

#OSSPARIS16



PARIS
OPEN
SOURCE
SUMMIT

ÉDITION 2016 | 16&17 NOVEMBRE

UN ÉVÉNEMENT



Vers une infrastructure pour l'open science

Odile Hologne, INRA,



@Holo_08

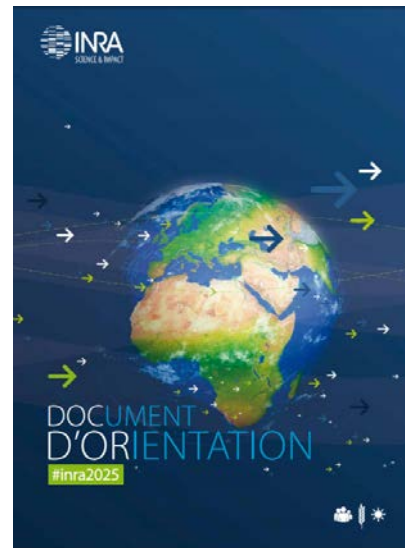


#OpenScience : une des 3 orientations de politique générale de l'INRA

Une science ouverte, numérique, collaborative pour répondre aux enjeux sociétaux

< #OpenScience >

Une science ouverte grâce au numérique



#OpenScience-1 : Des infrastructures de recherche connectées

#OpenScience-2 : Une organisation des données pour le partage et la réutilisation

#OpenScience-3 : Des approches prédictives en biologie

#OpenScience-4 : De nouveaux modes de diffusion de la connaissance

#OpenScience-5 : Le métier et l'environnement du chercheur adaptés au numérique

Charte du libre accès aux publications et aux données



- Données : Ouvertes par défaut sous réserve de droits
- Publications : ne pas céder les droits à l'éditeur, lien avec l'évaluation
- FAIR principes :
 - findable
 - Accessible
 - Interoperable
 - Re-usable

Le chercheur et « l'open »

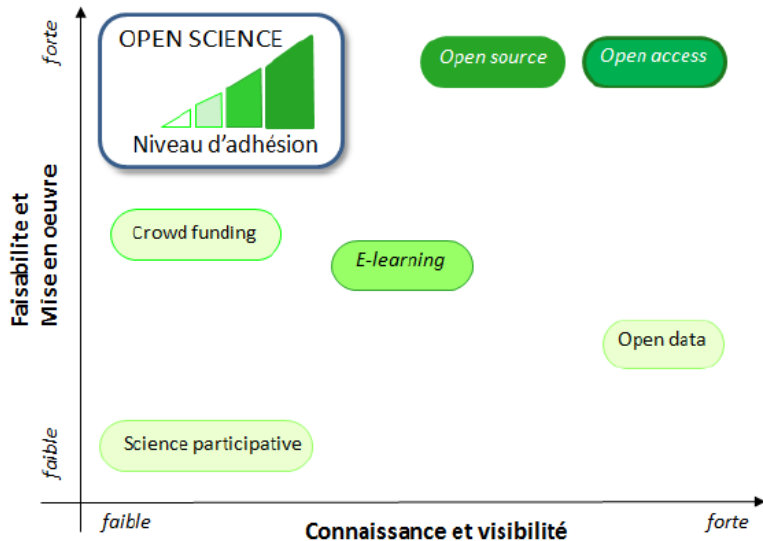


Fig. 5: Représentation du niveau de connaissance et de visibilité, de mise en œuvre et d'acceptabilité des différents modalités de l'Open au niveau du chercheur.

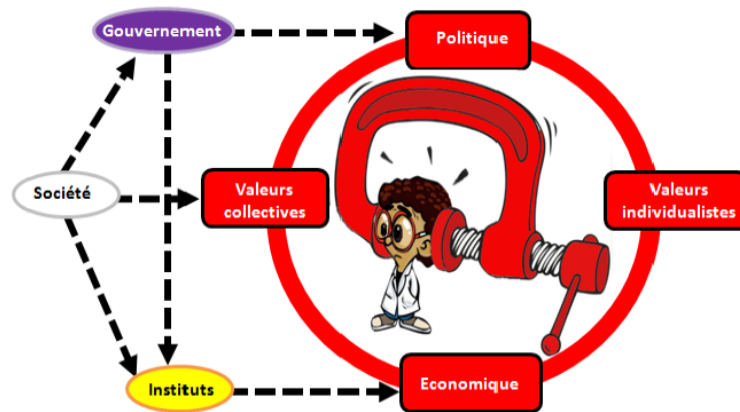


Fig. 6: Illustration des considérations personnelles et externes s'exerçant sur le chercheur concernant la démarche Open

Rapport EPMRA INRA 2014
(étude interne non diffusée)

Infrastructure pour les publi : <http://prodinra.inra.fr>

- Choix d'un logiciel propriétaire par la DSI dans une optique de développement de services multiples autour de la gestion de documents sur ce socle technique
- Interopérabilité in/out avec différents formats d'exposition des données et protocoles
- Aujourd'hui :
 - faible évolutivité de la solution et « cul de sac technologique »
 - 2 scénarii : HAL ou open source (Dspace, Islandora ...)

