

# Project LEARN in practice: a bottom up case study

Paola Galimberti Università degli Studi di Milano  
Marisol Occioni Università Ca'Foscari Venezia

Open Research Data Management: policies and tools  
May 24-25

Milano, Università Statale



# From the LEARN toolkit (It. Version)

Nell'era digitale i dati della ricerca valgono oro. Dai sonetti alle statistiche, dai dati genetici ai geodati: la quantità di materiali creati e memorizzati in formato digitale cresce in maniera esponenziale. La LERU Roadmap for Research Data identifica tuttavia **gravi lacune per quanto riguarda il livello di preparazione delle organizzazioni attive nel campo della ricerca**. Questo divario emerge in modo evidente per quanto riguarda lo **sviluppo di politiche** di gestione dei dati, la **consapevolezza di problemi** ancora aperti, lo **sviluppo di competenze**, la **formazione**, i **costi**, la creazione di comunità, la governance e le **differenze disciplinari/legali terminologiche e geografiche**.

# From the LEARN toolkit

- **Ogni organizzazione** che svolge attività di ricerca dovrebbe **dotarsi di una propria politica dei dati**, per poter definire un quadro di riferimento su come curare e gestire i dati della ricerca. **Anche i finanziatori** della ricerca **dovrebbero dotarsi di una tale policy**, per stabilire gli obblighi che i ricercatori sono tenuti ad adempiere come premessa per la concessione dei finanziamenti. LEARN ha creato un modello di “policy” per la politica di gestione dei dati della ricerca (Research Data Management) per le organizzazioni che svolgono attività di ricerca, oltre a delle linee guida per l’attuazione di questa policy. Il modello fornito da LEARN può essere adattato e adottato sia da singole organizzazioni di ricerca, che da associazioni attive nella ricerca a livello regionale, nazionale e/o internazionale.

# Needs

Our researchers work in the European Research Area

They need

- Policies
- Guidelines
- Instruments
- Skills



# RDM and the Italian context



# The Italian context

- Lack of policies
- Lack of guidelines
- Lack of instruments
- Lack of awareness of the importance of RDM
- Lack of awareness of the value of data
- No contact person

# An informal working group

- Politecnico di Milano
- Università Ca'Foscari Venezia
- Università di Milano
- Università di Torino
- Università di Trento

Different skills and different roles:

IT service, Digital library, Research Support Service, Legal service, Research evaluation

# Method

3 WG: main coordinator Paola Gargiulo (OpenAIRE- NOAD).

Group n. 1: deliverable → Policy model, (Università di Milano)

Group n. 2: deliverable → Template RDM, (Università Ca'Foscari Venezia)

Group n. 3: deliverable → E-Infrastructure, (Politecnico di Milano)

- Each member decided which group to belong to
- No meetings in presence, only on-line, mailing list, wiki, materials shared on google drive
- 6 Plenary sessions from April to October 2016



# Organisation

## Background:

- LERU Roadmap for Research Data
  - results of the LEARN project, policies of the University of Edinburgh and of UCL, Austrian group e- infrastructures
  - local interviews on research data management and researchers' needs and habits (data type, size, archiving, long term preservation)

## Timing:

- first draft: July 2016
- second draft: August 2016
- final document: October 2016

## Next steps:

- Internal discussion – Approval by SA (can take a long time)

# Main Issues (according to the LEARN toolkit)

- Introduction (why?)
- A definition of research data (what?)
- Processing of research data (where, how long?)
- Ownership and legal issues
- Responsibilities

# A definition of *research data*

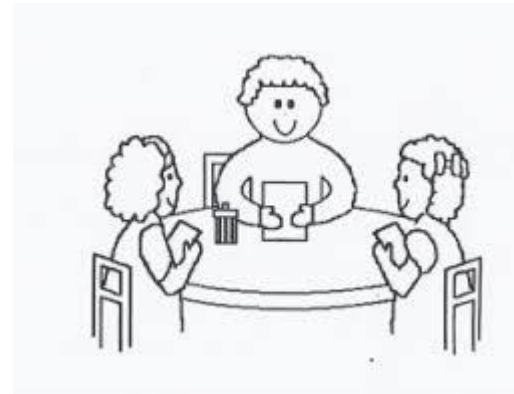
- Ai fini della presente policy, sono “dati della ricerca” le informazioni, in qualsiasi formato, utilizzate, secondo un protocollo definito nell’ambito di una specifica attività di ricerca condotta da afferenti all’università ..... (quali professori, ricercatori, dottorandi, assegnisti, borsisti, tecnici amministrativi coinvolti nelle attività di ricerca). A titolo meramente esemplificativo sono da intendersi dati della ricerca risultati (positivi o negativi) di tutti gli esperimenti rilevanti per la ricerca stessa fatti, osservazioni, esperienze, fonti edite ed inedite, riferimenti bibliografici, software e codice, testi, oggetti, del presente o del passato raccolti o creati in formato digitale e o cartaceo, secondo un protocollo definito nell’ambito di una specifica attività di ricerca, condotta da professori, ricercatori, dottorandi, assegnisti, borsisti personale tecnico amministrativo attivi nell’Università ..... I dati della ricerca possono essere espressi in formato numerico, descrittivo, audio o video. Possono essere grezzi o elaborati.

