

I linked open data per la valorizzazione dei beni culturali

Carlo Bianchini

Ciclo di incontri: *Di carte, di libri e dintorni*
Università di Bologna, Dipartimento di Beni Culturali
Ravenna, 8 aprile 2022



UNIVERSITÀ DI PAVIA

Sommario

- I dati di Musei, Archivi e Biblioteche
- I Linked (open) data
- I linked open data per la valorizzazione dei beni culturali



UNIVERSITÀ DI PAVIA

10/04/2022 •

La crisi delle istituzioni MAB

- La diffusione del web e l'enorme disponibilità di contenuti in rete sembra rendere marginale e forse obsoleto il ruolo delle istituzioni culturali (Musei, Archivi, Biblioteche)
- Le nuove abitudini degli cittadini hanno creato distanza tra i desideri percepiti e i servizi offerti, tra i *modelli di ricerca* di informazioni seguiti dai lettori e *gli strumenti messi a disposizione* dalle istituzioni culturali
- Si è creato *squilibrio* tra raccolte, strumenti di mediazione e utenti



UNIVERSITÀ DI PAVIA

10/04/2022 ● 3

I dati delle istituzioni MAB

- Milioni di (meta)dati sui beni culturali conservati negli opac delle biblioteche, archivi e musei di tutto il mondo non sono raggiungibili dai motori di ricerca;
- I (meta)dati MAB rimangono invisibili alla maggior parte degli utenti (84%) che cercano informazioni sulla rete a partire dai motori di ricerca



UNIVERSITÀ DI PAVIA

10/04/2022 ● 4

I (meta)dati MAB

Per esempio, se oggi cerchiamo un'opera letteraria tramite un motore otteniamo risposte dai **cataloghi di vendita** di libri in commercio (o di **antiquariato**).

I **cataloghi delle biblioteche**, invece, la cui redazione ha richiesto l'elaborazione di specifiche normative, l'uso di competenze professionali e l'investimento di consistenti finanziamenti pubblici, **non forniscono risposte ai motori di ricerca**.



10/04/2022 ● 5

opac.sbhut/bw5ne4/Opac.aspx?WEB=SRUD&SRUC=SBAS

Milano : Sperling & Kupfer, 1981
 ✓ Disponibili: 1

7 M
 Linea di sangue / Sidney Sheldon. - Milano : Club degli editori, stampa 1980. - 404 p. ; 21 cm
 Sheldon, Sidney
 Milano : Club degli editori, stampa 1980
 ✓ Disponibili: 6

8 M
 Linea di sangue / Sidney Sheldon ; [unica traduzione italiana autorizzata di Tilde Riva dall'originale inglese]. - Milano : Sperling & Kupfer, 1979. - 404 p. ; 21 cm.
 Sheldon, Sidney
 Milano : Sperling & Kupfer, c1979
 ✓ Disponibili: 1

9 M
 Linea di sangue / Sidney Sheldon. - Milano : CDE, c1979. - 404 p. ; 21 cm
 Sheldon, Sidney
 Milano : CDE, c1979
 ✓ Disponibili: 1

10 LINEADI
 Linea di sangue / Angela Marsons. - Roma : Newton Compton, 2019. - 381 p. ; 24 cm. - (Nuova narrativa Newton ; 0004)
 Marsons, Angela
 Roma : Newton Compton, 2019
 ✗ Disponibili: 0

Sistemi e biblioteche che partecipano al Polo SBNFVG:
 sbhu SB
 • Sistema Bibliotecario del Friuli • Sistema Bibliotecario Medio Friuli
 • Elenco biblioteche • Elenco biblioteche

Metamotore Regione Friuli Venezia Giulia:
 REGIONE AUTONOMA
 FRIULI VENEZIA GIULIA
 BIBLIOTECHE FVG
 motore di ricerca nei
 cataloghi delle biblioteche
 di Regione
 Friuli Venezia Giulia

© 2019 Polo SBN del Friuli Venezia Giulia Privacy & cookies - Powered by Bibliowin 5.0

eseguire la ricerca



10/04/2022 ● 6

https://www.europeana.eu/



Public Domain

+

♥

Share

Download

Pinturicchio. Retorica en Geometrie

KU Leuven. Glasdia's kunstgeschiedenis. Universit  de Louvain, tussen 1839 en 1939. Fotograaf onbekend.
 Toegevoegde informatie op dia. Strooming/Stijl: Renaissance. Creatie/Bouw: 1492-1494. Techniek/Materiaal: Fresco.
 Huidige locatie: Stato della Citt  del Vaticano, Palazzo Apostolico, Appartamento Borgia, Sala delle Arti Liberali.
 EuroPhot. Kunstgeschiedenis. 15de eeuw. Schilderkunst. Allegorie. EuroPhot. Art history, 15th century. Paintings.
 Allegory.
[Show less](#)

COLLECTIONS YOU MIGHT LIKE



UNIVERSIT  DI PAVIA

10/04/2022 ● 7

Descrivere le risorse (sul web)

- Malgrado molti detrattori, il *'catalogo'* – inteso come *complesso di dati e metadati* per la gestione e l'accesso alle risorse – rimane il **core business** delle istituzioni culturali
- Senza i dati non funzionano la ricerca, la consultazione, la circolazione, il reference, le acquisizioni, la conservazione, l'aggiornamento ...
- Ma soprattutto: i lettori, non trovando i nostri dati con i motori di ricerca, ignorano semplicemente le istituzioni culturali e *tutti i loro servizi*.



Contenitore vs contenuto

- Per comprensibili ragioni storiche, culturali, economiche, giuridiche ecc. MAB hanno un approccio ai propri dati per 'contenitore', per 'istituzione'
- Le domande di ricerca dei nostri utenti hanno invece un approccio per contenuto:
 - Quali sono le istituzioni che possiedono opere (documenti/testi) di Tiziano Vecellio e, quale, tra queste, ne possiede in numero maggiore?
 - Posso avere un elenco di tutte le opere (documenti/testi) che hanno come soggetto una scena della Bibbia (Antico Testamento, Nuovo Testamento, mitologia classica, una struttura architettonica, un oggetto specifico)?
 - Quali altre fonti ci sono su questo «Mario Rossi»?
 - Posso avere queste informazioni direttamente nella mia lingua?



UNIVERSITÀ DI PAVIA

10/04/2022 ● 11

Interoperabilità

- È necessario rendere i dati interoperabili, non solo tra istituzioni culturali (MAB), ma tra i silos culturali e qualsiasi altro produttore di dati esterno – pubblico e privato – attraverso il web aperto.
- Per riuscirci, è necessario registrare e pubblicare i dati (bibliografici e non) con modalità che consentano agli *elaboratori*
 - di *individuarli*,
 - di riconoscere che tra di essi sussistono *relazioni* e
 - di stabilire *che tipo* di relazioni sono (*tipizzazione*)