Infrastructure pour les données : Dataverse



Open source research data repository software

- Création d'un portail des données INRA (disponible en 2017)
- Basé sur <http://dataverse.org/>, open source développé par Harvard Univ.



Researchers

Enjoy full control over your data. Receive web visibility, academic credit, and increased citation counts. A personal dataverse is easy to set up, allows you to display your data on your personal website, can be branded uniquely as your research program, makes your data more discoverable to the research community, and satisfies data management plans. [Want to set up your personal dataverse?](#)



Journals

Seamlessly manage the submission, review, and publication of data associated with published articles. Establish an unbreakable link between articles in your journal and associated data. Participate in the open data movement by using Dataverse as part of your journal data policy or list of repository recommendations. [Want to find out more about journal dataverses?](#)



Institutions

Establish a research data management solution for your community. Federate with a growing list of Dataverse repositories worldwide for increased discoverability of your community's data. Participate in the drive to set norms for sharing, preserving, citing, exploring, and analyzing research data. [Want to install a Dataverse repository?](#)



Developers

Participate in a vibrant and growing community that is helping to drive the norms for sharing, preserving, citing, exploring, and analyzing research data. Contribute code extensions, documentation, testing, and/or standards. Integrate research analysis, visualization and exploration tools, or other research and data archival systems with Dataverse. [Want to contribute?](#)

Maquette Dataverse INRA



INRA Dataverse (INRA) Unpublished

INRA dataverse tagline to be validated as soon as possible.

Harvard Dataverse > INRA Dataverse

Search this dataverse...

Dataverses (3)

Datasets (1)

Files (1)

Publication Status

Unpublished (4)

Draft (1)

Dataverse Category

Organization or Institution (3)

Subject

Medicine, Health and Life Sciences (1)

Keyword Term

Transposable elements (1)

consensus sequence (1)

fasta file (1)

1 to 4 of 4 Results

Transposable element consensus library generated from Blumeria graminis hordei genome assembly Draft Unpublished

Nov 9, 2015 - INRA Omic Dataverse

Amselem, Joelle; Quesneville, hadi; Oberhaensli, Simone, 2015, "Transposable element consensus library generated from Blumeria graminis hordei genome assembly", <http://dx.doi.org/10.7910/DVN/BFFBNK>, Harvard Dataverse, DRAFT VERSION

Transposable elements, fasta file, consensus sequence

INRA Social Sciences Dataverse (INRA) Unpublished

Nov 9, 2015

INRA datasets in social Sciences

INRA Ecology Dataverse (INRA) Unpublished

Nov 9, 2015

INRA datasets in Ecology

D'autres projets

- Forge INRA pour partager le code
- Portail des vocabulaires (nécessaires pour l'interopérabilité) <http://lovinra.agroportal.lirmm.fr/>
 - Basé sur bioportal (open source du NCBI)

The screenshot shows the AgroPortal website interface. At the top, there is a green navigation bar with the AgroPortal logo and menu items: Browse, Search, Mappings, Recommender, Annotator, and Projects. On the right side of the bar are links for Sign In, Help, and Feedback. Below the navigation bar, a yellow banner states: "You are viewing the INRA Linked Open Vocabularies AgroPortalSlice at lovindra.agroportal.lirmm.fr. To see all of the ontologies in AgroPortal visit the full site." Below this, a paragraph explains the portal's purpose: "Use AgroPortal to access and share ontologies. You can create ontology-based annotations for your own text, link your own project that uses ontologies to the description of those ontologies, find and create relations between terms in different ontologies, review and comment on ontologies and their components as you browse them, sign in to AgroPortal to submit a new ontology or ontology-based project, provide comments on ontologies or add ontology mappings." To the right of this text is a "Login" button and a note about the "Current Release: NCBO Appliance 2.4" with a link to "Issue tracking on GitHub".

The main content area is divided into several sections:

- Search all ontologies:** A search box with the placeholder "Enter concept, e.g. Melanoma" and a green "Search" button. Below the search box is a link for "Advanced Search".
- Find an ontology:** A search box with the placeholder "Enter ontology name, e.g. NCI Thesaurus" and a green "Explore" button.
- Search resources:** A search box with the placeholder "Enter a concept, e.g. Melanoma" and a green "Search" button. Below the search box is a link for "Advanced Resource Search".
- Ontology Visits in full AgroPortal (October 2016):** A table listing ontology visits:

