

Outils et circulation des données en Europe : État des lieux par le consortium euroCRIS

Joachim Schöpfel (Université de Lille)

euroCRIS Board member

Pablo de Castro (University of Strathclyde)

euroCRIS Technical Secretary

<http://hdl.handle.net/11366/2719>

Plan

- L'information sur la recherche
- Les systèmes d'information sur la recherche
- La diversité sur le terrain
- Les dispositifs nationaux
- Les enjeux
- L'intelligence artificielle
- Références

L'information sur la recherche

Le Comité pour la Science Ouverte définit l'information sur la recherche comme des « métadonnées relatives à la conduite et à la communication de la recherche ».

Cette information se trouve dispersée dans des bases de données hétérogènes (ressources humaines, finances, bibliothèques, archives institutionnelles etc.), souvent cloisonnées en silos aux formats et standards variés.

Il s'agit notamment

1. des métadonnées bibliographiques telles que les titres, les résumés, les références, les données sur les auteurs, les données sur leurs affiliations et les données sur les lieux de publication,
2. des métadonnées sur les logiciels de recherche, les données de la recherche, les échantillons et les instruments,
3. des informations sur les financements et subventions
4. et des informations sur les organisations et les acteurs de la recherche.

Les systèmes d'information sur la recherche

- Des plateformes conçues pour agréger et stocker des métadonnées liées aux chercheurs et leurs affiliations, aux publications, aux brevets, aux projets et subventions, aux distinctions, aux retombées médiatiques et aux indicateurs d'impact (OCLC).
- Un concept générique, englobant toutes les variantes, indépendamment du modèle de développement, de la licence ou de l'architecture.
- Des portails d'agrégation, des logiciels intermédiaires (*middleware*), des systèmes intégrés (*CRIS*)...
- Convergences avec les archives ouvertes (cf. de Castro et al., 2014)

Les deux points communs

Les données exploitées

- Informations bibliographiques (titres, auteurs, affiliations, lieux de publication),
- métadonnées sur les logiciels et les instruments scientifiques, financements et subventions, acteurs institutionnels...

Les objectifs poursuivis

- Faciliter la gestion des ressources,
- produire des indicateurs,
- normaliser les informations
- et garantir leur interopérabilité à l'échelle nationale et internationale.

Intérêt d'un format standard comme CERIF pour assurer l'interopérabilité des différents systèmes

La diversité sur le terrain

Convergences

- Une présence mondiale
- Un rôle commun
 - *Tous ces dispositifs ont pour mission essentielle la gestion de l'information sur la recherche*
- Une centralisation des données
 - *La plupart des SI recherche servent à rassembler et structurer les données institutionnelles sur les chercheurs et les projets*

Divergences

- Leur périmètre d'application
 - *Institutions, régions, nations, agences de financement...*
- Le type de logiciel utilisé
 - *Solutions commerciales, SI open source, développements internes...*
- L'architecture globale
 - *SI locaux, dispositifs mutualisés, plateformes d'agrégation...*

Les dispositifs nationaux

Architecture globale

- Un système central mutualisé
- Un système mutualisé avec des instances locales
- Un système d'agrégation (*hub*)

Gouvernance/pilotage

- Ministère de la recherche
 - Finlande
- Agence nationale de financement (FCT)
 - Portugal

research.fi

 Publications	879 390	 Funding calls	807	 Granted funding	17 258	 People	1 398
 Research data	7 356	 Infrastructures	224	 Organizations	86	 Projects - BETA	11 251

Home Search Science and Innovation Policy ▼ Science and research news In English ▼

Search for information on research in Finland

Search target ▼

For example, publication, field of science, keyword

 SEARCH

Search help ▼

Publications

Funding calls

Granted funding

Show more ▼

Science and research in Finland



Latest science and research news

Perusopetuksen rehtori, osallistu Koulun kompassi -tutkimukseen

Open funding calls

France-Nokia Chaire d'Excellence (ICV)

Hakuaika näättv 17.4.2025

GIVE FEEDBACK!

