

ESUP-DAYS 36

VERS UN SYSTÈME D'INFORMATION RECHERCHE LIBRE ET MUTUALISÉ PAR LA CONSTRUCTION D'UN GRAPHE DE CONNAISSANCES INSTITUTIONNEL

REYMOND David -- Université de Toulon
DORNBUSCH Joachim -- Université Paris 1 Panthéon-Sorbonne



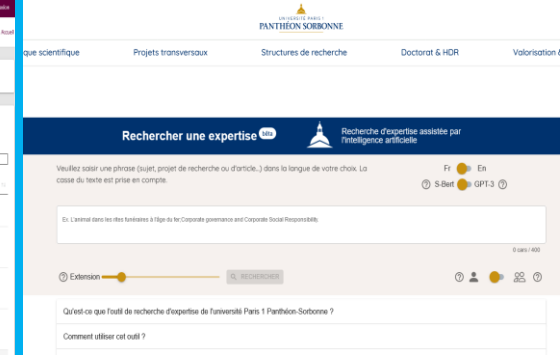
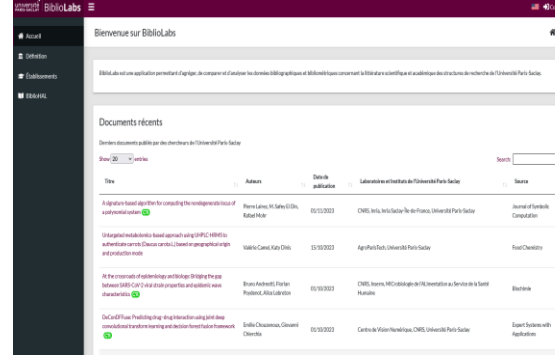
ESUP-DAYS 36

De SoVisu à SoVisu+

ORIGINES

Applications et usages de la documentation

- Le logiciel BiblioLabs (Bellier, Bretel et Gandois 2021), développé par Paris-Saclay focalise sur les fonctions de bibliométrie et de gestion documentaire des productions (**sans les chercheurs**)
- *L'Expert Finding System* de Paris 1 Panthéon Sorbonne prouve le potentiel d'utilisation des GML sur l'archive Hal pour l'identification d'expert et l'intérêt de mettre **les chercheurs** au cœur de ce processus
- SoVisu a été conçu, pour accompagner **les chercheurs de UTLN** à l'autoarchivage et cartographier les expertises



1

[BiblioLabs](#)

2

[Expert Finding System](#)

3

[SoVisu \(démonstration\)](#)

GESTION DE SES PROPRES Ids

Vérification des données > D. REYMOND

DÉPÔT
2023

on des données

de bord

es

e

extrait

S

s labora

s cherc

N

he d'exp

RMATIC

documen

Identifiants Notices HAL Expertises identifiées Domaines d'expertise

Dans cet onglet, vous devez vérifier, renseigner ou corriger vos identifiants chercheur.
L'IdHal est votre identifiant sur l'archive nationale Hal. Il permet de regrouper vos publications et de
L'idRef s'obtient dès la publication de votre thèse. Vous pouvez le vérifier : [sur le site prévu à cet e](#)
L'Orcid est un identifiant chercheur international. [Nous vous recommandons d'en créer un.](#) Vous av
Si vous avez **d'autres identifiants** (Arxiv, Mendeley, Research Gate...) leurs moissonneurs se charg
Merci de vérifier ces identifiants avant de les renseigner ici, puis validez. Vous accédez alors :

ID HAL (texte)
david-reymond

aurehal ID
7000

domain_shs.info.conf

Informatique

domain_shs.info.orga

domain_shs.info.bibl

domain_shs.info.hype

domain_shs.info.gest

domain_shs.info.inge

domain_shs.info.hist

Le chercheur vérifie :

- L'alignement de ses identifiants,
- Son statut,
- Son(s) affiliation(s)
- Son(s) domaines d'expertises
- Les données de l'annuaire ?

