

# IG : Professionalising Data Stewardship



Soizick Lesteven  
9 février 2023

Discussions récurrentes la gestion des données de la recherche dans le contexte de la science ouverte (FAIR)

→ dimension supplémentaire : qualité, pérennité, accessibilité

- L'idée de créer un groupe d'intérêt a été soulevée lors d'une session BoF pendant la 14ème plénière de RDA (Helsinki oct. 2019)
- Le groupe a été formellement approuvé lors de la 16ème plénière de RDA (Costa Rica nov. 2020)

## **Objectif :**

Le partage des bonnes pratiques, des initiatives et des projets qui pourraient aider les organismes de recherche à professionnaliser la gestion des données au sein et en dehors de la communauté RDA

- 130 membres

→ **charte :**

[https://www.rd-alliance.org/sites/default/files/case\\_statement/Professionalising\\_data\\_stewardship\\_charter\\_proposal\\_20200515.pdf](https://www.rd-alliance.org/sites/default/files/case_statement/Professionalising_data_stewardship_charter_proposal_20200515.pdf)

- Huit défis ont été identifiés et discutés (task groups) :
  - une **analyse** de la gestion des données ;
  - la **terminologie** de la gestion des données ;
  - l'**intégration** de la gestion des données dans une organisation ;
  - les profils d'emploi des gestionnaires de données ;
  - la **formation** des gestionnaires de données ;
  - les **parcours professionnels** ;
  - la **mise en réseau** et l'échange de connaissances ;
  - la **certification**.

- 4 ans avec beaucoup d'échanges (communauté slack : visio)
- Une recherche documentaire (2021)
- Une enquête sur les modèles de Data Stewardships (2021)
  - 30 questions
  - 136 réponses, 85 complètes (surtout en Europe)

Très variable en fonction du service (local, national ou international), de l'expérience, de la maturité, et de la nature des données (commerciales ou de recherche)

# Professionalising Data Stewardship enquête (30 questions)



- Q1 : What organisational level is your service at?
- Q2 : How specific is your service to given research domains?
- Q3 : What service type keywords (scope) best describe your service/provision?
- Q4 : Which global regions are your main activities focused within?
- Q5 : Is your service based in a specific country
- Q6 : What are the job titles of people in your organisation involved in "traditional" RDM support and services?
- Q7 : What are the job titles of people in your organisation involved in RDM, FAIR and open science ?
- ....

