

l'atelier Données

« DoreMITI »

*groupe de travail inter-réseaux de
la MITI*

*Un exemple de
« compétences collaboratives nécessaires pour
structurer, partager et ouvrir les données de la
recherche. »*

Plan

- Qui suis-je ?
- Présentation de DoReMITI
 - La genèse du Groupe de Travail
 - Objectifs
 - Composition du GT
 - Fonctionnement
 - Les Réalisations
 - Conclusions :
 - difficultés et bénéfices, les ingrédients de réussite

Préambule – Qui suis-je ?

- Maurice Libes IR OSU Pytheas
 - OSU = Observatoire des Sciences de l'Univers = Institut de l'INSU CNRS, regroupe 6 UMR
 - Doctorat en Biologie marine, puis DEA Informatique
 - 1993 - 2017 Ingénieur Système et Réseau, CSSI : Sécurité du Système d'Information
 - Depuis 2018 : En charge de la **constitution d'un portail de données Environnementales « FAIR » pour l'OSU Pytheas**
 - Membre fondateur du réseau « **SIST** » *Séries Interopérables et Système de Traitement.. des données*
 - Réseau technologique soutenu et labellisé par l'INSU
 - <http://sist.cnrs.fr>



NB : C'est au titre du réseau *SIST* que je participe au Groupe de Travail DoReMITI sur les aspects de gestion des données environnementales

DoReMITI : la Génèse ...

- Rencontres des réseaux professionnels de la MITI CNRS (01/2016)
 - Atelier : le jeu « inter-réseaux » : *la donnée de l'acquisition à la valorisation*
 - Présentation du cycle de vie de la donnée
 - Deux questions : besoins à court-terme et enjeux à long-terme
 - 19 inscrits de plusieurs réseaux de la MITI
 - Restitution : https://webcast.in2p3.fr/video/restitution_des_ateliers
 - De nombreux besoins recensés autour de la gestion des données
 - Pour aller plus loin : création d'un GT inter-réseau « *Atelier Données* »
 - Mettre en place une journée thématique sur l'interopérabilité
 - Rédiger un guide pratique sur la traçabilité des activités de recherche
- } NB : ce GT initialement appelé "Atelier Données" a été rebaptisé **DoReMITI** en 2023

Les missions et objectifs du GT Atelier Données :

- Construire et diffuser une vision transversale de la gestion des données
 - ✓ enrichir la pratique de chaque réseau dans le domaine des données
 - ✓ permettre le développement de la complémentarité entre réseaux
- Valoriser l'apport des expériences et expertises « métier » dans une *vision transversale* de gestion de données
- Sensibiliser les communautés professionnelles, au soutien à la gestion des données (organisation de journées thématiques par exemple) ;
- Identifier les problématiques concernant les « data » dans chaque réseau métier.
- Mettre en commun et partager de nouvelles pratiques en réseau et au sein de chaque réseau.

DoReMITI : composition

- 20 membres
- 12 réseaux et « structures »
 - Réseaux MITI : *Calcul, Devlog, Medici, QeR, rBDD, Renatis, Resinfo, RIS*
 - Réseaux Institut : *SIST*
 - Autres réseaux : *Relier*
 - Structures : *DDOR, INIST*
- Des animateurs en binôme – mandat 2 ans:
 - Emmanuelle Morlock & Caroline Martin (Janv. 2016 – Juin 2019)
 - Pierre Brochard & Marie-Claude Quido (Juillet 2019 – Janvier 2023)
 - Pierre Brochard & Violaine Louvet (Fév. 2023)

DoReMITI : Fonctionnement - HowTo ?

