



La Journée de
l'Innovation
Collaborative

28 mars 2013

Innover c'est réussir le pari de lancer quelque chose de nouveau sur le marché. Dans le monde numérique, le concept ou l'objet nouveau est le fruit d'une combinaison et d'une spécialisation de technologies génériques pour agréger et transformer des données provenant de sources diverses (données publiques, commentaires sur les réseaux sociaux, photos personnelles,...)

Les cinq années de travaux de R&D de Quaero ont été un laboratoire d'expérimentation sur les règles et méthodes pour faire cohabiter le monde industriel – jugé sur sa capacité à proposer de nouveaux concepts de produits – et le monde des inventeurs de technologies – jugé sur l'excellence scientifique et technique des modules produits. Ces deux mondes étant réunis pour qu'émergent des innovations au bénéfice des usages de services et plateformes numériques.

Dans la suite de ce livre blanc nous partageons notre retour d'expérience sur ces années, en mettant en évidence les méthodes qui nous ont paru essentielles pour réussir l'alchimie du partenariat dans le domaine du traitement automatique des contenus multimédias et multilingues.

#1 Atteindre une masse critique pour structurer efficacement une filière

Le programme Quaero évolue dans un domaine dominé par un nombre restreint de grands acteurs de l'Internet. Il est peu vraisemblable que des acteurs européens, pris individuellement, puissent contrer la force de frappe de ces géants de l'Internet, y compris dans le domaine de la recherche. Le pari de Quaero a été de réussir la structuration au sein d'une filière de l'information numérique multimédia et multilingue d'un groupe multidisciplinaire et multiculturel d'acteurs choisis à la fois sur leur réputation technique et leur volonté et capacité à coopérer. Réunis au sein de Quaero, les laboratoires de recherche, les PME, les industriels qui composent le partenariat ont développé plus de 35 prototypes industriels et ont obtenu, lors de défis scientifiques et techniques, des résultats qui les classent le plus souvent parmi les trois premiers.

Parmi les bénéficiaires principaux de ce regroupement en masse critique, citons :

- La diminution du risque : pour les industriels, la présence d'un grand nombre de fournisseurs permet de diminuer le risque technique, en bénéficiant de la possibilité de choisir le partenaire le plus apte à répondre aux besoins du marché visé.
- L'accroissement des opportunités de valorisation : pour les fournisseurs de technologies génériques la collaboration autour de plusieurs domaines d'application permet de diminuer le risque d'échec commercial et d'augmenter le potentiel de valorisation des technologies produites.
- L'émulation positive : la proximité avec les industriels a permis aux organismes de recherche d'affiner les besoins dans les différents domaines, et de disposer de données réelles et fiables pour concevoir et comparer les algorithmes et méthodes les plus adaptés.
- Les économies d'échelle : des économies substantielles ont pu être réalisées en mutualisant certains coûts. Ainsi une partie des données requises pour la mise au point des algorithmes d'analyse de contenus a été réutilisée sur plusieurs technologies et plusieurs cas d'usage.

Afin de pérenniser les acquis technologiques et méthodologiques de Quaero, et son impact sur le monde de la recherche scientifique, le programme a été à l'origine de la création, au sein du Laboratoire national de métrologie et d'essais¹ (LNE), d'une équipe dédiée à la coordination de campagnes d'évaluation, et de l'Institut des technologies Multilingues et Multimédias de l'Information² (IMMI), qui associe trois partenaires français et allemands (LIMSIS-CNRS (Orsay), RWTH (Aix-la-Chapelle), KIT (Karlsruhe)), pour soutenir un effort de recherche majeur pour développer les technologies de la langue au meilleur niveau international.