| | |
|--|----|
| Biorefinery (BIREFINERY) | 23 |
| Animal Disease Ontology (ADO) | 16 |
| Ontobiotope (ONTOBIOTOPE) | 15 |
| Multi-scale Multi-step ontology (MS2O) | 10 |
| Wheat Trait Ontology (WHEATPHENOTYPE) | 8 |
| More | |
- Latest Notes:** A section titled "Can measure be mapped to another ontology? (biorefinery)" with a sub-note "11 months ago by jonquet" and "Such as Unit of Measurement?".
- Latest Mappings:** A list of ontology mappings:
 - Season (CO_715) <=> Season (http://dbpedia.org) External Mapping 01/08/2015 by jonquet
 - Season (CO_715) <=> season (http://data.bioontology.org/ontologies/CCON) Interportal Mapping ncbo 01/08/2016 by jonquet
 - plant organ (PO) <=> Plant organ (http://dbpedia.org/ontology/) External Mapping 11/30/2015 by jonquet
 - Plant (STY) <=> Plant (http://dbpedia.org/ontology) External Mapping 11/05/2015 by jonquet
 - tissue (BT) <=> tissue (CI) REST Mapping 06/24/2015 by jonquet
- Statistics in full AgroPortal:** A table showing overall statistics:

| | |
|------------|-----------|
| Ontologies | 8 |
| Classes | 1,165,672 |
| Projects | 14 |
| Users | 49 |

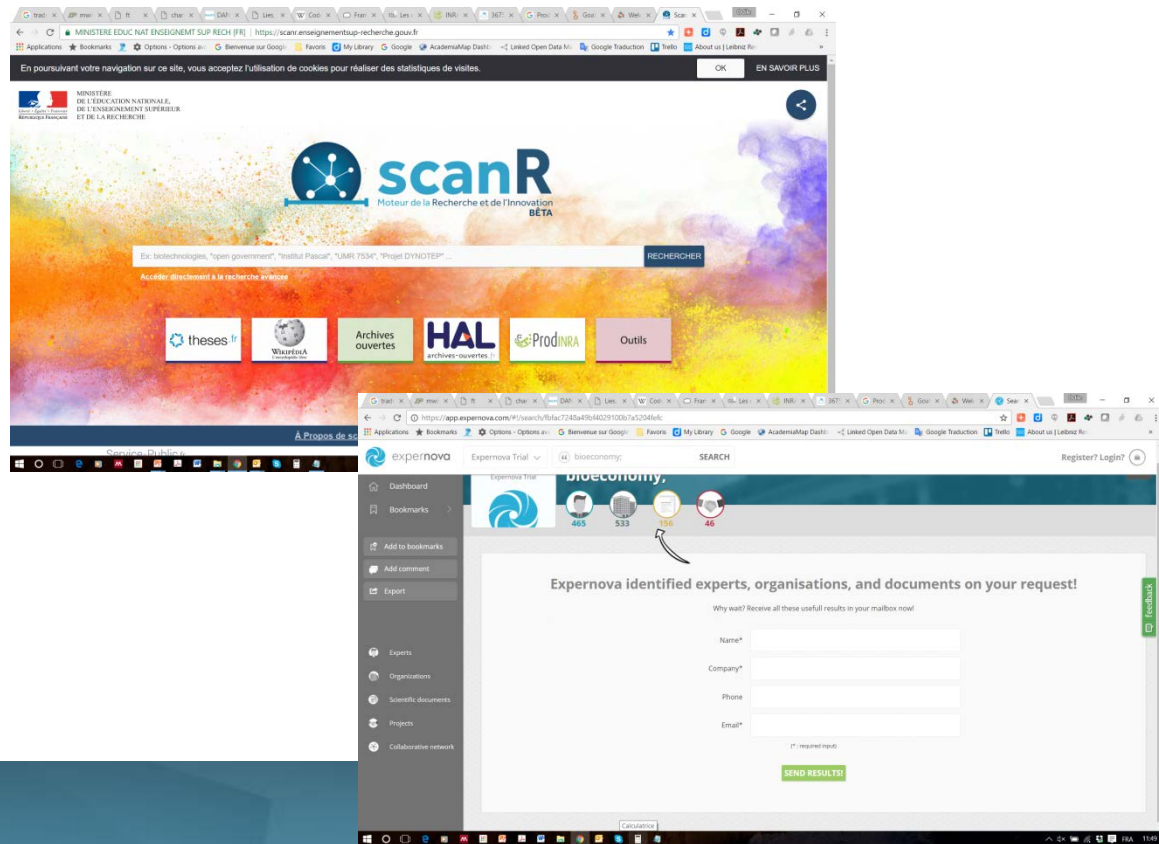
« Open data, open access, open code » penser à d'autres façons de créer de la valeur

- Des peurs : qui va cueillir les fruits (de l'ouverture)
 - Crainte d'une science à 2 vitesses entre les producteurs de données et les consommateurs



Des exemples de réutilisation sur les archives ouvertes

- Des services gratuits
- Des services payants



Conclusion

- A l'INRA : un cadre politique pour l'open science
- Plusieurs projets pour développer les composants d'une «e-infrastructure» basés sur des outils open source
- Un accompagnement des scientifiques <http://datapartage.inra.fr>
- Imaginer de nouvelles formes d'innovation et de création de la valeur
- Perspective : projets européens « e- ROSA » e- infrastructure roadmap for open science in agriculture et « AglInfra+ » prototypage de services innovants. Début 2017