

## Séminaire EUROLIO 2010 « Les indicateurs d'innovation localisés »

### **L'évolution des pratiques d'observation avec les pôles de compétitivité : l'apport de l'outil Servist pour mesurer le potentiel scientifique**

Valérie Bonvallot [valerie.bonvallot@inist.fr](mailto:valerie.bonvallot@inist.fr)

INIST-CNRS 2, Allée du Parc de Brabois CS 10310 F-54519 Vandoeuvre-lès-Nancy

Sylvie Chalaye [sylvie.chalaye@eurolio.eu](mailto:sylvie.chalaye@eurolio.eu)

Eurolio - Bâtiment des Hautes Technologies, 20 rue Benoit Lauras 42000 Saint-Etienne

Philippe Houdry [philippe.houdry@inist.fr](mailto:philippe.houdry@inist.fr)

INIST-CNRS 2, Allée du Parc de Brabois CS 10310 F-54519 Vandoeuvre-lès-Nancy

#### **Résumé**

La politique des pôles de compétitivité fait l'objet de financements élevés de la part de l'Etat et des collectivités territoriales. Le besoin croissant d'indicateurs à l'échelle de ces entités fait évoluer les pratiques d'observation territoriale et amène les observateurs à utiliser de nouveaux outils reliés à la veille, comme par exemple Servist. Ce Système d'Exploration, de Recherche et de Visualisation de l'Information Scientifique et Technique, développé par le service Veille de l'INIST-CNRS, propose le calcul et la mise à disposition d'indicateurs liés à la production scientifique, qu'elle soit liée à une thématique, un organisme ou un territoire.

En présentant dans quelle mesure l'observation des pôles de compétitivité fait évoluer les exercices d'observation territoriale, nous montrons comment le potentiel scientifique des pôles peut être mis en évidence grâce aux principales fonctionnalités de Servist. Des scénarios sont décrits à partir des données du pôle de compétitivité Xylofutur étudié dans le cadre d'une étude réalisée par EuroLIO pour la DATAR concernant la conception d'un tableau de bord territorialisé de l'innovation et le suivi des clusters.

**INTRODUCTION..... 3**

**1. L'OBSERVATION DES POLES DE COMPETITIVITE : LES CHANGEMENTS DANS LES ACTIVITES D'OBSERVATION TERRITORIALE..... 3**

**1.1. LES POLES : UN OBJET D'OBSERVATION COMPLEXE LIE A SON CONTENU MULTIDIMENSIONNEL ..... 4**

**1.2. LES POLES : UN NOUVEAU PERIMETRE D'ANALYSE FAISANT EVOLUER LES PRATIQUES D'OBSERVATION TERRITORIALE..... 7**

**2. LA VEILLE AU SERVICE DE L'OBSERVATION DES POLES DE COMPETITIVITE : EXEMPLE DE MESURE DE LEUR POTENTIEL SCIENTIFIQUE A L'AIDE DE SERVIST 9**

**2.1. COLLECTE DES DONNEES ..... 9**

2.1.1. SOURCES D'INFORMATION..... 9

2.1.2. REQUETE POUR CONSTITUER LE CORPUS ..... 9

2.1.3. DEDOUBLONNAGE DES NOTICES ..... 10

**2.2. PRETRAITEMENTS ..... 10**

2.2.1. REFORMATAGE EN XML..... 10

2.2.2. TRAITEMENTS SUR LA DESCRIPTION DOCUMENTAIRE (CHAMPS) ..... 10

2.2.3. TRAITEMENTS SUR LES DONNEES (CONTENUS)..... 11

2.2.4. TRANSCODAGE (JEUX DE CARACTERES) ..... 12

**2.3. FONCTIONNALITES DE SERVIST..... 13**

2.3.1. EXPLORATION DU CORPUS DE NOTICES ..... 13

2.3.2. CALCULS BIBLIOMETRIQUES..... 13

2.3.3. VISUALISATION DES RESULTATS..... 13

**2.4. SCENARIOS : ILLUSTRATION AVEC LE POLE XYLOFUTUR ..... 13**

2.4.1. COMPTAGES ..... 14

2.4.2. ANALYSE STATISTIQUE PAR TRIS CROISES..... 16

2.4.3. VISUALISATION GEOGRAPHIQUE..... 20

2.4.4. VISUALISATION DE GRAPHS RESEAUX A L'AIDE DE MATHEO ANALYZER ..... 21

**CONCLUSION..... 22**

**BIBLIOGRAPHIE ..... 22**

## Introduction

Avec la mise en place de la politique des pôles de compétitivité, de nouveaux besoins d'indicateurs apparaissent à cette échelle. Habituellement, dans les observatoires territoriaux, les indicateurs disponibles suivent, dans la grande majorité des cas, des limites géographiques prédéfinies (de la commune à la région). Ils sont donc construits selon une logique géographique et non organisationnelle. Cette dernière consisterait à prendre en compte les acteurs très en amont dans la démarche de production d'indicateurs.

Désormais, la demande émanant des acteurs publics concerne des indicateurs qui ont pour contour stricto sensu les pôles de compétitivité (ou autres clusters labellisés type SPL<sup>1</sup> ou clusters régionaux) au-delà d'un périmètre géographique prédéfini. Cette nouveauté entraîne une forte évolution des pratiques d'observation territoriale et le développement des démarches de veille.

C'est dans ce contexte qu'une étude a été réalisée par EuroLIO<sup>2</sup> pour la DATAR<sup>3</sup> en 2009 (DATAR/EuroLIO, 2009). L'objectif était de concevoir une grille d'indicateurs pertinents pour analyser les caractéristiques et l'impact des pôles de compétitivité. Il s'agissait également d'en étudier la faisabilité à partir des sources de données disponibles et à l'aide de 3 clusters tests (Pôle de compétitivité Xylofutur, SPL Mécaloire, SPL Paris Mix).

L'objet de cet article est de présenter en quoi les pôles de compétitivité sont difficiles à observer et en quoi ils font évoluer les pratiques d'observation territoriale. Ces évolutions sont illustrées avec le cas de Xylofutur<sup>4</sup> en mesurant plus précisément son potentiel scientifique. Pour cela, le logiciel Servist<sup>5</sup>, développé par l'INIST-CNRS<sup>6</sup>, a été utilisé. Ses fonctionnalités et les résultats produits sont présentés dans une logique de veille scientifique pour le pôle.

## 1. L'observation des pôles de compétitivité : les changements dans les activités d'observation territoriale

Les pôles de compétitivité sont fortement financés par l'Etat et les collectivités territoriales. Ces dernières ont besoin d'indicateurs à la fois de caractérisation et d'impact à l'échelle de ces pôles difficiles à observer.

---

<sup>1</sup> SPL : Système productif local

<sup>2</sup> EuroLIO : EuroLIO est l'Observatoire Européen des Données Localisées de l'Innovation. Il s'agit d'un réseau qui comprend 5 laboratoires français (CREUSET Université Jean Monnet de Saint Etienne, BETA Université de Strasbourg, GREThA Université de Bordeaux, IV INIST-CNRS, LEREPS Université Toulouse I.  
<http://www.eurolio.eu/>

<sup>3</sup> DATAR : Délégation interministérielle à l'aménagement du territoire et à l'attractivité régionale.

<sup>4</sup> Xylofutur : Pôle de compétitivité national appuyé par le Conseil régional d'Aquitaine, le pôle industries et pin maritime du futur a été labellisé le 12 juillet 2005. Les membres du pôle comporte 18 grandes entreprises, 55 entreprises, 14 centres de recherche, 9 centres de formation et 36 autres membres.

<sup>5</sup> Le Système d'exploration, de recherche et de visualisation de l'information scientifique et technique, est développé par le service Veille de l'INIST-CNRS. Optimisé pour Firefox, il propose le calcul et la mise à disposition d'indicateurs liés à la production scientifique, qu'elle soit liée à une thématique, à un organisme ou à un territoire. Ses fonctionnalités sont développées plus loin dans l'article.

<sup>6</sup> INIST-CNRS : Institut de l'information scientifique et technique, unité propre du Centre national de la recherche scientifique : <http://www.inist.fr/> ; <http://www.cnrs.fr>

En effet, ils représentent un objet complexe pour lequel une analyse multidimensionnelle est nécessaire pour bien comprendre le positionnement et les performances des clusters (§ 1.1). De nombreux indicateurs doivent ainsi être construits pour mieux cerner leur dynamique. Cependant, il n'est pas évident de construire des indicateurs qui concernent réellement le périmètre du pôle et non une échelle géographique ni une nomenclature sectorielle prédéfinies. En cela, l'évolution des pratiques d'observation territoriale est nécessaire (§ 1.2).

### **1.1. Les pôles : un objet d'observation complexe lié à son contenu multidimensionnel**

Pour comprendre les dynamiques des clusters et les différences de performance, il apparaît nécessaire d'adopter une approche multidimensionnelle (Chalaye et Massard, 2009) distinguant 6 grands thèmes (cf Figure 1).

1. **la structure générale** : on s'intéresse ici aux déterminants structurels du fonctionnement du cluster qu'ils soient d'ordre économique ou géographique ou qu'il s'agisse des structures de gouvernance et d'animation du cluster.

La structure économique renvoie aux organisations impliquées dans le pôle et à leurs statuts (entreprises, universités, centres techniques, de formation ...). Selon la composition du pôle, sa dimension sera plus ou moins industrielle, technologique et/ou scientifique.

Dans la population d'entreprises, le poids des grandes entreprises, des PME, des TPE et le statut des entreprises (indépendante ou filiale d'un groupe) est assez déterminant en termes d'innovation : chaque type d'entreprises met en œuvre des pratiques d'innovation (Lefebvre, 2008) et de coopérations différentes, en particulier concernant les relations avec la recherche publique (Audretsch et Vivarelli, 1994 ; Pacitto et Tordjman, 1999).

