



Besoins et attentes des Communautés? Le cas du CNRS

Atelier EOSC au MESRI

4 et 5 Février 2021

Sylvie ROUSSET , sylvie.rousset@cnrs.fr

Directrice DDOR – Direction des Données Ouvertes de la Recherche

Implication du CNRS dans la gouvernance de EOSC

- Membre de l'association AISBL depuis fin 2020
- Participation à la première AG 17/12/2020
- Suzanne Dumouchel membre du board of directors depuis le 17 décembre 2020

EOSC – CNRS – Journée du 22 janvier 2020 organisé par Volker Beckmann, coordinateur du GT EOSC CNRS

Participation du CNRS à l'EOSC

- **Projets où le CNRS est partie prenante** : **EOSC-pilot** (INSU, IN2P3, InS2I), **EOSC-hub** (IN2P3), **EOSC-Pillar** (INSU, IN2P3, InEE, InS2I, InSHS, InSB, InSMI), **XDC** (IN2P3), **TRIPLE** (InSHS), **ESCAPE** (INSU, IN2P3), **ENVRI-FAIR** (INSU), **SSHOC** (InSHS), **AENEAS** (INSU)
- **Projets où le CNRS participe** : **ELIXIR** (InSB dans le futur?), **PHIDIAS** (INSU), **GO-FAIR** (InSHS), **OpenAIRE NEXUS** (InSHS, DIST), **EGI** (IN2P3), **PaNOSC** (InP), **ExPaNDS** (InP)

EOSC – Domaines d'intérêt – état des lieux des Instituts

Institut	Trusted repositories	FAIRisation of data, DMPs	Service uptake/ consolidation	(heterogeneous) e-infrastructures, HTC/HPC	Grid/cloud computing	Software repositories	training
IN2P3			X	X	X		
INC	X	X		X			
INEE	X	X	X				X
INP	X	X		X			
INS2I			X	X		X	
INSB	X	X	X				
INSHS	X	X	X				X
INSIS	X	X	X				
INSMI	X						
INSU	X	X	X	X	X	X	
DIST	X	X	X			X	X

EOSC – European Open Science Cloud

Accompagner au niveau européen la montée en puissance sur les données de la recherche

- Cela concerne les services mais aussi tous les contenus scientifiques (publications, données, logiciels, méthodes)
- Tous les instituts du CNRS concernés dans des projets EOSC

Communauté des infrastructures de calculs, mais aussi intégrer la dimension Open Science :

Composer avec d'autres structures

- au niveau international pour les données: RDA avec sa déclinaison RDA France (pilotage CNRS)
- au niveau européen pour la science ouverte : OpenAIRE, projet OpenAIRE Nexus avec le CCSD (HAL)

La POLITIQUE SCIENCE OUVERTE DU CNRS

<https://www.science-ouverte.cnrs.fr/>



Feuille de route science ouverte du CNRS –
Novembre 2019 [Consulter le PDF \(version française\)](#)



Plan "Données de la recherche" du CNRS -
Novembre 2020
[Consulter le PDF \(version française\)](#)

UN PLAN DONNEES AU CNRS

Constat de la mission MICADO dans le Livre Blanc 2018 de promouvoir une « culture de la donnée » et de doter le CNRS d'une stratégie des données

Mise en place d'une gouvernance pour EOSC (European open science cloud)

- au niveau international, une association AISBL avec le CNRS comme membre (2020),
- au niveau français un « EOSC français » (DGRI)

Accélération des actions pour développer la Science Ouverte :

- Plan National Science Ouverte 2018 et demande d'un administrateur des données dans les établissements
- Feuille de Route Science Ouverte CNRS 2019 avec un volet « données de la recherche»

Enjeu scientifique fort pour la gestion et le partage des données au CNRS

- Voir les motivations scientifiques, juridiques et de moyens

MOTIVATIONS

Rendre la recherche plus efficace et non redondante (pas de duplication inutile)

Assurer l'intégrité scientifique (reproductibilité et validation des résultats)

Etre en capacité de réutiliser les données même sans en être à l'origine

Croiser les données pour favoriser de nouvelles analyses, voire faire émerger de nouvelles thématiques

Satisfaire le cadre légal d'ouverture des données a priori : « Ouvert autant que possible, fermé autant que nécessaire » (loi pour une république numérique 2016)

Mutualiser et rationaliser les infrastructures informatiques, les moyens RH et identifier les nouveaux métiers (datastewardship ...)

