

Collaboration entre GT5 des ateliers de la donnée et le service OPIDoR



RÉPUBLIQUE
FRANÇAISE

recherche.data.gouv.fr



Contexte du travail : Recherche Data Gouv - Ateliers de la donnée

Structures d'accompagnement à destination des équipes de recherche pour toute question relative à la gestion et à la diffusion de leurs données

Plusieurs groupes de travail sont actifs :

- GT1 : outils applicatifs communs pour le fonctionnement en réseau des ateliers
- GT2 : formation professionnelle des personnels des ateliers
- GT3 : boîte à outils et fiches pratiques à destination des chercheurs
- GT4 : mise en œuvre d'un réseau de correspondants données dans les laboratoires
- **GT5 : interaction données et codes**
- GT6 : interaction entre les ateliers de la donnée et les centres de référence thématiques
- GT7 : archivage pérenne des données scientifiques



Contexte du travail : l'importance des codes et des logiciels dans la recherche scientifique

Les logiciels constituent aujourd'hui l'un des piliers fondamentaux de la recherche, au même titre que les publications et les données. Ils sont devenus indispensables dans tous les domaines de la recherche scientifique, à la fois comme ***outils pour la recherche***, comme ***produits de la recherche*** et comme ***objets*** de recherche.

Contexte du travail : l'importance des codes et des logiciels dans la recherche scientifique

De plus en plus, le logiciel est pris en compte dans les politiques nationales et internationales de la science ouverte

MINISTÈRE
DE L'ENSEIGNEMENT
SUPÉRIEUR,
DE LA RECHERCHE
ET DE L'INNOVATION



Troisième axe : ouvrir et promouvoir les codes sources produits par la recherche

Contexte du travail : l'importance des codes et des logiciels dans la recherche scientifique

De plus en plus, le logiciel est pris en compte dans les politiques nationales et internationales de la science ouverte



Brussels, 21.10.2020
C(2020) 7149 final

Actions concrètes :

- mise en place d'un Open Source Programme Office (OSPO) à la commission
- améliorer le référentiel de logiciels
- réviser les pratiques de distribution de logiciels

COMMUNICATION TO THE COMMISSION

OPEN SOURCE SOFTWARE STRATEGY 2020 – 2023

Think Open

Le GT FAIR4RS (2020 - 2022)
 piloté par la [RDA](#), [FORCE11](#) et la
[Research Software Alliance](#)
 (ReSA) a rassemblé une
 communauté pour réviser et
 étendre les principes FAIR aux
 logiciels de recherche

FAIR4RS principles

| FINDABLE Software, and its associated metadata, is easy for both humans and machines to find | ACCESSIBLE Software, and its metadata, is retrievable via standardized protocols | INTEROPERABLE Software interoperates with other software by exchanging data and/or metadata | REUSABLE Software is both usable (can be executed) and reusable (can be understood, modified, built upon, or incorporated into other software) |
|--|---|--|---|
| F1. Software is assigned a globally unique and persistent identifier | A1. Software is retrievable by its identifier using a standardized communications protocol | I1. Software reads, writes and exchanges data in a way that meets domain- relevant community standards | R1. Software is described with a plurality of accurate and relevant attributes |
| F2. Software is described with rich metadata | A2. Metadata are accessible, even when the software is no longer available | I2. Software includes qualified references to other objects | R2. Software includes qualified references to other software |
| F3. Metadata clearly and explicitly include the identifier of the software they describe | | | R3. Software meets domain-relevant community standards |
| F4. Metadata are FAIR, searchable and indexable | | | |

Contexte du travail : dispositifs d'accompagnement

L'attention vers le logiciel se traduit également par la mise en place de dispositifs d'accompagnement à destination des équipes de recherche pour toute question relative à la gestion et l'ouverture du logiciel