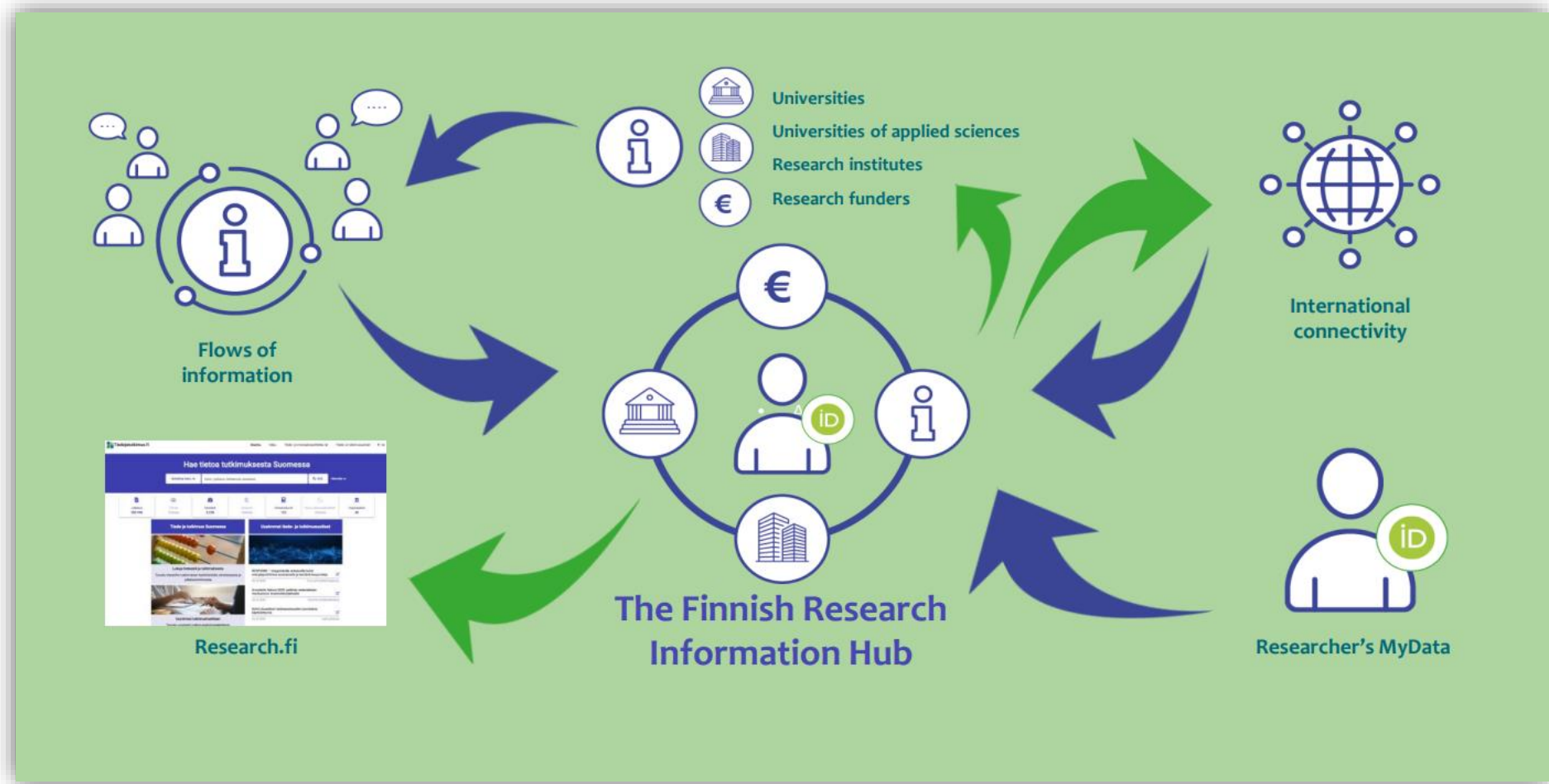
22 mai 2025

Journées d'études ABES 2025

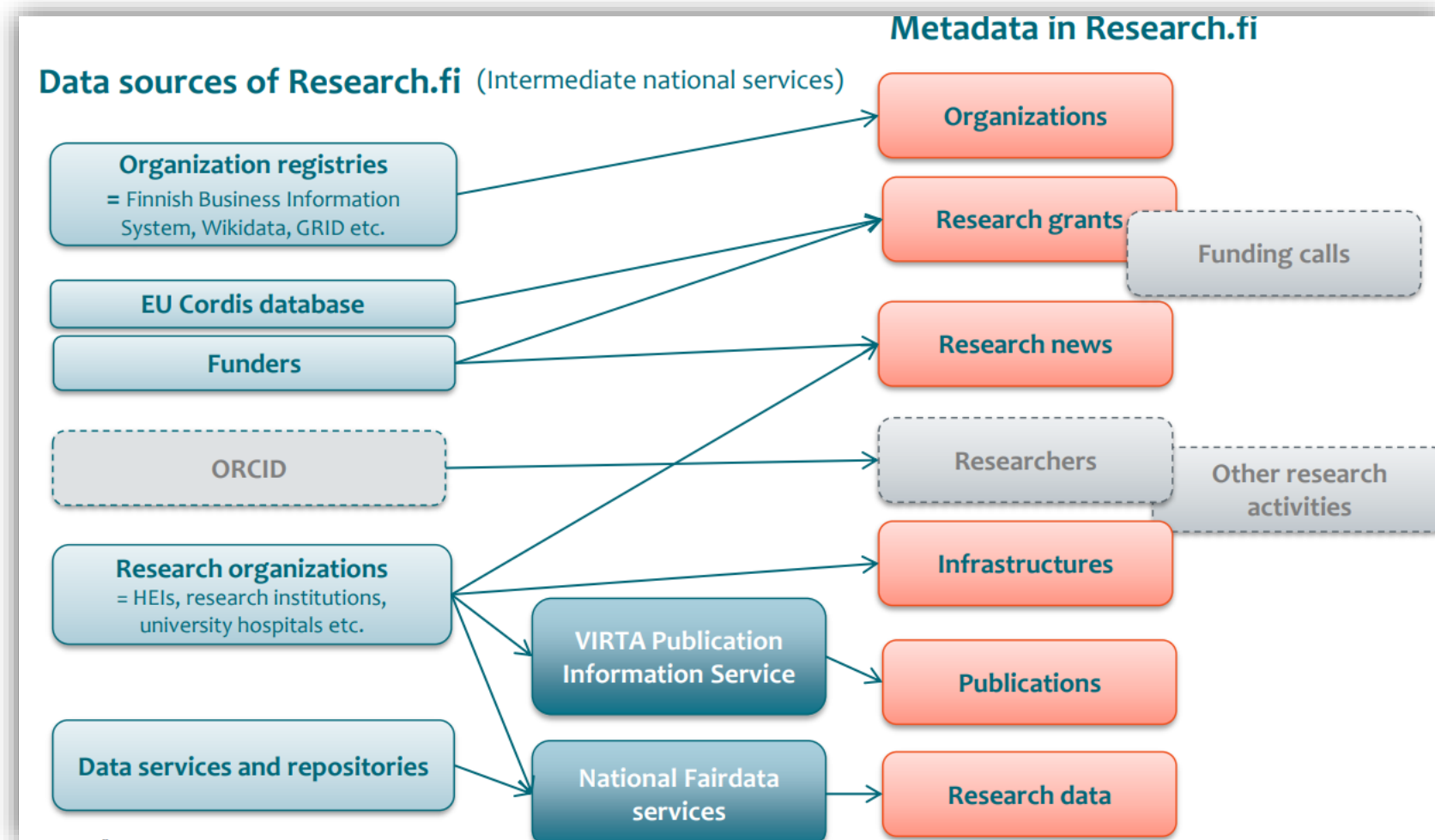
8

Les sources de l'information

- Le service met automatiquement à jour les informations déjà présentes dans les systèmes des établissements, organismes et agences de financement de la recherche.
 - Chaque organisation a son propre SI recherche (Pure, SoleCRIS, système maison, feuilles xls...).
- Les informations sont transférées selon des schémas XML convenus d'un commun accord et basés sur des modèles de données, des listes de codes et des concepts communs.
 - Pas de saisie manuelle dans le service, donc pas de charge administrative supplémentaire des chercheurs.
- L'exhaustivité et l'actualité des informations dépendent des organismes fournisseurs.
 - Chaque organisation est responsable de l'exactitude de ses propres informations.
- Le transfert d'information institutionnelle est obligatoire...
 - ...pour obtenir des subventions de l'Etat.
- Research.fi est un service du ministère de l'Éducation et de la Culture qui centralise et diffuse des informations sur la recherche en Finlande.
