

Dariah

Une infrastructure européenne pour les sciences humaines et sociales

www.dariah.eu et www.dariah.fr

Sophie David (CNRS, Huma-Num)
Nicolas Larrousse (CNRS, Huma-Num)

Plan

1. Qu'est-ce que DARIAH?

Objectifs

Gouvernance

Réseaux

2. Deux exemples de collaboration

Un modèle pour la visibilité des contributions

TADIRAH, décrire les activités de recherche pour les HN

3. Contribuer à DARIAH

4. DARIAH et le programme Horizon 2020

Qu'est-ce que DARIAH ?

Digital Research Infrastructure for the Arts and Humanities

Une infrastructure distribuée pour développer et soutenir la recherche et la formation dans toutes les disciplines des sciences humaines et sociales ...

... objets numériques privilégiés : textes, sons, images, vidéos ...

... un réseau d'expertises, d'outils, de services, d'acteurs ...

... pour les chercheurs, par les chercheurs ...

Des questions...

- **Accès à des quantités (inédites) de données**
Intégrité - Origine et contexte de production -
Préservation sur le long terme ?
- **Formats standardisés et interopérabilité ?**
- **Respect de la vie privée ?**
- **Propriété intellectuelle ?**
- **Formations ?**
- **Etc.**

- **Utilisation et la réutilisation de données**
- **Contexte technologique toujours en mouvement**

Des éléments de réponse...

- La promotion du libre accès pour les matériaux de la recherche (données, publications, etc.)
- La certification des entrepôts
- La fourniture d'identifiants pérennes
- Le développement de l'archivage à long terme
- Des recommandations (formats, bonnes pratiques)
- La promotion de la décentralisation

Bref historique

- **2006** : DARIAH inclus dans la feuille de route ESFRI



- **2008 – 2011** : Phase préparatoire
Preparing DARIAH (établir le projet scientifique et les statuts)



- **2011 – 2014** : Phase de transition
(établir DARIAH ERIC en intégrant les activités nationales)



- **2014+** : DARIAH-ERIC



Etablir une organisation internationale de recherche

DARIAH = un ERIC (*European Research Infrastructure Consortium*)

- * Pays
- * Assemblée générale

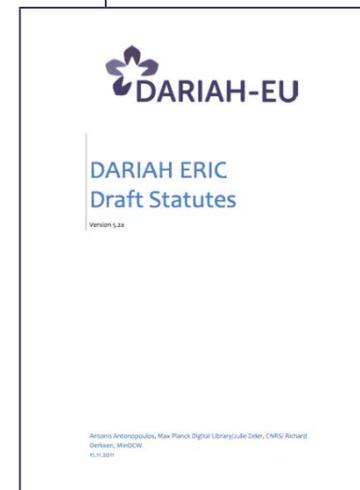
Intérêt : faciliter l'établissement d'infrastructures sur le long terme

