

Rapport d'évaluation

Nom de l'unité :	Entrepôts, Représentation & Ingénierie des Connaissances
Acronyme de l'unité :	ERIC
Label demandé :	EA
N° actuel :	3083
Nom du directeur (2012-2013) :	Jérôme Darmont
Nom du porteur de projet (2014-2018) :	

Membres du comité d'experts

Président :	Monsieur Jin-Kao Hao, Université d'Angers
Experts :	Monsieur Bernard Espinasse, Aix-Marseille Université (AMU)
	Monsieur Guy Melançon, Université de Bordeaux I

Représentant(s) des établissements et organismes tutelles de l'unité :	
	Madame Isabelle Lefort, Vice-Présidente du Conseil Scientifique, et de la recherche, de l'Université Lyon 2

1 • Introduction

La visite s'est déroulée le 14 janvier 2013 à l'ERIC dans les locaux de l'Université de Lyon 2. Il s'agit d'une évaluation à mi-parcours sollicitée par les établissements de tutelle (Universités de Lyon 1 et 2), la dernière évaluation AERES ayant eu lieu en février 2010.

La journée s'est déroulée selon le programme suivant qui a été défini conjointement par le comité de visite et la direction de l'ERIC:

9h30 : Présentation du bilan et du projet scientifique du laboratoire, discussion

12h : Repas

13h30 : Entretiens (direction des universités de tutelle, doctorants, personnel administratif et technique, direction du laboratoire

15h : Délibération du comité

17h : Compte-rendu du comité

Historique et localisation géographique de l'unité :

Fondé à Lyon 2 en 1995, le laboratoire ERIC a opéré à partir de 2009 un regroupement, officialisé en 2010, avec l'équipe MA²D (méthodes et algorithmes pour l'aide à la décision) de Lyon 1, sur la base de complémentarités thématiques en matière de recherche dans le domaine du décisionnel, d'une sensibilité commune au domaine des sciences humaines et sociales et de nombreuses collaborations sur les plans scientifiques et pédagogiques. Finalement, ERIC a intégré en 2012 l'Institut des Sciences de l'Homme (ISH), Unité de Service et de Recherche du CNRS et de l'Université de Lyon. La période 2009-2012 a donc été un moment charnière d'intégration de nouveaux collègues au sein d'ERIC.

Le laboratoire ERIC occupe une position originale dans le paysage informatique lyonnais, par son positionnement scientifique ciblé sur l'informatique décisionnelle, alors que les trois autres laboratoires d'informatique de la place de Lyon, le LIRIS, le LIP et le DISP, sont positionnés sur des créneaux plus généraliste pour le premier (image, données, connaissances, services) et également très spécialisés (parallélisme et systèmes de production, respectivement) pour les deux autres. Toutefois, des proximités thématiques existent avec des équipes du LIRIS et du DISP, ainsi qu'avec l'équipe MODEME du Centre Magellan (laboratoire de gestion) de Lyon 3, qui se traduisent par des participations croisées à des jurys de thèse ou l'organisation commune de manifestations scientifiques.

Équipe de Direction :

Compte-tenu de la double tutelle d'ERIC, le laboratoire est dirigé par un directeur d'un des deux établissements de tutelle (J. Darmont, Lyon 2) et un directeur adjoint de l'autre (S. Bonnevey, Lyon 1). Un conseil de direction du laboratoire, constitué du directeur, du directeur adjoint, des responsables d'équipes de recherche (F. Bentayeb et J. Velcin) et de la responsable administrative et financière (V. Gabriele Pietroforte), se réunit tous les mois entre les conseils de laboratoire pléniers.

Effectifs de l'unité :

Effectifs de l'unité	Nombre au 30/06/2012¹	Nombre au 01/01/2014²	2014-2018 Nombre de produisants du projet³
N1 : Enseignants-chercheurs titulaires et assimilés	22	22	20
N2 : Chercheurs des EPST ou EPIC titulaires et assimilés	0	0	0
N3 : Autres personnels titulaires (n'ayant pas d'obligation de recherche)	0	0	0
N4 : Autres enseignants-chercheurs (PREM, ECC, etc.)	0	0	0
N5 : Autres chercheurs des EPST ou EPIC (DREM, Post-doctorants, visiteurs etc.)	0	0	0
N6 : Autres personnels contractuels (n'ayant pas d'obligation de recherche)	0	0	0
TOTAL N1 à N6	22	22	20
Taux de producteurs	91 %		