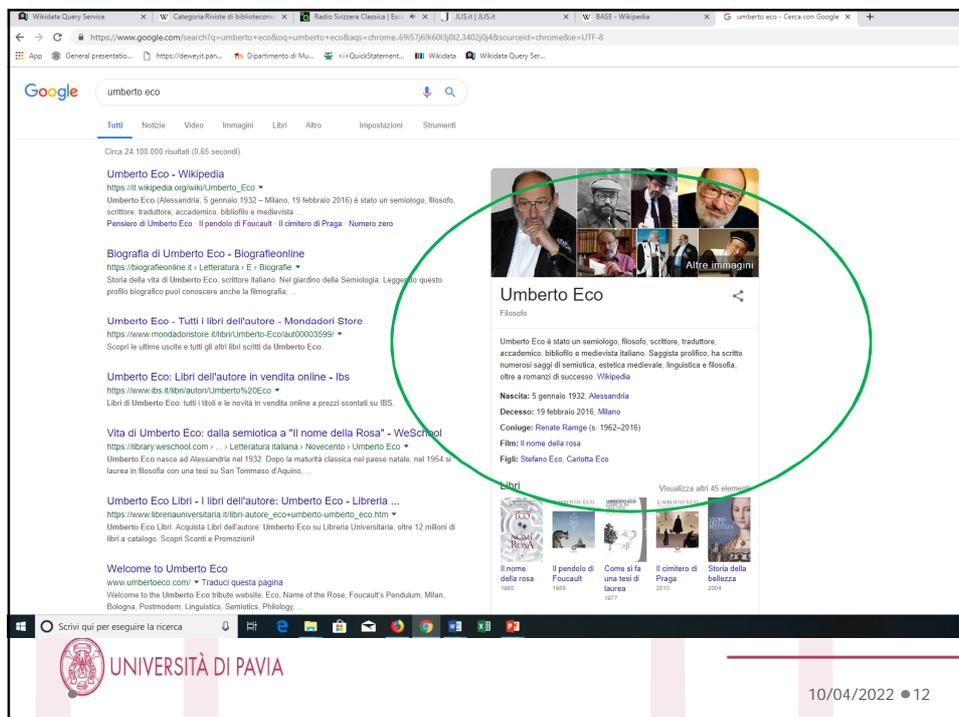
UNIVERSITÀ DI PAVIA

10/04/2022 ● 10

Oltre il web dei documenti

Non solo i dati di MAB e della PA hanno problemi di visibilità sul web. Anche i siti web che offrono contenuti informativi hanno bisogno di un modello diverso da quello attuale (web dei documenti) per almeno due esigenze:

1. Collegare le informazioni del web in modo 'più stretto'
2. Creare servizi di informazione che si aggiornino in modo automatico



The screenshot shows a Google search for "umberto eco". The search results include:

- Umberto Eco - Wikipedia**: https://it.wikipedia.org/wiki/Umberto_Eco. Description: Umberto Eco (Alessandria, 5 gennaio 1932 – Milano, 19 febbraio 2016) è stato un semiologo, filosofo, scrittore, traduttore, accademico, bibliofilo e medievista.
- Biografia di Umberto Eco - Biografieonline**: https://biografieonline.it/Letteratura/E/Biografie. Description: Storia della vita di Umberto Eco, scrittore italiano. Nel giardino della Semiologia. Leggendolo questo profilo biografico puoi conoscere anche la filmografia.
- Umberto Eco - Tutti i libri dell'autore - Mondadori Store**: https://www.mondadoristore.it/letteratura/italiana/Novocento/Umberto_Eco. Description: Scopri le ultime uscite e tutti gli altri libri scritti da Umberto Eco.
- Umberto Eco: Libri dell'autore in vendita online - Ibs**: https://www.ibs.it/libri/autori/Umberto%20Eco. Description: Libri di Umberto Eco: tutti i titoli e le novità in vendita online a prezzi scontati su IBS.
- Vita di Umberto Eco: dalla semiologia a "Il nome della Rosa" - WeSchool**: https://library.weschool.com/.../Letteratura/italiana/Novocento/Umberto_Eco. Description: Umberto Eco nasce ad Alessandria nel 1932. Dopo la maturità classica nel paese natale, nel 1954 si laurea in filosofia con una tesi su San Tommaso d'Aquino.
- Umberto Eco Libri - I libri dell'autore: Umberto Eco - Libreria ...**: https://www.libreriauniversitaria.it/libri-autore_eco/umberto_umberto_eco.htm. Description: Umberto Eco Libri. Acquista i libri dell'autore Umberto Eco su Libreria Universitaria, oltre 12 milioni di libri a catalogo. Scopri Scenari e Promozioni!
- Welcome to Umberto Eco**: www.umbertoeco.com/ Traduci questa pagina. Description: Welcome to the Umberto Eco tribute website. Eco, Name of the Rose, Foucault's Pendulum, Milan, Bologna, Postmodern, Linguistics, Semiotics, Philology.

A knowledge panel for Umberto Eco is displayed on the right, featuring a photo of him and the following information:

- Umberto Eco**: Filosofo
- Umberto Eco è stato un semiologo, filosofo, scrittore, traduttore, accademico, bibliofilo e medievista italiano. Saggista prolifico, ha scritto numerosi saggi di semiologia, estetica medievale, linguistica e filosofia, oltre a romanzi di successo. Wikipedia
- Nascita**: 5 gennaio 1932, Alessandria
- Decesso**: 19 febbraio 2016, Milano
- Coniuge**: Renata Rampe (n. 1962–2016)
- Film**: Il nome della rosa
- Figli**: Stefano Eco, Carlotta Eco

Below the knowledge panel, there is a section for books, including "Il nome della rosa" (1980), "Il pendolo di Foucault" (1988), "Come si fa una tesi di laurea" (1977), "Il cimitero di Praga" (2007), and "Storia della bellezza" (2004).

AGENDA Summer school 'Link' x Docenti - Linked Data per i Beni x Posta - carlabianchini@univd.it x BBC - Help - Web x umberto eco - Buscar con Google x

https://www.google.com/search?q=umberto+eco&oeq=umberto&saq=chrome.0.0.6915704.313767&sourceid=chrome&ie=UTF-8

Google umberto eco

Todo Imágenes Noticias Libros Videos Más Configuración Herramientas

Aproximadamente 24 000 000 resultados (0,44 segundos)

Umberto Eco - Wikipedia, la enciclopedia libre
https://es.wikipedia.org/wiki/Umberto_Eco
 Umberto Eco (Alessandria, Piamonte, Italia, 5 de enero de 1932 - Milán, Lombardia, Italia, 19 de febrero de 2016) fue un escritor, filósofo y profesor de ...
 Obras notables: *El nombre de la rosa* ... **Nacimiento:** 5 de enero de 1932, **Alessandria**, ...
 Causa de la muerte: Cáncer de páncreas ... **Nacionalidad:** Italiana

Umberto Eco - Wikipedia
https://it.wikipedia.org/wiki/Umberto_Eco • Traducir esta página
 Umberto Eco (Alessandria, 5 gennaio 1932 - Milano, 19 febbraio 2016) è stato un semiologo, filosofo, scrittore, traduttore, accademico, bibliofilo e medievista ...
 Pensiero di Umberto Eco - Il pendolo di Foucault - Il cimitero di Praga - Numero zero

Biografía di Umberto Eco - Biografieonline
<https://biografieonline.it/Letteratura/E/Biografie> • Traducir esta página
 Storia della vita di Umberto Eco, scrittore italiano. Nel giardino della Semiologia. Leggendo questo profilo biografico puoi conoscere anche la filmografia, ...

Welcome to Umberto Eco
www.umberteco.com • Traducir esta página
 Welcome to the Umberto Eco tribute website, Eco, Name of the Rose, Foucault's Pendulum, Milan, Bologna, Postmodern, Linguistics, Semiotics, Philology, ...