- Bibliothèque électronique
- Recherche d'information
- Web
- Traitement du texte et du document
- Intelligence artificielle

CONTRÔLE DE LA REPRÉSENTATION LEXICALE DE SES PRODUCTIONS

The screenshot displays the SoVisu+ interface. At the top left, there is a search bar with a magnifying glass icon and the text "validés X". Below this, there are two word clouds. The left word cloud is titled "Mots-clés sur notices Hal (FR) - Décompte" and features the word "Manquants" in large green letters. The right word cloud is titled "Termes extraits (résumés FR) - Décompte" and features the phrase "Résumés manquants" in large green letters. Below the word clouds is a table with two columns: "ID Hal" and "Titre". The table contains several rows of data, each with an "Editer la notice" link to the left of the "Titre" text.

ID Hal	Titre
	Approche extensive des métadonnées pour un site web : principes d'élaboration e
	La communication engageante, pour sensibiliser les visiteurs du parc national de f
	The use of engaging communication to make visitors of the Port-Cros National Pa
	Analyse lexicale d'une page web
	Lexical analysis of a web page
	Se former aux humanités numériques par les données ouvertes : les chaînes infon
	Training in Digital Humanities through Open Data: The Infometrics Chains of Paten

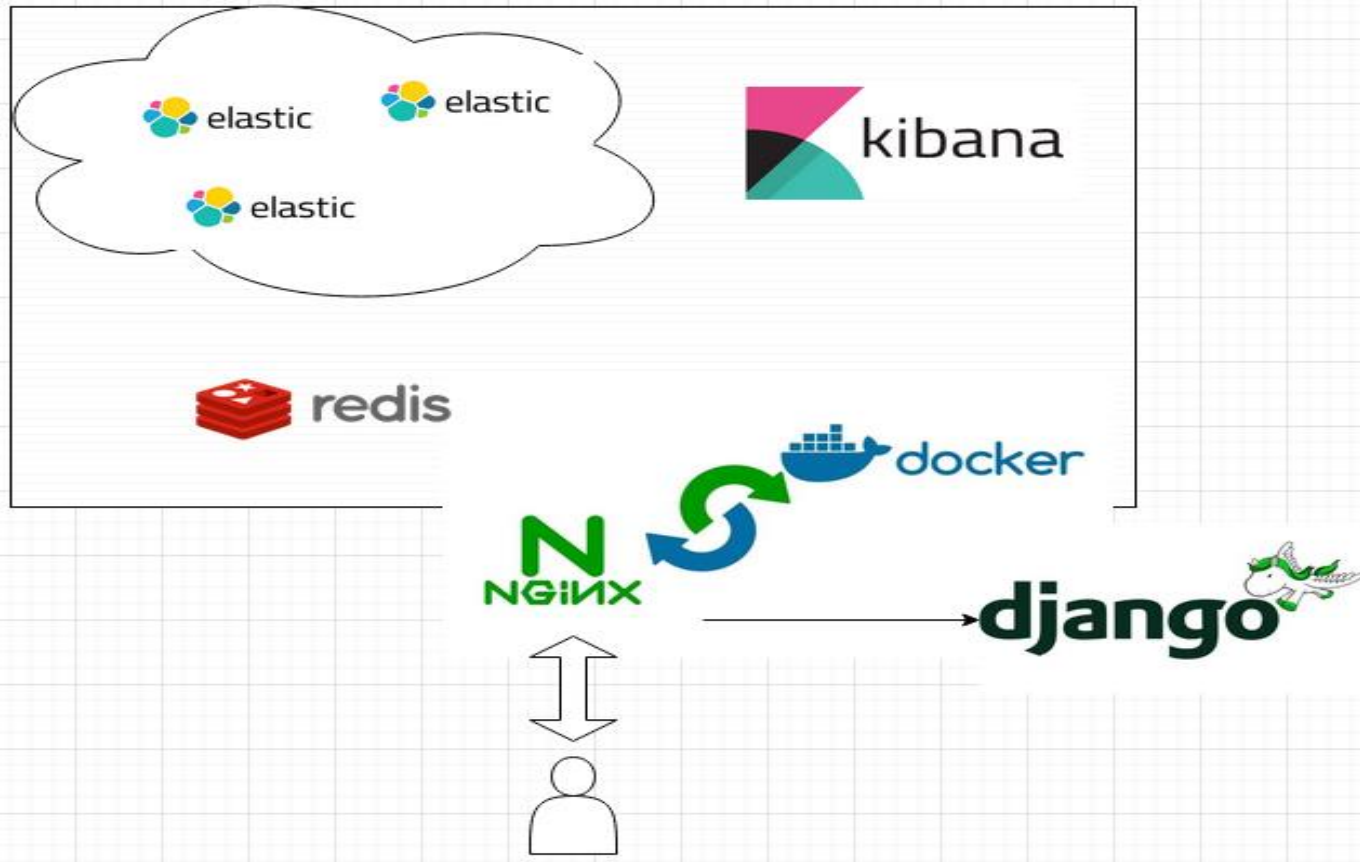
La publication n'est plus la seule affaire de la diffusion :

- mots-clés,
- résumé,
- texte complet : un graal s'il est bien décrit !

