

Vers un entrepôt de gestion des données pour l'université : retours d'expérience du projet Data Paris 8

Goran Sekulovski, SCD Paris 8

Johann Holland, TGIR Huma-Num

Journée d'études : Documents et données : l'université à l'ère de la transition numérique

26 mai 2021

1) Une politique nationale

The screenshot shows the top part of a website. At the top left is a black square with three white horizontal lines. To its right is the text 'OUVRIR LA SCIENCE' in bold, with 'LA SCIENCE' underlined. Further right is a navigation menu with the following items: 'SCIENCE OUVERTE', 'COMITÉ', 'GROUPES', 'BLOG', 'AGENDA', 'RESSOURCES', and 'LETTRES'. At the top right, there is a language selector 'FR - EN' and a magnifying glass icon. Below the navigation is a dark purple horizontal bar. The main content area has a light green background. It features the text 'Axe 2' in bold, followed by the main heading 'Structurer et ouvrir les données de recherche' in a larger bold font. Below this is a paragraph of text: 'Notre ambition est de faire en sorte que les données produites par la recherche publique française soient progressivement structurées en conformité avec les principes FAIR (Facile à trouver, Accessible, Interopérable, Réutilisable), préservées et, quand cela est possible, ouvertes.' In the bottom right corner of the green area, there is a small logo with a circular arrow and the text 'Confidentialité - Conditions'.

[Axe 2 du Plan national pour la Science Ouverte \(2018\)](#)

1) Une politique nationale

2 | LES DONNÉES DE LA RECHERCHE

Objectif: *Les données (données brutes, textes et documents, codes sources et logiciels) produites par les chercheurs et les chercheuses CNRS ou avec des moyens mis en œuvre par le CNRS doivent être, dans la mesure du possible, rendues accessibles et ré-utilisables selon les principes FAIR pour une consolidation des connaissances essentielle au développement d'une science plus efficiente. « Les données doivent être aussi ouvertes que possibles, et fermées autant que nécessaire ».*



1) Une politique nationale

La science ouverte



- Contribuer à l'ouverture des données quand cela est possible (Open Research Data)
- Élaboration obligatoire d'un Plan de Gestion des Données (PGD) pour les projets financés à partir de 2019

2) Une politique européenne



[About](#) [Services & Resources](#) [Policy](#) [Use Cases](#) [Media](#) [For providers](#) [Subscribe](#) [Using the Portal](#) [Q](#)

EOSC relies on 6 action lines:

[Learn more in the EC Staff Working Document](#)

**ARCHITECTURE,
DATA,
SERVICES,
ACCESS & INTERFACE,
RULES,
GOVERNANCE**

- L'European Open Science Cloud vise à relier les infrastructures/plateformes de données scientifiques européennes existantes, à y joindre des solutions cloud à haute capacité et/ou performance pour les exploiter.
- L'EOSC vise aussi à rendre les données scientifiques FAIR (Findable, Accessible, Interoperable, Reusable) et ouvertes.

[European Open Science Cloud, une initiative européenne](#)

3) Une déclinaison locale : Projet Data Paris 8

Comité de pilotage

- Chefs de projet : chargé de mission et VP Numérique
- VP Recherche
- Direction de la recherche
- Service Commun de la Documentation
- Référents « données de la recherche et science ouverte » des laboratoires
- Délégué à la protection des données (DPO)
- Délégué à l'intégration scientifique
- Direction du système d'information et du numérique

Partenaire

- TGIR Huma-Num (infrastructure de recherche dédiée aux lettres, sciences humaines et sociales) : Pôle Données / Accompagnement utilisateurs

3) Une déclinaison locale : Projet Data Paris 8

Objectifs du Comité de pilotage :

- Recenser les gisements de données existants et faire un état des lieux des pratiques et des besoins au sein de l'Université
- Faire un état des lieux du cadre légal applicable aux données numériques, dans un contexte complexe et mouvant (loi République Numérique, RGPD, etc.)
- Situer l'Université Paris 8 par rapport aux infrastructures locales, nationales et internationales en matière d'archivage, de traitement et de diffusion de données
- Dynamiser la recherche grâce à un entrepôt permettant le dépôt, la diffusion et le stockage à long terme des données collectées par les laboratoires de l'université

3) Une déclinaison locale : Projet Data Paris 8

Documents et données : l'université à l'ère de la transition numérique

Questionnaire - données de la recherche

Nom du laboratoire et nom du/de la répondant.e

Quelle est la nature des données produites ?

