



Analyse des publications par la bibliométrie

Nicolas Barts, IRD

Odile Bédu, Inra - Agreenium

Nathalie Darbon, BIU Sciences, Montpellier 2

Dominique Fournier, Inra

Chantal Salson, Agropolis International

Rencontres IST 2014 : L'IST en marche
16 décembre 2014, Agropolis International, Montpellier



Analyse des publications par la bibliométrie

- Définitions
- **La bibliométrie, pour quoi faire ?**
- **Quelles ressources pour quels résultats ?**
- **Résultats**
 - Étude des acteurs : pays, auteurs, institutions
 - Etude des revues : accessibilité, diffusion, notoriété
 - Etude thématique et les limites de l'analyse

Définitions :

Bibliométrie : Application de méthodes mathématiques et statistiques aux livres et autres moyens de communication (Pritchard 1969).

Scientométrie : Application de méthodes statistiques à des données quantitatives (économiques, humaines, bibliographiques) caractéristiques de l'état de la science (revue *Scientometrics*).

Infométrie : l'ensemble des activités métriques relatives à l'information, couvrant aussi bien la bibliométrie que la scientométrie (Egghe et Rousseau 1988).

3 familles d'indicateurs bibliométriques*

Indicateurs de **production et de spécialisation** : part de publications (dans le monde, dans un pays...) , indice de spécialisation...

Indicateurs de **visibilité** : part de citations, indice d'impact relatif...

Indicateurs de **collaboration** : part des co-publications nationales, internationales...

* définitions de l'OST : *Observatoire des Sciences et des Techniques* : <http://www.obs-ost.fr/>

La bibliométrie, pour quoi faire...

Pour la recherche :

estimer le *volume* et *l'impact* des publications

évolution d'un thème de recherche (pays, institutions, sous-thématiques...)

positionner un laboratoire, une institution, un pays

analyser une politique scientifique : déterminer forces et faiblesses et orienter les actions

Pour la politique documentaire :

qualifier et *choisir* les "meilleures" revues auxquelles s'abonner

Pour la veille technologique et sociologie des sciences :

repérer les liens entre disciplines, voire l'apparition de disciplines

Par exemple (1)

Bilans quantitatifs et qualitatifs

d'une production (à différentes échelles : unité, département, centre, institut) et de la comparer éventuellement à un autre niveau d'organisation, à un pays

Démarche exploratoire

dans le corpus à explorer, il s'agit d'identifier des thématiques et des acteurs et de repérer des liens entre ces différents éléments

Par exemple (2)

Les thématiques

La représentation relative de différentes thématiques, les technologies (sujets) émergent(e)s, les «signaux faibles», les évolutions au cours de la période, le repérage des experts sur les thématiques...

Les acteurs

auteurs, institutions, pays/régions (liens entre recherches et intérêts économiques)

leurs domaines de compétences (liens avec les thématiques)

leurs collaborations (liens entre les acteurs)

Par exemple (3)

Les revues

titre de revues (les titres de revues dans lesquelles un groupe publie)

Editeurs (les éditeurs, et les politiques d'accès – Open Access)

Exemple de cahier des charges

- **Objectif**

Commanditaire, date de la demande, date de fin

- **Méthodologie**

Constitution du corpus : choix de la bdd

Equation de recherche proposée

- **Plan de l'étude**

Livrables à la date...

- **Format**

- **Calendrier, réunions de travail**

Quelles ressources pour quel résultat...

Les données de l'Observatoire des sciences et des techniques (OST) pour les demandes institutionnelles

Des bases de données scientifiques pour repérer :
dates, revues, auteurs, affiliations, disciplines
scientifiques

Ressources disponibles dans les institutions

Téléchargement ou non des références

Pourquoi les données de l'OST ?

- La loi organique relative aux lois de finances (**LOLF**) du 1er août 2001 institue de nouvelles règles d'élaboration et d'exécution du budget de l'État.
- L'Inra, IRD, CIRAD, IRSTEA, IFREMER, BRGM dépendent du programme 187 : **Recherche dans le domaine de la gestion des milieux et des ressources.**
- Pour vérifier que ces objectifs sont atteints et que chaque euro alloué est dépensé à dessein, des indicateurs de performance ont été élaborés pour mesurer ces programmes.
- L'OST a été chargé de concevoir et produire des indicateurs relatifs à la recherche et à l'innovation, en particulier les indicateurs consolidés des programmes de la LOLF.
- L'OST utilise la base de données Web of Science™ pour produire les indicateurs de production scientifique.
- L'OST demande aux instituts de recherche de valider la liste de leurs publications ; **cette liste devient la liste de référence de chaque institut.**

Quelles ressources pour quels résultats (1)

Des bases institutionnelles pour travailler sur des corpus validées : auteurs, affiliations

<http://www.ist.agropolis.fr/les-ressources/production-scientifique/archives-ouvertes>

Pour les études sur les acteurs et pour les études thématiques (Horizon, Agritrop, Prodnra, export de l'OST utilisé par l'UM2...)

– Le + : données validées et WOS consolidé par les instituts (80% pour l'IRD)



Exhaustivité si lien avec l'évaluation

Indexation thématique et géographique



– Le – : périmètre limité à l'institution



Quelles ressources pour quels résultats (2)

Des bases de données scientifiques internationales pour repérer : dates, revues, auteurs, affiliations, disciplines scientifiques, citations

Voir les ressources affichées sur le site IST d'Agropolis : <http://www.ist.agropolis.fr/les-ressources/ressources-electroniques>


Pour les études thématiques (CAB, Pubmed, FSTA...)

Le + : classifications, descripteurs, thésaurus 

Le – : une seule adresse par notice 

Pour les acteurs (WoS, Scopus)

– Le + : les adresses de tous les auteurs 

– Le – : les classements thématiques concernent les revues (WoS categories, champs disciplinaires ESI) et non le contenu de l'article 

Types d'analyse (3)

1- Etudes exploratoires sans télécharger les références :

fonction Analyze et Refine du WoS et Scopus : production de listes, exploitables en tableaux Excel

2 – Analyses plus poussées : téléchargement des références :

Téléchargement par lots de références, WoS 500 max

Harmonisation des graphies, affectation des références par unités, réseaux.

Analyse des publications par la bibliométrie

- Définitions
- La bibliométrie, pour quoi faire ?
- Quelles ressources pour quels résultats ?
- **Résultats**
 - Étude des acteurs : pays, auteurs, institutions
 - Etude des revues : accessibilité, diffusion, notoriété
 - Etude thématique et les limites de l'analyse



EXEMPLES DE RÉSULTATS : ÉTUDES DES ACTEURS

Rencontres IST 2014 : L'IST en marche
16 décembre 2014, Agropolis International, Montpellier