ERIC DARIAH = 20 ans

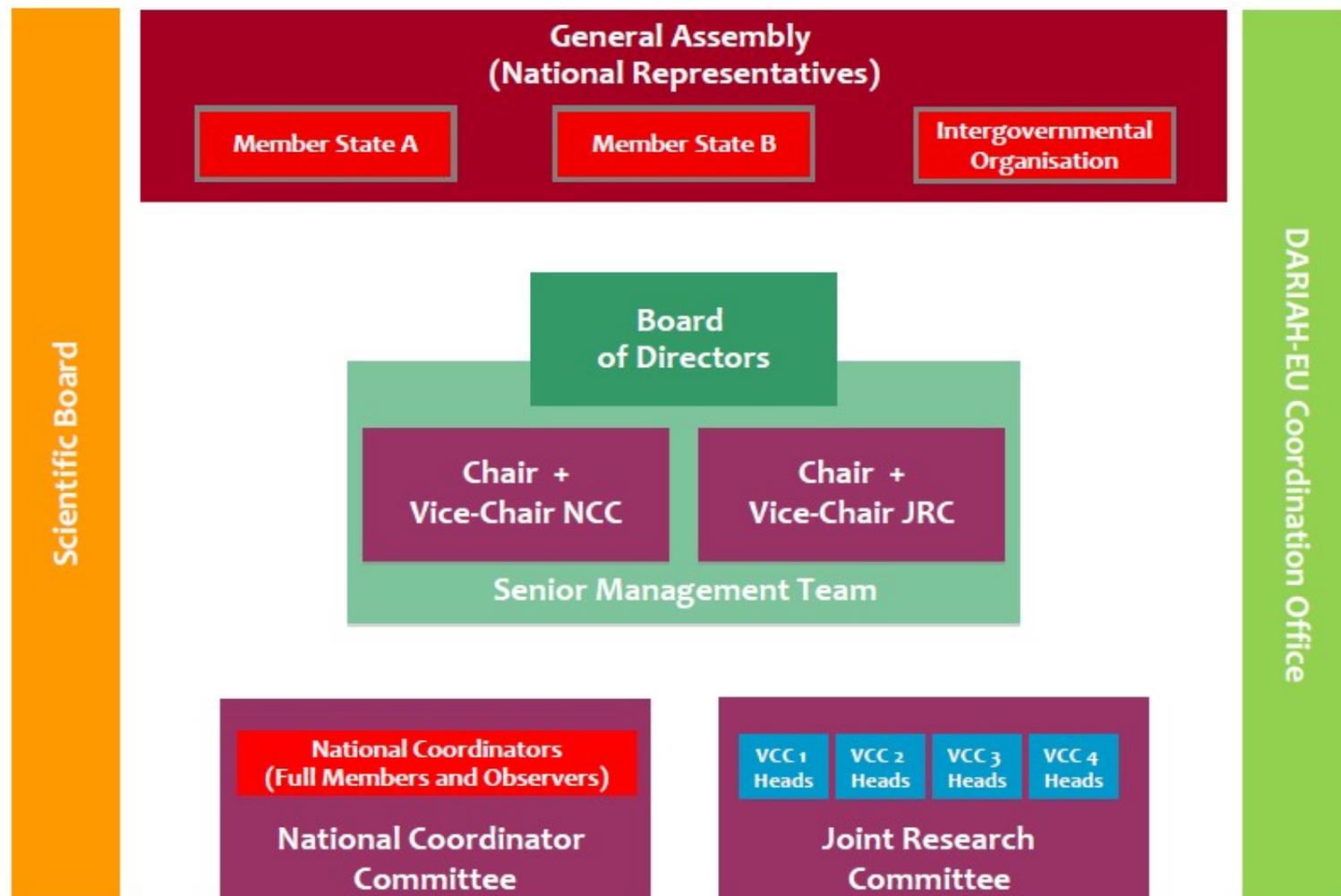


**Legal framework
for a European
Research
Infrastructure
Consortium - ERIC**

Practical Guidelines



La gouvernance



Le modèle budgétaire

- 4 000 000 euros / année
environ 10 % liquide
environ 90 % en nature
- La participation des pays est calculée en fonction de leur PIB
- **Liquide** : gestion de l'ERIC DARIAH
- **Nature** : activités existantes

DARIAH-EU Financial Statement



Report to October 2012

	Budget	Actual	Forecast
	€	€	€
Income			
Cash contributions	0	0	0
Total income	0	0	0
Expenditure			
	Budget	Actual	Forecast
Personnel costs			
Staffing	0	0	0
Material costs			
IT related costs	0	0	0
External Consultancy	0	0	0
Marketing & Communications	0	0	0
Travel	0	0	0
Other	0	0	0
Overhead	0	0	0
Total Expenditure	0	0	0