Par exemple : données d'observation, expérimentales, de simulation, dérivées, de référence, etc.

Quelles sont les types de données que vous voulez héberger ?

Par exemple : image ; audio ; vidéo ; texte ; données issues d'enquête quantitative (statistique) ; données issus d'enquête qualitative ; relevés ; données structurées et déjà exploitées ; autre (à préciser).

	B	C	D	E	F
1	Nom du laboratoire et nom du/de la répondant.e	Quelle est la nature des données produites ?	Quelles sont les types de données que vous voulez héberger ?	Utilisez-vous des formats de données non-ouverts et est-ce nécessaire de le préserver ?	Auriez-vous besoin d'une protection de données particulière (données sensibles) ?
2		Attention, je réponds pour moi par pour le laboratoire. Articles d'une part, images produites par nos propres logiciels.	Images, textes illustrés d'images	Non	Non
3		non concerné	non concerné		non
4		Données d'observation et expérimentales	image ; audio ; vidéo ; texte ; données issues d'enquête quantitative (statistique) ; données issus d'enquête qualitative ; relevés de terrains	Non	Oui
5	Anne E. Berger / UMR LEGS	Nous produisons essentiellement des données d'observation et de référence.	Images, audio, vidéos, textes. Données issues d'enquête quantitatives et qualitatives.	Nous utilisons en majorité des formats ouverts, mais, dans certains domaines touchant aux sujets humains, l'obligation de signer des chartes éthiques assurant l'exclusivité de l'accès aux données ou la limitation de leur diffusion entre en contradiction avec les objectifs d'une science ouverte.	Certain.e.s de nos collègues anthropologues par exemple sur des questions touchant à la migration encore aux migrations doivent prendre des précautions qui peuvent aller jusqu'à la destruction des données après leur analyse.
6	Section Clinique	Résultats de calculs avec le software MAGMA, articles, textes (dans le sens des fichiers avec des chiffres et symbole).	Texte (dans le sens des fichiers avec des chiffres et symbole).	Non.	Non.
		Données d'observation et d'entretien issues d'enquêtes qualitatives: documents	Données d'observation et d'entretien issues d'enquêtes qualitatives: documents, papiers, documents		

Enquête préliminaire auprès des laboratoires (des UMR et des EA) de Paris 8

3) Une déclinaison locale : Projet Data Paris 8

Données primaires collectées à l'Université Paris 8 :

- Documents textuels
- Documents audio et vidéo
- Archives
- Documents graphiques
- Entretiens
- Données d'enquêtes
- Données Web

Données dérivées produites à l'Université Paris 8 :

- Prises de note
- Tableurs
- Bases de données
- Analyses de données
- Traitements vidéo ou sonore
- Cartes

3) Une déclinaison locale : Projet Data Paris 8

Les documents non achevés comme les carnets de laboratoire, les carnets de recherche, les analyses préliminaires, les programmes de travaux futurs ainsi que les données déjà publiées sous forme de résultats dans des articles, les communications à un congrès ou bien les supports de formation ne relèvent pas des données de la recherche.

Attention à ce que l'on appelle « données de la recherche »...

3) Une déclinaison locale : Projet Data Paris 8

Laboratoires impliqués dans le projet pilote :

(la liste n'est pas exhaustive et peut être amenée à évoluer)

- CRESPPA UMR 7217, centre de recherches sociologiques et politiques de Paris
- LEGS UMR 8238, laboratoire d'études de genre et sexualité
- IFG Lab EA 353, centre de recherches et d'analyses géopolitiques
- Paragraphe EA 349, sciences de l'information et de la communication, psychologie
- SFL UMR 7023, structures formelles du langage
- CEMTI EA 3388, centre d'études sur les médias, les technologies et l'internationalisation

