

FAIRE PARLER LES DONNÉES SUR LA RECHERCHE GRÂCE AU WEB SÉMANTIQUE: LE PROJET VIVO

MARIE-DOMINIQUE HEUSSE, IRIT, UNIVERSITÉ DE TOULOUSE

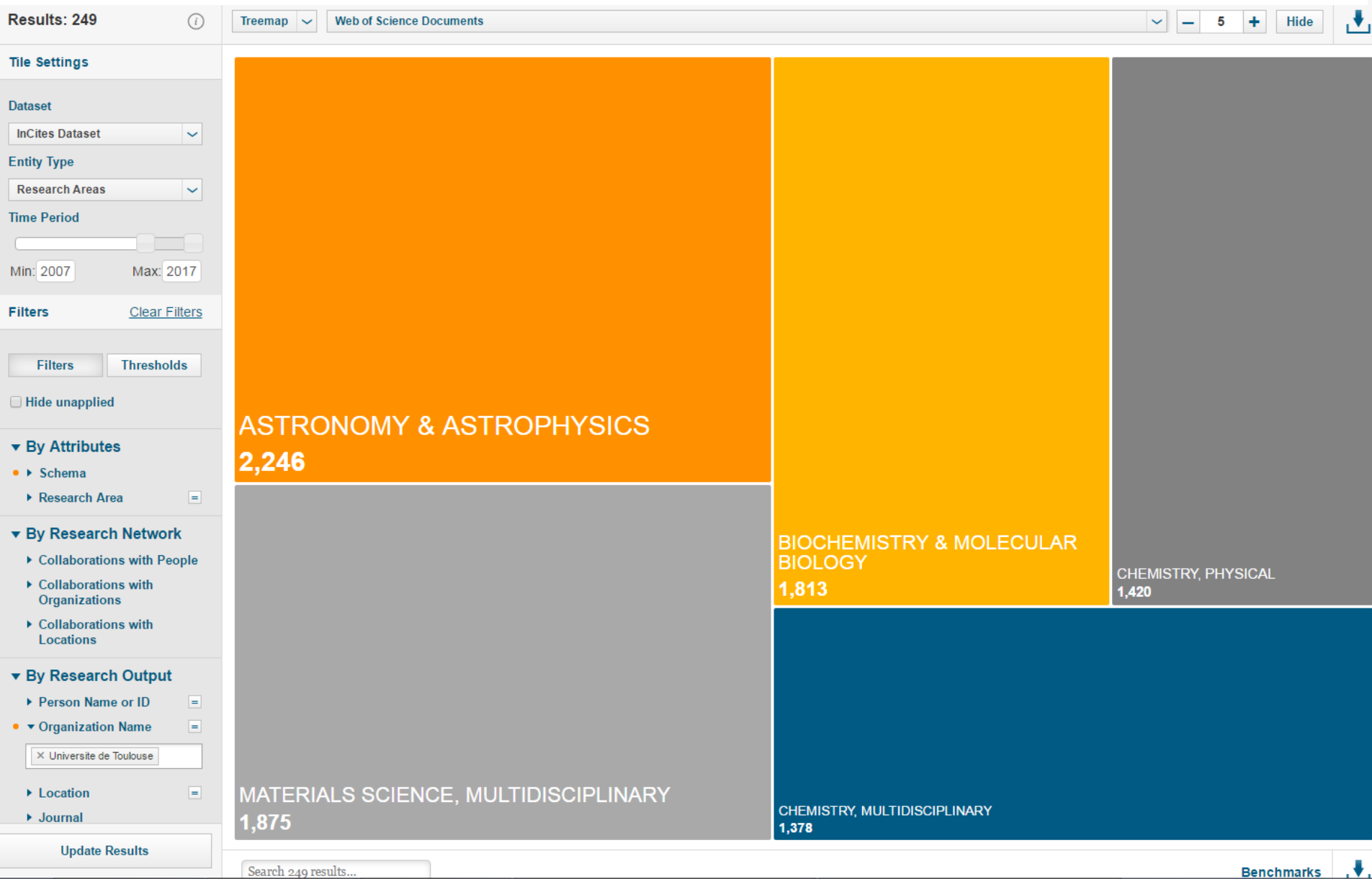
1er Atelier Valorisation et analyse des
données de la recherche (VADOR),
Toulouse, 2017

Comment caractériser l'activité de recherche du site Université de Toulouse?