D'un point de vue territorial, les collectivités locales s'intéressent au poids économique que représente le cluster. Au-delà du nombre d'emplois, les travaux d'économie géographique montrent l'importance de la répartition des emplois par qualifications (cadres techniques, techniciens, ouvriers...). Elle permet de repérer des fonctions dominantes au sein du cluster : fonction de production, de conception, ...

Dans la structure générale du pôle, la structure géographique est une autre composante discriminante : de la concentration des activités au sein d'une grande métropole (notamment observée pour les services intensifs en savoir) à une localisation interrégionale plus diffuse, les cas de figure sont nombreux.

Enfin, la gouvernance et l'animation peuvent impacter la performance du cluster comme l'ont montré O. Sölvell et al. (2003) dans le cadre d'une enquête mondiale réalisée auprès de 233 clusters. Les qualités de l'animateur du cluster (étendue des contacts, niveau de connaissances du cluster...) et la formalisation d'un cadre explicite, clair, basé sur les forces du cluster et partagé entre toutes les parties font notamment partie des facteurs de succès. L'absence d'un cadre partagé, de consensus en termes d'objectifs et d'initiatives ressort comme la principale raison des échecs.

2. **la capacité de R&D** : les déterminants liés au potentiel de R&D (niveau de ressources affectées à la production de nouvelles connaissances).

Le montant des dépenses de R&D, dans les secteurs privé et public, constitue une variable clé pour caractériser les clusters. Nombre de travaux économétriques ont en effet montré la présence d'externalités de connaissances<sup>7</sup> issues à la fois de la recherche privée et universitaire sur le niveau local d'innovation.

Selon Lefebvre (2008), à partir du niveau faible ou fort des potentiels privé et public de recherche, quatre types de pôles peuvent être identifiés ; ils offrent chacun des potentialités très différentes en matière de coopération et de processus d'innovation.

Enfin, le financement de la R&D des entreprises (autofinancement, accès aux prêts bancaires, crédit impôt recherche) doit également être pris en compte.

### **3. la capacité d'innovation**

Au-delà des ressources humaines et financières affectées à la production de connaissances, la performance du cluster dépend de sa capacité à transformer ces ressources en innovations technologiques ou non technologiques.

Le rôle des innovations technologiques est aujourd'hui largement reconnu dans la compétitivité des entreprises, même si les effets de l'innovation sur la productivité des entreprises semblent dépendre assez fortement du type d'innovations mises en œuvre (Duguet, 1999). Les innovations se développent aussi de plus en plus dans le domaine non technologique, que ce soit au niveau organisationnel ou marketing. Les principaux types d'innovations organisationnelles et administratives sont : l'adoption de techniques avancées de gestion, la modification des structures organisationnelles ou l'adoption d'orientations stratégiques nouvelles (OCDE, 2005). La dimension non technologique des innovations est importante à prendre en compte, non seulement parce qu'elle constitue la voie principale d'innovation dans certains secteurs moins centrés sur la R&D, mais aussi parce que les changements organisationnels et le progrès technologique sont étroitement liés.

### **4. les coopérations ou le niveau de connectivité**

Pour Andersson et Karlsson (2004), la production régionale d'innovation dépend :

- ✓ De la vitesse à laquelle les nouvelles connaissances sont introduites dans la région en provenance d'autres régions ;
- ✓ De la facilité des échanges de connaissances à l'intérieur même de la région.

Ces deux dimensions seront d'autant plus importantes que la capacité des acteurs à développer des coopérations, locales ou plus globales, est forte.

Ainsi, le niveau de connectivité du cluster doit s'apprécier de deux façons :

- ✓ A l'extérieur de celui-ci pour bénéficier des connaissances produites au niveau global ;
- ✓ Au sein du cluster pour favoriser la diffusion des connaissances entre les différents acteurs impliqués.

---

<sup>7</sup> Les externalités désignent des phénomènes d'interactions entre individus ou firmes qui ne sont pas internalisés par le marché c'est-à-dire qui ne donnent pas lieu à une contrepartie financière. Le terme externalités de connaissances renvoie au fait qu'une connaissance produite par un individu se diffuse et peut être appropriée par des individus autres que celui qui en est à l'origine.

Les clusters peuvent également se distinguer selon la nature même des coopérations : s'effectuent-elles pour l'essentiel entre entreprises, entre laboratoires ou entre entreprises et laboratoires ?

L'historique des relations est une autre dimension à prendre en compte. En effet, la coopération est un processus complexe, qui ne s'invente pas. Ainsi, au-delà des mesures incitatives pour développer des coopérations, l'un des facteurs déterminants pour expliquer l'intensité des échanges est l'existence de liens de partenariats établis par le passé (Rallet et Torre, 2001).

Enfin, la structure relationnelle du réseau est fondamentale pour mieux comprendre le niveau d'innovation. Pour Owen-Smith et Powell (2004), l'impact sur l'innovation diffère selon le niveau de concentration géographique du réseau, la position des agents dans le réseau et les caractéristiques institutionnelles des agents. Dans un réseau géographiquement concentré, les liens formels et informels entre acteurs ont tendance à se superposer. Dans ce cas, la centralité de l'organisation dans le réseau n'a que peu d'effet sur l'innovation. La situation est en revanche différente dans un réseau géographiquement dispersé où des liens informels plus faibles rendent difficile le contrôle de la diffusion. Ceci se traduit par des liens contractuels resserrés avec une limitation de l'accessibilité aux informations. Dans ce cas, la centralité dans le réseau est importante pour obtenir les informations et aura un impact sur l'innovation. Par ailleurs, la dimension institutionnelle est importante. Plus le réseau est dominé par des organismes publics, moins la centralité est importante car la diffusion des connaissances est largement plus étendue lorsque des institutions basées sur la « science ouverte » (« open science ») sont très présentes.

## **5. le secteur et le marché**

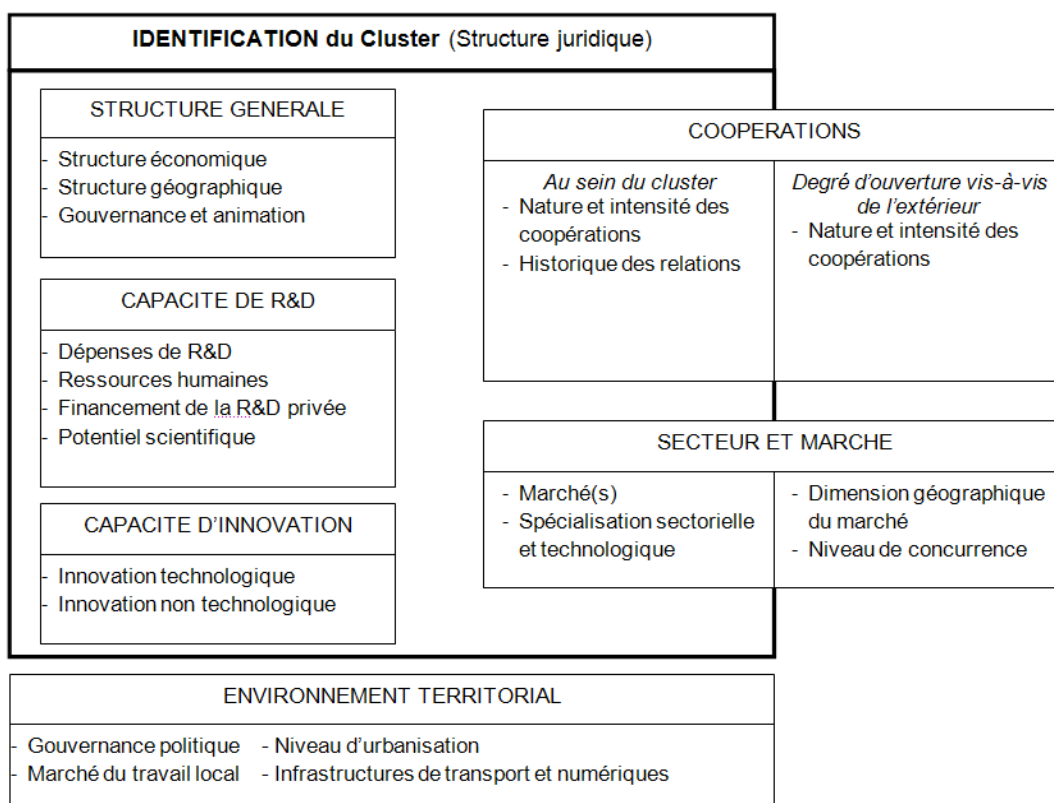
La performance du cluster dépend aussi des opportunités technologiques plus ou moins fortes selon son positionnement sectoriel. La spécialisation versus la diversité sectorielle et technologique peut impacter le niveau d'innovation. Néanmoins, ces effets ne font pas consensus dans la littérature (Autant-Bernard et al., 2008). Globalement, on peut avancer que, si la diversité des compétences disponibles sur un territoire a un impact plutôt favorable sur la production d'innovations du fait d'une plus grande possibilité de combinaisons de savoirs, il est important de développer des spécialisations pour pouvoir capter des connaissances de pointe produites à distance, les assimiler et les exploiter.

Les clusters peuvent se distinguer également par rapport au potentiel de marché auquel ils s'adressent. Les performances ne seront pas les mêmes selon que le cluster se positionne sur un ou plusieurs marchés au sein desquels il adapte et diffuse les technologies développées. Pour Hendry et Brown (2006), les stimuli et les trajectoires technologiques des marchés constituent un élément fortement explicatif des différences de performance des clusters.

## 6. L'environnement territorial

En plus des caractéristiques propres au cluster, nous proposons d'introduire des indicateurs caractérisant l'environnement territorial au sein duquel se trouve le cluster.

**Figure n° 1 : Tableau de bord Identification et Caractérisation du Cluster (ICC) : structuration générale**



### 1.2. Les pôles : un nouveau périmètre d'analyse faisant évoluer les pratiques d'observation territoriale

Les besoins d'indicateurs à l'échelle des pôles de compétitivité font évoluer les démarches d'observation territoriale.