3) Une déclinaison locale : Projet Data Paris 8

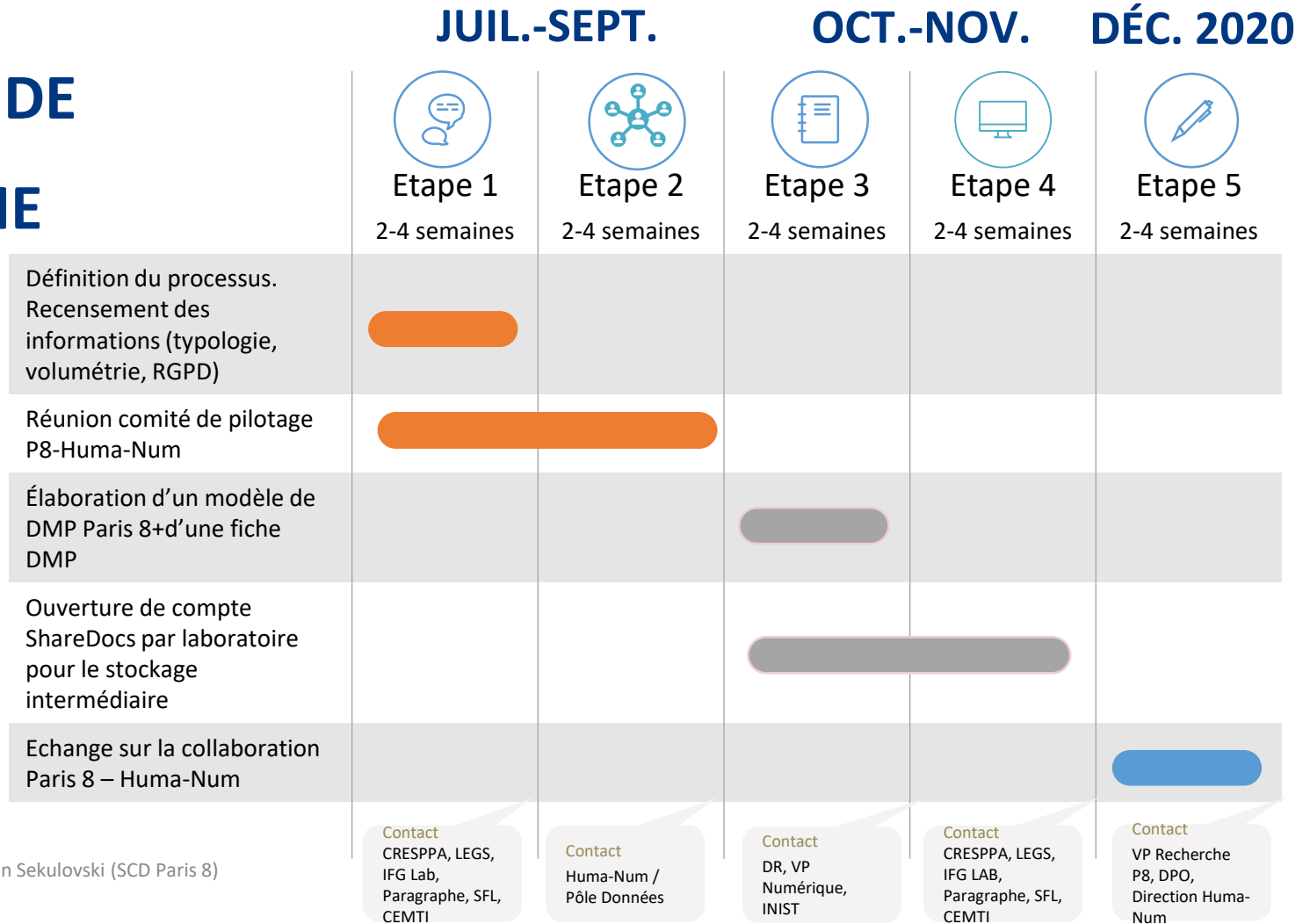
Période du projet : l'année universitaire 2020-2021

Etapas prévisionnelles du projet :

- Identifier les données (typologie, volumétrie)
- Elaborer un modèle de Data Management Plan Paris 8 applicable aux données des laboratoires
- Lancer la collecte des données auprès des laboratoires pilotes
- Ouverture des comptes ShareDocs par laboratoire pour le stockage intermédiaire
- Ouverture des comptes Nakala pour dépôt et stockage de long terme des données
- Constituer des collections de données dans Nakala par laboratoire / projet
- Créer des collections par laboratoire dans la bibliothèque numérique de Paris 8 (Octaviana) pour le signalement / moissonnage des données déposées dans Nakala
- Bilan du projet

