

«Data Steward» per i dati FAIR

Valentina Pasquale¹, Emma Lazzeri², Elena Giglia³

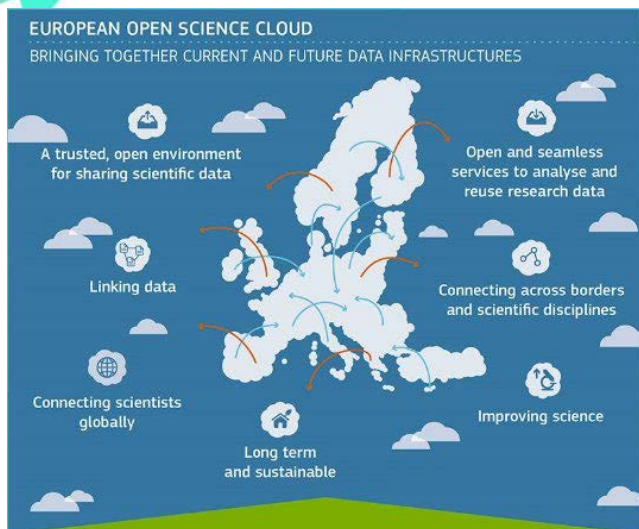
¹Istituto Italiano di Tecnologia, ²GARR, ³Università di Torino



Perché servono i «data steward»?

F indable
A ccessible
I nteroperable
R eusable

For humans
And
Machines*



U. von der Leyen's speech at #davos2020 (22nd January 2020)
https://www.youtube.com/watch?v=_A7Q514z_dw

We are creating a European Open Science Cloud now.
It is a trusted space for researchers to store their data and to access data from researchers from all other disciplines. [...]

Every researcher will be able to better use not only their own data, but also those of others. They will thus come to **new insights, new findings and new solutions.** [...]



Strategic Research and Innovation Agenda (SRIA)
of the
European Open Science Cloud (EOSC)
Version 1.0 15 February 2021

[EOSC SRIA v1.0](#)

7.4. Critical success factors

The developments and expected impacts described above will not happen spontaneously. For these benefits to materialise a number of critical success factors (CSFs) must be in place. The following CSFs have been identified for EOSC:

- Researchers performing publicly funded research make relevant results available as openly as possible;
- Professional data stewards are available in research-performing organisations in Europe to help implement FAIR principles and support Open Science;

*Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016).
<https://doi.org/10.1038/sdata.2016.18>

News Blog

◀ We need 500.000 respected data stewards to operate the European Open Science Cloud

[Interview to Barend Mons, Chairman of the High Level Expert Group on EOSC, May 2016](#)

04/05/16 09:08

At the e-IRG workshop in Amsterdam, we had the opportunity to talk to Barend Mons who is chairing the High Level Expert Group on the European Open Science Cloud, an advisory group to the European Commission. To be successful, the European Science Cloud needs a lot of experts to operate it, Barend Mons told us. Data stewards that have a lot of knowledge about managing and maintaining data. Experts who are well respected with a solid career path. Barend Mons also discussed several other findings of the Expert group, whose report will be published very soon.

