

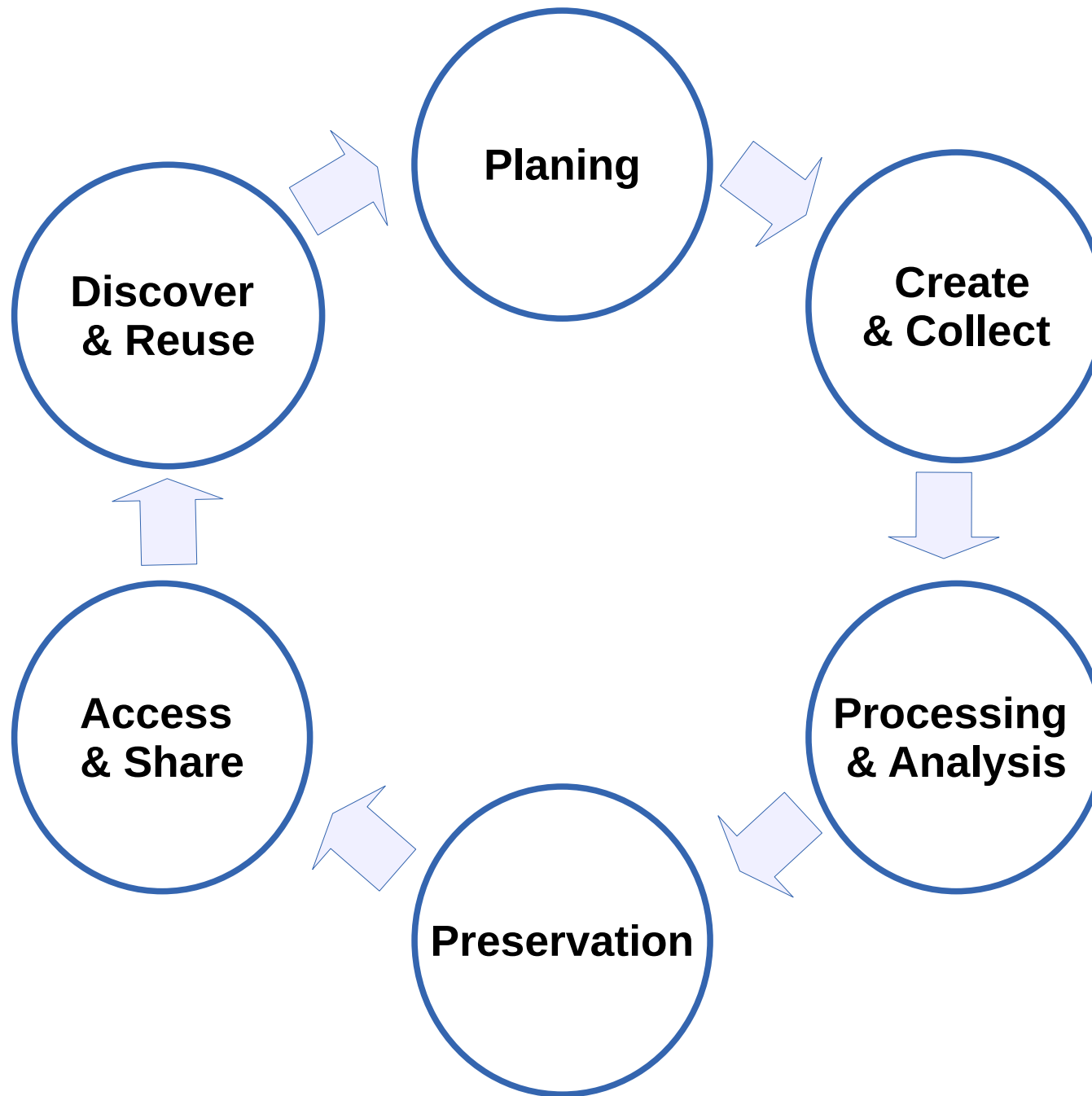
Using DMP machine actionable with research data repositories

