

**Séminaire
Ressources Technologiques
et Innovation**

*organisé grâce aux parrains
de l'École de Paris :*

Algoé²
Alstom
ANRT
AREVA²
Cabinet Regimbeau¹
CEA
Chaire "management de l'innovation"
de l'École polytechnique
Chaire "management multiculturel
et performances de l'entreprise"
(Renault-X-HEC)
Chambre de Commerce
et d'Industrie de Paris
Conseil Supérieur de l'Ordre
des Experts Comptables
Crédit Agricole SA
Danone
Deloitte
École des mines de Paris
EDF
Entreprise & Personnel
ESCP Europe
Fondation Charles Léopold Mayer
pour le Progrès de l'Homme
Fondation Crédit Coopératif
France Télécom
FVA Management
Roger Godino
Groupe ESSEC
HRA Pharma
IDRH
IdVectoR¹
La Poste
Lafarge
Ministère de l'Industrie,
direction générale de la compétitivité,
de l'industrie et des services
OCP SA
Paris-Ile de France Capitale
Economique
PSA Peugeot Citroën
Reims Management School
Renault
Saint-Gobain
Schneider Electric Industries
SNCF¹
Thales
Total
Ylios

¹ pour le séminaire
Ressources Technologiques et Innovation
² pour le séminaire Vie des Affaires

(liste au 1^{er} janvier 2010)

EXALEAD, NEUF ANS APRÈS

par

François BOURDONCLE
Cofondateur d'Exalead

Séance du 23 septembre 2009
Compte rendu rédigé par Élisabeth Bourguinat

En bref

Exalead, start-up issue de l'École des mines, est devenue en neuf ans l'un des leaders mondiaux des moteurs de recherche d'entreprise. Le marché des moteurs de recherche grand public, où elle a tenté de se faire une place avec Exalead.com, est dominé par deux géants, Google et Microsoft-Yahoo, et il est très difficile d'y pénétrer. En revanche, Exalead est en train de prendre une place importante à la fois dans le secteur des applications en ligne avec des clients comme Skyrock, ViaMichelin ou Pages Jaunes, et dans le domaine émergent des SBAs (*search-based applications*), qui reposent sur le découplage entre la gestion des bases de données et la mise en œuvre d'applications permettant d'utiliser ces données. En proposant de généraliser le recours aux moteurs de recherche pour les applications d'entreprise, Exalead concurrence les ERP (*Enterprise Resource Planning*) sur une partie de leur marché. Le développement des logiciels devient beaucoup plus agile et rapide, il peut être pris en charge directement par les métiers, et l'utilisation de ces applicatifs coûte jusqu'à six fois moins cher.

*L'Association des Amis de l'École de Paris du management organise des débats et en diffuse
des comptes rendus ; les idées restent de la seule responsabilité de leurs auteurs.
Elle peut également diffuser les commentaires que suscitent ces documents.*

EXPOSÉ de François BOURDONCLE

Bien que la société Exalead ait aujourd'hui pignon sur rue dans le monde de la technologie, elle reste paradoxalement assez mal connue. Fondée en septembre 2000, Exalead est aujourd'hui devenue le leader européen et l'un des leaders mondiaux des moteurs de recherche d'entreprise. En termes de chiffre d'affaires, nous arrivons en troisième position après nos concurrents Autonomy et Endeca. Nous avons réalisé 15,5 millions d'euros de revenus en 2008 et notre croissance s'est fortement accélérée ces deux dernières années. Elle a été de 215 % entre 2006 et 2007 et de 80 % entre 2007 et 2008, ce qui nous a valu plusieurs récompenses. Exalead compte aujourd'hui 150 collaborateurs et 6 filiales, dont une aux États-Unis.

Nous avons plus de 200 clients, des sociétés auxquelles nous vendons notre technologie avec un modèle d'affaires assez intéressant : outre les droits pour les licences des logiciels, nous percevons chaque année 15 à 20 % du chiffre d'affaires en maintenance. Si nous décidions, comme Oracle, d'arrêter de vendre nos produits et de licencier tous nos commerciaux, et même si nous le faisons un 1^{er} janvier, à la fin de l'année nous aurions malgré tout réalisé 70 % du chiffre d'affaires de l'année précédente. Cependant, compte tenu de notre croissance, ces revenus récurrents représentent actuellement moins de 20 % de notre chiffre d'affaires.

Le site Exalead.com

Notre société est connue le plus souvent à travers le site Exalead.com, d'où le fait que beaucoup de gens pensent que nous développons des moteurs de recherche pour le grand public, alors que notre modèle d'affaires est tourné vers les entreprises. Exalead.com est en réalité une sorte de rayonnement fossile de notre tentative, en 2005, de lancer un moteur de recherche grand public.

Les moteurs de recherche grand public nécessitent de maîtriser deux métiers : la technologie des moteurs de recherche et la publicité, ou art de vendre des liens sponsorisés. C'est pourquoi il est très difficile d'imposer un nouveau moteur de recherche. Le secteur est aujourd'hui dominé par un duopole de fait, avec Google d'un côté et l'alliance Microsoft-Yahoo de l'autre. La campagne publicitaire menée par Microsoft aux États-Unis pour lancer Bing lui coûte d'ailleurs très cher (100 millions de dollars) et ses parts de marché ne progressent que lentement. Un moteur de recherche est un outil et non un site de destination : les sites de destination peuvent s'ajouter les uns aux autres, mais une fois que les gens se sont habitués à un moteur de recherche, ils n'en changent pas facilement.