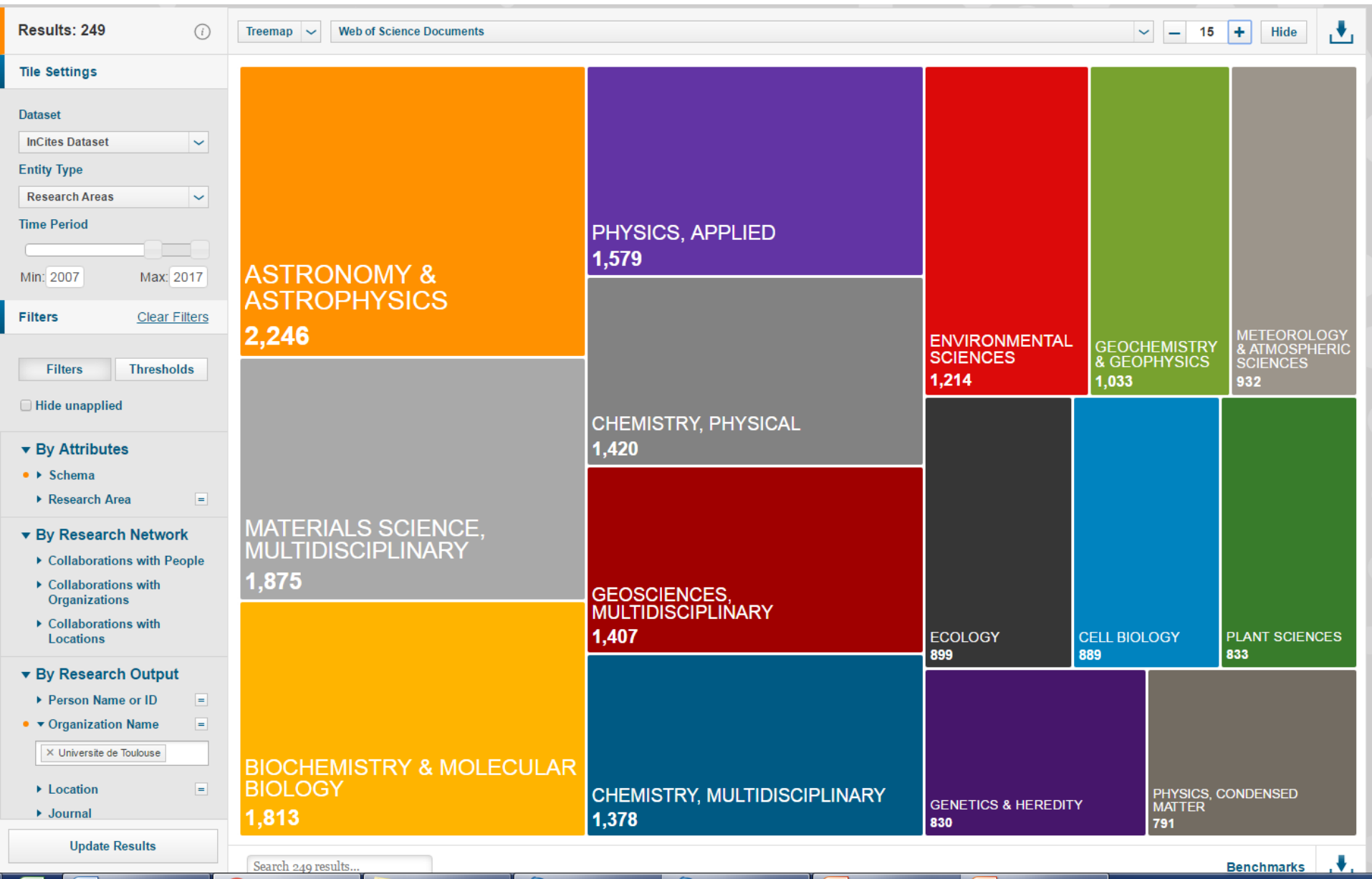
Des questions:

- Quelles sont les principales forces de recherche du site?
- Quel est le nombre annuel de publications des chercheurs de l'UT?
- Avec qui (et quels établissements français et étrangers) les chercheurs de l'UT co-signent?
- Combien y a-t-il de chercheurs par domaine disciplinaire?
- Comment repérer les axes stratégiques transversaux et l'interdisciplinarité?
- Les thématiques émergentes, les signaux faibles...

Chercher dans Incites?



Chercher dans Incites?



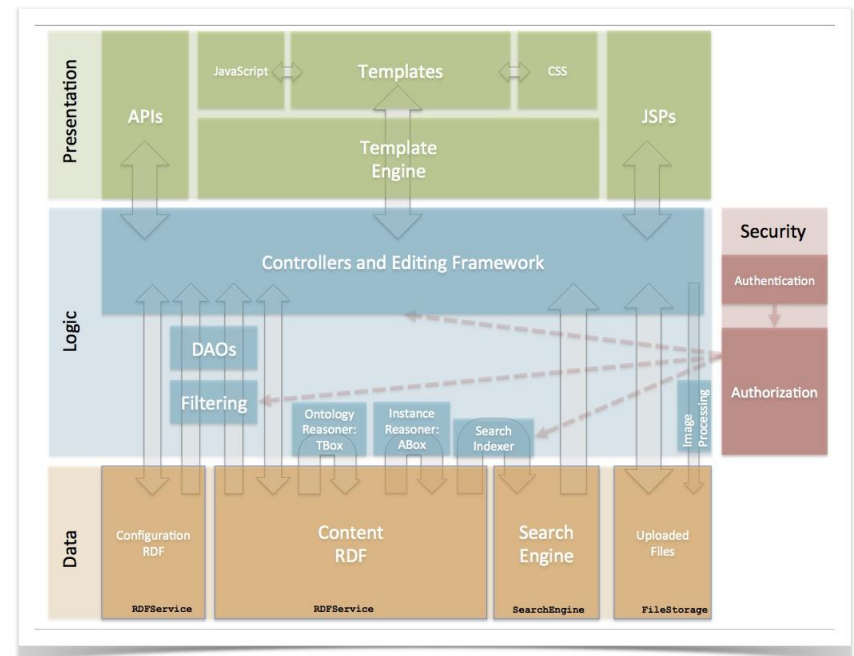
Le - double - pari du projet

- Rassembler les données de manière plus exhaustive et plus précise, par ex.
 - Liens chercheurs/labos, disciplines, pôles de coordination
 - Publications : articles RICL mais aussi les livres, en français, en archives ouvertes...
- Utiliser un outil qui permette de les exploiter pour atteindre les objectifs souhaités
 - Utiliser des sources externes sans refaire un entrepôt de données
 - Logique des données liées

VIVO

connect • share • discover

VIVO est développé par l'Université de Cornell, sous licence libre, et s'appuie sur les technologies sémantiques (RDF, SPARQL, OWL). VIVO est une application J2EE, intégrant SolR et une base de données RDF, et utilisant un langage de template (Freemarker) pour la création des écrans.



