



Communiqué de presse

Grenoble, le 27/06/2014

Découvrir, expérimenter, mettre en œuvre et appliquer le Big Data : les partenaires industriels et universitaires associés dans Datalyse, projet d'Investissements d'Avenir en Big Data, font le point à l'issue d'une année de travaux en commun.

Datalyse est un projet financé dans le cadre du Programme d'Etat des Investissements d'Avenir, Développement de l'Economie Numérique, Appel à projets « Cloud Computing » n°3 – Big Data, labellisé par les centres de compétitivité MINALOGIC et PICOM. Il a été lancé le 1er mai 2013 pour une durée de 42 mois.

***Business & Decision**, sa filiale **Eolas**, le **Groupe des Mousquetaires**, le **Laboratoire Informatique de Grenoble (LIG)** avec ses équipes **SLIDE** (ScaLable Information Discovery and Exploitation), **ERODS** (Efficient and Robust Distributed System) et **TYREX** (Types and REasoning for the Web), le **Laboratoire d'Informatique Fondamentale de Lille (LIFL)** avec l'équipe-projet **SPIRALS**, **Inria Saclay** avec l'équipe **OAK**, et le **Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier (LIRMM)** ont présenté les premiers prototypes devant la **DGCIS** au bout d'une année d'un travail collaboratif au sein du programme **Datalyse**.*

Datalyse est un projet de recherche en Big Data aux objectifs ambitieux.

L'objectif de Datalyse est de créer une méthodologie Big Data et une collection de prototypes et d'outils de traitement de données, les *datalyseurs*, qui permettent le passage à l'échelle des gros volumes de données rencontrés dans la pratique. Datalyse définit plusieurs étapes nécessaires dans l'exploitation efficace des Big Data : la préparation, la transformation, l'extraction de valeur et la visualisation.

La première phase est la création du concept de *datalyse* et de *datalyseurs*. Dans l'ère du Big Data, le sésame appartient à celui qui sait transformer le volume, la variété, la véracité et la vitesse en valeur. Une *datalyse* est définie comme un processus dynamique donnant une information raffinée, enrichie et agrégée par une ou plusieurs sources de données. Ces données peuvent être produites par des traitements, des analyses, des élargissements appliqués et choisies avec un angle particulier pour en tirer ou en partager de nouvelles connaissances, des décisions et des actions afférentes.

Les solutions scientifiques qui permettent la production de *datalyses* sont appelées des *datalyseurs*. Datalyse a pour objectif la formalisation et la réalisation de *datalyseurs* qui transforment, en mode juste-à-temps, des données brutes créées par des humains ou par des machines en information actionnable. Tout comme la catalyse est l'action d'une substance sur une transformation chimique dans le but d'augmenter sa vitesse de réaction, une *datalyse* est l'action d'opérations d'analyse et de transformation de données dans le but d'augmenter leur qualité et leur vitesse de traitement et de diffusion.

Datalyse introduit plusieurs innovations. Le premier objectif est de faire évoluer positivement le développement d'applications Big Data afin de bénéficier des possibilités nouvelles qu'elles offrent. La seconde phase est l'utilisation de ces *datalyseurs* pour faire évoluer le

développement d'applications Big Data avec la création d'interfaces dans le cadre d'une architecture dédiée.

La méthode de création de *datalyseurs* part des besoins des traitements intelligents pour Big Data hétérogènes, en particulier dans les domaines de l'Open Data, du commerce de détail, du marketing digital, du pilotage de Datacenter. Dans un premier temps, les *datalyseurs* nécessaires à chaque cas d'usage ont été identifiés. Ensuite, dans un effort d'unification, une architecture générique pour le traitement des Big Data a été proposée.

L'architecture de Datalyse, d'un point de vue aussi bien conceptuel que logiciel, est la toute première proposée en la matière, tant dans le domaine de la recherche que dans le domaine industriel. Elle témoigne d'une véritable collaboration entre Business & Decision Eolas, le Groupement des Mousquetaires et les partenaires académiques. Enfin, cette architecture dicte une méthodologie de traitement de données Big Data dont l'applicabilité va au-delà des cas d'usage étudiés.