Dès l'origine, nous savions que si nous ne parvenions pas à attirer un trafic suffisant sur Exalead.com, nous pourrions nous en servir comme d'un démonstrateur technologique, et c'est ce qu'il est devenu aujourd'hui. Ce site nous rapporte quelques centaines de milliers d'euros, qui couvrent à peu près ce qu'il nous coûte. C'est un bon outil de marketing à un prix relativement modeste.

Exalead.com tourne sur 8 milliards de pages indexées (contre 30 à 40 milliards pour Google), ce qui en fait la troisième base de données mondiale. En revanche, il est beaucoup plus sophistiqué que Google et propose une expérience d'utilisateur plus riche. En fait, il repose sur une application grand public d'une technologie développée pour les entreprises, avec de nombreuses fonctionnalités qui n'existent généralement pas dans les moteurs de recherche grand public : recherche phonétique, recherche tolérante aux fautes d'orthographe, recherche sur des mots à trous, etc.

Les applications destinées à l'Internet

Nous proposons aux entreprises deux types d'applications, les unes destinées aux sites en ligne, les autres conçues pour les usages internes.

Le marché des applications Internet représente un tiers de nos revenus. Parmi nos grands clients figurent Skyrock, qui indexe un demi-milliard de blogs et qui est un des tout premiers sites mondiaux en termes de trafic ; des sites d'annuaires téléphoniques comme Pages Jaunes et son concurrent 118-218 ; ViaMichelin, qui propose des itinéraires et des recherches d'adresses touristiques ; ou encore Rightmove, un site de petites annonces au Royaume-Uni. Ces sites sont en concurrence directe avec Google et Microsoft-Yahoo, et il est vital pour eux d'innover radicalement et rapidement. Ils répercutent cette contrainte sur nous et constituent notre principale source d'inspiration et d'innovation.

Le site Exalead.com n'attire qu'un million de visiteurs uniques par mois, mais le trafic agrégé des différents sites équipés de notre moteur de recherche dans le monde représente 100 millions de visiteurs uniques par mois, ce qui correspond à peu près à 1 000 requêtes par seconde. Le trafic de Google est de 11 000 requêtes par seconde ; nous représentons donc un petit dixième de Google, ce qui commence à être significatif.

Nous sommes en train de prendre une place importante sur le marché des moteurs de recherche en ligne. La société norvégienne Fast, qui était leader du domaine, a été rachetée par Microsoft et ce dernier est alors devenu un concurrent pour nos clients. Ils préfèrent désormais avoir affaire à un vendeur indépendant comme Exalead plutôt que de s'adresser à Microsoft, car ils n'ont aucune garantie que Microsoft confiera à Fast les innovations dont ils ont besoin pour lutter contre Microsoft et Google...

Les applications destinées aux usages internes

Les applications que nous proposons aux entreprises pour leurs usages internes n'ont rien à voir avec, par exemple, l'indexation des intranets. Ce marché n'a jamais prouvé sa valeur ajoutée car, en général, les pages qui se trouvent sur l'intranet sont les moins intéressantes de l'entreprise. Les données vivantes sont celles qui se trouvent sur les messageries, sur les disques durs, sur les serveurs partagés, dans les bases de données. De plus, c'est un marché très encombré, avec des logiciels *open source*, un logiciel gratuit offert par Microsoft avec Sharepoint, ou encore une offre Google en entrée de gamme.

Un marché émergent

En revanche, nous avons été les premiers à identifier le marché émergent des *search-based applications* (SBAs), qui consistent à utiliser les technologies des moteurs de recherche en lieu et place des bases de données traditionnelles, afin de proposer des applications qui soient plus accessibles à l'utilisateur final. Cette appellation est aujourd'hui reprise par nos concurrents et par les analystes américains, ce qui prouve que nous avons identifié un vrai marché.

Découpler archivage et utilisation des données

Historiquement, les applications d'entreprise se sont toujours appuyées sur les bases de données. Il ne pouvait pas en être autrement car il n'existait aucune autre infrastructure disponible. Progressivement, on est passé de la gestion de données à la gestion des applications sur les données, et on a proposé aux utilisateurs de se servir de ces applications.

Nous partons d'une idée un peu révolutionnaire, qui consiste à distinguer deux sortes d'applications dans les entreprises : d'un côté, tout ce qui concerne la gestion des données, les échanges de données, leur archivage, la garantie de leur pérennité ; de l'autre, toutes les

applications impliquant des interactions avec les êtres humains, qu'il s'agisse de chercher des informations ou de modifier les données. Même la terminologie est éloquente : *base de données* évoque quelque chose de statique ; alors que *moteur de recherche* évoque une action, une interaction. Le premier type d'application est naturellement associé à des technologies de bases de données. En revanche, les bases de données sont très mal adaptées au deuxième type d'application.