# Policy and Guidelines

- Work in progress
- Guidelines are necessary for the implementation of the policy

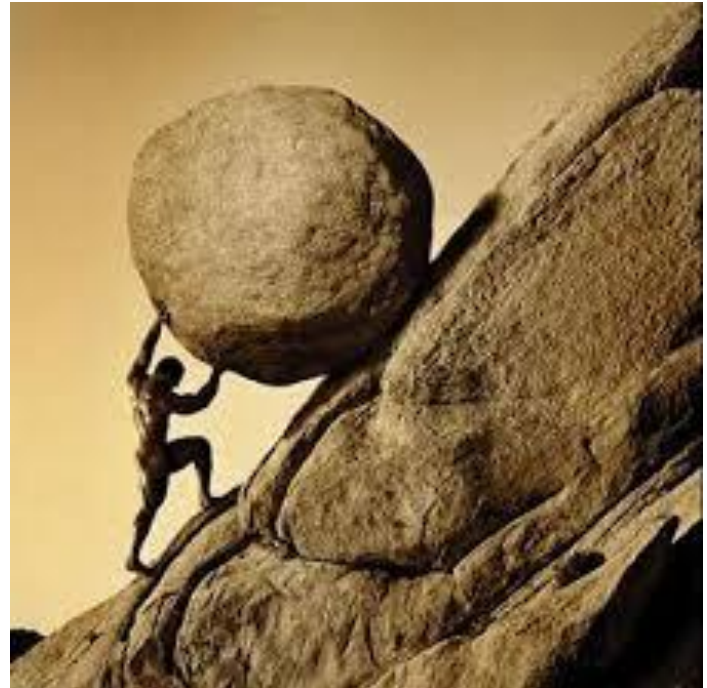
# Strengths

- Bottom up process
- It is easier to work in a small group without restrictions or obligations



# Weakness

- Having worked without endorsement makes the next steps very difficult



# Why adopt a Data Management Plan?

**FUNDERS** around the world have introduced

**POLICIES** requiring that their funded researchers make data available to others in a timely and responsible manner, and

**PLAN** their approach for managing data as an **integral part** of planning their research.

LEARN Toolkit of Best Practice for RDM, april 2017

# DMP: different worries same troubles





# Funder and cooperation

Same goal



Achived through  
cooperation



# DMP checklist

## **Administrative plan details**

Project ID: University logo, Project name, description, acronym, PersistentID, DMP creator, ORCID, Version & date, etc.

## **Dataset Description (useful 4 \$)**

Provenance of data & metadata, type, nature, formats & consistency, requirement for sw & hw, data re-use & integration, new dataset value, audio-visual materials, sw application

## **Standards & Metadata (quality assurance, data cleaning)**

Metadata standard & data documentation, methodologies for data collection generation, data quality & standards, metadata quality (re-use)

## **Data Management, documentation & curation (preservation plan, useful 4 \$)**

Managing, storing & curating data, data preservation strategy & standards (ex: Personal Data Protection Code, Legisl. Decree 30.6.2003 n. 196; Annex B Technical Specifications concerning minimum security measures; European Parliament Directive 2002/58/EC, Data protection and privacy ethical guidelines CE 2009)

## **Data security & confidentiality of potentially disclosive information (useful 4 \$)**

Confidential & sensitive data, data security standards, risks level, ethics & legal compliance, legal permission to obtain and process data, informed consent documents, national security (ex: Data protection and privacy ethical guidelines CE 2009, Data Protection Directive 95/46/EC)

## **Data sharing and access**

Sharing, discoverability, access, exclusive use, restrictions/delays, responsibilities, licences, repository

## **Responsibilities & resources (useful 4 \$)**

Who is responsible for: data capture, metadata creation, quality assurance, data security, archiving & sharing, policy.

Is the project compliant with: the policy of research data, legal and ethical requirements. Additional resources planned for the project

## **Relevant institutional policies on data sharing/security**

Research Data Policy, OA Repository. Report: where, how, who DMP is deposited, accessible, responsible for preservation

# Glossary

- Data definition
- Metadata quality
- What is DMP (living document: should be updated at least every six months, machine-actionable format)
- Machine-actionable DMPs: FAIR for machine as well a people
- The goal of DMP

# Link

- The costs of data curation
- Data Management costs
- European Science Cloud Initiative
- H2020
- Legal Aspects
- Machine-actionable DMPs
- Metadata standards
- OpenAIRE
- Wiki OA-Italia

# Where to?

Italian Policy & DMP Checklist:

[http://wikimedia.sp.unipi.it/index.php/OA\\_Italia/Risorse  
sugli open research data#Italy](http://wikimedia.sp.unipi.it/index.php/OA_Italia/Risorse_sugli_open_research_data#Italy)

# Desirable Future Steps

**Research  
Integrity Code  
of Conduct**

**Research Data  
Services**

**RDM website**

**Costing tools  
for DMP**

**Support**

**Training  
Programme:**  
staff members,  
researchers, Phd  
students, Librarians

**E-Infrastructure  
for underlying  
research data**

**Data  
Stewardship**

# THANKS

Paola Galimberti [paola.galimberti@unimi.it](mailto:paola.galimberti@unimi.it)

Marisol Occioni [occioni@unive.it](mailto:occioni@unive.it)

