherche d'information non seulement au niveau du web search mais également dans les moteurs de recherche spécialisés dans les données de l'entreprise.

\*



## Web Services et Grid Computing : la nouvelle ère des systèmes distribués

Par Rémy Amouroux,  
Ensimag 1992, TEOREM - co-fondateur de Kelkoo

*Le Grid Computing ou Cloud Computing est le dernier aboutissement d'une longue évolution d'un concept assez ancien : cela fait référence à l'utilisation de la mémoire et des capacités de calcul des ordinateurs et des serveurs répartis dans le monde entier, et liés par un réseau, tel Internet. Les utilisateurs de la grille ou du nuage pourraient ainsi disposer d'une puissance informatique considérable et modulable.*

**Rémy Amouroux** est un des co-fondateurs de Kelkoo et l'inventeur de la technologie qui a permis la création. Chez Kelkoo, il a dirigé et orienté le développement d'applications Internet s'adressant aussi bien au public qu'à ses clients, jusqu'à l'acquisition de la société par Yahoo en 2004. Il est devenu Chief Architect pour l'Europe chez Yahoo, supervisant le développement d'applications destinées à l'une des plus grosses audiences du Web. En 2008, il a créé Teorem, un cabinet de conseil, afin d'aider les entreprises à dépasser les obstacles techniques empêchant leur croissance.

### S'affranchir du matériel

L'idée d'utiliser les ordinateurs alors même qu'ils réalisent déjà une tâche plus ou moins gourmande et de regrouper cette puissance grâce à Internet n'est pas nouvelle. Entre 1997 et 2000, les projets de calculs distribués Distributed.Net<sup>1</sup> (répondre au concours lancé par la société RSA pour casser des clés de cryptage) ou Seti@Home<sup>2</sup> (analyse des signaux électromagnétiques pour découvrir ainsi de la vie extra-terrestre) ont démontré que le concept était valide.

Le rapport coût/puissance des ordinateurs a continué à évoluer dans un sens positif ces dernières années. Le coût principal des infrastructures n'est donc pas dans le matériel, mais dans tout ce qui l'entoure : l'hébergement, le réseau, mais aussi tout ce qui va permettre la surveillance et la réparation de ces machines. Ces coûts évoluent en fonction du nombre de machines installées.

Quelques soient les besoins (production, développement et tests, data-mining), les machines ne sont pas utilisées à leur pleine capacité : un site web n'utilisera la pleine puissance de ses serveurs que lors de pics de trafic, les machines de développement ne fonctionneront que pendant les heures de travail ou pour de nouvelles versions à tester. Cela pose le problème de plus en plus sensible de l'impact écologique et cela fait enrager tout gérant de société voyant qu'une partie de ses frais fixes ne servent en fait à rien du tout.

Il est donc extrêmement important de mettre en place des mécanismes permettant d'optimiser l'utilisation des machines en tenant compte du temps d'utilisation, en ayant la capacité d'agrandir ou de réduire la voilure de manière souple et rapide.

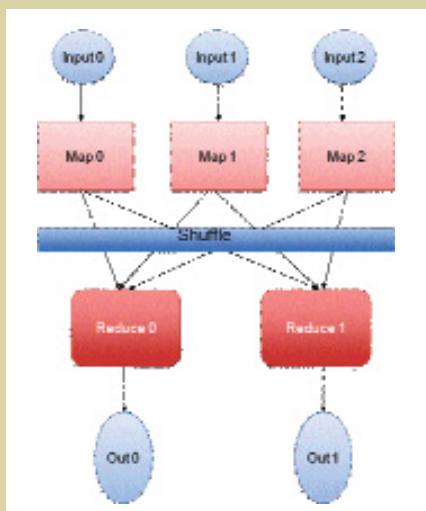
### La virtualisation mise à profit

Les progrès des systèmes de virtualisation sont en partie à l'origine de la montée en puissance des premières offres de Cloud Computing. Dans ce domaine, la technologie open-source Xen est utilisée pour répondre à ce besoin de flexibilité.

La société Gandi, connue en tant que Domain Name Registrar, base son offre d'hébergement sur cette technologie et les outils qu'elle a développés autour de celle-ci. Gandi vous propose<sup>3</sup> de louer des parts de serveurs puissants et vous



## LE MOTEUR MAP/REDUCE



Le mécanisme Map/Reduce demande aux développeurs de définir un certain nombre de fichiers comme input (ceux-ci étant stockés dans HDFS) et deux fonctions, l'une pour la phase Map, et l'autre pour la phase Reduce.

La phase d'Input du moteur se charge de séparer les fichiers en parts plus ou moins égales et de les affecter à des tâches Map.

La fonction définie en tant que Map est exécutée durant cette phase pour transformer une entrée considérée comme un ensemble de paires clé-valeur en un autre ensemble de paires clé-valeur. Par exemple, une ligne de log d'un site web sera considérée

comme une seule paire et pourra être transformée en ses différents constituants : IP du client, page accédée, code http, etc.

La phase Shuffle (fournie par Hadoop) trie les paires clé-valeur et les redistribue vers les tâches Reduce. La fonction définie pour cette phase regroupe les paires ayant la même clé pour en créer de nouvelles. Par exemple, c'est dans cette phase que l'on pourra compter le nombre de hits par seconde dans le cas de l'analyse des logs d'un site web.

La dernière phase est en charge de l'écriture de fichiers de sortie, en général vers HDFS.

Le moteur est en charge de créer et de distribuer les tâches sur le cluster utilisé. Par défaut, il crée autant de tâches Map qu'il y a de blocs HDFS à traiter, mais cela peut être contrôlé si l'on définit la taille minimale de partition des fichiers. Le nombre de tâches Reduce dépend de la taille des données à traiter bien sûr, mais le moteur va essayer d'utiliser le plus possible de nœuds disponibles dans le cluster.

permet de moduler, en fonction de contraintes horaires, le nombre de serveurs et le nombre de parts (donc la puissance) effectivement loués, sans que cette montée en puissance nécessite de redémarrer les applications. L'autre point intéressant de cette offre est l'assurance qu'en cas de problème sur la machine sur laquelle s'exécutent vos parts, ces dernières sont automatiquement transférées sur d'autres machines avec le moins d'impact possible pour les utilisateurs.

Un autre utilisateur de la technologie Xen est le système EC2<sup>4</sup> d'Amazon. L'offre vous permet de créer des images de votre serveur contenant les

applications et les données qui sont nécessaires à son fonctionnement, tout en paramétrant le processeur, la mémoire et le stockage qui leur permettront de s'exécuter de manière optimale. Une fois l'image créée, il est possible de la démarrer et elle se comporte comme un serveur réel. Amazon vous fera payer le temps d'exécution (l'unité étant l'heure) et les transferts de données entrant et sortant du réseau d'Amazon. A cette modularité des coûts qu'apprécieront les responsables financiers s'ajoutent trois fonctionnalités importantes qu'apprécieront les concepteurs d'applications. Pouvoir définir des images des serveurs

à l'avance permet de déployer très rapidement de nouvelles instances, mais facilite aussi grandement l'automatisation de tests. La gestion par web services du déploiement des images rend possible l'automatisation de la taille de la plateforme en fonction des besoins réels, comme le montre ScalR<sup>5</sup>. Le service d'Availability Zone, associé au service d'Elastic IP Address, permet de répondre aux besoins de haute disponibilité.

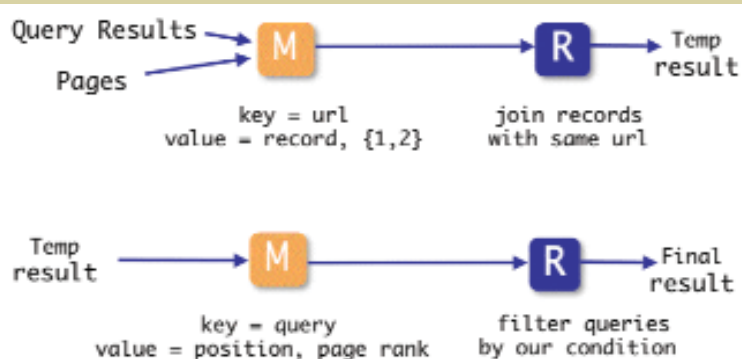
Les nombreux services fournis par Amazon sont une première base pour s'affranchir des problèmes matériels mais ils restent au niveau du serveur : le développeur doit toujours concevoir des applications scalables, résilientes et dont l'architecture est évolutive et modulaire pour en profiter.

### Aller plus loin avec Hadoop<sup>6</sup>

En juillet dernier, Yahoo, Hewlett-Packard et Intel ont annoncé l'initiative Cloud Computing Test Bed. Chacun des 6 centres de recherches associés au projet aura à sa disposition une ferme de 1000 à 4000 machines afin de procéder à des expériences de Cloud Computing. Le matériel sera d'origine HP et Intel. Yahoo apportera son savoir-faire autour du projet Hadoop. Ce dernier est issu de la fondation apache ; il fut créé par Doug Cutting, le développeur à l'origine du projet Lucene, pour supporter la distribution dans le moteur de recherche Nutch. Il est composé d'un moteur au dessus d'un système de stockage. HDFS (Hadoop Distributed File System) stocke les données sur lesquelles travailleront les applications. Il assure leur distribution sur plusieurs nœuds et leur haute disponibilité. Les nœuds utilisent le réseau pour servir des blocs de données et ils se parlent pour assurer une bonne distribution des blocs et assurer que leur répliquon reste à un haut niveau. Le moteur est une implémentation du modèle de programmation Map/Reduce, que Google a présenté en 2004 à la conférence OSDI<sup>7</sup>. C'est un framework, écrit en Java, permettant le traitement distribué de très gros jeux

## PIG : LARGE-SCALE DATA ANALYSIS<sup>8</sup>

Query Results			Pages	
query	position	url	url	pagerank
"news"	1	cnn.com	bbc.com	0.9
"news"	2	bbc.com	cnn.com	0.9
-	-	-	pigwheels.com	0.3
"pig"	1	pigwheels.com	-	-
-	-	-	-	-



```

query_results = load '/data/query_results';
pages         = load '/data/pages';

A = join query_results by url, pages by url;
B = group A by query;
C = filter B by checkTop5(A);
D = foreach C generate A.query;

store D into '/data/interesting_queries';
    
```

Pig est un environnement de programmation fournissant aux développeurs un langage de haut niveau (Pig Latin) pour exprimer des traitements de données. Une fois un programme écrit, il peut être compilé et être exécuté sur un cluster Hadoop. Pig Latin est présenté comme étant à Hadoop ce que SQL est aux bases relationnelles. Un de ses points forts est que sa structure est très facilement parallélisable, tout en laissant cet aspect de la programmation complètement transparent pour le développeur. Il simplifie aussi la vie du développeur en lui permettant de définir en un seul programme ce qui correspond à plusieurs itérations map-reduce.