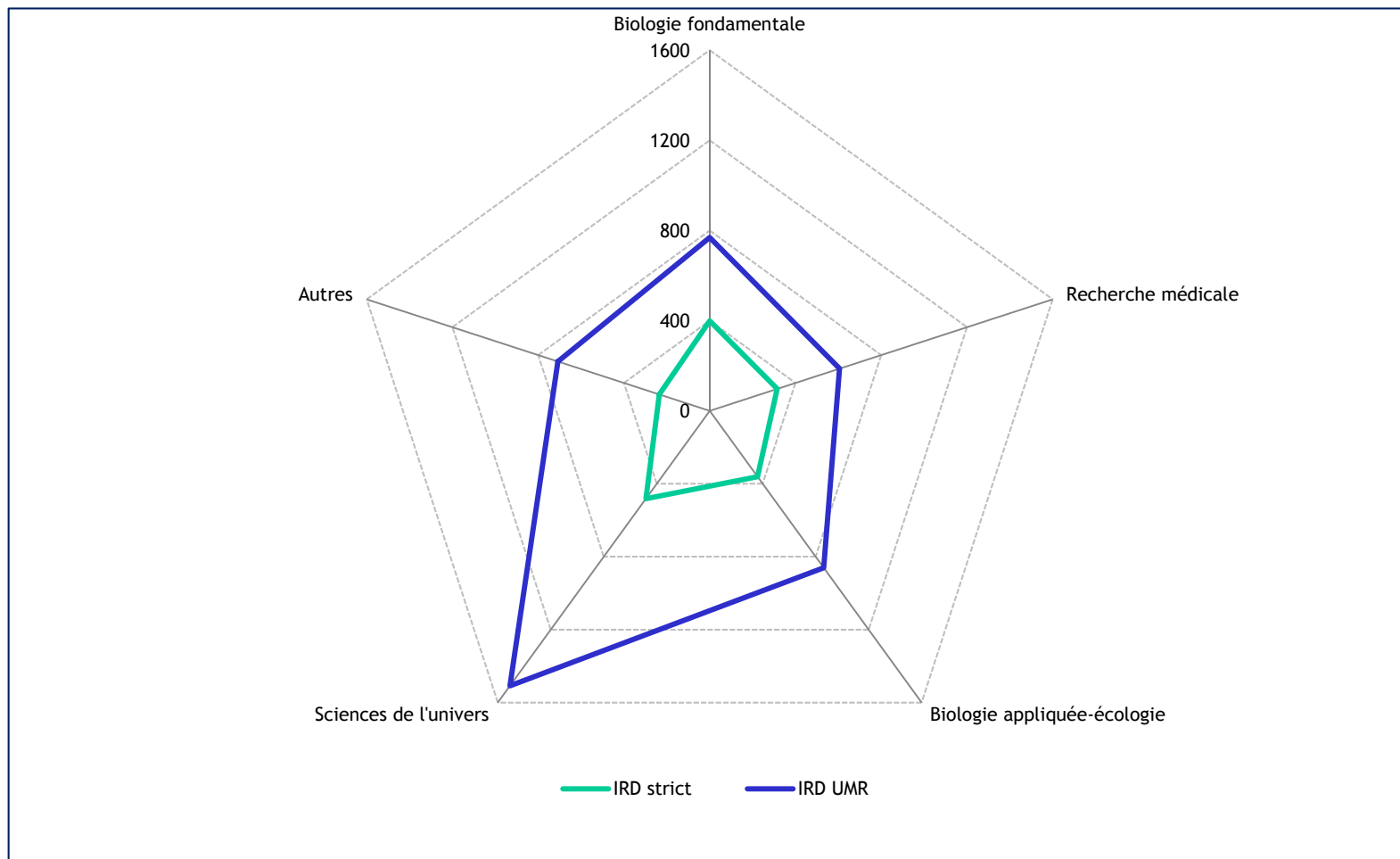
Exemple : Etude exploratoire à l'IRD - Pays

- Etude Flash (online) « PAYS » (ex Perou)
- Ressources utilisés : WoS, fonction Analyze
- Indicateurs :
 - Production du Pérou
 - Copublications France / Pérou
 - Copublications France / Pérou / IRD
- Type d'analyse et logiciels utilisés : WoS online
- Représentations des résultats : tableaux Excel

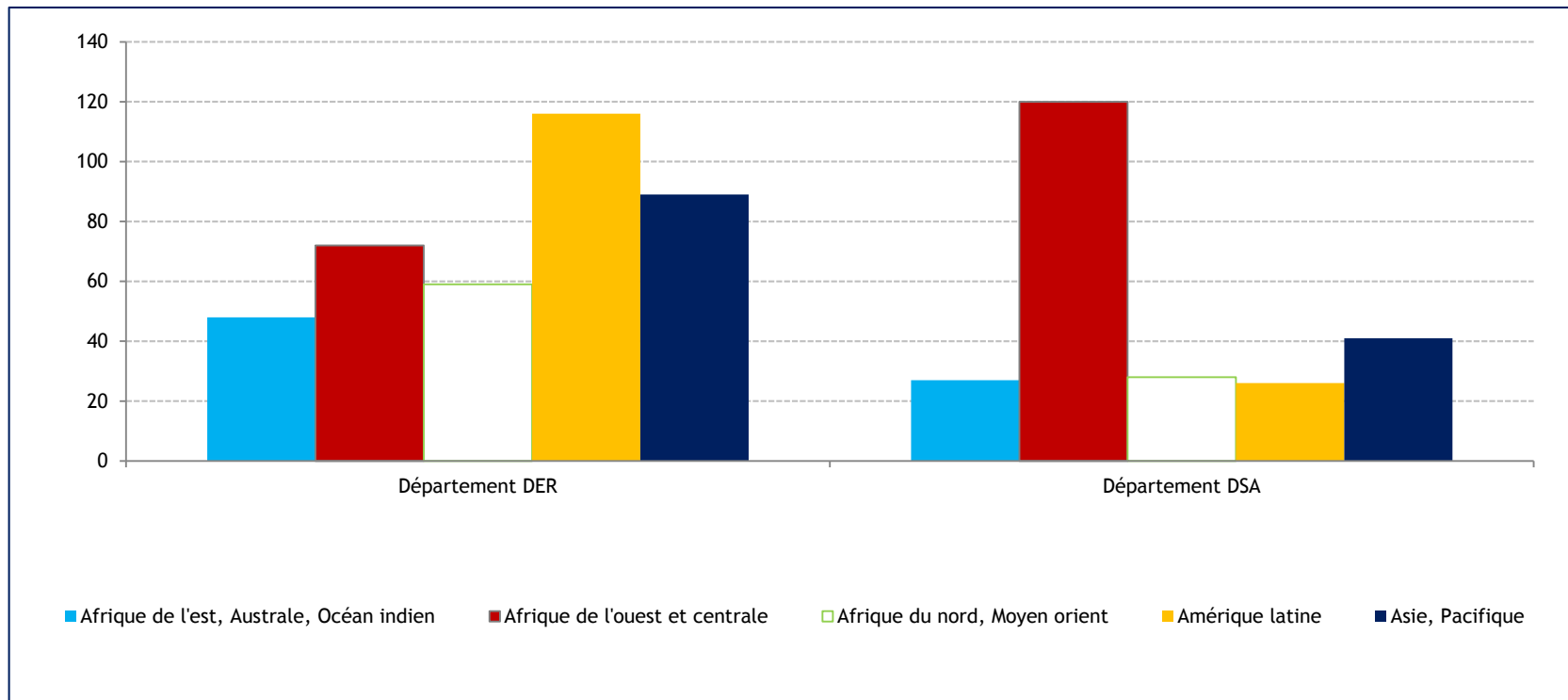
Exemple : Rapport indicateurs annuels - IRD

- Ressources utilisés (WOS, bases bibliométriques de l'IRD : périmètre UMR, périmètre IRD)
- Indicateurs (production scientifique, partenaires et co-publications, principales revues, positionnement)
- Type d'analyse et logiciels utilisés (excel, ...)
- Représentations des résultats

