La gestione degli *Open* (*Big*) *Data* nei progetti di ricerca

Augusto Celentano, DAIS
Delegato alla gestione delle banche dati di Ateneo

Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike

- availability and access

- re-use and distribution

- universal participation

- interoperability

The focus is on non-personal data, that is, data which does not contain information about specific individuals

For some kinds of government data, national security restrictions may apply

opendatahandbook.org

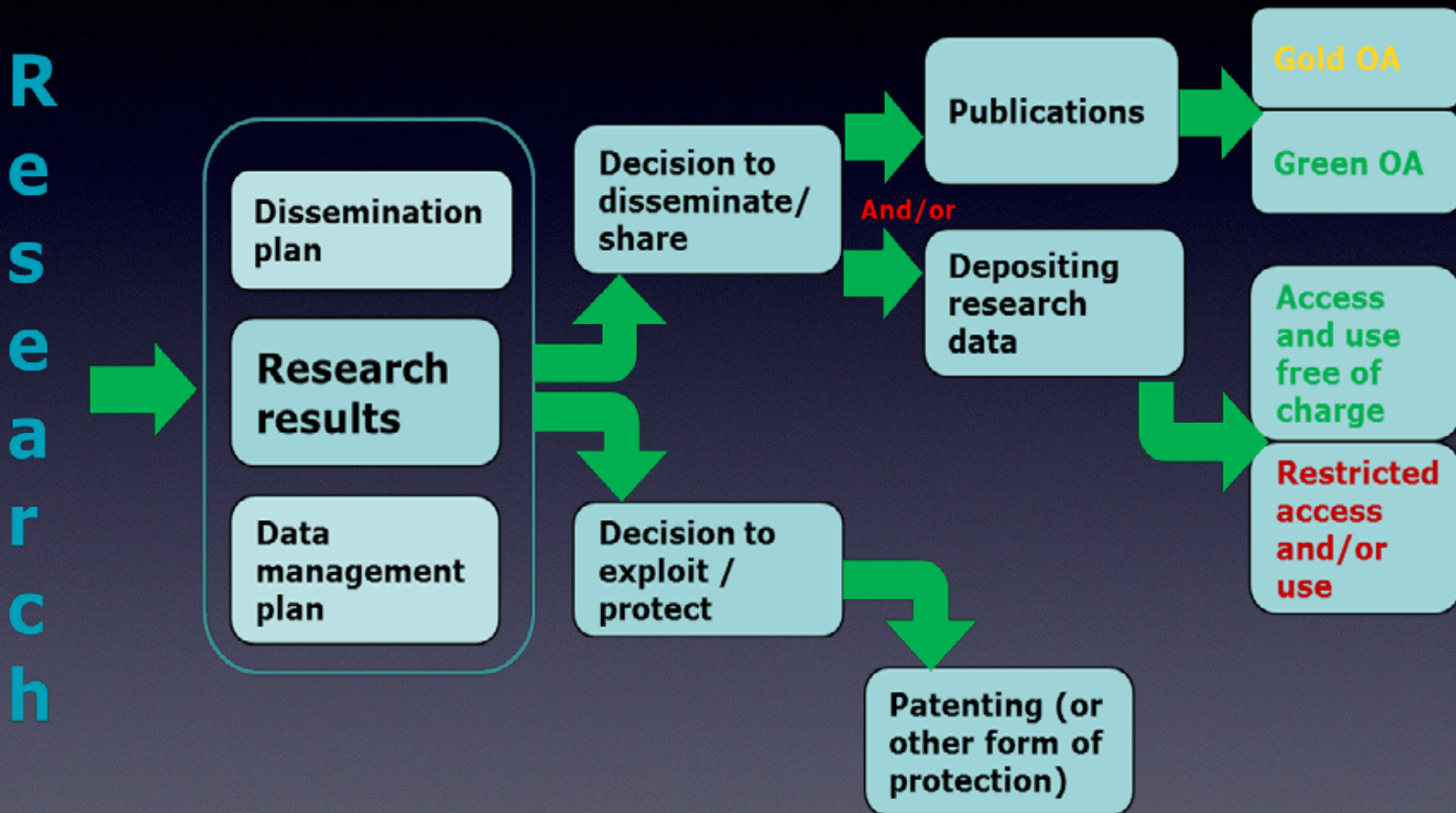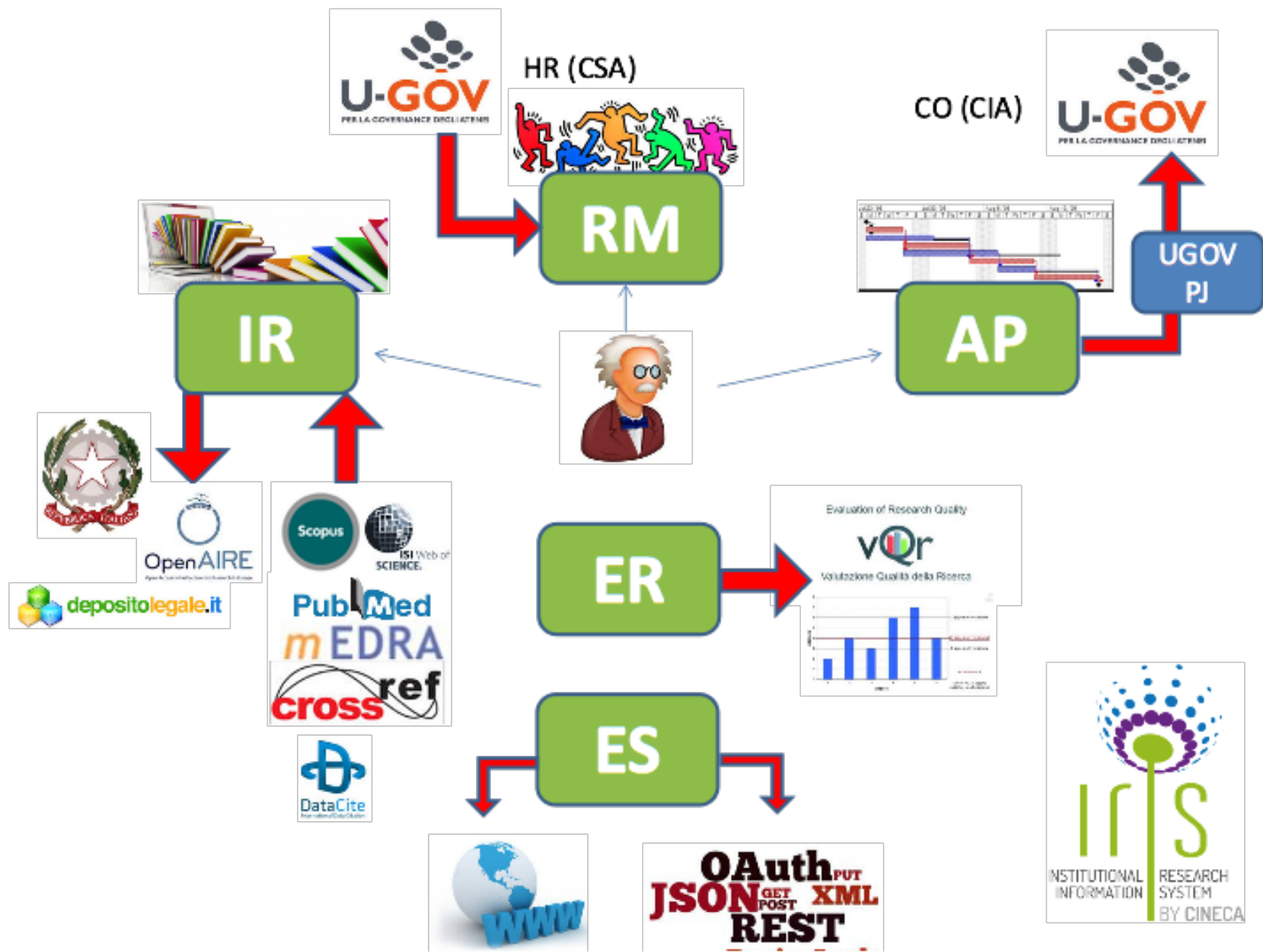# *Il modello di gestione dei risultati di Horizon 2020*

HR (CSA)

U-GOV

RM

CO (CIA)

U-GOV

IR

AP

UGOV
PJ

OpenAIRE

depositolegale.it

Scopus
ISI Web of SCIENCE.
Pub Med
mEDRA
cross ref
DataCite

ER

Evaluation of Research Quality
vQr
Valutazione Qualità della Ricerca

ES

WWW

OAuth PUT
JSON GET POST XML
REST

iris
INSTITUTIONAL RESEARCH
INFORMATION SYSTEM
BY CINECA

Cerca

Tutte le collezioni

Ricerca libera

28.000 oggetti digitali: immagini, documenti, libri, risorse per la didattica.