Auparavant, ces dernières consistaient essentiellement à collecter des données agrégées à des échelles géographiques prédéfinies, couplées ou non à des échelles sectorielles (principalement définies par le code NAF des entreprises<sup>8</sup>). Sur les territoires, on observe ainsi, assez souvent, des tentatives de construction d'indicateurs à l'échelle des pôles en rapprochant les nomenclatures NAF avec les activités de ceux-ci. Mais ces analyses basées sur une **entrée géographique et sectorielle prédéfinies ne permettent pas réellement d'observer la dynamique des pôles de compétitivité** pour deux raisons :

<sup>8</sup> Code NAF : un des codes de l'INSEE consacré à la nomenclature des activités françaises.

- ✓ Le contour géographique des pôles ne suit pas les limites territoriales administratives ou les zonages d'études prédéfinies comme les zones d'emploi. L'étendue géographique des pôles peut être très variée et dépasse souvent les limites régionales ;
- ✓ Le contour sectoriel des pôles est souvent difficile à cerner car ces derniers peuvent regrouper des membres aux activités très différentes (industrielles ou tertiaires). C'est notamment vrai pour des pôles davantage constitués sur une logique de marché que sur une logique sectorielle.

De nouvelles démarches d'observation doivent alors être mises en œuvre pour s'affranchir des périmètres géographiques et sectoriels prédéfinis. La construction d'indicateurs à l'échelle des pôles de compétitivité stricto sensu impose de disposer au départ de la liste des membres (entreprises, laboratoires, autres) et de collecter ensuite des données individuelles, d'où le développement de démarches de veille territoriale.

Méthodologiquement, plusieurs voies sont envisageables :

1. Les démarches de veille internet permettent de collecter des informations qualitatives sur les membres.
2. La collecte directe d'informations auprès des entreprises ou auprès des animateurs des pôles peut être réalisée via des enquêtes. Le problème de cette procédure est la grande sollicitation des entreprises et animateurs entre les initiatives nationales et locales de demande d'information.
3. La dernière voie envisageable et privilégiée par EuroLIO est d'utiliser des bases de données produites au niveau national ou européen, de manière systématique par des organismes de traitement d'information qu'ils soient publics (Insee, Ministères, OCDE, etc.) ou privés (Bureau van Dijk...). Les avantages de cette procédure sont les suivants :
  - ✓ Elle sollicite très peu les animateurs des clusters (seule la liste des membres leur est demandée). Les membres, entreprises ou laboratoires, ne sont pas sollicités ;
  - ✓ Elle permet une plus grande objectivité des indicateurs car la source d'information est la même quel que soit le pôle. Elle facilite ainsi les comparaisons entre pôles ;
  - ✓ Elle permet une mise à jour plus facile dans le temps car les bases sont régulièrement alimentées.

Les bases de données utilisées dans cette démarche peuvent être soit numériques (qui peuvent être traitées « classiquement » dans des tableurs ou logiciels de statistiques), soit textuelles. Le traitement des données textuelles est plus difficile mais il offre une très grande richesse d'information. Dans ce cas, des outils spécifiques peuvent être utilisés pour faciliter leur traitement, comme l'outil Servist développé par l'INIST-CNRS.



## **2. La veille au service de l'observation des pôles de compétitivité : exemple de mesure de leur potentiel scientifique à l'aide de Servist**

L'objectif de cette partie est de montrer en quoi une démarche de veille, via l'utilisation d'un outil comme Servist, peut fournir des indicateurs concernant les clusters.

Le potentiel scientifique, dont l'importance est surtout liée à la présence de laboratoires de recherche, est une dimension à analyser au sein des pôles de compétitivité. Généralement mesuré à partir des publications scientifiques, il est source d'innovation notamment dans le cadre de partenariats science-industrie.

Tout le processus, de la collecte de données à la production des indicateurs, va être présenté. La démarche méthodologique sera illustrée avec le pôle Xylofutur sur la période 1995-2009.

### **2.1. Collecte des données**

#### **2.1.1. Sources d'information**

La production d'indicateurs est liée au corpus initial. Il est constitué à partir d'une collecte des données. Cette étape, essentielle, doit être réalisée minutieusement.

Les sources d'informations retenues sont les bases de données bibliographiques. Elles ont pour avantage d'être régulièrement mises à jour et permettent ainsi de développer une démarche de veille scientifique. Nombreuses et aux contenus scientifiques et structurels différents, les sources à interroger ne sont pas clairement définies pour calculer les indicateurs bibliométriques.

La première étape de sélection est naturellement liée aux domaines traités. Il faut ensuite prendre en compte la structure de la notice bibliographique (voir exemple §2.2.3) en fonction du type d'indicateurs à calculer :

- ✓ Présence d'un champ interrogeable dédié aux affiliations car seul le nom des organismes des clusters est fourni ;
- ✓ Présence de l'affiliation de tous les auteurs pour mettre en évidence les coopérations scientifiques. L'affiliation du premier auteur est insuffisante ;
- ✓ Présence d'une classification pour identifier les spécialisations des clusters.

Afin d'éviter de multiplier les phases de prétraitement (sur lesquelles nous reviendrons), et notamment celles de reformatage et d'homogénéisation des classifications, sont retenues des bases multidisciplinaires dont le service Veille connaît parfaitement le fond et la forme : PASCAL<sup>9</sup> et le Science Citation Index<sup>10</sup>.

#### **2.1.2. Requête pour constituer le corpus**

La recherche des notices est faite à partir du nom des organismes appartenant au cluster. Comme il n'existe ni identifiant unique (type SIRET) ni de règles d'écriture, la complexité de cette étape est liée aux points suivants :

---

<sup>9</sup> Base de données bibliographiques, multidisciplinaire et multilingue qui couvre l'essentiel de la littérature mondiale en sciences, technologie et médecine, produite par l'INIST-CNRS.

<sup>10</sup> Base de données bibliographiques, multidisciplinaire et multilingue qui couvre l'essentiel de la littérature mondiale en sciences, technologie et médecine, produite par ISI-Thomson.

- ✓ Le nom d'un organisme se présente sous différentes formes d'écriture à un temps T mais également sur la période étudiée<sup>11</sup> (fusion, renumérotation, intégration, éclatement d'un laboratoire par exemple). Toutes ces formes doivent être recensées ;
- ✓ Les auteurs déclarent le siège social de l'entreprise ou l'établissement secondaire. Pour les entreprises multi-établissements, la publication peut donc être adressée au siège alors que la recherche s'est faite dans un établissement secondaire.
- ✓ L'adresse précise n'est pas toujours indiquée, ce qui est une gêne pour les entreprises multi-établissements<sup>12</sup>. Dans ce cas, une recherche sur internet avec le nom des auteurs et l'entreprise est réalisée pour localiser la ville des auteurs de manière à décider si la notice est gardée ou pas. Lorsque la vérification est impossible, la notice n'est pas conservée.

En plus de ces difficultés, une limite de la démarche apparaît. Comme seul le nom des organismes du cluster est connu, la recherche bibliographique ne s'effectue pas à partir des noms des personnes. Ainsi, des publications peuvent être comptabilisées alors que leurs auteurs, appartenant à l'organisme, ne font pas directement partis du cluster. Il ne s'agit donc pas de la production scientifique du pôle stricto sensu mais de celle des membres le composant, d'où le terme de potentiel scientifique.

### **2.1.3. Dédoublonnage des notices**

Une fois chaque source interrogée, les notices sont regroupées en un corpus puis dédoublonnées semi-automatiquement, dans Excel, sur les champs "Source"<sup>13</sup> et "Titre"<sup>14</sup> à l'aide d'une macro.

## **2.2. Prétraitements**

### **2.2.1. Reformatage en XML**

Le format informatique des données est converti en XML car les prétraitements des données ainsi que Servist s'appuient sur des outils d'ingénierie XML. Effectivement, les données collectées sont généralement dans un autre formalisme structurel. Cela nécessite quelquefois l'écriture de programmes à façon (en partant de formats type serveur, de formats tabulés ou autres).

### **2.2.2. Traitements sur la description documentaire (champs)**

La connaissance de la description documentaire des données à traiter est importante pour connaître le découpage en champs, la sémantique et les contraintes associées. Cela peut passer par une liste des champs avec leurs intitulés et tous les séparateurs structurant l'information. L'utilisateur précise les champs sur lesquels il souhaite faire des analyses pour identifier les index à générer dans Servist.

Dans le cadre de l'étude pour le pôle Xylofutur, en plus des index habituels<sup>15</sup>, les suivants ont été créés pour répondre au besoin d'indicateurs particuliers :

<sup>11</sup> Exemple : pour le "Laboratoire d'ingénierie des systèmes biologiques et des procédés" les différentes formes d'écriture suivantes ont été recherchées, entre autres : LISBP, UMR792, UMR 792, UMR5504, UMR 5504, LBB, Laboratoire biotechnologie-bioprocédés, LIPE, Laboratoire d'ingénierie des procédés de l'environnement, EA 833. En effet, ce laboratoire est le résultat de la fusion en 2007 des laboratoires LBB et LIPE.

<sup>12</sup> Par exemple, le SPL Mécaloire compte parmi ses membres le bureau Veritas de Villars (Loire). Or, dans les publications, le Bureau Veritas apparaît souvent sans indicatif d'adresse, de la manière suivante : Bureau Veritas, France. C'est également le cas avec d'autres établissements : Aubert et Duval, Idestyle, IMS ...

<sup>13</sup> Source : titre, numéro et volume de la revue

<sup>14</sup> Titre de l'article

- ✓ Pays d'affiliation pour mettre en évidence des coopérations internationales ;
- ✓ Marquage pour les affiliations et mettre en évidence les acteurs et leurs coopérations ;
- ✓ Année pour mettre en évidence les évolutions ;
- ✓ Code de classement PASCAL pour mettre en évidence les spécialisations du pôle.