¹ Le LNE est un établissement public sous tutelle du ministère en charge de l'industrie

² L'IMMI est une Unité Mixte Internationale du CNRS

#2 Bâtir une vision à moyen et long terme : la R&D, ça prend du temps

Bien que les acteurs du monde numérique se doivent d'être très réactifs pour répondre aux évolutions très rapides des marchés, tendances et apparitions des innovations de l'Internet, nous pensons que la durée de cinq ans du programme a été un bénéfice et un facteur de réussite.

Pour mener à bien les recherches qui permettent de faire progresser l'état de l'art, la nécessité de disposer de suffisamment de temps est certainement l'un des principaux enseignements de Quaero. En effet dans le domaine de l'accès à l'information, les chercheurs et les ingénieurs s'attaquent à des problèmes très ardues, parfois difficiles à définir, nécessitant des recherches lourdes et consommatrices en ressources humaines très pointues. Pour certains domaines technologiques, les recherches s'appuient sur des travaux initiés il y a des décennies. D'autres domaines de recherche sont plus récents. Dans un cas comme dans l'autre, il faut du temps pour faire progresser significativement l'état de l'art, afin de l'amener vers un niveau performance et de fiabilité qui permette d'envisager des applications industrielles.

L'enjeu alors est de faire cohabiter les besoins, souvent à court terme des industriels, avec la nécessité de disposer de suffisamment de temps pour effectuer les recherches. Au sein de Quaero, cette question a été traitée par la mise en place d'un cycle de R&D en spirale avec des échéances annuelles dont les objectifs sont discutés et convenus par toutes les parties prenantes. Pour encadrer ce processus le programme a fait appel à une institution neutre et respectée – la DGA.

En outre le facteur temps a joué un rôle très positif dans l'approfondissement des coopérations entre partenaires. La durée a permis aux équipes issues de milieux différents d'apprendre à travailler ensemble et de tisser des relations de confiance. Ceci s'est manifesté par une diffusion plus large et plus rapide, au sein des prototypes industriels, des résultats obtenus par les chercheurs.

Pour que certains acquis de la filière de R&D, créée au sein de Quaero, perdurent au-delà du terme du programme, une structure pérenne d'encadrement a été créée au sein du LNE. Elle a pour mission d'animer le monde de la recherche par l'organisation de campagnes d'évaluation qui permettent de mesurer les progrès technologiques.

#3 Inscrire le transfert technologique dans les gènes du programme de recherche

«Technology transfer is a contact sport. It depends upon collaboration, teamwork, and interpersonal networking much more than fundamental research does. So the communication skills of the team are critical.» – Dr. Jim Mitchell, Director of Sun Labs.

Le transfert technologique a été une préoccupation centrale des partenaires dès la mise en place du programme. Les acteurs de Quaero sont convaincus que pour être efficace, le transfert technologique doit intervenir dès le démarrage du programme. Puis aller en s'intensifiant tout au long de son exécution.

Au sein du programme, différentes règles et procédures ont été mises en place pour favoriser le transfert technologique :

- Pour commencer, les partenaires de Quaero ont été sélectionnés en fonction de leur capacité et volonté à s'impliquer dans des transferts d'inventions vers des cas d'usage industriels.
- Les technologies retenues dans le programme de R&D du programme devaient au départ répondre à un besoin exprimé par au moins un industriel, avec une prévision de transfert dans un prototype industriel développé dans Quaero ; l'évolution normale des besoins industriels et des perspectives de transfert pouvait conduire à la suspension des travaux sur une technologie, ou au démarrage de travaux sur une nouvelle technologie.
- Les dispositions relatives à la propriété intellectuelle et industrielles, définies par l'accord de consortium, ont été conçues pour permettre le transfert de savoir-faire entre partenaires, tout en protégeant les droits des inventeurs et développeurs.
- Le transfert a été animé et organisé tout au long de l'exécution du programme.

La Direction technique a maintenu un tableau de bord des technologies produites et des transferts en cours ou à venir. Ce tableau a été présenté et discuté avec les partenaires à l'occasion de toutes les réunions plénières du programme (deux réunions annuelles regroupant environ 120 participants).

