

Centre de Calcul
de l'Institut National de Physique Nucléaire
et de Physique des Particules



Portail pour le Plan de Gestion de Données

15èmes Journées Informatiques IN2P3/IRFU

Gautier Debaecker, Gino Marchetti

Aperçu



- Le Plan de Gestion des Données
- **❖** Le portail DMP de l'IN2P3
- Développements accomplis
 - Personnalisation du portail
 - Gestion des utilisateurs
 - Administration de l'outil
- Conclusions et travaux futurs



Le Plan de Gestion des Données

Fonction et Besoins

PGD: définition



❖ Document synthétique qui aide à organiser/anticiper les étapes du cycle de vie de la donnée

- Exprime le point de vue de l'expérience sur ses données
- Propose des indications de gestion pour chaque jeu de données depuis leur création/collecte jusqu'à leur partage/archivage
- ➤ Plus d'infos dans https://doranum.fr/plan-gestion-donnees-dmp/

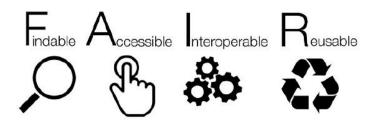
- Document évolutif qui s'adapte aux exigences de l'expérience
 - ❖ Il aura besoin de :
 - Révisions régulières
 - Plusieurs contributeurs



PGD: fonction



- ❖ Plan de Gestion des Données = PGD = Pour Générer du Dialogue
 - > Point de départ pour échanger avec le centre hébergeur sur les besoins en ressources
 - Unique à l'expérience, l'hébergeur s'adapte
- Il existe diffèrents modèles adaptés a diffèrents projets
 - Ex: Horizon 2020, cirad, etc...
- Standardisé par les « RDA DMP Common Standard »
- S'inscrit dans le standard « FAIR » de la Science Ouverte



faciles à trouver (F), accessibles (A), interopérables (I) et réutilisables (R)

Le PGD DMP est nécessaire au CC-IN2P3



En 2024, 219 collaborations avec du stockage alloué

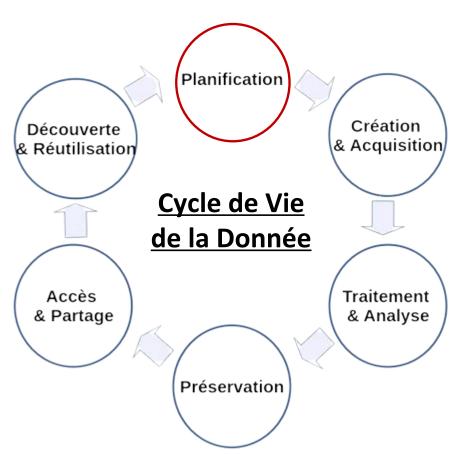
Plus de 360 Po de données

Gestion du cycle de vie des données

- Réattribution des données « orphelines »
- Programmation des choix technologiques
- Prévision d'achats

Quel document ?

- ❖ Si vous avez déjà un DMP, dialoguons!
- ❖ Si vous ne l'avez pas, nous avons un outil pour le rédiger





Le portail DMP de l'IN2P3

Choix technologique

Outils existants





Digital Curation Centre (DCC) in FR



Leibniz Institute for Astrophysics Potsdam (AIP)
Department of Information Sciences of the University
of Applied Sciences Potsdam



Digital Curation Centre (DCC) in UK



California Digital Library (CDL)

Choix de RDMO



Avantages

- Intégration de la notion de dataset
- ➤ La gestion de plusieurs langues
- ➤ Gestion des questionnaires "universels" ET customisés
- Support du modèle commun maDMP du RDA
- Modulable
- > Implémentation de fonctionnalités machine-actionnable

Mis en place depuis 2017

- Proposé aux collaborations volontaires
- Accompagnées durant la rédaction et validation du DMP



https://rdmorganiser.github.io/

Rédaction du DMP



Actualité:

- > L'IN2P3 demande un DMP pour chaque collaboration supportée
- > Lors de la demande des ressources: soumission et/ou révision
- ❖ Il ne sera pas possible d'accompagner « personnellement » les 200+ collaborations

Objectif:

- L'outil doit accompagner les utilisateurs:
 - Apparence « IN2P3 »
 - Adapter le questionnaire
 - Fournir une documentation associée

Développements accomplis



- Personnalisation du portail
- Accompagnement des utilisateurs
- Provisionnement/authentification des comptes
- Formalisation du projet



Customisation graphique du portail

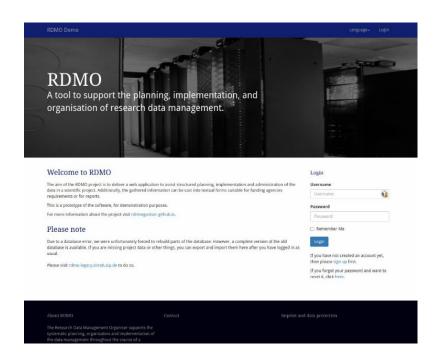
De RDMO à DMP IN2P3

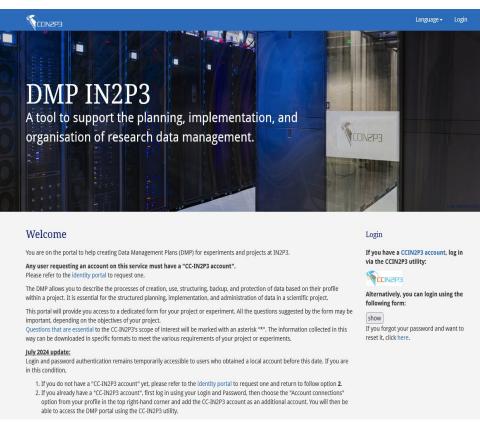
Thème personnalisé

CCIN2P3

https://dmp.in2p3.fr/

- ❖ Adaptation des CSS à la chartes graphique du CNRS/IN2P3
- Ajout de photos des projets hébergés par le CC
- ❖ Lien vers la doc + support utilisateurs + CC ID
- Ajout du logo CC





Nouvelle traduction française

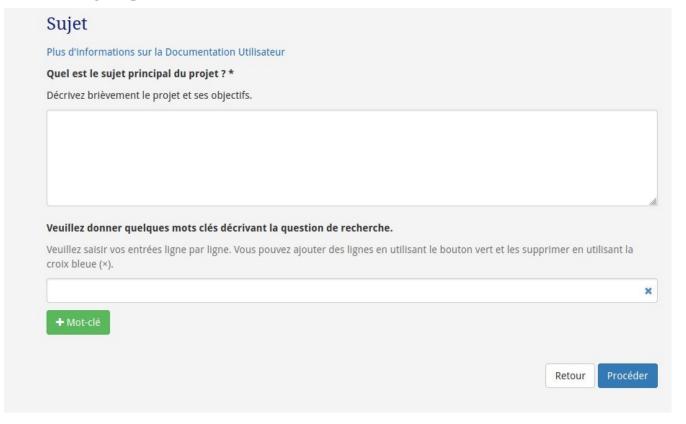


- 1) Première traduction via script
- 2) Relecture et correction par un humain
- 3) Proposition de la traduction aux développeurs de RDMO (intégrée à la distribution officielle depuis février!)