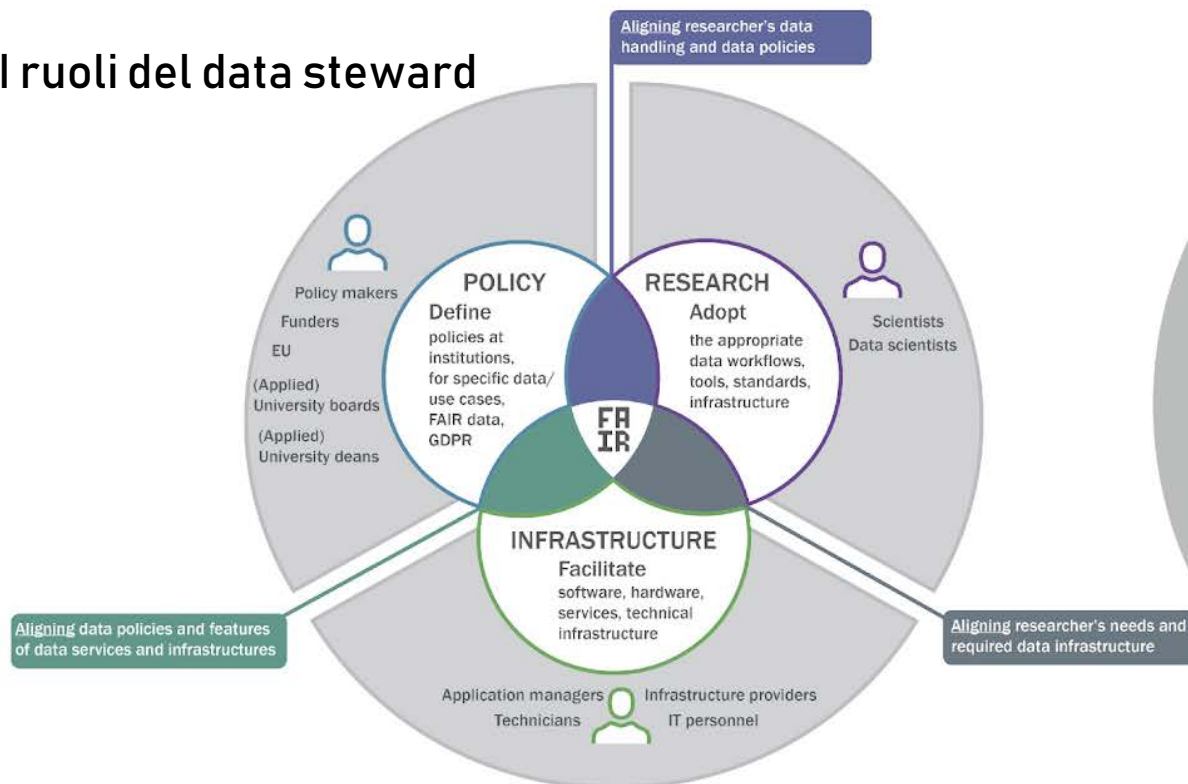
We are here at the e-IRG workshop in Amsterdam and we are talking with Barend Mons. Welcome. You just had a presentation here, and you were also part of the panel. One of the things that you focused on, was the European Open Science Cloud, because you are chairman of the Expert Group of the European Commission. So can you tell a little bit about the progress? What is the status of the European Science Cloud?



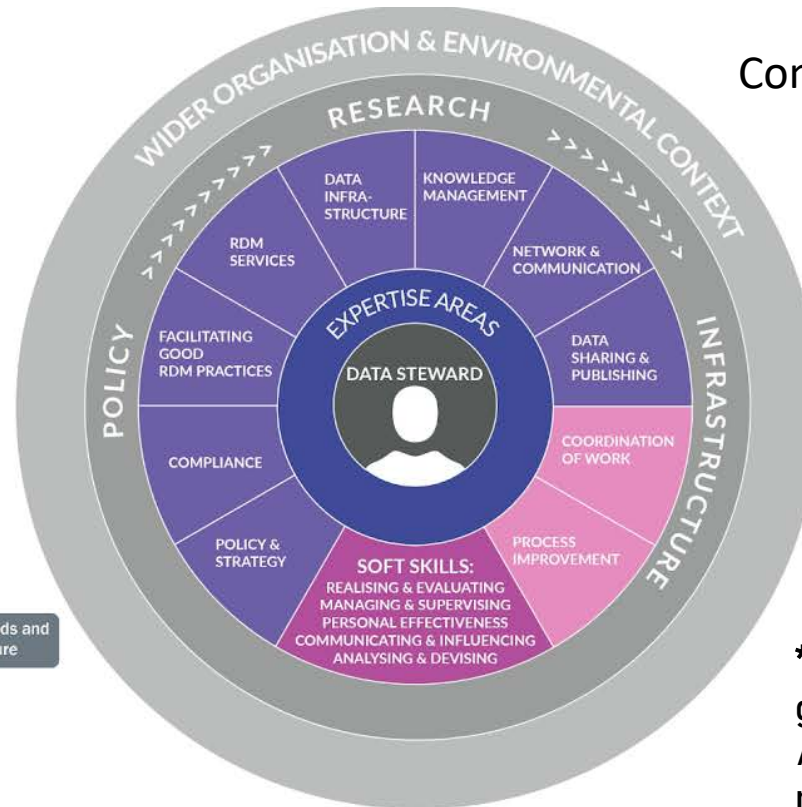
Chi è il «data steward*» (1)