Collège codes sources et logiciels
du Comité pour la science ouverte



**Community for University and Research Institution
OSPOs**



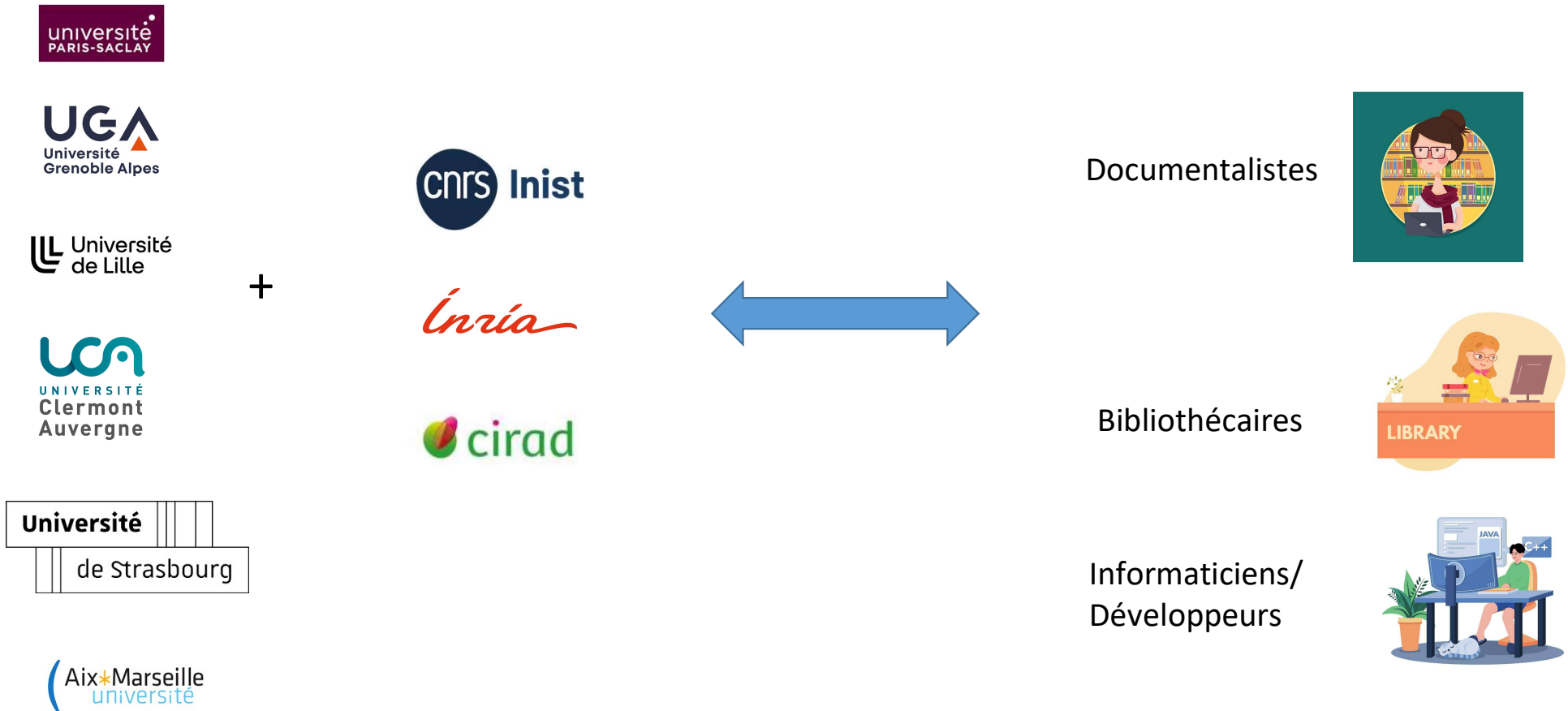
Le réseau de la recherche reproductible



*La cellule Codes Données Grenoble Alpes a été
reconnue OSPO UGA en 2025.*



GT5 : un groupe mixte profils au profit de l'accompagnement autour de logiciels

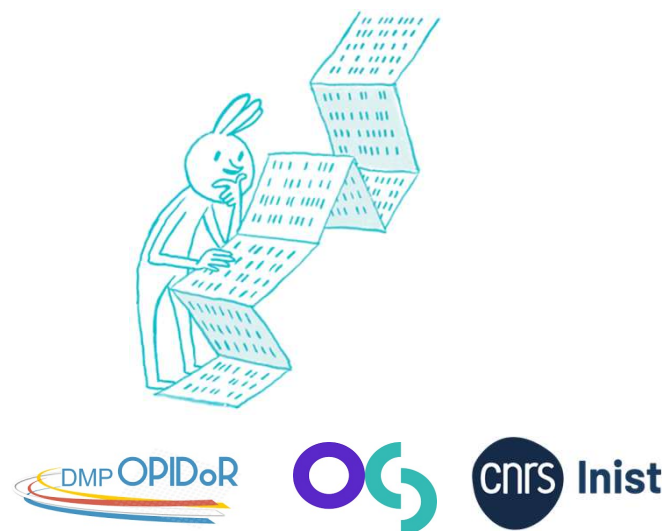


GT5 : focus sur 2 axes de travail

1. Signalement des forges logicielles dans Cat OPIDoR



2. La nouvelle fonctionnalité pour décrire la gestion des codes et logiciels dans DMP OPIDoR



1. Signalement des forges logicielles dans Cat OPIDoR

GT5 dédié aux codes sources et logiciels

Équipe Cat OPIDoR

Référent national pour les forges logicielles de l'Enseignement Supérieur et de la Recherche