3) Une déclinaison locale : Projet Data Paris 8

PROJET DONNÉES DE LA RECHERCHE PARIS 8 « DATA PARIS 8 »



Projet Data Paris 8 | Goran Sekulovski (SCD Paris 8)

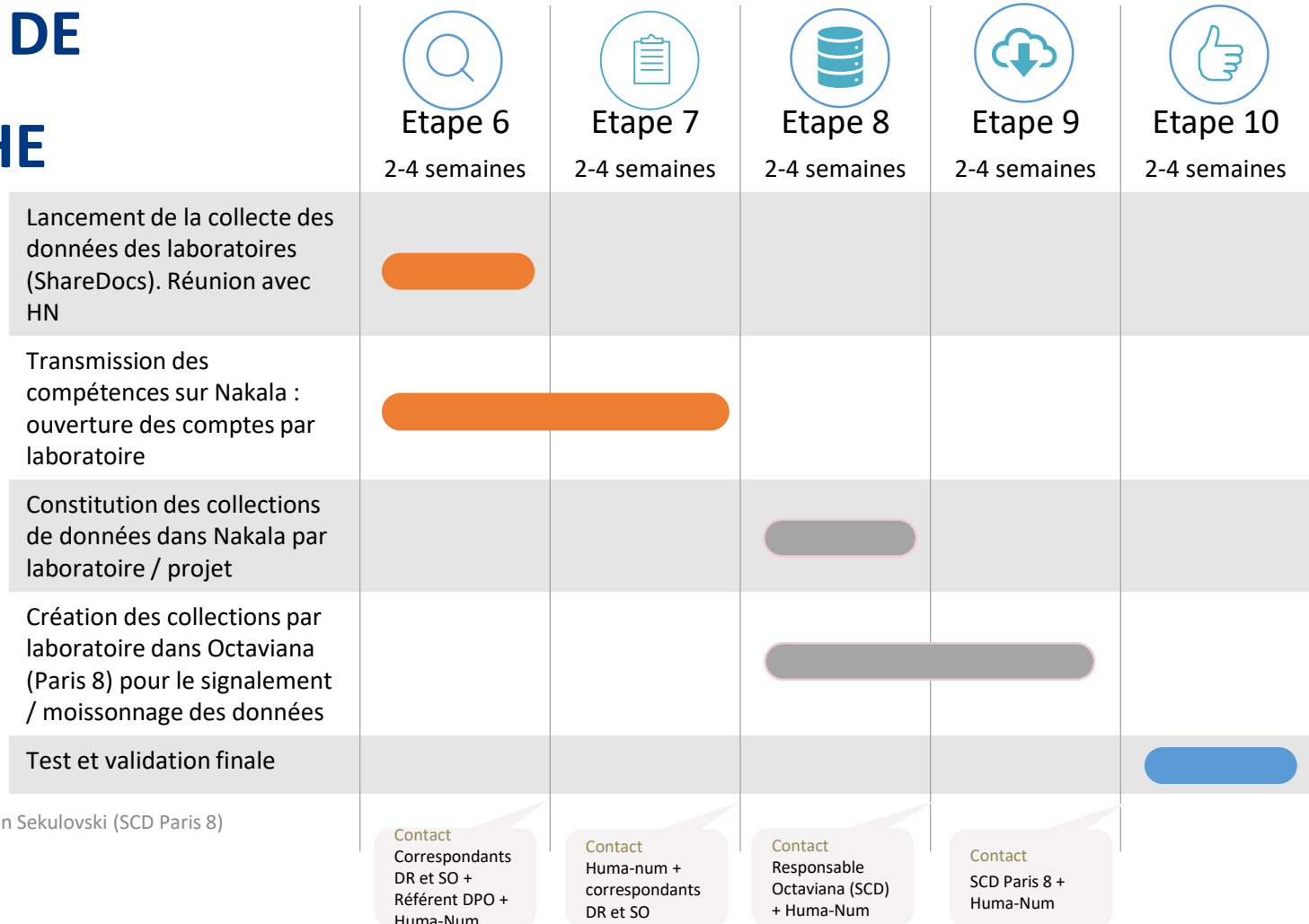
3) Une déclinaison locale : Projet Data Paris 8

PROJET DONNÉES DE LA RECHERCHE PARIS 8 « DATA PARIS 8 »

JAN. 2021-MARS

AVR.-JUN

JUIL.-SEP.