Il «data steward» è una figura professionale competente nella pianificazione e gestione responsabile dei dati della ricerca durante tutto il loro ciclo di vita.

I ruoli del data steward



Competenze e abilità del data steward



***stewardship:** «pianificazione e gestione responsabile ed etica delle risorse».

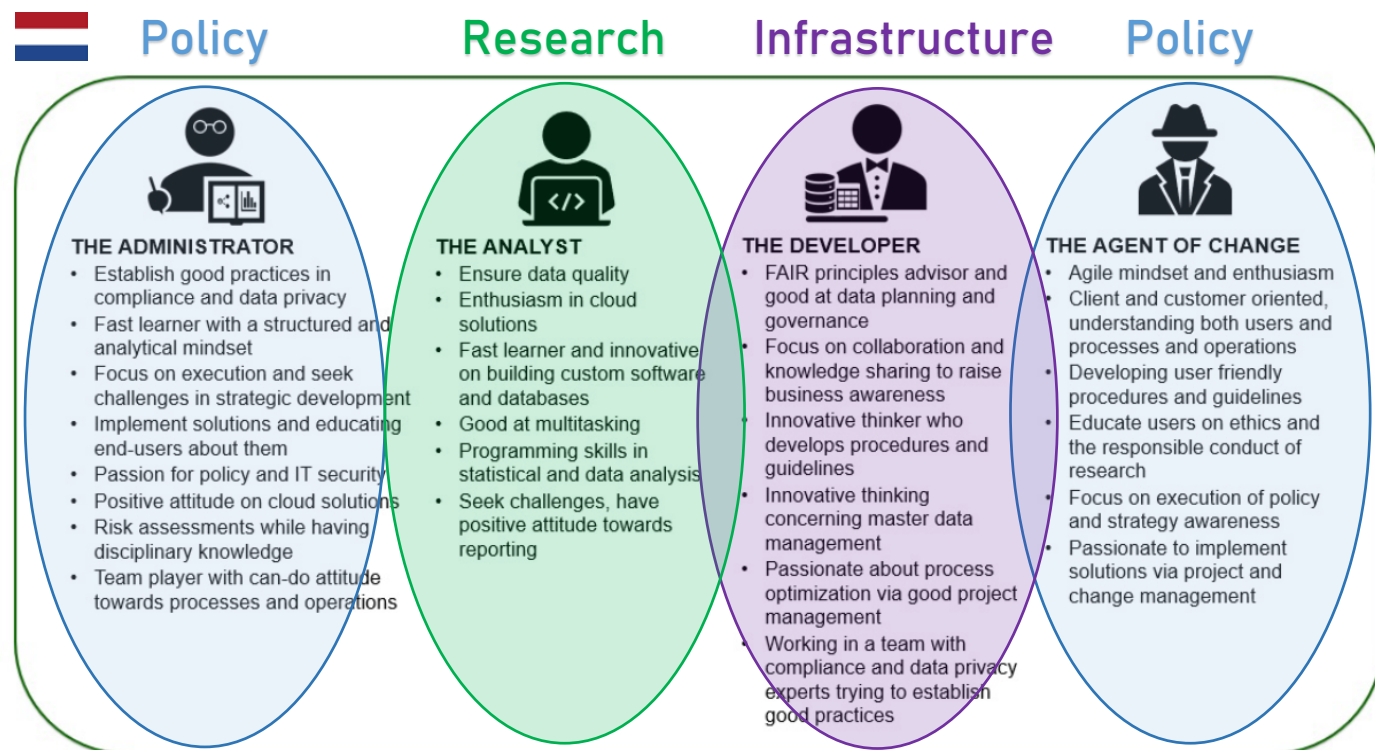
<https://doi.org/10.5281/zenodo.3243909>
<https://zenodo.org/communities/nl-ds-pd-ls>