Nous proposons de fonder désormais toutes les applications impliquant des interactions avec les êtres humains sur la technologie des moteurs de recherche. Il n'y a aucune raison que les informations interprétables par des êtres humains restent dépendantes des bases de données dans lesquelles elles sont stockées. On peut les rassembler, les agglomérer, les triturer. En découplant les fonctionnalités d'archivage et de structuration des données de l'utilisation de ces mêmes données, on libère énormément d'énergie dans l'entreprise pour développer de nouveaux applicatifs, et ce sans mettre en danger la pérennité des données. Cette approche redonne de l'oxygène aux entreprises par rapport à l'utilisation des ERP (*Enterprise Resource Planning*), qui, certes, ont facilité les opérations de fusions-acquisitions, mais ont aussi énormément rigidifié les entreprises. Avec notre dispositif, l'ERP continue à gérer les flux de données, leur structuration, leurs échanges, leur mise à jour, mais les applicatifs peuvent être développés séparément. On peut ainsi travailler sur 30 ans de conservation de données avec 20 générations successives d'applicatifs.

Des applications innombrables

Voici un exemple d'application que nous avons réalisée pour Gefco, sous-traitant des constructeurs automobiles pour la livraison de véhicules chez les concessionnaires. Gefco disposait d'une application gérée sous Oracle qui était tellement sollicitée que pendant pratiquement toute la durée journalière du travail légal, les concessionnaires ne parvenaient pas à y accéder. De plus, le formulaire de saisie comportait 40 champs : numéro de la commande, nom du client, type d'automobile, etc. Quand un client téléphonait pour savoir où en était la livraison de sa voiture, il était quasiment impossible de lui répondre. L'entreprise a adopté une architecture SBA tout en conservant l'application Oracle pour gérer les mises à jour de la base de données. Nous lui avons proposé un formulaire de recherche à entrée unique : soit le nom du client, soit le type de voiture, soit l'adresse du concessionnaire. Non seulement les concessionnaires ont retrouvé l'accès au système, mais Gefco a énormément réduit ses coûts, entre autres parce que la charge sur la base de données Oracle a fortement baissé et que la société a pu se dispenser de racheter certaines licences.

Les applications sont innombrables. Essayer de trouver un *pilote* sur le site internet de Dell, par exemple, est une expérience très pénible. Il est urgent d'inventer des systèmes beaucoup plus confortables pour les utilisateurs que ceux qui existent actuellement. C'est ce marché que nous sommes en train de défricher.

Un développement beaucoup plus agile

Dans la démarche classique, le développement des applications d'entreprise prend trois ans : un an pour l'élaboration des cahiers de spécifications et deux ans pour les implémentations. Notre démarche permet des modes de développement beaucoup plus agiles, qui s'inspirent des technologies du monde Internet. On peut réaliser des prototypes très rapidement, les tester auprès des utilisateurs, analyser la façon dont ces derniers réagissent, etc.

Le groupe La Poste, par exemple, a souhaité mettre en place un numéro d'appel unique, le 3631, pour que les clients rencontrant des problèmes de distribution du courrier ne soient pas obligés d'aller demander des informations à un guichet. Plutôt que de déployer un outil spécifique de gestion de la relation client (par exemple Siebel, dont La Poste était pourtant déjà cliente), la solution retenue a consisté à aller chercher les données sur les incidents de distribution dans la base où elles étaient stockées et à les fédérer grâce à un moteur de recherche avec le fichier des clients comportant leurs adresses, rues et codes postaux. Une expérimentation a été montée dans une petite commune de France : 10 opératrices ont été

recrutées et un *call center* a été créé. Une communication régionale a été lancée pour annoncer cette expérience. En trois mois, la démonstration a été faite que le dispositif marchait très bien et cela a même été un tel succès qu'il a été déployé en douze mois sur l'ensemble de la France.

Positionnement stratégique

Comme nous ne sommes pas des spécialistes de la logistique, de la relation client, ni de la Business Intelligence, nous passons par des SSII (Sociétés de services en ingénierie informatique) qui, elles, connaissent les clients et savent identifier les problèmes que nos technologies peuvent résoudre. Notre action consiste à éduquer le marché pour créer du *mindshare* (reconnaissance de la marque) chez les clients finaux potentiels, et à évangéliser les SSII et les intégrateurs qui, eux, réalisent les ventes. Toutes proportions gardées, notre positionnement, qui est celui de la vente indirecte, peut être comparé à celui d'Oracle dans les années 1970 : Oracle proposait des bases de données et SAP, People Soft et d'autres sociétés réalisaient des applicatifs pour ces bases de données. Un jour, Oracle a décidé de récupérer la valeur ajoutée en devenant concurrent de ses clients, ce qui a entraîné le basculement de SAP d'Oracle vers IBM.