Importance des forges logicielles

- Les forges logicielles sont devenues un outil incontournable pour tous les développeurs de logiciels
- L'usage d'un gestionnaire de version est particulièrement utile pour tracer tous les développements et pouvoir revenir en arrière le cas échéant. C'est la bonne pratique lorsqu'on développe du code
- Elles intègrent les outils de développement collaboratif pour le suivi des modifications du code, la gestion des demandes et des réponses d'utilisateurs, la gestion des contributions, la gestion du projet
- L'utilisation d'une forge logicielle permet également d'assurer la sauvegarde du code sur un serveur et est un outil indispensable dès qu'on développe à plusieurs
- Elles intègrent des outils de communication comme des forums

Rapport : Forges de l'ESR – Définition, usages, limitations rencontrées et analyse des besoins

Le rapport porte sur les forges logicielles utilisées dans les établissements de l'Enseignement supérieur et de la Recherche (ESR)

Son propos est de faire un premier état des lieux des forges logicielles utilisées dans l'ESR français et d'identifier les moyens de rendre plus visibles les productions logicielles issues de la science ouverte.



Daniel Le Berre, Jean-Yves Jeannas, Roberto Di Cosmo, François Pellegrini
([hal-04098702v7](https://hal.archives-ouvertes.fr/hal-04098702v7))

Besoin concernant les forges : comment choisir une forge ?

⚙ **Beaucoup de facteurs à prendre en compte :**

- Accès et périmètre de partage
- Sécurité et conformité
- Support d'utilisation
- Fonctionnalités disponibles

Besoin concernant les forges : comment choisir une forge ?

Beaucoup de facteurs à prendre en compte :

- Accès et périmètre de partage
- Sécurité et conformité
- Support d'utilisation
- Fonctionnalités disponibles

Questions à se poser :

- Y a-t-il une forge dans mon établissement ?
- Comment puis-je m'y connecter ? Avec quel identifiant ?
- Qu'est-ce que je peux faire dans la forge ?
- Puis-je inviter des collaborateurs externes ?
- Mon projet sera-t-il privé ou public ?
- La forge est-elle souveraine (hébergée en France / par l'établissement) ?
- Quelle est la visibilité de la forge ?

Comment construire un outil utile ? Étapes pour sa construction

étape 1

Recenser les besoins des utilisateurs et les fonctionnalités des forges



Comment construire un outil utile ? Étapes pour sa construction

étape 1

Recenser les besoins des utilisateurs et les fonctionnalités des forges

étape 2

Traduire les besoins et les fonctionnalités sous forme de questions équivalentes.



Pour les utilisateurs : Avez-vous besoin d'être aidé dans l'utilisation d'une forge ?



Pour les responsables des forges : La forge bénéficie d'un accompagnement personnalisé à l'utilisation ?

Comment construire un outil utile ? Étapes pour sa construction

étape 1

Recenser les besoins des utilisateurs et les fonctionnalités des forges

étape 2

Traduire les besoins et les fonctionnalités sous forme de questions équivalentes.

étape 3

Trouver l'outil le plus adapté pour le choix d'une forge selon les besoins et les fonctionnalités



Pour les utilisateurs : Avez-vous besoin d'être aidé dans l'utilisation d'une forge ?



Pour les responsables des forges : La forge bénéficie d'un accompagnement personnalisé à l'utilisation ?



Logigramme : difficile d'imposer une hiérarchie claire



Base des données : manque d'ergonomie

Comment construire un outil utile ? Étapes pour sa construction

étape 1

Recenser les besoins des utilisateurs et les fonctionnalités des forges

étape 2

Traduire les besoins et les fonctionnalités sous forme de questions équivalentes.

étape 3

Trouver l'outil le plus adapté pour le choix d'un forge selon les besoins et les fonctionnalités

étape 4

Rassembler les informations de la communauté sur l'outil Cat Opidor pour chercher sa forge avec des filtres



Pour les utilisateurs : Avez-vous besoin d'être aidé dans l'utilisation d'une forge ?



Pour les responsables des forges : La forge bénéficie d'un accompagnement personnalisé à l'utilisation ?



Logigramme : difficile d'imposer une hiérarchie claire



Base des données : manque d'ergonomie

Informations relayées dans des listes de diffusion à destination des responsables des forges et la communauté



Les étapes à venir

État actuel

25 forges sont actuellement référencées dans l'outil

Appel à contribution

 Si vous êtes responsable ou utilisateur d'une forge non encore signalée, n'hésitez pas à nous contacter :

 adld-gt5-forges@groupes.renater.fr

 cat.opidor.fr

2. La nouvelle fonctionnalité pour décrire la gestion des codes et logiciels dans DMP OPIDoR

GT5 dédié aux codes sources et logiciels

Équipe DMP OPIDoR

Soutien du Collège codes sources et logiciels du Comité pour la science ouverte

Plan de Gestion Logiciel : pourquoi ?