Si l'on cherche, par exemple, à partir des deux tables Query Results et Pages (voir ci-dessus) quelles sont les requêtes pour lesquelles la page de plus haut PageRank n'apparaît pas dans les cinq premiers résultats, le résultat Map-Reduce devra avoir deux itérations (voir ci-dessous), ce qui implique au minimum quatre classes Java plus ou moins complexes, plus la coordination de l'exécution de ces deux itérations.

Cela s'écrit en Pig Latin de la manière ci-après, sous une forme à la fois procédurale et déclarative. Nous avons là avant tout une suite de transformations facilitées par les opérateurs et les fonctions fournis par cet environnement.

Il est à noter que le modèle de données est assez évolué (bien plus que de simples paires clé-valeur) et que les fonctions disponibles sont assez complètes (opérations mathématiques, mais aussi somme, moyenne, etc) et qu'il est possible d'insérer ses propres fonctions dans le système. Il est très intéressant de voir aussi que la traduction de requête SQL en Pig Latin est somme toute assez facile.

de données, vus comme des ensembles de paires clé-valeur (voir l'encadré). Map/Reduce est un modèle de programmation simple utilisé pour résoudre de nombreux problèmes. Le traitement de fichiers log de toutes sortes vient naturellement à l'esprit, mais le projet Nutch a démontré que ce modèle était aussi très utile dans le cadre d'un moteur de recherche. C'est aussi un modèle au dessus duquel il est possible de construire des modèles plus complexes (voir l'encadré sur PIG). Bien qu'écrit

en Java, le framework permet d'utiliser des logiciels externes au travers de ses APIs dites de Streaming. L'un des avantages d'Hadoop est qu'il connaît l'emplacement des données dans HDFS et en tient compte dans la distribution des tâches sur les machines. Il est aussi capable dans sa gestion des tâches de ré-exécuter celles qui ont pu avoir un problème. De plus, sachant que dans un cluster les machines ne sont jamais complètement homogènes, le framework va utiliser la puissance disponible pour exé-

cuter la même tâche plusieurs fois, et utiliser le résultat de celle qui se finit en premier et interrompre les autres.

### L'avenir est aux applications distribuées

L'un des points forts de ce que Tim O'Reilly a appelé le Web 2.0 est la prolifération des interfaces de programmations pour des services web. Ceux-ci ont mis à la disposition des développeurs de nombreux services réutilisables, et

# dossier

## A L'HEURE DE L'INTERNET GLOBAL

en conjonction avec les techniques AJAX, nous avons vu l'apparition de nombreux mash-up combinant ces différents services pour en créer de nouveaux. Il est maintenant inimaginable d'avoir son propre système de cartographie pour un site web : on utilisera naturellement l'un des fournisseurs existants comme Mappy, Google ou Yahoo, services dont l'intégration sur un site est extrêmement aisée.

Cette approche de la conception d'applications Internet amène à la distribution sur plusieurs plateformes de ce qui compose réellement le service rendu à l'utilisateur final, celui-ci n'ayant plus aucune notion de ce qui est exécuté sur sa propre machine, sur les machines du site auquel il accède ou sur les machines de fournisseurs de celui-ci.

Les offres de Grid Computing associées aux technologies permettant la distribution d'une application au travers d'Internet vont encore accentuer la

dématérialisation et la virtualisation de celle-ci. Elles vont abaisser encore la barrière à l'entrée pour tous les développeurs ayant des idées innovantes à

réaliser et à faire partager, ce qui ouvrira une nouvelle ère d'opportunités très intéressantes. \*

## INDEX

1. Distributed.Net : projet de calcul distribué, <http://www.distributed.net/>
2. Seti@Home : projet de calcul distribué, <http://setiathome.berkeley.edu/index.php>
3. Gandi Hébergement avec Xen : <https://www.gandi.net/hebergement/offre/xen/>
4. Amazon Elastic Compute Cloud : <http://aws.amazon.com/ec2/>
5. Scalr : <http://code.google.com/p/scalr/> et <https://scalr.net>
6. Apache Hadoop : <http://hadoop.apache.org/>
7. MapReduce: Simplified Data Processing on Large Clusters \_par Jeffrey Dean et Sanjay Ghemawat, <http://labs.google.com/papers/mapreduce.html>
8. Pig, a language for large data analysis, <http://incubator.apache.org/pig/>



**cimpa**

- Process
- Configuration de Produit
- Applications et Methodes
- Software
- Simulation et Visualisation
- Entreprise Etendue
- In Services
- Formation et Training

**ILS NOUS ONT FAIT CONFIANCE**  
Airbus, IXXIS, L'Imagerie, Airbus, Siemens VIXI, Intel, IBM/berkeley, CEV, Renault, Peugeot, Bouygues Telecom, Air France...

## Votre partenaire PLM

Créé en 1995 à l'initiative d'AIRBUS, Cimpa est un acteur majeur et international du PLM (Product Lifecycle Management / Gestion du Cycle de vie du Produit) : maîtrise des technologies informatiques et des systèmes d'information.

Cimpa vous offre **des solutions personnalisées** en Consulting, développement et intégration d'applications, ainsi qu'un support technique et humain, pour chaque étape du cycle de vie de votre produit.

### CAD/CAM - PDM

► Consulting
► Solutions
► Support

**Venez participer à la croissance de CIMPA**  
Adressez votre candidature  
Madame Françoise Soguero (Responsable de Recrutement)  
Email : [contact@cimpa.com](mailto:contact@cimpa.com)



**PARIS**  
1, avenue de la  
Cordillère  
92310  
Sèvres  
Tel: +33 (0)1 29 41 70 00



**TOULOUSE**  
Cordère 1  
4, av. Didier Daurat  
31700  
Blagnac  
Tel: +33 (0)5 61 10 60 85



**BORDEAUX**  
Espace Wilpinux  
37, rue du Bonquet  
33100  
Saint-Aubin-de-Médoc  
Tel: +33 (0)5 47 35 66 00

[www.cimpa.com](http://www.cimpa.com)



## Le vote par Internet : quel avenir ?

Par Benoit Chenon  
Ensimag 1986 Co-fondateur de la Société Voxaly

*Le vote électronique, les boîtiers électroniques, le vote par Internet : tout le monde en parle, chacun a son mot à dire. Mais précisément de quoi parle-t-on ? Quelles sont les élections concernées ? Quelles sont les techniques utilisées et pourquoi ? Quels sont ses avantages ? Comment réussir ?*

### Le vote par Internet

Pour voter, l'usage classique est de se rendre dans l'école primaire la plus proche de son domicile, d'aller dans l'isoloir, de mettre dans l'enveloppe le bul-

#### Benoit Chenon.

Ensimag 1986, fut Architecte senior chez Atos Origin puis Caggemini, sur de grands projets nationaux, spécialisé dans la sécurité des SI et les architectures SOA (Architecture Orientée Service). A ce titre, il participa à la mise en place des premières élections par Internet en 2000 dans le monde bancaire. Co-fondateur de la société VOXALY en 2006, il est actuellement Directeur des Systèmes d'Informations et responsable du pôle recherche et développement chez VOXALY. Au sein de la Fédération Nationale des Tiers de Confiance (FNTC), il anime le groupe de travail sur le vote électronique par Internet.

letin de son choix, puis de mettre le tout dans l'urne en plastique transparent, tout en ayant décliné son identité et signé la liste d'émargement. Cette modalité de vote est nommée « vote physique à l'urne ».

Maintenant, avec la dématérialisation de l'information, d'autres modalités existent :

- Le vote avec boîtier électronique (utilisé par exemple lors des dernières élections municipales dans certaines communes)
- Le vote par correspondance papier
- Le vote par téléphone, avec l'usage d'un serveur vocal interactif
- Et enfin, le vote par Internet

De plus, le vote par Internet est souvent utilisé avec d'autres modalités de vote, notamment le vote par correspondance. Cette mixité des canaux permet d'augmenter les taux de participation en réduisant toute discrimination sur l'usage de l'outil informatique.

### Pour quelles élections ?

Le vote par internet s'adresse à une large variété d'élections tel que par exemple :

- L'élection de vos représentants aux

comités d'entreprises et vos délégués du personnel au sein de votre entreprise

- La désignation des résolutions en assemblée générale
- L'élection de vos représentants régionaux

Par contre, le vote par Internet n'a pas vocation à gérer les votes politiques en se substituant au vote physique, car certaines obligations électorales (présence de l'isoloir, authentification forte de l'électeur, ...) ne peuvent être mises en œuvre dans un budget raisonnable.