La concurrence

Nous avons quatre concurrents principaux. La société Autonomy s'est beaucoup diversifiée en rachetant des sociétés et elle est devenue une sorte de "petit IBM". Le rachat de Fast par Microsoft a été pour nous une excellente nouvelle, car Fast était notre concurrent sur notre domaine de prédilection. L'intégration au sein de Microsoft a duré un certain temps et a été compliquée par le fait que des procès visant le PDG de Fast ont éclaté en Norvège. Nous commençons cependant à voir les effets du lobbying de Microsoft à haut niveau dans certains grands comptes. Endeca a un chiffre d'affaires (100 millions de dollars) et un nombre de salariés (500 personnes) 5 fois supérieurs aux nôtres (21 millions de dollars – 100 salariés). La grande différence entre nos deux sociétés, outre la taille, est le fait que l'essentiel du marché d'Endeca se trouve aux États-Unis, tandis que le nôtre se trouve principalement en Europe, différence majeure aux yeux des analystes américains. Enfin, Sinequa est pour nous un concurrent important en France.

Retour sur l'histoire d'Exalead

J'ai commencé ma carrière comme chercheur en informatique théorique chez Digital puis comme adjoint au directeur de la recherche de l'École des mines. J'ai également enseigné pendant 15 ans à Polytechnique et à l'École normale supérieure. Je suis devenu entrepreneur en fondant Exalead, où j'ai exercé à peu près tous les types de métiers : architecture du système, codage de centaines de milliers de lignes de produit, vente, et aujourd'hui surtout marketing stratégique, communication, représentation et lobbying.

Le partenariat avec Altavista

Suite à mon passé de chercheur chez Digital Equipment, j'étais un ami personnel de Louis Monier, fondateur d'Altavista. En 1995, Il m'a proposé de l'aider à créer un moteur de recherche, concept très nouveau à l'époque. Comme je travaillais à l'École des mines, je savais que, pour monter un projet de recherche dans cette école, il fallait trouver un peu d'argent. J'ai donc réalisé un prototype que j'ai vendu à Altavista. Ce contrat m'a permis de monter une équipe chargée de la réalisation du produit pendant deux ans, entre 1996 et 1998. Au début de 1997, le produit a été déployé sur Altavista. En 1998, Altavista a perdu le contrat qui le liait à Yahoo et a changé de stratégie : il est devenu concurrent de Yahoo et s'est lancé dans le métier des médias, des portails, etc. N'ayant pas réussi à renouveler notre contrat avec Altavista, que Louis Monier avait quitté entre-temps, nous nous sommes retrouvés avec une équipe, de l'argent pour la faire fonctionner, et plus de lien particulier avec une entreprise. Nous avons décidé de continuer dans le domaine que nous connaissions bien, celui des moteurs de recherche.

Création de l'entreprise et premiers clients

À la fin de l'année 2000, nous avons essayé d'intéresser France Télécom et d'autres sociétés à notre outil, mais personne n'y comprenait rien. Nous avons alors décidé de créer notre entreprise, ce qui témoignait d'une certaine inconscience : je ne connaissais rien à la conduite d'une entreprise, pas plus qu'Éric Jeux, qui venait de Bouygues Telecom ; Patrice Bertin, le troisième associé, avait travaillé avec moi sur le projet Altavista.

Notre premier client a été Bouygues Telecom, en octobre 2000. Nous lui avons vendu un forfait et quand le trafic a explosé, le contrôleur de gestion a mis très longtemps à se rendre compte qu'au-delà du forfait, il payait des tarifs relativement élevés. Nous avons opéré notre première levée de fonds en janvier 2001. À la mi-2001, nous avons trouvé un deuxième gros client, Scoot, filiale de Vivendi et concurrent des Pages Jaunes. Ce client a été très important pour nous car il nous a fait changer de cap : nous sommes passés d'une logique de moteur de recherche web, caractérisé par une volumétrie énorme, un ensemble de pages indexées homogène et des fonctionnalités relativement simples, à une logique de moteur de recherche d'entreprise, avec une volumétrie très faible, des données hétérogènes mais structurées et des fonctionnalités très sophistiquées. Nous avons, depuis, conservé cet ADN hybride, et c'est ce qui fait notre particularité sur le marché mondial : aucune autre société ne sait concevoir des moteurs de recherche sophistiqués sur de très gros volumes.

Au début de 2002, AOL nous a confié son moteur de recherche pour la France, ce qui nous a mis le pied à l'étrier. Trois mois plus tard, la maison mère signait un accord avec Google, mais nous avons pu conserver notre contrat car nous avons réussi à démontrer la très bonne monétisation du trafic que permettait notre moteur. Au bout de quelques années, cependant, la maison mère d'AOL a exigé que le même produit soit déployé sur l'ensemble du réseau, et nous avons perdu le contrat.

À la fin de 2002, nous avons recruté un directeur commercial très expérimenté qui nous a fait progresser dans l'art de la vente. En 2003, Éric Jeux a quitté la société. En 2008, je suis passé du rôle de PDG à celui de directeur de la stratégie.

Le financement

Notre investisseur historique, Qualis SCA, est entré au capital en janvier 2001, et détient aujourd'hui 83 % des actions. Nous avons réalisé quatre tours de financement, en janvier 2001 (3 millions d'euros), en juillet 2005 (5 millions d'euros), en décembre 2006 (12 millions d'euros) et en mai 2008 (5 millions d'euros).

Le développement international

En 2002, nous avons créé une filiale en Italie, un peu par hasard : notre directeur commercial était très proche de l'ancien directeur d'Altavista software, qui se trouvait être italien.