Malgré leur rôle central dans la recherche, les logiciels restent souvent difficiles à trouver, à citer et à référencer correctement. En raison de l'absence de mécanismes appropriés, ils sont souvent négligés ou mal décrits dans les plans de gestion des données.

Plan de Gestion Logiciel : pourquoi ?

Malgré leur rôle central dans la recherche, les logiciels restent souvent difficiles à trouver, à citer et à référencer correctement. En raison de l'absence de mécanismes appropriés, ils sont souvent négligés ou mal décrits dans les plans de gestion des données.

Pour soutenir les scientifiques dans la gestion de la productions de logiciels, les **plans de gestion** sont un outil essentiel. Ils offrent un moyen simple de décrire le logiciel, d'identifier et de planifier les actions de développement et de conservation, tout en tenant compte des leurs spécificités et de leur diversité

Comment mettre en place un plan de gestion de logiciel ?

Vous pouvez utiliser un modèle de PGL pour décrire uniquement votre logiciel

- [PRESOFT](#) disponible sur DMP OPIDoR
 - Teresa Gomez-Diaz, Genevieve Romier. Research Software Management Plan template, V3.2. 2018. [\(hal-01802565\)](#)
- le modèle proposé par le Software Sustainability Institute, disponible sur [Zenodo](#) sous plusieurs formats (docx, odt, md)
- le modèle proposé par [Elixir](#) disponible sur Data Stewardship Wizard

Objectif du GT5

Considérer le logiciel comme un produit de recherche, créé dans le cadre d'un projet, et le décrire directement dans le plan de gestion des données du projet avec un formulaire dédié

Objectif du GT5

Considérer le logiciel comme un produit de recherche, créé dans le cadre d'un projet, et le décrire directement dans le plan de gestion des données du projet avec un formulaire dédié

Élaborer un formulaire :

- qui prendre en compte les spécificités du cycle de vie des logiciels
- simple et adapté à la description de tous types de logiciels (du plus simple au plus complexe)
- conforme aux principes FAIR pour les logiciels
- accessible et exploitable par machine → intégré dans DMP OPIDoR
- avoir la possibilité de décrire le logiciel et le données associées dans le même plan de gestion

L'objectif principal de cette démarche

- Aider les chercheurs à anticiper les enjeux liés au développement des logiciels
- En posant les bonnes questions, le formulaire met en lumière les bonnes pratiques à adopter tout au long du projet de développement
- Il permet aussi de définir les rôles et responsabilités au sein de l'équipe, de structurer le développement, et de choisir les infrastructures et outils adaptés à la production de logiciels de qualité et bien documentés
- Ce travail vise également à favoriser la trouvabilité, la préservation à long terme et la pérennité des logiciels, répondant ainsi aux enjeux de reproductibilité et de réutilisation

Objectif du GT5 : élaborer des recommandations associées au formulaire

- destinées à accompagner les chercheurs dans leurs réponses au formulaire
- précisent les attendus pour chaque question
- encouragent les bonnes pratiques à adopter tout au long du projet de développement

Objectif du GT5 : élaborer des recommandations associées au formulaire

- destinées à accompagner les chercheurs dans leurs réponses au formulaire
- précisent les attendus pour chaque question
- encouragent les bonnes pratiques à adopter tout au long du projet de développement

