

Archiviazione, disseminazione e riuso dei dati: a che punto siamo?

Gli archivi di dati per le scienze sociali: l'esperienza di UniData – Bicocca Data Archive

Carlo Pisano

UniData – Bicocca Data Archive
Università degli studi di Milano-Bicocca

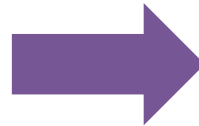
Torino, 23 Novembre 2017

Cosa sono gli archivi di dati per le scienze sociali?

Infrastrutture che operano nel campo della *condivisione* e *riutilizzo* dei dati.

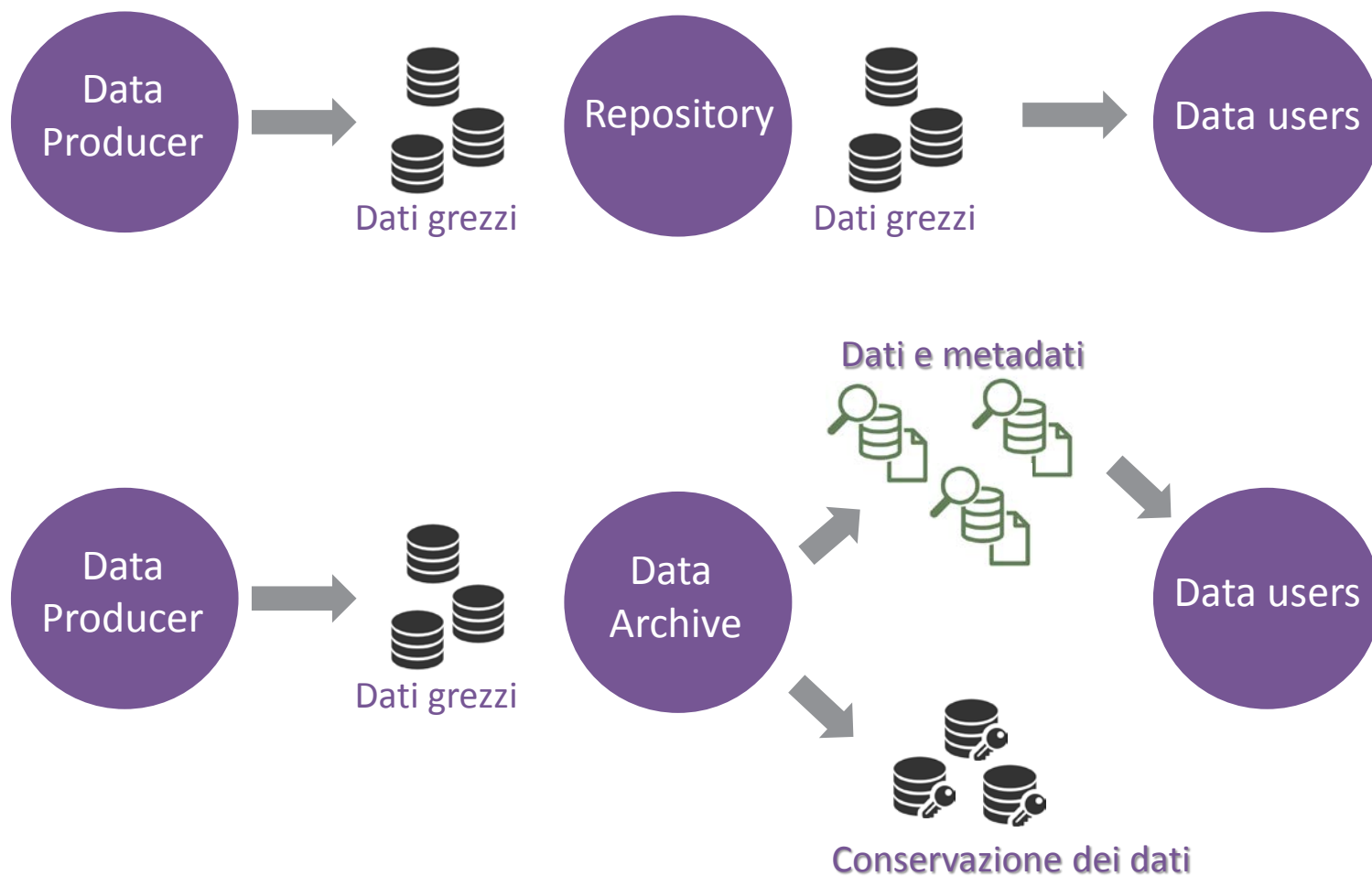
Un archivio di dati svolge le seguenti attività:

- **acquisizione**
- **conservazione**
- **documentazione**
- **trattamento**
- **catalogazione**
- **disseminazione**
- **supporto** nell'utilizzo



Attività finalizzate a **depositare** in modo sicuro, **descrivere**, **controllare**, **validare** e **distribuire** dati per la ricerca scientifica

Perché utilizzare un archivio di dati?