Projet Data Paris 8 | Goran Sekulovski (SCD Paris 8)

3) Une déclinaison locale : Projet Data Paris 8

Le Projet Data Paris 8 permettra de valoriser les données par :

- l'augmentation de leur visibilité via un accès libre
- leur standardisation
- leur interopérabilité
- la facilitation de l'échange des données
- la diversification des métadonnées

3) Une déclinaison locale : Projet Data Paris 8

Document et données : l'université de la transition numérique

Rechercher par titre	Filterer par : Statut	Année de création	Trier par : Date de dépôt (décroissant)
Titre		Date de création	Status
EA 349 : Paragraphe		26/02/2021	privé
UMR 7217 : Centre de recherches sociologiques et politiques de Paris (CRESPPA)		26/02/2021	privé
UMR 8238 : Laboratoire d'études de genre et de sexualité (LEGS)		26/02/2021	privé
EA 3388 : Centre d'études sur les médias, les technologies et l'internationalisation (CEMTI)		26/02/2021	privé
EA 353 : IFG-LAB		26/02/2021	privé
UMR 7023 : Structures formelles du langage (SFL)		26/02/2021	privé
EA 4004 : Cognitions humaine et artificielle (CHArt)		26/02/2021	privé
Data Paris 8		06/04/2021	privé

À propos | Contact | API | OAI | Mentions légales

3) Une déclinaison locale : Projet Data Paris 8

la transition numérique
l'université
Docu

The screenshot shows the Nakala interface for managing a collection. At the top, there is a search bar with the text "Rechercher des données dans Nakala...". Below this, the "Description" section contains a dropdown menu for "Pas d'information de langue". The "Mots-clés" section also has a dropdown menu for "Pas d'information de langue". The "Ajouter d'autres métadonnées" section is highlighted with an orange oval and contains a dropdown menu for "dcterms:isPartOf" with the value "Data Paris 8" selected. The "Type" dropdown is set to "Type" and the language dropdown is set to "Pas d'information de langue". A blue button labeled "Ajouter une métadonnée" is at the bottom of this section. The "Partager les droits de cette collection avec d'autres utilisateurs" section is also highlighted with an orange oval and shows a list of users. The user "Goran Sekulovski" is listed as the "Propriétaire". Below this, there is a search bar for "Ajouter un utilisateur ou une liste d'utilisateurs" with the text "Chercher un utilisateur ou une liste". A user "Samuel Szoniecky" is listed with a gear icon and a close icon. The footer of the page contains links for "À propos", "Contact", "API", "OAI", and "Mentions légales", and a note "Service développé par Huma-Num".

Ajout de métadonnée : dcterms:isPartOf

Attribution des droits

4) Accompagnement à la gestion des données

Accompagner les personnels (enseignants-chercheurs, BIATSS) à Paris 8 en développant une offre de services :

- Formations à la rédaction d'un DMP et ateliers pratiques DMP
- Transmission des compétences sur Nakala (métadonnées, formats, traitement, interopérabilité, exposition, valorisation)
- Aspects juridiques autour des données
- l'Open Access Week
- Le [cycle d'ateliers du numériques](#) au SCD
- Formations doctorales dédiées aux données

4) Accompagnement à la gestion des données

Documents et données : l'université à l'ère de la transition numérique

Janvier 2021

De la collecte sur le terrain à la citation dans les publications : parcours d'une source enregistrée à travers les fonds de la phonothèque de la MMSH

Véronique Ginouvès (MMSH)

Février 2021

Le RGPD appliqué dans le cadre de recherche en sciences humaines et sociales

Fabrice Leclere (Paris 8)