La participation de la France

- **Accueil de l'ERIC - Siège social à Paris**
- **Participation au Dariah Coordination Office**
- **Rôle d'Huma-Num**
Coordination de l'ensemble de la participation française
- **Participation à tous les groupes de travail**
 - « e-infrastructure, « recherche et formation », « gestion des contenus »
 - co-responsabilité avec DANS (Pays-Bas) du groupe de travail « gestion des contenus »

Un réseau de pays

Membres fondateurs		Membres fondateurs		Membres fondateurs	
Allemagne		France		Malte	
Autriche		Grèce		Pays-Bas	
Chypre		Irlande		Serbie	
Croatie		Italie		Slovénie	
Danemark		Luxembourg			

Un réseau de plateformes

Annonces d'événements →

Blogs scientifiques ↑

Des données ↓

Archive ouverte ↑

it en ▶ Il progetto ▶ Documentazione tecnica ▶ Mappa del sito

Un réseau pour les formations

GCDH GOTTINGEN CENTRE FOR
DIGITAL HUMANITIES

GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

RESEARCH TEACHING SERVICES EVENTS PROJECTS BLOG ABOUT PEOPLE PUBLICATIONS



2014 DARIAH INTERNATIONAL
DH SUMMER SCHOOL
17 - 30 August 2014

[READ MORE](#)

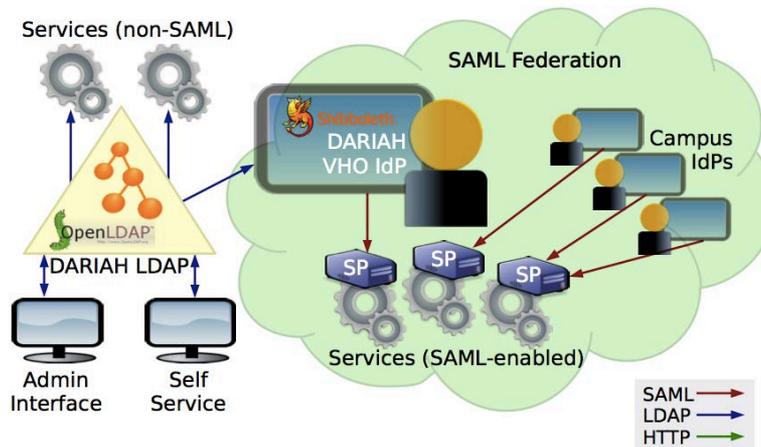
Universités d'été

Matériels
pédagogiques
multilingues



Un réseau technique

Infrastructure d'authentification ↓



The screenshot shows the DANS website header with the title 'Data Archiving and Networked Services' and the 'DANS' logo. The navigation menu includes: HOME, DATA ARCHIVE, PROJECTS, SERVICES, SYMPOSIA, PUBLICATIONS, NEWS ARCHIVE, CALENDAR, VACANCIES, ABOUT DANS, and CONTACT. The 'Citing data' section provides instructions on how to cite data sources correctly, listing required elements: Name of the creator of the dataset, Title of the dataset, Date created, DANS Data Archive, and The Persistent Identifier. An example citation is provided for a dataset from the Researchcentrum voor Onderwijs en Arbeidsmarkt - ROA, Universiteit Maastricht.

Identifiants pérennes ↑

The screenshot shows the 'Data Seal of Approval' website. A red seal logo is in the top left. The main content area features a photo of a person at a computer. Below the photo, text reads: 'Your online self-assessment will be peer-reviewed, after which you can earn the Data Seal of Approval and display the DSA logo.' and 'The application and peer review process is free of charge.'