### **2.2.3. Traitements sur les données (contenus)**

Les contenus de ces champs ne sont pas toujours manipulables tels quels et demandent :

- ✓ Le choix de la granularité de la structure. On peut simplifier une structure complexe, ou au contraire détailler certains champs : dans la présente étude, le pays d'affiliation a été extrait dans un nouveau champ.
- ✓ Le choix de la granularité des informations. Pour les notices issues de la base de données PASCAL, les codes de classement (champ CC dans le tableau ci-dessous) renvoient à des thématiques avec plusieurs niveaux de détails. Le niveau le plus agrégé comprend 4 caractères (trop général) et le niveau le plus fin 10 caractères (trop détaillé). La nomenclature à 6 caractères, bon compromis, est donc retenue. Les codes sont tous ramenés à cette racine de 6 caractères. Ils ont été verbalisés (transformation des codes alphanumériques en intitulés français). Par exemple, pour le code de classement 001B80A10F, on retient le code à 6 caractères (001B80) dont l'intitulé est Physique, domaines interdisciplinaires : science des matériaux ; rhéologie.
- ✓ L'homogénéisation à l'intérieur d'un champ (ex : écriture des auteurs différente suivant le type de publication ou la base de données, variantes des affiliations, etc.). Ici, seuls les noms des organismes (champ AF dans le tableau ci-dessous) appartenant au cluster ont été homogénéisés. Par exemple, pour l'organisme : TREFLE-Site ENSCPB, 16 Avenue Pey Berland, 33607 Pessac, Bordeaux, France, le nom a été homogénéisé de la manière suivante : Trèfle – Pessac Talence.
- ✓ L'enrichissement à l'aide de tables extérieures. Les affiliations non membres du cluster ont été marquées de la manière suivante : Entreprise – Département géographique ; Laboratoire, grande école ou centre technique localisés – Département géographique; CEA – Département géographique; CETIM – Département géographique; Organisme – Pays. Par exemple, l'organisme Department of Mechanical Engineering Technology, Benha High Institute of Technology, Benha, 13 512, EGYPT est recodé de la manière suivante : Organisme\_Egypte.

---

<sup>15</sup> Base d'origine (Bases de données bibliographiques interrogées : ici PASCAL et SCI), Auteur, Affiliation, Périodique, Type de document, Descripteur, Mot du titre anglais

## Exemple de notice originale simplifiée issue de la base PASCAL<sup>16</sup>

NO : PASCAL 05-0213181 INIST

ET : A fixed domain model for microsegregation during solidification of binary alloys

AU : ARQUIS (Eric); RADY (Mohamed)

AF : TREFLE-Site ENSCPB, 16 Avenue Pey Berland/33607 Pessac, Bordeaux/France (1 aut.); Department of Mechanical Engineering Technology, Benha High Institute of Technology/Benha, 13 512/Egypte (2 aut.)

SO : Heat and mass transfer; ISSN 0947-7411; Allemagne; Da. 2005; Vol. 41; No. 6; Pp. 545-558; Bibl. 26 ref.

LA : Anglais

EA : A fixed-domain numerical model for microsegregation during alloy solidification is developed. The phenomena of solute partitioning at the moving solid/liquid interface and subsequent redistribution by diffusion in the solid and liquid phases have been formulated using volumetric terms. A solute balance equation valid for the whole domain comprising the solid and liquid phases has been obtained in terms of the liquid concentration. The effects of microstructure coarsening on microsegregation has been described and included in the present model. Numerical experiments and comparisons have been carried out between the present fixed-domain model, previous deforming-domain models, and the exact analytical solutions available in the literature

CC : 001B80A10F; 001B80A30F

FD : Croissance cristalline en phase fondue; Interface liquide solide; Solidification; Partage phase; Alliage binaire; Ségrégation; Microstructure; Modélisation; Simulation numérique; 8110F; 8130F

ED : Crystal growth from melts; Liquid solid interface; Solidification; Phase partition; Binary alloys; Segregation; Microstructure; Modelling; Digital simulation

SD : Interfase líquido sólido; Repartición fase

**A fixed domain model for microsegregation during solidification of binary alloys**

Base origine : *Webspirs ISD*

Titre : A fixed domain model for microsegregation during solidification of binary alloys

Auteur : *ARQUIS Eric ; RADY Mohamed*

Affiliation : TREFLE Site ENSCPB, 16 Avenue Pey Berland, 33607 Pessac, Bordeaux ; Department of Mechanical Engineering Technology, Benha High Institute of Technology, Benha 13 512

Pays d'affiliation : *France ; Egypt*

Marquage : *TREFLE - Pessac Talence ; Organisme Egypte*

Source : Heat and mass transfer, 2005; 41 (6): 545-558

Date de publication : 2005

Nom journal : *Heat and mass transfer*

ISSN : 0947-7411

Résumé : A fixed domain numerical model for microsegregation during alloy solidification is developed. The phenomena of solute partitioning at the moving solid/liquid interface and subsequent redistribution by diffusion in the solid and liquid phases have been formulated using volumetric terms. A solute balance equation valid for the whole domain comprising the solid and liquid phases has been obtained in terms of the liquid concentration. The effects of microstructure coarsening on microsegregation has been described and included in the present model. Numerical experiments and comparisons have been carried out between the present fixed domain model, previous deforming domain models, and the exact analytical solutions available in the literature

Code de classement : 001B80A10F ; 001B80A30F ; 8110F ; 8130F

Code classement PASCAL (6 car.) : 001B80 ; 001B80 ; 8110F ; 8130F

Code classement PASCAL : *PHYSIQUE. DOMAINES INTERDISCIPLINAIRES: SCIENCE DES MATERIAUX; RHEOLOGIE. ; PHYSIQUE. DOMAINES INTERDISCIPLINAIRES: SCIENCE DES MATERIAUX; RHEOLOGIE.*

Numéro : 16650110; 050213181

Copyright : 2005 INIST CNRS. All rights reserved.

Figure 1 : Exemple d'une notice "finalisée", telle qu'affichée par Servist

### 2.2.4. Transcodage (jeux de caractères)

Enfin, il ne faut pas oublier la conversion des jeux de caractères. Cette opération est importante pour préserver les caractères spéciaux (accentués, mathématiques et autres) lors des reformatages nécessaires aux prétraitements des données. De plus, Servist utilise le codage UTF-8 qui doit donc être le codage du corpus final pour le chargement du serveur.

<sup>16</sup> NO : identifiant / ET : Titre de l'article / AU: Auteurs / AF: Affiliations / SO : référencement du périodique / LA : langue de l'article / EA : résumé anglais / CC: Code de classement / FD : mots-clés français / ED : mots-clés anglais / SD : mots-clés espagnols /

## 2.3. Fonctionnalités de Servist

Servist repose sur la volonté de prendre en compte des besoins spécifiques, en partie pris en charge aujourd'hui par les différentes solutions à la disposition des veilleurs. Cette plateforme est le résultat d'une collaboration engagée avec le LORIA en 1999 (Bonvallot et al. 2007). Elle peut être utilisée pour calculer et mettre à disposition des indicateurs liés à la production scientifique que cette dernière soit liée à une thématique, un organisme ou un territoire. Elle peut l'être également pour un pôle de compétitivité. Nous allons voir de quelle manière à travers l'exemple du pôle de compétitivité Xylofutur.

### 2.3.1. Exploration du corpus de notices

Servist permet une exploitation facile du corpus de notices bibliographiques grâce aux fonctionnalités suivantes :

- ✓ **Recherche rapide** : permet une recherche quel que soit l'index (présent sur toutes les pages) ;
- ✓ **Recherche** : permet une recherche dans tout index en laissant le choix du mode (strict, par troncature à droite, par expression régulière – sous-chaîne) ;
- ✓ **Liste des termes d'un index** : affiche le contenu d'un champ en liste de termes, avec le nombre de notices contenant ce terme. On peut accéder à toutes ces notices ;
- ✓ **Nuage des termes d'un index** : affiche le contenu d'un champ comme un nuage de termes, le nombre de notices contenant un terme étant affiché au survol de celui-ci par le curseur de la souris. Les termes du nuage sont également combinables en une requête booléenne à l'aide des opérateurs ET – OU – SAUF ;
- ✓ **Notices** : affiche une liste de notices correspondant à une requête, permet aussi de visualiser chacune de ces notices ;
- ✓ **Historique** : liste l'historique des requêtes et le nombre de notices trouvées, permet la combinaison deux à deux de chacune de ces requêtes avec les opérateurs booléens ET, OU et SAUF. Tout clic sur un terme est historisé.

### 2.3.2. Calculs bibliométriques

Servist permet également d'effectuer des calculs bibliométriques grâce aux fonctionnalités suivantes :

- ✓ **Bibliométrie avancée** : affiche la répartition de termes, avec tout ou partie du corpus, d'un croisement entre deux ou trois index ;
- ✓ **Cooccurrences** : affiche les cooccurrences entre des termes d'un même index pour tout ou partie du corpus.

### 2.3.3. Visualisation des résultats

Les résultats peuvent être visualisés de différentes manières :

- ✓ **Croisement** : affiche, à partir du choix de deux requêtes dans le module *Historique*, les volumes correspondant à chaque requête et à leur conjonction (opérateurs ET – OU - SAUF) ;
- ✓ **Cartographie** : projette une répartition calculée dans le module *Bibliométrie avancée* sur une carte (par exemple, une mappemonde). En général, il s'agit de données géographiques (les pays de publication ou d'affiliation) ;
- ✓ **Histogramme** : projette une répartition calculée dans le module *Bibliométrie avancée* sur un diagramme en barres.

Quelques unes de ces fonctionnalités seront illustrées dans les scénarios suivants.

## 2.4. Scénarios : illustration avec le pôle Xylofutur

Les fonctionnalités proposées par Servist permettent de répondre à des besoins classiques de veille comme le calcul d'indicateurs. Ainsi, il est possible de décrire divers protocoles opérationnels – les scénarios.