Yonny CARDENAS
cardenas@cc.in2p3.fr

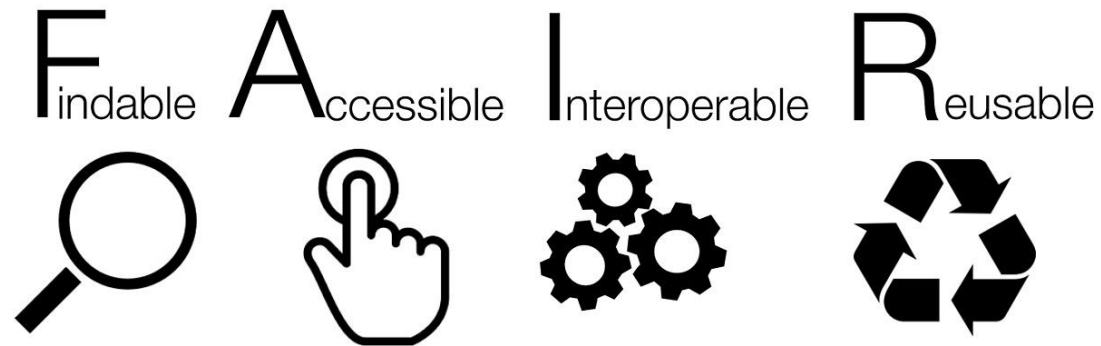
Journées Calcul Données

15 December 2021

Life cycle data research management



- Principles “FAIR” research data



Data Management Plan

Document that describes how the data will be obtained, documented, analyzed, preserved and disseminated during and after a research project.

Decisions to manage data throughout the project and beyond.



Data Management Plan

- For simplicity, the DMP is usually presented in the form of a questionnaire or form to be answered.
- It can take a lot of upstream work to complete it.



RDMO Project

- Research Data Management Organiser
- Proposed by the Astrophysics community
 - Leibniz Institute for Astrophysics Potsdam (AIP)
- Objectives
 - Planning, organization and implementation of research data management
- Contributions
 - Flexible, generic and multilingual DMP questionnaire
 - Tools for information management



Leibniz-Institut für
Astrophysik Potsdam



DMP Questionnaire

Objectives

- Encourage a global reflection on the different aspects of research data management.
- Support the data lifecycle management planning process.

Features:

- It is a generic questionnaire (can be used by all scientific communities)
- It does not refer to specific technologies or services.
- 123 questions organized in 7 sections and 42 subsections

On line version:

- <https://dmp.in2p3.fr> deployed on RDMO at CC-IN2P3
- <https://dmp.opidor.fr> deployed on DMPOpidor at l'INIST

Printing Version:

- <http://irods.in2p3.fr/dmp>

DMP Example online form

dmp.in2p3.fr - Mozilla Firefox

dmp.in2p3.fr

https://dmp.in2p3.fr/projects/62/questions/

dmp.in2p3.fr Retour au projet Gestion Admin Langue Yonny CARDENAS

Question

Stockage et conservation à long terme / Conservation à long terme

Veillez remplir le formulaire pour chaque jeu de données. Les différents jeux de données seront mentionnés dans les questions suivantes. Vous pouvez ajouter un nouveau jeu de données en utilisant le bouton vert. Une fois créé, vous pouvez éditer ou supprimer des jeux de données en utilisant les boutons dans le coin supérieur droit.

données « source » données « résultats » base de données mongodb « vigiechiro »

[Ajouter jeu de données](#)

Ce jeu de données doit-il être conservé à long terme?

Oui Non

Quelles sont les raisons pour lesquelles ce jeu de données doit être préservé à long terme?

- Utilisé dans une publication / preuve de bonnes pratiques scientifiques
- Réutilisation dans des projets ultérieurs ou par d'autres
- Obligations légales
- La documentation, car elle est pertinente pour la société
- Engagement personnel
- Autre:

Combien de temps les données seront-elles stockées?