Tri des questions



- Choix de questions essentielles signalées par un * (35/123)
 - Quoi? Qui ? Quand ? Comment ? Où ?
- Retrait de questions jugées superflues





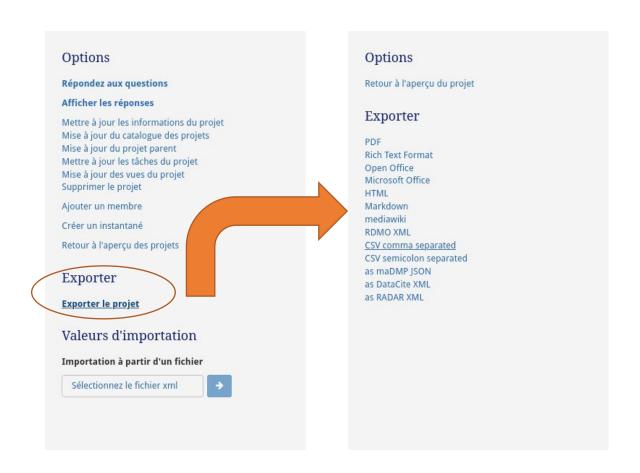
Accompagnement des utilisateurs

Documentation

Ajout de fonctionnalités

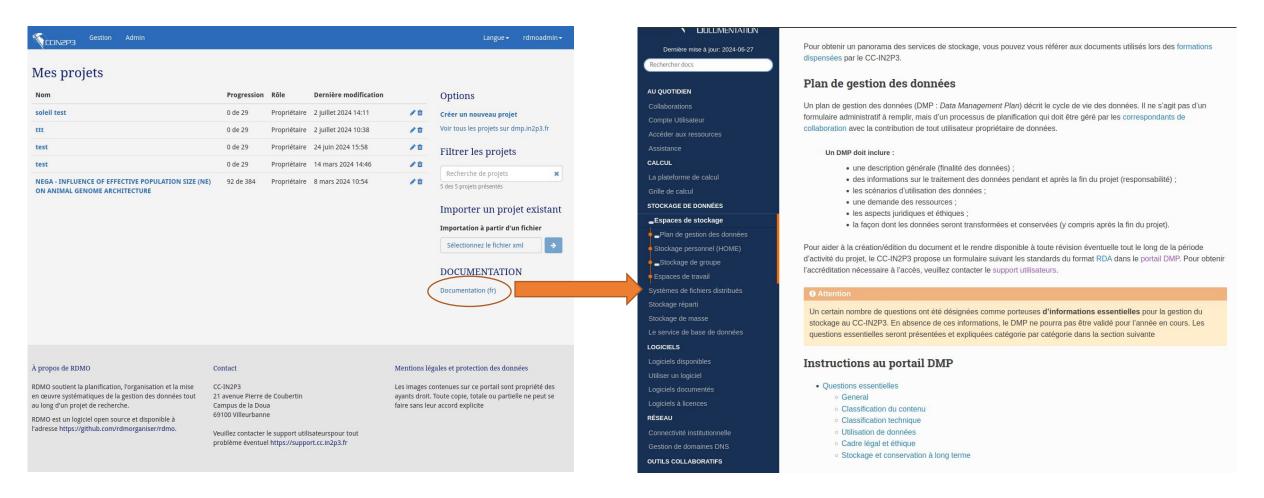


- Clarification de l'export du projet
- **❖** Ajout de l'export PDF des DMP
- Ajout des interactions mail
 - > Demande de nouveau mdp
 - > Invitation à rejoindre un projet