Umberto Eco: Libri dell'autore in vendita online - IBS
<https://www.ibs.it/libri/autori/Umberto%20Eco> • Traducir esta página
 Libri di Umberto Eco: tutti i libri e le novità in vendita online a prezzi scontati su IBS.

Umberto Eco - Tutti i libri dell'autore - Mondadori Store
<https://www.mondadoristore.it/libri/Umberto-Eco/au00003059/> • Traducir esta página
 Scopri le ultime uscite e tutti gli altri libri scritti da Umberto Eco.

Vita di Umberto Eco: dalla [semiotica](#) a ["Il nome della Rosa"](#) - WeSchool

Umberto Eco
 Escritor

Umberto Eco fue un escritor, filósofo y profesor de universidad italiano. Fue autor de numerosos ensayos sobre semiótica, estética, lingüística y filosofía, así como de varias novelas. *El nombre de la rosa* es la más conocida. *Wikipedia*

Fecha de nacimiento: 5 de enero de 1932, **Alessandria**
Fallecimiento: 19 de febrero de 2016, **Milán**
Género: Ensayo y novela
Películas: *El nombre de la rosa*
Premios: Premio Strega, Premio Médica Extranjero, Legión de Honor, Orden de las Artes y las Letras

Libros Ver otros 45

El nombre de la rosa 1980	El pendulo de Foucault 1988	Como se hace una tesis 1977	El cementerio de Praga 2010	Historia de la belleza 2004

Otras personas también buscan Ver otros 15

UNIVERSITÀ DI PAVIA

10/04/2022 ● 13

Creare il web dei dati

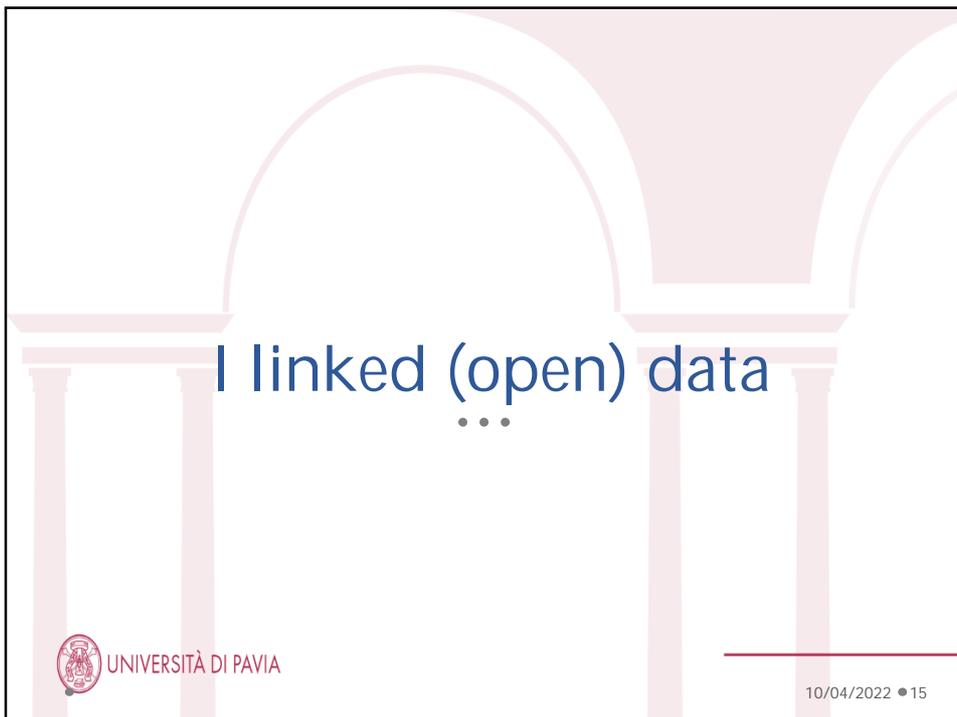
Knowledge panel di Google è un esempio di passaggio dal web dei documenti al web semantico

Il contenuto informativo di una pagina non è statico, ma si aggiorna automaticamente grazie a *dati* liberamente disponibili sul web

Perché ciò sia possibile, è necessaria una infrastruttura adatta a costruire una *rete di dati* nella quale

- I dati siano accessibili sul web, tramite tecnologie web standard
- I dati siano interconnessi sul web, ovvero possano essere completamente integrati nel web.

Qui entrano in gioco i *linked (open) data*

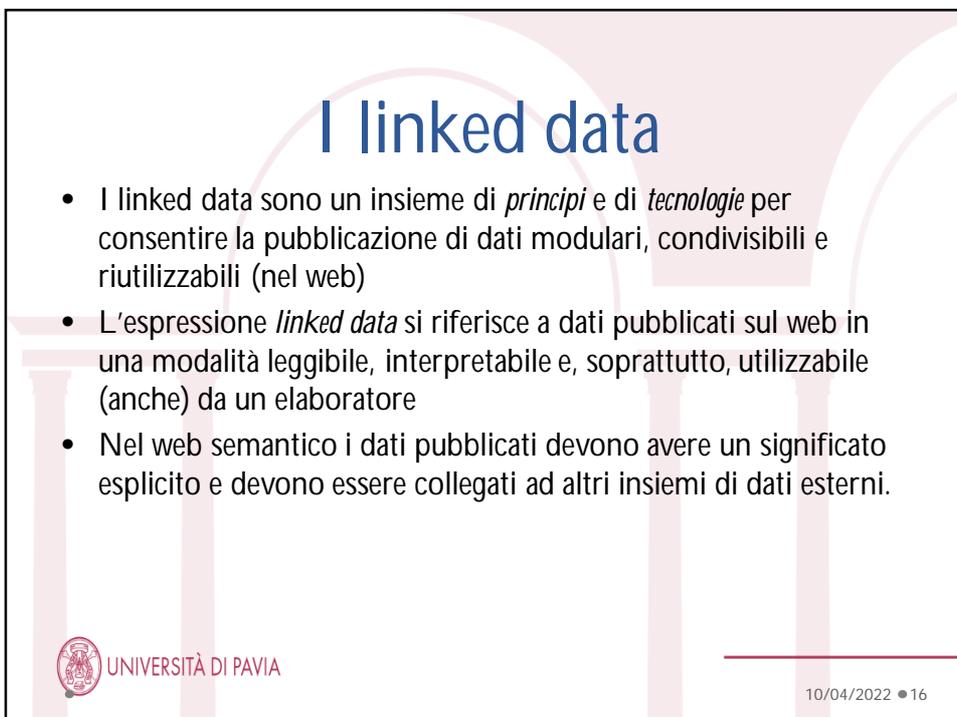


I linked (open) data

...

 UNIVERSITÀ DI PAVIA

10/04/2022 ● 15



I linked data

- I linked data sono un insieme di *principi* e di *tecnologie* per consentire la pubblicazione di dati modulari, condivisibili e riutilizzabili (nel web)
- L'espressione *linked data* si riferisce a dati pubblicati sul web in una modalità leggibile, interpretabile e, soprattutto, utilizzabile (anche) da un elaboratore
- Nel web semantico i dati pubblicati devono avere un significato esplicito e devono essere collegati ad altri insiemi di dati esterni.

 UNIVERSITÀ DI PAVIA

10/04/2022 ● 16

I linked *open* data

- I linked *open* data sono dati realizzati con la tecnologia dei linked data e, oltre ad averne tutte le caratteristiche tecniche, sono aperti, accessibili e riutilizzabili *liberamente*.