Aussi nous soulignons l'importance de la présence des PME, spécialistes de domaines technologiques pour établir un continuum allant de l'invention à la mise en œuvre au sein de prototypes industriels. En effet, ces PME ont eu pour rôle de transformer algorithmes et méthodes en systèmes fiabilisés applicables aux domaines d'applications visés.

#4 Etablir une gouvernance équilibrée et transparente

L'innovation naît de la rencontre entre une technologie, un besoin et une idée d'usage. Pour éviter que se créent des barrières entre inventeurs et fournisseurs de technologies d'une part, et intégrateurs de ces technologies d'autre part, il est primordial que ces acteurs puissent parler d'égal à égal. Ainsi, il est essentiel que les décisions qui engagent l'avenir et l'exécution du programme prennent en compte les considérations, opinions et souhaits de toutes les composantes du partenariat.

Au sein de Quaero, le Comité de pilotage regroupe les partenaires industriels, académiques et les acteurs institutionnels dans une structure en charge de toutes les décisions stratégiques, techniques, budgétaires et opérationnelles. En interaction étroite avec OSEO, l'autorité de financement et de suivi, ce comité a notamment encadré le processus d'élargissement du partenariat (voir le point #5).

Aussi, en maintenant un dialogue stratégique équilibré entre concepteurs de technologies (les laboratoires de recherche et certaines PME) et les intégrateurs de ces technologies (les industriels comprenant grands groupes et PME), le Comité de pilotage a eu un rôle essentiel dans l'établissement des orientations techniques et scientifiques des recherches et développements.

Au final, depuis le lancement du programme, le Comité de pilotage s'est réuni plus de 30 fois pour prendre des décisions sur des aspects stratégiques ou opérationnels.

#5 Garder une agilité en ouvrant régulièrement le programme à de nouveaux partenaires

A deux reprises le partenariat de Quaero s'est élargi pour intégrer de nouveaux partenaires. Yacast a rejoint le programme en 2009, suivi en 2011 par les sociétés A2iA, Easy Life Mobile, Itesoft, Movea et SYSTRAN.

En ouvrant des nouvelles perspectives tant industrielles que techniques, l'arrivée de ces nouveaux partenaires, tous des PME, a considérablement dynamisé le programme. D'une part, les nouveaux usages industriels visés ont permis d'élargir le champ d'application des recherches entreprises et des technologies développées. D'autre part, en apportant des nouvelles technologies, des nouvelles données, et en exprimant de nouveaux besoins, ces arrivants ont aidé à consolider le socle technologique du programme. Par exemple, les archives radio et télévision de Yacast ont été mises à profit pour améliorer certains algorithmes d'analyse de contenus, le savoir d'A2iA en reconnaissance de l'écriture manuscrite a permis de passer un cap technique en terme de traitement de documents papier.

De plus, l'apport de nouveaux savoir-faire et idées a contribué à améliorer certains des scénarios d'acteurs "historiques" du programme. Par exemple, l'association des modules de reconnaissance d'image de LTU technologies avec les bases de données audio-visuelles de Easy Life Mobile a permis d'investiguer de nouveaux modes de navigation et de sélection de vidéos dans les scénarios de la télévision connectée et sociale de Technicolor. L'arrivée des nouveaux partenaires a donc été une source de dynamisme très bénéfique pour le programme.

#6 Permettre de gérer avec réactivité des évolutions liées à l'environnement

Quaero a été confronté à des besoins d'évolution tout au long de sa phase d'exécution. Par exemple, des changements de périmètre et de priorités stratégiques chez un partenaire ont conduit à l'arrêt d'un prototype industriel, au profit de l'extension des objectifs pour un autre.