Ajout d'un lien vers la doc directement sur la page du projet CEINZPE

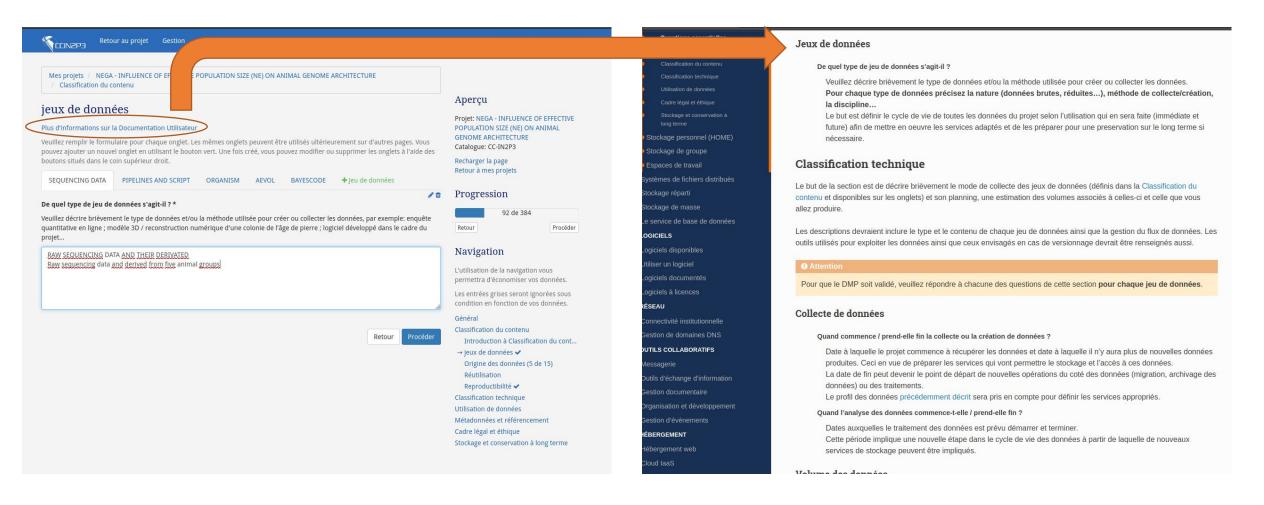
https://doc.cc.in2p3.fr/fr/Data-storage/storage-areas.html#dmp



Lien vers la doc dans les questions (Userdoc à jour)



https://doc.cc.in2p3.fr/fr/Data-storage/dmp-questions.html





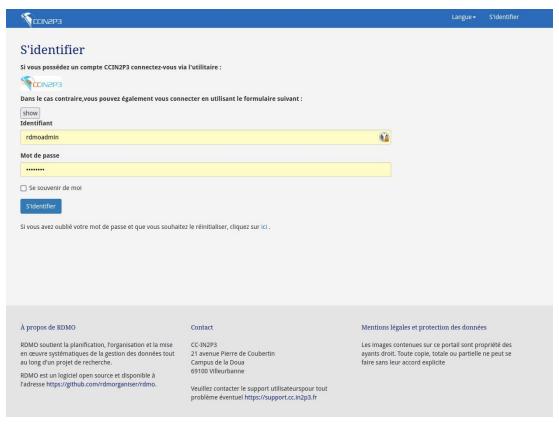
Comptes utilisateurs

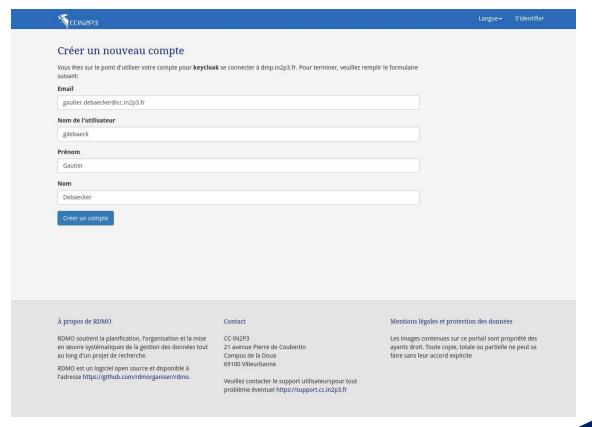
Provisionnement et Authentification

Authentification centralisée



- * « Compte CC » approvisionné directement depuis Portail ID
- Obligation de passer par Portail ID pour créer un nouveau compte
- Volonté de retirer user/psw dans le future







Administration du portail

Organisation - Mises à jour - Automatisations

Organisation du travail (GitLab)



Versioning des modifications via GitLab

- MàJ logicielles (RDMO)
- Fichiers de configuration (dockerfile+local.py)
- Thème & Questionnaire(s)



Suivie des bonnes pratiques de qualité logicielle

Gestion des tâches avec issues et milestones (GitLab)