Ces derniers vont être illustrés, dans le cadre de l'étude pour Xylofutur, par la construction d'indicateurs pour les pôles de compétitivité :

- ✓ Exploration de premier niveau (comptages) ;
- ✓ Analyse statistique par tris croisés sur deux ou trois champs (une particularité de Servist) ;
- ✓ Construction d'un indicateur de production scientifique avec visualisation géographique des résultats.

Pour le pôle Xylofutur, 2519 notices entre 1995 et 2009 ont été recensées puis reformatées en XML. Leur description documentaire a été modifiée. Le contenu de certains champs a été homogénéisé et/ou enrichi.

### 2.4.1. Comptages

Servist permet d'identifier les principaux publiants d'un pôle de compétitivité, l'évolution du nombre de publications et les principales spécialisations traitées.

### Production scientifique

L'index "Marquage" dédié aux affiliations du corpus permet d'afficher la production scientifique de chaque membre du pôle et de leurs copubliants. Les résultats sont déchargeables pour être manipulables dans Excel.

Nom	Nb d'items	Nb docs
Base origine	2	2519
Auteur	4985	2519
Affiliation	4846	2519
Pays d'affiliation	57	2519
Marquage	223	2519
Année	15	2519
Périodique	793	2432

Marquage	Nb docs
LPMC - Palaiseau	827
LISBP - Toulouse	739
TREFLE - Pessac Talence	468
US2B - Talence Cestas	330
Organisme Etats-Unis	145
Laboratoire - Gironde	132
Grandes Ecoles - Institut - Essonne	125
Grandes Ecoles - Institut - Haute Garonne	123
CTP - Grenoble	121
Organisme Royaume-Uni	116

Figure 2 Liste des publiants

Membres du cluster	Nombre de publications	Répartition en %
LPMC – Palaiseau	827	33%
LISBP – Toulouse	739	29%
TREFLE - Pessac Talence	468	19%
US2B - Talence Cestas	330	13%
CTP – Grenoble	121	5%
FCBA – Moulis Bordeaux	43	2%
Centre Régional de la Propriété Forestière d'Aquitaine - Bordeaux	3	0%
Xylomeca - Moulin Neuf	3	0%
Tembec R&D Tartas - Gradignan	2	0%
Vilmorin - La Mégnière	1	0%

Figure 3 Tableau après traitement sous Excel

## Evolution de la production scientifique

L'onglet "Bibliométrie" permet de sélectionner le champ consacré à l'année de publication afin de calculer la production scientifique puis de montrer son évolution à travers un histogramme.

La mise à jour très régulière de la base permet d'avoir des informations très récentes, contrairement à d'autres bases de données habituellement utilisées dans les observatoires, pour lesquelles on constate plusieurs années de retard.

Pour 2009, le nombre de publications est encore peu élevé car tous les articles parus pour cette année n'étaient pas encore recensés dans la base PASCAL au moment de l'extraction (juin 2009). Ils le seront à la fin du premier semestre 2010.

2519 notices

Bibliométrie		
Nom	Nb d'items	Nb docs
Base origine	2	2519
Auteur	4985	2519
Affiliation	4846	2519
Pays d'affiliation	57	2519
Marquage	223	2519
Année	15	2519
Périodique	793	2432

Descripteurs

Nom	Nb d'items	Nb docs
Type document	8	916
Descripteur	8674	2339
Code classement SCI	116	912
Code classement PASCAL	57	1553
Mot du titre anglais	5701	2519

Recherche: Base origine [v] Liste de Notices Liste de descripteurs OK

Index Notices Recherche Historique **Bibliométrie** Croisement Histo. Carte

Application sur: tout le corpus

Aucun

- Aucun
- Base origine (2 items)
- Auteur (4985 items)
- Affiliation (4846 items)
- Pays d'affiliation (57 items)
- Marquage (223 items)
- Année (15 items)**
- Périodique (793 items)
- Type document (8 items)
- Descripteur (8674 items)
- Code classement SCI (116 items)
- Code classement PASCAL (57 items)
- Mot du titre anglais (5701 items)

© 2006-2010 INIST-CNRS. Tous droits réservés.

Figure 4 Sélection du champ pour calculer l'indicateur lié à l'année

2519 notices

Bibliométrie		
Nom	Nb d'items	Nb docs
Base origine	2	2519
Auteur	4985	2519
Affiliation	4846	2519
Pays d'affiliation	57	2519
Marquage	223	2519
Année	15	2519
Périodique	793	2432

Descripteurs

Nom	Nb d'items	Nb docs
Type document	8	916
Descripteur	8674	2339
Code classement SCI	116	912
Code classement PASCAL	57	1553
Mot du titre anglais	5701	2519

Recherche: Base origine [v] Liste de Notices Liste de descripteurs OK

Index Notices Recherche Historique Bibliométrie **Croisement** Histo. Carte

Histogramme / tout le corpus (2519 notices)

Répartition par Année

217

174

130

87

43

0

1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009

© 2006-2010 INIST-CNRS. Tous droits réservés.

Figure 5 Graphique de l'évolution par année

## Répartition par domaine de la production scientifique

La répartition par domaine des publications a été calculée à partir des 1603 notices de la base de données PASCAL<sup>17</sup>.

L'index "*Code classement PASCAL*" permet d'avoir la liste des domaines dans lesquels les membres du cluster publient.

Comme notre démarche n'est pas centrée au départ sur un choix de nomenclature prédéfinie mais bien sur la liste des membres, il est ainsi possible d'analyser la diversité des domaines technologiques dans un pôle de compétitivité.

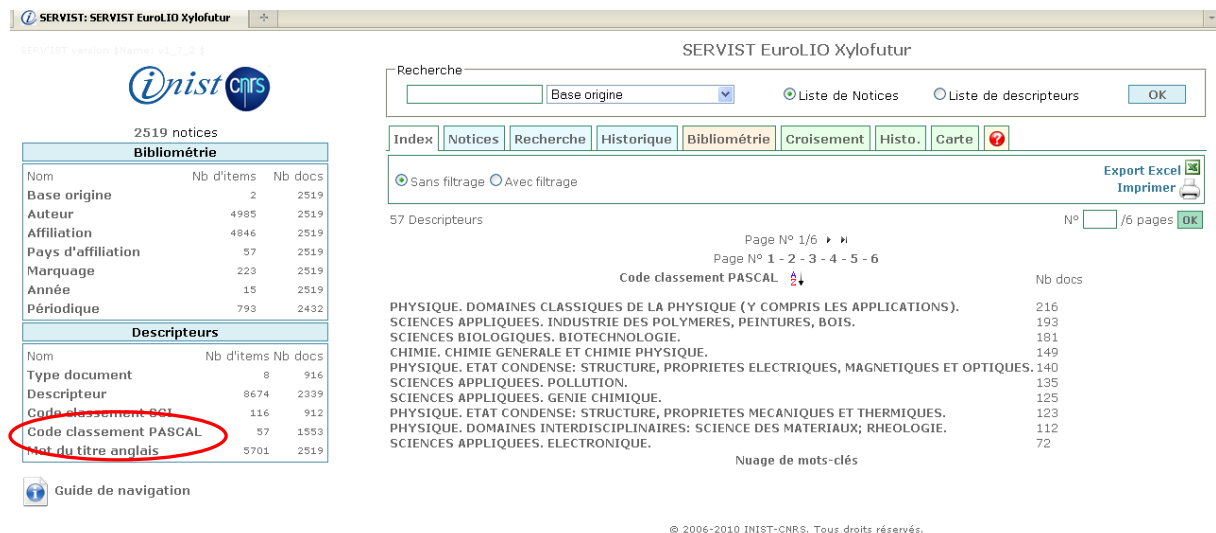


Figure 6 Principaux domaines scientifiques du cluster

#### 2.4.2. Analyse statistique par tris croisés

Servist permet, à l'aide de tris croisés intrachamps et de tris croisés entre deux ou trois champs, d'identifier les coopérations, l'évolution des spécialisations du pôle mais également de chaque publiant.

#### Analyse intrachamp : les coopérations

Par des tris croisés, il est possible de produire une matrice des copublications scientifiques entre les membres et de mieux appréhender le niveau de coopérations scientifiques selon deux points de vue :

- ✓ La nature et l'intensité des coopérations au sein du cluster (entre laboratoires et entre entreprises et laboratoires) ainsi que l'historique ;
- ✓ Le degré d'ouverture avec l'extérieur : coopérations scientifiques entre laboratoires et entre entreprises et laboratoires.

La sélection des index "*Marquage*" au niveau de l'onglet dédié à la Bibliométrie permet de lister les publiants et leurs coopérants.

<sup>17</sup> En effet, pour prendre en considération l'ensemble des 2519 notices, il aurait fallu réindexer les 916 références issues de l'ISI afin d'avoir des données homogènes. Cette démarche chronophage a été écartée d'autant plus que les "*subjects categories*" de l'ISI sont attribuées au périodique et non à l'article.