→ Accompagner le chercheur dans la gestion de son profil scientifique

- Faciliter,
- Ne faire qu'une seule fois,
- Capitaliser

ARCHITECTURE OUVERTE ET ÉPROUVÉE



- docker
- Django (vues sur données)
- Nginx (reverse proxy sécurité)
- Elastic (cluster indexations)
- Kibana (tableaux de bord)
- Redis (back office)

Avantages et limites

Intérêt de l'écosystème (chercheurs, valorisation, doc...)



Curation collective des données



Offre une potentielle interface ciblée de médiation



Empiète sur le SI et le périmètre BiblioLabs et EFS

Les graphes de connaissances représentent des concepts (par exemple, des personnes, des lieux, des événements) et leurs relations sémantiques. En tant que structure de données, ils soutiennent un système d'information numérique, aident les utilisateurs dans la découverte et la récupération de ressources, et sont utiles à des fins de navigation et de visualisation. Dans le domaine des bibliothèques et des sciences humaines, les graphes de connaissances sont généralement ancrés dans des systèmes d'organisation des connaissances, qui ont une tradition centenaire et ont subi leur transformation numérique avec l'avènement du Web et des données liées

Bernhard Haslhofer, Antoine Isaac, Rainer Simon, 2018
https://doi.org/10.1007/978-3-319-63962-8_291-1

L'ÉCOLE
DES HAUTES
ÉTUDES EN
SCIENCES
SOCIALES

Inria

UN Nantes
Université

Université Claude Bernard
 Lyon 1

 UNIVERSITÉ DE
MONTPELLIER


UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

université
PARIS-SACLAY


UPEC | UNIVERSITÉ
PARIS-EST CRÉTEIL
VAL DE MARNE
Connaissance - Action

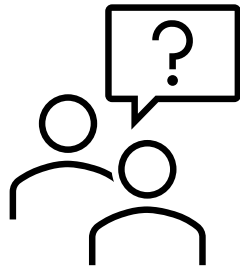
 UNIVERSITÉ DE
TOULON

SoVisu+

(Re)constitution du graphe de connaissance institutionnel

CRISALID

Comment aborder le domaine "recherche" du SI institutionnel...