10 ans

Aperçu

Projet: [Vigiechiro](#)
Catalogue: RDMO

[Retour à mes projets](#)

Progression

35 de 122

[Retour](#) [Sauter](#)

Navigation

Veillez noter que l'utilisation de la navigation supprimera toute entrée utilisée.

Général

- [Classification du contenu](#)
- [Classification technique](#)
- [Utilisation de données](#)
- [Métadonnées et référencement](#)
- [Cadre légal et éthique](#)
- [Stockage et conservation à long...](#)
- [Sélection](#)
- [Conservation à long terme](#)
- [Coûts de conservation à lon...](#)

DMP Example on line form

The screenshot shows a web browser window with the URL `https://dmp.in2p3.fr/projects/62/answ`. The page title is "Réponses pour Vigiechiro". The browser's address bar shows the URL and search icons. The page has a dark blue header with navigation links: "dmp.in2p3.fr", "Retour au projet", "Gestion", "Admin", "Langue", and "Yonny CARDENAS".

Réponses pour *Vigiechiro*

Dans ce qui suit, nous avons résumé les informations sur le projet telles que fournies par vous et vos collaborateurs.

Général

Sujet

Décrivez brièvement le projet et ses objectifs.

Il s'agit de stocker les données produites par le suivi acoustique automatique des chauves-souris, réalisé à l'échelle nationale par le réseau participatif animé par le CESCO (Centre d'Ecologie et des Sciences de la Conservation) du MNHN (Museum National d'Histoire Naturelle). Sont enregistrés à la fois les données brutes des enregistrements de sons, et les résultats des traitements d'identification automatique des espèces de chiroptères enregistrés.

Veillez donner quelques mots clés décrivant la question de recherche.

- chauves-souris
- sons
- chiroptères
- suivi acoustique
- mnhn
- CESCO
- identification automatique

Domaine de recherche

À quel(s) domaine(s) de recherche ce projet appartient-il?

- Sciences de la Vie / Zoologie

Quelles personnes ou institutions sont responsables de la coordination du projet?

- MNHN (Museum National d'Histoire Naturelle)
- CESCO (Centre d'Ecologie et des Sciences de la Conservation)
- Yves BAS

