

European Open Science Cloud: infrastrutture e servizi a supporto della ricerca all'Università degli Studi di Padova

Lorisa Andreoli, Elena Bianchi, Yuri Carrer, Vera Dean, Antonella Zane,
Michela Zorzi

Università degli Studi di Padova, Sistema Bibliotecario di Ateneo

Abstract. Questo documento presenta l'implementazione da parte del Sistema Bibliotecario dell'Università di Padova di un repository per archiviare, preservare e disseminare i dati di ricerca sviluppati all'interno dell'Ateneo. La relazione descrive le premesse del progetto e i passi che il Sistema Bibliotecario ha adottato per produrre l'infrastruttura tecnologica e corredarla di un adeguato servizio di supporto ai ricercatori.

Keywords. Data Repository, Open Data, Open Science, FAIR Data, Archivi aperti

Introduzione

L'Università di Padova è da tempo attiva nel promuovere l'accesso aperto, dotandosi di archivi aperti per la disseminazione e valorizzazione della propria produzione scientifica. A seguito delle recenti raccomandazioni della Commissione Europea per una gestione e un'apertura dei dati finalizzate al riuso, e tenuto conto dell'avvio dell'iniziativa European Open Science Cloud (EOSC), l'Ateneo patavino ha messo in atto delle azioni per dotarsi di un servizio a supporto dei ricercatori per la gestione e l'utilizzo dei dati della ricerca. L'esito di questo processo ha portato alla messa in produzione di un repository di Ateneo per i dati della ricerca conforme alle raccomandazioni della Commissione e all'organizzazione di servizi di supporto all'Open Science.

1. Progetto per lo sviluppo di un Servizio di Research Data Repository istituzionale

1.1 Analisi preliminare

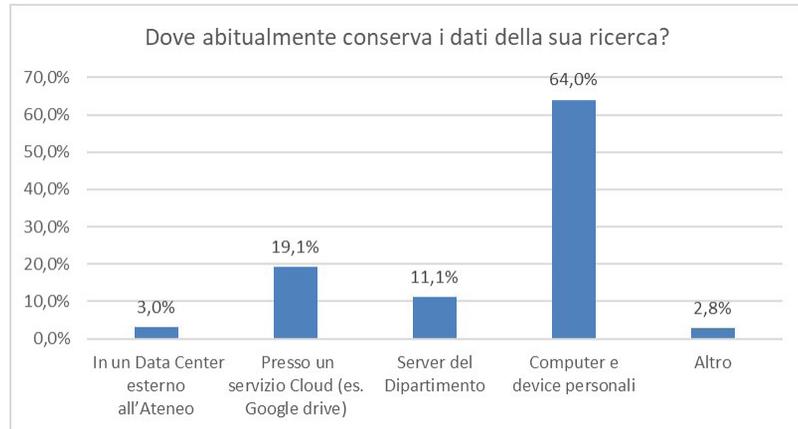
Per conoscere il livello di attuazione delle buone pratiche relative alla gestione dei dati della ricerca, il Sistema Bibliotecario e l'Area Ricerca e Rapporti con le Imprese dell'Ateneo hanno condotto un'indagine conoscitiva tramite interviste e questionari online rivolta alle segreterie di dipartimento e alle comunità di ricerca dell'Ateneo. Tali comunità comprendono ricercatori e docenti che necessitano di un servizio per il deposito e la conservazione dei dati di ricerca alternativo a Zenodo o ad altri repository disciplinari.

I risultati dell'indagine evidenziano che la conservazione dei dati della ricerca è gestita in modo disomogeneo da ricercatori e docenti. Nella maggior parte dei casi i dati sono ar-

chiviati in dispositivi e supporti personali; in minima parte vengono utilizzate infrastrutture messe a disposizione del dipartimento; in alcuni casi ci si avvale di servizi esterni, gratuiti o a pagamento.

Dall'indagine emerge inoltre l'esigenza da parte dei ricercatori di avere a disposizione adeguate infrastrutture tecnologiche e servizi di supporto per la gestione dei dati di ricerca, che vanno dall'intervento di primo livello (help desk) fino a un supporto specifico (Data Management Plan, questioni legali).