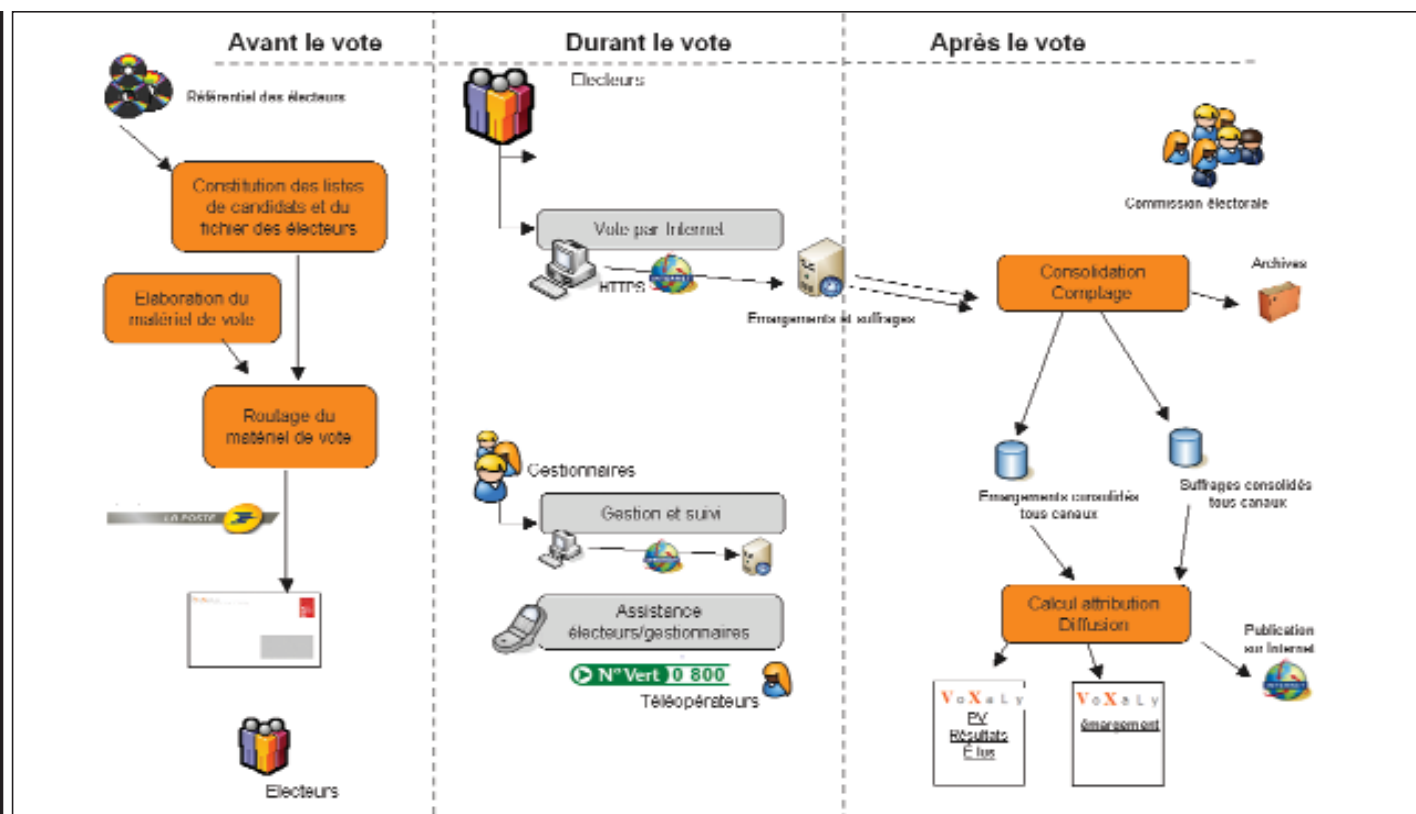
### Les technologies utilisées

Le vote par Internet met en jeu de nombreuses technologies, qui impactent tous les niveaux d'architecture du système d'information, tant applicatif, que logiciel ou matériel.

Les principales technologies utilisées sont :

- Chiffrements pour la protection de la confidentialité (RSA par exemple)
- Empreinte de code et fonction de hash (SHA-1)
- Signature électronique et authentification à base de certificat
- Redondance matérielle pour la haute disponibilité
- Protocole de transport sécurisé (SSL par exemple)
- Cloisonnement applicatif et matériel des données

...



## Des avantages certains

Le vote par Internet propose de nombreux avantages :

- simplifie la mise en place des élections pour l'organisateur en regroupant dans une unique urne virtuelle l'ensemble des différents bureaux de vote,
- est accessible de partout, et à toute heure,
- accélère le dépouillement tout en garantissant la parfaite exactitude des résultats,
- apporte une grande souplesse sur la durée du scrutin en permettant d'étaler des scrutins sur plusieurs semaines, voire plusieurs mois,
- réduit les coûts, tant au niveau de l'organisation que sur les matières premières et l'affranchissement,
- et ainsi participe au développement durable.

Par rapport au vote par téléphone, le vote par Internet

- apporte des niveaux de sécurité supérieurs,
- offre une ergonomie évoluée et plus intuitive,

- et en contrepartie, nécessite un équipement plus coûteux et la maîtrise, même minime, d'un ordinateur.

Des freins pour le vote par Internet ? Oui, pour le moment le principal frein est la confiance à démontrer auprès de l'électeur et des organisateurs.

## Un avenir prometteur

La conception et la mise en place d'un système de vote par Internet est une organisation longue, onéreuse et complexe, faisant intervenir un grand nombre d'acteurs et de prestataires et mettant en jeu des technologies pointues.

Par rapport à la « simple urne en plastique transparent », une telle complexité soulève davantage de doute, et peut être le siège de failles potentielles, d'erreurs ou de malversations.

Comment alors garantir la sincérité des résultats ?

Heureusement pour l'électeur et l'organisateur d'élection, il existe dorénavant un cadre législatif qui détaille les nombreuses exigences auxquelles doit

répondre un tel système de vote par Internet.

De plus, la Commission Nationale Informatique et Liberté (CNIL) a émis de nombreuses recommandations aussi bien au niveau organisationnel qu'au niveau applicatif et technique. Ainsi, le respect de ses différentes recommandations, la définition précise des rôles de chacun et une expertise du système de vote dans son ensemble (pas uniquement de la partie informatique mais incluant aussi l'organisation et les processus) par un organisme tiers indépendant du prestataire permet d'atteindre un niveau de sécurité en conformité avec le code électoral et d'apporter la confiance nécessaire.

Telles sont les conditions indispensables pour réussir une opération de vote par Internet.

Un prestataire répondant à tous ces critères est un Electionneur®. Un tel label garantit le succès des opérations électorales, dans une enveloppe budgétaire attractive.

\*



## Crédit de particulier à particulier L'argent sans la banque ?

par Marc Duleba  
Ensimag 1984 - Analyste ebusiness chez LCL

*Le crédit P2P paraît anecdotique dans le contexte actuel. Cette crise financière brutale révèle la fracture entre une élite qui joue de mécanismes économiques sophistiqués et une population désorientée devant leur complexité. Saurons nous en tirer la leçon et construire un monde économique viable et compréhensible de tout un chacun ?*

### Marc Duleba

Du développement de logiciels telecom au marketing Internet en passant par la direction de projets, je suis aujourd'hui analyste ebusiness chez LCL, en charge de l'innovation et des nouveaux usages de la Banque sur Internet. J'observe et analyse les innovations technologiques et les tendances sociétales dans cet univers en perpétuelle évolution. J'anime aussi un club de foot pour les jeunes un samedi sur deux. « L'important n'est pas d'être sur le palet mais là où le palet sera »  
Proverbe de Hockey Suédois.

### Quand les réseaux sociaux investissent la finance

#### L'Internaute veut ses finances en ligne ...

Les services bancaires de base sont parmi les plus utilisés sur Internet. Aujourd'hui, l'Internaute veut aller plus loin comme le montre le succès des demandes de crédit en ligne. Les établissements financiers s'adaptent à cette nouvelle demande mais peinent à prendre en compte un besoin plus humain d'implication que permettent les plateformes de réseau social (Facebook). Le monde du financement de

#### Crédit P2P n'est pas micro finance

Le crédit P2P diffère de la Micro Finance qui a une vocation plus sociale. Le crédit P2P est un crédit classique pouvant financer n'importe quel type de projet (jusqu'à quelques milliers d'euros sur quelques années), alors que la micro finance est l'offre de services financiers basiques à des pauvres qui n'ont pas accès au système bancaire classique, leur permettant ainsi d'améliorer leurs conditions de vie : les prêts consentis sont de l'ordre de 10 euros sur 6 à 12 mois. (d'après [www.babyloan.org](http://www.babyloan.org))

projets personnels n'échappe pas à la vague.

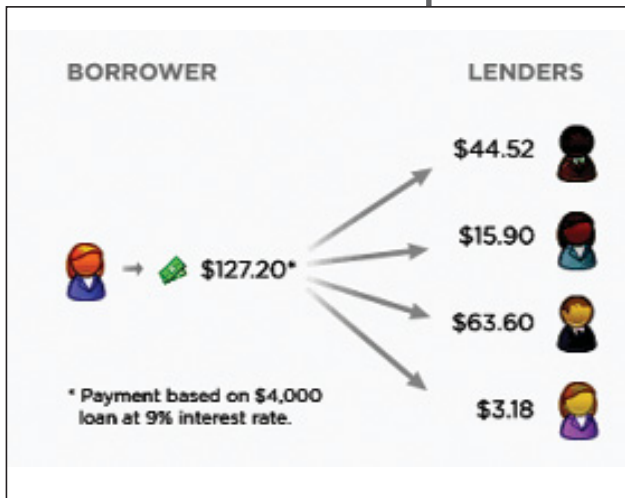
#### ... Mais aussi s'impliquer dans les projets

Les investisseurs/prêteurs éprouvent le besoin de donner un sens humain à leur investissement (fonds éthiques, écologiques). Cependant, les produits proposés par les institutions financières classiques répondent mal à une nouvelle demande: prêter à des personnes réelles que l'on peut voir, dont on peut suivre le projet... De nouveaux entrants du monde de la finance surfent sur cette tendance pour offrir des services qui bousculent le monde du crédit.