DOI : 10.5281/zenodo.6665146  
<https://zenodo.org/record/6665146#.Y-ImVtLMKf4>

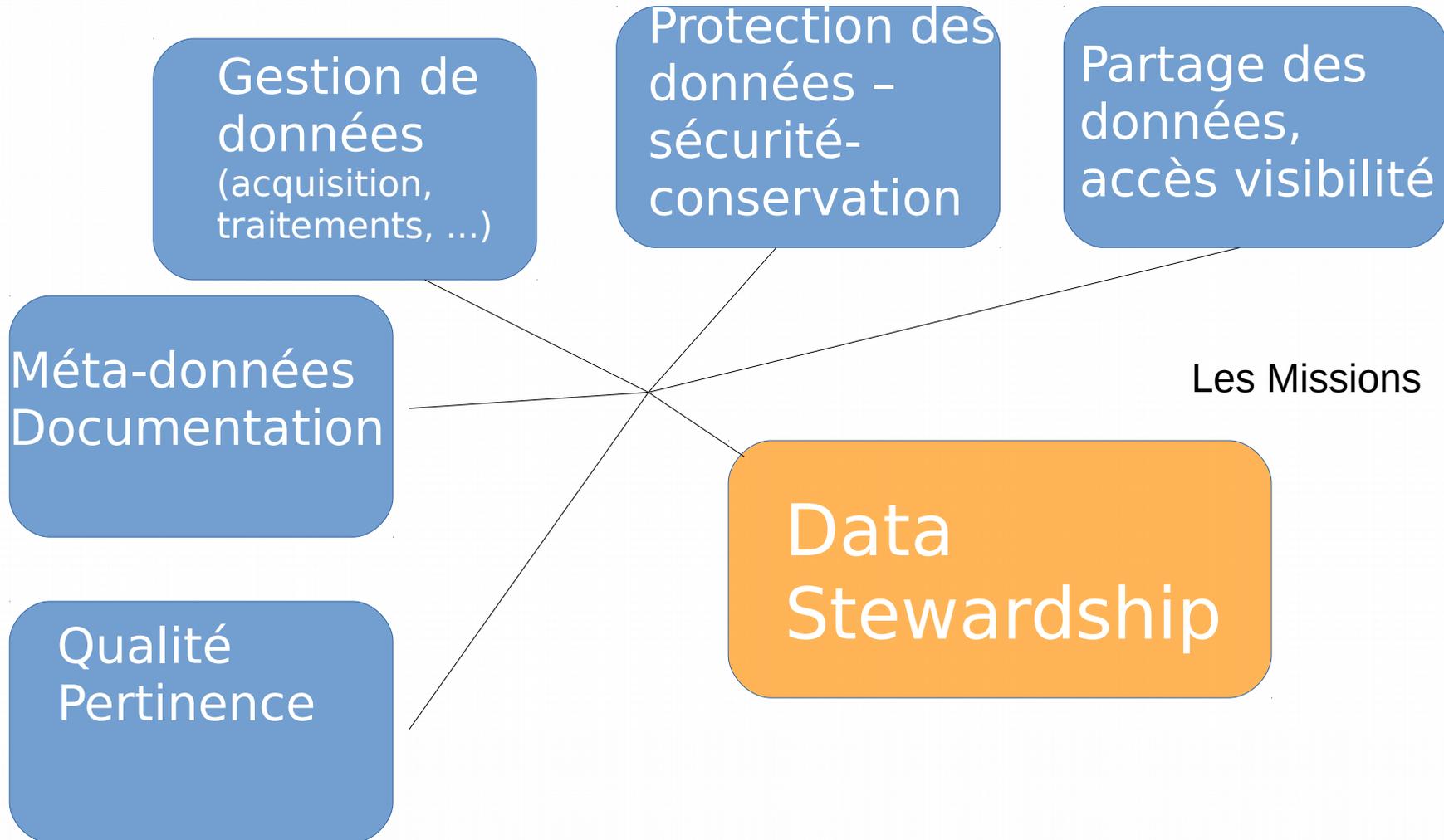
Rapport : DOI: 10.15497/RDA00075  
<https://zenodo.org/record/6665146#.Y-S4FUN7nVF>