❖ Mise en place du déploiement automatisé en ligne (WOK)

- Pipeline de déploiement
 - pré-prod/prod
- > Synchronisation avec doc interne



Adaptation de la configuration (1.10 → 2.1.3)



Mise à niveau des fichiers de configuration

- Changement de librairies obsolètes
- Ajout de nouvelles dépendances
- > Ajout de fonctionnalités liées au site

Dockerisation du service

- Choix de l'image et de l'environnement logiciel (Engine)
- > Ajout du dossier contenant le thème personnalisé (Custom)

Personnalisation

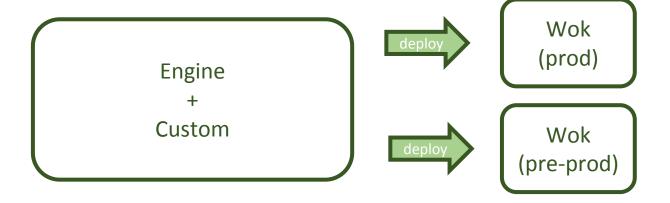
> Ajout d'un script maison pour le carrousel de photos

Automatisation GitLab



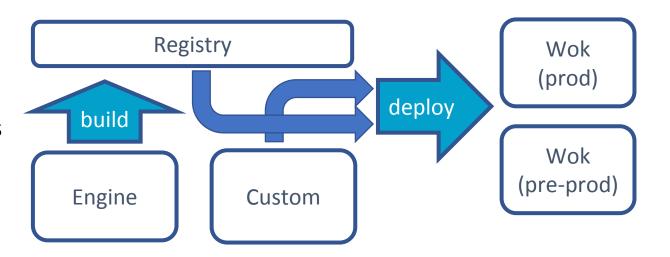
❖À présent

- > Tout dans un seul projet
 - ➤ Mises à jour de l'outil
 - Configuration
 - Changements des textes



❖ En étude

- Développement à 2 projets
 - Possibilité de plusieurs thèmes/projets
 - Gain de temps pour modifications mineures





Conclusions et perspectives

Travaux achevés, en cours et à faire

Conclusions



Le CC-IN2P3 offre un outil de rédaction et gestion des DMP aux communautés avec qui il collabore

- Accès centralisé avec le système d'identité
 - Nécessité d'un compte utilisateur du CC-IN2P3
- Le questionnaire
 - Conforme aux standards RDA
 - Documenté sur la documentation utilisateurs
- Collaboration avec la communauté RDMO
 - Autonome sur le thème et les questions proposées

Perspectives



Modèle de gestion des déploiements avec plusieurs projets GitLab

Séparer les personnalisations de la construction de l'image docker et du déploiement

Traitements automatiques des formulaires

- Analyse syntaxique des formulaires remplis
- Dépôt et versionnage des documents validés

Contrôles automatiques des textes proposés

Vérification conformité RDA en cas de modification du questionnaire



Merci pour votre temps

- Générique de fin
 - > Portail DMP:
 - https://dmp.in2p3.fr
 - > Portail ID :
 - https://id.cc.in2p3.fr
 - Documentation Plan de gestion des données :
 - https://doc.cc.in2p3.fr/fr/Data-storage/storage-areas.html#dmp
 - ➤ Documentation Questions essentielles :
 - https://doc.cc.in2p3.fr/fr/Data-storage/dmp-questions.html