Labellisation ← Data Seal of Approval

Un réseau pluridisciplinaire



Archéologie



CENDARI
COLLABORATIVE EUROPEAN DIGITAL
ARCHIVE INFRASTRUCTURE

Histoire médiévale
et contemporaine



EUROPEAN HOLOCAUST
RESEARCH INFRASTRUCTURE

Histoire de l'Holocauste



Digital Scholarly Editions
Initial Training Network

Bourses

dans le domaine
de l'édition de
textes



Ontologie sur les
méthodes (DH)



Visibilité des contributions

Objectifs

“Susciter de nouvelles recherches et améliorer leur résultat en reliant des ressources numériques de tous types”

Pour les contributeurs

Donner une meilleure visibilité aux contributions

Pour les participants

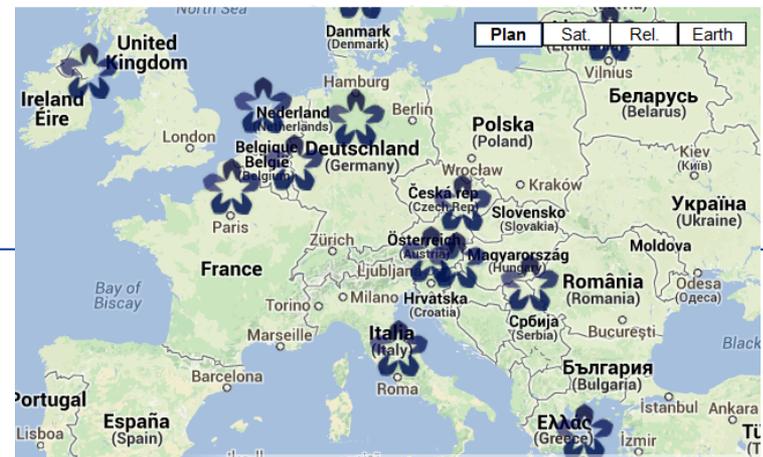
Donner des outils pour retrouver l'information



Quelques exemples de questions

- ⤴ Comment ouvrir un blog scientifique ?
- ⤴ Quels sont les experts en Open Archive ?
- ⤴ Quels sont les formats recommandés pour les images ?
- ⤴ Quelles sont les contributions hollandaises ?
- ⤴ Quels projets Européens sont associés à DARIAH ?
- ⤴ Dispose-t-on d'un corpus important de données accessibles librement (Open Access) en France ?

Principes généraux



- Les données sont gérées de manière décentralisée

Chaque contributeur est

responsable de la description de sa contribution

Chaque pays est responsable de rassembler et d'afficher ses contributions localement

- Utilisation d'outils standardisés

Le dispositif est basé sur les technologies du Web de données (RDF, SPARQL)

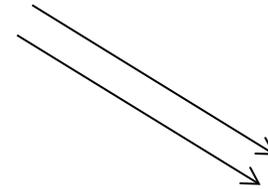
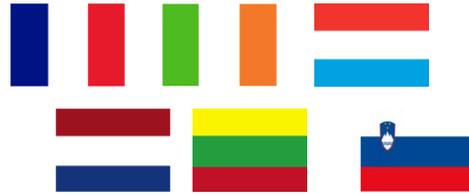
- Exploiter les possibilités du « Linked Open Data »

Réutiliser les données d'autres entrepôts pour les enrichir

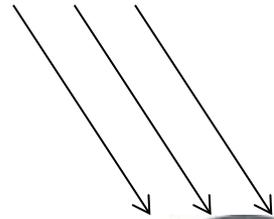
- Un investissement minimal en coût et en temps

Fonctionnement général

Descriptions des contributions
(RDFa + HTML)



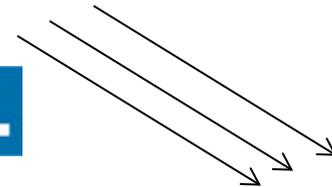
Sites Nationaux



Triple Store



Site



Autres produits

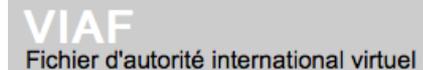
Qu'est-il possible de faire avec ces données ?