#### Principe du crédit entre particuliers sur Internet

##### Le prêt réinventé ?

Le crédit entre particuliers est une pratique vieille comme le monde. Prêts familiaux, prêts au sein d'une communauté, ce mode de financement est pratique mais n'offre pas le même niveau de sécurité que celui d'un prêt bancaire classique. Les services de cré-



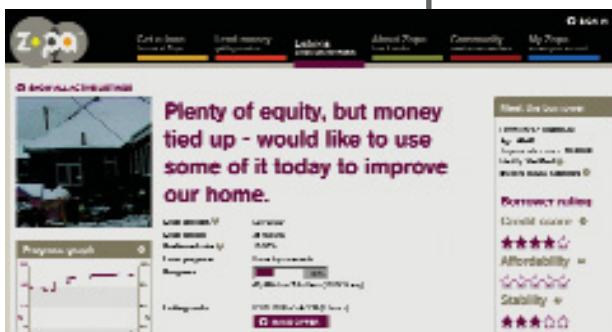
dit P2P mettent en relation des emprunteurs et des prêteurs via un site web spécialisé qui permet de :

- Proposer des taux attractifs pour l'emprunteur et le prêteur
- Simplifier les formalités (mettre en place le crédit, les prélèvements, les remboursements)
- Limiter les intermédiaires et les frais (e.g. supprimer l'intermédiaire de la banque)

Le montant et le taux du prêt est (ou peut être) déterminé par un système d'enchères. Le prêt est généralement établi entre un emprunteur et un ensemble de prêteurs afin de minimiser le risque de non recouvrement. Le service de crédit P2P se rémunère via une commission sur les transactions (en pourcentage du montant emprunté).

Suivons le parcours de l'emprunteur et du prêteur sur une plateforme de crédit P2P.

### **Vous êtes emprunteur ? Décrivez votre projet**



assurer un prêteur potentiel sur votre capacité à rembourser cet emprunt.

2. La plate-forme analyse le marché des prêts et détermine quels prêts sont les mieux adaptés à votre offre d'emprunt. Si des prêts sont susceptibles de coïncider avec votre offre d'emprunt, la plate forme vous prévient.

3. Vous choisissez quels prêts vous souhaitez prendre et envoyez une confirmation. Le prêteur est alors prévenu et confirme ou non votre offre. Une barre de progression vous permet de visualiser l'état d'avancement du projet en fonction des propositions des prêteurs. Le projet est présenté pour une durée limitée à quelques jours et peut déboucher ou non sur une offre de prêt.

4. Lorsque le montant des propositions de prêt atteint la somme que vous demandez, le contrat de prêt est signé, les paiements associés sont mis en place.

### **Vous êtes prêteur ?**

#### **Choisissez votre investissement**

Sur la plate forme, vous, prêteur/investisseur, éliez vos projets :

1. Vous parcourez la liste des projets et vous choisissez ceux dans lesquels vous voulez investir en

fonction de la nature du projet, de la situation financière des emprunteurs, du taux qu'ils souhaitent obtenir ainsi que de leur score de crédit.

2. Vous faites une offre, si vous pensez que l'emprunteur peut rembourser son crédit. vous choisissez alors le montant à prêter ainsi que le taux de rémunération de votre placement. Ne soyez pas trop gourmand sinon votre offre n'aboutira pas.

3. Les offres de prêt qui proposent le taux le plus bas sont sélectionnées à la date de fin fixée pour le projet qui dure de 5 à 10 jours. S'il n'y a pas assez d'offres pour couvrir le montant du projet, le prêt ne se fera pas.

4. Une fois votre argent ainsi placé, vous pouvez voir quelle somme vous prêtez et à qui, ainsi que la somme qui reste à investir.

Vous reversez une commission annuelle de l'ordre de 1 % du montant investi via le service P2P.

#### **Pour mieux emprunter, groupez-vous !**

Certains services de crédit P2P permettent de définir des communautés d'emprunteurs (ex : étudiants asiatiques de Californie ou les femmes entrepreneurs de Seattle) et ainsi de faire bénéficier leurs membres d'une meilleure note de réputation qui est synonyme de confiance accrue dans l'emprunteur et donc un taux plus avantageux pour ce dernier.

Les facteurs déterminants pour la réussite d'un projet de prêt sont, dans l'ordre décroissant (cf. [5]) : la note de crédit, les informations personnelles détaillées (sources de revenus, budget mensuel) l'explication claire du projet et enfin l'appartenance à un groupe d'emprunteurs et la note de réputation associée.

#### **Mais comment gérer le risque ?**

Ces plates-formes sont séduisantes pour les investisseurs (meilleure rémunération) et les emprunteurs (meilleur taux d'intérêt), cependant rien ne garantit au prêteur qu'il sera entièrement remboursé !

Le prêteur porte entièrement le risque.

La plate-forme n'est qu'un intermédiaire. Pourtant le taux de défaillance (non paiement) est très faible. Pourquoi ? Les plates-formes ne sont, contrairement à la légende, pas ouvertes aux mauvais payeurs ! Elles sélectionnent les emprunteurs ayant un bon score FICO (obligatoire aux USA). De plus, la pression des « pairs » (via les groupes entre autres) induit une plus grande responsabilisation de l'emprunteur.

### Premier retour d'expérience

- ...Où les banques s'invitent ZOPA a démarré son service aux USA en s'alliant à 5 Credit Unions (banque mutualiste étasunienne) pour distribuer leurs certificats de dépôt. En choisissant de se démarquer du modèle classique du leader PROSPER, ZOPA USA parie sur un modèle d'affaire hybride puisqu'il s'agit de proposer des produits garantis mais en servant un taux de rémunération plus faible. Cette symbiose entre banques traditionnelles et nouveaux entrants introduit une déclinaison intéressante du modèle. En Europe, ABN AMRO, une des principales banques des Pays Bas vient de proposer un service de crédit P2P (offline) à ses clients.

- ... Quelques ratés ...

Lending Club (USA) a arrêté temporairement ses activités après un an, faute d'être profitable. IOU Central (Canada) a dû suspendre ses activités pour se conformer à la législation locale. Ces tousotements sont inévitables dans l'exploration d'un nouveau marché. Mais, dans le même temps, ZOPA et PROSPER ont doublé le nombre de leurs abonnés avec, il est vrai, une modeste croissance des prêts (+ 15%), car à peine 10% des projets de prêt se concrétisent (cf. [8]).

- ... Mais une solidité du modèle L'épreuve de la crise des crédits hypothécaires (subprimes) a conduit les banques à resserrer leur offre de crédit et pourtant les plates-formes P2P n'ont pas connu de défaillance majeure. Si la forme et les variantes du modèle peuvent encore évoluer,



le cœur du principe semble solide.

### Une offre très diversifiée

Les plates-formes de crédit P2P couvrent un spectre très large allant de la plate-forme à but non lucratif (projets humanitaires de KIVA par exemple) aux plates-formes mercantiles comme PROSPER ou ZOPA qui sont adossées à des fonds de Capital Risque bien connus. Leur management est issu du monde de l'Internet et de la Finance. Certaines plates-formes favorisent le contact entre leurs membres en les incitant à créer une communauté, alors que d'autres préfèrent se présenter comme un service neutre et anonyme d'investissement et de prêt.

### L'Europe et la France à l'abri?

L'Europe commence à être investie : ZOPA est présent en Grande Bretagne et Italie, Boober aux Pays Bas. Mais le marché français possède des spécificités qui le rendront difficile à pénétrer :

1. La réglementation Française actuelle est un obstacle : seuls des établissements de crédit peuvent effectuer des opérations de banques à titre habituel (Code monétaire et financier). Le prêt entre particuliers est très encadré en France.

2. Les Français sont peu endettés et peu appétants au crédit (2 fois moins que la moyenne Européenne) ([2] et [3]).

3. Le marché est très concurrentiel, les taux sont faibles (les taux d'intérêt proposés aux USA par certains établissements de crédit sont inimaginables en France : ils dépassent le taux de l'usure!).

4. La notation systématique de l'emprunteur (score FICO) n'a pas d'équivalent en France où seuls les défauts de crédit sont notés. Or c'est un élément déterminant pour la plate-forme. Mais le contexte réglementaire Européen est en pleine évolution et un tel service apparaîtra tôt ou tard en France en s'adaptant aux spécificités nationales.

### Un marché naissant qui se construit

Plus de 20 services de crédit P2P sont opérationnels dans le monde en 2008 et plusieurs lancements sont prévus pour 2008/2009 (CommunityLend par exemple). PROSPER, le leader aux USA traite \$100 millions de prêts personnels. Selon le cabinet Celent cette forme de crédit devrait représenter 10% du marché bancaire en 2010 : Un marché prometteur mais encore naissant.

Les opérateurs de crédit P2P constituent-ils une menace pour les banques traditionnelles ? Oui, si les banques n'évoluent pas. Les plates-formes de crédit P2P vont probablement grignoter le marché des crédits de « petit » montant. Pour des institutions financières agiles, le crédit P2P constitue une opportunité d'élargissement de services. Innovation de rupture, ces services de crédit P2P peuvent déplacer les frontières du marché du crédit en rééquilibrant le pouvoir en faveur du





consommateur... Mais si des acteurs indépendants prospèrent et prennent un poids significatif, cela ne risque-t-il pas de poser un problème de régulation ?