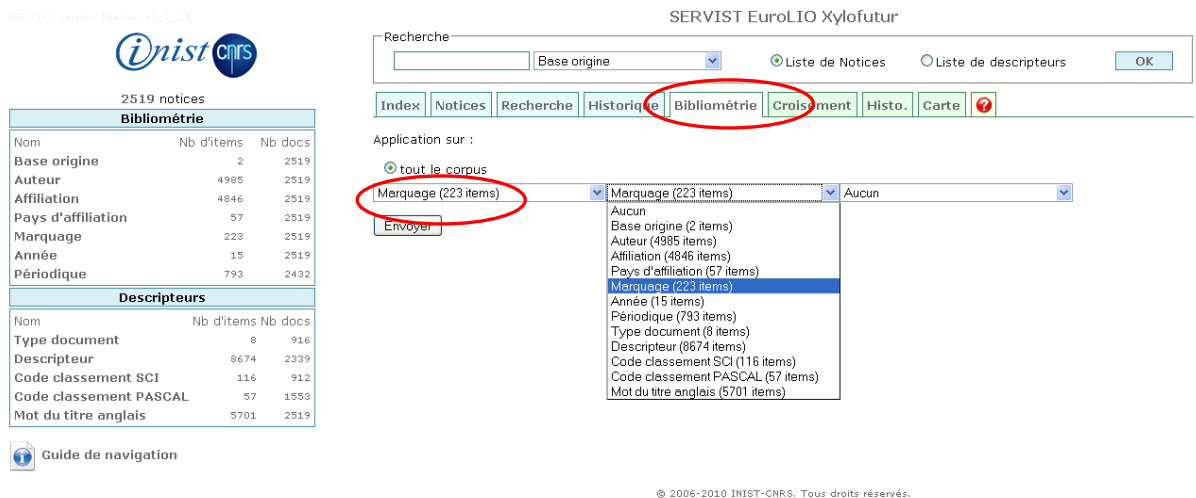


Figure 7 Sélection des champs pour les coopérations

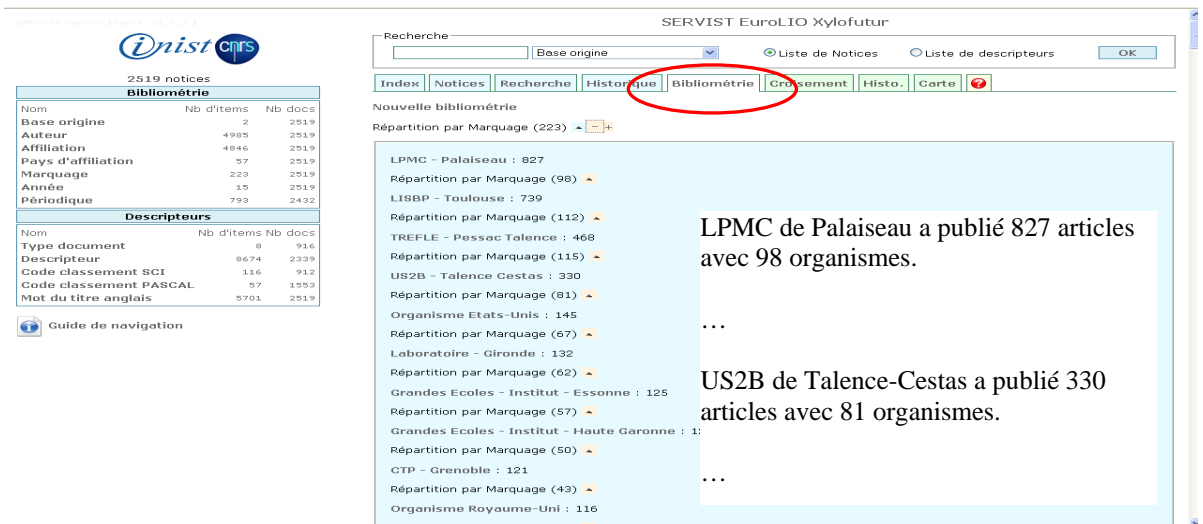


Figure 8 Extrait de la liste des publiants

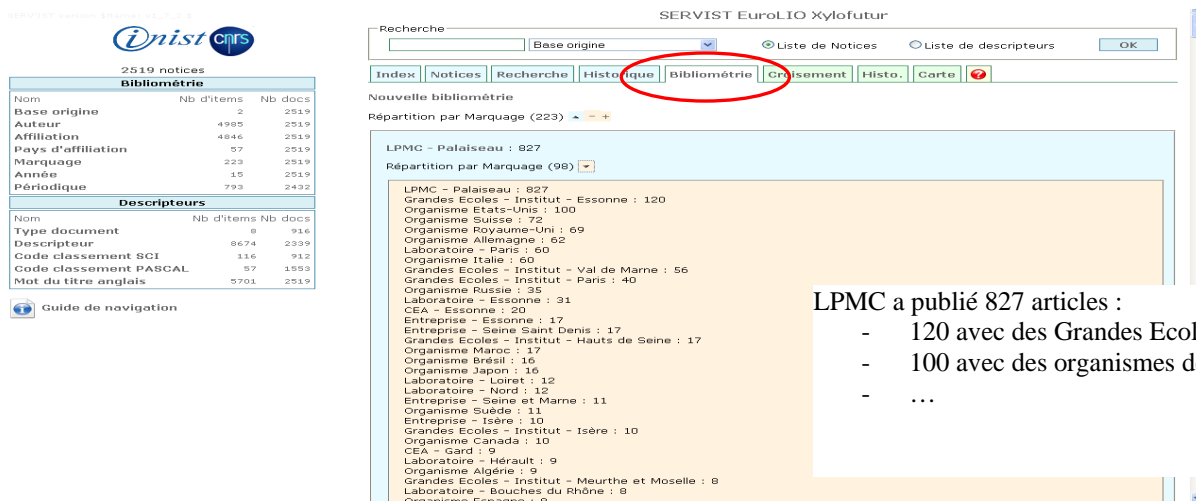


Figure 9 Extrait de la liste des coopérants du LPMC

## Analyse croisant deux champs : évolution des domaines scientifiques

A l'aide de tris croisés entre deux champs, il est possible, par exemple, de rendre compte de la progression annuelle des spécialisations en sélectionnant les index "Code classement PASCAL" et "Année".

SERVIST EuroLIO Xylofutur

Recherche

Base origine [v] Liste de Notices Liste de descripteurs OK

Index Notices Recherche Historique **Bibliométrie** Croisement Histo. Carte

Application sur :

tout le corpus  
 la requête (BASE PASCAL 1603 notices)

Code classement PASCAL (57 items) Année (15 items) Aucun

Envoyer

© 2006-2010 INIST-CNRS. Tous droits réservés.

Figure 10 Sélection des champs pour l'évolution des domaines scientifiques

PHYSIQUE. ETAT CONDENSE: STRUCTURE, PROPRIETES MECANIQUES ET THERMIQUES. : 123

Répartition par Année (14)

2003	16
1997	15
1998	13
2004	12
2001	11
1996	9
2000	9
2002	9
1999	8
2006	7
2005	5
1995	4
2007	4
2008	1

**PHYSIQUE. DOMAINES INTERDISCIPLINAIRES: SCIENCE DES MATERIAUX; RHEOLOGIE. : 112**

Répartition par Année (15)

2001	12
2003	12
2004	12
1997	9
2006	9
1998	8
2000	8
2002	8
1999	6
2005	6
2007	6
2008	6
1996	5
1995	4
2009	1

SCIENCES APPLIQUEES. ELECTRONIQUE. : 72

Répartition par Année (13)

2002	13
2001	11
1999	9
1998	7
2000	7
2003	5

Figure 11 Extrait des résultats de l'évolution annuelle des spécialisations (classement croissant du nombre de publications)

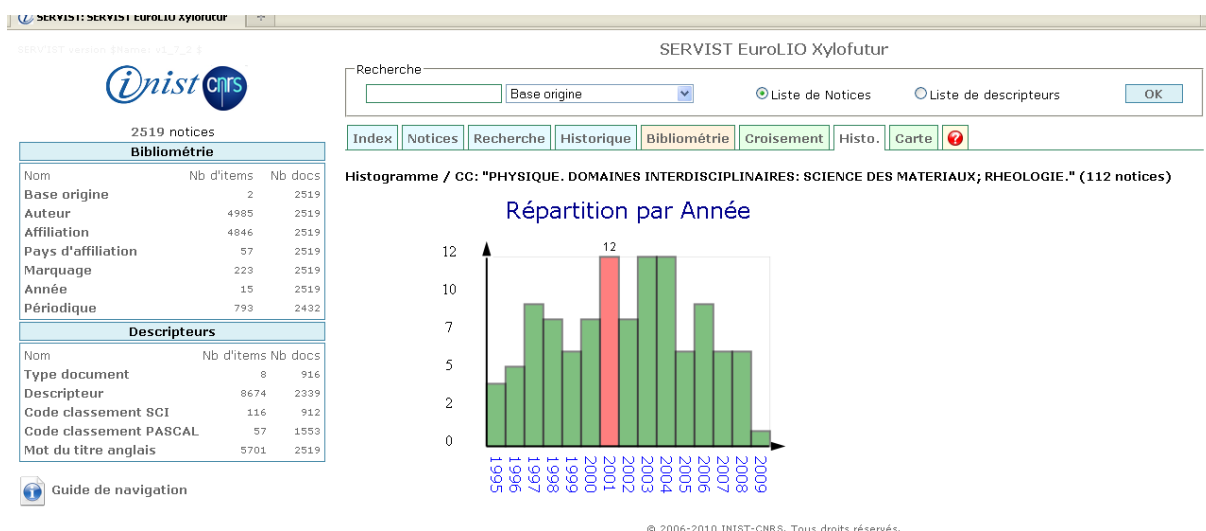


Figure 12 Graphe de l'évolution annuelle "Science des matériaux, Rhéologie"

### Analyse statistique croisant trois champs

Servist permet enfin de faire des analyses statistiques en croisant trois champs, ce qui est une de ses particularités. Ainsi, pourra-t-on, par exemple, isoler les notices PASCAL et voir pour chaque organisme publiant l'évolution des domaines scientifiques en sélectionnant les index "Marquage", "Code classement PASCAL" et "Année".