Effectifs de l'unité	Nombre au 30/06/2012	Nombre au 01/01/2014
Doctorants	23	
Thèses soutenues	10	
Post-doctorants ayant passé au moins 12 mois dans l'unité *	3	
Nombre d'HDR soutenues	2	
Personnes habilitées à diriger des recherches ou assimilées	11	9

2 • Appréciation sur l'unité

Points forts et possibilités liées au contexte :

1. Un positionnement original et singulier dans l'environnement lyonnais avec de nombreuses collaborations académiques (au niveau national et international) et industrielles.
2. Une implication très forte dans la formation de niveau master, avec deux parcours dont un européen Erasmus Mundus.
3. Une amélioration très nette des supports de publication et une augmentation significative de financement sur projets nationaux et européens depuis l'évaluation AERES de 2010.
4. Une bonne visibilité du laboratoire au niveau national et international.

Points à améliorer et risques liés au contexte :

1. Manque de locaux adaptés et de support technique.
2. Pour certains, l'effort de publication dans les meilleurs supports doit être poursuivi.
3. Surveiller que les services d'enseignement ne deviennent pas trop lourds pour certains enseignants-chercheurs.
4. Consolider la vision scientifique du laboratoire pour assurer son développement pour le long terme.

Recommandations :

1. Renforcer et consolider le positionnement original du laboratoire et l'ancrage avec les SHS.
2. Poursuivre et amplifier les travaux intra- et inter-équipes.
3. Continuer l'effort de publications dans les meilleurs supports.
4. Préparer le départ à la retraite de membre(s) senior(s) et définir le plus tôt possible une politique claire concernant le maintien et/ou le renouvellement de thématiques.
5. La tutelle devrait être attentive à :
 - Apporter les moyens nécessaires (postes EC, BIATOSS) à la politique de recherche de l'unité pour que les enseignant-chercheurs ne fassent pas office d'ingénieur.
 - Améliorer les conditions de travail des membres du laboratoire notamment en terme de locaux.

3 • Appréciations détaillées

Appréciation sur la production et la qualité scientifiques

Selon le dernier rapport d'évaluation AERES (2010), une des principales difficultés concerne "l'écart qui subsiste entre la qualité globale du travail effectué qui est satisfaisante, et le manque de résultats de la valorisation des résultats, que ce soit en terme de publications dans des conférences et revues de références ou de réussite à l'ANR ou au PCRD."

Grâce à une nouvelle politique de publications, cette difficulté a été largement palliée par une augmentation significative des articles dans des revues et de conférences de qualité du domaine. Plus de 90 % des membres du laboratoire peuvent être considérés producteurs selon les critères AERES. On ne peut qu'encourager le laboratoire à continuer ses efforts dans cette direction afin de valoriser au maximum ses résultats.

Cette période est aussi marquée par l'obtention de plusieurs projets importants incluant notamment deux projets européens et un projet ANR.

Appréciation sur le rayonnement et l'attractivité académiques

L'ERIC joue un rôle important dans son environnement. L'ERIC est d'abord un laboratoire très visible sur le plan national voire international dans ses domaines de compétences (les entrepôts et la fouille de données). Ensuite, certains membres du laboratoire jouent un rôle clé dans le développement de ces domaines en France et à l'international. Le laboratoire maintient de nombreux échanges internationaux. Il joue un rôle très actif dans les principales structures nationales du domaine à travers ses responsabilités dans l'association EGC et la revue RNTI, ses actions d'organisation de conférences/workshops/journées et ses collaborations industrielles.

Enfin, malgré le nombre très réduit d'allocations ministérielles accordées par l'Ecole Doctorale, il est remarquable que le laboratoire arrive à recruter plusieurs doctorants par an sur des financements très variés (grâce à ses collaborations internationales pour des boursiers étrangers, des projets industriels/conventions CIFRE...).

Appréciation sur l'interaction avec l'environnement social, économique et culturel

L'interaction du laboratoire avec le monde économique est très bonne avec de nombreuses collaborations industrielles (contrats, conventions CIFRE...). Il faut noter également l'initiative importante de certains membres du laboratoire dans l'édition, le maintien et la diffusion de plusieurs logiciels libres dont TANAGRA, un logiciel de fouille de données reconnu, est un excellent exemple.

Appréciation sur l'organisation et la vie de l'unité :

Le laboratoire a entendu les recommandations du précédent comité appelant à abandonner l'organisation matricielle et à revenir à une structuration plus classique par équipes.