# Professionalising Data Stewardship

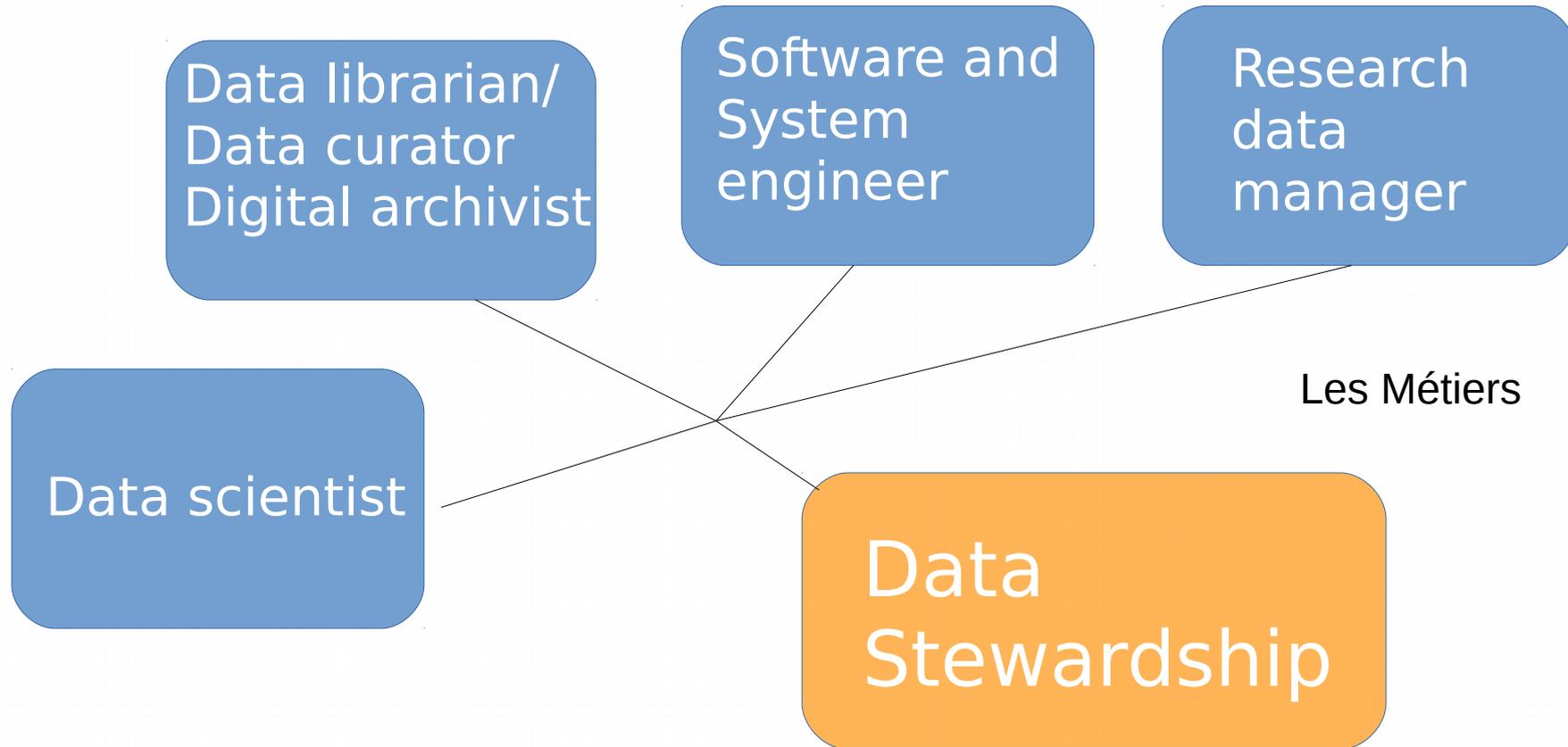


Nuage de mots pour les rôles liés au RDM et à la science ouverte et FAIR dans les organisations nationales et de niveau supérieur (Q7)

