



Open Data: la prospettiva di un ricercatore

Marco J. Morelli
Head of Bioinformatics
Center for Omics Sciences
IRCCS Ospedale San Raffaele



@CosrLab





Il ricercatore (in campo biomedico)



Teoria



Pratica





Il ricercatore (in campo biomedico)



Ricerca/clinica



Burocrazia
Articoli



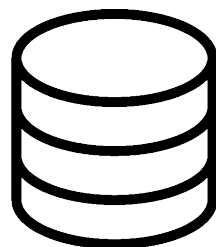
Fund raising
Rendicontazione



Supervisione
Mentoring
Insegnamento



Riunioni
Networking



DATI?

Pratica

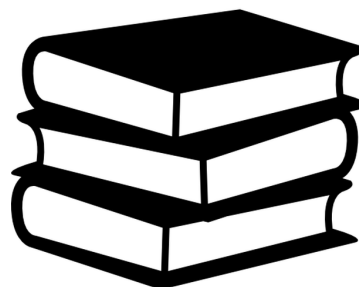
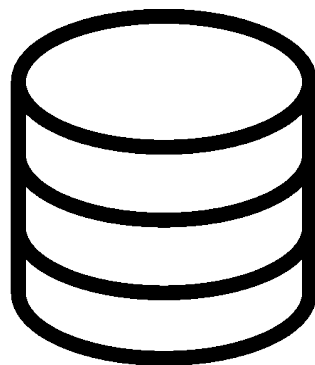




Molti uffici coinvolti in maniera tangenziale



GDPR

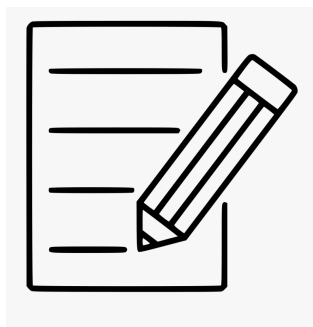




Prima di un progetto



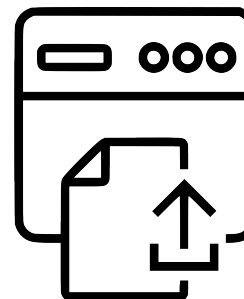
Idea



Scrittura



Feedback



Invio



Approvazione

Data Management Plan!

Richiede conoscenza dei dati molto profonda -> bisogna pensarci -> tempo
Coinvolge molti uffici -> no coordinazione -> tempo perso -> opt out

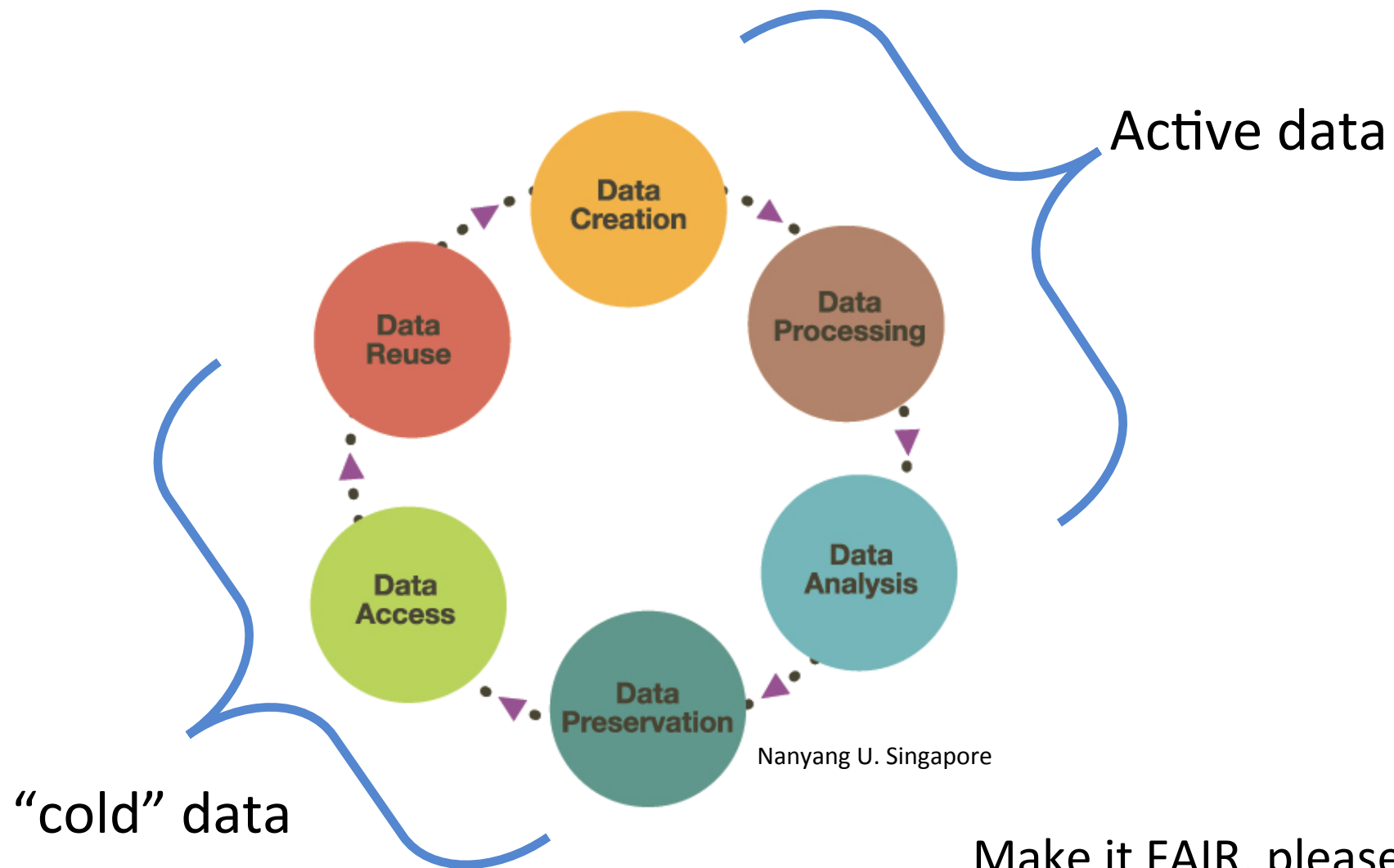
Opportunità mancata per mancanza di supporto