Élaboration collaborative du formulaire et des recommandations

GT5 – ateliers de la donnée

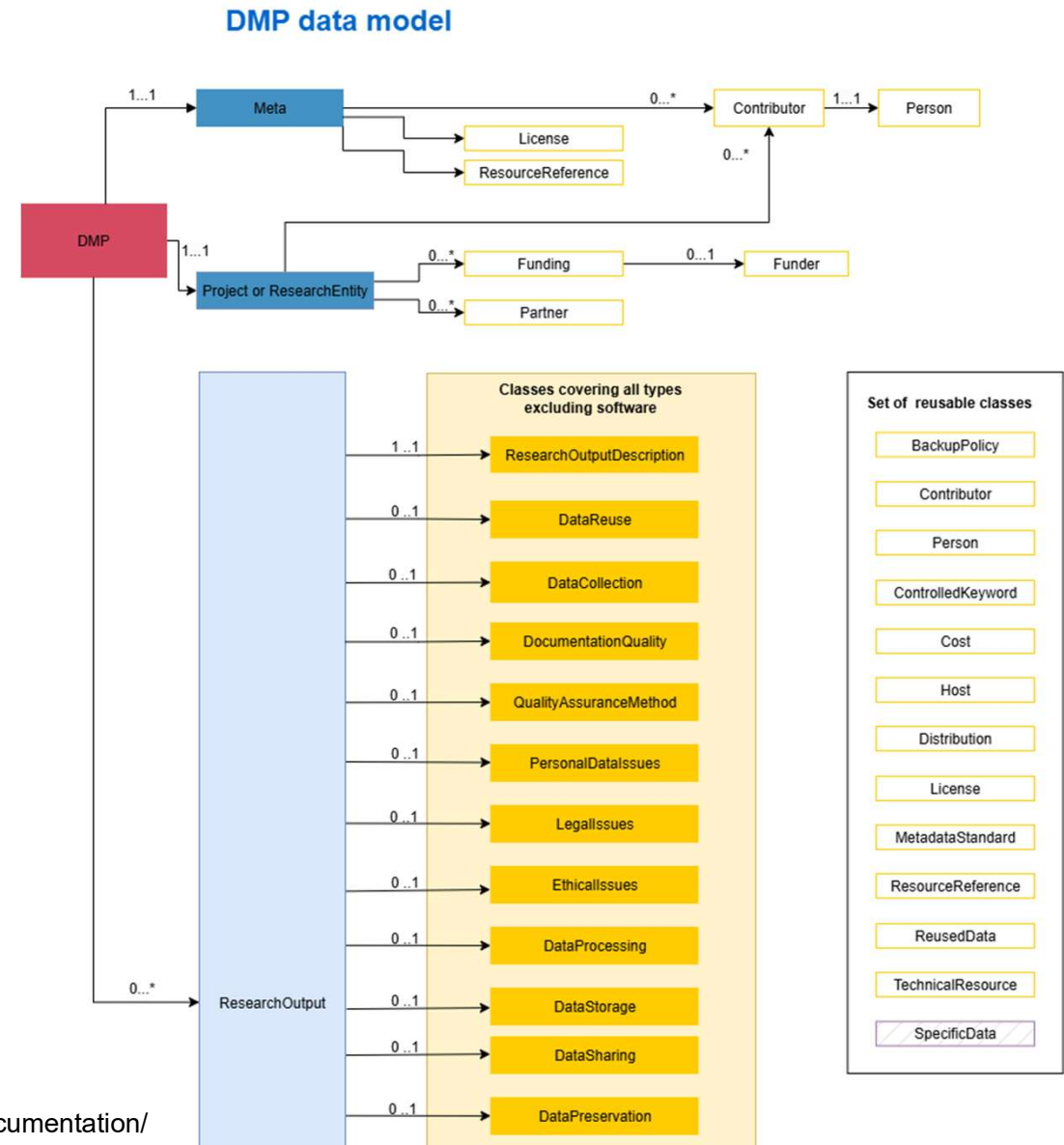
Équipe DMP OPIDoR

Avec le soutien du Collège codes sources et logiciels – Comité pour la science ouverte

Point de départ du travail

- Révision du diagramme simplifié du modèle Science Europe
- Modèle structuré dans DMP OPIDoR
- Composé de plusieurs blocs
- Objectif : intégrer la dimension “logiciel” de manière cohérente