Il contesto italiano tra divergenze e contraddizioni

PUNTI DI FORZA

Aumenta uso di **analisi secondaria**
(riduzione dei costi)

Incremento nella **produzione di dati**
(small, open, e big)

Vincoli nazionali ed europei a favore della **condivisione dei dati per la ricerca**
(H2020, SIR 2014)

Diffusione di una cultura del dato
(forza del movimento Open Access)

PUNTI DI DEBOLEZZA

Frammentazione, eterogeneità e incomparabilità dei dati

Scarso supporto alla comunità di ricerca (mancanza di pratiche comuni, inefficacia normativa)

Poca attenzione ai dati rispetto alle pubblicazioni nelle politiche nazionali di Open Access

UniData – Bicocca Data Archive



- eredita il lungo lavoro svolto da **ADPSS-Sociodata** a partire dal 1999
- eccellenza nella creazione e diffusione delle **pratiche di condivisione dei dati**, e nella promozione dell'**analisi secondaria** e di un **utilizzo informato dei dati per la ricerca**

UniData: l'archiviazione dei dati

Acquisizione

protezione del produttore dei dati (proprietà, copyright, licenza)
incoraggiare gli utenti al riutilizzo dei dati (ricerca e accesso facili)

Conservazione

custodire i dati in un ambiente sicuro
conservare i dati in un formato standardizzato accessibile

Documentazione

creazione di metadati secondo uno standard internazionale (DDI)
descrizione delle variabili



modello OAIS
(Open Archival Information System)

UniData: servizi

Data curation

- tutela a lungo termine del valore della ricerca
- mitigazione del rischio di obsolescenza dei dati
- controllo sulla qualità dei dati
- creazione di una valida documentazione dei dati (metadati)

Data dissemination

- costruzione di dati FAIR (Findable, Accessibile, Interoperable, Reusable)
- gestione delle procedure di accesso ai dati in accordo con la licenza
- assegnazione di un Digital Object Identifier (DOI)

Consulenza e supporto

- ricognizione delle fonti di dati
- progettazione e costruzione di *shareable survey data*
- predisposizione di un Data Management Plan

UniData: le sfide per il futuro

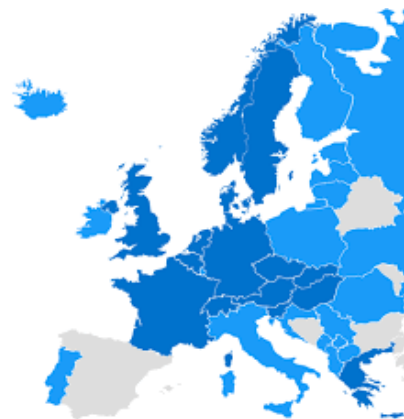
- rafforzare le basi per la ricerca **comparativa** e **interdisciplinare** nel campo delle scienze sociali
- incentivare una **cultura della condivisione** che favorisca l'**accessibilità** e il **riutilizzo** dei dati per la ricerca



L'infrastruttura **CESSDA** (Consortium of European Social Science Data Archives) promuovere l'uso di dati di alta qualità nell'ambito delle scienze sociali, e cura lo sviluppo di **standard**, **protocolli** e *best practice* in tema di archiviazione dei dati per la ricerca.



accesso a dati e servizi 'senza frontiere'



Gli archivi di dati nell'era dei Big Data

Questioni aperte

- ❑ Spesso di proprietà degli attori privati
- ❑ Aspetti legali relativi ai dati personali
- ❑ Dati non strutturati e non documentati



*Gli archivi di dati possono operare come **mediatori**, e facilitare il riutilizzo dei dati:*

- ✓ *curando la documentazione*
- ✓ *sviluppando workflows e servizi per la distribuzione di collezioni di Big Data*
- ✓ *adottando strategie di conservazione*

Alcuni esempi promettenti:


Administrative Data Research Network (in collaborazione con UK Data Archive)

Wearables project (in collaborazione con AUSSDA)



Università degli Studi di Milano-Bicocca

 unidata@unimib.it

 02 6448 7513

 www.unidata.unimib.it