En 2004, nous avons essayé de créer une filiale aux États-Unis, en y envoyant notre directeur commercial et quelques ingénieurs français, et en recrutant des commerciaux locaux. Cela n'a pas marché car il fallait éduquer le marché plutôt qu'essayer de passer en force.

En 2006, nous avons tenté l'Allemagne, avec beaucoup de difficultés aussi, car les Allemands ont tendance à acheter des produits allemands, en particulier dans le secteur du logiciel. Cette attitude est en train de changer et nos ventes commencent à décoller dans ce pays.

À la fin de 2006, nous nous sommes également lancés au Royaume-Uni, ce qui représentait un défi pour nous car nous allions affronter directement Autonomy, société anglaise à l'origine, ainsi que Fast, société norvégienne et leader du marché à l'époque. Nous avons découvert que les Britanniques sont réfractaires au marketing à l'américaine et en revanche très sensibles aux mérites technologiques des solutions qu'on leur propose. Nous avons

rapidement recruté des clients comme Rightmove, l'équivalent anglais de Seloger.com, ou comme Fish4, une application de ressources humaines qui fait du rapprochement automatique entre curriculum vitae et offres d'emplois. Ils ont préféré notre solution à celle de Fast, entre autres parce que la comparaison en termes de coûts était très flatteuse pour nous. Une recherche chez Rightmove faite avec notre outil, c'est-à-dire avec des technologies entièrement automatiques avec validation a posteriori, revenait 6 fois moins cher qu'avec Fast. Au bout de six mois, notre filière en Grande-Bretagne était à l'équilibre, et elle continue à être très rentable.

À l'été 2008, nous avons lancé une nouvelle tentative aux États-Unis, avec un manager américain et une filiale installée non à New York mais à San Francisco. Nous avons également commencé à nous développer dans le Benelux et en Espagne. Il y a deux ans, le rapport entre notre chiffre d'affaires en France et à l'international était de 50/50, mais comme nous avons beaucoup progressé en France, nous sommes revenus à un rapport de 60 % en France et 40 % à l'étranger.

Politique d'innovation

La R&D d'Exalead représente près de 30 % de nos dépenses. Il s'agit de la R&D entendue au sens anglo-saxon, qui est assez large et comprend par exemple des *proof of concepts* en avant-vente, mais même ainsi, c'est un taux très important par rapport aux pratiques des groupes américains du secteur, où la R&D n'excède généralement pas 5 à 10 %.

Nous considérons en effet que notre atout principal est notre capacité à innover et à mettre sur le marché des produits nouveaux avant Google ou Microsoft, avec des ressources pourtant dix fois moins importantes que les leurs. Nous avons par exemple été les premiers à proposer la recherche d'images sur le web avec détection du visage. Quinze jours plus tard, Google a annoncé dans son blog qu'il en était capable également, et Microsoft en a fait autant un mois après.

Les disciplines mobilisées

Les compétences nécessaires pour élaborer des moteurs de recherche sont extrêmement variées et englobent pratiquement toutes les disciplines informatiques. Nous recrutons des spécialistes en théorie des graphes, bases de données, informatique distribuée, langage de programmation, dictionnaires, structures de données avancées, allocation de mémoire, linguistique informatique, apprentissage par la machine, recherche documentaire. En général, nous ne recrutons jamais quelqu'un qui est spécialiste d'une seule de ces disciplines, car la plupart des problématiques sont transversales.

Les Exalabs

Il y a quelques années, les start-ups embauchaient des PhD qui préparaient des thèses et qu'elles faisaient travailler dans leurs laboratoires, ce qui ne donnait pas grand résultat. Très vite, et même dans des entreprises de 50 personnes, les chercheurs se mettaient à tenir des discours du type : « *Je trouve des solutions géniales et ces débiles d'ingénieurs sont incapables de les implanter.* »

Nous avons changé notre fusil d'épaule : maintenant, nos thésards sont directement intégrés aux groupes produits, et ceux qui travaillent dans les *Exalabs* ont plutôt des profils d'ingénieurs. Ces derniers ont deux grandes missions : ils s'intéressent aux innovations d'usage et conduisent des projets sophistiqués pour des clients exigeants ; ils gèrent l'interface avec la R&D interne et avec les programmes de recherche collaborative, et ils sont chargés d'intégrer les solutions de nos partenaires dans les programmes de R&D. En d'autres termes, ils s'occupent de l'innovation visible. L'innovation cachée est "enfouie" dans les groupes produits et non isolée dans les laboratoires. Ce dispositif fonctionne extrêmement bien.

Les brevets

Sur un marché qui est devenu mondial, on ne peut pas se permettre de décider qu'on n'aime pas les brevets et refuser de jouer le jeu. Quel que soit mon avis personnel sur les brevets concernant les logiciels, je me devais, à l'époque où j'étais responsable social de l'entreprise, de faire en sorte que nous déposions des brevets. C'est ce que nous avons fait depuis le début, et même avant la création de la société. Non seulement nous déposons des brevets, mais nous les défendons, sans quoi cela ne servirait à rien de dépenser des centaines de milliers d'euros.

La recherche collaborative

Nous consacrons beaucoup d'énergie aux programmes de recherche collaborative, ce qui nous permet de nous faire financer une partie de notre effort de R&D. La détection du visage dans les images web est un des résultats obtenus dans le cadre de ces programmes.