- Sa mise en œuvre est assurée par le CSC – Centre informatique pour la science, en partenariat avec les établissements d'enseignement supérieur, les instituts de recherche et les organismes de financement.
- Il s'agit de l'interface publique (portail) du National Research Information Hub.
- Le projet a débuté en 2017 et le service Research.fi a été lancé en juin 2020. Son développement est toujours en cours.

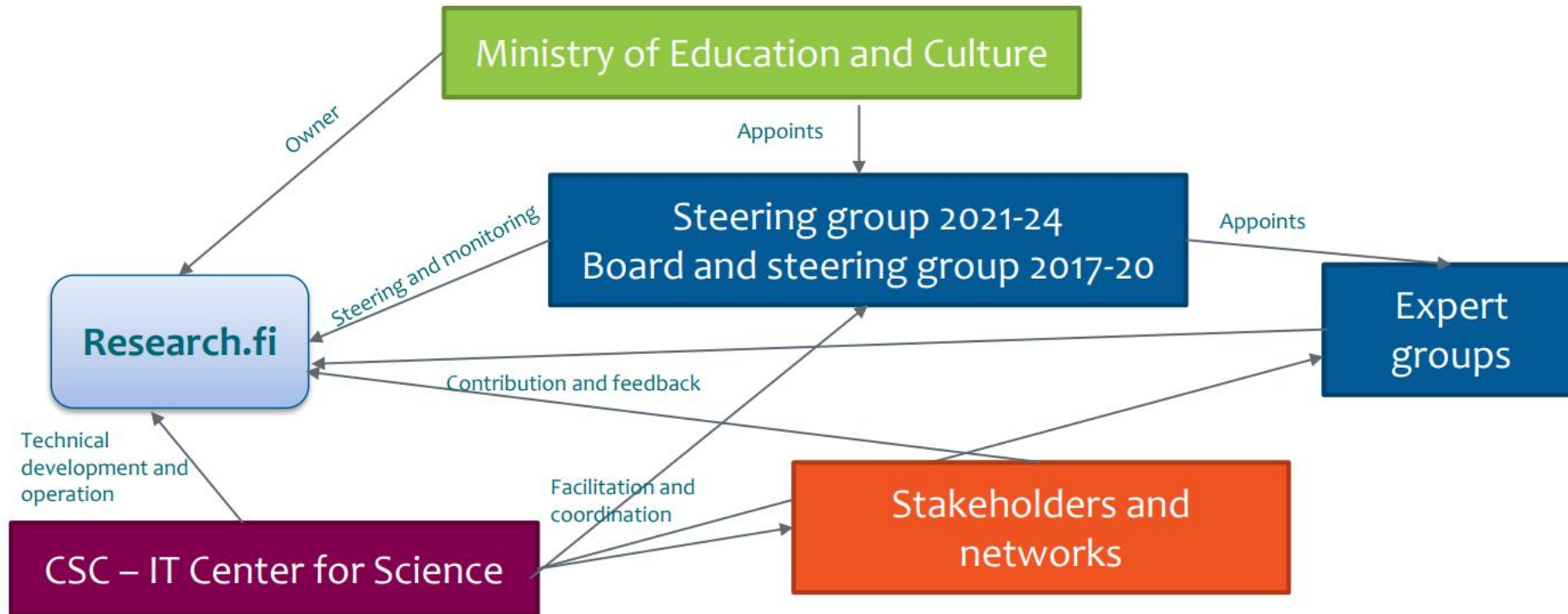


Puuska, H.M. (2021). Research.fi – the Finnish research information hub. <http://hdl.handle.net/11366/1790>



Puuska, H.M. (2021). Research.fi – the Finnish research information hub. <http://hdl.handle.net/11366/1790>

Research.fi - organization



Puuska, H.M. (2021). Research.fi – the Finnish research information hub. <http://hdl.handle.net/11366/1790>

Aims of the Finnish Research Information Hub



R&D analytics
Visualizations
Business
Intelligence
Landscape

To describe the Finnish
research ecosystem and
act as an information
base for science policy

To help to find Finnish
research and researchers

Research.fi portal
Semantic search
Linked data

APIs
Shared data
model

To support the reuse of
existing research
information in other
services and systems