Fig. 1
Conservazione
dei dati di ricerca

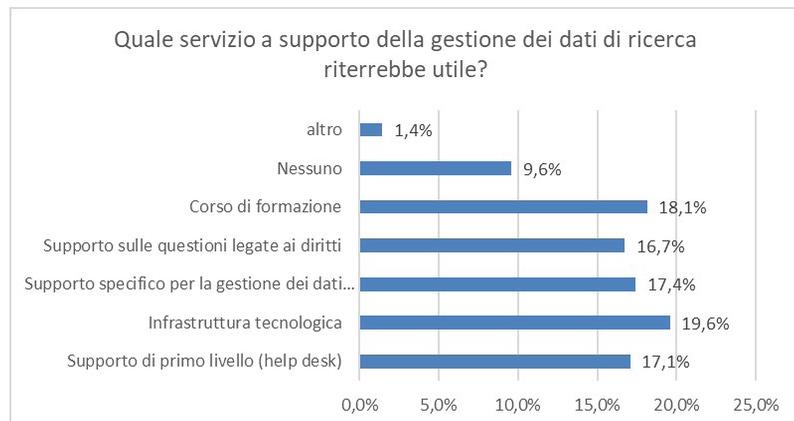


1.2 Scelta del software

Il Sistema Bibliotecario di Ateneo ha da anni una dimostrata esperienza nello sviluppo e nella gestione di repository ad Accesso aperto: risale al 2006 la realizzazione di due importanti archivi istituzionali nati per il deposito dei contributi di ricerca e delle tesi di dottorato (Padua@research) e per il deposito delle tesi di laurea ai fini della disseminazione (Padua@thesis).

Per identificare uno strumento che permetta ai ricercatori di archiviare, conservare e rendere fruibili i dati della ricerca, anche in risposta ai mandati degli enti finanziatori, nel 2017 il SBA ha intrapreso un'analisi dei principali software per repository.

Fig.2
Servizi richiesti
a supporto
della ricerca



Dallo studio della letteratura sul tema, si sono individuate le caratteristiche generali e le funzionalità tipiche di un repository istituzionale per i dati della ricerca ad accesso aperto, compresa la possibilità di integrare la nuova piattaforma con l'archivio istituzionale dei prodotti della ricerca (Padua Research Archive).

Tra i software free e open source per repository è stato scelto EPrints già utilizzato in Ateneo per i due archivi sopraccitati. EPrints, sviluppato dall'università di Southampton, è uno dei software per repository più diffuso al mondo.

La gestione del repository (implementazione e sviluppo del software) è curata dal Sistema Bibliotecario, mentre lo storage è ospitato presso il centro dati dell'ASIT (Area Servizi informatici e telematici di Ateneo).

1.3 Caratteristiche di Research Data Unipd

- Accesso aperto;
- Principi FAIR. Dati e metadati rispettano i principi FAIR (Findable, Accessible, Interoperable, Reusable) come richiesto dalla Commissione europea e dai principali enti finanziatori. La Fig. 3 riporta la misurazione della FAIRness del repository;
- Identificatore persistente. Il repository attribuisce in automatico un DOI (Digital Object Identifier) ai dataset depositati, grazie all'adesione del Sistema Bibliotecario di Ateneo al progetto DOI non commerciale della CRUI ;
- Licenze. Ai metadati è associata di default la Licenza Creative Commons 0. Chi deposita può scegliere tra un set di licenze da assegnare ai dataset;
- Soggetti. I dataset vengono classificati utilizzando i settori ERC (European Research Council);
- Integrazione con CRIS (Current Research Information System). Research Data Unipd è tra i primi casi in Italia di integrazione tra un archivio di dati e un archivio di prodotti della ricerca. Gli underlying data depositati in Research Data Unipd sono collegabili con le rispettive pubblicazioni presenti in Padua Research Archive (il CRIS dell'Università di Padova). Allo stesso modo è possibile collegare il dataset al documento pubblicato sul sito dell'editore;
- Deposito. Il deposito dei dataset è consentito agli utenti istituzionali attraverso autenticazione via Single Sign-On di Ateneo;
- Istruzioni per gli autori. La sezione informativa del repository comprende indicazioni generali sulle finalità, sulla policy istituzionale, sui formati dei file e sulle licenze applicabili. Fornisce inoltre istruzioni sulle operazioni preliminari al deposito (verifica dei requisiti previsti dai finanziatori, predisposizione di identificatori personali, URL degli articoli collegati ai dataset, etc.). Nel backend è inoltre disponibile un help contestuale per ogni campo editabile.

1.4 Integrazione con infrastrutture e servizi europei relativi ai dati di ricerca

Per garantire al repository e ai dataset caricati la massima visibilità a livello internazionale, l'archivio è stato indicizzato presso servizi internazionali, quali re3data.org e OpenDOAR. Essendo conforme agli standard richiesti da OpenAIRE , Research Data Unipd è attivo come content provider di questa iniziativa europea.