Datalyse présente ses premiers prototypes devant la DGCIS

Les premiers prototypes d'applications Big Data utilisant des *datalyseurs*, issus de la première année de travaux, sont présentés à la Direction générale de la compétitivité, de l'industrie et des services (DGCIS). Ils sont au nombre de quatre.

Prototype Pilotage de Datacenter dédié à l'analyse du réseau et du trafic, ainsi qu'à l'efficacité énergétique :

- ✓ statistiques de trafic de centres de données,
- ✓ données de fonctionnement du Datacenter par la collecte, le traitement et l'analyse des données issues des nombreux points de mesure (capteurs),
- ✓ données : les MBD (Monitoring Big Data) sont de gros volumes de données générées par des machines

Prototype Linked Open Data (LOD) :

- ✓ accéder à des données ouvertes, propres et enrichies (application testée au Musée dauphinois de Grenoble lors de Museomix),
- ✓ utiliser les données UBD (User Big Data), des règles sémantiques et les connecter avec d'autres solutions Open Data,
- ✓ produire les interfaces d'accès aux données (API) et des outils de visualisation.

Prototype Digital Marketing :

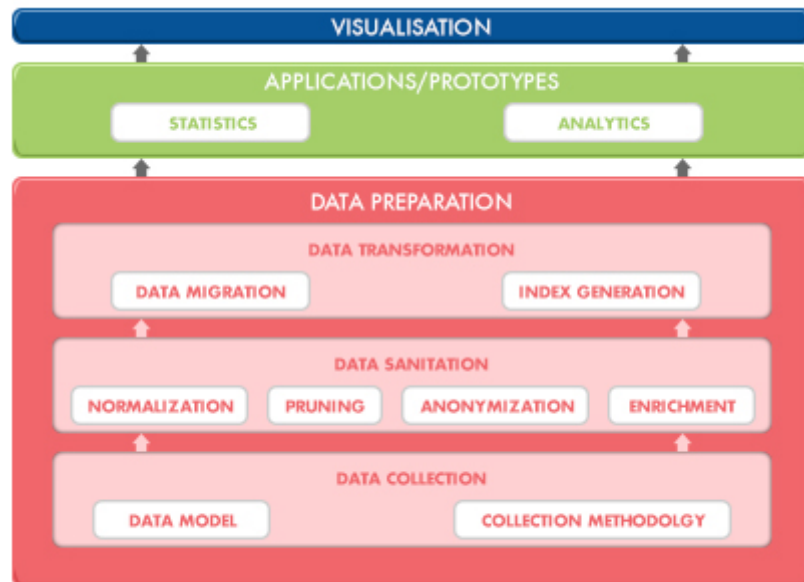
- ✓ fouiller les données de navigation issues des applications hébergées,
- ✓ utiliser les données UBD (User Big Data),
- ✓ produire des statistiques centrées clients pour l'optimisation des sites marchands en ligne.

Prototype Commerce de détail (retail) :

- ✓ déterminer les mécanismes qui conduisent à un comportement d'attrition,
- ✓ utiliser les données UBD (User Big Data),
- ✓ soutenir les offres promotionnelles,
- ✓ challenger les approches BI conventionnelles.

Datalyse est une architecture Big Data et une modélisation de *datalyseurs*

Les *datalyseurs* créés ont porté sur les trois niveaux de l'architecture : **Data Preparation** pour la collection, le nettoyage (cleaning, enrichment, segregation, linkage) et la transformation ; **Data Exploitation** pour les applications et prototypes analytiques et statistiques ; **Data Visualisation** pour mieux comprendre les données.



© Laboratoire d'Informatique de Grenoble (LIG)

Les premiers retours d'expérience ont confirmé l'approche novatrice de Datalyse qui a nécessité un important travail de coordination initiale pour créer la méthodologie liée avec, d'une part, la richesse et la diversité des besoins des différents cas d'usage et, d'autre part, l'hétérogénéité des données.

Les partenaires de Datalyse poursuivent leurs travaux.

Les travaux du groupe se poursuivent durant la seconde année du projet et portent sur des thèmes principaux suivants :

- extension de la méthodologie de développement d'applications Big Data (infrastructure, architecture, élasticité, streaming),
- *datalyseurs* (couplage plus fort des parties stockage et indexation avec les parties transformation et fouille),
- interface d'applications et de *datalyseurs*,
- montée en charge d'acquisition des données Big Data, en volume et en variété.