To support the merit of
researchers and increase
the visibility of their
research output and
activities

Researcher's Profiles
Traditional research outputs
Other research activities

Nikkanen, J. (2023). Research.fi and the lessons learned so far. <http://hdl.handle.net/11366/2694>

Future topics: Researcher's MyData in Research.fi

“The term MyData refers to: 1) a new approach, a paradigm shift in personal data management and processing that seeks to transform the current organisation-centric system to a human-centric system; 2) personal data as a resource that the individual can access and control”

<https://www.sitra.fi/en/dictionary/mydata/>

- Creating a profile at Research.fi is completely voluntary for researchers.
 - The researchers can decide what information they want to publish about themselves
 - The researchers may authorize the disclosure of the information to any party they want, such as funders or other researchers' services
- > Emphasizing the "My" in MyData,
- > Research.fi as MyData operator

Nikkanen, J. (2023). Research.fi and the lessons learned so far. <http://hdl.handle.net/11366/2694>

Future topics: eOSC | FAIRCORE4EOSC



- Developing EOSC-Core components to enable a FAIR EOSC ecosystem 2023-2025
- Realize the potential in combining CRIS-like information from national sources on European level, aggregating this information to RDGraph and OpenAIRE
- Showcase the CERIF data model as common for exchanging information between organizational and national CRIS systems & OpenAIRE
 - Included in the case study: Research.fi (Finland), HELIX (Greece), Open Knowledge Base (Netherlands)
 - To use refactored CERIF at <https://github.com/EuroCRIS/CERIF-Core>
- Introducing RAiDs to CRIS-domain + organizations
- Have a template for integration of further national research information systems with the EOSC Core

Nikkanen, J. (2023). Research.fi and the lessons learned so far. <http://hdl.handle.net/11366/2694>

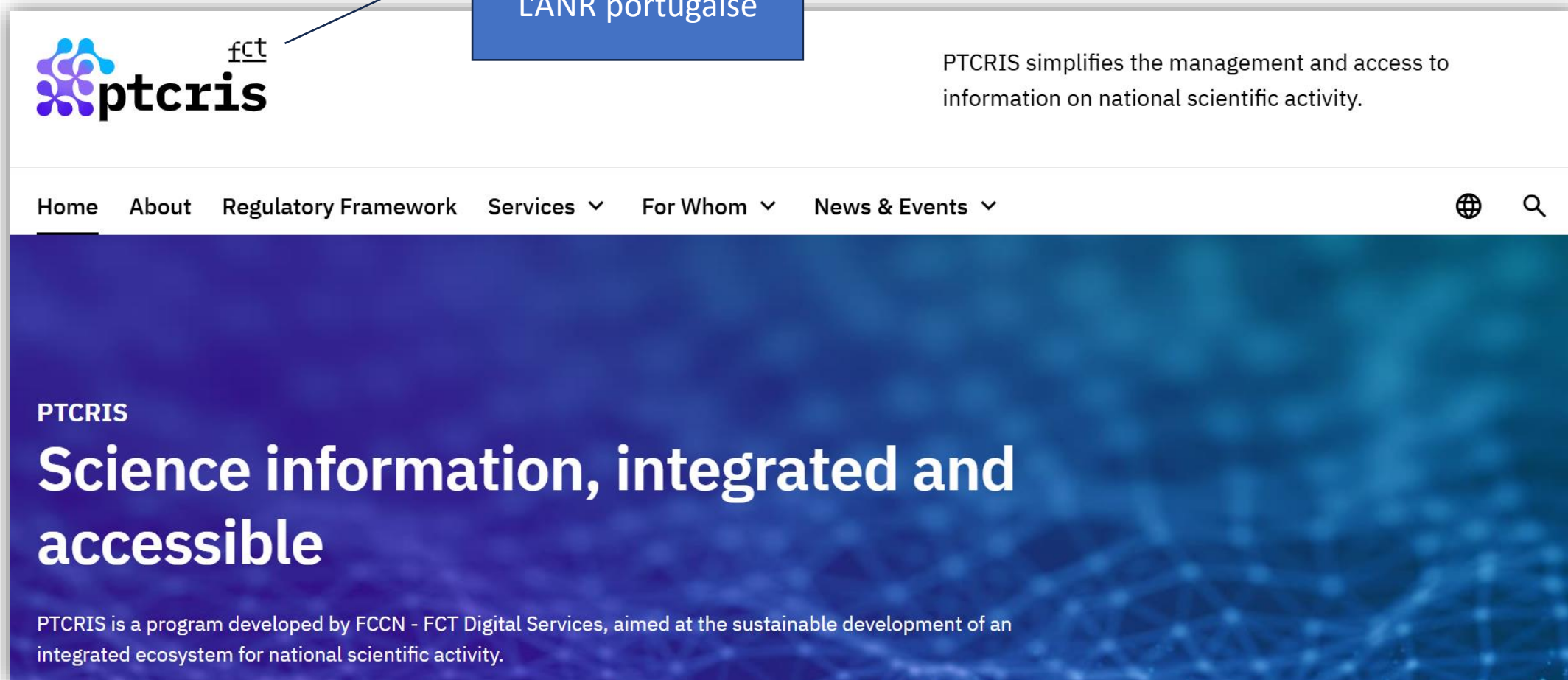
Lessons learned (Nikkanen, 2023)

1. La coopération est essentielle : research.fi s'appuie sur une équipe de projet clairement structurée ainsi que sur des groupes de travail.
2. Il faut maintenir (et améliorer, mais pas remplacer) les processus existants dans les institutions partenaires.
3. La durabilité et la longévité sont importantes : Les objectifs initiaux génèrent continuellement de nouveaux objectifs, c'est pourquoi l'établissement de priorités stratégiques et la gestion de projet restent pertinents même pendant le fonctionnement régulier et après que les objectifs initiaux ont été atteints.