FINDABLE		assessment	
F1	(meta)data are assigned a globally unique and eternally persistent identifier	✓	F Findable
F2	data are described with rich metadata	✓	
F3	(meta)data are registered or indexed in a searchable resource (able to google data-objects)	✓	
F4	metadata specify the data identifier	✓	
ACCESSIBLE			
A1	(meta)data are retrievable by their identifier using a standardized communications protocol	✓	A Accessible
A1.1	the protocol is open, free, and universally implementable	✓	
A1.2	the protocol allows for an authentication and authorization procedure, where necessary	✓	
A2	metadata are accessible, even when the data are no longer available		
INTEROPERABLE			
I1	(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation	✓	I Interoperable
I2	(meta)data use vocabularies that follow FAIR principles	unclear	
I3	(meta)data include qualified references to other (meta)data	unclear	
RE-USABLE			
R1	(meta)data have a plurality of accurate and relevant attributes	✓	R Reusable
R1.1	(meta)data are released with a clear and accessible data usage license	✓	
R1.2	(meta)data are associated with their provenance	✓	

Fig. 3
FAIRness di
Research Data Unipd

1.5 La policy per la gestione dei dati della ricerca dell'Università di Padova

Contestualmente al rilascio del repository, il 1 dicembre 2018 il Senato Accademico ha approvato la Policy istituzionale per la gestione dei dati della ricerca, che identifica esplicitamente Research Data Unipd come l'archivio digitale dell'Università di Padova per i dati della ricerca. Il testo della policy si basa sul modello elaborato dal gruppo di lavoro interuniversitario IOSSG (Italian Open Science Support Group).

2. La diffusione e la promozione dell'Open Science all'Università di Padova

Attualmente, presso l'Università di Padova, i servizi di supporto alla ricerca per le tematiche relative all'Open Access, agli Open Data e alla normativa sul diritto d'autore vengono garantiti dal Sistema Bibliotecario.

Attraverso il suo sito web fornisce informazioni e approfondimenti sull'Open Access e l'Open Science (Sezione "Per chi pubblica") e il "Servizio Aiuto SBA". La finalità di questo servizio è di offrire al ricercatore un unico punto di accesso a servizi di supporto coordinati tra le varie unità organizzative collegate alla ricerca, come descritto in "A view on the implementation of the European Open Science Cloud".

Inoltre, il Sistema Bibliotecario offre a tutti gli utenti istituzionali (dottorandi, ricercatori, docenti e staff) occasioni di formazione e informazione, in presenza e on-line, sui temi dell'Open Science.

3. Conclusioni

Le biblioteche accademiche sono di fronte a un periodo di grandi sfide innovative, quali la gestione e condivisione dei dati prodotti dalla ricerca nell'ambito dell'iniziativa European Open Science Cloud promossa dalla Commissione Europea. In questo contesto, le com-

petenze specialistiche sui dati e metadati, proprie della professionalità del bibliotecario, diventano strategiche per le attività di supporto al “Research data planning and management” essenziali alla ricerca accademica per accedere ai finanziamenti europei.

Lo sviluppo di una infrastruttura per la gestione e archiviazione FAIR dei dati della ricerca e l'offerta di una gamma di servizi a supporto dei ricercatori, il cui accesso è agevolato da un unico punto di contatto, sono i primi necessari passi verso la European Open Science Cloud.

Per l'Università di Padova lo sviluppo della nuova infrastruttura è stata accompagnata e sostenuta da un'azione decisa da parte della governance di Ateneo concretizzata nella adozione di una policy istituzionale specificatamente dedicata ai dati. La formulazione e l'adozione di un documento di policy è infatti il punto di partenza per una piena attuazione della European Open Science Cloud da parte di un'istituzione di ricerca.

Riferimenti bibliografici

Bianchi E., Budroni, P., Celentano, A., Occioni, M., Toniolo S. (2017), A view on the implementation of the European Open Science Cloud. Conferenza GARR_2017 Selected Papers. The data way to Science. Venezia, 15-17 novembre 2017, Associazione Consortium GARR, Roma

Solodovnik, Iryna (2015), Repository Istituzionali Open Access e strategie Linked Open Data, Firenze University Press, Firenze http://www.fupress.com/archivio/pdf/3091_9271.pdf

<http://bibliotecadigitale.cab.unipd.it/aiuto>

http://bibliotecadigitale.cab.unipd.it/bd/per_chi_pubblica/documenti-e-materiali/dichiarazionedimessina2004.pdf

http://bibliotecadigitale.cab.unipd.it/bd/per_chi_pubblica/documenti-e-materiali/Policy_accesso_aperto1.pdf