Les données sur la recherche

Les familles de données

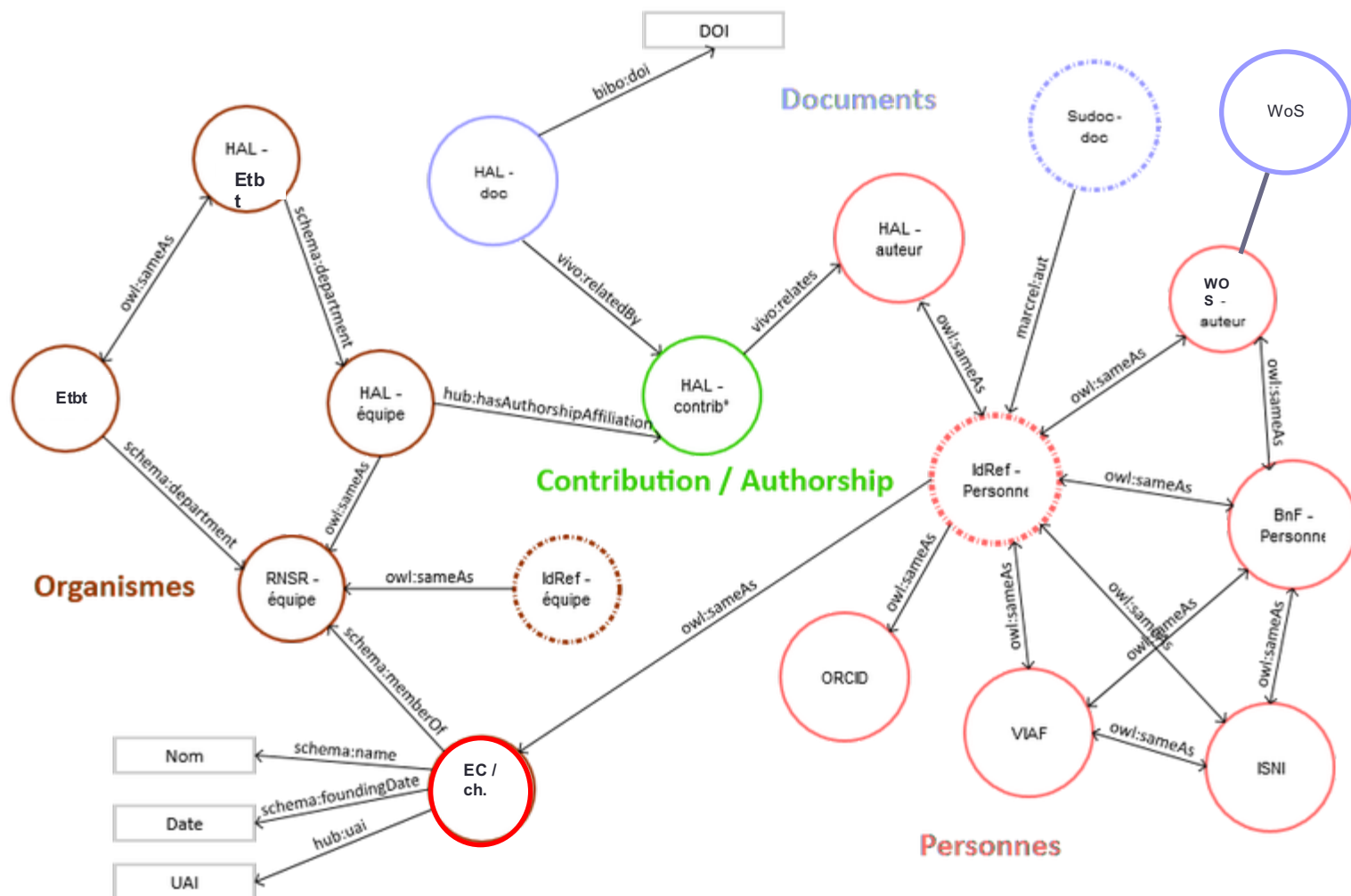
- **Les structures:** établissements, organismes de recherche, labos, équipes, écoles doctorales, pôles de coordination
- **Les auteurs:** nom, prénom, année de naissance, discipline, rattachement / structures
- **Les productions de la recherche:** publications, brevets, autres formes de production scientifique
- **Les moyens:** contrats, équipements

Les exigences

- **Qualité des données**
- **Identifiant pour chaque type de donnée**
- **Source de donnée valide**
- **Non redondance (si possible...)**

[illegible]