Rapport bibliométrique annuel - IRD



Rapport bibliométrique annuel - IRD



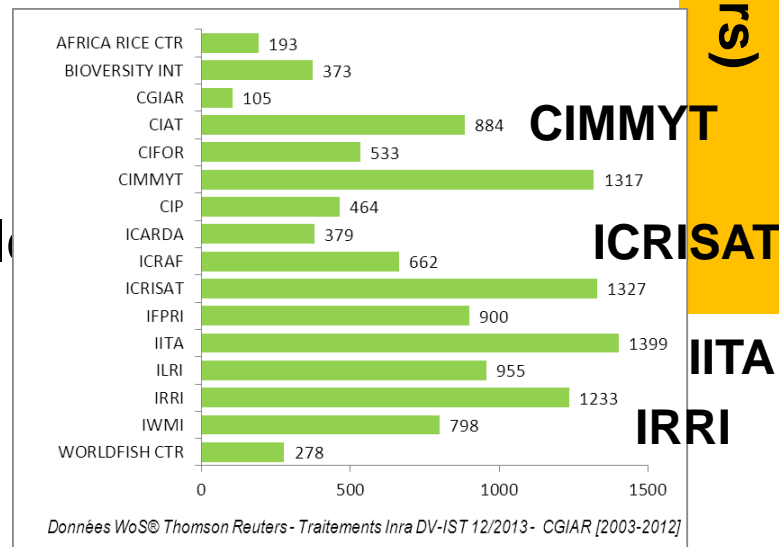
Exemple : Etude CGIAR

Objectifs :

- Etude du CGIAR au travers de ses publications scientifiques
- Cartographie des instituts du CGIAR (implantations, thématiques et collaborations)
- Cartographie des collaborations du CGIAR
- Place de la France dans les collaborations internationales

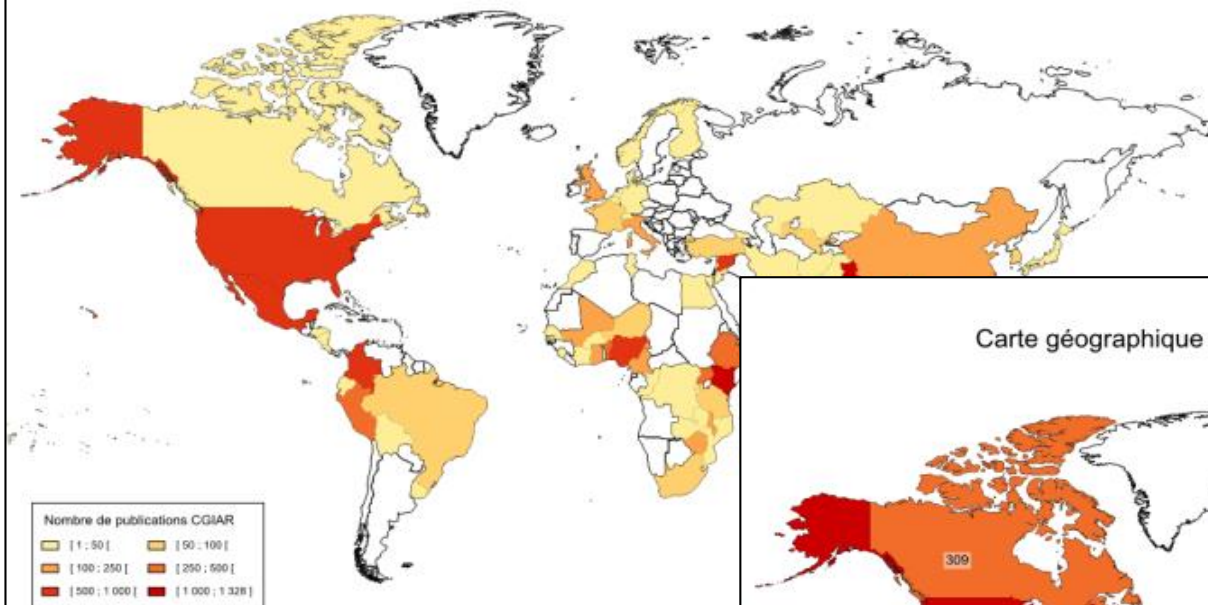
- Ressources utilisés : Web of Science™ : 2003-2012

- 10 785 publications de 2003 à 2012
- Variation sur 10 ans : 71%
- 71 % de co-publications internationales



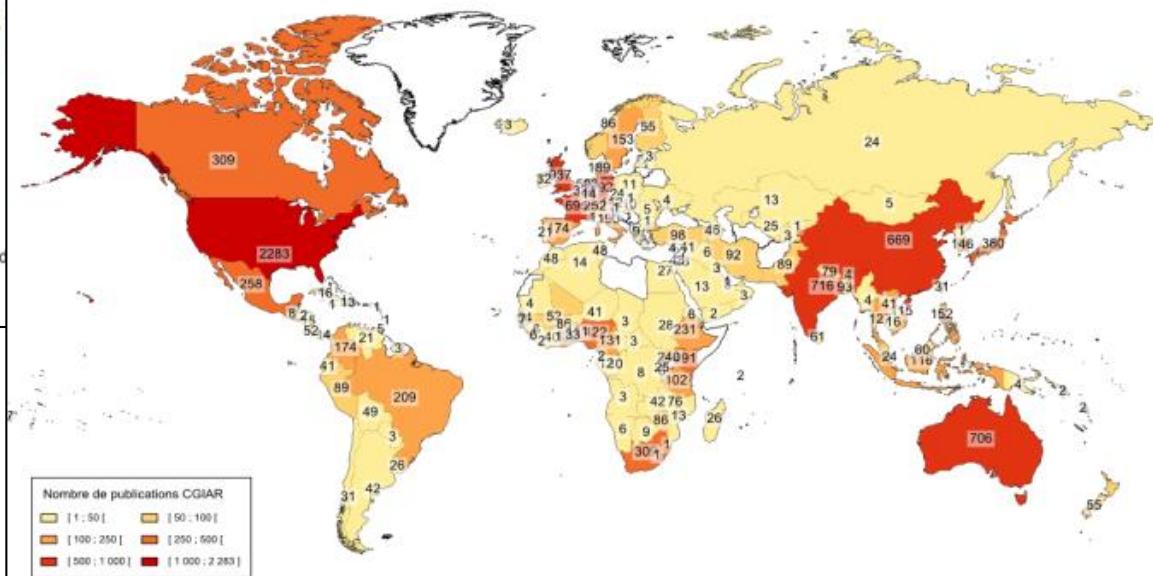
Cartes des implantations du CGIAR et de ses pays partenaires

Carte géographique des implantations du CGIAR 2003-2012



Données WoS® Thomson Reuters – Traitement Inra DV-IST c

Carte géographique des partenaires du CGIAR 2003-2012



Données WoS® Thomson Reuters – Traitement Inra DV-IST décembre 2013 - CGIAR 2003-2012