Le laboratoire apparaît être une unité cohérente et soudée. La proximité des thématiques du laboratoire et les travaux communs entre ses membres facilitent les communications internes et favorisent la production scientifique. L'organisation du laboratoire en deux équipes semble appropriée. Il faut veiller à préserver cette unité et continuer des travaux entre les deux équipes.

Appréciation sur l'implication dans la formation par la recherche

Le laboratoire est très fortement impliqué dans la spécialité "Fouille de données et gestion des connaissances" du Master d'Informatique de Lyon 2: parcours "Extraction des connaissances à partir de données (ECD)" en partenariat avec Polytech'Nantes (une trentaine d'étudiants en M2) et parcours "Erasmus Mundus DMKM" en partenariat avec 5 universités européennes (aussi une trentaine d'étudiants

en M2). Un projet de Doctorat Erasmus Mundus a été également déposé en 2012. Ces formations constituent un excellent facteur pour la visibilité et l'attractivité du laboratoire et un riche vivier pour le recrutement de doctorants.

Appréciation sur la stratégie et le projet à cinq ans

Peu de laboratoires évoluant dans un environnement académique en LLASHS ont une production scientifique et une visibilité pouvant se comparer à ERIC. Il faut reconnaître à cette équipe son talent, et son dynamisme. ERIC est, en particulier, moteur dans l'animation de la communauté nationale sur ses thèmes de prédilection. La taille du laboratoire lui permet sans doute de fédérer la totalité des forces en présence ; c'est une chance que ERIC a su saisir au fil des ans.

Pour assurer son développement, le laboratoire devrait affirmer davantage sa vision scientifique de son avenir. Dans un environnement lyonnais où plusieurs laboratoires d'informatique de taille imposante sont présents, il est indispensable de promouvoir et défendre sa "couleur" scientifique originale tout en renforçant ses liens déjà existants avec les autres laboratoires locaux en informatique et LLASHS sur la base de la complémentarité thématique.

Enfin, cette vision doit également alimenter la politique de recrutement et de renouvellement des EC, avec l'appui et l'engagement des établissements de tutelle.

Remarques additionnelles sur la facture du rapport d'activité

Ces remarques sont formulées dans le but d'aider à la préparation du prochain rendez-vous d'évaluation.

Concernant la rédaction de l'annexe détaillant la production scientifique (Annexe 5), il eut été judicieux de fournir une liste par équipe - tout en soulignant les productions émanant d'un travail inter-équipes (quitte à dupliquer les entrées). La numérotation [aa-xyz] n'est pas utile et brouille la lecture des références ; on aurait pu préférer des citations de type apacite (LaTeX) rappelant les noms des auteurs et l'année de publication.

Il faut veiller à fournir des références complètes : les maisons d'éditions et les collections des ouvrages édités sont essentielles (fournir une URL cliquable dans le document d'accompagnement peut s'avérer utile) ; les références des publications en revue doivent être complétées en fournissant les numéros de volume et de pages.

Il eut été utile de décrire en une phrase le résultat qui est au cœur de chacune des 10 publications retenues parmi la liste des publications marquantes de l'équipe. Incidemment, cette liste présente certains écarts avec les références proposés au bilan (section 2.3 « Contributions majeures ») alors qu'on s'attend à un recoupement plus fort (peut-être s'agit-il seulement de pouvoir proposer une liste de plus de 10 publications majeures ?).

Par ailleurs, il eut été utile de connaître les personnes impliquées dans chacun des axes (sections 1.3 et 2.3 « Contributions majeures ») admettant naturellement des chevauchements ; ces thèmes ne semblent pas cloisonnés - puisqu'on peut s'attendre à ce que la fouille de graphes intersecte les travaux sur la fouille d'opinions, par exemple - mais servent à structurer et à cibler les efforts de l'équipe.

4 • Analyse équipe par équipe

Équipe 1 :	Equipe SID (Systèmes d'Information Décisionnels)
Nom du responsable :	Fadila Bentayeb (MCF HDR)

Effectifs

Effectifs de l'équipe	Nombre au 30/06/2012 ¹	Nombre au 01/01/2014 ²	2014-2018 Nombre de produisants du projet ³
N1 : Enseignants-chercheurs titulaires et assimilés	8	8	8
N2 : Chercheurs des EPST ou EPIC titulaires et assimilés	0	0	0
N3 : Autres personnels titulaires (n'ayant pas d'obligation de recherche)	0	0	0
N4 : Autres enseignants-chercheurs (PREM, ECC, etc.)	0	0	0
N5 : Autres chercheurs des EPST ou EPIC (DREM, Post-doctorants, visiteurs etc.)	0	0	0
N6 : Autres personnels contractuels (n'ayant pas d'obligation de recherche)	0	0	0
TOTAL N1 à N6	8	8	8