- Des moyens de diffusion et communication
 - Un site WEB : <https://gt-atelier-donnees.miti.cnrs.fr/> (technologie Quarto + Gitlab)
 - Une liste de mail : donnees-inter-reseaux@services.cnrs.fr
 - Une chaîne Canal-U pour les vidéos
 - Veille techno : Une Lettre d'information mensuelle
 - <https://gt-atelier-donnees.miti.cnrs.fr/lettres.html>
- Réunions bi mensuelles du comité de pilotage
 - › Définition des actions à venir

DoReMITI : Réalisations

- **Partages d'expériences**

- Séminaires (FAIR, Traçabilité)
- Webinaires (datapaper, qualité des données, wikidata)
- Hackathon (entrepôts de données, TP Dataverse)

- **Ouvrages de Synthèse**

- « Cartographie des actions des réseaux métiers sur la gestion des données »
- « Guide de bonnes pratiques sur la gestion des données de la recherche »
 - <https://mi-gt-donnees.pages.math.unistra.fr/guide/00-introduction.html>

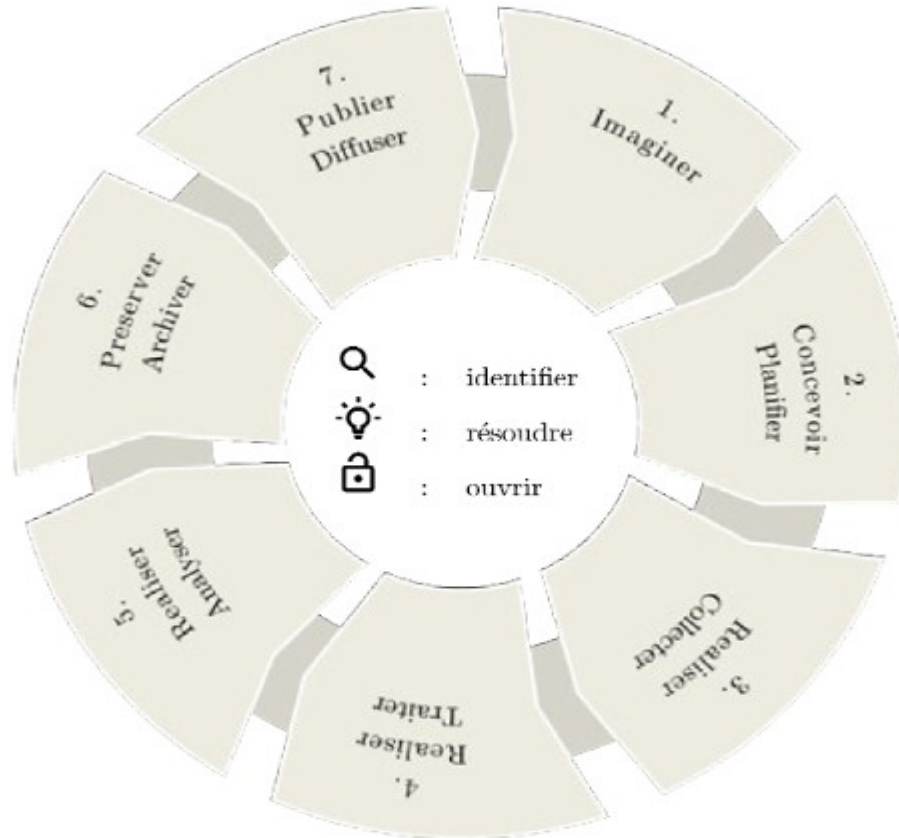
DoReMITI : Réalisations

- **Interopérabilité et pérennisation des données de la recherche. Comment FAIR en pratique ?**
 - ✓ Séminaire, Paris, Novembre 2018, 150 personnes <https://gt-donnees2018.sciencesconf.org/>
- **Comment améliorer le dépôt de données de recherche ? (métadonnées, entrepôt de confiance...)**
 - ✓ Hackathon, Grenoble, Janvier 2020, 50 personnes
- **Data paper, une incitation à la qualification et à la réutilisation des jeux de données**
 - ✓ Webinaire, Novembre 2020, 150 personnes connectées
 - ✓ https://www.canal-u.tv/video/medici/introduction_du_webinaire.58157
- **Qualité des données**
 - ✓ Webinaire juillet 2021
 - ✓ <https://gt-atelier-donnees.miti.cnrs.fr/qualite.html>
- **Foire Aux Questions**
 - ✓ 2 grands thèmes : « entrepôts de données » et les « data papers »
 - ✓ <https://gt-atelier-donnees.miti.cnrs.fr/faq.html>