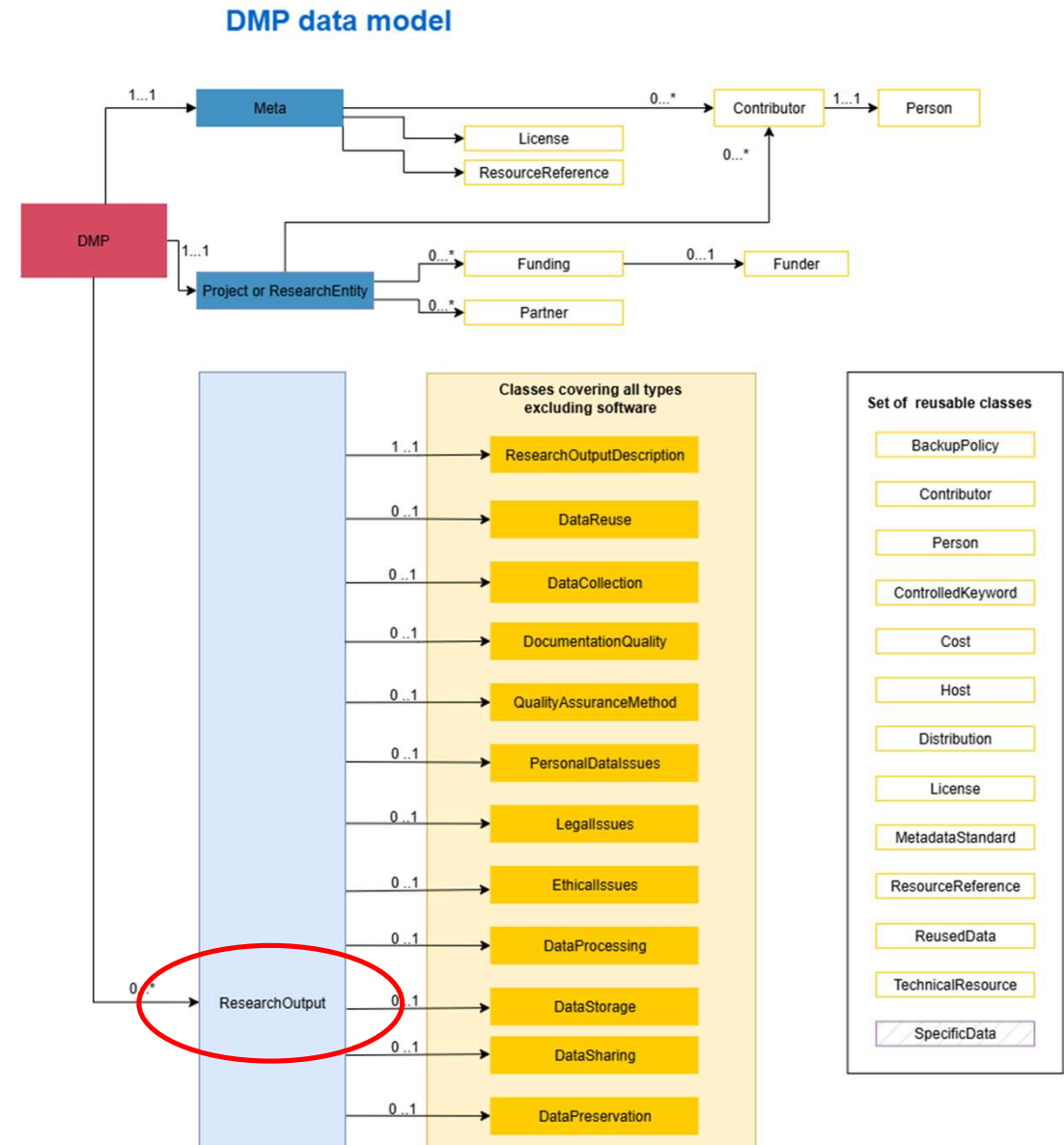
https://opidor.github.io/maDMP-OPIDoR_documentation/



Point de départ du travail

Bloc “Produit de recherche”

- Contient les classes et sous-classes pour la description des produits de recherche
- **Objectif GT5 :**
 - réutiliser les classes existantes
 - créer de nouvelles classes si nécessaire pour les logiciels



Point de départ du travail

Pour adapter les classes aux spécificités du cycle de vie des codes sources et logiciels, nous nous sommes appuyé sur le format CodeMeta qui fournit un standard de métadonnées logicielles largement reconnu

CodeMeta Generator v3.0

Most fields are optional. Mandatory fields will be highlighted when generating CodeMeta.

| | | | |
|--|--|---|--|
| The software itself Name My Software <small>the software title</small> Description My Software computes ephemerides and orbit propagation. It has been developed from early '80. Creation date YYYY-MM-DD First release date YYYY-MM-DD License(s) LGPL-3.0-or-later <small>from SPDX License List; type or select from the list, one by one</small> | Discoverability and citation Unique identifier 10.151.xxxxx <small>The name of the software N codes, UUIDs, etc. see https://schema.org/identifier</small> Application category Astronomy Keywords ephemerides, orbit, astronomy <small>separated by commas (,)</small> Funding PRA_2018_73 grant funding software development Funder Università di Pisa organization funding software development Authors and contributors can be added below | Development community / tools Code repository git+https://github.com/You/RepoName.git Continuous integration https://travis-ci.org/You/RepoName Issue tracker https://github.com/You/RepoName/issues Related links https://www.example.com https://www.example.org <small>URL(s), one URL per line</small> | Run-time environment Programming language C#, Java, Python 3 <small>separated by commas (,)</small> Runtime platform .NET, JVM <small>separated by commas (,)</small> Operating system Android 1.6, Linux, Windows, macOS <small>separated by commas (,)</small> Other software requirements https://www.python.org/downloads/release/python-3130/ https://github.com/psf/requests <small>URL(s), one URL per line</small> |
| Current version of the software Version number 1.0.0 Release date YYYY-MM-DD Download URL https://example.org/MySoftware.tar.gz Release notes Change log: this and that; Bugfixes: that and this. | Editorial review Reference publication https://doi.org/10.1000/xyz123 Review aspect Object facet Review body Review about my software. | Additional information Development status <small>see www.repostatus.org for details</small> Is source code of Bigger Application Is part of https://The.Bigger.Framework.org | |