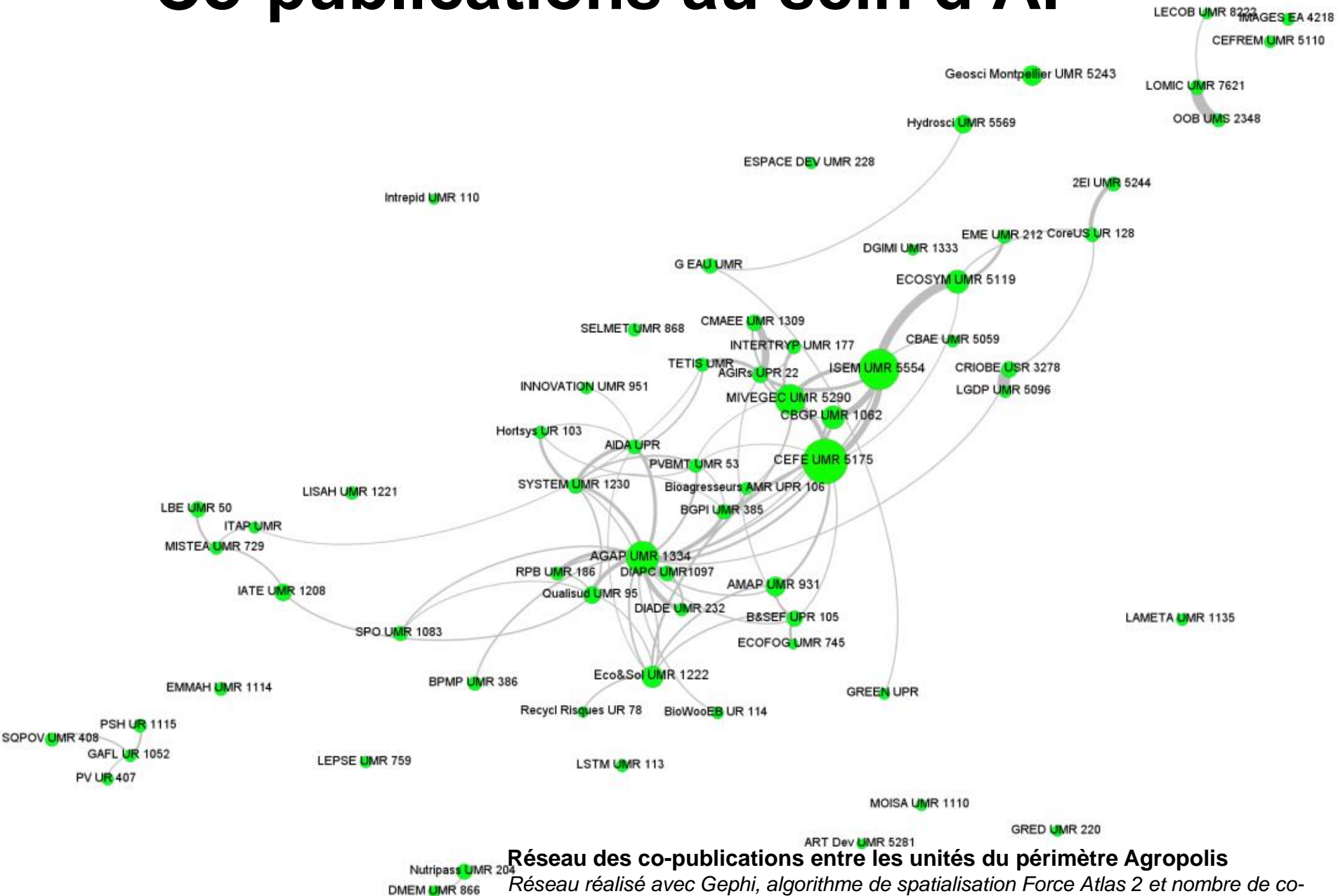
Exemple : Unités du périmètre Agropolis

Agriculture Alimentation Biodiversité Environnement

- Ressources utilisés
 - Interrogation du WOS sur les villes de la région LR
- Indicateurs
 - 13 811 publications des unités du pôle Agropolis (2008-2013), soit le tiers des publications en région LR (35%)
 - Indicateurs de collaborations :
 - intra-unités : 18 % des publications impliquent au moins 2 unités
 - internationales

Collaborations internationales	%
publis avec au moins un pays étranger	66%
publis avec au moins un pays EU28	28%
publis avec au moins un pays méditerranéen	16%
publis avec au moins un pays PSEM	6%
publis avec au moins un pays du SUD	37%
Publis avec au moins un pays BRICS	10%