Effectifs de l'équipe	Nombre au 30/06/2012	Nombre au 01/01/2014
Doctorants	11	
Thèses soutenues	2	
Post-doctorants ayant passé au moins 12 mois dans l'unité	2	
Nombre d'HDR soutenues	2	
Personnes habilitées à diriger des recherches ou assimilées	4	4

• Appréciations détaillées

La thématique centrale de l'équipe SID (Systèmes d'Information Décisionnels) porte sur les entrepôts de données (ED) et l'analyse en ligne (OLAP), d'une façon générale, l'informatique décisionnelle. Avec l'avènement des données complexes (textes, images, son, vidéo, ...), l'analyse en ligne doit s'adapter à la nature spécifique de ces données tout en gardant l'esprit de l'OLAP. Ainsi les opérateurs OLAP classiques sont définis pour des données classiques et s'avèrent souvent inadaptés pour des données complexes. Ces dernières portent généralement plus de sémantique que les données classiques, sémantiques qu'il s'agit de prendre en compte dans la modélisation multidimensionnelle et dans leur analyse en ligne. Les activités de recherche de l'équipe poursuivent deux objectifs majeurs :

- La proposition de nouveaux modèles d'ED et d'analyse en ligne pour répondre aux nouveaux défis liés à l'entreposage de données complexes, à la prise en compte de l'utilisateur et à la sécurité des données et des résultats ;
- La conception et le développement d'architectures de SID dans différents domaines d'application comme la linguistique, l'histoire, la médecine, etc.

De par sa proximité avec l'équipe DMD, une des spécificités des travaux développés par l'équipe SID réside dans la combinaison de l'OLAP avec plusieurs méthodes de traitement de données comme la fouille de données, la statistique et la recherche d'information. Cette synergie doit permettre de franchir le fossé sémantique pouvant exister entre des données complexes riches en information et des analyses OLAP simples.

Pour conclure, l'activité de l'équipe SID essaie d'apporter des solutions relatives à : (1) l'intégration et la représentation des données complexes, (2) l'analyse en ligne (OLAP) de données complexes, (3) les entrepôts centrés utilisateur, et enfin (4) la sécurité et la qualité des données.

Appréciation sur la production et la qualité scientifiques

L'effort de l'équipe depuis 2010 a surtout porté sur les publications internationales, dont le volume a augmenté de 73 % (le nombre de revues internationales a, notamment, plus que doublé), tandis que le volume de publications nationales a baissé de 26 % du fait de cet effort. L'équipe SID a suivi les recommandations du comité d'évaluation AERES de 2010, en privilégiant dans le choix des revues et des conférences internationales la qualité et non la quantité, même si comme elle le reconnaît elle-même, des efforts restent à faire pour publier dans les revues et conférences de rang A.

Tout d'abord tous les permanents de l'équipe sont publiants. Ensuite les publications tant au niveau des revues que des conférences, impliquent souvent d'autres auteurs appartenant à d'autres équipes en pointe de la communauté française des ED (Toulouse-IRIT, LI-Tours, LIRMM-Montpellier), ainsi de nombreux auteurs d'universités étrangères. Enfin il y a plusieurs publications de qualité impliquant des membres des deux équipes SID et DMD, montrant ainsi un certain décloisonnement entre les deux équipes.

Notons que parmi les publications sur cette période 2009-2012, il y a eu la soutenance de 2 HDR de chercheurs de l'équipe, celle de Fadila Bentayeb (2011) et celle de Sabine Loudcher (2011). Ces HDR sont de grande qualité, et souvent référencées dans la communauté francophone des ED.

Enfin, l'équipe SID produit aussi des logiciels libres téléchargeable autour des entrepôts de données et l'analyse en ligne.

Appréciation sur le rayonnement et l'attractivité académiques

Comme déjà évoqué, l'équipe SID est bien intégrée dans la communauté des ED, de par de nombreuses collaborations cristallisées dans des publications impliquant des chercheurs de diverses équipes, ceci tant à un niveau national qu'international. Son rayonnement est ainsi effectif aux deux niveaux. Notons que l'équipe SID est fondatrice de la conférence francophone reconnue EDA (Entrepôt de Données et Analyse en ligne), dans laquelle elle est toujours fortement impliquée. Elle est par ailleurs impliquée dans l'organisation de sessions de conférences ou workshop internationaux dans le domaine des entrepôts de données. Les membres de l'équipe SID sont aussi présents dans les comités de programmes de nombreuses conférences internationales ou nationales reconnues dans le domaine des entrepôts, notamment DOLAP.