Schéma de données



Ontologies

VIVO connect • share • discover

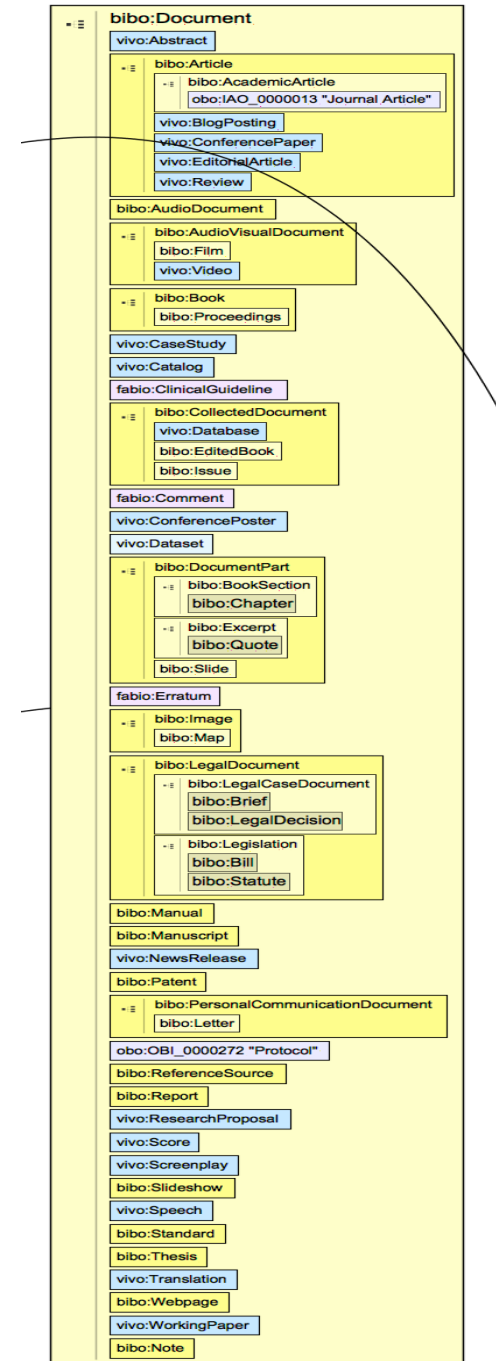
Index Site Admin Marie-Dominique ▼

Home People Organizations Research Events Capability Map

Ontologies

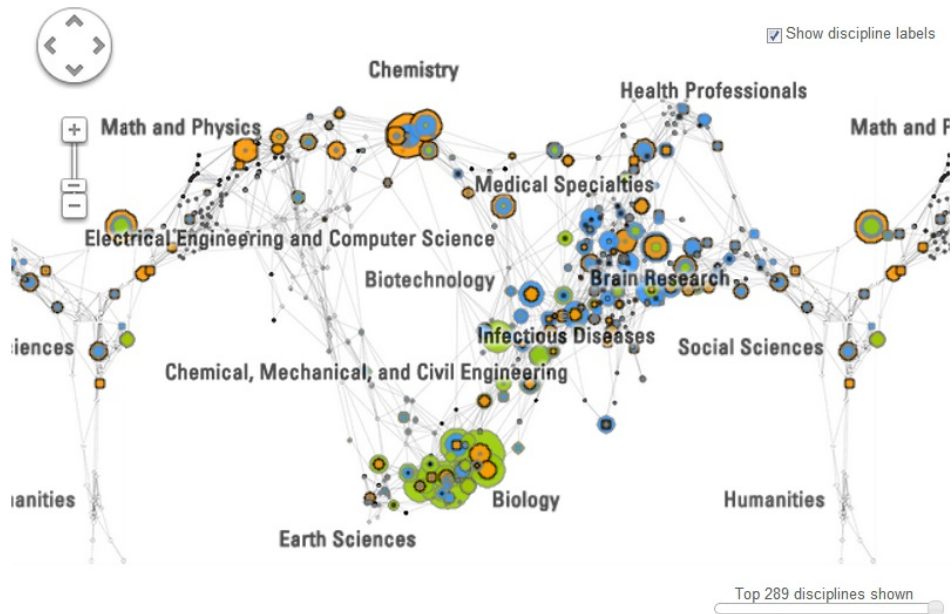
Add new ontology

#	Ontology	Namespace	Prefix
1	Bibontology	http://purl.org/ontology/bibo/	bibo
2	Citation Counting and Context Characterization Ontology	http://purl.org/spar/c4o/	c4o
3	CiTO (Citation Typing Ontology)	http://purl.org/spar/cito/	cito
4	Event Ontology	http://purl.org/NET/c4dm/event.owl#	event
5	FaBio (FRBR-Aligned Bibliographic Ontology)	http://purl.org/spar/fabio/	fabio
6	FOAF (Friend of a Friend)	http://xmlns.com/foaf/0.1/	foaf
7	Geopolitical Ontology	http://aims.fao.org/aos/geopolitical.owl#	geo
8	OBO Foundry	http://purl.obolibrary.org/obo/	obo
9	OCRe research	http://purl.org/net/OCRe/research.owl#	ocrer
10	OCRe study design	http://purl.org/net/OCRe/study_design.owl#	ocresd
11	SKOS (Simple Knowledge Organization System)	http://www.w3.org/2004/02/skos/core#	skos
12	VCard	http://www.w3.org/2006/vcard/ns#	vcard
13	Vitro Public Ontology	http://vitro.mannlib.cornell.edu/ns/vitro/public#	vitro-public