DoReMITI : Réalisations

- « **Cartographie des actions des réseaux métiers autour de la gestion des données** »
 - ✓ Mai 2017
 - ✓ <https://mi-gt-donnees.pages.math.unistra.fr/site/download/GTInterreseaux-CartoSynteseV6-optimize.pdf>
- « **Guide de bonnes pratiques sur la gestion des données de la recherche** »
 - ✓ V1 : Juillet 2019 -> Février 2021
 - ✓ V2 : Janvier 2023
 - ▶ <https://mi-gt-donnees.pages.math.unistra.fr/guide/00-introduction.html>

Exemple 1 : Cartographie des actions des réseaux métiers autour de la gestion des données (2017)




Définition du cycle de vie des données du GT

Complémentarité des métiers

Exemple 2 : Guide des bonnes pratiques sur la gestion des données de la recherche (2021)



Rechercher dans ce livre ...

1. Imaginer et préparer
 2. Concevoir et planifier
 3. Collecter
 4. Traiter
 5. Analyser
 6. Préserver et archiver
 7. Publier et diffuser
- Conclusion
- Glossaire
- Infrastructures
- Reproductibilité
- Autres guides de bonnes pratiques
- Crédits
- Document pdf 

[contact](#)

Pour formaliser les différentes étapes de gestion des données, nous nous sommes servis du "cycle de vie des données" élaboré au sein de l'Atelier Données. Il s'agit d'un cercle vertueux que l'on peut faire correspondre aux différentes phases d'un projet scientifique.

L'apport des réseaux métiers du CNRS

Dans leurs différentes pratiques, les réseaux métiers du CNRS, regroupés au sein de la [Mission pour les Initiatives Transverses Interdisciplinaires](#) (MITI) ou soutenus par les Instituts sont en première ligne pour participer à ce mouvement d'ouverture et de partage des données. Les personnels des organismes de recherche qui les constituent, œuvrent pour mettre en place de bonnes pratiques de gestion et participent également au processus de production des données scientifiques aux côtés des équipes de recherche. C'est aussi dans ce cadre qu'ils interviennent en appui et soutien à la recherche scientifique.

C'est précisément ce travail de soutien que nous nous proposons de rendre compte dans ce document qui, à travers de nombreux séminaires communications et formations, vise à fournir les meilleures pratiques du moment en matière de gestion des données, et peut ainsi s'apparenter à un "guide de bonnes pratiques".

Certes de nombreux guides existent déjà dans le domaine, mais l'originalité de ce document réside dans son application aux données de la recherche sous l'angle de différents métiers de la recherche. Il fournit donc un point de vue transversal intéressant et traduit les efforts et le soutien mis en place par les personnels d'appui à la recherche au sein des réseaux, dans la gestion et la valorisation des données scientifiques.

Objectifs du guide

Ce guide est la production du groupe de travail inter-réseaux "Atelier Données". Il s'agit d'un groupe composé de plusieurs réseaux métiers de la MITI ([Calcul](#), [Devlog](#), [QeR](#), [rBDD](#), [Renatis](#), [Resinfo](#), [Medici](#)), du [réseau SIST](#), (labellisé par l'INSU et regroupant les gestionnaires de données environnementales), de l'INIST, et de la Direction des données ouvertes de la recherche (DDOR-CNRS ^[1]). [Les activités de cet atelier](#) ont été présentées lors du séminaire SIST 2019.