Mijke Jetten et al. (2021). Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship. Zenodo.
<http://doi.org/10.5281/zenodo.4623713>

Chi è il «data steward» (2)

Profili professionali del data steward



Lorna Wildgaard et al. (2020). National Coordination of Data Steward Education in Denmark: Final report to the National Forum for Research Data Management. Zenodo.
<https://doi.org/10.5281/zenodo.3609515>

I «data steward»...

... hanno (preferibilmente) un PhD e possiedono nozioni su come i dati vengono gestiti in un dominio di ricerca specifico

... possibilmente hanno esperienza pregressa in programmazione, sviluppo software, gestione di database e infrastrutture di ricerca, sicurezza dei dati

... hanno buone capacità comunicative, di insegnamento e organizzative

... possiedono nozioni su aspetti legali della gestione dei dati (privacy, proprietà intellettuale) ed etici

... comprendono la psicologia dei ricercatori e parlano lo stesso linguaggio specifico

... desiderano intraprendere un percorso di carriera che non è né puramente scientifico né tecnico

Educazione dei «data steward» (1)

Modelli di formazione ed educazione dei data steward

Students with Bachelor degree

Directed corporate employment

- *One-year master programme*
For students who fulfill requirements to basic programming skills, study skills, subject knowledge and academic language level.
- *Two year master programme*
For students who do not meet the requirement for programming and study skills follow a *pre-master's year before the master's* and then continue with the one-year master.
- *Two-year candidatus*
Traditional university candidatus combining theory, methods and internships. A dissertation in the area of DS is required.

Students with a PhD or equivalent

Directed corporate or academic employment

For PhDs from any field at university faculties or part of research teams in industry. Educational programmes are a collaborative endeavor between faculties, library or knowledge centers, Centers for Information Security, Data Steward Community, Research coordinators, Project PI's, System Developers, Communication and Teaching teams, the Graduate School (for PhD training) & Human Research Ethics Committees.

- *Requirements:* Short courses, workshops, mentorships, online modules, summer school programmes & MOOCs, that immerse the student in the DS Community.

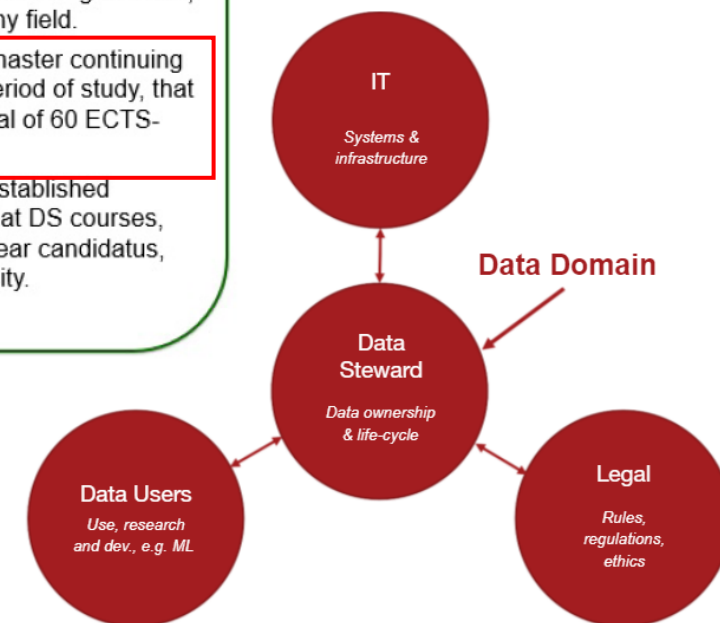
Continuing and professional education

Directed corporate or academic employment

Flexible master programme

- For students who have professional experience and wish to improve their DS skills but keep working full-time, or already have a master's degree in any field.
The programme is a part-time flexible master continuing education program with a prescribed period of study, that the student individually plans, and a total of 60 ECTS-points.
- The student combines elements from established programmes, thus it is a requirement that DS courses, such as the one year master and two year candidatus, are available through the Open University.