## Références

- [1] « Financial Services 2.0 », Deutsche Bank Research, août 2006
- [2] « L'accès des ménages au crédit en France », Rapport du Sénat, Joël Bourdin, mars 2006
- [3] « Le marché du crédit à la consommation en Europe », SOFINCO, Nicolas Pecourt, juillet 2007
- [4] « Peer-to-peer lending, Crunchless credit » The Economist, Oct 25th 2007
- [5] « The democratization of personal consumer loans ?.. » M.Herzenstein, R.L.Andrews, U.M.Dholakia, E.Lyandres, University of Delaware, Rice

## THÈME DU PROCHAIN NUMÉRO

# Green IT

Rédacteur :

*Georges Allemand, Ensimag 1986,  
Directeur de projet, AtosOrigin*

### Rejoignez le leader de la téléphonie sur le web !

**Meilleurmobil.com** est une PME de 40 personnes, leader français de la vente en ligne de téléphonie avec abonnement.

**Meilleurmobil.com**  
votre allié dans la jungle des mobiles

Accompagné par un fond d'investissement, l'entreprise double son Chiffre d'Affaires tous les ans. Créé en 2003 par son Président Jean-Vital de Rufz, l'équipe a entre 20 et 30 ans, dans une ambiance intense et dynamique.

Vous êtes familier avec le web, attiré par les nouvelles technologies ? Vous voulez participer au développement d'une entreprise en pleine expansion, vous investir dans une aventure motivante et partager l'esprit créatif de l'équipe en place ?

Meilleurmobil propose aux étudiants de l'ENSIMAG deux stages de dernière année :

Refonte du comparateur de forfait	Module d'un Projet CRM
<p>L'objectif de Meilleurmobil est de conserver son leadership en refondant l'IHM de son comparateur de forfait pour le rendre plus interactif.</p> <ul style="list-style-type: none"><li>- Outil plébiscité par la presse !</li><li>- Gère plusieurs dizaines de combinaisons</li><li>- <b>Environnement</b> : J2EE, Eclipse, SQL, Jboss, Linux</li><li>- Découverte des frameworks Hibernata, <b>Struts</b> 1/2</li><li>- Stage : analyse, conception, développement</li><li>- En collaboration avec un développeur <b>Ajax</b></li></ul>	<p>Meilleurmobil s'est doté d'un outil de gestion de la relation client, construit sur mesure en interne.</p> <p><b>Objectif</b> : créer de nouvelles sources de données, doter l'outil de ciblage d'une nouvelle IHM</p> <ul style="list-style-type: none"><li>- Outil gérant plusieurs millions de contacts</li><li>- Découverte des outils TALEND (ETL), outils de <b>Datamining</b>, enrichissement du Datawarehouse.</li><li>- <b>Environnement</b> : J2FF, Postgresql, Eclipse, Linux</li><li>- Stage : analyse, conception, développement</li></ul>

**Rémunération** : Montant à négocier - **Lieu** : Boulogne (92), Métro ligne 10 (Boulogne-Jean-Jaurès), 30 min de Paris-centre.  
**Plus d'informations ?** Consulter la liste des stages de l'ENSIMAG,  
Ou contactez Arnaud Lapiere, DSI Meilleurmobil, 01 41 12 80 74 / [alapiere@meilleurmobil.com](mailto:alapiere@meilleurmobil.com)

*S'il est un entrepreneur français symbole de l'internet, c'est bien Pierre Chappaz, cofondateur de Kelkoo et ardent défenseur du web 2.0 sur son blog, qui figure régulièrement dans le top 20 français. Wikio, le service d'information qu'il a créé, est un exemple parfait d'application au quotidien des principes du web 2.0, qui peut être vu comme une plateforme informatique à part entière, fournissant des applications web aux utilisateurs, à travers des technologies mûres que se sont appropriées les développeurs, en particulier de RSS. Dans ce court article, Pierre nous donne un peu plus de détails sur ce service.*

Jean-Marc Darrigol.

# Wikio, un portail d'information 2.0

par Pierre Chappaz, centralien, fondateur de Kelkoo puis de Wikio

Wikio est un portail d'information de nouvelle génération, et un service shopping associé.

Basé sur la technologie RSS (Real Simple Syndication), et des algorithmes sémantiques (analyse de texte par extraction de concepts), Wikio facilite la navigation dans les informations produites par les medias et les meilleurs blogs.

Seuls les blogs de qualité sont indexés par Wikio, et la totalité des médias. Le service s'adresse au grand public autant qu'aux internautes avancés, gros consommateurs d'information.

Grâce à sa technologie d'analyse sémantique et d'analyse des liens hypertextes, Wikio fait ressortir automatiquement les tendances des informations, des discussions blogosphériques, ou encore les vidéos ou les produits les plus buzzés sur Internet. Wikio comporte aussi un moteur de recherche spécialisé pour la recherche d'informations dans les blogs, les medias, les vidéos et les produits.

Pour les internautes désireux de publier un article directement (tout le monde n'a pas un blog), le site dispose aussi d'une plateforme de publication collective.

Enfin, le classement des blogs les plus influents publié par Wikio fait référence dans la blogosphère (c'est un excellent moyen de découvrir le monde des blogs!).

## **Des technologies avancées rendues accessibles pour tous**

Wikio, dont l'index est constitué par les flux RSS des medias et des blogs référencés, propose un lecteur de flux RSS débarrassé de toute complexité technique: pour recevoir automatiquement le flux d'informations d'un media ou d'un blog, il suffit de cliquer sur le lien 's'abonner' placé sous le titre de n'importe quel article. L'abonnement est gratuit bien sûr. De même, vous pouvez recevoir toutes les informations publiées sur un sujet particulier qui vous intéresse, (faites une recherche sur le mot clé correspon-

dant, puis cliquez sur le bouton "s'abonner") Vous créez ainsi des pages d'informations personnelles, pour suivre en permanence les sujets ou les sources qui vous intéressent.

Cette notion de personnalisation de l'information va plus loin: des pages personnelles d'information plus sophistiquées peuvent être créées sur n'importe quel sujet, avec l'assistance d'un système de suggestions dynamiques.


## **Shopping: agrégation de toutes les informations sur les produits**

Les technologies développées pour l'analyse de l'information permettent aussi à Wikio d'offrir un service shopping très utile pour le consommateur, en agrégeant toutes les informations disponibles sur le Net sur chaque produit: articles, tests, avis de consommateurs, photos du produit, vidéos, et bien sûr offres des marchands en ligne. Ce service shopping complète le modèle économique de Wikio. Une entreprise dans le secteur de l'information comme Wikio ne peut pas vivre que de l'information ... En effet, la publicité ne rapporte pas assez, car le service - entièrement gratuit, et produit en 5 langues - est coûteux à produire, avec une équipe de 35 personnes. Il faut des services associés comme le shopping pour rentabiliser l'investissement.

## **Une ambition européenne, voire mondiale**

A peine plus de deux ans après son lancement, Wikio est déjà utilisé par 3 millions de français (www.Wikio.fr) et le double d'euro-péens. Un site américain a également été lancé. Le Web 2.0 permet à de jeunes entreprises Internet de se développer à l'échelle internationale, et le secteur de l'information qu'a choisi Wikio est sans aucun doute l'un de ceux où il y a le plus à inventer.



A photograph of the Golden Gate Bridge in San Francisco, California, viewed from a low angle looking up at one of its massive towers. The bridge's red-orange steel structure is prominent against a clear blue sky. The water of the bay is visible in the foreground, and rolling hills are in the background.

Dexia est une banque européenne, leader mondial du financement public local. Avec une capitalisation boursière de EUR 20,3 milliards au 31 décembre 2007. Dexia se classe parmi les quinze plus grands établissements financiers de la zone euro. À la même date, le groupe compte 35 200 collaborateurs, répartis dans 37 pays. Dexia jouit d'une des meilleures notations de crédit du secteur bancaire. La stratégie de développement de Dexia repose sur deux piliers : la banque universelle en Europe (Belgique, Luxembourg, Slovaquie et Turquie), et les services financiers au secteur public local et financements de projets au niveau mondial.

Retrouvez toutes nos opportunités  
sur [www.workatdexia-cl.com](http://www.workatdexia-cl.com)

*sans durée, on ne construit rien*

**DEXIA**



## L'Internet mobile

Par Laurent Gatignol  
Ensimag 1994 - Directeur exécutif de Kelkoo France

*L'Internet mobile, à savoir l'accès au réseau Internet depuis un terminal autre qu'une machine type PC et via une liaison non fixe, a fait naître à la fin des années 90 et au début des années 2000 de grands espoirs puis des déceptions. Aujourd'hui, on est loin des prémisses et promesses, les liaisons proposent un débit en constante augmentation et les terminaux sont de plus en plus sophistiqués, à l'image de l'iPhone d'Apple, emblème de cette évolution. En parallèle, les services associés devront suivre l'évolution des terminaux et des débits et s'adapter à de nouveaux usages pour le mobile connecté.*

### Laurent Gatignol

Passionné de mobilité et éditeur de plusieurs blogs et sites web dans le domaine, Laurent Gatignol, fraîchement diplômé de l'Ensimag en 1994, débute sa carrière chez Microsoft France avant de rejoindre le département Télécom de Sema Group où il assure la direction de projets notamment pour France Télécom.

Arrivé en 2000 chez Kelkoo France, il prend en charge la direction technique France, lance les sites Belges, il travaille plus récemment à la stratégie européenne des équipes de développements locales, lance le service Kelkoo mobile en 2007 en France puis en 2008 en Europe. Laurent est désormais Directeur Exécutif de Kelkoo France.

### Evolution technique de l'Internet mobile !

Il y a bientôt une dizaine d'années, malgré des terminaux monochromes et une navigation via des pages WAP formatées pour de tous petits écrans, nombreux étaient ceux qui pariaient sur un développement rapide et généralisé de l'accès aux informations et aux sites mobiles depuis son téléphone. Dans les faits, les limitations en performance et les coûts ont généré de nombreuses déceptions et une adoption plus lente et moins massive que prévue. On retiendra tout de même le succès de l'i-mode, extrêmement populaire au Japon et qui a trouvé son public également en Europe.