Dai documenti ai dati

- Tim Berners Lee ha pensato che si potesse *applicare* ai singoli dati la medesima *struttura* con cui è stato creato il *World Wide Web*, allo scopo di condividere i *dati* su scala globale (es. Knowledge panel)
- Quale struttura? Il World Wide Web oggi disponibile, inteso come spazio unico di pubblicazione di documenti, è cresciuto e si è sviluppato grazie a *tre* strumenti fondamentali

Dai documenti ai dati

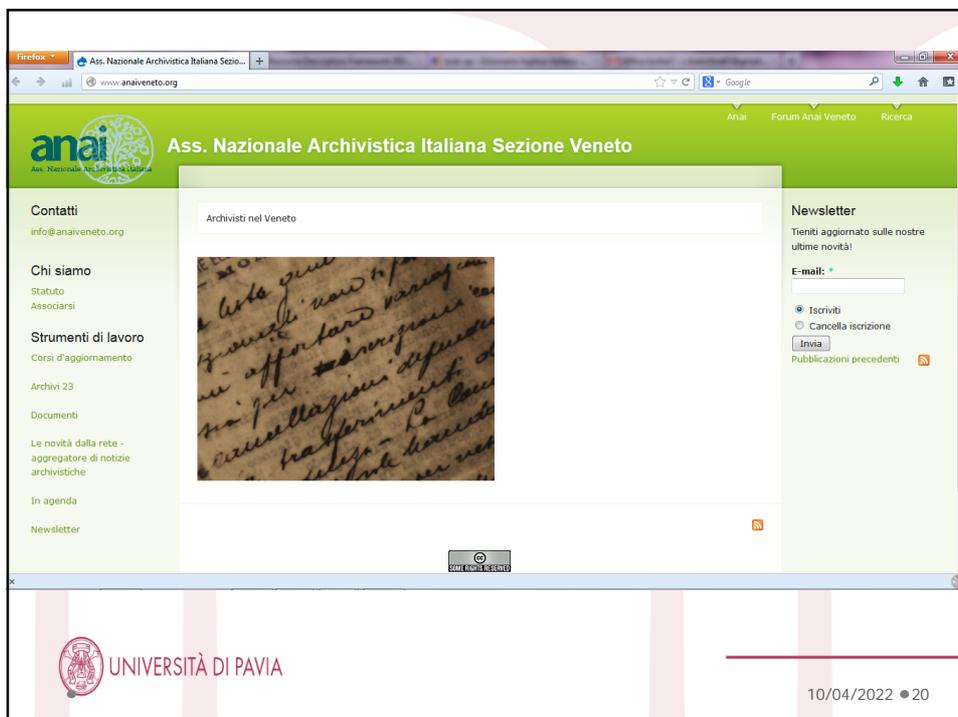
Uno strumento di *identificazione* globale
URI – Uniform Resource Identifier

Uno strumento di *accesso* globale
HTTP – HyperText Transfer Protocol

Uno strumento di *formattazione dei contenuti*
HTML – HyperText Markup Language

 UNIVERSITÀ DI PAVIA

10/04/2022 ● 19



The screenshot shows the website of the Ass. Nazionale Archivistica Italiana Sezione Veneto. The browser address bar displays "www.anaiveneto.org". The page features a green header with the "anai" logo and navigation links for "Anai", "Forum Anai Veneto", and "Ricerca". The main content area is titled "Archivisti nel Veneto" and includes a search bar and a large image of a handwritten document. The left sidebar contains sections for "Contatti" (info@anaiveneto.org), "Chi siamo" (Statuto, Associarsi), "Strumenti di lavoro" (Corsi d'aggiornamento, Archivi 23, Documenti, Le novità dalla rete - aggregatore di notizie archivistiche, In agenda, Newsletter), and "Newsletter" (Iscriviti, Cancella iscrizione, Invia, Pubblicazioni precedenti).

 UNIVERSITÀ DI PAVIA

10/04/2022 ● 20

Dai documenti ai dati

Lo sviluppo e l'uso di standard (HTTP e HTML) consente di superare le differenze tra le architetture tecnologiche con le quali vengono prodotti i documenti

I link (HyperLink) che si intrecciano tra i diversi documenti consentono agli utenti di navigare nella rete.

I link sono stati e sono determinanti per la creazione e il mantenimento di un *unico spazio di informazioni globale* (il *docuverso*)



UNIVERSITÀ DI PAVIA

10/04/2022 ● 21

I principi dei linked data

...



UNIVERSITÀ DI PAVIA

10/04/2022 ● 22

Principi per i linked data

Tim Berners-Lee (2006) ha definito quali devono essere i requisiti dei linked data; questi sono noti come *Principi* dei linked data

1. Usare gli URI come *nome* per le cose
2. Usare degli URI HTTP, in modo che i nomi delle cose si possano *cercare*
3. Fornire informazioni utili in forma standard (RDF) in risposta alle richieste HTTP
4. Inserire collegamenti ad altri URI, in modo che si possano *scoprire cose collegate* tra loro



UNIVERSITÀ DI PAVIA

10/04/2022 ● 23

RDF per il contenuto

- Per pubblicare linked data è necessario esprimere i dati secondo il modello **Resource Description Framework (RDF)**
- RDF è un *modello di dati* estremamente semplice e fortemente orientato al web



UNIVERSITÀ DI PAVIA

10/04/2022 ● 24

Il modello di dati RDF

- L'essenza di RDF è quella di esprimere entità, attributi e relazioni tra le cose sotto forma di *triple*, cioè affermazioni semplici

Esempi

SculturaX – èIn – Marmo

MaglioneY – èDiColore – Rosso

PersonaZ – LavoraPresso – DittaW



10/04/2022 ● 25

A che servono le triple?

- Mediante una concatenazione di triple è possibile esprimere dati molto complessi, in modo che siano leggibili, modulari e riutilizzabili anche da una macchina.

Esempio

Robinson Crusoe / Daniel Defoe



10/04/2022 ● 26

A che servono le triple?

- In base alle triple esistenti, un elaboratore può anche 'dedurre' dati nuovi. Per esempio, è in grado – a certe condizioni – di inferire nuove triple corrette

Esempi:

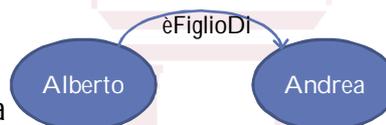
Alberto – èFiglioDi – Andrea

Agata – èFigliaDi – Andrea

→

Alberto – èFratelloDi – Agata

Agata – èSorellaDi – Alberto (i)



