



Area Data Curators

I'esperienza di AreaSciencePark

| Stefano Cozzini |



PREMESSA: ISTITUTO RICERCA E TECNOLOGIE (RET)

Obiettivo

- gestire le piattaforme di ricerca

Investimento

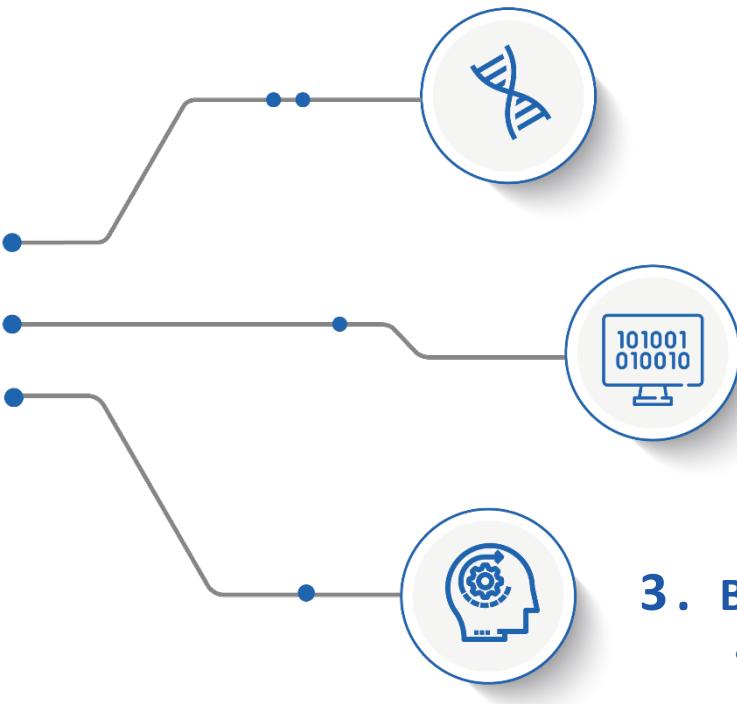
- 3.5 Mln euro

Visione

- creare un **ecosistema** virtuoso per ricerca e impresa



ELEMENTI DELL'ECOSISTEMA



1. Hardware

- Piattaforma biologia strutturale
- Piattaforma materiali innovativi
- Piattaforma Genomica e Epigenomica