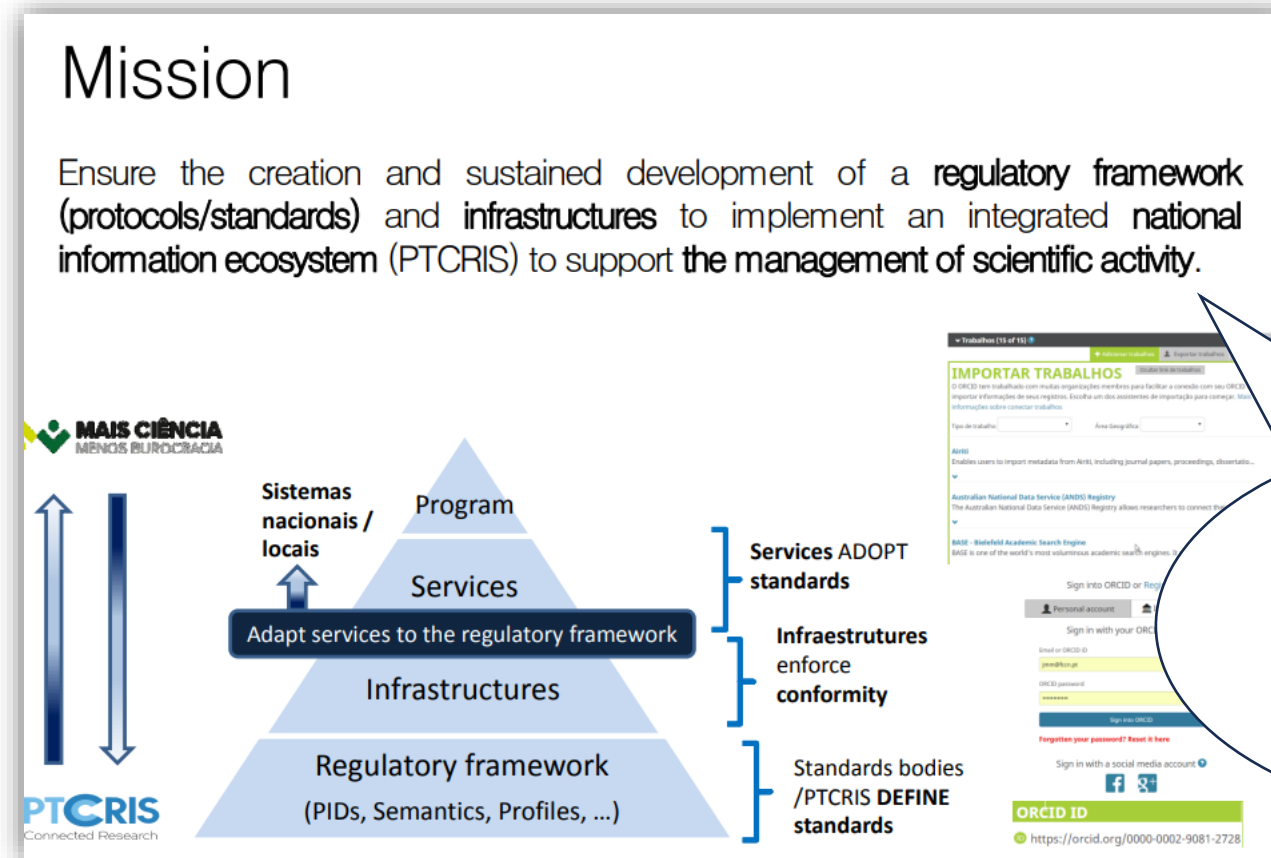
PT CRIS

"more science, less bureaucracy"

L'ANR portugaise



PTCRIS, un écosystème d'information national



PTCRIS va implémenter la carte CRIS allemande https://kfid-online.de/fis_landkarte_visual/

Moreira, J.M., & Laranjeira, C. (2018). PTCRIS: Planning and implementing a national CRIS ecosystem.

<http://hdl.handle.net/11366/664>

Le cadre normatif de PTCRIS

Data model

Interoperability
dictionary

Ontologies

Unique individual
identifiers

Organization identifiers

DOI identifier for
publications and funding

Scientific fields

- Le modèle de données de référence est CERIF (*Common European Research Information Format*), recommandé par l'UE pour la gestion et l'échange d'information sur la recherche.
- L'interopérabilité entre PTCRIS et d'autres infrastructures utilise le profil OpenAIRE de CERIF, qui permet d'exporter, de collecter et d'importer des métadonnées, garantissant ainsi le respect des principes de la science ouverte.

Cadre normatif (suite)

- PTCRIS adopte le dictionnaire CASRAI et utilise CERIF-XML pour le transfert.
- PTCRIS considère l'ontologie VIVO comme une référence internationale.
- La référence établie pour l'identification des chercheurs est le système CIÊNCIA ID+. Dans le cadre de ce système, toute personne opérant dans le système académique et scientifique dispose d'un CIÊNCIA ID et d'un ORCID.
- La référence adoptée pour l'identification des organisations est ISNI+. PTCRIS accompagne également l'initiative ROR.
- L'identifiant DOI (Digital Object Identifier) a été adopté.
- PTCRIS adopte la taxonomie FOS/FORD (Fields of Science/Fields of Research and Development) de l'UNESCO.