UNIVERSITÀ DI PAVIA

10/04/2022 ● 27

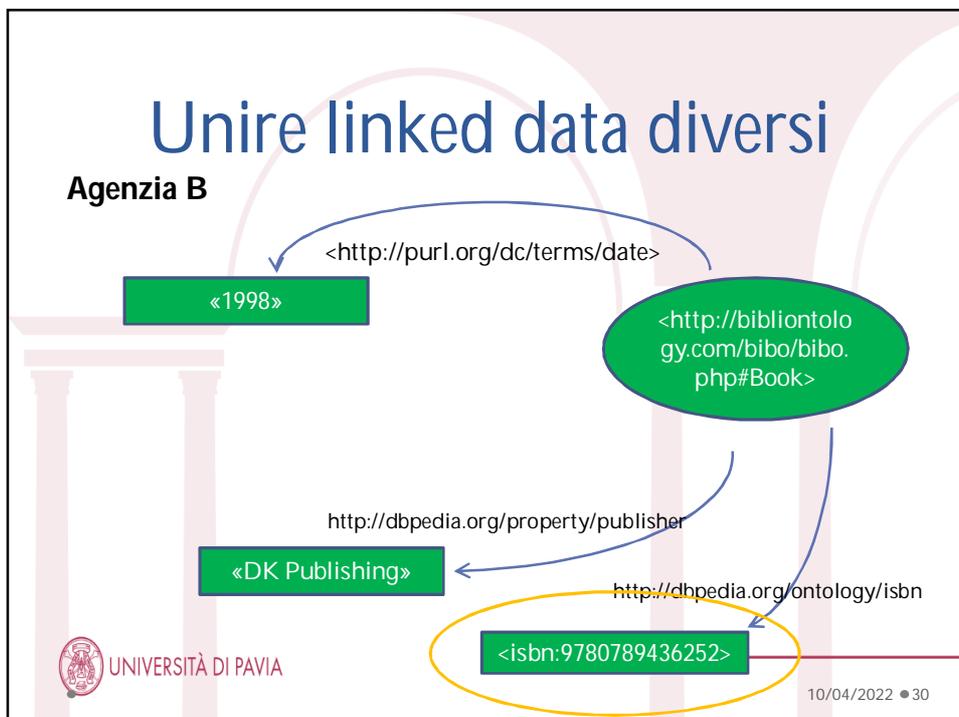
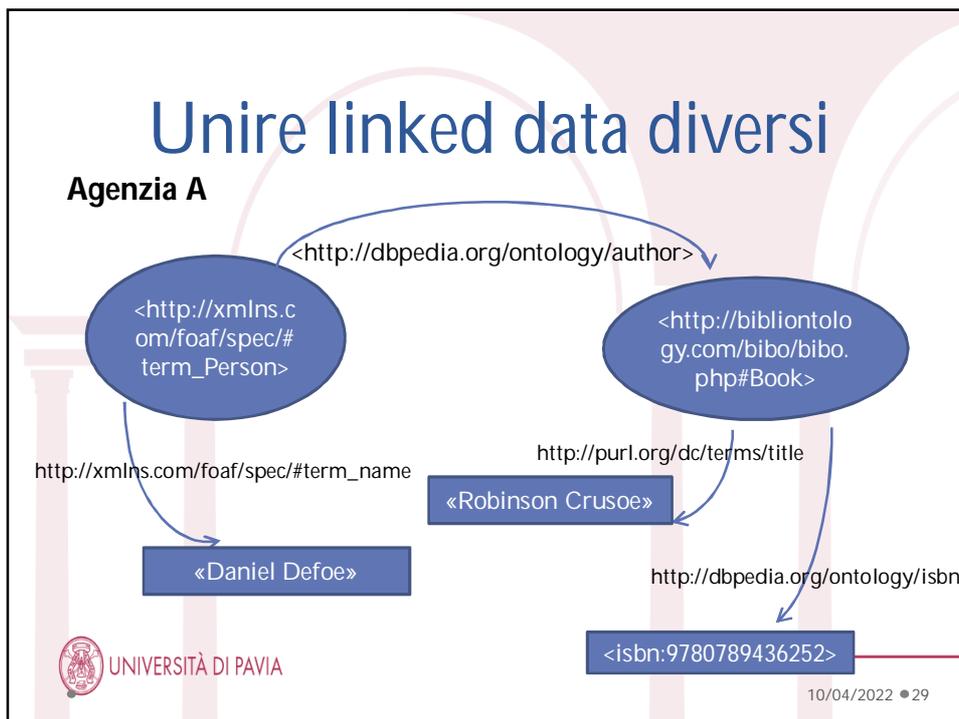
Un esempio

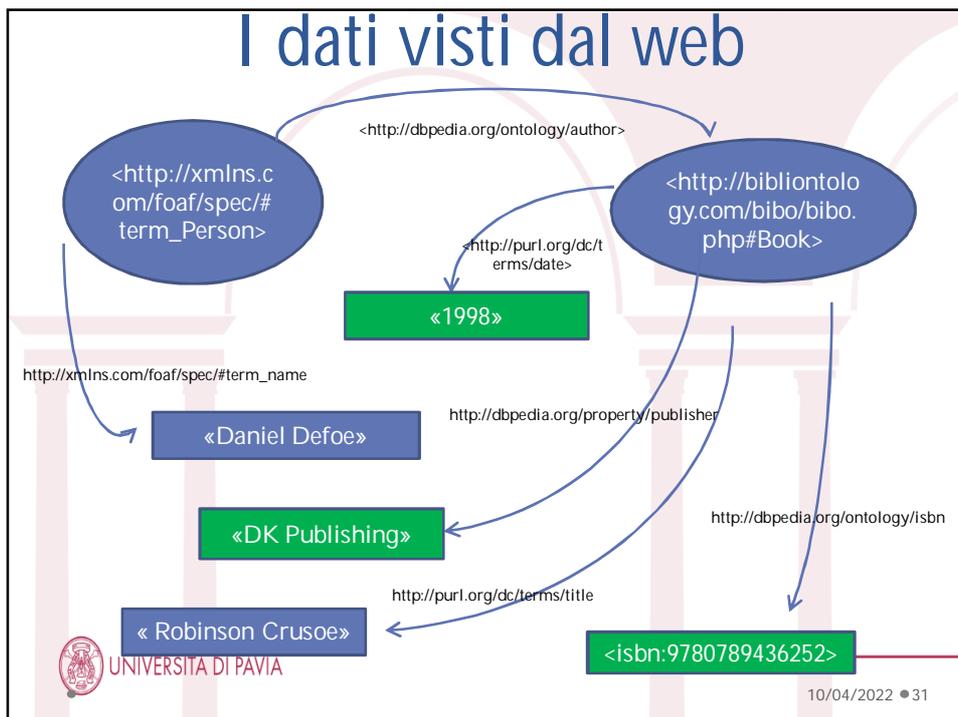
...



UNIVERSITÀ DI PAVIA

10/04/2022 ● 28





Alcuni altri vantaggi di RDF

Il modello logico RDF presenta numerosi altri vantaggi:

1. Il modello di dati RDF consente di creare collegamenti tra dati provenienti da *fonti* diverse
2. RDF consente di rappresentare in un unico grafo informazioni in origine espresse mediante diversi *modelli*, mescolando ontologie e vocabolari diversi
3. Il modello dei dati e i dati possono essere prodotti nella scala preferita (pochi dati/molti dati; poco dettagliato/molto dettagliato).
4. È possibile creare un grafo unico a partire da due insiemi di triple diverse

UNIVERSITÀ DI PAVIA

10/04/2022 ● 32

Web semantico

- Nel web semantico, i computer hanno accesso a raccolte strutturate di informazioni e a insiemi di regole di inferenza da potere utilizzare per sviluppare un «ragionamento automatico».
- Se i dati sono pubblicati come linked data, i computer sono in grado di recuperare dati prodotti da agenzie diverse, interpretarli (cioè riconoscere identità e relazioni) e riutilizzarli correttamente

Il web semantico

- Oltre al web dei documenti, un web di dati.
- I linked data non sostituiscono le tecnologie web, ma ne costituiscono uno strato superiore: sono una trama che si aggiunge alla rete esistente.
- I linked data offrono una modalità di pubblicazione per i dati più generale e più flessibile di quelle adottate finora nel web
- I linked data sono una tecnologia che rende i dati controllati, integrabili, interoperabili, modulari, riutilizzabili, esplorabili

La valorizzazione dei beni culturali con i LOD

...