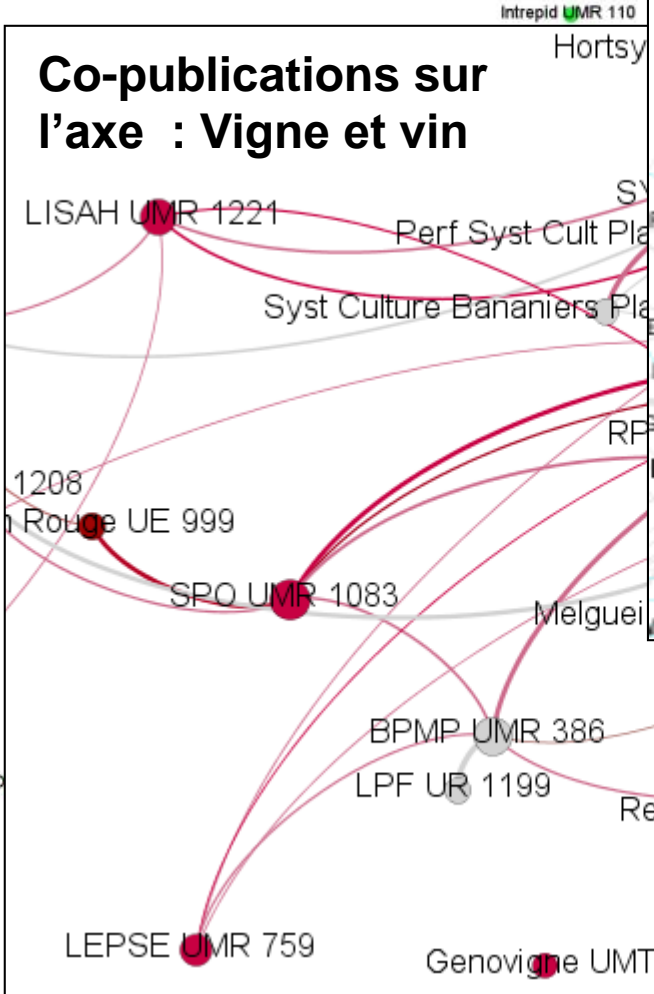
Co-publications au sein d'AI



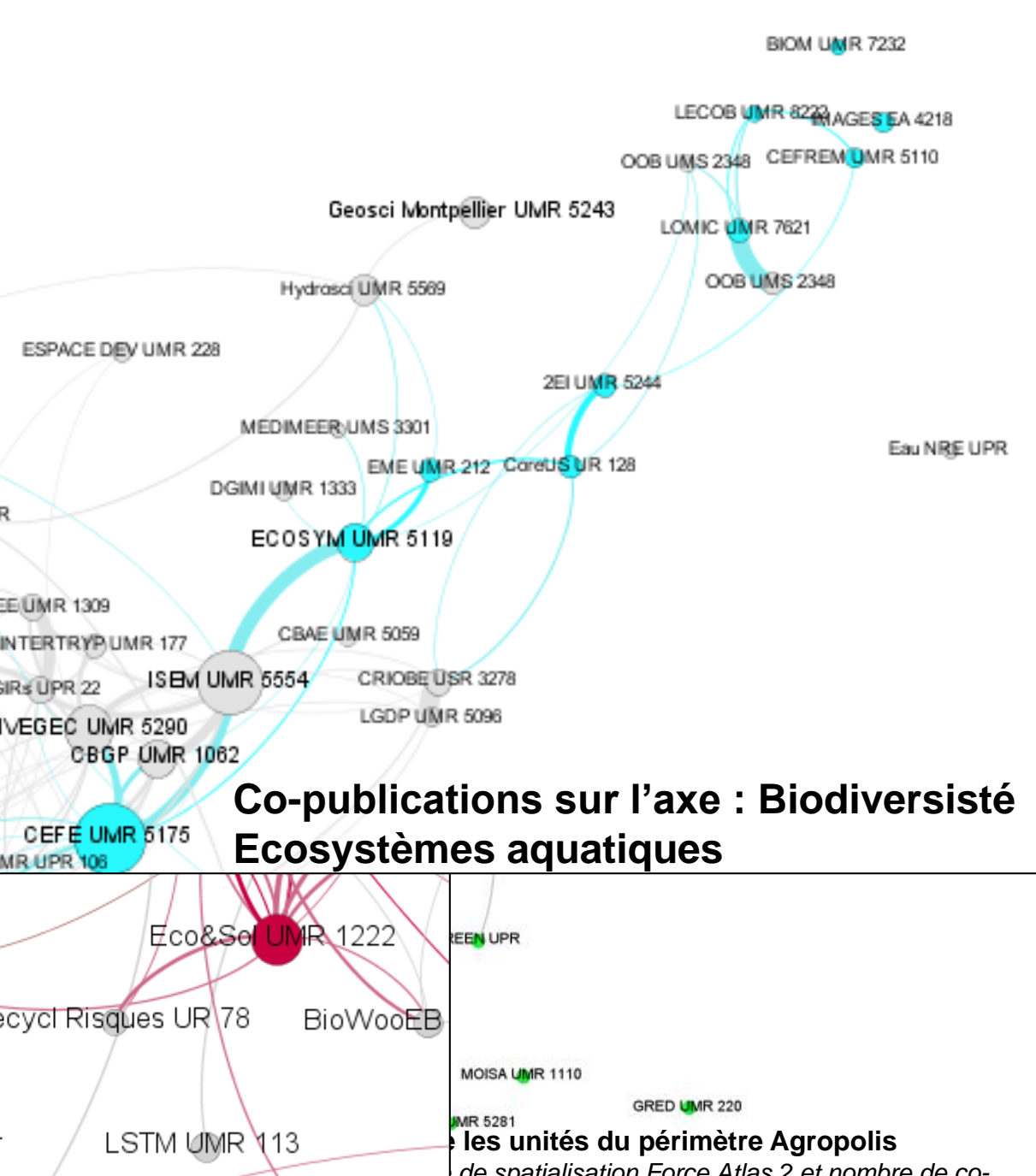
Réseau des co-publications entre les unités du périmètre Agropolis
Réseau réalisé avec Gephi, algorithme de spatialisation Force Atlas 2 et nombre de co-publications entre unités limité à 5 - Source : Données du WOS 2008 - 2013

Co-publications

Co-publications sur l'axe : Vigne et vin



Co-publications sur l'axe : Biodiversité Ecosystèmes aquatiques



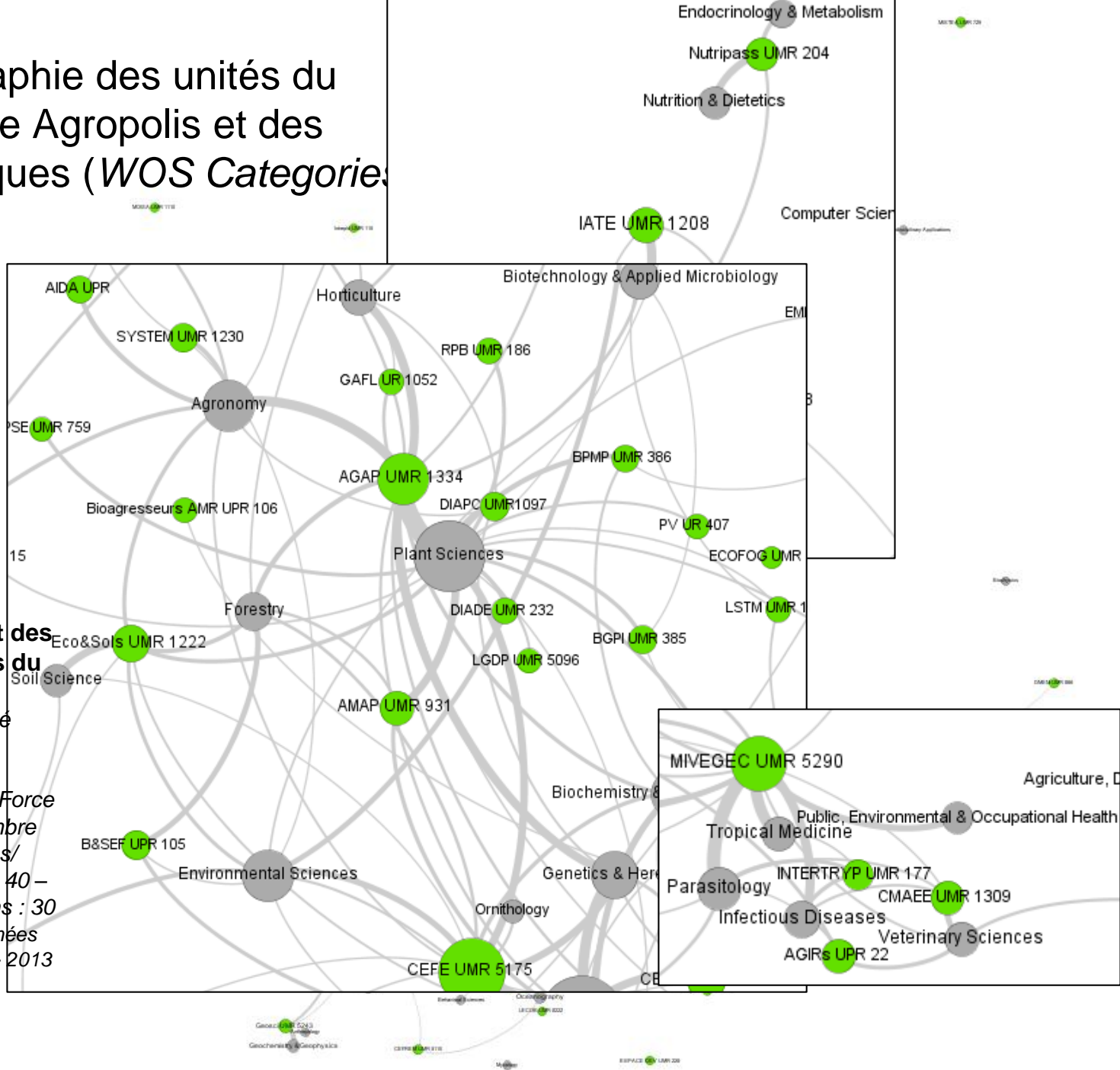
les unités du périmètre Agropolis

de spatialisation Force Atlas 2 et nombre de co-publications entre unités limité à 5 - Source : Données du WOS 2008 - 2013

Cartographie des unités du périmètre Agropolis et des thématiques (WOS Categories)

Réseau des unités du périmètre Agropolis et des thématiques du WOS

Réseau réalisé avec Gephi, algorithme de spatialisation Force Atlas 2 et nombre de publications/ unités limité à 40 – Poids des liens : 30 – Source : Données du WOS 2008 - 2013