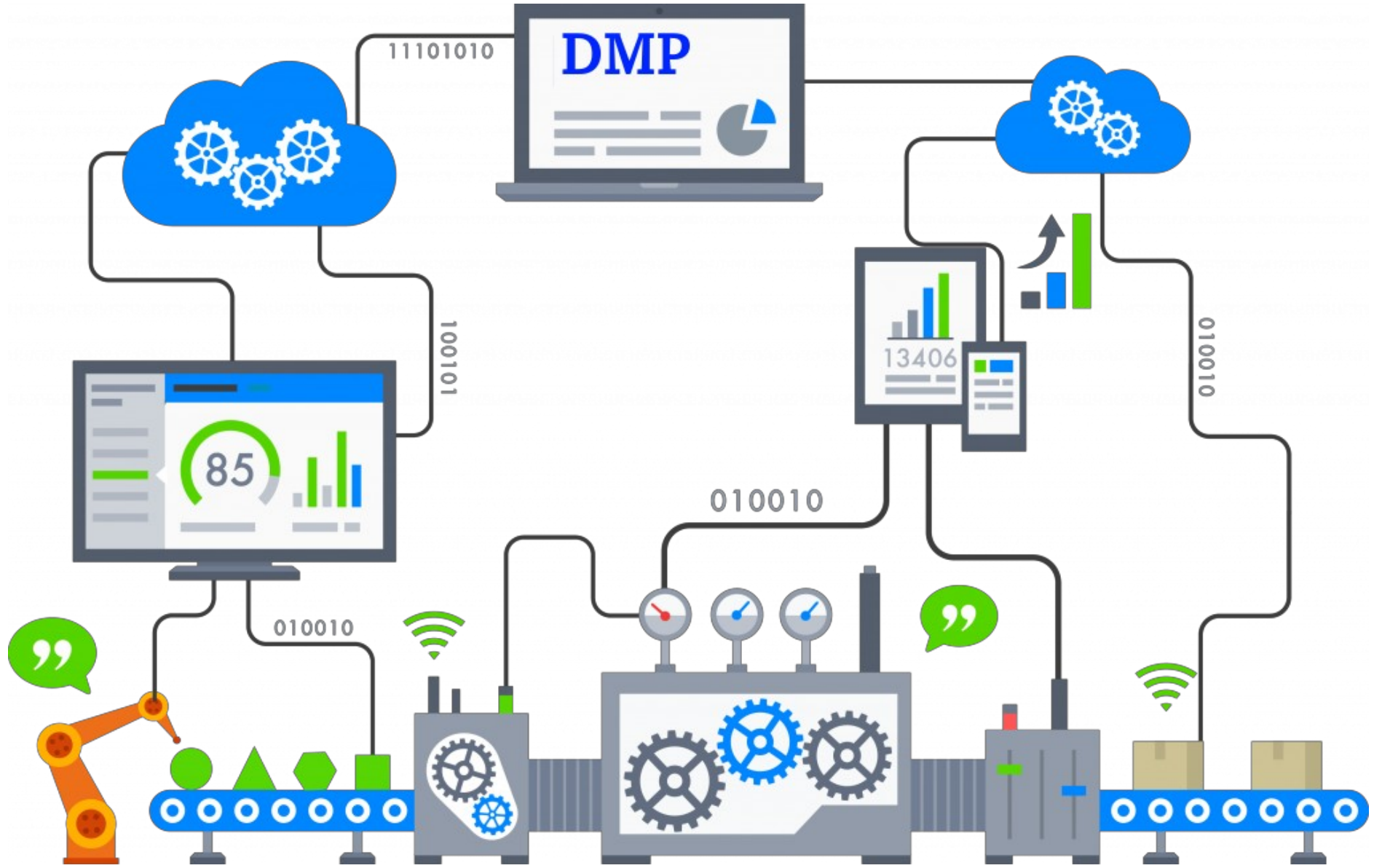
Options

[Retour à l'aperçu du projet](#)

