

CNRS et EOSC – opportunitéS ?

Journée EOSC au CNRS

Auditorium Marie Curie - 22 Janvier 2020

ALAIN SCHUHL - DGDS

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)

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
- ➤ au niveau européen pour la science ouverte : OpenAIRE



CNRS – Construction d'un plan données (2020)

Acteurs réunis autour du DGDS

- ➤ MiCaDo mission calcul données au CNRS (inter-instituts)
- ➤ DIST → mise en œuvre de la feuille de route science ouverte du CNRS
- ➤ Infrastructures de recherche productrices de données (TGIR et IR)
- > Particularité des infrastructures SHS

Dans la continuité de la Feuille de route Science Ouverte du CNRS (Axe 2 données de recherche) annoncée par Antoine Petit (JNSO 2019) et Alain Schuhl (Montréal) le 18 novembre 2019



Feuille de route du CNRS pour la Science Ouverte

- 1. 100% de publications en accès ouvert et ré-utilisables
- 2. Des données de la recherche FAIR-isées
- 3. Promouvoir l'utilisation d'outils pour l'analyse et la fouille des textes et de données
- 4. L'évaluation individuelle des chercheurs



Publications: 100% accessibles et ré-utilisables

- Soutien à l'archive ouverte nationale HAL et incitation à y déposer les publications scientifiques :
 - Soutien à HAL, moyens exceptionnels du CNRS en 2019 et modèle économique validé ensuite
 - Obligation du dépôt dans HAL : CRAC 2019 alimenté par HAL (Ribac 2020)
 - Simplification, moissonnage, interopérable autres archives ouvertes
 - Fonctionnement avec nos partenaires universitaires, écoles et organismes
- Dépôt dans les serveurs de preprint
 - Serveurs de preprint hébergeant des manuscrits soumis à des revues
 - Diffusion rapide en accès ouvert via des plateformes à but non lucratif
- Négociations avec les éditeurs conventionnels
 - Modèle de transition avec une diminution du coût des abonnements
 - Conserver les droits de ré-utilisation et de fouille sur les textes
 - Obtenir l'archivage des textes version éditeur dans les archives ouvertes
- Les plateformes d'édition électroniques « vertueuses »
 - Plateformes de publications à but non lucratif dans la plupart des cas
 - Plateformes de publications en accès ouvert proposant différents types d'objets
- Création de nouvelles revues avec d'autres organismes étrangers (G6)



FAIR-isation des données de la recherche Faciles à trouver, Accessibles, Interopérables, Réutilisables

- Développer une culture FAIR de la gestion et du partage des données chez tous les acteurs du cycle de vie de la donnée.
- Placer les instituts au centre de la stratégie, car les pratiques des communautés scientifiques sont différentes d'une discipline à l'autre.
- Aider les infrastructures de recherche à mettre en place une politique des données
- Soutenir et accompagner les entrepôts de données y compris pour les « petites data » (aussi appelée longue traîne des données)
- Accompagner les chercheurs dans les outils de gestion des données et le dépôt conjoint publication/données, notamment grâce à l'outil DMP OPIDOR pour remplir les plans de gestion des données, et l'attribution des DOI via Datacite (INIST)









Fouille et analyse des textes et données

- Faciliter la fouille et l'analyse des résultats de la science avec le développement des infrastructures, des outils et des compétences permettant l'analyse de contenus scientifiques en toute indépendance
- Soutenir les infrastructures permettant l'analyse des contenus : ISTEX, VisaTM etc..
- Développer l'usage d'outils et de techniques de traitement : établissement d'une trousse à outils TDM à l'usage des chercheurs de toutes disciplines
- Veiller au cadre législatif qui doit autoriser la fouille des textes et données, en relation avec la loi numérique (2016) et la directive européenne (mars 2019) : accompagner, traduire, informer