UNIVERSITÀ DI PAVIA

10/04/2022 ● 35

LOD e i beni culturali

I LOD possono contribuire a promuovere la valorizzazione dei beni culturali garantendo ai (meta)dati relativi alcune caratteristiche non disponibili nel web tradizionale:

- Identificazione
- Scoperta
- Analisi
- Riutilizzo



UNIVERSITÀ DI PAVIA

10/04/2022 ● 36

Risolvere le ambiguità

- I motori di ricerca di oggi (web dei documenti) consentono di effettuare ricerche per stringa, cioè in base a sequenze di caratteri, senza tenere conto del loro significato
 - Per esempio: «pesca»; «gemma»; «Mario Rossi», «Giuseppe Verdi»
- Nei cataloghi delle biblioteche e delle altre istituzioni culturali, questa ambiguità non esiste, perché – di norma – viene risolta con il lavoro di «controllo di autorità»



10/04/2022 ● 37

Il controllo di autorità

Catalogo del Servizio Bibliotecario Nazionale

HOME | RICERCA AVANZATA | **VOCI CONTROLLATE** | LE MIE RICERCHE | ALTRI CATALOGHI | SERVIZI | BIBLIOTECHE SBN | AIUTO

Ricerca: Nome = platone

Stampa E-mail

Scheda 277

Tipo autore Persona

Nome autore Plato

Forme varianti - Platone-ca. 428-347 a.C.>

- Platon
- Platonas
- Platone

Datazione ca. 428 - ca. 347 a. C.

Nota informativa Filosofo greco, unico pensatore antico di cui siano rimaste le opere integrali, fondò l'Accademia, prima scuola filosofica dell'antichità. Nato ad Atene nel 428 o 427 a. C. e morto ivi nel 348 o 347.

Lingua GRECO ANTICO (FINO AL 1453)

Paese GRECIA

Fonti - Censimento delle edizioni italiane del secolo XVI. Roma, ICCU: <http://edit16.iccu.sbn.it/>
 - Enciclopedia italiana di scienze lettere ed arti. Roma, Istituto della Enciclopedia italiana. 1929-
 - World biographical Index. Internet edition. K. G. Saur Electronic Publishing München. www.saur-wbi.de
 - Bibliografia nazionale italiana: nuova serie del bollettino delle pubblicazioni italiane ricevute per diritto di stampa a cura della Biblioteca nazionale centrale di Firenze. A. 1, n. 1 (gen. 1958). Firenze, Centro nazionale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche, 1958- (CDROM)

Regole di catalogazione REICAT

Identificativo SBN IT\ICCU\CFIV\007350

Data di aggiornamento 27/04/2021

10/04/2022 ● 38

Il controllo di autorità nel VIAF

The screenshot displays the VIAF search results for 'Eco, Umberto, 1932-2016'. The interface shows a search bar at the top with the query 'Eco, Umberto, 1932-2016'. Below the search bar, there is a list of preferred forms for the name, including 'Eco, Umberto', 'Eco, U.', 'Eco, V.', and 'Eco, Umberto, 1932-2016'. To the right of the list, there is a circular visualization of the data points, showing a dense cluster of small icons representing different national authorities.

Il valore del controllo di autorità

- Il controllo di autorità è un'attività che gli istituti culturali hanno svolto per decenni, con grandi sforzi e grandi risultati, nei loro rispettivi cataloghi.
- Grazie al VIAF, i diversi «silos» nazionali sono stati collegati tra loro a livello internazionale e il VIAF è diventata una risorsa molto importante per il web semantico
- Ma il VIAF è progettato per le persone-autori, mentre nel web semantico c'è molto di più, come abbiamo visto.
- I dati di autorità delle biblioteche e delle altre istituzioni culturali sono molto preziosi per il web semantico; ma allo stesso tempo sono parziali, perché riguardano soltanto certi tipi di entità (le persone, gli enti, le risorse bibliografiche ...)

La gestione delle identità

- Nel web semantico *ogni* persona, luogo, oggetto, concetto, relazione, proprietà ecc. deve essere identificato in modo assoluto (URI), cioè in modo da essere completamente distinto da qualsiasi altra entità che abbia come nome la stessa stringa di caratteri
- Ciò vale anche se l'entità è descritta più volte in fonti (dataset) diversi. Infatti ogni produttore di LOD può assegnare un proprio URI a un'entità da identificare
- Quindi esisteranno potenzialmente molti URI – tutti validi – per la stessa entità (una città, una persona famosa ecc.)
- Quarto principio: collegare!



UNIVERSITÀ DI PAVIA

10/04/2022 ● 41

La gestione delle identità 2

- Nel momento in cui si pubblicano i propri dati nel web semantico si deve svolgere un lavoro nuovo, che è la gestione delle identità
- Si tratta di creare e pubblicare delle corrispondenze (relazioni di identità) tra i propri dati che si vanno a pubblicare e i dati già presenti e disponibili nel web semantico
- Alcuni dataset nella nuvola dei LOD – come, per esempio Wikidata – svolgono un ruolo determinante per la gestione delle identità in tutto il web semantico
- Wikidata si è trasformato in un «identity hub» del web semantico e la sua importanza è indiscussa e in continua crescita.



UNIVERSITÀ DI PAVIA

10/04/2022 ● 42

Wikidata come «identity hub»

- In Wikidata le forme varianti del nome di un'entità vengono controllate

The screenshot shows the Wikidata profile for Carlo Collodi (Q199588). A table lists name variants in different languages, and another table lists aliases. Red boxes highlight these sections.

Lingua	Etichetta	Descrizione	Anche conosciuto come
italiano	Carlo Collodi	scrittore italiano	Carlo Lorenzini Carlo Lorenzini Collodi
tedesco	Carlo Collodi	italienischer Schriftsteller (Pinocchio)	Carlo Lorenzini
inglese	Carlo Collodi	Italian writer	Carlo Lorenzini
spagnolo	Carlo Collodi	escritor italiano y cantante	Carlo Lorenzini Collodi Carlo Lorenzo Filippo Giovanni ...
francese	Carlo Collodi	écrivain italien	Carlo Lorenzini

Wikidata come «identity hub»

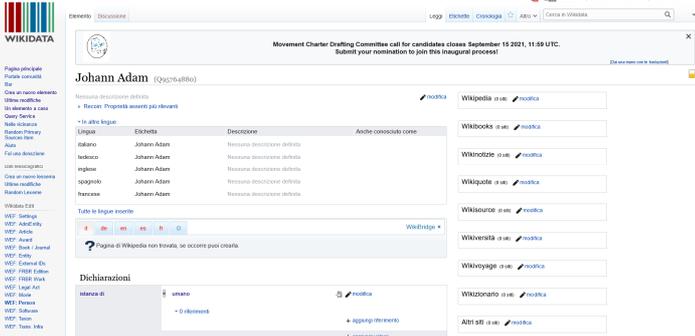
- In Wikidata concetti diversi rappresentati dalla stessa stringa, vengono nettamente distinti

The screenshot shows the Wikidata profile for Carlo Collodi (Q199588). A dropdown menu is open for the label 'Caravaggio', showing several distinct entities with the same label. A red box highlights this dropdown menu.