Avril 2021

Nakala : une plateforme pour dépôt et diffusion des données de la recherche

L. Capelli (Huma-Num)/
G. Sekulovski (Paris 8)

Mai 2021

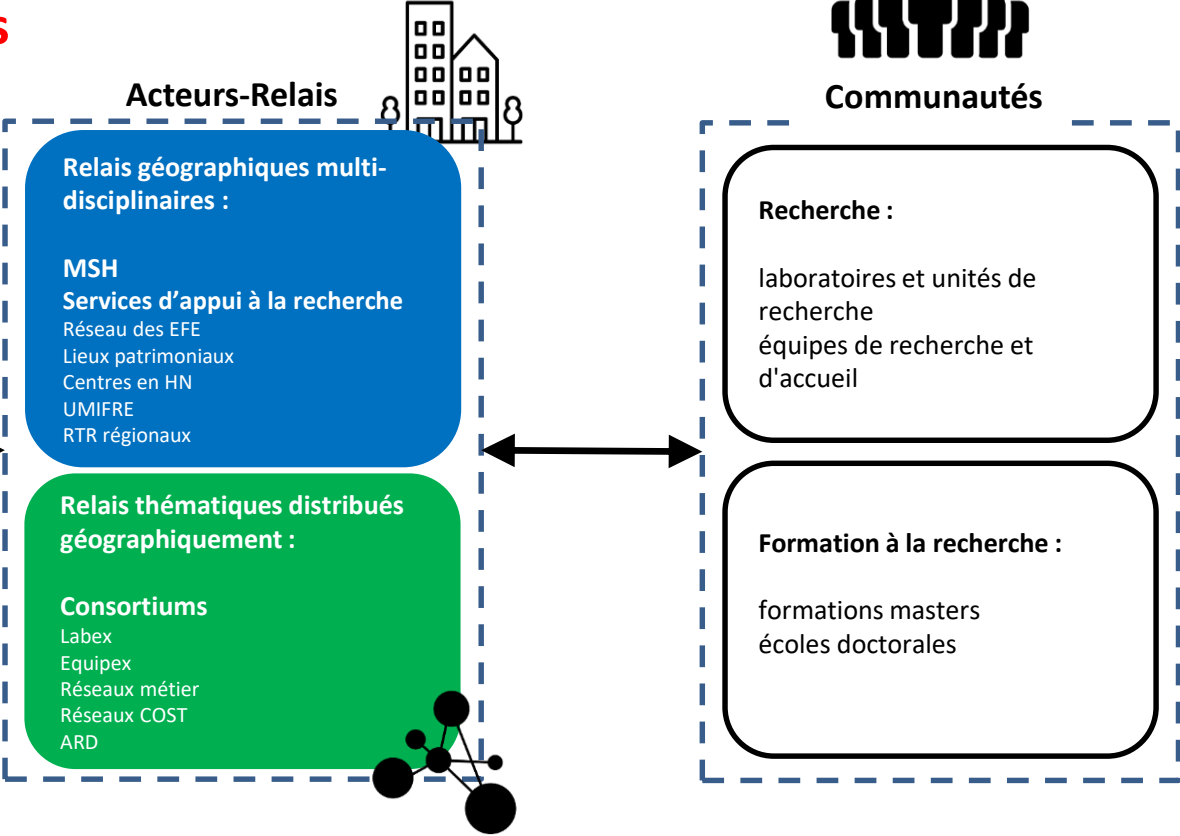
Propriété intellectuelle et droits sur les données de la recherche

Bertrand Warusfel (Paris 8)