Le programme Quaero, lancé par Jacques Chirac et tourné vers les problématiques d'indexation multimédia, a fait couler beaucoup d'encre et a été très mal compris : il a été perçu comme une volonté de concurrencer Google, ce qui n'a jamais été dans les intentions des concepteurs. Ce programme devait être franco-allemand, mais les industriels allemands se sont finalement retirés. Il comprend aujourd'hui 25 participants, dont un tiers d'industriels et deux tiers de laboratoires publics, et dispose d'un budget de 200 millions d'euros, dont 100 sous forme de subventions ou d'aides remboursables. Les industriels sont financés nettement en dessous du seuil de 50 %, et les laboratoires publics à 100 %.

Grâce à ce programme, nous avons pu créer le moteur de recherche vidéo le plus performant au monde, récompensé par la prestigieuse société savante ACM aux USA. Il est capable d'indexer, avec un décalage de quelques heures seulement, les flux vidéo d'informations en français, anglais, arabe et chinois, sur toutes les chaînes de télévision diffusées sur Internet. Il transcrit la bande son, traduit le texte obtenu et permet d'en extraire des noms de personnes, d'organisations ou encore de localisations. De grandes agences internationales peuvent ainsi "traquer" les discours de tel ou tel patron, et les groupes industriels peuvent savoir quelles informations ont été diffusées à travers le monde sur leur compte.

Le *Global Magic Quadrant* de Gartner

Le diagramme de Gartner appelé *Global Magic Quadrant* repose sur deux axes, la vision d'ensemble en abscisses et la capacité opérationnelle en ordonnées, ce qui détermine quatre cases : en bas à gauche les *acteurs de niche*, puis à droite les *visionnaires* ; en haut à gauche les *challengers*, puis à droite les *leaders*. Gartner a beau dire que ce diagramme n'est pas un outil d'aide à l'achat, les grands acheteurs s'y réfèrent en permanence.

Exalead a commencé par être situé par Gartner dans le "purgatoire", c'est-à-dire la catégorie des *acteurs de niche*. Au bout de la troisième année, nous sommes passés dans celle des *visionnaires*, ou sociétés qui sont en train de réussir. Nous sommes aujourd'hui rangés dans la classe des *challengers*, c'est-à-dire des entreprises qui ont démontré, par leur chiffre d'affaires, que leur vision théorique correspondait vraiment à un marché. Avec un peu de chance (et beaucoup de travail), nous figurerons l'an prochain parmi les *leaders*, c'est-à-dire les sociétés qui bénéficient d'une croissance très forte sur le marché américain, l'adjectif *américain* pesant bien sûr de tout son poids.

DÉBAT

Google et Oracle, des concurrents ?

Un intervenant : *Considérez-vous Google comme un concurrent ?*

François Bourdoncle : Ce n'est pas un concurrent, car il n'intervient chez nos clients que pour des applications très simples. Google propose par exemple des indexations de sites Internet, ou d'intranet, et quelques applications de type SBA, mais avec des fonctionnalités très limitées. Compte tenu de la plateforme dont Google dispose, se lancer réellement dans les SBAs reviendrait à vouloir construire une cathédrale sur des sables mouvants.

Cela dit, il n'est pas exclu que d'ici deux ou trois ans Google monte en puissance avec des solutions d'entreprise beaucoup plus sophistiquées. Microsoft a fortement progressé dans ce domaine grâce au rachat de Fast, et comme Google et Microsoft sont en perpétuelle compétition, Google va probablement s'y mettre aussi.

Int. : *Qu'en est-il d'Oracle et d'IBM ?*

F. B. : De l'aveu même de la personne qui était en charge du moteur de recherche d'Oracle, l'entreprise ne l'a vendu que quand elle l'a donné, voire même quand elle a payé le client pour l'adopter... IBM, de son côté, offre cinq ou six moteurs différents, mais aucun ne répond à l'ensemble des besoins des clients. Ces entreprises se doivent d'être présentes rapidement sur ce marché et elles feront ce qu'il faudra pour cela. Le fait qu'Oracle n'y soit pas encore parvenu est sans doute lié à des raisons technologiques, mais aussi à des raisons culturelles et à un problème d'auto-cannibalisation, comme on le voit à travers l'exemple de Gefco : le produit que nous avons vendu à Gefco lui a permis de réduire fortement le nombre de licences achetées à Oracle. Peut-être la solution passera-t-elle par une croissance externe : il y a eu des rumeurs concernant le rachat d'Autonomy par Oracle ; aujourd'hui, on parle plutôt de Hewlett-Packard. Le fait que Microsoft ait racheté Fast oblige tous les grands groupes à évoluer dans cette direction.