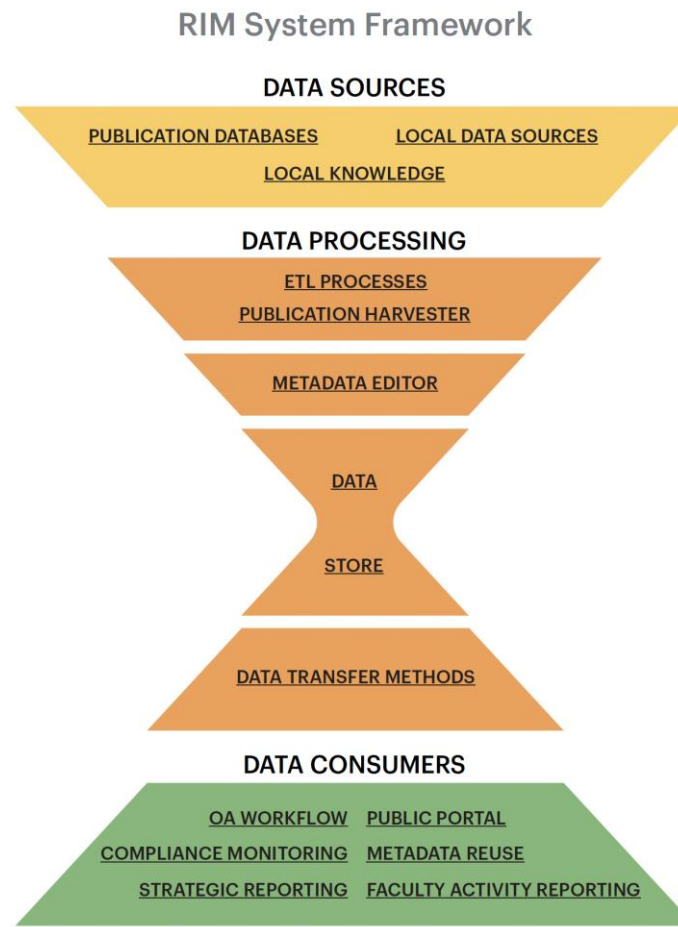


... sans créer de la dette technique pour les 20 ans à venir

Un peu de recul

- « Research information management (RIM) systems support the transparent aggregation, curation, and utilization of data about institutional research activities. »

Bryant, Rebecca, Jan Fransen, Pablo de Castro, Brenna Helmstutler, and David Scherer. 2021. *Research Information Management in the United States: Part 1—Findings and Recommendations*. Dublin, OH: OCLC Research. <https://doi.org/10.25333/8hgy-s428>.



Anticiper la diversité



Base de données projets ?



Progiciels recherche ?



Wos / Scopus ?



Stratégie identifiants ?



Web / portail recherche ?



État annuaire institutionnel ?