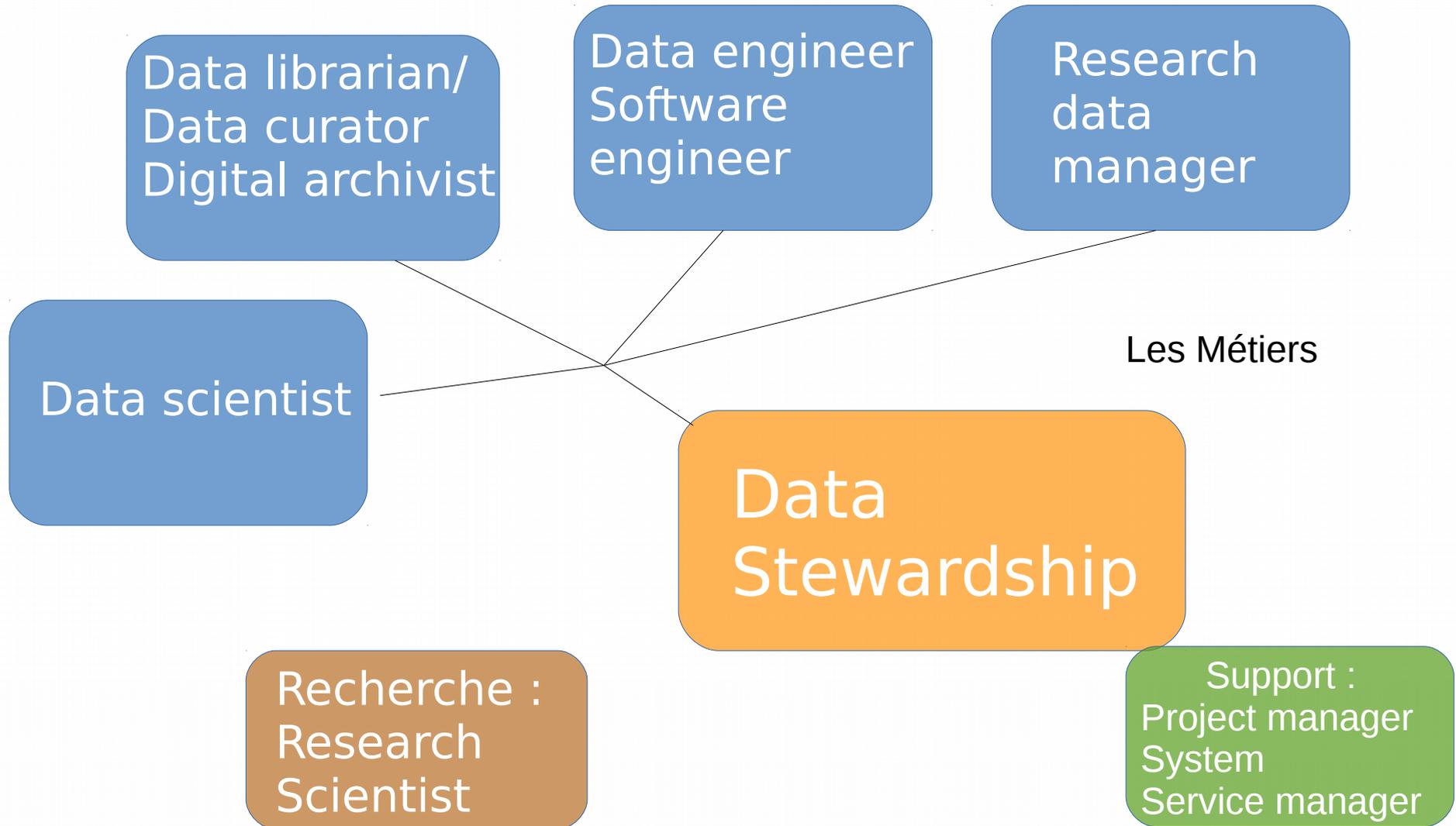
# □ Professionalising Data Stewardship



# □ Professionalising Data Stewardship



# □ Professionalising Data Stewardship



# □ Data .... qu'est-ce que c'est ?



- Le **Data Scientist** conçoit des modèles et des algorithmes pour collecter, stocker, traiter et restituer de la Data (organisation)
- Le **Data Engineer** développe, construit, teste et maintient des logiciels et des systèmes de traitement de données à grande échelle.
- Le **Data Analyst** analyse les données (exploiter, interpréter, ...)
- Le **Data Curator** traite et valorise les données pour les rendre accessibles, partageables et réutilisables de façon pérenne et assurer leur qualité.
- Le **Data Manager** est responsable de la politique des données, de leur exploitation et de leur protection

- **Le Data Steward**



## Data Steward, c'est qui ?

- Responsable de la planification et de l'exécution de toutes les actions sur les données (numériques) avant, pendant et après un projet de recherche, dans le but d'optimiser l'utilisabilité, la réutilisation et la reproductibilité des données résultantes.
- Le responsable du maintien de la qualité, de l'intégrité et des modalités d'accès aux données et aux métadonnées d'une manière conforme à la politique institutionnelle.
- Le coordonnateur de la gestion des données et assure le lien entre les différentes équipes (informatique, recherche et curateur de données).

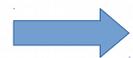
**Supervision, valorisation et animation**

Data  
Steward,  
c'est qui ?



Manque de consensus au niveau mondial sur les responsabilités, les connaissances, les compétences ce qui entraîne une certaine confusion quant à leur rôle et entrave le développement de capacités adéquates en la matière.

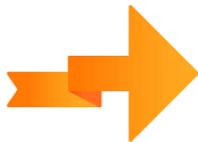
Un rôle peu reconnu



Nécessité de professionnaliser : rôle, compétence, formations, ..

## Data Steward, c'est qui ?

- Dépend surtout du service : labo, national ou international et du type de données et du contexte



Une personne  
ou  
Un **sport d'équipe** ?

# □ CDS : Ses objectifs



- 50 ans au service de la communauté scientifique
- Depuis 1972, les objectifs du CDS sont :
  - Collecter l'information « utile » sur les objets astronomiques
  - Améliorer cette information par évaluations critiques et comparaisons
  - Distribuer cette information à la communauté astronomique internationale
  - Conduire des recherches en utilisant ces données
- Approche « **bottom-up** »

# CDS : Ses activités