Au fil des années, les paramètres de l'Internet mobile ont changé, et tout semble maintenant réuni pour que l'adoption aille en s'accéléralant. Pour s'en rendre compte, un témoin emblématique de ce changement est apparu

en 2007, il s'agit de l'iPhone. Avec ce terminal et ses forfaits associés, Apple et les opérateurs n'ont pas inventé l'Internet mobile, mais cela a incarné et incontestablement accéléré le changement qui s'opérait dans ce domaine.

Ainsi, la limitation liée au terminal lui-même est en train d'être réduite peu à peu au fur et à mesure des progrès sur les résolutions d'écran, sur l'interface utilisateur (via les écrans tactiles et claviers miniaturisés) et enfin sur la consommation électrique. Les smartphones, qui ont remplacé les PDA pour l'accès mobile à Internet, proposent des écrans d'une dimension et d'une résolution croissante. Des résolutions de 320 x 240 pixels voire VGA (640 x 480) ne sont plus rares. Il devient ainsi possible d'afficher à l'écran de manière lisible une portion d'une page Internet standard, voire une page complète. L'iPhone (résolution de 320 x 480) a lui proposé le premier navigateur embarqué (Safari) qui soit capable d'afficher

une page web complète avec une compatibilité HTML semblable à celle d'un navigateur bureautique type Internet Explorer ou Firefox. Mieux encore, il permet un zoom rapide sur une portion de la page, nécessaire pour la lisibilité sur un terminal de taille réduite et est compatible avec les technologies regroupées sous le terme Ajax, utilisées sur les sites web récents (web 2.0). Cela permet l'utilisation de Webapps, véritables applications connectées, proposant une interface riche et ne nécessitant pas de recharger la page. D'autres navigateurs empruntent une voie semblable en terme de visualisation, citons notamment Opera mobile sous Windows Mobile et Nokia S60. Pour l'internaute, c'est la garantie d'un accès à l'ensemble des sites Internet. Pour les éditeurs de sites, cela signifie que leur site standard est lisible par des terminaux mobiles, là où auparavant une version dédiée devait être prévue, augmentant les coûts, les versions et la maintenance.

L'autre limitation concernait la faiblesse des débits de données et donc le temps d'attente engendré pour l'utilisateur. Pour pouvoir afficher des pages de sites standards, il est vital de disposer d'un débit suffisant. Après le GPRS puis le Edge(2G), le lancement des réseaux de troisième génération (3G /UMTS et 3G+ / HSPDA), capables d'atteindre des débits de plusieurs Mbit/s, offrent désormais le cadre technique suffisant pour un véritable confort d'utilisation de l'Internet mobile. En parallèle, les terminaux sont de plus en plus équipés de la technologie sans-fil WiFi.

Mais l'aspect technique n'est pas suffisant, restait en effet le paramètre du coût de connexion. Pour faire une analogie avec l'Internet fixe : qui n'a pas apprécié le passage d'un Internet par liaison téléphonique RTC, facturé souvent à la minute de connexion, à l'ADSL ou la connexion devenait illimitée et permanente ? Le même changement est en cours pour les connexions mobiles. Les forfaits proposant du 'data illimité', selon le jar-

gon des opérateurs, sont en train de se généraliser. En France, SFR avec les forfaits Illimythics, mais également Orange avec ses forfaits iPhone (qui a été l'un des premiers terminaux commercialisés aux Etats-Unis avec un forfait données illimitées) offrent la possibilité d'être connecté en permanence à Internet via son téléphone mobile, pour un surcoût d'abonnement forfaitaire et non plus facturé à la quantité de données échangées.



Mixez un débit conséquent, une connexion permanente et des terminaux offrant un confort d'utilisation et de visualisation suffisant, et vous avez l'environnement parfait pour l'adoption croissante de l'Internet mobile. Reste à savoir ce que les utilisateurs attendent et ce que les opérateurs et les éditeurs de sites peuvent proposer

### Usage de l'Internet mobile

IDC compte en 2008 pas moins de 546 millions d'utilisateurs de mobiles connectés à Internet, soit le double de 2006. En 2012, le nombre de mobiles connectés à Internet sera supérieur au nombre d'ordinateurs personnels connectés. Il est alors légitime de se poser la question de ce que font et feront les « mobinautes ». Les usages du mobile dans le grand public sont actuellement liés à la recherche d'information : actualité, sport, météo, bourse. On trouve donc des moteurs spécialisés pour mobile, édités par les grands de la recherche

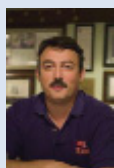
mondiale que sont Google (avec des versions mobiles de ses services), Yahoo ! (avec OneSearch) ou MSN. La consultation d'email et la messagerie instantanée figurent également dans les usages les plus développés actuellement.

Les tendances générales de l'Internet « classique » devraient également avoir leur équivalent sur le mobile. Ainsi, le « m-commerce », encore peu développé faute de sites marchands compatibles, devrait progresser, poussé par des terminaux performants permettant de surfer et d'acheter depuis son mobile mais également par le développement croissant du nombre de sites marchands adaptés à l'utilisation depuis un mobile. En 2007, ce sont plusieurs sites e-commerce français d'importance qui ont lancé une déclinaison mobile de leur site : Voyages-sncf, Alapage, Fnac ou encore Kelkoo.

La dimension sociale du web actuel (web dit « 2.0 ») devrait trouver un prolongement naturel et même s'enrichir avec le mobile. Ainsi, tenir son blog depuis son mobile, le « moblogging », faire du « microBlogging » (envoi de textes courts vers son réseau d'amis avec des applications type Twitter) voire du « photoblogging » en envoyant directement ses clichés en ligne, en temps réel et de manière géo-localisée (grâce à la présence d'un GPS embarqué de plus en plus généralisé) sont des utilisations en plein essor. Bien entendu, l'outil mobile, porté sur soi et connecté en permanence devrait ouvrir la voie à des services liés à la géo-localisation tels le signalement de boutiques, restaurants, centres d'intérêts mais aussi rencontres liés au lieu où l'on se trouve.

On peut anticiper également que le mobile, au delà du mail, servira à retrouver sur son petit écran, où que l'on soit, l'ensemble de son univers numérique personnel. Celui-ci, stocké sur des serveurs accessibles au choix depuis son ordinateur de bureau ou depuis son mobile permettra aux documents, photos, musique et vidéos

d'être accessibles en permanence. Enfin, selon IDC, dans les années à venir, avec l'augmentation des débits et la réduction des coûts, la consommation de contenu multimédia (musique, vidéos et sonneries), devrait-elle devenir prédominante depuis les mobiles. En attendant, les éditeurs de sites mobiles doivent continuer à composer avec des débits variables et un parc de terminaux hétérogènes. Tous les terminaux en circulation ne proposant pas une navigation intuitive et à la compatibilité avancée, il est raisonnable de déployer des sites dans un format lisible par une majorité de mobiles. Une norme basée sur XHTML (voire la récente XHTMLBasic 1.1) permet de proposer un Internet mobile compatible avec la plus grande variété de terminaux. Il convient également de veiller à conserver un poids des pages le plus réduit possible, une taille et une quantité d'images réduites au strict nécessaire tout en travaillant à inventer des applications mobiles connectées de plus en plus ambitieuses. Elles devraient pouvoir trouver un financement avec la publicité, qui va bien entendu accompagner l'augmentation de l'audience et des usages mobiles.



## L'Agilité au Service du Web

par Alexandre Boutin  
Ensimag 1989  
Agile Coach – Praticant Scrum Certifié – [www.agilex.fr](http://www.agilex.fr)

*Les approches dites 'Agile' ont vraiment le vent en poupe ces dernières années car elles offrent une alternative réaliste et plus efficace à l'approche basée sur des 'Cycle en V' associée à un référentiel normatif plus ou moins complexe (ISO, CMMi ...).*

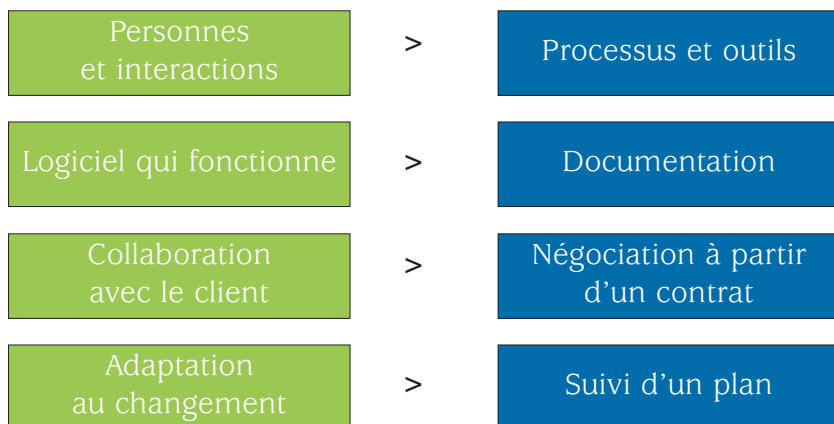
*Cet article vous éclairera sur les raisons qui font des entreprises du Web les leaders mondiaux incontestés des pratiques Agile.*

### Alexandre Boutin

Diplômé de l'Ensimag en 1989, je travaille sur la région parisienne pendant 9 ans pour différentes entreprises au développement de logiciels embarqués civils et militaires. Durant cette période, j'occupe les postes de développeur, chef de projet et chef de produit. En 1998, je rejoins une société de service en informatique grenobloise pour devenir Directeur France des réalisations au forfait (avant-vente et réalisation - Département de 140 personnes en 2003). Je contribue également à l'obtention de certifications ISO9001 et CMMi. Début 2004, je rejoins Yahoo pour manager les équipes de développement (70 personnes à Grenoble), implémenter des méthodes et outils pour les développeurs et unifier le processus de développement logiciel pour les équipes locales puis au niveau Européen. Enfin, en 2006, je suis muté dans l'équipe stratégie internationale pour définir et implémenter un processus de développement logiciel commun à l'Europe, l'Asie et l'Inde. Dès cette époque je commence à me documenter sur les méthodes agiles, me forme pour devenir "Scrum Practitioner", devient préconisateur et coach agile en interne, tout en contribuant à la communauté Agile mondiale en présentant mon retour d'expérience lors des conférences de Londres (2007) et Toronto (2008).

