

L'archiviazione delle risorse web

Chiara Storti

bnc-fi.magazzinidigitali@cultura.gov.it

**Web Archiving: conservare la memoria digitale
tra sfide culturali e nuove responsabilità**

16 marzo 2026



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Che cos'è il Web archiving

The screenshot shows the Bologna Biblioteche website with a 404 error message. The header includes the logo, navigation links, and a search bar. The main content area displays the error message and a link to the home page. The footer contains contact information and a list of services.

Logo: bologna biblioteche

Navigation: CHI SIAMO ▾ BIBLIOTECHE NEWS ATTIVITÀ ▾ COME FARE PER ▾ CHIEDILO ALLA BIBLIOTECA ▾ PROPOSTE DI LETTURA

Search: Cerca nel sito CERCA

Error Message: QUESTO CONTENUTO NON È PIÙ DISPONIBILE.
PUÒ CAPITARE CHE ALCUNE PAGINE VENGANO STRAPPATE DAI LIBRI, ANCHE DA QUELLI DIGITALI.

Link: > [Vai alla home page](#)

Footer:

- Comune di Bologna**
- Settore Biblioteche e Welfare culturale**
- Piazza Maggiore, 6, Bologna
- telefono +39.051.2194400 c/o
- [Biblioteca Salaborsa](#)
- [email](#)

Services:

- Informazioni**
 - Statistiche
 - Area stampa
 - Documenti
 - Seguici sui social
 - Settore Biblioteche e Welfare culturale
 - Trattamento dei dati personali
- Attività**
 - Gruppi di lettura
 - Proposte per le scuole
 - Facilitazione digitale
 - Progetti speciali
 - Archivio newsletter
- Come fare per**
 - Iscriversi
 - Iscriversi a Emilib
 - Leggere, studiare, consultare
 - Cercare un libro
 - Prendere un libro in prestito
 - Riprodurre un documento
 - Navigare in Internet
 - Giocare
- Chiedilo alla biblioteca**
 - Come funziona
 - Fai la tua domanda

Footer: Webdes: ign+CMS [chialab](#)

STORIES —

SPOTLIGHTING THE WORLD FACTBOOK AS WE BID A FOND FAREWELL

February 4, 2026



One of CIA's oldest and most recognizable intelligence publications, The World Factbook, has sunset. The World Factbook served the Intelligence Community and the general public as a longstanding, one-stop basic reference about countries and communities around the globe. Let's take a quick look into the history of The World Factbook.

<https://www.cia.gov/stories/story/spotlighting-the-world-factbook-as-we-bid-a-fond-farewell/>

INTERNET ARCHIVE

WEB TEXTS VIDEO AUDIO SOFTWARE IMAGES

ABOUT BLOG EVENTS PROJECTS HELP

INTERNET ARCHIVE

Wayback Machine

Explore more the

Enter a URL

Tools

Wayback Machine Availability API

Chrome Extension

Firefox Add-on

Safari Extension

MS Edge Add-on

Subscription Services

Archive-It enables you to capture search collections of digital content, technical expertise or hosting for it to build and browse the collection.

<https://web.archive.org/>

THE WORLD FACTBOOK (2021 ARCHIVE)

Country Flag

Country Map

Download Country Documents

Locator Map

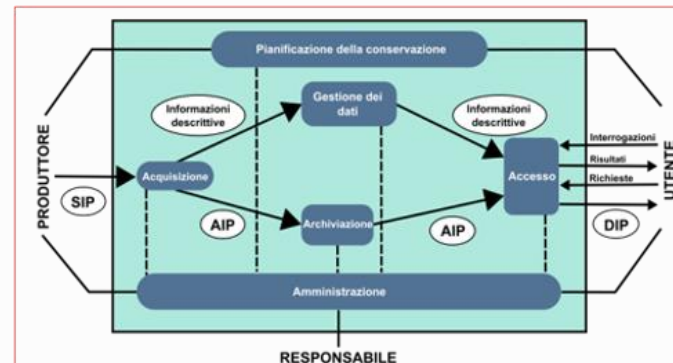
Photos of Italy

November 24, 2021

<https://web.archive.org/web/20221009214139/https://www.cia.gov/the-world-factbook/about/archives/2021/countries/italy/>

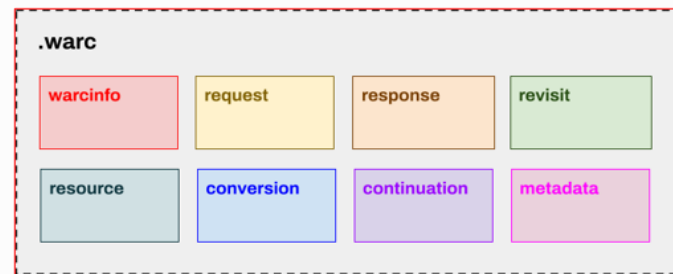
Che cos'è il Web archiving?