🔍 Ricerca avanzata »
Ricerche precedenti »

Iconoteca dei geologi

Comprende ritratti di geologi, naturalisti e scienziati italiani e stranieri. La collezione è conservata presso l'archivio della Biblioteca di Geoscienze dell'Università di Padova ed è costituita da 115 fotografie raccolte nell'album dedicato a Giovanni Omboni con 42 dediche manoscritte, oltre a un numero ancora non quantificato di ritratti e immagini a stampa.

Ex libris e segni di possesso dalle collezioni della Biblioteca medica "V. Pinali" antica

Archivio di immagini relative ai segni materiali lasciati nel corso dei secoli da diversi possessori sui libri che fanno oggi parte del patrimonio della biblioteca, costituitasi a partire dal 1875 coi fondi donati all'Università di Padova dal clinico Vincenzo Pinali: ex libris e super libris, timbri, note manoscritte di dedica o di possesso, di lettura, di acquisto e altri elementi distintivi.

Guidelines on Data Management in Horizon 2020

Where relevant, applicants must provide a short, general outline of their policy for data management, including the following issues:

- What types of data will the project generate/collect?

- What standards will be used?

- How will this data be exploited and/or shared/made accessible for verification and re-use? If data cannot be made available, explain why.

- How will this data be curated and preserved?

The Open Research Data Pilot applies to two types of data:

1. the data, including associated metadata*, needed to validate the results presented in scientific publications as soon as possible;

2. other data, including associated metadata**, as specified and within the deadlines laid down in the data management plan.

   * describing the research data deposited

   ** not directly attributable to a publication, or raw data

Horizon 2020 Open Research Data Pilot

Step 1: participating projects are required to deposit the research data described above, preferably into a research data repository. 'Research data repositories' are online archives for research data. They can be subject-based/thematic, institutional or centralised. … In addition, it is expected that the Open Access Infrastructure for Research in Europe (OpenAIRE) will become an entry point for linking publications to underlying research data.

Step 2: as far as possible, projects must then take measures to enable for third parties to access, mine, exploit, reproduce and disseminate (free of charge for any user) this research data. One straightforward and effective way of doing this is to attach Creative Commons Licence (CC-BY or CC0 tool) to the data deposited.

Horizon 2020 Open Research Data Pilot

At the same time, projects should provide information via the chosen repository about tools and instruments at the disposal of the beneficiaries and necessary for validating the results, for instance specialised software or software code, algorithms, analysis protocols, etc. Where possible, they should provide the tools and instruments themselves.

Projects are required to deposit the research data …

- tipologia dei dati (es. testi, tabelle, dati numerici, immagini, suoni, video, stream, serie, dati eterogenei)

- formato dei dati (elettronici, fisici)

- standardizzazione (ufficiale, de facto)

- compatibilità

- persistenza, coerenza temporale

- numerosità dei dati ➤ Big Data

Alcuni problemi

Projects must then take measures to enable for third parties to access, mine, exploit, reproduce and disseminate …

- access: come presentare dati eterogenei

- mine: dipende da tipologia e formato

- exploit: in base a quali obiettivi?

- reproduce: ripetizione degli stessi esperimenti?

- disseminate: i dati originali o le repliche?

Projects should provide information via the chosen repository about tools and instruments at the disposal of the beneficiaries …

- disponibilità

- durata

- compatiblità

- operatività

- strumentazione

- standard

Where possible, they should provide the tools and instruments themselves …

- qualità degli strumenti

- trasportabilità in altri ambienti operativi

- compatibilità

- sicurezza

- normative locali

Open Access vs Open Data

By *open access* to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

budapestopenaccessinitiative.org

Open Access
=
Open Data

Data

Open Access