Ateliers à destination des enseignants-chercheurs

5) Le partenariat avec Huma-Num

Structurer les usages à l'écoute des pratiques

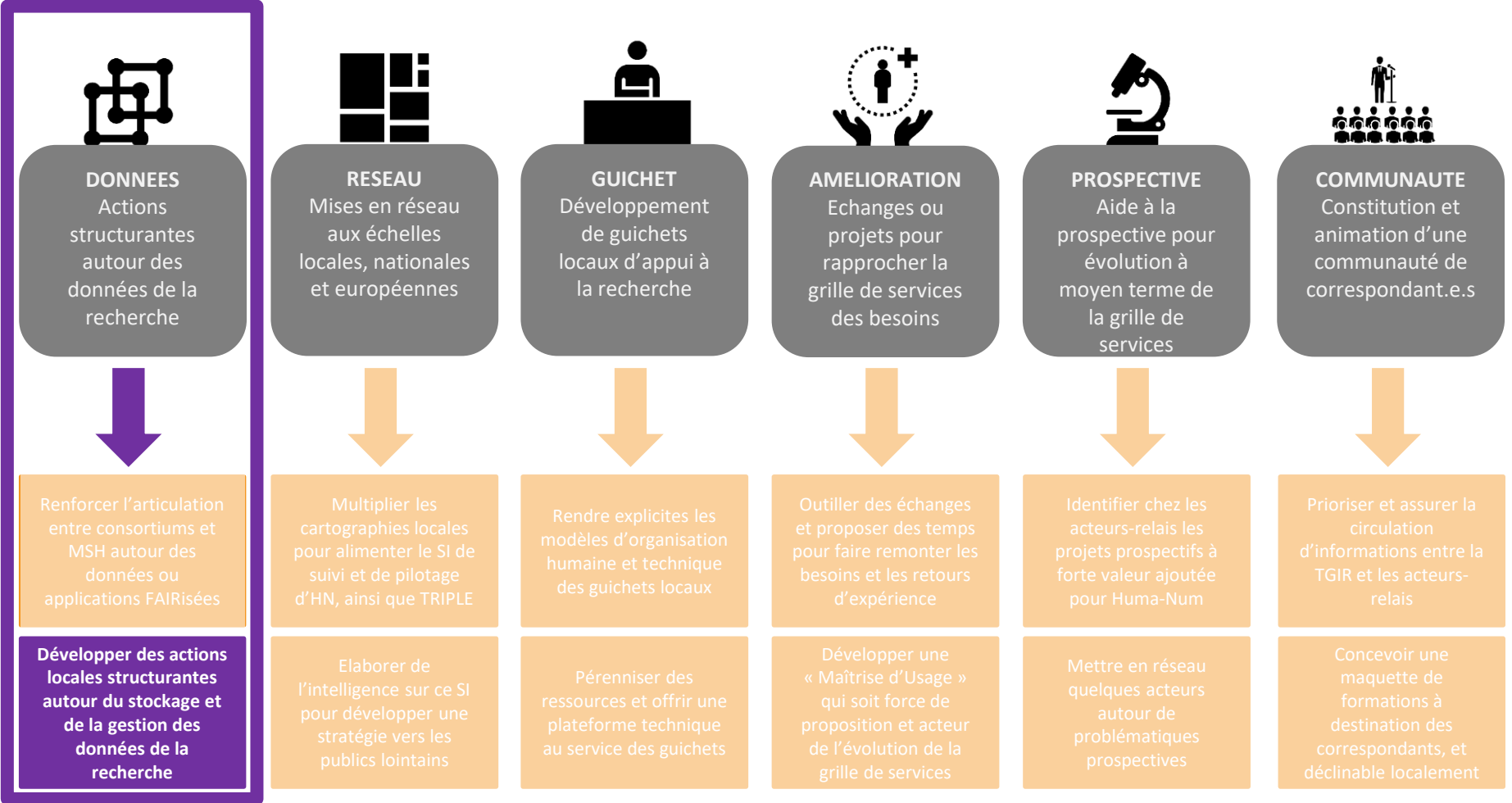


Un cadre stratégique pour Huma-Num : les collaborations avec les acteurs-relais

5) Le partenariat avec Huma-Num

- Les MSH et le RnMSH
 - 22 MSH sur le territoire national
 - Démarche *bottom up*
 - Tour de France des MSH
 - Typologie des collaborations
 - Vers une stratégie concertée avec le RnMSH
- Partenariats spécifiques d'appui à la recherche
 - Data Lab de la BNF
 - Labex Les Passés dans le Présent
- Collaborations en développement
 - Campus Condorcet
 - Le réseau des EFE
 - Université de Paris (Centre HN)
 - *Université Paris 8 : PoC data Paris 8*

5) Le partenariat avec Huma-Num



5) Le partenariat avec Huma-Num

- Valider la pertinence d'un tel partenariat avec au moins deux critères
 - Accompagnement dans la chaîne de services d'Huma-Num
 - FAIRisation des données
- Etudier la faisabilité technico-financière d'un passage à l'échelle éventuel
 - Quel impact sur l'infrastructure de stockage ?
 - Quel impact sur l'organisation interne à la TGIR ?
 - Quel surcoût cela engendrerait, et comment le financer ?
- En profiter pour faire du SCD de Paris 8 un acteur-relais à part entière
 - Valorisation locale de l'offre de services d'Huma-Num
 - Retours d'expériences et de besoins sur ces services

Contacts

Goran Sekulovski, goran.sekulovski@univ-paris8.fr

Johann Holland, johann.holland@huma-num.fr