<http://bibliotecadigitale.cab.unipd.it/bd/per-chi-pubblica>

<http://paduaresearch.cab.unipd.it/>

<http://tesi.cab.unipd.it/>

<http://v2.sherpa.ac.uk/opensoar/>

<http://www.unipd.it/download/file/fid/55401>

<https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>

<https://elearning.unipd.it/sba/course/view.php?id=21>

<https://erc.europa.eu/>

<https://sites.google.com/view/iossg/home>

<https://www.coar-repositories.org/files/NGR-Final-Formatted-Report-cc.pdf>

<https://www.crui.it/biblioteche-didattica/progetto-doi.html>

<https://www.go-fair.org/fair-principles/>

<https://www.openaire.eu/>

<https://www.re3data.org/>

<https://www.research.unipd.it/>

<https://www.unipd.it/sites/unipd.it/files/2018/policy%20dati%20ricerca.pdf>

<https://zenodo.org/>

Autori

Lorisa Andreoli - lorisa.andreoli@unipd.it

Lorisa Andreoli lavora presso il Sistema bibliotecario dell'Università degli Studi di Padova. È responsabile del Settore Servizi all'utenza e Comunicazione dell'Ufficio Biblioteca digitale. Si occupa dei servizi della biblioteca digitale, in particolare del coordinamento dei progetti di digitalizzazione e della piattaforma "Phaidra - Collezioni digitali", della formazione e supporto alla predisposizione di mostre virtuali, delle attività di supporto relative ai dati della ricerca.

Elena Bianchi - elena.bianchi@unipd.it

Elena Bianchi lavora dal 1999 per il Sistema bibliotecario dell'Università degli studi di Padova. Dopo alcuni anni come responsabile della Biblioteca Interdipartimentale di Psicologia F. Metelli, dal 2012 al 2018 gestisce il settore Comunicazione del Sistema Bibliotecario. Dal 2018 è a capo dell'Ufficio Biblioteca digitale di Ateneo. Si occupa in particolare di servizi ed infrastrutture tecnologiche a supporto dei ricercatori per l'Accesso aperto a dati e pubblicazioni. E' membro dell'Italian Open Science Support Group (IOSSG).

Yuri Carrer - yuri.carrer@unipd.it

Yuri Carrer si laurea in Ingegneria Informatica presso l'Università di Padova; nella sua tesi di laurea si occupa di oggetti digitali, attività che continua dal 2003 presso il Sistema Bibliotecario dell'Università di Padova. Sviluppa gli archivi istituzionali Padua@Thesis/Padua@Research e gestisce come responsabile tecnico le collezioni digitali. Recentemente è occupato di dati della ricerca con la piattaforma Research Data Unipd. Ha creato diversi siti con CMS Plone a supporto delle attività istituzionali.

Vera Dean - vera.dean@unipd.it

Vera Dean ha un Master in Biblioteconomia e Scienze dell'Informazione. Dopo aver lavorato presso Infologic s.r.l. e Atlantis (oggi ExLibris Italy), lavora dal 2007 per il Sistema Bibliotecario dell'Università degli studi di Padova, dove si occupa del sistema di automazione gestionale del Sistema Bibliotecario. Dal 2018 di occupa anche del Servizio di Supporto alla Ricerca sui temi dell'Open Science e dell'Open Access all'interno dell'Ufficio Biblioteca digitale di Ateneo.

Antonella Zane - antonella.zane@unipd.it

Antonella Zane ha un PhD in Scienze della Terra e 10 anni di ricerca presso l'Università di Padova. Dal 1998 lavora per il Sistema bibliotecario di Ateneo. Dal 2002 collabora con il Digital Library Group dell'ISTI-CNR, Pisa, ed è stata capoprogetto del team padovano del progetto EU "Linked Heritage" (CIP, 2007-2013). Dal 2011 al 2018 è responsabile del Settore Biblioteca digitale con un focus sui repository per i dati della ricerca. Oggi è il responsabile tecnico della Biblioteca biomedica "A. Vallisneri" e membro dell'Italian Open Science Support Group (IOSSG).

Michela Zorzi - michela.zorzi@unipd.it

Michela Zorzi lavora dal 1999 per il Sistema bibliotecario dell'Università degli studi di Padova. Dal 1999 al 2008 si occupa del Servizio di Prestito Interbibliotecario di Ateneo e successivamente lavora presso la Biblioteca di Scienze Statistiche e le altre biblioteche del Polo di Scienze Sociali. Dal 2018 di occupa del Servizio di Supporto alla Ricerca sui temi dell'Open Science e dell'Open Access all'interno dell'Ufficio Biblioteca digitale di Ateneo.