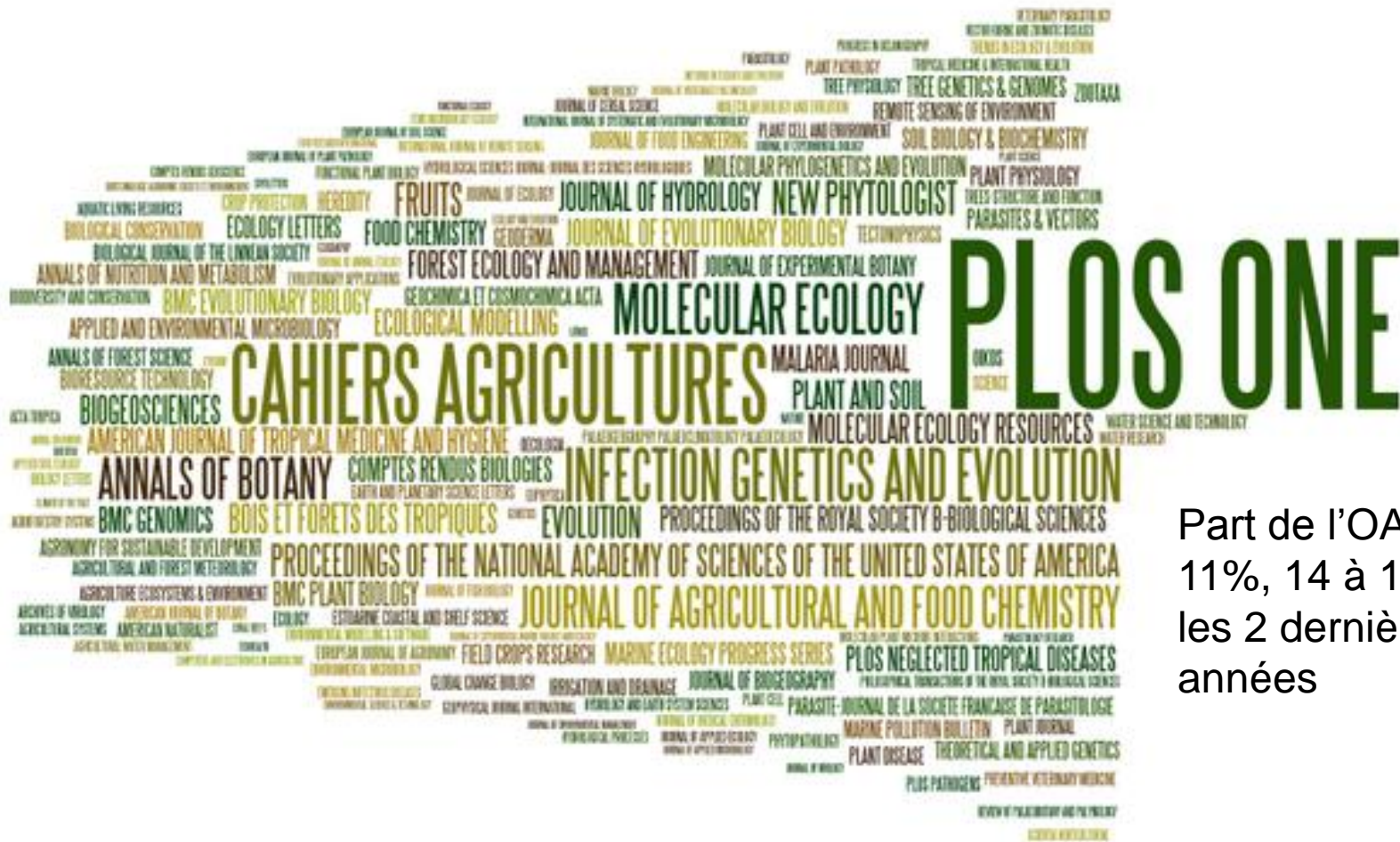
Infométrie
 disciplines scientifiques
 Des bases de données scientifiques
 Les données de l'Observatoire des sciences et des techniques
 Définitions
 Analyse des publications
 réseaux
 Ucinet
 Gephi
 bibliométrie
 dates
 auteurs
 affiliations
 Excel
 Indicateurs de visibilité
 Etude des revues
 auxqueltes s'abonner ou à promouvoir
 Qualifier et choisir les revues auxquelles s'abonner ou à promouvoir
 MySQL/Access
 Définitions
 auteurs
 réseaux
 Ucinet
 Ressources disponibles dans les institutions
 Répondre à un commanditaire :
 logiciels
 Représentation visuelle des données
 Cartes & Données
 Etude thématique
 Gephi
 Les revues
 dates
 Excel
 Analyser une thématique
 revues
 La bibliométrie, pour quoi faire ?
 logiciels
 orienter les actions
 Demarche exploratoire
 Quelles ressources pour quels résultats ?
 Datavrapper
 Nuage de mots
 revues
 Scientométrie

EXEMPLES DE RÉSULTATS : ÉTUDES DES REVUES

Rencontres IST 2014 : L'IST en marche
16 décembre 2014, Agropolis International, Montpellier

Exemple : Unités du périmètre Agropolis

Agriculture Alimentation Biodiversité Environnement



Part de l'OA :
 11%, 14 à 15%
 les 2 dernières années

Représentation des principales revues dans lesquelles les unités du périmètre Agropolis publient (>=10 articles)
 La taille est proportionnelle au nombre d'articles publiés. Source : Données du WOS 2008 – 2013 – Nuage réalisé avec Wordle.net

Exemple : La part de l'Open Acces dans les publications

BIU Sciences de Montpellier 2



Éditeurs principaux UM2

Editeurs	Nombre distincts	d'articles
Elsevier		3337
Wiley-Blackwell		1924
Springer		1313
American Chemical Society (ACS)		711
Oxford University Press		344
American Physical Society (APS)		322
Royal Society of Chemistry		320
Public Library of Science		315

Départements scientifiques publiant dans des revues en *Open Access*

Pôles	Articles distincts en accès libre	Articles distinct tous éditeurs confondus, d'accès libre ou payant	Accès libre par rapport au reste des publications du pôle
BS	1089	2679	40,6 %
CHIMIE	142	2938	4,8 %
EVAP	1225	5514	22,2 %
MIPS	447	3194	14 %
SC	8	17	47 %
SHS	5	45	11,1 %
Total	2916	14387	100 %

28 % des publications du pôle XXX ne sont pas couvertes par les abonnements de la BU

Éditeurs	Nbre articles distincts
IEEE	206
Springer	130
IOP Institute of Physics	113
SPIE	81
#N/A	61
Scientific.net	23
Cambridge University Press	19
Taylor & Francis Group	19
Optical Society of America	18
Gordon and Breach Science Publishers	14
EDP Sciences	12
Wiley-Blackwell	11

76
éditeurs

28 % des publications du pôle XXX ne sont pas couvertes par les abonnements de la BU

Est-ce significatif par rapport à tout ce que l'UM2 publie chez ces éditeurs ?