Deux possibilités principales

Interroger
les données

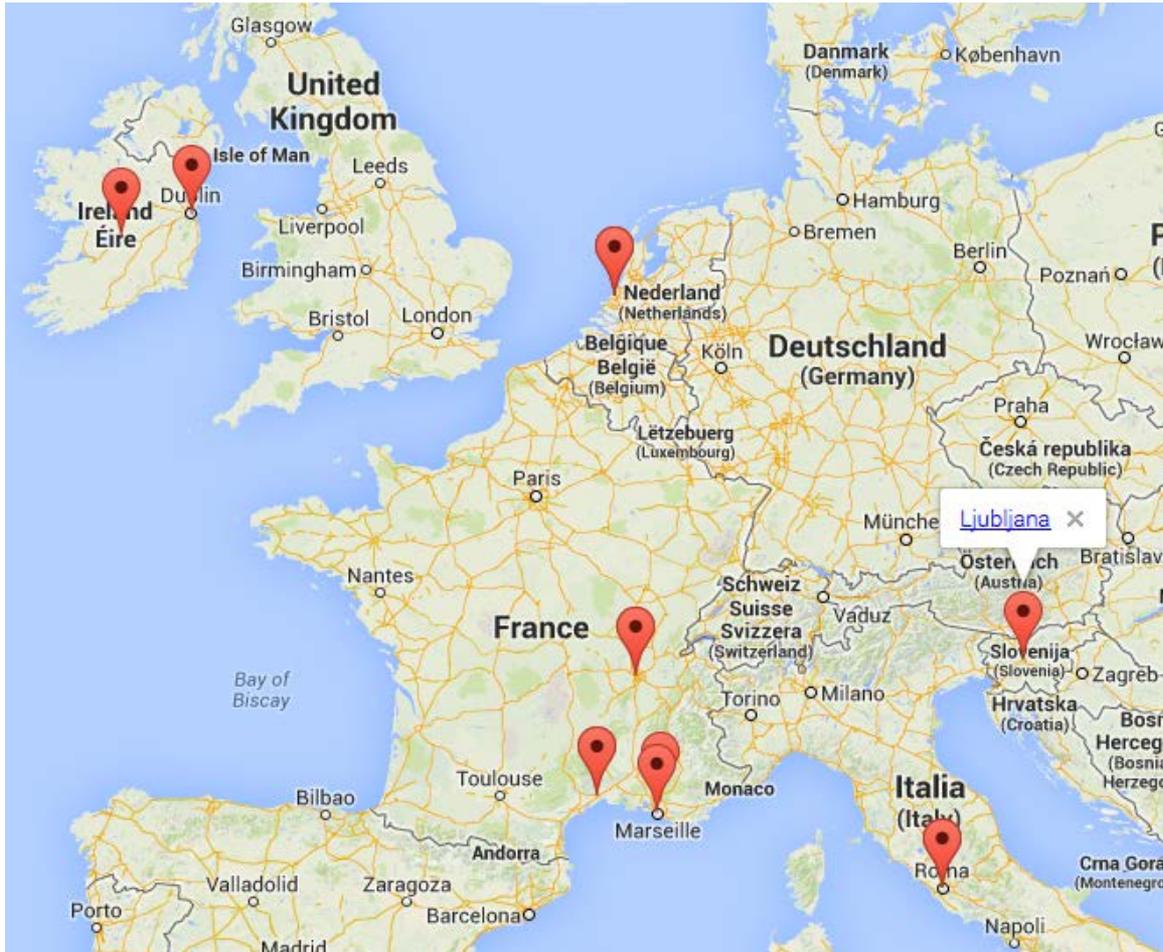


Enrichir les données
en utilisant d'autres
entrepôts



Enrichir les données

Enrichir les données avec GeoNames



On peut récupérer les données de localisation des contributions et

...

construire une carte

Enrichir les données

Depuis les données de GeoNames ... on peut accéder à d'autres entrepôts ... "follow your nose !"