Exporter

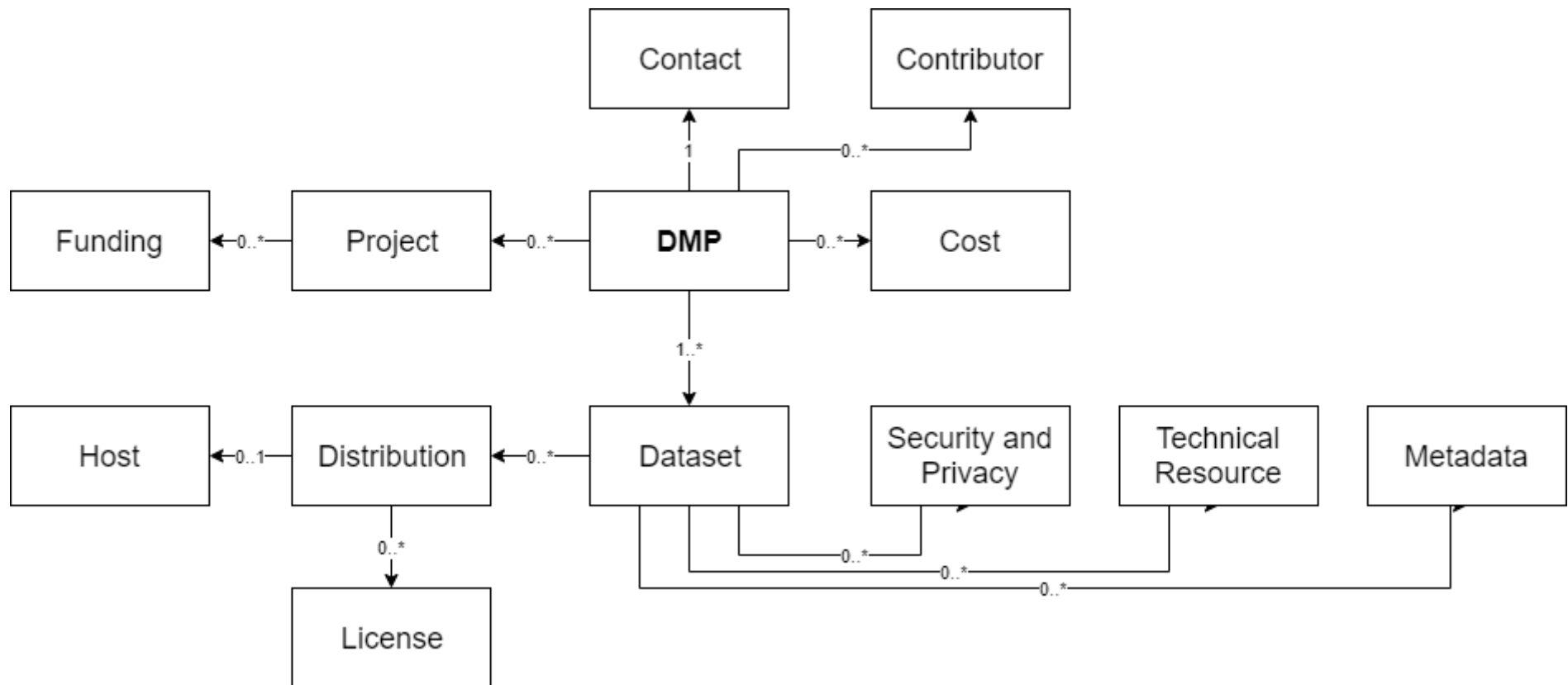
- PDF
- Rich Text Format
- Open Office
- Microsoft Office
- HTML
- Markdown
- mediawiki
- LaTeX