2. Software

- Codice e Dati

3. Brainware

- Persone

PIATTAFORMA GENOMICA ED EPIGENOMICA: 2 COMPONENTI

PIATTAFORMA GENOMICA ED EPIGENOMICA



1. LABORATORIO GENOMICA
ED EPIGENOMICA (LAGE)

2. ORFEO DATA CENTER

DATI PRODOTTI



NOVASEQ 6000

145
corse

~100
TB dati
prodotti

15
corse

~60
TB dati
prodotti



PROMETHION 24

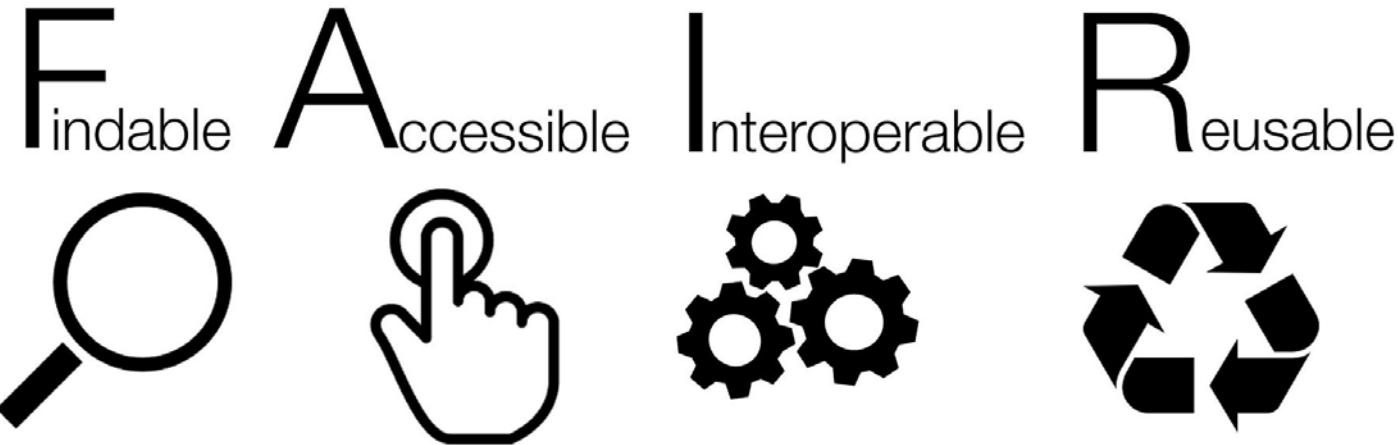


THE ECOSYSTEM

GENOMIC DATA SCIENCE

- **HPC & DATA INFRASTRUCTURE MANAGEMENT**
(Advanced services for Users)
- **SCIENTIFIC DATA MANAGEMENT**
(FAIR access to genomic data)
- **ADVANCED DATA ANALYSIS SERVICE**
(by means of ML/DL algorithms)

“FAIR PRINCIPLES” TO THE RESCUE



Door SangyaPundir - Eigen werk, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=53414062>

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

“FAIR” definitions from : <https://www.go-fair.org/fair-principles/>



LA SFIDA PIU' COMPLESSA...

formazione delle risorse umane interne

creare una squadra interna in grado di indirizzare le esigenze contingenti e future relative alla gestione FAIR del dato genomico.

"BOL - BIO OPEN LAB" PROJECT - *STRENGTHENING THE HUMAN CAPITAL OF RESEARCH INFRASTRUCTURES*

- **RESEARCH AND INNOVATION PLAN 2015-2017**
- **"PNIR - Programma Nazionale Infrastrutture di Ricerca"
(National Research Infrastructure Programme)**

PARTNERS:

- *Consortium for the area of scientific and technological research, Trieste (LP)*
- *University of Salerno (PP)*
- *University of Salento (PP)*

Start date: **18.11.2020**

Closing date: **by 31/12/2024**

Duration: **50 months**

- Funding granted: **1,963,575.20 €**

PROJECT OBJECTIVES (OF)

1

to conduct technical and scientific activities aimed at enhancing the facility relating to each OR

2

to create a team of individuals with transversal professional skills relevant to the OR, capable of managing a distributed research infrastructure

3

to create specialist profiles to develop cooperation networks aimed at enhancing strengthened infrastructure

4

to train a class of researchers familiar with the FAIR approach to the scientific data produced, common to all the ORs of the distributed infrastructure

5

to train a new professional data curator, who supports the infrastructure and manages and handles the scientific data produced

BRAINWARE: LE PERSONE

2 Project Manager

(Olga Durì
Elizabeth Visintin)

3 data curators

(Edith Natalia Villegas
Federica Bazzocchi
Samuele Maria Semi)

4 Data Scientist

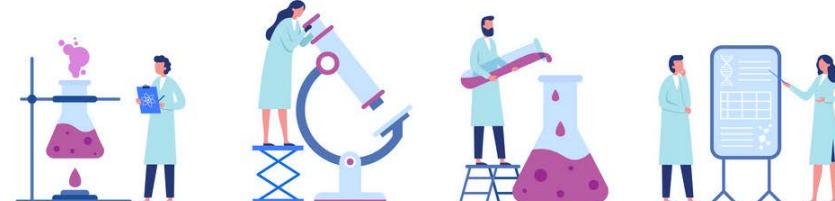
(Alessio Ansuini
Elena Tea Russo
Alberto Cazzaniga
Francesca Cuturello)

4 Biotecnologi

(Danilo Licastro/
Simeone Dalmonego
Margherita Degasperi
Marco De Bortoli)

4 data engineers

(Tommaso Rodani
Eros Fabrici
Federico Barone
Ezio Corso)



easydmp – creazione di DMP



Questionnaire TODOS List Metrics Preview Documents Settings

Current Phase: Before Submitting the Proposal

Chapters:

- I. General Informations
- II. Describe your research
- III. Storage and Security
- IV. Archiving and Sharing**
- V. Resourcing