On récupère depuis GeoNames, une entrée DpPedia, dont on affiche les données

... et ainsi de suite !!

TADIRAH

Comment décrire **les activités de recherche** liées aux humanités numériques avec une **structure commune** ?

1) Un découpage basé sur les différents stades du cycle de la recherche

2) Trois catégories

- **Activités**

- **Objets**

- **Techniques**

« ta-DEE-ra » basé sur les anagrammes DiRT and DARIAH

Définir une « taxonomie »

Basée sur le concept de « scholarly primitives »

Modélisation des étapes du processus de la recherche

Utilisation de l'existant

- arts-humanities.net
- Catégories définies par Bamboo / DiRT

Utiliser des outils collaboratifs (GIT)

dhtaxonomy / TaDiRAH

Taxonomy of Digital Research Activities in the Humanities

12 commits 1 branch 0 releases 3 contributors

branch: master TaDiRAH / +

Update activities.md ...

quinnanya authored on May 14 latest commit 6c7df7352b

reading	Reading and reference sections; naming scheme	12 days ago
reference	Update activities.md	12 days ago
introduction.md	Reading and reference sections; naming scheme	12 days ago
readme.md	Reading and reference sections; naming scheme	12 days ago
references.md	Reading and reference sections; naming scheme	12 days ago

readme.md

TaDiRAH - Taxonomy of Digital Research Activities in the Humanities (v. 0.5.1, 05/2014)

Taxonomie utilisée pour décrire les contributions

Contribuer à Dariah (1)

- **Les principes**
 - Ouverture
 - Volontariat
 - Transparence
 - Une « administration » la plus légère possible
- **Une contribution annuelle**

Contribuer à Dariah (2)

- Identifier son offre

Accès

Expertise

Interopérabilité

Accueil de contenus

Outils / logiciels

Formations

Evénements

Ecoles d'été

Collaboration

Ressources pédagogiques...

(voir www.dariah.fr)

- Rédiger un texte présentant la contribution
- Respecter des critères DARIAH (formats standardisés, gratuité, disponibilité du service, open source, etc.)
- Chiffrer sa contribution (personnels, matériels, traductions, etc.)

Contribuer à Dariah (3)

- Comité de Coordination

Aurélien Berra, Nadine Dardenne, Sophie David, Anne Joly, Nicolas Larrousse, Marie-Véronique Leroi, Jean-Luc Minel (Huma-Num, Univ. Paris 10, Ministère de la Culture et de la Communication)

- Circuit de validation

- Conseil scientifique de la TGIR Huma-Num
- Validation par le *Joint Research Committee*

- Convention entre Huma-Num et les partenaires français

DARIAH et le programme H2020



- Le programme H2020 comprend un volet e-infrastructures
- Pour les données de la recherche produites par des fonds publics, un accent fort est mis sur :
 - l'accès
 - la conservation
- Par exemple, obligation de définir un *Data Management Plan*



Data management plan (DMP)

- Description du jeu de données
- Standards du domaine
- Accès & Partage
- Archivage à long terme

DARIAH et le programme H2020



1) 2014 – Réponse coordonnée des partenaires de DARIAH à un projet de cluster (INFRADEV-4-2014/2015)

2) 2015 – Réponse de l'ERIC DARIAH à l'appel INFRADEV-3-2015

-> Intérêt de participer à DARIAH pour H2020

- 1) Etre associé aux réponses aux appels d'offres
- 2) Faire la preuve de son inscription dans l'écosystème

De quelques intérêts à participer à DARIAH

- Accroître la visibilité de ses recherches au niveau européen
- Accroître l'accès aux données de la recherche, aux outils et aux services dans le domaine des HN
- Bénéficier de ces services au-delà du temps de financement de son projet
- Accroître l'accès aux données sur le long terme
- Bénéficier de formations
- Augmenter ses chances de financement (e.g. H2020)
- Développer son réseau de collaborations