La spécificité des SBAs

Int. : *Je n'arrive pas à saisir la spécificité des SBAs par rapport à ce que fait Google. S'agit-il seulement d'une sorte de customisation du moteur de recherche, que vous adaptez aux besoins de vos clients ?*

F. B. : Une mobylette et un 4x4 sont tous deux pourvus d'un moteur à explosion, mais il est difficile d'affirmer qu'ils remplissent les mêmes fonctions. L'offre de Google est plus pauvre fonctionnellement que celle d'Exalead. Pour fabriquer un produit qui réponde aux attentes du client à partir du moteur de Google, il faut lui ajouter de très nombreux modules, ce qui exige beaucoup de main-d'œuvre et peut être source d'erreurs et de désynchronisation lorsque certains des modules évoluent. C'est l'intérêt d'une approche plateforme comme la nôtre, par opposition à une approche produit. Sur des projets critiques, un client préfère se voir proposer une carte mère sur laquelle on ajoutera deux ou trois modules, qu'une série de composants à assembler.

Le point fort d'Exalead

Int. : *Quel sera, selon vous, le point fort d'Exalead dans la durée ? Vous avez créé un moteur de recherche très puissant mais il peut en exister d'autres. Vous proposez des services sur mesure, mais d'autres entreprises peuvent s'y mettre aussi. Vous avez une forte capacité d'innovation, mais les innovations d'usage sont celles qui se copient le plus vite.*

F. B. : Il serait tellement pratique de pouvoir isoler un principe actif unique qui explique l'ensemble de la réussite de l'entreprise... Si la purée de pommes de terre préparée par

Robuchon est meilleure que celle que vous cuisinez chez vous, ce n'est pas simplement parce qu'il y met plus de beurre : la mise en œuvre est aussi importante que les ingrédients.

Un de nos points forts aux yeux de nos clients on-line est incontestablement la capacité que nous leur offrons d'innover extrêmement vite, et de se différencier ainsi de Google et de Microsoft. ViaMichelin utilise notre moteur de recherche non seulement pour trouver le nom d'un restaurant ou d'un hôtel, mais pour positionner les points d'intérêt autour de ce que vous regardez sur la carte : si vous vous déplacez avec la souris, de nouveaux points d'intérêt s'affichent. Google Maps offre aussi ce service, mais pour son propre compte : il ne sait pas le faire pour ses clients.

Lors de la ruée vers l'or, certains ont fait fortune en trouvant des pépites, d'autres en vendant des pelles ; nous sommes plutôt dans le deuxième cas, c'est-à-dire dans une logique d'équipementier. Et plus Google offre des produits performants, plus nos clients sont stressés et nous demandent de nouveaux équipements.

Pour les moteurs de recherche internes, le cocktail qui explique notre succès est plus sophistiqué. Nous offrons à nos clients la possibilité de trancher le nœud gordien qui liait étroitement les applicatifs à la gestion de données, et de libérer ainsi leur énergie pour la création de nouvelles applications. Jusqu'alors, celle-ci était réservée aux départements informatiques ; avec nos moteurs de recherche, les métiers peuvent s'en emparer.

Le logiciel libre

Int. : *Quelle est votre stratégie par rapport aux logiciels libres ?*

F. B. : Dans notre domaine, les logiciels libres sont utilisés surtout pour des sites on-line par des éditeurs qui ont besoin d'ajouter à leurs logiciels des fonctionnalités de recherche et préfèrent ne pas faire appel à nous afin de ne pas perdre une partie de leur valeur ajoutée. Quand la fonctionnalité de recherche n'est pas très stratégique, en revanche, il leur arrive de se tourner vers nous. Nous avons des accords, par exemple, avec des sociétés qui font de l'archivage de messagerie : les volumes à traiter sont tels que le parc informatique qu'elles devraient déployer pour utiliser du logiciel libre serait beaucoup trop coûteux.

En réalité, il s'agit d'un calcul économique de type *make or buy*. Contrairement à ce qu'on peut penser, le logiciel libre n'est pas une solution magique : on ne se contente pas de télécharger un programme et d'appuyer sur "Entrée". Pour intégrer un logiciel libre à un programme et assurer sa synchronisation avec une dizaine de modules, il faut embaucher une dizaine d'ingénieurs, ce qui a un coût. Les logiciels libres sont donc pour nous des concurrents comme les autres, qui présentent des avantages mais aussi des inconvénients.

Pour ce qui est des applications internes, les clients préfèrent généralement s'adresser à un éditeur qui prend la responsabilité du fait que le logiciel fonctionnera.

Les perspectives de développement

Int. : *Vous nous avez montré divers usages des moteurs de recherche, y compris la recherche dans des documents multimédias. Quels sont les nouveaux domaines dans lesquels les applications des moteurs de recherche pourraient se développer ?*

F. B. : Tous les domaines : seuls les applicatifs purement transactionnels pourront échapper à cette évolution. Mais même dans une banque, ils ne représentent qu'environ 10 % des applicatifs : tous les autres ont une interface avec l'être humain et ont donc vocation à être traités sous forme de moteur de recherche.

Ce changement radical de perspective a un coût : l'éducation du marché, qui va être longue. Nous commençons à pouvoir citer de premières références, et les analystes ont perçu le potentiel de cette nouvelle approche. À terme, le marché est aussi infini que celui des bases de données. La vraie question, c'est de savoir comment une petite entreprise comme la nôtre peut s'estimer crédible pour équiper l'ensemble des entreprises. Le bénéfice pour le client final est tout à fait évident, mais ce n'est pas forcément très rassurant de faire appel à une entreprise qui n'a pas encore atteint quelques milliards de dollars de chiffre d'affaires, même si on a déjà vu disparaître des sociétés qui avaient des chiffres d'affaires colossaux, comme Nortel.