En ce qui concerne les doctorants, on ne peut que constater une forte attractivité de l'équipe, au vu des nombreuses thèses financées par des bourses de pays étrangers, tant du Magreb, que d'Europe (Ukraine), et d'Asie (Thaïlande).

Appréciation sur l'interaction avec l'environnement social, économique et culturel

Principalement de par son attache assumée aux Sciences Humaines et Sociales, l'équipe est très bien intégrée dans l'environnement social, économique et culturel. Signalons aussi son interaction avec le monde industriel au travers de son implication dans notamment des conventions CIFRE.

Appréciation sur l'organisation et la vie de l'équipe

L'équipe apparaît très dynamique, soudée et il semble y avoir une très bonne ambiance de travail en son sein, comme le confirme d'ailleurs les doctorants de l'équipe rencontrés lors de cette visite. Notons que l'équipe se réunit chaque vendredi après-midi, avec des réunions d'une durée de deux à trois heures. Ce qui est remarquable.

Appréciation sur l'implication dans la formation par la recherche

C'est sans nul doute un des points forts du laboratoire et aussi de l'équipe. L'équipe prend en effet sa part dans l'opérationnalisation des formations de niveau master : la spécialité "Fouille de données et gestion des connaissances" du Master d'Informatique de Lyon 2: parcours "Extraction des connaissances à partir de données (ECD)" avec Polytech'Nantes, et le parcours "Erasmus Mundus DMKM" en partenariat avec 5 universités européennes (aussi une trentaine d'étudiants en M2). Notons qu'un projet de Doctorat Erasmus Mundus a été également déposé en 2012.

Appréciation sur la stratégie et le projet à cinq ans

Le projet à cinq ans (2013-2017) de l'équipe SID porte bien évidemment sur l'informatique décisionnelle. Deux axes sont privilégiés :

Le premier axe privilégié concerne les entrepôts et l'analyse en ligne et s'intéresse plus particulièrement aux données textuelles et aux réseaux sociaux. Plus globalement, Il s'agit pour l'équipe de concevoir des entrepôts d'informations et/ou de connaissance avec une analyse en ligne adaptée aux informations et/ou aux connaissances, avec des opérateurs OLAP spécifiques. L'équipe s'intéresse ainsi à l'intégration d'une ontologie dans le processus même d'entreposage. Signalons d'autres ouvertures comme l'élaboration d'entrepôts Web et la combinaison de technologies Web sémantique et OLAP, ou encore les entrepôts de données issues de simulation. Cet axe concerne aussi l'analyse en ligne collaborative avec plus particulièrement l'enrichissement des modèles d'entrepôts de données et des cubes OLAP par des connaissances (profils, annotations...).

Le second axe privilégié s'inscrit dans le « Cloud Computing », c'est l'informatique décisionnelle dans les nuages. Cette dernière doit permettre notamment de diffuser plus largement l'usage de la "Business Intelligence - BI", permettant à plus d'utilisateurs, de collaborateurs, de partager dans la limite de leur rôle et fonctions, des informations, et des résultats d'analyses. L'équipe s'intéresse ainsi à la BI à la demande ou BI collaborative devant fournir des outils d'analyse en ligne pour tous, de permettre aux citoyens de s'approprier les outils décisionnels. Associés à ces objectifs l'équipe souhaite contribuer sur : le définition de structure et de langages pour la modélisation multidimensionnelle dans les nuages, la personnalisation des analyses selon des profils d'utilisateurs, la qualité et la sécurité des données, l'utilisation d'agents intelligents pour l'intégration, l'analyse en ligne et la sécurité des données dans un environnement collaboratif dans le nuage.

Ces deux axes apparaissent d'actualité et très ambitieux. Cependant si l'équipe semble posséder toutes les compétences pour développer, dans les trois années prochaines, des contributions significatives sur chacun de ces axes, il n'est pas certain qu'elle dispose de suffisamment de forces vives. Aussi faudrait-il peut être réduire un peu l'ambition du projet.

Conclusion :

● Points forts et possibilités liées au contexte :

L'équipe possède une parfaite maîtrise de l'objet d'étude, à savoir les entrepôts de données, tant en ce qui concerne leur modélisation que leur exploitation, notamment leur analyse en ligne (OLAP). Le projet que l'équipe s'est définie est pertinent et ambitieux. Ses collaborations tant nationales qu'internationales sont un atout dans la réalisation de ce projet. Son enracinement dans le domaine des SHS permettra aussi d'appréhender certains aspects de ce projet, notamment l'étude des réseaux sociaux, mais aussi les aspects collaboratifs.