Interoperability DMP machine actionable



Interoperability DMP machine actionable

RDA Common Standard Model

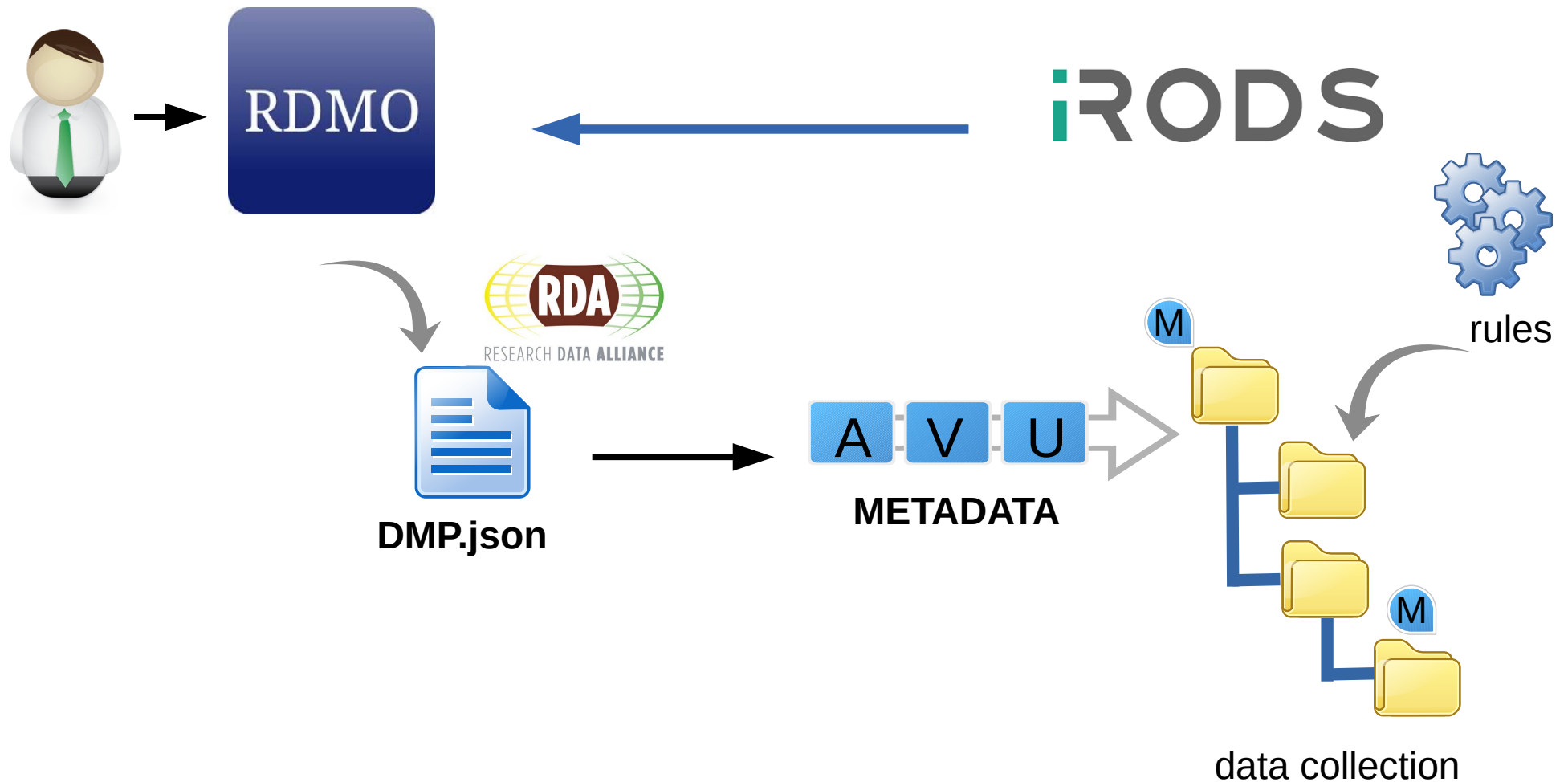


DMP machine actionable (RDA)

The screenshot shows a web browser window with the URL `https://dmp.in2p3.fr/projects/62/`. The page title is 'Vigiechiro'. The interface includes a navigation bar with 'Gestion' and 'Admin' menus, and a user profile 'Yonny CARDENAS'. The main content area is divided into several sections:

- Description:** A paragraph explaining the project's goal: to store data from automatic acoustic monitoring of bats, implemented at the national scale by the CESCO network.
- Catalogue:** A table with one entry: 'RDMO'.
- Tâches:** A section explaining that tasks are generated automatically from project responses. Below it is a table with columns: 'Tâche', 'Description', 'Plage de temps', and 'Statut'. One task is listed: 'Contact IT department to acquire Infrastructure resources' with a status of 'ouvrir'.
- Vues:** A section explaining that views are created from project responses and can be exported. Below it is a table with columns: 'Vue' and 'Description'. Two views are listed: 'Bielefeld' and 'CITEC DMP'.
- Options:** A list of actions including 'Répondre aux questions', 'Afficher les réponses', 'Mettre à jour les Informations du projet', 'Ajouter un membre', 'Créer un instantané', and 'Retour à l'aperçu des projets'.
- Exporter:** A list of export formats: 'RDMO XML', 'CSV séparées par des virgules', 'CSV point-virgule séparé', 'as maDMP JSON', 'as DataCite XML', and 'as RADAR XML'. Two red arrows point to 'CSV séparées par des virgules' and 'as maDMP JSON'.
- Valeurs d'importation:** A section titled 'Importation à partir d'un fichier' with a button 'Sélectionnez le fichi...' and a right arrow.