Une évolution continue ou discontinue ?

Int. : *Votre moteur de recherche a-t-il évolué de façon continue ou avec des ruptures technologiques ?*

F. B. : Un logiciel doit rester vivant en permanence dans toutes ses ramifications : toute fossilisation d'une partie du code serait dramatique. D'où la nécessité de faire tourner les développeurs. L'évolution du logiciel a été continue jusqu'en 2006, date à laquelle nous avons décidé de revoir entièrement l'architecture que j'avais conçue, ce qui nous a pris trois ans. L'objectif était de supprimer les derniers verrous qui empêchaient notre plateforme d'être déployée sur des fermes de serveurs de grande taille. Certains modules ont ensuite dû être complètement redéveloppés, d'autres ont simplement été corrigés.

Cette opération a constitué une véritable rupture, d'autant plus délicate à gérer que pendant trois ans, nous avons dû continuer à servir nos clients avec le produit existant et ses limitations. Nous devions à la fois les faire rêver sur la future architecture et les convaincre de se contenter provisoirement de celle que nous leur proposons. Cette étape est maintenant derrière nous, et notre produit dispose de l'architecture qui lui permettra de tenir toutes ses promesses.

Int. : *Quelle sera la prochaine rupture ?*

F. B. : Nous allons créer des modules "métiers" spécifiques. Jusqu'ici, nous proposons toujours des produits ad hoc pour nos clients mais, avec le retour d'expérience dont nous disposons aujourd'hui, nous pouvons factoriser un certain nombre d'opérations. Une partie d'entre elles, qui étaient réalisées à l'intérieur des modules, pourront désormais l'être à l'extérieur, et être réutilisées sur de nombreux autres modules.

Les disciplines maîtresses

Int. : *Parmi les différentes disciplines mobilisées pour développer vos moteurs de recherche, quelles sont celles qui comptent le plus aujourd'hui ?*

F. B. : La plus cruciale pour nous est l'apprentissage par la machine (*machine learning*). Jusqu'à présent, le déploiement de nos logiciels s'est fait de façon artisanale : dès que l'on crée des fonctions un peu sophistiquées d'extraction, de compréhension de texte, ou encore de terminologie, on arrive à des projets très lourds. Sans être réellement indispensables pour toutes les applications, ces fonctions constitueraient cependant une plus-value intéressante pour nombre d'entre elles. Notre objectif est d'industrialiser ces processus artisanaux, ce qui exige une forte capacité d'apprentissage des machines. Elles doivent être capables d'analyser les corpus, d'identifier les terminologies et de comprendre quelles sont les relations entre elles. Par exemple, en chimie, il existe des "objets métier" comme *molécule* et *réaction*, et entre les deux une relation de type "*être le catalyseur de*". À partir d'une série d'exemples de ce type, la machine doit identifier automatiquement les notions de *catalyseur*, de *composé chimique* et de *réaction*. Ce genre de chose se fait aujourd'hui à la main, coûte très cher et ne peut concerner que des applications à très forte valeur ajoutée, et donc un marché très restreint. Notre objectif est de rendre ces techniques plus abordables pour pouvoir les étendre à de très nombreuses applications.

Le destin des créateurs de start-ups

Int. : *Au bout de quelques années, les créateurs de start-ups sont souvent un peu dépassés par l'essor de leur entreprise et se retrouvent en général en retrait des tâches de gestion, avec un poste de CTO (Chief Technology Officer), de directeur de la stratégie ou de Business Developer. Vous avez franchi cette étape en devenant directeur de la stratégie. Comment avez-vous vécu cette transition ?*

F. B. : Je n'ai jamais été vraiment chargé de la gestion de l'entreprise : à l'origine, j'avais mon Steve Ballmer en la personne d'Éric Jeux. Puis j'ai dû exercer ce rôle de façon transitoire, jusqu'à ce qu'un nouveau directeur général prenne le relais. Je suis capable de m'occuper de gestion, mais cela me passionne tellement peu que j'en attrape des migraines. J'aurais souhaité être déchargé beaucoup plus tôt du rôle de PDG, car ce qui m'intéresse est vraiment le positionnement stratégique, et c'est d'ailleurs ce que je faisais en tant que chercheur : imaginer des solutions harmonieuses dans des environnements contraints. Mais nos investisseurs n'étaient pas de cet avis et craignaient, si nous recrutions un manager, que je ne m'entende pas avec lui. D'où le fait que la transition a été plus tardive que prévu.

Cela dit, je reconnais que passer d'un statut où on a la décision finale sur tout à une position où on est "conseiller technique de luxe" n'est pas si facile à vivre. Mais je suis heureux de l'avoir fait et j'ai maintenant une existence beaucoup plus sereine.

Présentation de l'orateur :

François Bourdoncle : est l'un des pionniers du marché du logiciel de recherche ; polytechnicien, ingénieur du Corps des Mines ; il a été chercheur dans l'industrie ; professeur associé à l'École polytechnique et l'École normale supérieure, et il est cofondateur en 2000 de la société Exalead, leader européen des moteurs de recherche en entreprise.

Diffusion janvier 2010