IV. Archiving and Sharing

Chapter text

1 What resources should be kept as evidence of your research?

2 Where will these resources be hosted? + Add

3 When will the resources be made available? (choose one or more)

4 How will you make other researchers aware that the resources exist?

5 What steps will you take to ensure resources are easy to analyse and use in future research?

6 If resources can be made available, but not openly, what conditions on access/use must be met? + Add

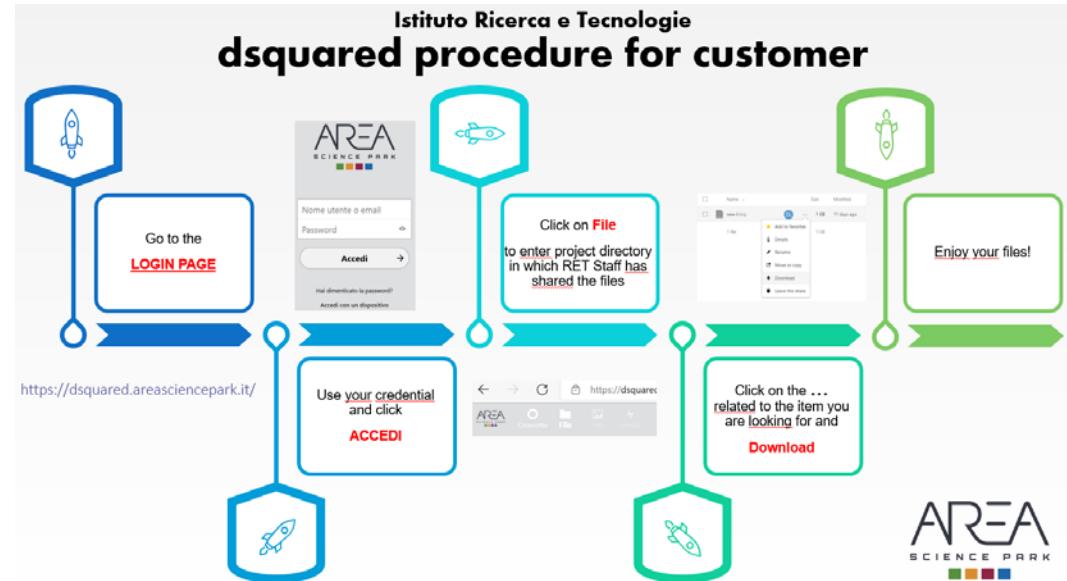
easydmp.areasciencepark.it è il tool per la redazione di DMP in maniera guidata, che guida i gruppi di ricerca nella generazione e nell'analisi del DATO.

Il risultato è un documento «live», di informazione e consapevolezza, conforme ai programmi Horizon Europe.

- ✓ Administrative
 - ✓ Metadata
 - ✓ Type of upload
 - ✓ Data origin
 - ✓ Data informations
 - ✓ Preserving data
 - ✓ Giving access to data
- 

dsquared – condivisione sartoriale

**dsquared è
condivisione di file personalizzata
garantendo i massimi livelli di
cybersecurity per l'organizzazione
e partner.**