\*

## MANIFESTE AGILE : VALEURS



### Les approches dites 'Agile'

Les approches Agile sont toutes basées sur le « Manifeste Agile » qui définit les principes de commun à toutes les méthodes de développement dites Agile ([www.agilemanifesto.org](http://www.agilemanifesto.org)). Les plus connues et utilisées sont SCRUM et XP, mais d'autres méthodes telles que DSDM, RAD ou CRYSTAL ont également leur fervents supporters.

Une façon simple, quoique restrictive, pour appréhender les approches Agile est d'analyser les valeurs partagées par ces approches.

### Les valeurs de l'agilité

L'agilité se reconnaît dans les valeurs de gauche (en vert) plutôt que les valeurs de droite (en bleu). Cela ne veut pas dire que les valeurs de droite sont de mauvaises valeurs mais simplement que pour tirer le maximum de bénéfices de l'agilité, il faut faire des choix et donc accepter de sacrifier certains avantages apportés par les valeurs de droite pour en obtenir beaucoup plus des valeurs de gauche.

#### Personnes et Interactions

L'Agilité considère que le facteur humain est une plus grande source de succès que les moyens matériels ou les procédures. Il est préférable d'avoir une équipe soudée, et qui commu-

nique beaucoup en toute confiance plutôt que des individualistes même brillants

Un processus détaillé ne garantit qu'un minimum (ce qui est décrit par le processus) qu'un groupe se limitera souvent à produire. A l'opposé, une équipe agile aura à cœur de dépasser les limites individuelles en s'entraînant et la performance finale sera bien souvent au-delà des espérances initiales.

#### Logiciel qui fonctionne

L'utilisateur final sera satisfait uniquement si le produit qu'il utilise est fonctionnel, et il n'y a rien de mieux que d'avoir en permanence un produit fonctionnel pour s'assurer que le produit final le sera également.

Pour l'Agilité, l'effort doit être porté sur le produit qui fonctionne plutôt que sur les artefacts qui y sont associés (dont la documentation). Par exemple, il vaut mieux commenter le code que de rédiger un document de conception détaillée très coûteux à maintenir. De même, la distribution des compétences au sein de l'équipe est la meilleure garantie de faire perdurer la connaissance.

#### Collaboration avec le client

La contractualisation est liée à l'inquiétude du client de ne pas recevoir un produit conforme à ses attentes et donc il définit le détail de ce qu'il attend au risque de ne plus pouvoir chan-

ger d'avis en fonction de l'évolution de son marché.

Pour l'Agilité, le client doit être impliqué dans le développement du produit afin que ses besoins, même ceux non identifiables au démarrage, soient satisfaits. La notion de client et de fournisseur s'estompe pour devenir une collaboration basée sur un respect mutuel et une confiance réelle.

#### Adaptation au changement

Lors de l'utilisation d'une première version d'un produit, nous fourmillons tous d'idées nouvelles ou de modifications des besoins initiaux. Si telle est la réalité, pourquoi la nier ?

Pour l'Agilité, le changement est normal et bienvenu, car le suivi d'un plan établi à l'avance est moins efficace que la re-planification régulière en fonction des nouvelles informations disponibles. Mais pour ce faire il est nécessaire que la structure du logiciel et la méthode de production de ce logiciel soient flexibles.

### Le Web : Leader Agile

Les études actuelles faites par le cabinet Forrester montrent que toutes les grandes sociétés du Web font partie des leaders mondiaux de l'Agilité, en compagnie de grands groupes tels que British Telecom et Borland.

Ce constat s'explique par l'aisance des sociétés du Web à implémenter plus rapidement et facilement les valeurs de l'Agilité que d'autres entreprises.



Qui voudrait ne pas être Agile ?

#### Personnes et Interactions

Le Web est un endroit idéal pour créer une nouvelle société car pour un investissement raisonnable (des machines

# dossier

## A L'HEURE DE L'INTERNET GLOBAL

et quelques concepteurs) vous pouvez accéder à des millions de clients. Une startup du Web met l'accent sur l'esprit d'équipe et des personnes polyvalentes plutôt que sur des processus établis.

Cette façon de faire est une valeur Agile qui reste profondément ancrée dans les esprits, même lorsque la société est amenée à grandir du fait de sa réussite.

### Logiciel qui fonctionne

Une autre caractéristique des sociétés du Web est leur maîtrise complète du déploiement du produit final chez le client qui paye lui-même la connexion pour y accéder. A l'opposé, un logiciel de téléphone portable se déploie sur un matériel physique et il faut avion pour embarquer un calculateur de vol avec son logiciel.

Les sociétés du Web ont donc l'avantage de pouvoir vérifier plus facilement, tout au long du développement, le produit final.

Pour ce qui est la documentation, et

sans affirmer qu'elle est inutile, je vous poserais simplement la question : Avez-vous vraiment lu le manuel utilisateur pour lire les news sur Yahoo.fr ou acheter sur eBay ?

### Collaboration avec le client

L'accès direct à l'utilisateur est vraiment une grande force du Web. En effet, nul besoin de déployer une nouvelle version sur un parc de machines dans les locaux de vos clients, ce sont les clients qui viennent à vous pour voir et utiliser les nouveaux produits.

Le feedback client est obtenu directement par les « avis utilisateurs » rédigés par les utilisateurs eux-mêmes ou indirectement par l'analyse du taux de fréquentation du site web avant et après la mise en ligne de nouveaux produits.

### Adaptation au changement

Le Web est très concurrentiel, certains diront plus qu'ailleurs, car il n'y a qu'à regarder la rapidité du succès de certains produits comme Facebook, ou de se souvenir de quels produits Web

étaient utilisés il y a 10 ans (AltaVista) pour se rendre compte de la vitesse incroyable de l'évolution du marché du Web.

Les entreprises du Web ont donc l'absolue nécessité d'être très réactives et de s'adapter sur des délais très courts.

### Conclusion

La similitude des valeurs de l'Agilité avec les valeurs des entreprises du Web explique en grande partie les raisons du succès des méthodes Agile au sein de ces entreprises.

Lorsque votre culture, votre indépendance, vos utilisateurs et votre marché vous poussent vers l'Agilité, l'appliquer devient du bon sens plus que de la stratégie et en devenir les leaders semble naturel, et si en bonus cela contribue à l'amélioration de votre image de marque, c'est parfait !

\*

## Voyage d'études 2008 : destination Maroc

*Le voyage d'étude est de retour dans la vie associative de l'Ensimag. Organisé en octobre par le BDE-Entreprise, il a permis à 21 étudiants de partir à la découverte du Maroc.*

Lorsque la semaine banalisée réservée au Séminaire Carrières, a été supprimée dans la nouvelle maquette pédagogique de l'école, le BDE-Entreprise a décidé de faire renaître le voyage d'étude - tradition perdue au milieu des années 90- et a choisi le Maroc comme destination.

Ce pays, qui connaît un fort développement économique couplé à des plans ambitieux dans de nombreux secteurs (transports, TIC, tourisme, etc.) a été choisi aussi pour sa stratégie visant au développement des activités offshore.

Les partenariats (dont certains ont été établis avec l'aide de l'AAE et de l'association

AMET) ont permis de visiter 5 sociétés : Atos Origin (Centre de services de Casablanca) ; GFI Maroc ; Omnidata, (1ère SSII marocaine



créée en 1989 par Monsieur Khalid El Hariry, Ensimag 1987) ; S2M (Société Maghrébine de Monétique) et l'opérateur téléphonique Wana. La semaine a comporté une rencontre avec le Centre d'Investissement de Tanger (visite du port Tanger-Méditerranée, investissement d'un milliard d'euros), ainsi qu'une réception au Ministère de l'Industrie, du Commerce, et des Nouvelles Technologies

(présentation du plan de développement des TIC "IMPACT 2013").

Les étudiants ont présenté l'Ensimag au lycée Mohammed V de Casablanca, et ont remis les livres de prépa récoltés avant leur départ à la bibliothèque du lycée. A l'Ecole Nationale Supérieure d'Informatique et d'Analyse des Systèmes, ils ont posé les bases d'un partenariat entre la Junior-Entreprise de l'Ensimag (NSIGMA) et la future Junior-Entreprise de l'ENSISAS.

Mais parce qu'un voyage d'étude ne se résume pas au mot "étude", le petit groupe a pleinement profité de la richesse culturelle du Maroc, notamment dans les villes de Marrakech et de Rabat, et apprécié la gastronomie au travers de succulents et... interminables repas, dont un fameux partagé avec les anciens de Casablanca !

*L'équipe organisatrice du voyage d'étude*  
*Jérôme Daydé et Aness Meski*



# la vie de l'association

## L'INRIA, partenaire de l'AAE Ensimag

Nous avons le plaisir et l'honneur de vous annoncer le partenariat que nous venons de signer avec l'INRIA.



## Les événements parisiens

par Nadia Robinet, Ensimag 1990.

Cet automne à Paris, l'AAE ENSIMAG est co-organisatrice de deux événements majeurs de l'institut G9+ dont notre formation est membre actif :

### La treizième rencontre annuelle de l'institut G9+

abordera la période actuelle avec une double interrogation :

- Quel impact d'une période difficile sur l'écosystème DSI / fournisseurs ?
- Face aux nouveaux défis, quelles sont les opportunités à saisir ?

Une matinée d'exception le **23 octobre** où nous comptons parmi les intervenants Ludovic Le Moan ENSIMAG 88, serial entrepreneur et créateur de Goojet (prix de la meilleure start-up lors du WEB3 2007).