Authors

Travail étroit entre GT5 et DMP OPIDoR

Première étape : Avancement sur le formulaire

- Définition des sections du formulaire
- Rédaction des questions pour chaque section
- Caractérisation de chaque question (champ texte libre, liste déroulante contrôlée, etc.)
- Réutilisation de listes contrôlées déjà implémentées sur DMP OPIDoR ou identifiées dans d'autres sources
- Rédaction de recommandations pour certaines questions

Travail étroit entre GT5 et DMP OPIDoR

Deuxième étape : validation du formulaire et des recommandations par le Collège codes sources et logiciels du Comité de la science ouverte

Travail étroit entre GT5 et DMP OPIDoR

Deuxième étape : validation du formulaire et des recommandations par le Collège codes sources et logiciels du Comité de la science ouverte

Troisième étape : intégration sur DMP OPIDOR

- Plusieurs **allers-retours** pour adapter notre document aux contraintes techniques de la plateforme
- Réflexion sur la structuration du document dans l'interface DMP OPIDoR
- Objectif : assurer l'interopérabilité et la lisibilité par machine
- Recherche d'un compromis entre nos attentes et les contraintes techniques

Structure du formulaire

En français et en anglais

Intégré dans le modèle de plan gestion de données Science Europe structuré

Produit de recherche de type "Logiciel"   

- Nom abrégé : LOG
- Nom : Produit de recherche de type "Logiciel"
- Type : Logiciel