Contact

Gérald Dulac

e-mail gerald.dulac@businessdecision.com

Tél. 00 (4) 76 44 50 50

Plus d'informations : www.datalyse.fr

A propos de Business & Decision et de Business & Decision Eolas

Business & Decision est un Groupe international de services numériques, leader de la Business Intelligence (BI) et du CRM, acteur majeur de l'e-Business. Le Groupe contribue à la réussite des projets à forte valeur ajoutée des entreprises et accompagne ses clients dans des domaines d'innovation tels que le Big Data et le Digital. Il est reconnu pour son expertise fonctionnelle et technologique par les plus grands éditeurs de logiciels du marché avec lesquels il a noué des partenariats. Fort d'une expertise unique dans ses domaines de spécialisation, Business & Decision offre des solutions adaptées à des secteurs d'activité ainsi qu'à des directions métiers. Présent dans 15 pays, Business & Decision emploie plus de 2 500 personnes en France et dans le monde.

Au sein de Datalyse, Business & Decision a pour rôle de développer les démonstrateurs relatif au décisionnels sur l'entrepôt intelligent et de promouvoir le projet et ses solutions dans son ensemble.

Business & Decision Eolas est spécialiste de l'Internet. Expert en hébergement, digital marketing et développement d'applications, Eolas assure la réussite de nombreux projets en e-commerce, e-administration et e-business. Eolas emploie 125 personnes en Rhône Alpes. L'entreprise est fondateur de la Green Ethiquette (<http://www.greenethiquette.fr>, un engagement pour les usages des NTIC éco-responsables), et a participé à deux programmes de recherche sur les Datacenters : Energetic (dans le cadre de l'appel FUI) et CtrlGreen (projet ANR). Eolas a créé un des premiers Green DataCenters en France, à Grenoble. Eolas est le chef de file du projet Datalyse et a la responsabilité de la mise en mode « PaaS » (Platform as a Service) de l'Entrepôt intelligent et du développement des démonstrations de pilotage des processus, de territoires de données ouvertes sur le territoire et du marketing digital.

Contact : frederic.dulac@businessdecision.com, safa.baghai@businessdecision.com

Plus d'informations : www.businessdecision-eolas.com et www.businessdecision.com

A propos du LIG et de ses équipes HADAS, ERODS et TYREX

Le Laboratoire d'Informatique de Grenoble (LIG) est un laboratoire d'envergure, dont les partenaires académiques sont : le CNRS, Grenoble INP, l'Université Joseph Fourier, Inria Grenoble Rhône-Alpes, l'Université Pierre-Mendès-France et l'Université Stendhal. Le LIG rassemble près de 500 chercheurs, enseignants-chercheurs, doctorants et personnels en support à la recherche. Le projet scientifique du LIG est l'Informatique ambiante et durable". L'ambition est de s'appuyer sur la complémentarité et la qualité reconnue des 22 équipes de recherche du LIG pour contribuer au développement des aspects fondamentaux de la discipline (modèles, langages, méthodes, algorithmes, systèmes) et pour développer une synergie entre les défis conceptuels, technologiques et sociétaux associés à cette thématique.

Le LIG est structuré en 5 axes de recherche et regroupent 22 équipes. Les 3 équipes qui participeront au projet Datalyse sont l'équipe HADAS, l'équipe ERODS et l'équipe-projet commune avec Inria TYREX.

Contact Sihem.Amer-Yahia@imag.fr

Plus d'informations : www.liglab.fr

A propos du LIFL et de l'équipe Spirals

Le laboratoire d'informatique fondamentale de Lille (LIFL) est le laboratoire de recherche en informatique situé à Lille. C'est une unité mixte de recherche avec le CNRS et l'Université Lille 1. Il est rattaché à l'Institut des Sciences de l'Information et de leurs Interactions (INS2I) du CNRS. Le LIFL est également partenaire de l'Université Lille 3 et du centre de recherche Inria Lille - Nord Europe avec lequel il partage dix équipes et équipes-projets, dont l'équipe Spirals, précédemment ADAM.