KØBENHAVNS
UNIVERSITET



Flessibilità, Integrazione, Formazione continua



Lorna Wildgaard et al. (2020). National Coordination of Data Steward Education in Denmark: Final report to the National Forum for Research Data Management. Zenodo.
<https://doi.org/10.5281/zenodo.3609515>

Martin Lillholm, *Data Steward Education*
Copenhagen University, *Challenges in*
professionalizing data stewardship BoF RDA P15

Educazione dei «data steward» (2)



Programme de la formation

Comprendre l'environnement des données de la science

- Module 1 : Territoires, innovation et financement
- Module 2 : Recherche, gouvernance et données
- Module 3 : Données, innovation et droit

En savoir +

Maîtriser les outils d'analyse des données de la science

- Module 1 : Extraire les données
- Module 2 : Gérer les données et les métadonnées
- Module 3 : Explorer et fouiller les données
- Module 4 : Sécuriser les données
- Module 5 : Protéger les données personnelles

En savoir +

Gérer l'ouverture des données de la science

- Module 1 : Construire un plan de gestion de données (PGD)
- Module 2 : Stocker et archiver les données
- Module 3 : Diffuser et partager les données
- Module 4 : Valoriser et réutiliser les données

En savoir +



- ✓ How to find data more efficiently
- ✓ How to publish data Software and data carpentry
- ✓ How to use Github
- ✓ Introduction to theory and practice of visualization
- ✓ Introduction to machine learning
- ✓ Grounding in use of computational research infrastructures

- ✓ Principles and practices of Open Science
- ✓ FAIR data and using data ethically
- ✓ An introduction to research data management and the research data lifecycle
- ✓ How to prepare a data management plan
- ✓ Introduction to persistent identifiers and licensing
- ✓ A grounding in how to compile bibliographies

Università di Trieste, Università di Udine, SISSA, ICTP

- Artificial Intelligence and Machine Learning
- Data Science and Applications
- Computational Science and Engineering