1. Description générale du logiciel

Tout développer | Tout réduire

1.1 Dénomination et description du logiciel 

2. Outils et environnement d'exécution

Tout développer | Tout réduire

2.1 Environnement de développement 

2.2 Documentation 

2.3 Environnement d'exécution 

3. Préservation du logiciel

Tout développer | Tout réduire

3.1 Référencement 

3.2 Archivage pérenne 

4. Questions juridiques

Tout développer | Tout réduire

4.1 Quels sont les aspects juridiques en lien avec la gestion du logiciel ? 

5. Valorisation scientifique

Tout développer | Tout réduire

5.1 Valorisation scientifique 

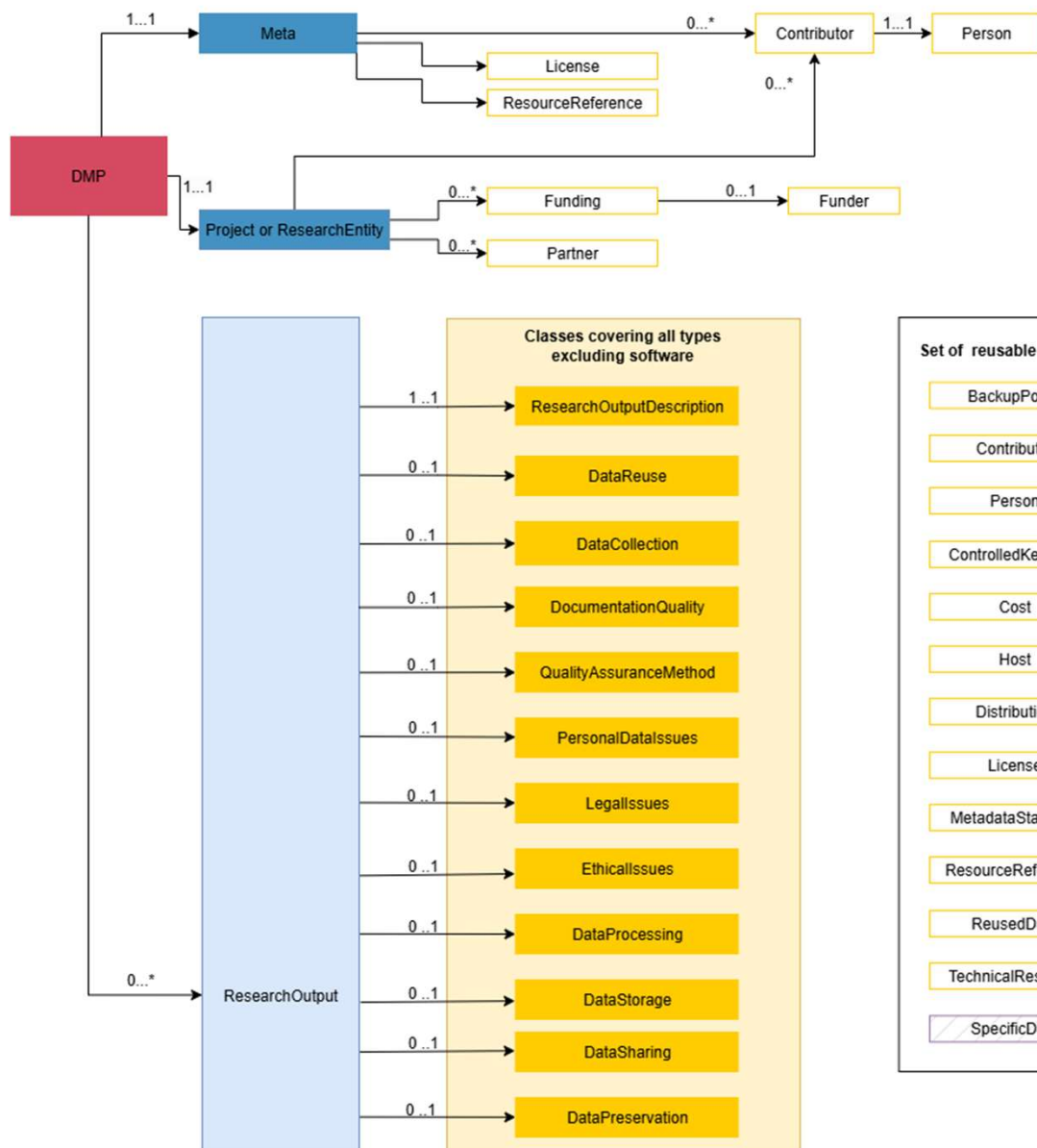
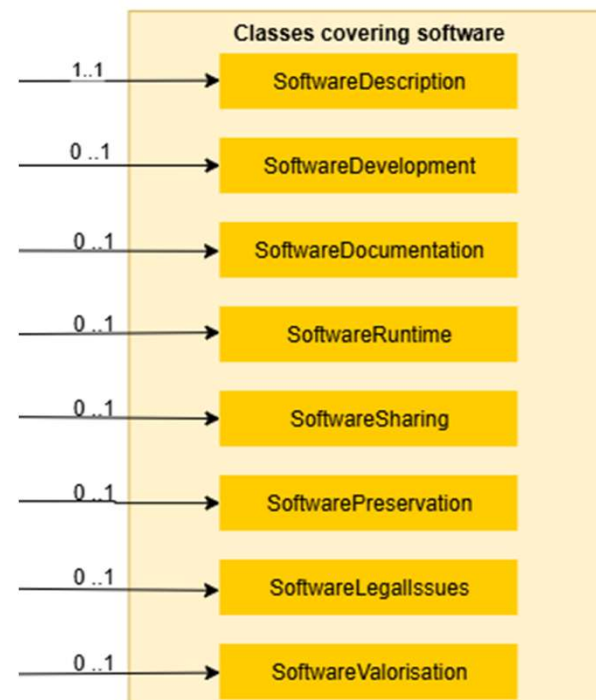


Diagramme simplifié décrivant l'arborescence globale du modèle Science Europe structuré dans DMP OPIDoR avec les classes pour le logiciel



Conforme aux principes FAIR4RS

Project information

Description, funding/funder (ROR, grantId), contributors (ORCID, ROR, role)

DMP information

Scope, **contributors**, references to related documents and DMPs

Q1.1 Definition and name of the software

Description, disciplines, keywords, references to related outputs, version

Q4.1 Legal aspects

Licences : SPDX references, authors and contributors

Q2.2 Documentation

Readme, documentation for users and developers

Q2.1 Development environment

Forges (public depot), versioning, tests

Q2.3 Runtime environment

Programming languages, operating systems, dependencies

F2. Software is described with rich metadata.

R1. Software is described with a plurality of accurate and relevant attributes.

I2. Software includes qualified references to other objects.

R2. Software includes qualified references to other software.

R3. Software meets domain-relevant community standards.

I1. Software reads, writes and exchanges data in a way that meets domain-relevant community standards.

Q3.1 Referencing

PID for a version :
HAL/Zenodo

Q3.2 Long-term archiving

Software Heritage: SWHID

Q5.1 Scientific promotion

Publications, citation

F3. Metadata clearly and explicitly include the identifier of the software they describe.

F4. Metadata are FAIR, searchable and indexable.

F1. Software is assigned a globally unique and persistent identifier.

A2. Metadata are accessible, even when the software is no longer available.

A1. Software is retrievable by its identifier using a standardized communications protocol.

[FAIR4RS principles](#)

Les étapes à venir

- Sous-groupe du GT5 qui travaille sur le mapping entre le formulaire et le format codemeta
- Import / Export avec codemeta
- Rapport qui sera déposé sur Hal