Pour d'autres partenaires industriels, l'évolution des perspectives de marché les a conduit à modifier les priorités et les objectifs au sein de leur développement de prototypes, nécessitant de réduire et ralentir certaines activités, et au contraire d'en accélérer et d'en renforcer d'autres. Enfin, nous avons pu saisir des opportunités d'innovation exploitant les technologies développées, grâce à l'implication de nouveaux partenaires, comme cela a été souligné dans le point # 5. Pour optimiser l'efficacité d'un programme tel que Quaero, la possibilité de bâtir une vision à moyen et long terme, grâce à la durée de cinq ans de la phase de R&D, doit s'accompagner d'une flexibilité dans la gestion du contenu et des priorités du programme. L'environnement évolue rapidement dans le monde numérique, et les besoins et perspectives des partenaires industriels suivent ces évolutions.

Quaero a pu mettre en œuvre ces évolutions, et cela a été une source majeure de dynamisme du programme. Cependant les mécanismes permettant de valider et d'implémenter des évolutions dans le partenariat et le contenu du programme peuvent encore être améliorés.

#7 Construire une structure de recherche partagée : le pôle de technologies pluridisciplinaires

Le programme Quaero a mis en place une structure de recherche relativement indépendante des projets applicatifs, mais avec une forte capacité de transfert technologique et une organisation favorisant la « coopétition » et l'évaluation comparative.

Quaero a rassemblé des acteurs du monde de la recherche publique ou privée ayant, d'une part, la capacité à faire progresser l'état de l'art sur les domaines technologiques du programme et, d'autre part, la capacité de valoriser leurs résultats. En particulier, le programme s'est appuyé sur les résultats de campagnes d'évaluation nationales et internationales pour identifier les acteurs. Les laboratoires ainsi identifiés ont adhéré au principe de l'évaluation comparative qui, dans le contexte Quaero, s'appuie sur des données représentatives des besoins des projets applicatifs (aussi appelées corpus).

La structure de recherche partagée est organisée en une dizaine de domaines technologiques (moteur de recherche, traitement de la langue, image, vidéo, musique, ...) couvrant une quarantaine d'actions de recherche. Plusieurs acteurs peuvent contribuer à chaque action de recherche, au moins l'un d'entre eux ayant la mission de livrer un ou plusieurs modules technologiques aux projets applicatifs. Chaque acteur doit avoir la capacité de participer aux campagnes d'évaluation (voir #8).

La disponibilité de grands corpus est vitale pour le programme Quaero, qui a consacré environ 10% de son budget pour collecter et annoter des données. Ces données servent, d'une part, à l'évaluation des résultats des travaux de recherche et, d'autre part, à l'élaboration des technologies du programme qui sont pour la plupart fondées sur des méthodes statistiques. Ces corpus et les protocoles d'évaluation associés constituent le principal vecteur pour introduire les besoins des projets applicatifs dans les objectifs de recherche, amenant les projets applicatifs à fortement contribuer à leur préparation.

Les communications scientifiques présentées lors des ateliers semestriels, permettent de maximiser les échanges scientifiques et techniques. Les partenaires sont également fortement encouragés à publier leurs résultats (plus de 700 publications à ce jour) et à prendre part aux campagnes d'évaluation nationales et internationales sur les thématiques du programme.

#8 Choisir les meilleures approches et mesurer les progrès en organisant régulièrement des campagnes d'expérimentations scientifiques

L'expérimentation scientifique dans le domaine du traitement sémantique des données non structurées suppose une organisation particulière. Il faut que les systèmes développés par les différentes équipes de recherche soient testés sur les mêmes données, selon des protocoles communs, et de manière simultanée, afin que les résultats soient comparables et ne comportent pas de biais. Cette organisation communément appelée « campagne d'évaluation » permet à la fois de comparer les approches imaginées par différents chercheurs et de mesurer les progrès accomplis.

Cette organisation, indispensable à la rigueur scientifique des développements et à la transparence des échanges entre créateurs et intégrateurs de technologies, suppose un environnement adapté. Il faut en effet disposer des infrastructures de recherche et de coordination nécessaires, à savoir des entités dédiées à l'organisation de telles campagnes, et fédérer les chercheurs autour de ces infrastructures.