Systeme

modulaire

paramétrable

compatible cloud

Modularité de l'architecture fonctionnelle

SoVisu +

Gestion/curation
production scientifique,
identité numérique

Brique **chercheur**

Idyia

Moteur de recherche
d'expertise

Brique **publique (+portail)**

Agrégation
production
scientifique, suivi
science ouverte,
bibliométrie

Brique **bibliothèque**

Référentiel
laboratoire,
rapports Hceres,
utilitaires

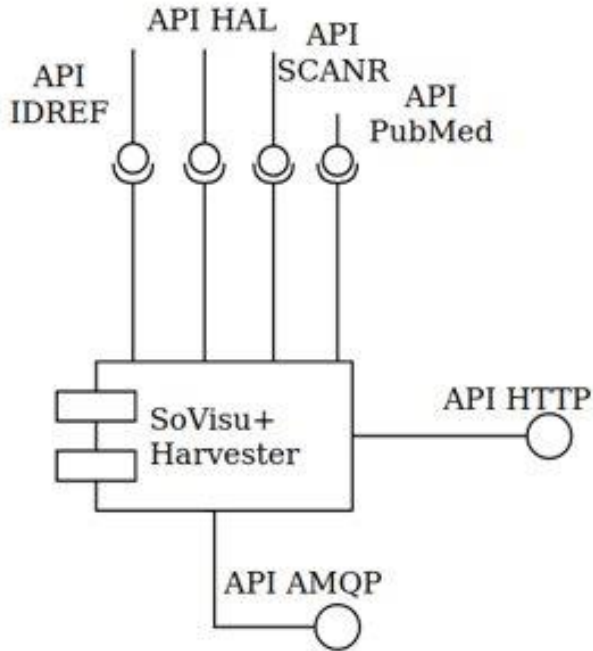
Brique **laboratoire**

Référentiel
recherche,
tableaux de bord,
pilotage

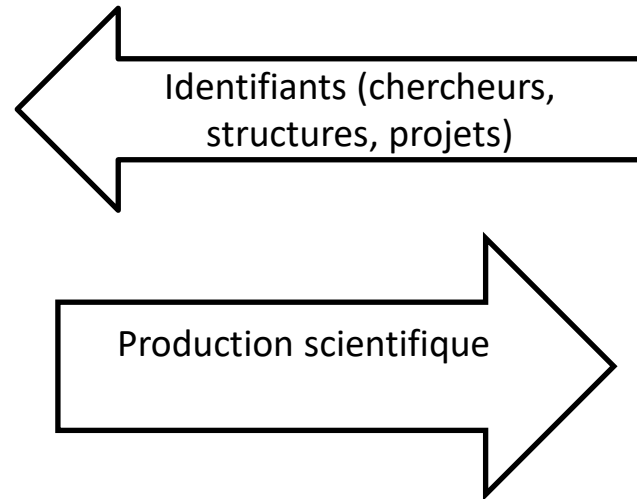
Brique **direction
recherche**

Modularité de l'architecture technique

Ex. SVP Harvester



<https://svp-harvester.readthedocs.io>



Objectif qualité

Démarche
orientée
utilisateur

Outils de
contrôle qualité

Couverture par
les tests

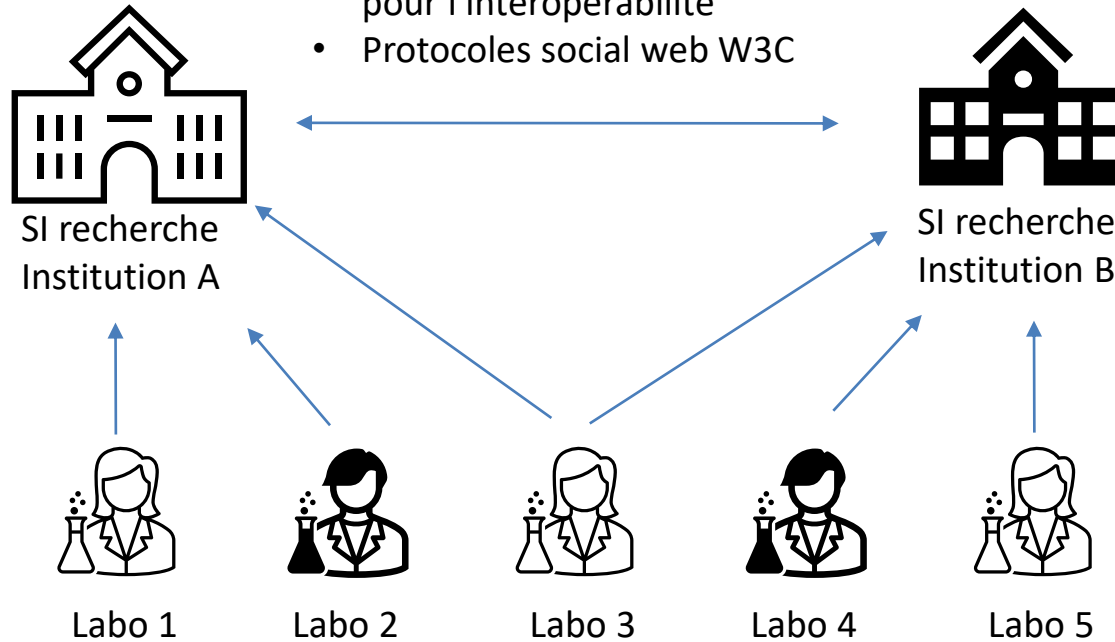
Automatisation

Documentation

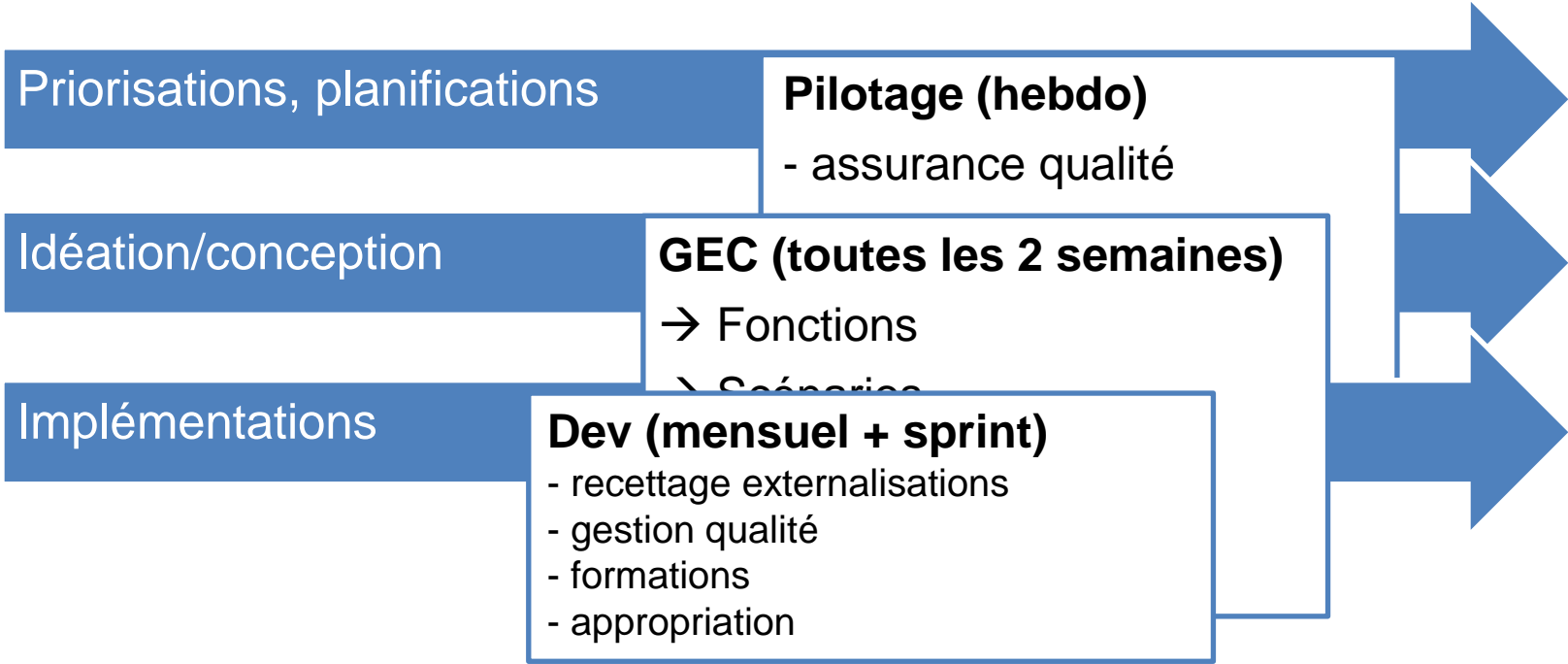
I18n,
responsive,
RGAA

Objectif interoperabilité

- Cadre de cohérence recherche
- Standards web sémantique pour l'interopérabilité
- Protocoles social web W3C



Organisation (2023/2024)



démonstrateur

Prototype

UNE COMMUNAUTÉ NATIONALE

- ◆ Documentalistes
- ◆ Services de valorisation
- ◆ Chercheurs
- ◆ Informaticiens

- ◆ Du code libre, ouvert, modulaire, testé et de qualité...
- ◆ Des formats interopérables avec l'existant à l'international

- ◆ Documentation
- ◆ Accompagnement,
- ◆ Maintenance,
- ◆ Suivi

Intérêt de l'intégration dans la communauté e-sup :

- visibilité,
- utilisation des services messagerie, confluence, git
- traçage et suivi des travaux,
- partage et éventuelle collaboration et participation,