Clôturée par Louis Schweitzer, président de la Halde et président du Conseil d'administration de Renault, la manifestation va réunir à nouveau des intervenants de haut niveau, dans l'ambiance « zéro langue de bois » qui explique son succès constant.

Une conférence sur le thème des **nouvelles stratégies économiques et commerciales Open Source des éditeurs de logiciel**. D'ores et déjà, réservez la soirée du **2 décembre 2008** dans votre agenda.

Nous sommes à la recherche de nouveaux sujets et intervenants pour les dîners-débats et nous cherchons toujours à enrichir l'offre destinée aux anciens élèves alors n'hésitez pas à nous soumettre vos idées et pourquoi pas, vous impliquer dans l'association.

Contact : [nadia.robinet@aae-ensimag.com](mailto:nadia.robinet@aae-ensimag.com)

Groupe : [imaginactionparis@yahoogroupes.fr](mailto:imaginactionparis@yahoogroupes.fr)

## Un nouveau site emploi pour 2009 !

Vous êtes nombreux à être abonnés au site emploi de l'AAE ENSIMAG, aujourd'hui géré par une liste de diffusion, et il nous semblait important de rendre ce site plus fonctionnel: capacités de recherche par région, mise en place d'alertes automatiques, etc...

Ainsi, dans le cadre des projets 2A (projets pour les étudiants ENSIMAG 2<sup>ème</sup> année), l'AAE ENSIMAG a proposé un sujet de refonte du site d'offres d'emploi qui a suscité un réel intérêt auprès des étudiants car huit d'entre eux ont demandé à travailler dessus !

Nous avons rédigé les spécifications du projet puis avec l'aide précieuse de Laurent Goujon (ENSIMAG 2003), nous avons suivi l'avancement du projet jusqu'aux soutenance en juin. Les étudiants se sont répartis les rôles sur le projet, ont planifié leur travail, développé le site et géré la QA.

Au final, le projet est une réussite car ils ont pu livrer un site fonctionnel qui est en bêta-test jusqu'à la fin de l'année 2008 pour une mise en production prévue pour début 2009.

Merci à Antoine UBERGER, Olivier BITTINGER, Christian HUGUES, Damien BRUNEAU, Baptiste LEPERS, Maxim OKOROKOV, Bogdan POPA et Ekaterina SHEVELEVA pour leur travail, et à Yves Denneulin et Marie Laure Potet pour l'encadrement ! Vu les résultats obtenus, nous espérons continuer la collaboration avec les 2A sur d'autres sujets: à discuter lors des réunions du groupe grenoblois !

En tant que premiers utilisateurs, toutes vos remarques sur le nouveau site seront les bienvenues, n'hésitez pas à nous contacter (plus d'infos sur :

<http://wiki.aae-ensimag.com/services>).

Laurent Testard (ENSIMAG 1993) & Fanny Strudel (ENSIMAG 1999)

## XMP Business Angels : investisseur privé, fond d'investissement et formation

XMP BUSINESS-ANGELS a pour vocation la mise en relation entre des **investisseurs privés personnes physiques (business angels)** et des **entrepreneurs** à la recherche de financement.

Les membres de cette association sont des ingénieurs diplômés issus de l'Ecole Polytechnique, des Ecoles Nationales des Mines, de l'Ecole Nationale des Ponts et des

# la vie de l'association

autres Ecoles d'Application traditionnellement associées. A ce titre les ingénieurs ENSIMAG peuvent adhérer à cette association.

Chaque adhérent moyennant une cotisation annuelle se voit proposer au cours de l'année des opportunités d'investissement dans de jeunes entreprises lancées et développées par des ingénieurs en majorité issus des écoles cités ci dessus mais sans exclusive. L'investissement se fait à titre personnel ou par un groupe d'investisseurs via un pacte d'actionnaires. Le processus de sélection des entreprises est décrit sur le site de XMP Business Angels. Dans le cadre de la loi TEPA, **XMP Business Angels a ouvert en Juin 2008 pour ses adhérents un fond d'investissement éligible à la déduction de l'ISF.**

Le fond investira dans les jeunes entreprises sélectionnées par les adhérents.

Du fait de cette déductibilité de nombreux fonds se développent rapidement et le nombre d'adhérents croit très

rapidement. Aujourd'hui les frais d'accès sont réduits. Cependant la gestion des fonds levés va demander de l'intelligence et du temps. Les gestionnaires de fond vont être amenés prochainement à monter les cotisations et-ou créer un droit d'entrée. XMP est probablement un excellent observatoire des créateurs d'entreprise et leur fond une opportunité plus un lieu d'échanges.

**L'association organise également en partenariat avec OSEO des journées de formation** dont l'objectif est de présenter les modes opératoires du Business Angel et de créer du lien entre les différents acteurs : investisseurs, réseaux, entrepreneurs.

Pour mieux connaître XMP Business Angels, vous pouvez consulter les informations sur leur site : <http://xmp-ba.m4x.org/>

Etant adhérent de ce réseau vous pouvez aussi me **contacter** : [xavier.tariel@wanadoo.fr](mailto:xavier.tariel@wanadoo.fr)

**Xavier Tariel (Ensimag 1974).**

## Jean-Pierre Ansart

« **Jean Pierre ANSART** nous a quitté le 22 octobre 2008.

Ingénieur ENSIMAG 1972,

expert reconnu en télécommunications et intégration des systèmes hétérogènes, il demeure une référence cumulant plus de 30 ans d'expérience dans le domaine des systèmes d'information ».

Jean Pierre Ansart démarre sa carrière en France au sein du CNRS et d'IBM. En 1979, il rejoint MARBEN et crée la division « Réseaux et Télécom ». En 1986, il crée MARBEN PRODUCTS, société spécialisée dans les logiciels télécom OSI et équipe la plupart des grands constructeurs mon-

diaux. Jean Pierre Ansart poursuit sa carrière chez ATOS en tant que directeur des divisions Télécom, Produit et Technologie. En 1996, il développe le premier prototype du logiciel SCORT et fonde la société avec ses partenaires. Il continue à collaborer avec ATOS jusqu'en octobre 1999, en qualité de Directeur Technique d'ATOS Intégration.

Depuis octobre 1999, Jean Pierre Ansart était président de SCORT.

Nous nous associons à la peine de son épouse Salwa et de ses deux filles pour saluer la mémoire d'un homme hors du commun.»



## BON D'ADHESION 2009

Tous les Anciens Elèves de l'ENSIMAG sont invités à adhérer à l'Association. La cotisation annuelle est valable du 1er janvier au 31 décembre 2009 et est établie selon les catégories suivantes :

Catégorie	Montant de la cotisation (en euros)
Bon de soutien Promotion 2009 à 2011	5
Promotion 2008	15
Promotion 2007	30
Promotion 1957 à 2006	60
Diplômé en recherche d'emploi	30
Soutien complémentaire à l'AAE	....
<b>TOTAL</b>	

Le taux de cotisations pour un couple d'une promotion < – 2007 s'élève à 1,5 fois le montant de la cotisation individuelle soit:  
- 90 Euros pour les promotions <= 2006  
- 45 Euros pour la promotion 2007

**La cotisation peut être réglée par CB, chèque, virement ou prélèvement bancaire.**

- **CB** : <http://annu.aac-cnsimag.com/coti.html>

- **Chèque** : complétez le coupon réponse ci-dessous et libellez votre chèque à l'ordre de **l'Association des Anciens Elèves de l'ENSIMAG**. Envoyez l'ensemble à l'adresse suivante  
**AAE Ensimag - 68, Bld de Port Royal - 75005 Paris Tél : 06 28 50 08 85**

- Pour effectuer un **Virement**, écrivez à [contact@aae-ensimag.com](mailto:contact@aae-ensimag.com) pour recevoir les coordonnées bancaires de l'AAE.

- **Prélèvement Bancaire** : Annuel ou bien périodique, écrivez à [contact@aae-ensimag.com](mailto:contact@aae-ensimag.com) pour recevoir le formulaire approprié ou bien rendez vous sur le site <http://wiki.aae-ensimag.com/adhesion>.

### COUPON-REPONSE : AAE ENSIMAG COTISATION 2009

Nom : ..... Promotion de sortie : .....

Prénom : .....

Montant de la cotisation : .....

Cette cotisation n'est pas déductible des impôts.

Reçu : Oui Non

Date : .....

Signature :



# Sophis recrute têtes bien faites



**V**ous êtes ingénieur de formation. Vous recherchez, loin des schémas tout tracés, une entreprise où l'audace, le non-conformisme, l'ouverture sur l'international sont des valeurs-clés. Sophis va vous intéresser. Acteur majeur du marché des logiciels pour les produits dérivés (sur actions, crédit, matières premières, taux, hybrides...), Sophis poursuit

son développement rapide dans les grandes capitales financières. Pour soutenir cette croissance, Sophis recrute.

**Ingénieur**, vous intégrez notre département Recherche et Développement pour concevoir des logiciels financiers innovants, de très haute technicité.

**Consultant**, vous participez au plus haut niveau à des projets d'intégration

complexes et de grande envergure, de l'avant-vente à l'après-vente.

Réactif, avec une grande capacité d'adaptation, maîtrisant l'anglais, vous avez de fortes compétences en mathématiques appliquées ou en développement et un intérêt prononcé pour les marchés financiers.

Bienvenue chez Sophis.

Découvrez les carrières Sophis sur : [www.sophis.net](http://www.sophis.net).

Merci d'adresser votre candidature sous la référence **MAG** à [drh@sophis.net](mailto:drh@sophis.net)  
ou à Sophis - Ressources Humaines - 24-26 Place de la Madeleine - 75008 Paris



**SOPHIS**

New York - Dublin - Londres - Paris - Francfort - Milan - Dubai - Singapour - Hong Kong - Seoul - Tokyo