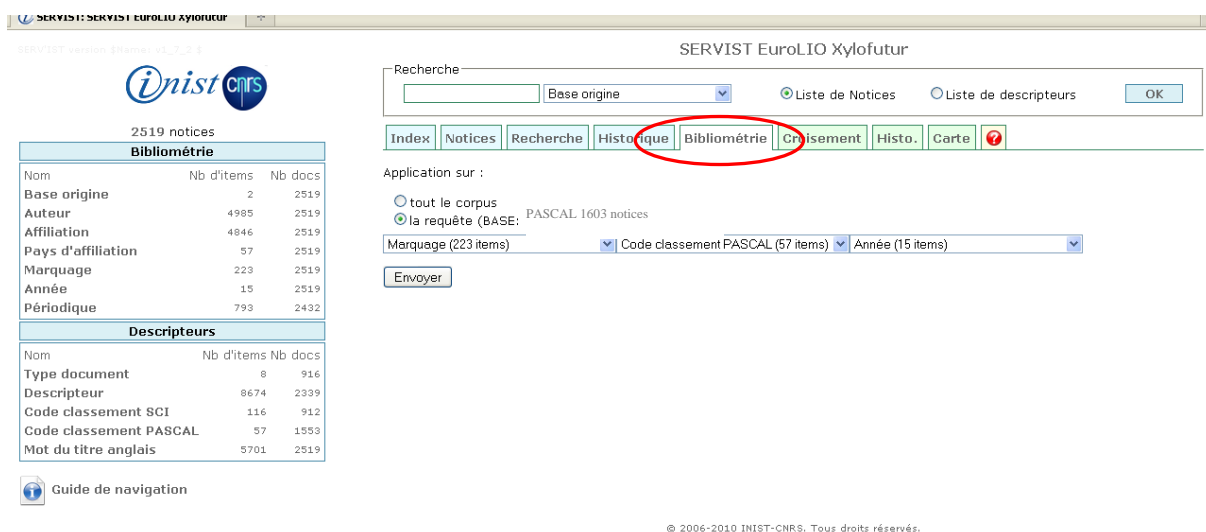


Figure 13 Sélection des champs pour l'évolution annuelle des domaines scientifiques pour chaque publiant

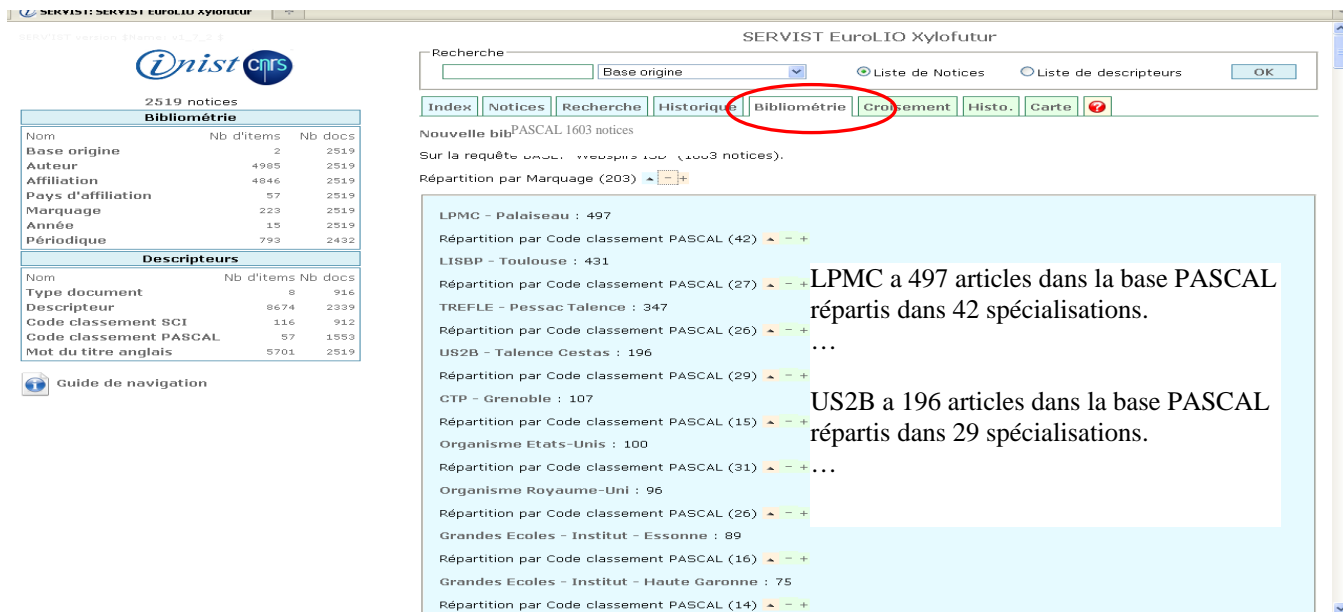


Figure 14 Extrait de la liste des publiants

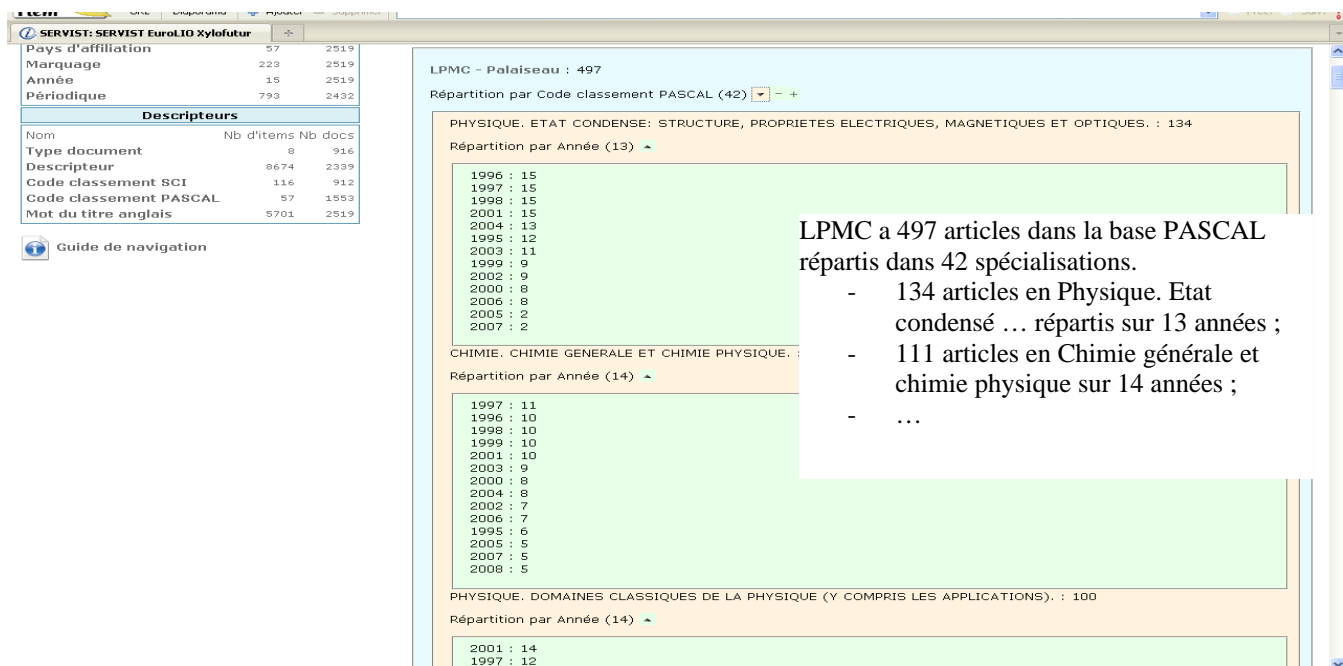


Figure 15 Extrait de l'évolution annuelle des domaines scientifiques pour LPMC

### 2.4.3. Visualisation géographique

Servist permet pour finir une visualisation géographique des résultats à l'échelle des pays en sélectionnant, par exemple, pour un publiant donné l'index "Pays d'affiliation".

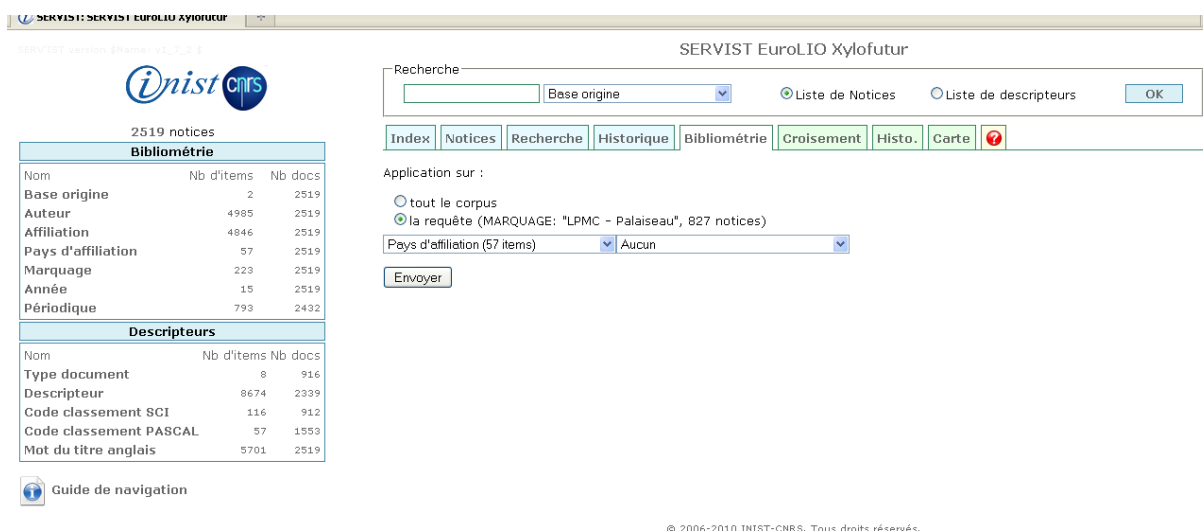


Figure 16 Sélection du champ "Pays d'affiliation"