Difficultés rencontrées liées à l'inter-réseaux

Un exercice de transversalité et de décloisonnement... inhabituels

Nous sommes spécialisés dans un domaine thématique

- Information scientifique et technique
- Science des données
- Science de l'ingénierie

Des Actions diversifiées

- Membres aux profils très différents
- Centres d'intérêt parfois très éloignés

**Nécessité, établissement d'une base d'un langage commun,
comprendre les concepts des autres**



Ce GT est une structure d'échanges, de confrontation d'idées, de partage d'expériences, ...

avec des personnes de profils et métiers différents animés par le besoin de procéder au mieux à la gestion des données et d'en diffuser les bonnes pratiques métiers

Meilleure connaissance des autres métiers de la recherche

- Connaissance et prise en compte de leurs besoins et/ou de leurs contraintes
- De leurs langages, vocabulaires et standards

Élargissement du domaine de compétence

- Drainer des idées et des concepts nouveaux provenant de sphère professionnelles différentes

Bénéfices de l'inter-réseaux liés à la thématique

Un véritable plus dans le domaine de la science ouverte

- Connaître les sigles / concepts
- Connaître le paysage national (IR, DataGouv, CoSO, EOSC, RDA, ...)
- Connaître les solutions techniques possibles



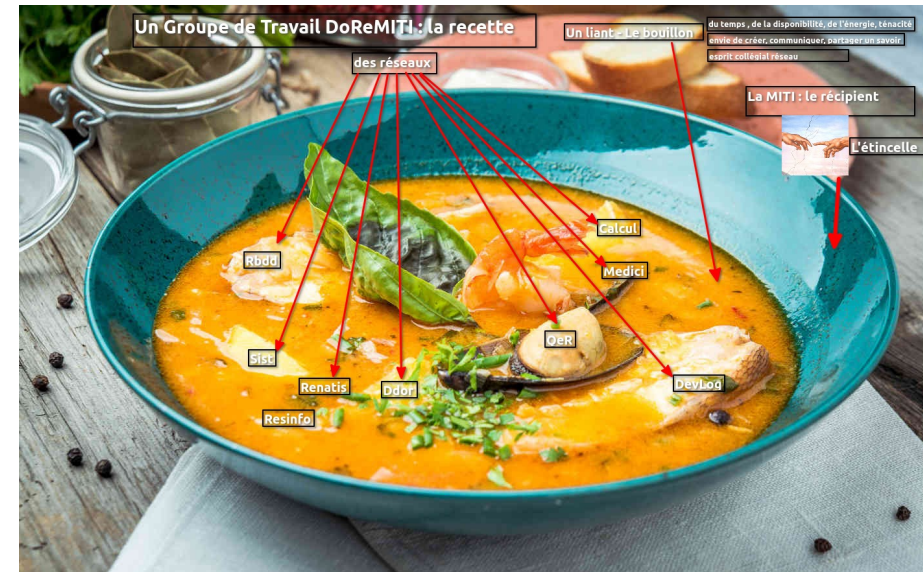
Un véritable plus pour accompagner les chercheurs



Conclusions

*“ La folie, c'est de faire toujours la même chose et de s'attendre à ce que les choses changent” ...
X?*

- **No pain ... No gain** : une mise en commun difficile mais utile
- **Une belle aventure collective basée sur**
 - L'envie de partager un savoir... et d'apprendre des autres
 - Un « esprit réseau » collégial et amical
 - Un pilotage induit par des objectifs partagés
 - Un support Institutionnel de la MITI, nécessaire
- **Beaucoup de gain et d'apprentissage**
 - Effet drainant et énergisant dans les réseaux
- **Nécessite du temps de la disponibilité de l'énergie, de la ténacité**



• Les réseaux sont des vecteurs de production nécessaires, des tremplins pour la connaissance

Contacts :

[gt-donnees-inter-reseaux @groupes.renater.fr](mailto:gt-donnees-inter-reseaux@groupes.renater.fr)

Site web :

<https://gt-atelier-donnees.miti.cnrs.fr/>

Liste de discussion :

donnees-inter-reseaux@services.cnrs.fr