LES DONNÉES DE LA RECHERCHE

Une donnée doit être ouverte ou protégée:

« ouvert autant que possible, fermé autant que nécessaire ».

Toutes les données de la recherche n'ont pas vocation à être ouvertes ou divulguées. Il existe des exceptions évidentes telles que les données spécifiques à caractère confidentiel, que cela soit du fait de leur caractère personnel, pour des raisons de concurrence industrielle ou pour des intérêts fondamentaux ou réglementaires des États.

La décision d'ouverture ou de protection des données de la recherche doit être prise avec les services compétents du CNRS :

- les Services Partenariat Valorisation pour la propriété intellectuelle,
- la Délégation à la protection des données pour les données à caractère personnel et de
- la Direction de la sûreté pour les questions relatives à la souveraineté.

PLAN D' ACTIONS : VALORISER LES DONNEES

Promouvoir une « culture de la donnée », ainsi que les infrastructures, les outils et services, en incluant la question des moyens (humains et financiers) et de la formation

Travailler avec les instituts en prenant en compte les spécificités disciplinaires et l'existant, et en s'appuyant sur une gouvernance transverse des données

Soutenir les communautés scientifiques pour définir les éléments spécifiques de la gestion de leurs données à FAIRiser

Encourager le dépôt des données dans des entrepôts thématiques en priorité

Valoriser les données en lien avec les publications, et les « datapapers »

PLAN D' ACTIONS : OUTILS ET SERVICES

Proposer **un annuaire** pour identifier les outils et services existants (Cat-OPIDoR)

Accompagner la réalisation des **Plan de Gestion des Données** (PGD- OPIDoR)

Aider à **la certification des dispositifs** de prise en charge des données (ex. CoreTrustSeal)

S'assurer que toutes les données à FAIRiser trouvent **un entrepôt des données** (métadonnées + jeux de données) et **un espace physique de stockage**

Fournir des **identifiants pérennes** pour les jeux de données (DOI – Datacite – INIST)

Encourager la FAIRisation des données avec les **Cahiers de Laboratoire Electroniques**

Travailler sur les **portails de référencement** pour accéder aux données disponibles

PLAN D' ACTIONS : RH ET FORMATION

Recenser les nouveaux métiers liés au traitement des données

Modifier les référentiels de postes IT (Referens) pour inclure les nouvelles compétences

Nourrir les référentiels d'évaluation pour inclure l'implication dans la gestion et le partage des données

Former les personnels à la culture des données et aux nouveaux outils et services par la formation permanente notamment (très forte demande)

Faire connaître les succès du partage des données aux communautés réticentes

Une direction fonctionnelle des données ouvertes de la recherche : la DDOR

Ecosystème des données de la recherche :

Infrastructures numériques, Big Data, services et principes FAIR, Science Ouverte.

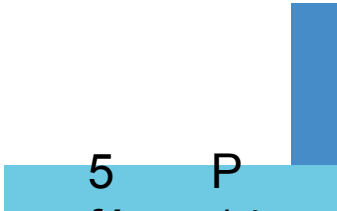
Rapprocher deux communautés IST et Calcul : des grandes différences y compris le vocabulaire mais les sujets sont liés et on ne peut pas les dissocier.

Interlocuteur pour les relations avec les partenaires de l'ESR et EOSC et RDA au niveau européen et international

Direction DDOR à l'intérieur de la DGDS avec le périmètre suivant:

- Calcul (calcul haute performance, grilles, "cloud", données massives, représentation du CNRS dans les "objets" nationaux : INFRANUM (Infrastructures numériques, DGRI), GENCI, RENATER, ...)
- Données (Stockage, FAIRisation, Annuaire, DMP, DOI, Aide à la certification des entrepôts, représentation CoSO, RDA, OPenAIRE, EOSC)
- Publications (Négociations et politique documentaire, Archives ouvertes, Nouvelles plateforme d'édition, Pilotage GIS FNSO, CoSO)

Administratrice des données de la recherche (directrice de la DDOR)



EOSC – SRIA Priorities

Additional priorities

- The licenses, usage terms and conditions and user authentication and authorisation methods set by service providers (machine readable)
- Foster skills and training
- A common European evaluating system needs to be implemented. If Open science is recognized as a premise for good science, the rewarding system needs to be updated.
- Urgent, additional resources (financial, but also skills and training) should be dedicated specifically to communities with less developed or with no community standards to mitigate the risk of the EOSC becoming inaccessible to the majority of researchers within academic institutions.