Éditeurs	Articles sans accès	Publis. totales	% sans accès
IEEE	206	222	92,8 %
Springer	130	1313	9,9 %
IOP Institute of Physics	113	267	42,3 %
SPIE	81	87	93,1 %
#N/A	61	210	29 %
Scientific.net	23	26	88,5 %
Cambridge University Press	19	150	12,7 %
Taylor & Francis Group	19	159	11,9 %
Optical Society of America	18	37	48,6 %
Gordon and Breach Science Publishers	14	20	70 %
EDP Sciences	12	254	4,7 %
Wiley-Blackwell	11	1924	0,6 %
American Mathematical Society	10	10	100 %
Yokohama Publishers	10	10	100 %

Exemple : La part de l'Open Acces dans les publications

BIU Sciences de Montpellier 2



Infométrie
 disciplines scientifiques
 Des bases de données scientifiques
 Les données de l'Observatoire des sciences et des techniques
 Définitions
 Analyse des publications
 réseaux
 Ucinet
 Gephi
 bibliométrie
 dates
 auteurs
 affiliations
 Excel
 Indicateurs de visibilité
 Etude des revues
 auxquel les s'abonner ou à promouvoir
 Qualifier et choisir les revues auxquelles s'abonner ou à promouvoir
 MySQL/Access
 Définitions
 auteurs
 réseaux
 Ucinet
 Ressources disponibles dans les institutions
 Répondre à un commanditaire :
 logiciels
 Représentation visuelle des données
 Cartes & Données
 Etude thématique
 Gephi
 Les revues
 dates
 Excel
 Analyser une thématique
 revues
 La bibliométrie, pour quoi faire ?
 logiciels
 orienter les actions
 Demarche exploratoire
 Quelles ressources pour quels résultats ?
 Datavrapper
 Nuage de mots
 revues
 Scientométrie

EXEMPLES DE RÉSULTATS : ÉTUDES THÉMATIQUES

Rencontres IST 2014 : L'IST en marche
 16 décembre 2014, Agropolis International, Montpellier



Études thématiques et les limites de l'analyse

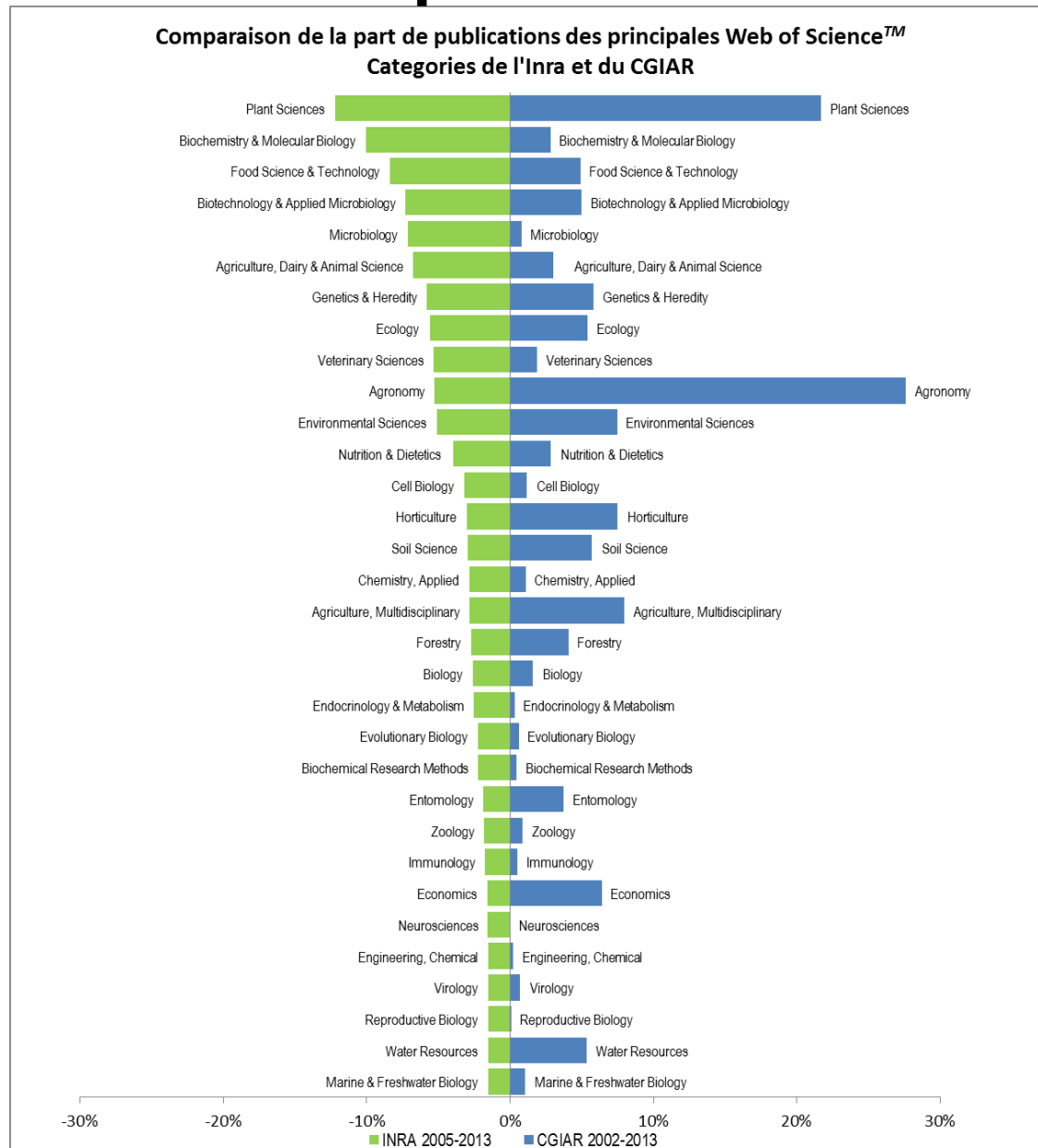
Classification thématiques des revues : *WoS Categories*,
Subject Areas de Scopus

Mots-clés auteurs / descripteurs (thésaurus)

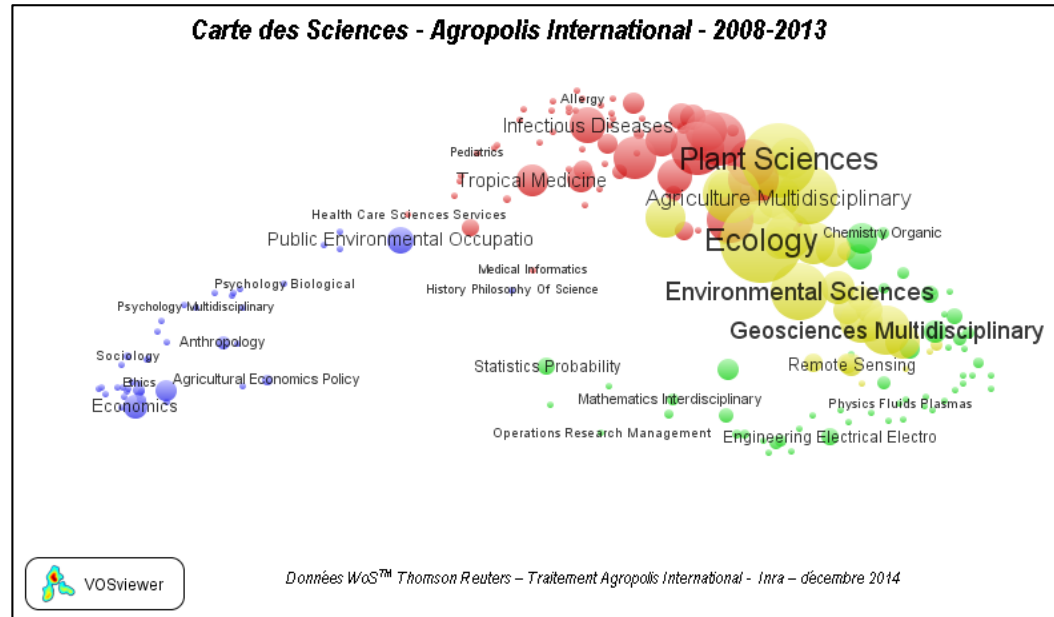
keywordPlus du Wos : risque d'apporter du bruit dans la recherche thématique