14	VIVO Core	http://vivoweb.org/ontology/core#	vivo
15	VIVO Scientific Research Ontology	http://vivoweb.org/ontology/scientific-research#	scires



Cartographies

Réseau de co-auteurs



Chabrol Henri

Co-author Network ([GraphML file](#))

Profile

Chabrol Henri

[VVO profile](#)

32 Publication(s)

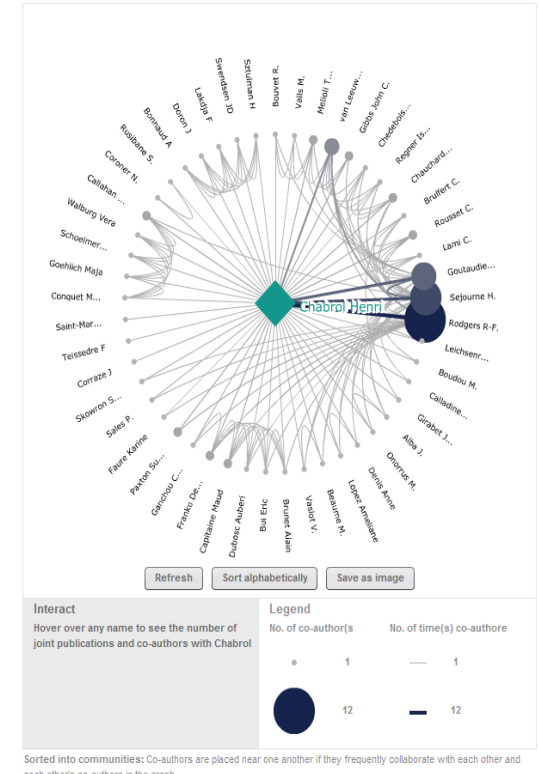
49 Co-author(s)

2000 First Publication

2014 Last Publication

Note: This information is based solely on publications that have been loaded into the VIVO system. This may only be a small sample of the person's total work.

Log in to enter additional details about your publications on your profile page.



Map of Science

Conclusion (provisoire)

- L'accueil du projet
 - Vif intérêt des établissements
 - Adhésion des chercheurs consultés
- Les difficultés du projet
 - Besoin de compétences diverses: informatique, données
 - Techniques et formats encore peu répandus
 - Long travail d'amélioration de la qualité des données
- Les atouts du projet
 - Occasion de fédérer des savoir-faire et des expertises
 - En faire un démonstrateur de l'importance des données et de leur qualité

Merci de votre attention

marie-
dominique.heusse@univ-toulouse.fr

12 October 2016

12 October 2016