dcubed – pubblicazione dei dati

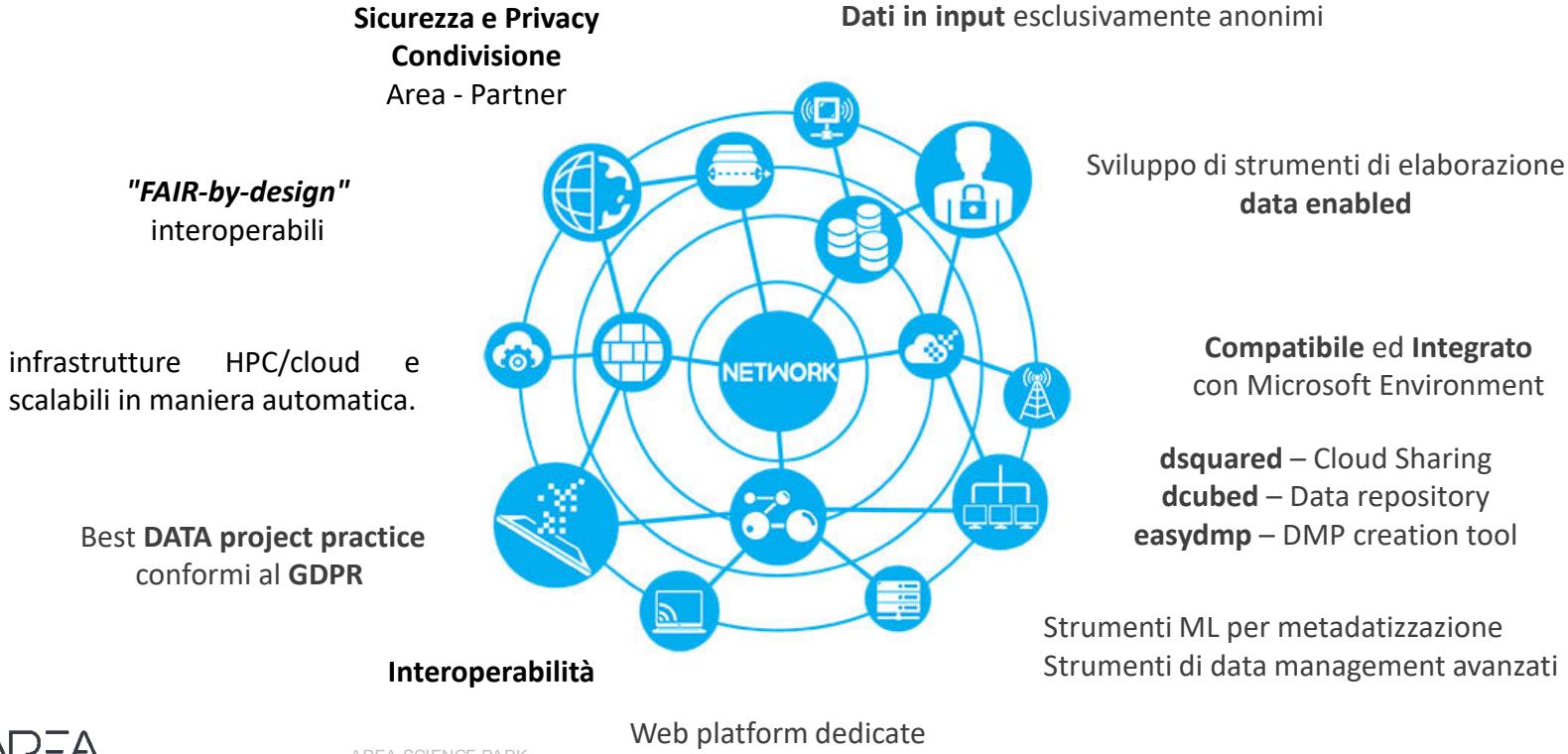
Il sistema di **gestione dati e metadati**, per lo sviluppo agile di ricerche future.

Il servizio permette di:

- corredare i dataset con un pacchetto di **metadati** conforme a **Dublin Core™ Metadata**
- attribuire un ***persistent identifier DOI*** a Dataset
- **pubblicare Dataset**

The screenshot shows the dcubed data repository interface. At the top, there is a header with the AREA Science Park logo and navigation links for Search, User Guide, Support, and a user profile icon. Below the header, the AREA Science Park logo is displayed again. The main content area has tabs for Metrics and 21 Downloads, with Contact, Share, and Edit buttons. A search bar with placeholder "Search this dataverse..." and an Advanced Search button is present. On the left, there are filters for Dataverses (2), Datasets (11), and Files (24). Under Dataverse Category, there is a link to Department (2). Under Publication Year, there is a link to 2021 (2). Under Subject, there are links for Other (2), Arts and Humanities (1), Computer and Information Science (1), Earth and Environmental Sciences (1), and Engineering (1). The results section shows two entries: "ISC (AREA SCIENCE PARK) Sep 30, 2021" and "RET (AREA SCIENCE PARK) Sep 30, 2021". Each entry includes a thumbnail image of a starry sky, the dataset name, the date, and a brief description. There are also "Sort" and "More..." buttons at the bottom of the results table.

Gli ambiti di sviluppo





GRAZIE PER L'ATTENZIONE!