- retours critiques, compléments, ...
- construit pour durer dans l'intérêt de la communauté

RÉFÉRENCES

- ◆ Tabariès, A., & Reymond, D. (2021, January 27). Comparaison de méthodes d'extraction de mots-clés non supervisées. *Atelier TextMine - 21ème édition conférence Extraction et Gestion des Connaissances (EGC)*, Montpellier, France.
- ◆ Reymond, D. 2021b. « SoVisu, terrains de recherches SIC en cours ». présenté à Réunion IMSIC AXE 1 Enjeux et usages des dispositifs sociotechniques numériques et des mutations informationnelles Captas, documentation et numérique, juin 18.
- ◆ —. 2022a. « SoViSu : visibilité et lisibilité en SO ». présenté à Séminaire IMSIC AXE 1 *Brevets, captas et Science Ouverte*, février 24.
- ◆ —. 2022b. « Involving researchers in Open Science: the SoVisu innovative solution ». Dans *EOSC Symposium, session MVE beyond EOSC Future and the INFRAEOSCO7*. Prague, Tchéquie.
- ◆ —. 2023. « Les chercheurs dans la boucle de la Science Ouverte. Et l'IA ? » présenté à Séminaire ObTIC, Sorbonne Université, Paris, France, mai 11.
- ◆ —.(2023, Juin) Le dispositif “SoVisu”. Séance publique de l’atelier “cartographie de la recherche”, Université Paris 1 Panthéon-Sorbonne, Paris, France. [⟨hal-04141467⟩](#)
- ◆ —., Lapôtre, R. « SoVisu+: starting point and foundations of a national CRIS » présenté à la conférence *Semantic Web in Libraries (SWIB) 23*, Berlin, Allemagne, sept 12.
- ◆ —. , Buccheri, A., Dornbush, J., Dubernard, H., Lapôtre, R., (2023, sep) SoVisu+ (SVP). 4^e séance publique de l’atelier “cartographie de la recherche”, Université Paris 1 Panthéon-Sorbonne, Paris, France