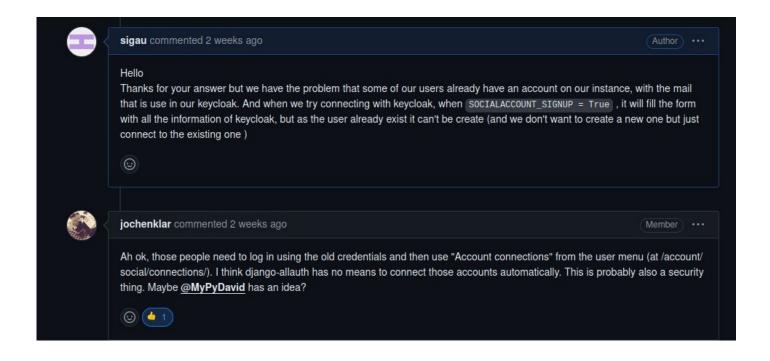


Backslides

Lier un compte CC à un compte DMP existant



Obligation de passer par une méthode manuelle



Lier un compte CC à un compte DMP existant





Bienvenue

Vous êtes sur le portail pour la création des Plans de Gestion des Données (DMP : Data Management Plan) des expériences et projets de l'IN2P3.

Tout utilisateur demandant un compte sur ce service doit avoir un « compte CC-IN2P3 ».

Veuillez vous référer au portail des identités pour en faire demande.

Un DMP permet de décrire les processus de création, utilisation, structuration, sauvegarde et protection des données en fonction de leur profil au sein d'un projet. Il est indispensable pour une planification structurée, la mise en œuvre et l'administration des données d'un projet scientifique.

Ce portail vous donnera accès à un formulaire dédié à votre projet ou expérience. Toutes les questions proposées par le formulaire peuvent être importantes selon les objectifs de votre projet.

Les questions essentielles pour le périmètre d'intérêt du CC-IN2P3 seront marquées par un astérisque « * ». Les informations ainsi recueillies pourront être téléchargées dans les formats spécifiques pour répondre aux différentes exigences de votre projet ou expérience.

Mise à jour de juillet 2024

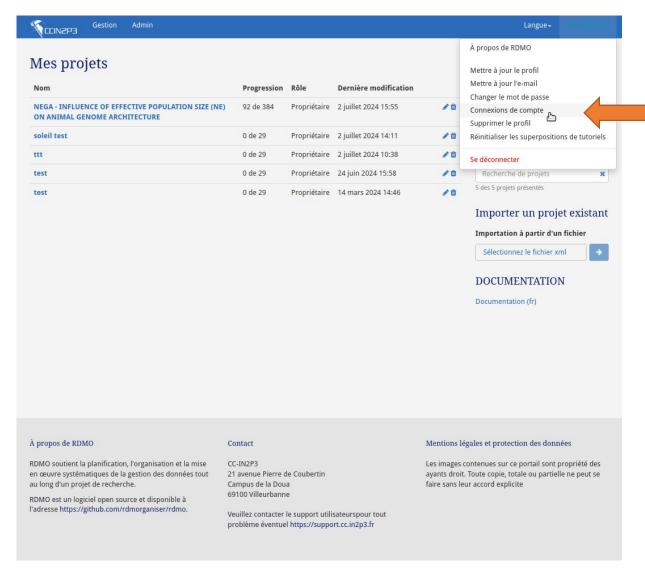
L'authentification par Identifiant et Mot de passe reste temporairement accessible aux utilisateurs ayant obtenu un compte locale avant cette date. Si vous êtes dans cette condition,

- Si vous n'avez pas encore un « compte CC-IN2P3 », veuillez vous référer au portail des identités pour en faire demande, et revenez pour suivre l'option 2.
- 2. Si vous avez déjà un « compte CC-IN2P3 », connectez-vous une première fois par Identifiant et Mot de passe, ensuite choisissez l'option "Connexions de compte" depuis votre profil en haut à droite et ajoutez le compte CC-IN2P3 en tant que compte supplémentaire. Par la suite vous pourrez accéder au portail DMP utilisant l'utilitaire CC-IN2P3.

S'identifier Si vous avez un compte CCIN2P3, connectez-vous via l'utilitaire : Dans le cas contraire,vous pouvez également vous connecter en utilisant le formulaire suivant : show Identifiant rdmoadmin Mot de passe Si vous avez oublié votre mot de passe et que vous souhaitez le réinitialiser, cliquez

S'authentifier une première fois avec Identifiant et Mot de passe





Choisir « Connexions de compte » du menu