Analyse textuelle : limitée au titre / résumé / mots-clés.
Nécessite des outils de text-mining ou d'analyse sémantique et éventuellement le texte des publications

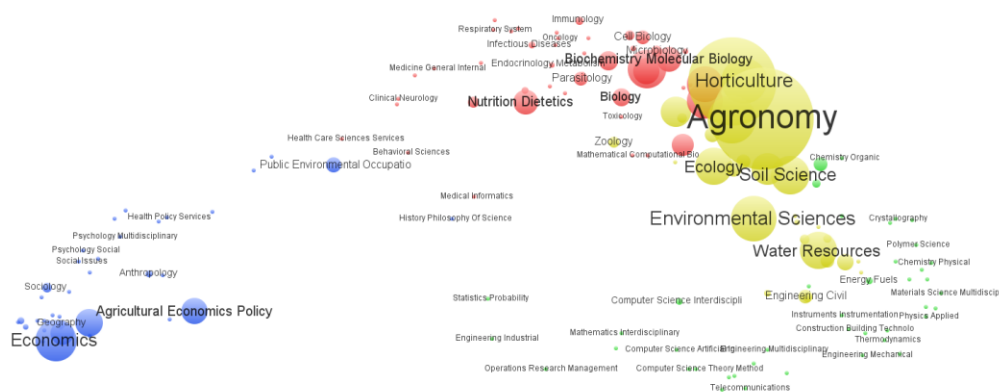
Comparaison de profils thématiques



Carte des sciences



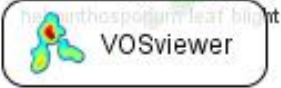
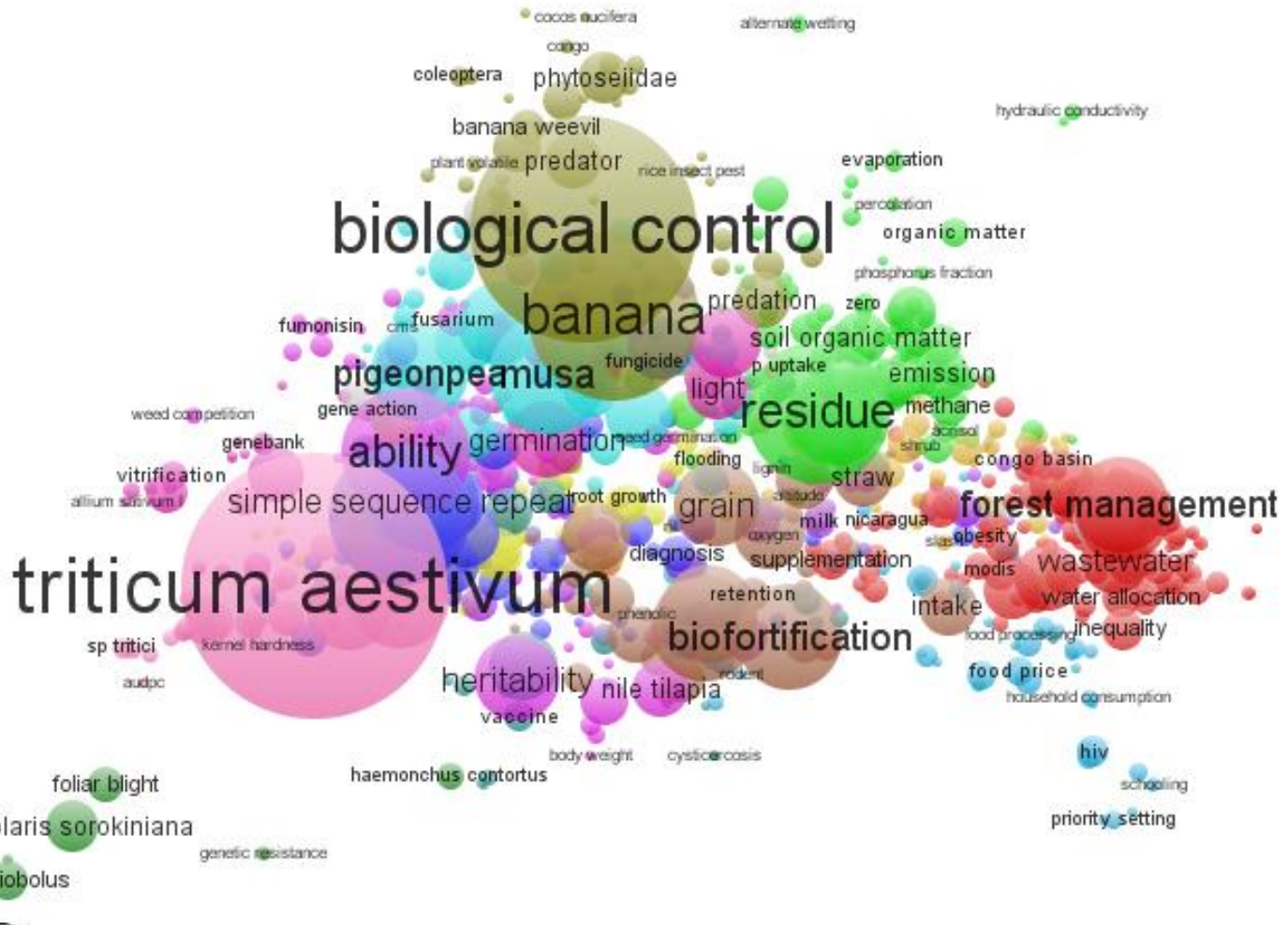
Carte des sciences - CGIAR 2003-2012



A user-friendly method for generating overlay maps (2012 update)

<http://www.leydesdorff.net/overlaytoolkit/>

Carte de co-occurrence de mots-clés



Données Web of Science™, traitement Inra et Agreenium 2013, CGIAR (2003-2012)

Notre boîte à outils

Logiciels	
MySql/Access	Requêtes spécifiques
Excel	tableaux croisés dynamiques, Histogrammes, radar, fonction RECHERCHEV,
Google Spreadsheets	Histogrammes, cartes géographiques
Le Sphinx	Histogrammes, analyse lexicale
Cartes & Données	Cartes géographiques
Wordle	Nuage de mots
Gephi	Réseaux (co-occurrences)
Ucinet	Réseaux (co-occurrences)
Vosviewer	Réseaux (text mining)
iMacro	Automatisation d'opérations
Datawrapper	Datavisualisation web
Tableau public	Datavisualisation web

Merci de votre attention

Des questions ?

Nicolas Barts, IRD

Odile Bédu, Inra - Agreenium

Nathalie Darbon, BIU Sciences, Montpellier 2

Dominique Fournier, Inra

Chantal Salson, Agropolis

Rencontres IST 2014 : L'IST en marche
16 décembre 2014, Agropolis International, Montpellier