Enfin, rappelons l'excellente ambiance de travail régnant dans cette équipe, tant au niveau des permanents que des doctorants, équipe qui arrive à se réunir une fois par semaine.

Un éventuel rapprochement avec d'autres équipes de petites tailles dans la région lyonnaise et privilégiant déjà des thématiques de recherche voisines pourrait être envisagée, on pense notamment à l'équipe MODEME de l'IAE de Lyon.

● Points à améliorer et risques liés au contexte :

Le projet est à l'égal de l'excellence des travaux de l'équipe, même s'il peut paraître en première lecture trop ambitieux au vu des forces disponibles

Les risques principaux apparaissent être tout d'abord liés à la disponibilité des chercheurs, essentiellement des enseignants-chercheurs, fortement impliqués dans des programmes de formation originaux et attractifs. Le risque est ainsi que les chercheurs n'aient pas assez de temps à consacrer à leurs recherches, étant fortement impliqués dans des charges d'enseignements et d'administration.

Un autre risque est l'implication dans des projets contractualisés, eux aussi gourmands en temps, et parfois éloignés des axes de recherche privilégiés par l'équipe, mais cependant nécessaires au fonctionnement de l'équipe sur le plan financier.

● Recommandations :

Pour son projet ambitieux, fortement associé au SHS, l'équipe SID a besoin de moyens tant matériels (locaux et matériel notamment), qu'humain. Sur le côté matériel, l'équipe nous apparaît à l'étroit dans ses locaux actuels, locaux qui nécessiteraient d'ailleurs des travaux (notamment pour l'accessibilité à des toilettes).

Sur le volet humain, la présence d'un ingénieur d'étude ou de recherche, pouvant être mutualisé avec l'équipe DMD, nous apparaît indispensable pour supporter les développements informatiques associés au développement du projet.

En ce qui concerne les postes d'enseignant-chercheur, les institutions de tutelles pourraient prioriser un poste de Professeur et/ou de Maître de Conférences en 27^{ième} section (informatique) qui seraient nécessaires au développement de l'équipe, ces enseignants-chercheurs recrutés pouvant apporter à l'équipe plus de compétences notamment dans des domaines comme les ontologies, Web sémantique, et les réseaux sociaux, compétences impliquées dans le projet de l'équipe. Il serait également souhaitable que l'équipe puisse pouvoir attirer des chercheurs à temps plein.

L'intégration dans le milieu SHS doit encore être renforcée, elle apporte à l'équipe d'innombrables domaines d'application, ceci en interdisciplinarité.

4 • Analyse équipe par équipe

Équipe 2 :	DMD (Data Mining et Décision)
Nom du responsable :	Julien Velcin (MCF)

Effectifs

Effectifs de l'équipe	Nombre au 30/06/2012 ¹	Nombre au 01/01/2014 ²	2014-2018 Nombre de produisants du projet ³
N1 : Enseignants-chercheurs titulaires et assimilés	14	14	12
N2 : Chercheurs des EPST ou EPIC titulaires et assimilés			
N3 : Autres personnels titulaires (n'ayant pas d'obligation de recherche)			
N4 : Autres enseignants-chercheurs (PREM, ECC, etc.)			
N5 : Autres chercheurs des EPST ou EPIC (DREM, Post-doctorants, visiteurs etc.)			
N6 : Autres personnels contractuels (n'ayant pas d'obligation de recherche)			
TOTAL N1 à N6	14	14	12

Effectifs de l'équipe	Nombre au 30/06/2012	Nombre au 01/01/2014
Doctorants	11	
Thèses soutenues	8	
Post-doctorants ayant passé au moins 12 mois dans l'unité	1	
Nombre d'HDR soutenues	0	
Personnes habilitées à diriger des recherches ou assimilées	7	5-

• Appréciations détaillées

L'appréciation de l'équipe DMD se base sur le contenu du rapport, sur la présentation du bilan et du projet (en matinée), complétée par les entretiens de l'après-midi. Les commentaires formulés ici apportent notamment des éléments de réponses au dernier rapport d'évaluation (2010).