- Caravaggio pittore italiano
- Caravaggio comune italiano
- Caravaggio pagina di disambiguazione di un progett...
- Caravaggio film del 1986 diretto da Derek Jarman
- Caravaggio 2007 Italian television movie
- Hitachi Caravaggio (Caravaggio) Elettrotreno regionale italiano
- Caravaggio cognome

Wikidata come «identity hub»

- Wikidata consente anche di individuare gli eventuali errori di identificazione presenti in dataset esterni (come il VIAF!)



Wikidata

Johann Adam (Q92714816)

Lingua	Etichetta	Descrizione	Anche conosciuto come
italiano	Johann Adam	nessuna descrizione definita	
tedesco	Johann Adam	nessuna descrizione definita	
inglese	Johann Adam	nessuna descrizione definita	
spagnolo	Johann Adam	nessuna descrizione definita	
francese	Johann Adam	nessuna descrizione definita	

Tutte le lingue in cui

Wikipedia: non trovata

Dichiarazioni

Altre lingue in cui

UNIVERSITÀ DI PAVIA

10/04/2022 ● 45

Analisi

- La pubblicazione dei dati in una forma più granulare consente di effettuare con estrema facilità operazioni di analisi dei dati in risposta a quesiti anche complessi
- Avere il proprio dataset in forma di LOD consente di effettuare *ricerche esplorative* che gli stessi dati, in una struttura più rigida, non consentono di fare
- Da notare che i dati, nella maggioranza dei casi, non sono diversi o maggiori di quelli già creati e a disposizione delle istituzioni MAB; sono semplicemente *strutturati e pubblicati* in modo diverso

Analisi. Esempio del CoBiS

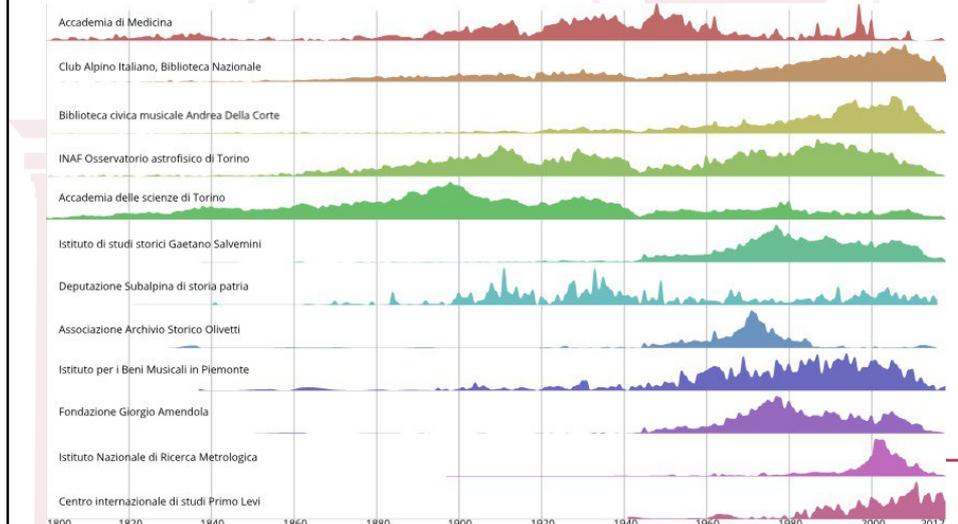
- Il progetto Linked Open Data del Coordinamento Biblioteche Speciali e Specialistiche di Torino (CoBiS), realizzato grazie ad una convenzione con la Regione Piemonte, crea collegamenti tra i dati bibliografici delle biblioteche che aderiscono alla sperimentazione e la comunità del Web
- I dati catalografici provenienti da differenti software gestionali sono ora accessibili tramite un'unica ricerca e possono essere arricchiti con informazioni provenienti dal web semantico.



UNIVERSITÀ DI PAVIA

10/04/2022 ● 47

Analisi. CoBiS



Analisi. Scholia

- Scholia è un strumento gratuito messo a disposizione dalla comunità di Wikimedia per l'analisi delle pubblicazioni scientifiche
- Consente di analizzare la produzione di un ricercatore, oppure di un dipartimento universitario, oppure di una rivista, anche mettendoli a confronto tra loro
- Si basa sui dati presenti su Wikidata.



UNIVERSITÀ DI PAVIA

10/04/2022 ● 49

Analisi. Scholia

Esempi

- Studioso: Denny Vrandečić ([Q18618629](#))
- Dipartimento: Technical University of Denmark ([Q1269766](#))
- Rivista: JLIS.it (Italian Journal of Library, Archives and Information Sciences) ([Q3805294](#))
- Confronto tra riviste: *Biogeochemistry* vs *Biogeosciences* ([confronta](#))



UNIVERSITÀ DI PAVIA

10/04/2022 ●

Scoperta

La scoperta resa possibile dai LOD consiste nel:

- 1. poter porre domande in precedenza impossibili o inedite
- 2. scoprendo *dataset rilevanti* che non si conoscevano in anticipo

Esempi di domande:

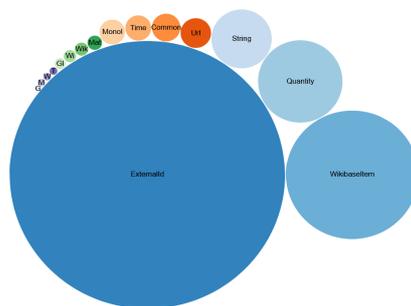
1. Posso avere un elenco di giornalisti nati a Parigi e morti a New York?
2. Posso avere una mappa degli ospedali vicini a Pavia nel raggio massimo di 50 km?
3. Posso avere una linea del tempo con le tele di Tiziano?
4. Quali opere di autori nati a Venezia nella prima metà dell'800 sono state pubblicate in Germania?
5. Quali opere figurative del '700 rappresentano il tema (barocco) dello scorrere del tempo?



UNIVERSITÀ DI PAVIA

10/04/2022 ● 51

Scoperta 2



Wikidata Query Service



UNIVERSITÀ DI PAVIA

<http://w.wiki/fl8> by Toby Hudson

10/04/2022 ●

Scoperta 3

The screenshot shows the iNaturalist profile for 'Porcino (Boletus edulis)'. The main content area features a large photograph of several porcini mushrooms on a bed of autumn leaves. To the right of the image is a sidebar with various details:

- MIGLIOR OSSERVATORE:** viktoribilous (Classifica A 94)
- ULTIMA OSSERVAZIONE:** 12 Settembre, 2021 (YouTuber osservatore)
- MIGLIOR IDENTIFICATORE:** tmaximo (Classifica A)
- OSSERVAZIONI TOTALI:** 7.713 (Mostra tutti)

The sidebar also contains a 'Stagionalità' section with a line graph showing observation frequency from October to December. Below the graph are tabs for 'Mappa', 'Informazioni', 'Tassonomia', 'Stato', and 'Specie simili'. At the bottom, there is a 'Fonte: Wikipedia' link and a button for 'Ulteriori informazioni'.