L'adoption de ce cadre normatif

- L'adoption du cadre normatif de PTCRIS dans le contexte du *PTCRIS Data Catalog* et par les différents systèmes opérant dans l'écosystème ESR national est fondamentale pour garantir une intégration large et efficace et, par conséquent, la réalisation du principe de base de PTCRIS : *register once, reuse always*.

D'autres exemples

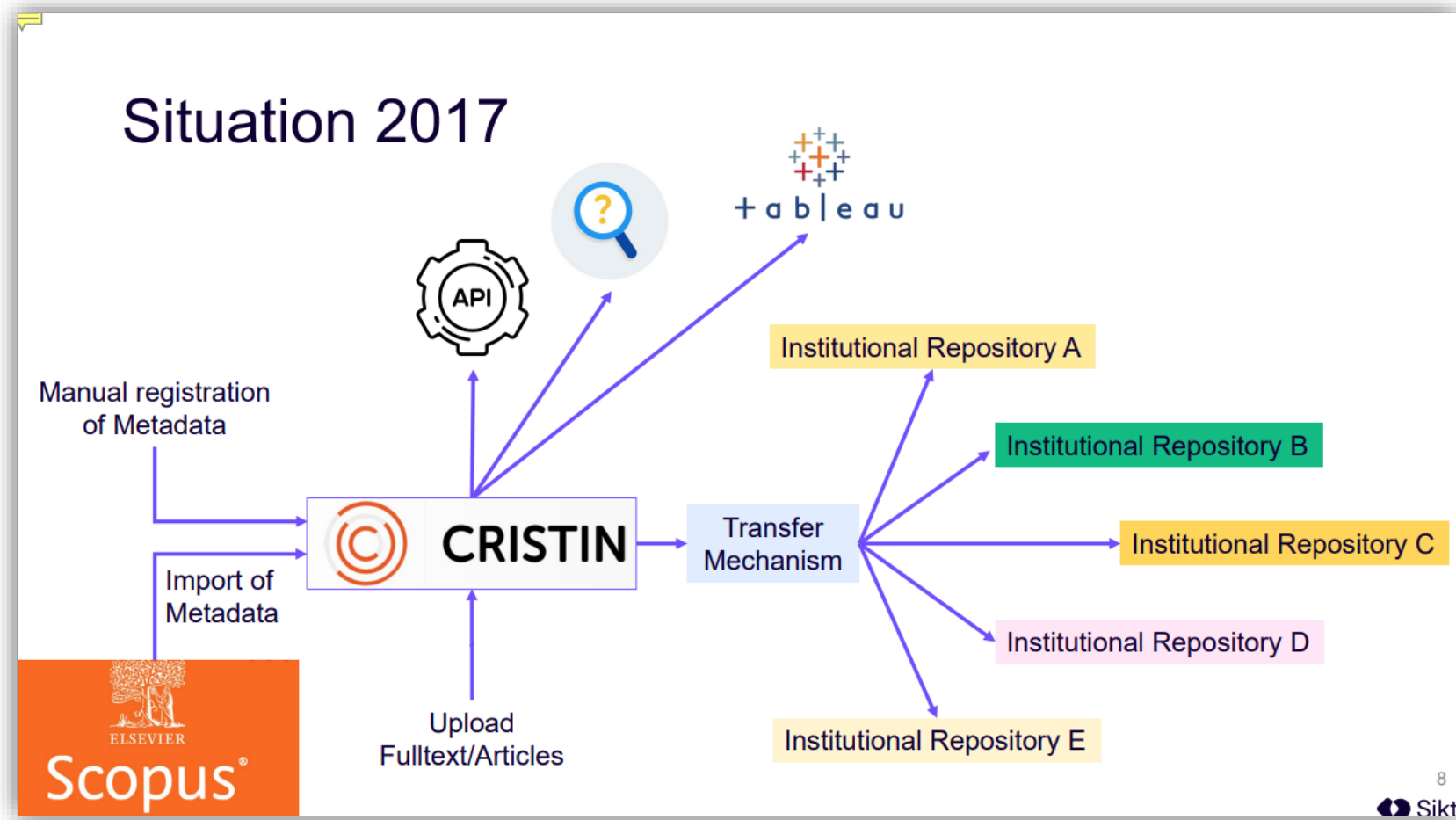
Pologne « [Ludzie Nauka](#) »

- [Pilotage par le Ministère](#)
- Alimentation par différentes sources
 - Bases de données, archives...
 - Systèmes institutionnels
- Transfert de données vers d'autres dispositifs
 - Agences de financement
 - Portail d'évaluation RADON

Norvège « [Nasjonalt Vitenarkiv](#) »