Appréciation sur la production et la qualité scientifiques

L'équipe DMD développe ses travaux dans le domaine de la fouille de données et de l'apprentissage, et depuis plusieurs années. Les résultats obtenus se situent autant sur le volet théorique (par exemple : apprentissage et forêts aléatoires menant à une publication PKDD 2010 ; propriétés d'antimonotonie pour l'extraction de règles d'association, publié chez Wiley en 2012) que sur le volet plus « appliqué » (évaluation des modèles de « topics », publié à IJCAI en 2011) engageant la transversalité des approches de fouille et d'IA. La thématique d' « Aide à la décision », prise dans un sens large (Business Intelligence), est une cible définissant la direction dans laquelle les efforts sont fournis - et cela concerne tout autant DMD que l'équipe SID de ERIC qui en fait une priorité forte de son projet (« Devenir leader dans le domaine des SID en France »).

La présentation des travaux se déploie sur trois axes : « Modélisation, caractérisation, fouille dans les graphes » ; « Modèles pour l'aide à la décision » ; « Analyse de données complexes et fouille d'opinions ». Les thèmes de recherche de DMD sont donc bien en phase à la fois avec les défis qui se posent à cette communauté de recherche, et avec les besoins engendrés par la production croissante de données numériques (le « Big data »). L'équipe DMD est naturellement bien placée pour apporter des réponses aux questions et aux défis qui se posent aujourd'hui.

Par ailleurs, l'équipe s'est résolument attaquée aux exigences de publication à l'international qui avaient été formulées lors de la dernière campagne d'évaluation. On trouve de nombreuses publications de qualité parues dans des revues dont la visibilité est acquise et qui assurent aux travaux de l'équipe une portée scientifique accrue (on note, parmi les publications en revue internationale : Journal of Applied Statistics - Taylor & Francis ; Expert Systems with Applications - Elsevier ; Web Intelligence and Agent Systems - IOS Press ; Computational Intelligence - Wiley...). L'activité d'édition d'ouvrages est à l'égal du travail d'animation qu'assurent les membres de DMD depuis des années. Ce changement de stratégie de publication s'est fait tout en maintenant une présence dans les forums les plus sélectifs de la discipline, tout comme dans les forums nationaux.

Appréciation sur le rayonnement et l'attractivité académiques

Nous commenterons ici les aspects liés à la contractualisation de la recherche, tant ils peuvent être vu comme un prisme au travers duquel on peut lire le rayonnement scientifique de l'équipe. C'est par sa visibilité et son rayonnement qu'une équipe suscite des opportunités de contrats. DMD a su intéresser des partenaires nationaux et s'engager dans le projet ANR d'envergure ImagiWeb. Elle est aussi partie prenante dans un appel récent autour de la thématique « Big Data » (projet Investissement d'avenir). Son implication dans deux projets européens, lié aux domaines de la santé plus proche de l'équipe de Lyon 1 venue récemment grandir ses rangs, apporte une solide réponse aux critiques du dernier rapport d'évaluation. Il y a certainement une « culture de projets » à développer dans l'équipe, et dans tout ERIC, pour pérenniser cette façon de financer ses activités.

Appréciation sur l'interaction avec l'environnement social, économique et culturel

De nombreuses thématiques de recherche de DMD, qu'elles soient décrites dans le bilan ou annoncé au projet, placent l'équipe dans le contexte de l'étude des systèmes complexes. Sur ce point, le rapport dit peu de choses sur le positionnement de DMD par rapport à son environnement immédiat (on pense à l'institut des systèmes complexes lyonnais IXXI). Cela dit, le projet inscrit résolument DMD dans une démarche la rapprochant, au niveau national, de l'institut des sciences humaines et sociales InSHS du CNRS. Il faut y voir la volonté de DMD de s'accorder avec son environnement proche - c'est en partie ce contexte universitaire SHS qui lui confère sa spécificité.

Appréciation sur l'organisation et la vie de l'équipe

L'équipe adopte un rythme de réunion favorisant sa cohésion. Des séminaires thématiques lui permettent d'accueillir sur son site des collaborateurs du PRES lyonnais. L'unité thématique de DMD est palpable - on peut lire le

résultat de la réorganisation d'ERIC en deux équipes, sur un schéma plus traditionnel comme on le lui avait proposé lors de la dernière évaluation.

Appréciation sur l'implication dans la formation par la recherche

L'implication de l'équipe dans le montage et l'animation du Master d'Informatique - parcours "Extraction des connaissances à partir de données (ECD)" (Lyon 2) témoigne de son dynamisme dans la formation par la recherche - cette remarque vaut pour tout ERIC. Par ailleurs, DMD (et ERIC tout entier) a développé un réseau de collaborateurs qui lui permet de recruter ses doctorants dans le vivier international - c'est là un atout pour une équipe évoluant en contexte universitaire SHS.