Durante un progetto



- Il data management è generalmente delegato a ricercatori junior
- Crisi: il mio postdoc se n'è andato!!!
- Metadata
- Responsabilità e riproducibilità



Make it FAIR, please!



Dopo un progetto: dubbi (real life stories)



Open data: obbligo vs opportunità (teoria dei giochi)

Non bastano i supplementary material?

Dove li metto? E come?

Ma anche quelli unpublished? Li vorrei usare in futuro (sono miei!)

E se li voglio usare per un brevetto? Se sono dati sensibili di pazienti?

PS: 'puoi scaricare e analizzarmi i dati di questo paper?'





Riluttanza al condividere dati:



- Perché? Che benefici ne ho?
- Mancanza di nozioni e abilità in data management



- A disagio nel condividere dati potenzialmente confidenziali
- Dataset di bassa qualità
- Paura delle critiche
- Conseguenze legali e mancanza di chiarezza sulle responsabilità



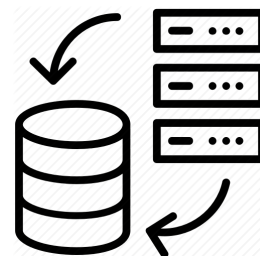
- Il dato è mio e me lo gestisco io



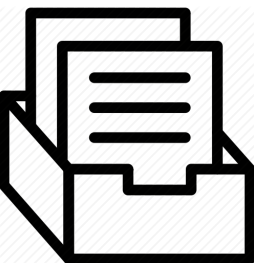
Necessità di policies chiare per la gestione dei dati

- Organizzazione
 - Nomi file e formattazione
 - Formati dati e software
 - Trasferimento file, sharing e accesso remoto
 - Controllo delle versioni

- Amministrazione
 - Back-up
 - Documentazione e metadati
 - Controllo degli accessi
 - Sicurezza



- Archiviazione e condivisione
- Aspetti etici e legali della manipolazione e possesso dei dati
- Protezione della proprietà individuale





L'inizio di un lungo viaggio:

1. Sensibilizzare e ottemperare alle indicazioni, in essere e future, delle funding/policy agencies e delle istituzioni pubbliche. Iniziative attivate (a vari stadi di completamento):

- Messa a punto criteri Data Management Plan (DMPs) -> sharing/reusability
- Registrazione, trasparenza e comunicazione dei risultati dei trial clinici OSR
- Procedure di research data management (RDM) vincolanti, stringenti e ispirate ad OS
- Formazione ed informazione su data sharing, preprint posting, social networking, open educational resources
- Trasparenza e rigore nella metodologia scientifica - Politiche di Research Integrity
- Disseminazione e outreach
- Sito web con FAQ



2. Introdurre elementi di OS nelle procedure interne di valutazione della ricerca:
 - Valutazione delle attività di science policy, outreach e disseminazione, contatti con organizzazioni di pazienti, data sharing, preprint posting, social networking, open educational resources, etc.
3. Attenzione alle disclosure in tutte le comunicazioni scientifiche (paper, comunicazioni a congresso, disseminazione, etc.)
4. Arruolamento di un RDM team e di 'data champions' che dia supporto pratico e operativo + responsabilizzazione uffici competenti.

VS.





Grazie per l'attenzione!