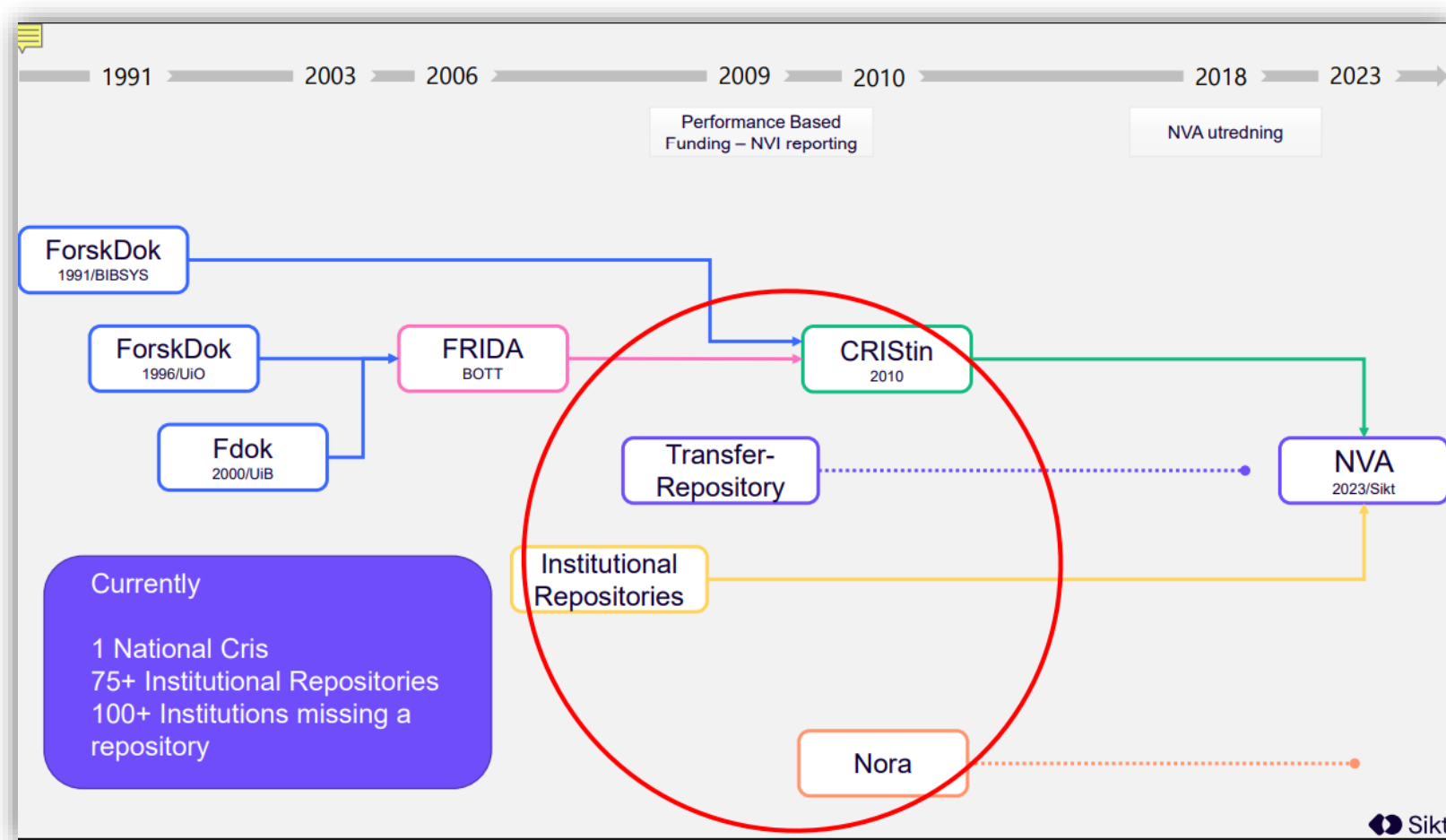
- [Pilotage par le Ministère](#)
- Double alimentation
 - Import semi-manuel de métadonnées et de fichiers à partir de sources externes validées (Scopus...).
 - Dépôt manuel

Norvège : d'un SI recherche national...



Garshol, J.E. (2023). NVA: A National Repository and Research Information System. <http://hdl.handle.net/11366/2449>

... vers une plateforme d'agrégation



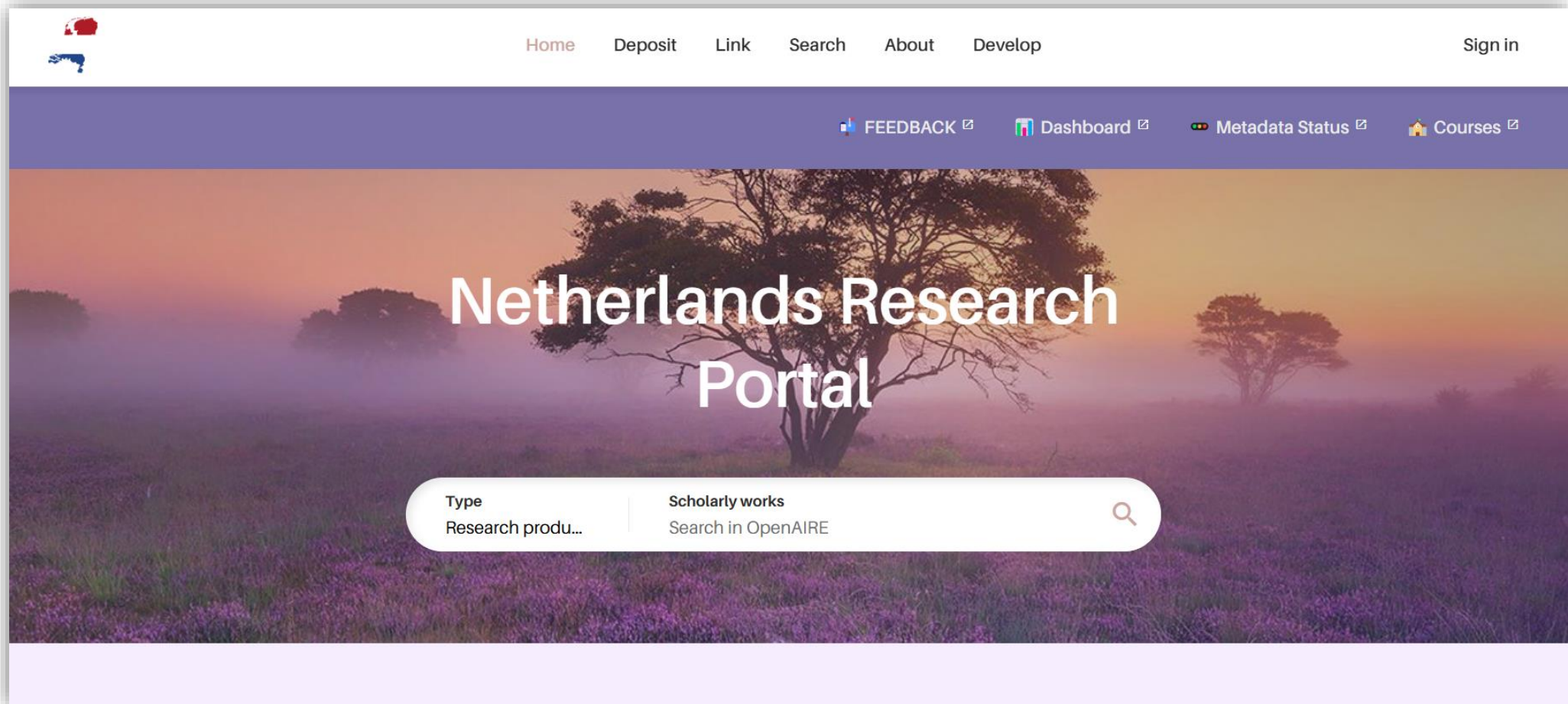
Garshol, J.E. (2023). NVA: A National Repository and Research Information System. <http://hdl.handle.net/11366/2449>