Scoperta 4

The screenshot shows the Spotify artist profile for Leonard Cohen. The main content area features a large black and white portrait of the artist and a list of popular songs:

Rank	Song	Streams	Duration
1	Hallelujah	66,219,637	
2	Suzanne	91,406,445	3:49
3	You Want It Darker	48,231,775	4:44
4	Dance Me to the End of Love	34,303,239	4:40
5	Famous Blue Raincoat	44,142,729	5:10

To the right of the song list is a detailed information sidebar for Leonard Cohen, including his biographical details, a list of genres (blues, folk, musica pop, rock, world music, soft rock, folk rock, pop rock), and a link to the 'Love and Hate 50th Anniversary edition' playlist.

Riuso. CoBiS



[HOME](#) [ABOUT](#) [ONTOLOGIA](#) [COBIS](#) [BIBLIOTECHE](#) [OPERE](#) [AUTORI](#) [PERCORSI](#)



AUTORE
Angelo Mosso

Torino, 1846
Torino, 1910
MEDICO, FISILOGO E ARCHEOLOGO ITALIANO

LINKED DATA

Interlinking
[Wikipedia](#) [Virtual International Authority Base](#) [Deutsche Nationalbibliothek](#) [BnF Catalogue Général](#) [Open SBN](#) [DBI Treccani](#) [Open Library](#)

Professione: medico
Istruzione: Università degli Studi di Torino
Membro di: Accademia delle Scienze di Torino

Biografia: Angelo Mosso (Torino, 30 maggio 1846 – Torino, 24 novembre 1910) è stato un medico, fisiologo e archeologo italiano.

OPERE DATI/DE

Cerca un'opera per titolo, autore o data di edizione...


UNIVERSITÀ DI PAVIA
10/04/2022 ● 55

Alcune delle informazioni visualizzate in questa pagina provengono da fonti esterne distribuite nel Web. Il counter sulla sinistra mostra il numero di informazioni che sono recuperate in tempo reale grazie al Linked Open Data. Le singole informazioni sono evidenziate da un riquadro giallo.

(Catalogo Pontificio Università della Santa Croce)

Riuso. Il catalogo PUSC

Home / Dettagli di: Semiotica e filosofia del linguaggio /

Visualizzazione normale Vista MARC Visualizzazione ISBD

Semiotica e filosofia del linguaggio / Umberto Eco.

Per: Eco, Umberto (autore)

Serie: Biblioteca studio ; 29.
 Dettagli di pubblicazione: Torino : G. Einaudi, 1996
 Descrizione: XVII, 318 p. : ill. ; 21 cm
 ISBN: 8806140655
 Classificazione Dewey: 401.41
 Dewey: 401.41 SEMIOTICA

Dati di copia (1) Note (1) Commenti (0) QR codes

N. Collocazione Localizzazione Copia Barcode In prestito Fino al Note di copia
 1 401.41 / 2 Generale Copy 001 100401 Disponibile

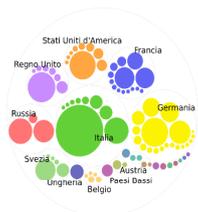
AuthorityBox ↕

Eco, Umberto

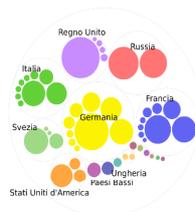
 05/01/1932 Alessandria - 19/02/2016 Milano
 Italiano, francese
 Eco, Umberto Filosofo, scrittore, semiologo, professore emerito e presidente della Scuola Superiore di Studi Umanistici dell'Università di Bologna (2008), studioso di estetica

 UNIVERSITÀ DI PAVIA

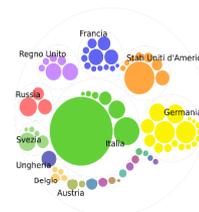
Riuso. CoBiS e Wikidata



Affiliazioni accademie
1500/oggi

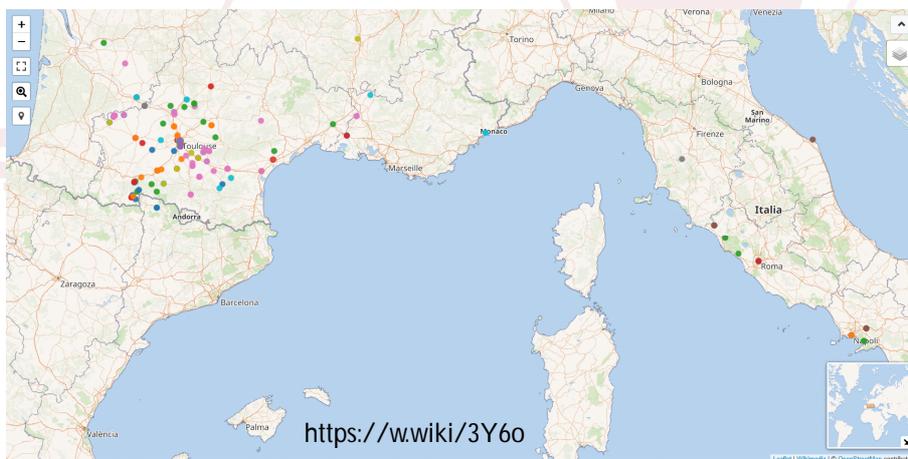


Affiliazioni accademie
1500/1900



Affiliazioni accademie
1900/oggi

Riuso. Musée Saint Raymond di Toulouse e Wikidata



I LOD con Wikidata

Pubblicare i propri dati su Wikidata ha inoltre altri vantaggi

1. Visibilità (Metropolitan Museum; BEIC)
2. Cooperazione
3. Controllo di qualità dei dati (in entrambe le direzioni)



UNIVERSITÀ DI PAVIA

10/04/2022 ● 59

GWMAB

Per promuovere l'incontro tra professionisti della cultura e Wikidata è nato il *Gruppo Wikidata per Musei Archivi e Biblioteche*, promosso dall'Università di Pavia. Ha lo scopo di:

1. promuovere la conoscenza di Wikidata all'interno del mondo MAB;
2. diffondere la conoscenza e l'uso degli strumenti gratuiti disponibili per la produzione dei dati su Wikidata;
3. aumentare il numero di persone che, a titolo personale o istituzionale, collaborano alla costruzione di Wikidata;
4. promuovere progetti di ricerca e sul campo di istituzioni MAB per la pubblicazione di LOD tramite Wikidata;
5. favorire il confronto, la diffusione e lo scambio di punti di vista e *best practises*, di esperienze teoriche e pratiche e di progetti di ricerca sviluppati dai partecipanti, a titolo personale o istituzionale.



UNIVERSITÀ DI PAVIA

10/04/2022 ●

Per concludere

I LOD rappresentano uno strumento fondamentale per la valorizzazione dei beni culturali:

1. L'identificazione consente di collegare, integrare, rendere interoperabili i propri dati con quelli di altre istituzioni culturali e non
2. L'analisi permette una visualizzazione efficace dei dati delle collezioni utile sia per la loro gestione che per il loro utilizzo
3. La scoperta consente un approccio di ricerca ai dati delle collezioni e del SW davvero *esplorativo* e basato sul *contenuto*
4. Il riuso consente di arricchire e diffondere i propri dati, promuovendoli in una logica di economia di scala e lungo periodo
5. I progetti come Wikidata aumentano la visibilità, la cooperazione e il controllo sulla qualità dei dati



UNIVERSITÀ DI PAVIA

10/04/2022 ● 61

Grazie!

• • •

carlo.bianchini@unipv.it

https://www.wikidata.org/wiki/Wikidata:Gruppo_Wikidata_per_Musei,_Archivi_e_Biblioteche



UNIVERSITÀ DI PAVIA

10/04/2022 ● 62