Master degree in



DATA SCIENCE & SCIENTIFIC COMPUTING

• Data Management and Engineering



Data stewardship in IIT

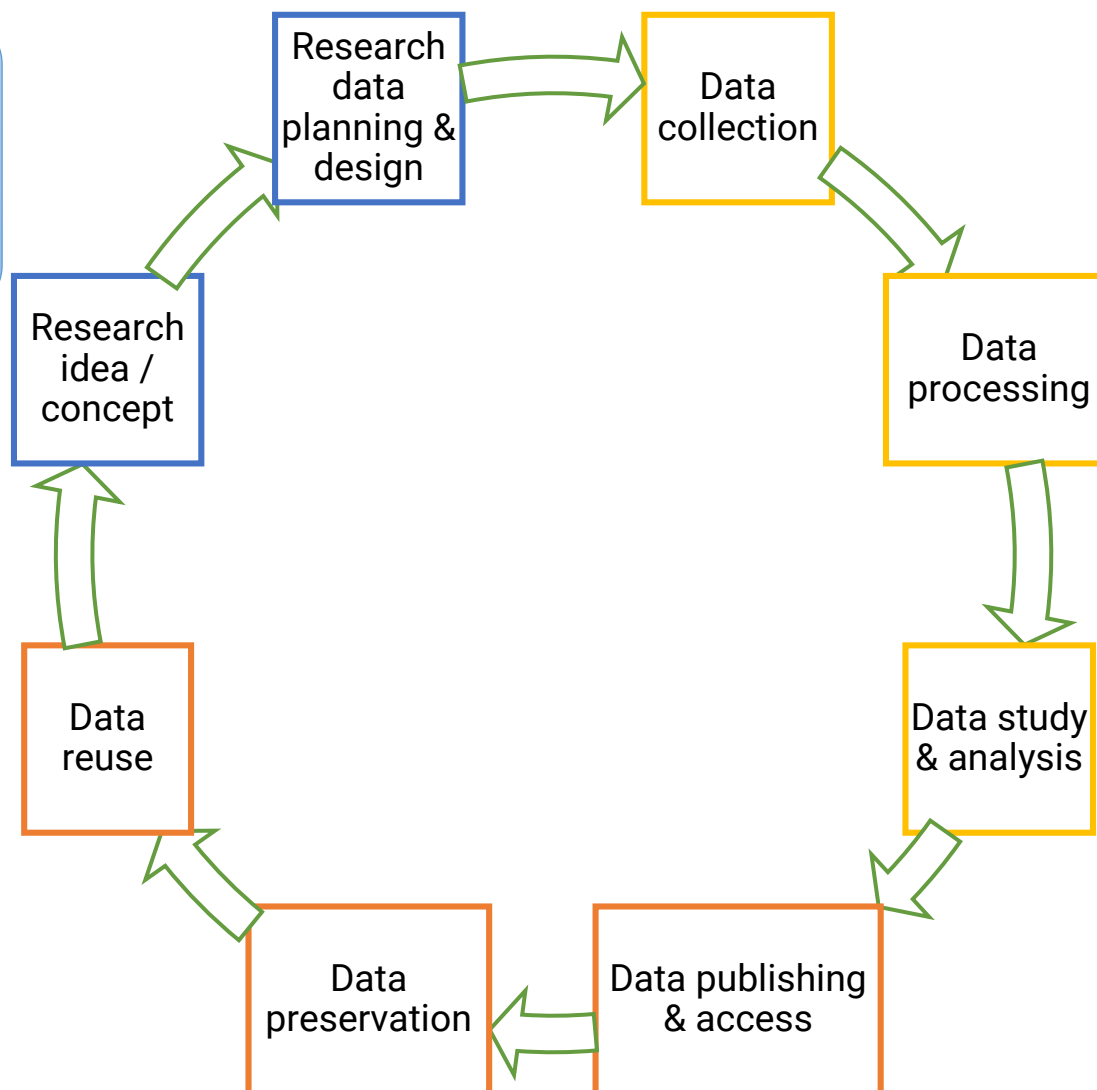
Support scientists in data planning

- Data management planning
- Collaboration tools
- Data protection / privacy

The **Dataverse**[®] Project
<https://dataverse.org/>

Institutional data repository

- Long-term data storage & preservation
- FAIR Data sharing
- Data curation & metadata



Active data support

- Short-term data storage
- Private file sharing

Foster a “data culture” with trainings and events

- Intranet guidelines
- Webinars
- PhD training

Prospettive italiane



6.2 IL PIANO NAZIONALE PER LA SCIENZA APERTA

Il Piano nazionale per la scienza aperta si struttura in quattro assi di intervento centrati sulle pubblicazioni scientifiche, sui dati della ricerca scientifica, sulla valutazione della ricerca e sul coinvolgimento dei ricercatori, enti di ricerca, infrastrutture per l'adozione delle pratiche di scienza aperta.

<https://www.mur.gov.it/sites/default/files/2021-05/PNR2021-2027.pdf>



La Task Force EOSC Competence Centre

<https://www.icdi.it/it/attivita/tf-cc>

Centro di competenza nazionale sui temi della Scienza Aperta, dei dati FAIR per la ricerca scientifica e per la partecipazione italiana in EOSC. Punto di riferimento nazionale per la formazione e la condivisione di buone pratiche in tema di data stewardship.

<https://www.icdi.it/it/attivita/tf-cc/open-science-cafe>



Serie di eventi formativi in tema Open Science