Appréciation sur la stratégie et le projet à cinq ans

Le projet scientifique de DMD met en exergue un nombre de questions et de verrous à lever qui entre dans son champ thématique : données massives et scalabilité des algorithmes ; traçabilité des données, fouille d'opinions ; approche par apprentissage topologique intégrant données multi-attribuées et méta-données ; etc. Ce sont là de réelles opportunités pour DMD ; l'équipe possède les compétences pour attaquer ces questions, et se distinguer en apportant des réponses originales.

Le format de l'équipe doit lui permettre d'apporter une réponse « intégrée » à ces questions qui émanent pour beaucoup de besoins sociétaux. L'exposé nous fait moins sentir le besoin d'avancées théoriques que la nécessité de produire un effort transversal (informatique / statistiques / mathématiques appliquées) pour apporter des solutions aux problèmes évoqués. On pourrait comparer DMD à une EPI (INRIA), par sa taille et par les perspectives annoncées au projet, pour lui suggérer d'articuler le projet autour de jalons (milestones).

La vision qui se dégage du projet est en phase avec la stratégie de contractualisation qui est en marche dans DMD. L'assise scientifique de DMD, sa notoriété et les compétences in-house, sont de réelles garanties de succès. La projection à 5 ans doit servir de moteur pour les enrichir.

Conclusion :

- Points forts et possibilités liées au contexte :

L'équipe DMD (et le laboratoire ERIC tout entier) se trouve dans un environnement de proximité avec les SHS. Ses thématiques de recherche en font un acteur clé des défis posés par la numérisation grandissante des données de toute sorte, alors que tout un pan des SHS se « digitalise ». Il faut y voir une occasion unique pour DMD de s'affirmer dans son environnement proche, comme sur la place de Lyon et au niveau national.

L'équipe DMD est, dans une certaine mesure, une équipe « historique » de la fouille de données. Sa notoriété nationale, depuis longtemps acquise, doit rayonner au-delà. L'effort de publication observé aujourd'hui va dans ce sens et doit être maintenu. Aussi, il faut veiller à associer pleinement les doctorants à cette politique de publication à l'international.

- Points à améliorer et risques liés au contexte :

La place de la problématique santé dans DMD, et dans son articulation avec l'environnement lyonnais doit être réfléchi. Cette double portée Santé / SHS ne doit pas forcer une si petite unité à faire le grand écart.

L'implication dans des actions de recherche contractualisée doit être poursuivie. Les compétences et le dynamisme de DMD sont des atouts pour réussir dans cette voie. L'établissement doit en prendre la pleine mesure, puisque ce mode de fonctionnement exige un soutien administratif et logistique certain - notamment lorsqu'il s'agit de s'investir au niveau européen.

La place du logiciel TANAGRA, son développement, son évolution, doit entrer dans la stratégie du laboratoire dans la mesure où il lui apporte visibilité et notoriété. Les ressources nécessaires au maintien et au développement d'un tel outil sont à l'égal de sa qualité et de son impact scientifique (sur les communautés académiques, mais aussi sur l'industrie).

Nous avons souligné à plus d'une fois combien DMD rassemble un patrimoine scientifique unique dans son domaine d'activités. Il faut veiller à le préserver - pour assurer le succès futur de l'équipe, la bonne implémentation

de la stratégie et l'atteinte des objectifs, notamment au travers des recrutements (renouvellement ou nouveau recrutement).

- **Recommandations :**

La taille de chacune des équipes de ERIC, voire ERIC tout entier, peut être comparée à une équipe d'un plus grand laboratoire ou à une EPI INRIA. De plus, sur le plan thématique, ERIC présente à la fois un spectre de compétences et un niveau d'intégration que l'on peut envier. De ce point de vue, ERIC pourrait développer une vision plus proactive de ses forces tant auprès des collègues SHS de Lyon (et ailleurs) que vers le monde industriel en terme de transfert de compétences et de technologie. Il peut être un acteur incontournable pour affirmer sur Lyon la place de son établissement dans l'univers des SHS.

Ayant en vue la prochaine campagne d'évaluation, la formulation du projet à 5 ans saura s'articuler et mieux intégrer les aspects stratégiques que nous soulignons ici. Les objectifs de recherche sont clairement identifiés et sont en phase avec les courants de recherche nationaux et internationaux. Il faut mûrir la dimension de politique scientifique du projet, en termes de positionnement et de développement futur.