Evaluation et Science Ouverte

Généralisation de 4 principes communs pour les évaluations individuelles

- 1.Ce sont les résultats eux-mêmes qui doivent être évalués, et non pas le fait qu'ils aient pu être publiés dans une revue prestigieuse ou autre media réputé
- 2. Pour chacune des productions citées dans les dossiers d'évaluation, es chercheurs doivent en expliquer la portée, l'impact, et la contribution personnelle qu'ils y ont apportée
- 3. Tous les types de production doivent pouvoir être des éléments de l'évaluation
- 4. Toutes les productions citées dans les dossiers d'évaluation doivent être accessibles dans HAL ou éventuellement dans une autre archive ouverte*

*Trois exceptions à cette règle sont recevables :

- Les résultats trop récents peuvent être sous embargo. Auquel cas ils doivent quand même avoir été déposés dans HAL, avec une durée d'embargo ne dépassant pas ceux prévus par la loi (6 mois en STM, et 12 mois pour les SHS). Ils sont alors fournis par un lien privé dans HAL (ou alors dans le dossier).
- 2. Pour les recrutements, cette règle ne peut pas être absolue pour les candidats exerçant à l'étranger dans des institutions étrangères ou internationales, ou des institutions privées.
- 3. Le type de production peut ne pas être accepté dans HAL.



Feuille de route du CNRS pour la Science Ouverte : c'est aussi

- 100% de publications en accès ouvert et ré-utilisables
- Des données de la recherche FAIR-isées
- Promouvoir l'utilisation d'outils pour l'analyse et la fouille des textes et de données
- L'évaluation individuelle des chercheurs
- Un plan de formation et compétences
- Une stratégie internationale
- Un plan d'implication des chercheurs et des chercheuses



EOSC – La genèse de cette journée

Signature EOSC?

- **≻**Opportunité
- **≻**Risque
- **≻**Faiblesse
 - Lettre de mission Volker Beckmann

Attendus:

- ➤ Recueil avis tous les acteurs du siège
- Recenser l'engagement actuel du CNRS dans l'EOSC
- ➤ Déterminer les besoins de services pouvant relever de l'EOSC
- ➤ Quelles communautés pourraient être concernées (en dehors de celles qui le sont déjà) ?
- Comment assurer une collaboration avec l'existant : RDA France Comité pour la science ouverte, Go FAIR et les autres parties prenantes de l'EOSC en France ?
- ➤ Quels modèles économiques et de gouvernance possibles ?



Construction de l'EOSC : qu'est-ce que c'est?

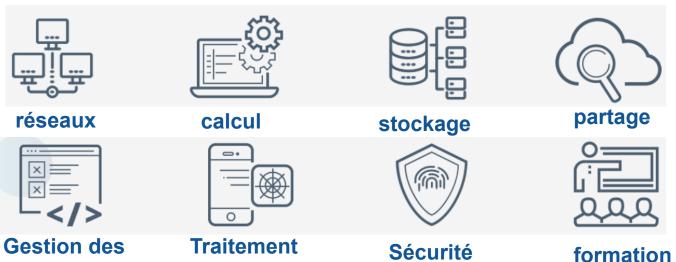
Une gouvernance en construction :

- ➤ Comité de pilotage (Laurent Crouzet, MESRI)
- ➤ Comité exécutif (JF Abramatic)
- ➤ 5 groupes de travail → participants français

et analyse

- ➤ Rules of Participation (O. Hologne, INRAE)
- FAIR (F. Genova, CNRS DIST INSU/RDA)
- Architecture (J.F. Abramatic INRIA, J. Pansanel, CNRS IN2P3)
- ➤ Landscape/Paysage (V. Beckmann CNRS IN2P3)
- ➤ Sustainability/viabilité (JY Berthou, INRIA)

Un portail à alimenter (EOSC portal)





données

EOSC – Projets CNRS – état des lieux des Instituts

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), ENVRIFAIR (INSU), SSHOC (InSHS), AENEAS (INSU)
- Projets où le CNRS participe : ELIXIR (InSB dans le futur?), PHIDIAS (INSU), GO-FAIR (InSHS), OpenAIRE (InSHS, DIST), EGI (IN2P3), PaNOSC (InP), ExPaNDS (InP)
- Objets d'intérêt pour l'EOSC dans les Instituts : TGIR ESRF, Soleil, ILL, IFB
 - Prace