DMP machine actionable and IRODS



DMP as metadata IRODS

The screenshot shows the iRODS User Interface. The browser address bar displays the URL: `ccirdli205.in2p3.fr:5500/metalnx/collectionInfo?path=%2Farchive%2Fvigiechiro%2Fpackages`. The interface includes a sidebar with 'Favorites', 'Public', and 'Trash' options. The main content area has tabs for 'Details', 'Metadata', and 'Permissions', with 'Metadata' selected. A 'Delete selected' button and a search input are at the top of the table. The table lists 75 metadata entries, showing the first 13. Each entry has a checkbox, an 'Attribute' column, a 'Value' column, a 'View' button, and 'Edit' and 'Delete' buttons.

<input type="checkbox"/>	Attribute	Value	View	Edit	Delete
<input type="checkbox"/>	dmp/title	maDMP for Vigiechiro	View	Edit	Delete
<input type="checkbox"/>	dmp/modified	2021-11-30T14:32:37.31450...	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/title	base de données mongodb « vi...	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/sensitive_data	unknown	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/security_and_...	Security measures	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/security_and_...	Access permissions	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/security_and_...	Et, il faut rappeler que les don...	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/personal_data	no	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/keyword/6	identification automatique	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/keyword/5	CESCO	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/keyword/4	mnhn	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/keyword/3	suivi acoustique	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/keyword/2	chiroptères	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/keyword/1	sons	View	Edit	Delete
<input type="checkbox"/>	dmp/dataset/2/keyword/0	chauves-souris	View	Edit	Delete

DMP as metadata IRODS

The screenshot displays the iRODS User Interface with a table of metadata entries. A modal window titled "View Attribute and Value" is open, showing the details for the selected entry. The table lists various attributes such as keywords, security, personal data, and distribution, each with a corresponding value and actions like View, Edit, and Delete.

Attribute	Value
dmp/dataset/0/keyword/2	chiroptères

Attribute	Value	View	Edit	Delete
dmp/dataset/0/keyword/2	chiroptères	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dmp/dataset/0/keyword/1	sons	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dmp/dataset/0/keyword/0	chauves-souris	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dmp/dataset/0/distribution/1/...	Preservation after the project	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dmp/dataset/0/distribution/1/...	closed	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dmp/dataset/0/distribution/0/...	Storage during the project	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dmp/dataset/0/distribution/0/...	enregistrements de sons fichie...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dmp/dataset/0/distribution/0/...	shared	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dmp/dataset/0/description	Il s'agit des fichiers « .wav » co...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

maDMP et IRODS

<input type="checkbox"/>	dmp/dataset/0/distribution/1/...	closed		
<input type="checkbox"/>	dmp/dataset/0/distribution/0/...	Storage during the project		
<input checked="" type="checkbox"/>	dmp/dataset/0/distribution/0/...	tape		
<input checked="" type="checkbox"/>	dmp/dataset/0/distribution/0/...	disk		
<input type="checkbox"/>	dmp/dataset/0/distribution/0/...	enregistrements de sons fiche...		