La méthodologie, conçue il y a déjà une trentaine d'années dans le cadre des programmes de recherche de la défense américaine avec le soutien du NIST¹, a été systématisé au sein de Quaero, avec le soutien du LNE et d'autres entités. Les moyens nécessaires ont été alloués et les équipes ont été mises en place pour que tous les développements scientifiques du programme puissent être évalués.

Cette infrastructure est portée par une vingtaine de personnes au sein de 6 entités, représentant environ 8 équivalents temps plein et travaillant au service des centaines de chercheurs du programme. Grâce à elle, plus de 30 campagnes d'évaluation sont organisées chaque année et produisent des centaines de mesures scientifiques, que les chercheurs comme les industriels utilisateurs et les équipes de management utilisent pour orienter leurs choix.

Cette infrastructure a également permis de mesurer les nombreux progrès accomplis par les chercheurs de Quaero pendant la durée du projet. À titre d'exemple, dans le domaine de la transcription automatique de la parole, outre les progrès facilement mesurables comme l'augmentation du nombre de langues couvertes et la réduction des temps de développement, les taux d'erreurs ont été réduits de moitié en seulement cinq ans, ce qui est deux fois plus rapide que les progrès habituels.

¹ National Institute of Standards and Technology, une agence du Département du Commerce des États-Unis

Fournir les technologies qui permettront le multilinguisme

L'extension du modèle Quaero au niveau Européen

Par Joseph Mariani – Directeur, Institut des technologies Multilingues et Mutimédias de l'Information (IMMI-CNRS)

Permettre à chacun de s'exprimer dans sa langue et de préserver sa culture, tout en lui permettant de communiquer avec les autres, tel est l'enjeu du multilinguisme au temps de la globalisation et de la construction d'une Union Européenne fière de sa diversité linguistique. Du fait du coût humain et financier qu'il représente, le multilinguisme ne sera possible que s'il existe des technologies de qualité qui apportent leur aide à relever cet enjeu. Mais peut-on accepter que ces solutions technologiques à un enjeu stratégique pour l'Europe soient seulement apportées par Google avec Google Translate, Apple avec SIRI ou IBM avec Watson ? La recherche et l'industrie européennes sont capables de fournir leurs propres réponses, adaptées aux besoins spécifiques de la culture et du marché communautaires, à la condition d'une volonté politique des Etats-Membres et de la Commission, et de la mise en place d'un programme de taille suffisante et d'architecture adaptée et efficace.

Quaero est le modèle qui pourrait inspirer un tel programme en reprenant les points forts qui ont fait son succès et qu'on ne trouve pas actuellement dans les mécanismes des projets de la Commission Européenne : conduite dans un même programme, d'une durée et d'un financement suffisants, de plusieurs projets applicatifs dirigés par des entreprises et développement, associant les meilleurs laboratoires de recherche, de nombreuses technologies concernant plusieurs médias et plusieurs langues ; développement de technologies dans

un contexte de coopération pour alimenter les applications, une même technologie pouvant être utilisée par plusieurs applications et plusieurs technologies pouvant nourrir l'une d'entre elles ; utilisation systématique et régulière de l'évaluation des technologies et production des corpus nécessaires à leur développement et à leur test ; adaptation du programme au fil du temps, avec l'arrivée de nouvelles applications et de nouveaux partenaires, ou leur départ.

Ceci en élargissant l'actuelle participation franco-allemande de Quaero à l'ensemble de la communauté de la recherche et de l'industrie européenne, en étendant les domaines applicatifs, en s'attaquant au verrou scientifique du traitement du sens à travers l'utilisation de données multimédias et en veillant à couvrir l'ensemble des langues qui font le patrimoine commun de l'Europe.



Contact

techday@quaero.org

www.quaero.org