Les identifiants au cœur du dispositif

Identifiers   	• Persistent IDs is used where possible • Every person can connect their ORCID • All organizations have a ROR ID • DOIs used wherever relevant • Connection between research data, research results, projects, grants etc. • Biobanks will be connected • Investigating RAiD for Projects • Pilot on DOI for Grants • Connect Approvals originating from Ethical assessments as well as Privacy assessments relating to Research projects.
--	--

Garshol, J.E. (2023). NVA: A National Repository and Research Information System. <http://hdl.handle.net/11366/2449>

Une autre option : OpenAIRE



OpenAIRE et SI recherche

- Le graphe OpenAIRE s'inspire du modèle CERIF (auquel OpenAIRE est conforme, au moins partiellement)
- Les SI recherche sont interopérables avec l'agrégation OpenAIRE via les OpenAIRE *guidelines*, développées conjointement par euroCRIS et OpenAIRE.
- OpenAIRE peut servir de base pour un SI recherche.
 - Cf. exemple des Pays Bas.
- Mais le modèle d'OpenAIRE a pour l'instant certaines lacunes (ex. instruments, projets...).

Les enjeux

- Depuis 2022, un [groupe de travail d'euroCRIS](#) a recueilli les retours d'expérience de plus de 20 projets de SI recherche nationaux et régionaux.
 - Le constat majeur : « *a scattered landscape* »
 - Des réunions avec des représentants finlandais, danois, portugais et français depuis 2023.
- Malgré des différences liées à la taille des infrastructures, aux spécificités locales et aux niveaux de maturité variables, les enjeux sont similaires :
 - Standardisation des formats et des données (CERIF, KDSF...)
 - Stratégie sur les identifiants pérennes (DOI, ORCID, ROR, grant IDs...)
 - Problématique : organisations, infrastructures de recherche et autres équipements
 - Sélection, nettoyage et déduplication des données
 - Création de communautés d'utilisateurs partageant des pratiques et valeurs
 - Transfert des données (incitation, obligation, conditions...)

Une dernière remarque

- L'avenir des SI recherche est lié à l'intelligence artificielle.
 - Automatisation de l'intégration des données
 - Amélioration de la qualité des données
 - Recommandations
 - ...
- Groupe de travail d'euroCRIS
 - [AI4CRIS](#)
 - Piloté par António Luís Lopes (U Lisbonne)
 - Réunion à [Leuven en mai 2025](#)
- Travail sur la transposition du *EU AI Act* dans le domaine des SI recherche

Références

- Bryant, R., Clements, A., Feltes, C., Groenewegen, D., Huggard, S., Mercer, H., ... Wright, J. (2017). *Research Information Management: Defining RIM and the Library's Role*. Retrieved from <https://www.oclc.org/research/publications/2017/oclcresearch-defining-rim.html>
- Bryant, R., Clements, A., de Castro, P., Cantrell, J., Dortmund, A., Fransen, J., ... Mennielli, M. (2018). *Practices and Patterns in Research Information Management. Findings from a Global Survey*. Retrieved from <https://www.oclc.org/research/publications/2018/oclcresearch-practices-patterns-research-information-management.html>
- Bryant, R., Fransen, J., Castro, P. de, Helmstutler, B., & Scherer, D. (2021). *Research Information Management in the United States*. <https://doi.org/10.25333/8hgy-s428>
- de Castro, P. (2018). The role of Current Research Information Systems (CRIS) in supporting Open Science implementation: the case of Strathclyde. *ITLib*, (5), 21–30. <https://doi.org/https://dx.doi.org/10.25610/itlib-2018-0003>
- de Castro, P., Shearer, K., & Summann, F. (2014). The Gradual Merging of Repository and CRIS Solutions to Meet Institutional Research Information Management Requirements. *Procedia Computer Science*, 33, 39–46. <https://doi.org/10.1016/j.procs.2014.06.007>
- Ribeiro, L., de Castro, P., & Mennielli, M. (2016). *EUNIS – EUROCRIS joint survey on CRIS and IR. Final report*. Retrieved from <http://www.eunis.org/wp-content/uploads/2016/03/cris-report-ED.pdf>

Merci !

Joachim Schöpfel schopfel@eurocris.org <https://cv.hal.science/joachim-schopfel>

Pablo de Castro pablo.de-castro@eurocris.org

euroCRIS eurocris@eurocris.org