EOSC – Projets CNRS – état des lieux des Instituts

Participation du CNRS à l'EOSC ?

| | | Projets EOSC | | | | | | Related to EOSC | | | | | | | | |
|-------|-------------------------------|----------------|--------------|-----------------|-----|--------|--------|-----------------|-------|--------|--------|--------|-------------|--------------|-----|----|
| | | EOSC- pilot | EOSC- hub | EOSC- Pillar | XDC | TRIPLE | ESCAPE | ENVRI- FAIR | SSHOC | AENEAS | ELIXIR | PHIDIA | GO- FAIR | OpenAIR E | EGI | WG |
| | Nb partenaires français | 4 | 4 | 6 | 1 | 1 | 1 | 5 | 1 | 1 | 4 | 3 | 1 | 4 | 1 | 3 |
| IN2P3 | 6 | Χ | X | X | Χ | | Χ | | | | | | | | Χ | 2 |
| INC | 0 | | | | | | | | | | | | | | | |
| INEE | 2 | | | Χ | | | | | | | | | Χ | | | 1 |
| INP | 0 | | | | | | | | | | | | | | | |
| INS2I | 2 | Χ | | X | | | | | | | | | | | | |
| INSB | 2 | | | Χ | | | | | | | X | | | | | |
| INSHS | 5 | | | Χ | | Χ | | | Χ | | | | Χ | Χ | | 1 |
| INSIS | 0 | | | | | | | | | | | | | | | |
| INSMI | 1 | | | Χ | | | | | | | | | | | | |
| INSU | 7 | Χ | | Χ | | | Χ | Χ | | Χ | | Χ | Χ | | | 1 |
| DIST | 1 | | | | | | | | | | | | | Χ | | 1 |
| CNRS | 15 | Χ | Χ | Χ | Χ | Χ | Χ | Χ | Χ | Χ | Х | Χ | Χ | Χ | Χ | 6 |



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 | | | Χ | X | X | | |
| INC | Χ | X | | X | | | |
| INEE | X | X | X | | | | X |
| INP | Χ | X | | X | | | |
| INS2I | | | X | X | | X | |
| INSB | Χ | X | X | | | | |
| INSHS | X | X | X | | | | X |
| INSIS | Χ | X | Χ | | | | |
| INSMI | X | | | | | | |
| INSU | X | X | X | X | X | X | |
| DIST | X | X | X | | | X | X |



Appels INFRAEOSC 2020 – Horizon 2020

https://www.horizon2020.gouv.fr/cid148149/appels-infraeosc-2020.html

Appel INFRAEOSC03: « Integration and consolidation of the existing pan-European access mechanism to public research infrastructures and commercial services through the EOSC Portal"

- a. Exploitation, maintenance et amélioration du portail EOSC (site Web, catalogue des services et marché)
- b. Favoriser et permettre la composabilité de services sécurisés
- c. Expérience améliorée de l'utilisateur grâce aux techniques d'intelligence artificielle
- d. Élargir la base d'utilisateurs de l'EOSC
- e. Élargir l'offre de services avec des services commerciaux
- f. Activités de soutien
- i. Sensibilisation et compétences
- ii. Soutien à la contribution de la Research Data Alliance (RDA) à l'EOSC





Ouverture: 19.11.2019

Clôture : étape 1 le 22.04.2020

Budget: 64,9 millions €



- a1. Ressources distribuées et cloud computing (lien avec EGI laaS, PaaS et services Saa)
- a2. Services pour les données de recherche
- a3. Services pour l'accès ouvert et l'accès à la production scientifique (lien avec Openaire)
- a4. Services au-delà d'internet
- a5. Services et ressources de fournisseurs de données du secteur privé
- a6. Autres sujets de recherche fournissant des services