View Attribute and Value

Attribute	dmp/dataset/0/distribution/0/host/1/backup_type
Value	tape

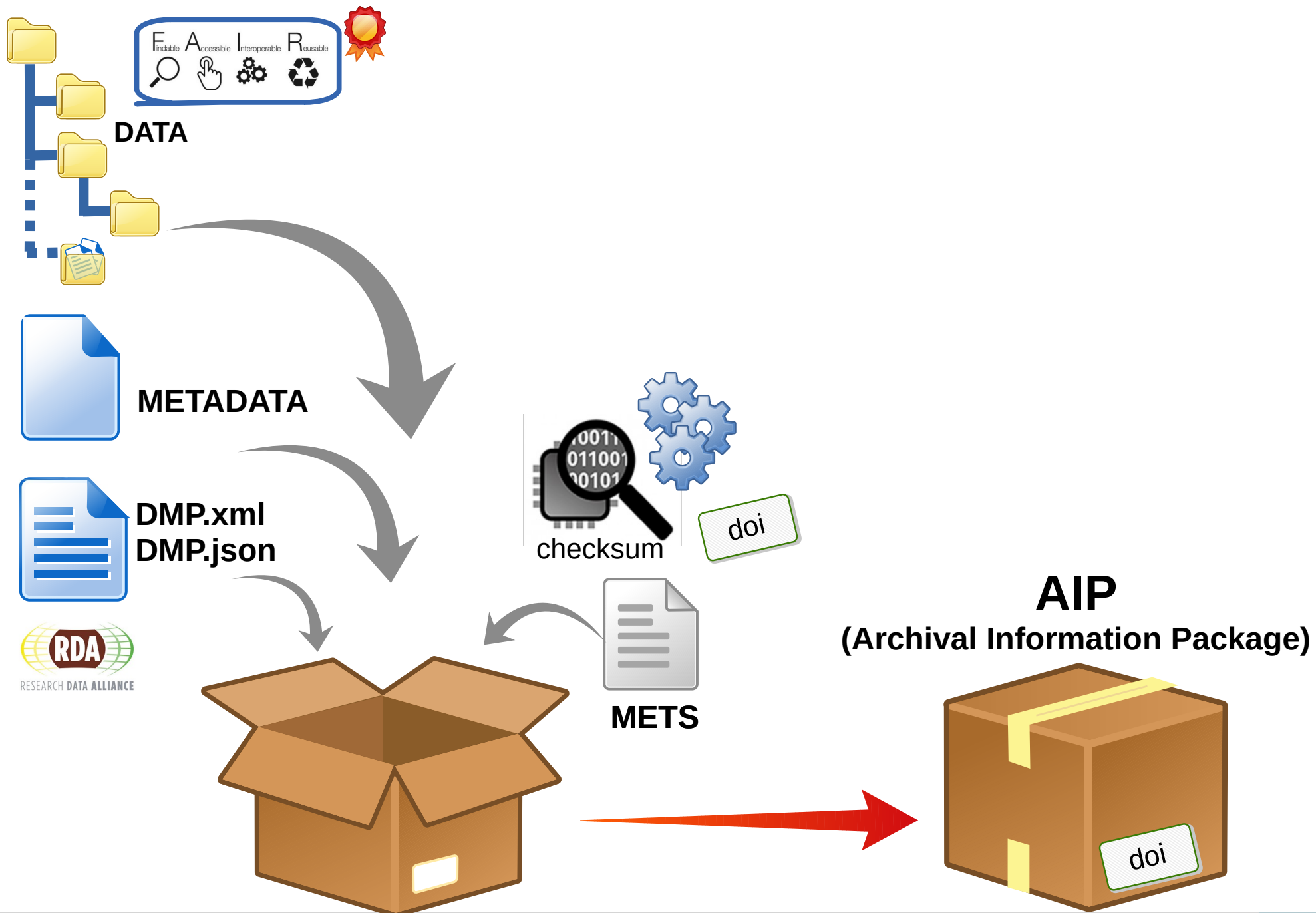
Close

View Attribute and Value

Attribute	dmp/dataset/0/distribution/0/host/0/backup_type
Value	disk

Close

Another Example: OAIS Preservation



METS: Metadata Encoding and Transmission Standard

Summary

- DMP is essential to improve data life cycle
- It exposes key decisions to manage dataset during project and beyond
- DMP machine-actionable information can be useful to assist real data management
- Making DMP interoperability is mandatory to data management
- RDA Common Standard maDMP to provide basic interoperability between systems producing or consuming machine-actionable data management plans (maDMPS).
- IRODS can used maDMP to automatize data management policies

Using DMP machine actionable with research data repositories

Yonny CARDENAS
cardenas@cc.in2p3.fr

Journées Calcul Données

15 December 2021