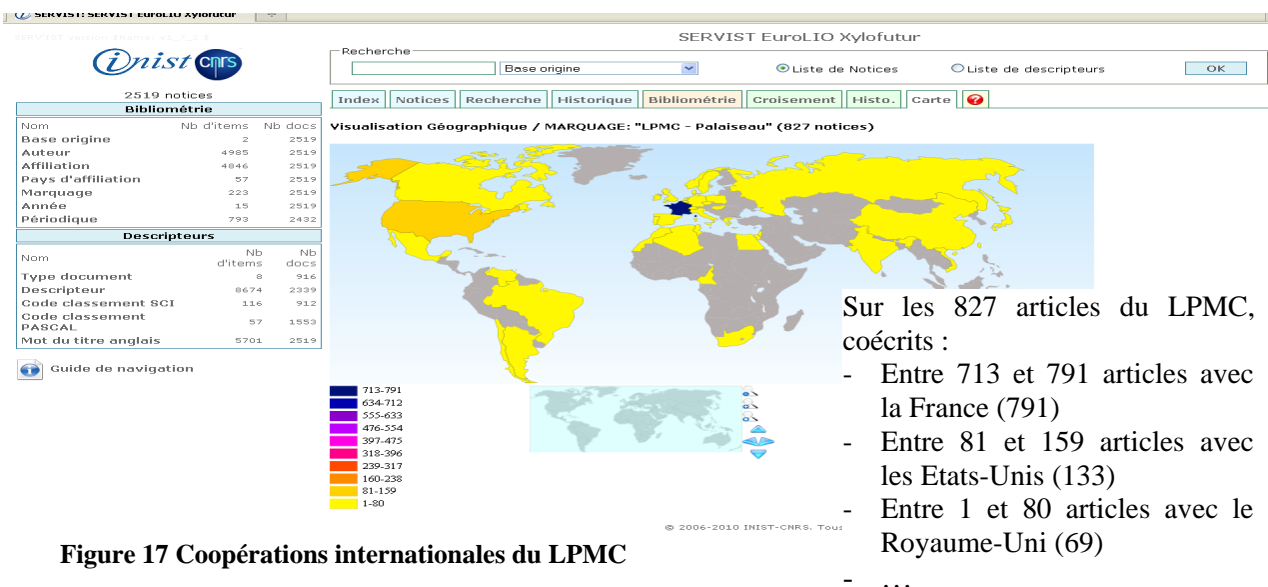


Figure 17 Coopérations internationales du LPMC

#### 2.4.4. Visualisation de graphes réseaux à l'aide de Matheo Analyzer

Outil d'analyse de documents structurés, Matheo Analyzer<sup>18</sup> permet la réalisation de statistiques simples sur un champ, de statistiques croisées (croisement de 2 champs) et de classifications. Le résultat des différentes statistiques est représenté sous formes de graphes : histogramme, camembert, réseau et matrice. Il a été utilisé ici pour mettre en évidence les réseaux entre les différents types d'acteurs. Le graphe montre que les relations scientifiques, de 1995 à 2009, entre les laboratoires de Xylofutur d'une part, et les entreprises de Xylofutur d'autre part, ont été peu nombreuses (seulement 4 co-publications).

Ce type de graphe sera intéressant à analyser dans le temps. La mise en place de la politique des pôles de compétitivité est encore très récente et les dynamiques à l'œuvre notamment entre la science et l'industrie n'apparaissent sûrement pas encore dans les bases de données.

<sup>18</sup> <http://www.matheo-analyzer.com/>

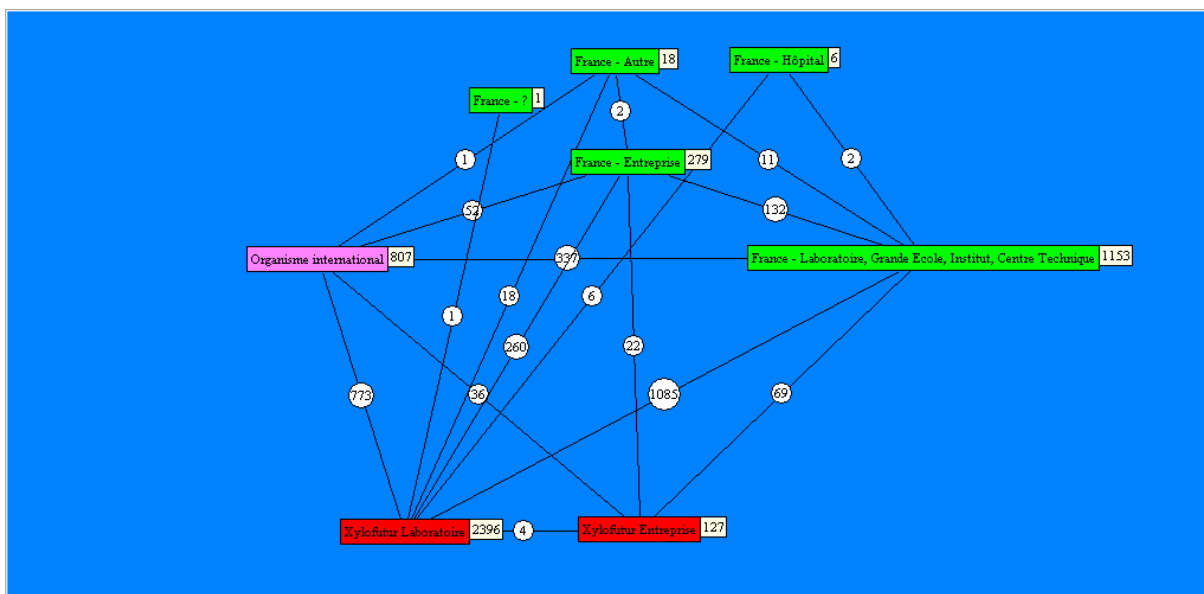


Figure 18 Les relations scientifiques du pôle Xylofutur sur la période 1995 à 2009 par grandes catégories d'organisme

## Conclusion

Cet article montre en quoi les pôles de compétitivité sont difficiles à observer : ils nécessitent d'adopter une grille multidimensionnelle d'indicateurs pour bien cerner leurs caractéristiques. Pour produire des indicateurs à cette échelle, la démarche demande de connaître au préalable la liste des membres pour ensuite rechercher les informations dans les bases de données. Les pratiques d'observation territoriale sont en train d'évoluer et le recours à de nouveaux outils est nécessaire pour traiter des informations autres que celles structurées dans des tableaux.

Dans notre démarche appliquée au pôle Xylofutur, l'utilisation du logiciel Servist pour traiter des informations de la base de données PASCAL, a permis de mesurer le potentiel scientifique de ce pôle. Il s'agit de premiers indicateurs dont il convient de suivre l'évolution pour repérer des tendances relatives :

- ✓ Aux domaines scientifiques ;
- ✓ Aux principaux publiants ;
- ✓ Aux relations scientifiques entre les membres du pôle et à l'extérieur.

Afin de procéder à des comparaisons avec d'autres pôles de compétitivité, l'enjeu va être de généraliser ces indicateurs à tous les pôles. La démarche est cependant assez fastidieuse car la recherche des informations concernant les entreprises et les laboratoires dans la base ne peut se faire qu'à partir du nom de l'organisme et non à partir d'un numéro identifiant unique (type SIRET). L'automatisation de la procédure pour généraliser ces indicateurs à l'ensemble des pôles nécessite donc le développement d'algorithmes pour une meilleure identification des affiliations.

## Bibliographie

Andersson, M., & Karlsson, C., 2004. The role of accessibility for the performance of regional innovation systems. *Working Paper Series in Economics and Institutions of Innovation* (9).

Audretsch, D., & Vivarelli, M. , 1994. Small firms and R&D spillovers : evidence from Italy. *Revue d'économie industrielle* , 67 (1), 225-237.

Autant-Bernard, C., Massard, N., et Charlot, S. (2008), « Les déterminants géographiques de l'innovation : diffusion spatiale des connaissances et choix de localisation » dans T. Madiès, & J.-C. Prager, *Innovation et compétitivité des régions*, Paris: La Documentation Française.

Bonvallot V., Houdry P., Parmentier F., Tramonti A., Vidal S., 2007, Servist : Plateforme d'exploration, de recherche et de visualisation de données multimédias, colloque VSST 2007, 21-25 octobre 2007, Marrakech.

Chalaye S. et Massard N., 2009, Les clusters : diversité des pratiques et mesures de performance, *Revue d'Economie Industrielle*, n°128, 4<sup>ème</sup> trimestre 2009

DATAR / EuroLIO, 2009, Conception d'un tableau de bord territorialisé de l'innovation – suivi des clusters, Rapport d'étude, Octobre 2009.

Duguet, E., 1999. Innovation, diffusion des connaissances et croissance : une analyse micro-économétrique du lien entre la productivité globale des facteurs et l'innovation sur données d'entreprises. Dans *Les chiffres clés de l'industrie, collection "Analyse"* (Vol. 210).

Hendry, C., et Brown, J. (2006), “Dynamics of clustering and performance in the UK Optoelectronics Industry”, *Regional Studies* , vol.40, n°7, pp.707-725.

Lefebvre, P., 2008. Pour une typologie des trajectoires de développement des pôles. *Séminaire Pôles de compétitivité. Comment appréhender la diversité des pôles ?*

OCDE., 2005. *Manuel d'Oslo - Principes directeurs pour le recueil et l'interprétation des données sur l'innovation*. Editions OCDE, 3<sup>ème</sup> édition.

Owen-Smith, J. et Powell, W. W. (2004), “Knowledge Networks as Channels and Conduits : The Effects of Spillovers in the Boston Biotechnology Community”, *Organization Science*, vol.15, n°1, January-February.

Pacitto, J.-C., & Tordjman, F., 1999. L'innovation technologique dans la très petite entreprises industrielle française : ce que disent les statistiques. *Revue Internationale PME* , 12 (3), 59-90.

Rallet, A., & Torre, A., 2001. Proximité géographique ou proximité organisationnelle ? Une analyse spatiale des coopérations technologiques dans les réseaux localisés d'innovation. *Economie appliquée* , 44 (1), 147-171.

Sölvell, O., Lindqvist, G., & Ketels, C., 2003. *The Cluster Initiative Greenbook*. Ivory Tower, Sweden.