Les recherches de l'équipe Spirals portent sur la notion d'adaptation, en particulier pour les logiciels de type intergiciels et pour les applications réparties. Deux défis sont ainsi visés par l'équipe : l'adaptabilité dans toutes les phases du cycle de vie du logiciel, de la conception à l'exécution de celui-ci, et l'adaptation à multi-échelles, i.e., la capacité à produire du logiciel qui s'adapte à différentes granularités d'environnement distribué, typiquement des plates-formes de cloud computing aux réseaux de capteurs.

Contact Lionel.Seinturier@lifl.fr

Plus d'informations sur www.lifl.fr

A propos l'Inria Saclay et de l'équipe Oak

Aujourd'hui, les technologies numériques rendent les transports plus autonomes et plus sûrs, les maisons plus intelligentes, l'agriculture plus respectueuse de l'environnement...Elles sont à l'origine de nouveaux services, transforment en profondeur nos modes de vie et enrichissent notre quotidien.

Pour se développer, notre société compte toujours plus sur ces technologies numériques qui restent souvent invisibles. Elles sont issues de travaux de recherche longs et complexes associant sciences informatiques et mathématiques. Inria, institut public de recherche, a pour mission de faire progresser la connaissance que des industriels partenaires transforment en produits et services accessibles au plus grand nombre.

Créé il y a 45 ans, Inria réunit aujourd'hui 3000 chercheurs, inventeurs du monde numérique

Contact ioana.manolescu@inria.fr

Plus d'informations: <https://team.inria.fr/oak/>

A propos du LIRMM

Le Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier - LIRMM - est une unité mixte de recherche, dépendant conjointement de l'Université Montpellier 2 et du Centre National de la Recherche Scientifique. Il est situé sur le Campus Saint-Priest de l'UM2. Ses activités de recherche positionnent pleinement le LIRMM au cœur des sciences et technologies de l'information, de la communication et des systèmes. Ainsi, de l'information aux systèmes, de la technologie à l'humain et aux usages, les activités de recherche du LIRMM concernent : la conception et la vérification de systèmes intégrés, mobiles, communicants; la modélisation de systèmes complexes à base d'agents; les études en algorithmique, bioinformatique, interactions homme-machine, robotique, et le web.

Le groupe de recherche sur les données ouvertes et liées prenant part au projet Datalyse étudie des méthodes et des techniques permettant de faciliter la publication et la réutilisation des données ouvertes. Dans le cadre de Datalyse, l'expertise du groupe sur les technologies du web sémantique sera mise à disposition du projet pour les *datalyseurs* sur l'Open Data et l'E-commerce, et en s'appuyant sur les résultats du projet ANR Datalift (<http://datalift.org>). Les activités de recherche porteront aussi sur des techniques permettant d'identifier des liens entre données similaires à l'échelle du Web, un vaste problème qui revient comme le dit l'expression à chercher une aiguille dans une botte de foin.

Contact : zohra.bellahsene@lirmm.fr

Plus d'informations : www.lirmm.fr

A propos du Groupement des Mousquetaires

Avec plus de 3 500 points de vente en Europe et un chiffre d'affaires de 39,9 milliards d'euros en 2013, les Mousquetaires sont des acteurs majeurs de la grande distribution. Créé en France en 1969, ce Groupement, fondé sur l'initiative privée, rassemble aujourd'hui 2 921 chefs d'entreprise indépendants et 146 000 collaborateurs. Il intègre également un important pôle industriel. Onzième groupe agroalimentaire en France avec 61 usines implantées dans nos régions, ce pôle est présent dans de nombreuses activités : boissons, céréales, fruits et légumes, huiles et condiments, produits carnés, produits de la mer et produits laitiers. Les Mousquetaires sont également présents au Portugal, en Belgique, en Pologne et dans les Balkans. Ses enseignes sont : Intermarché, Netto, Bricomarché, Brico Cash, Rody et Poivre Rouge. Les Mousquetaires placent le consommateur au cœur de leurs préoccupations en lui assurant le meilleur rapport qualité/prix.

Contact : slery@mousquetaires.com

Plus d'information : www.mousquetaires.com