- Il processo che consente di **catturare delle istantanee**, più o meno profonde, di porzioni di Web, ad intervalli di tempo definiti, di **archiviarle e renderle accessibili nel lungo periodo**.
- la tecnica di cattura più comune* si basa sull'utilizzo di **Web crawler**, il più diffuso è Heritrix (Open source).
- Il formato di archiviazione più diffuso attualmente è il **WARC** che, nel 2009, viene codificato nello standard ISO 28500 (versione corrente aggiornata ISO 28500:2017).



Modello OAIS

Di Mathieualexhache (original work) -
File:Schema fonctionnel modele oais.jpg,
CC BY-SA 4.0



File WARC Schema

<https://archive-it.org/post/the-stack-warc-file/>

Che cos'è il Web archiving?

- Lo strumento di accesso più usato è la **Wayback Machine** (Open source), tool sviluppato e mantenuto principalmente, ma non solo, da Internet Archive.
- L'istanza più nota di Wayback Machine è la **Wayback Machine di Internet Archive!**


**INTERNET
ARCHIVE**

 WEB
  TEXTS
  VIDEO
  AUDIO
  SOFTWARE
  IMAGES

[SIGN UP](#) | [LOG IN](#) | [UPLOAD](#)

[ABOUT](#)
[BLOG](#)
[EVENTS](#)
[PROJECTS](#)
[HELP](#)
[DONATE](#)
[CONTACT](#)
[JOBS](#)
[VOLUNTEER](#)



Explore more than 1 trillion [web pages](#) saved over time




Tools

- [Wayback Machine Availability API](#)
- [Chrome Extension](#)**
- [Firefox Add-on](#)
- [Safari Extension](#)
- [MS Edge Add-on](#)


Subscription Service

Archive-It enables you to capture, manage and search collections of digital content without any technical expertise or hosting facilities. [Visit Archive-It](#) to build and browse the collections.


Collection Search

This service is based on indexes of specific data from [selected Collections](#).


Save Page Now

Capture a web page as it appears now for use as a trusted citation in the future.

<https://web.archive.org/>

Cosa archiviamo

WEBSITE STATISTICS 2025

How Many Websites Are There in the World?

UPDATED DAILY

All you ever wanted to learn about websites in one place. Number of websites worldwide (updated), stats, facts, data, charts, and much more.

Explore Website Tools

Read Full Analysis

1,354,989,060

TOTAL WEBSITES

Currently, there are around 1.4 billion websites in the World

206,211,019

ACTIVE WEBSITES

15% of these websites are active, 85% are inactive

Il [dominio.it](https://www.dominio.it)
registra
circa 3.500.000
siti

<https://siteefy.com/how-many-websites-are-there/> (dati al 20/1/2026)




THE LIFESPAN OF A WEBSITE

How long does a website last?

In general, a typical web page can be expected to last **-90-100 days** before changing, moving, or disappearing completely.

- > In 2013, our colleagues at Old Dominion University determined that **over 10%** of event related content posted to social media platforms is lost after one year.
- > In 2014, a study by UCLA determined that **7-in-10** scholarly articles that include citations with hyperlinks suffer from *reference rot*.



Brewster Kahle
 @brewster_kahle

This is a book that has every footnote archived. So good. (and the book is free and wonderful to read).

Traduci il Tweet



Glyn Moody @glynmoody · 17h

the ebook version of "Walled Culture" uses 750 links to the indispensable @internetarchive to avoid link rot; so it seems only right that you can download "free" copies from the same great site, forever: archive.org/details/walled...

07:03 · 09 ott 22 · Twitter Web App



Pew Research Center

RESEARCH TOPICS · PUBLICATIONS · OUR METH

[Home](#) · [Research Topics](#) · [Internet & Technology](#)

REPORT | MAY 17, 2024

When Online Content Disappears

38% of webpages that existed in 2013 are no longer accessible a decade later

BY ATHENA CHAPERIS, SAMUEL BESTVATER, EMMA BERRY AND GONZALO RIVERO

How we did this

The internet is an unimaginably vast repository of modern life, with hundreds of billions of indexed webpages. But even as users across the world rely on the web to access books, images, news articles and other resources, this content sometimes disappears from view.

A new Pew Research Center analysis shows just how fleeting online content actually is:

- A quarter of all webpages that existed at one point between 2013 and 2023 are no longer accessible, as of October 2023. In most cases, this is because an individual page was deleted or removed on an otherwise functional website.

<https://www.pewresearch.org/data-labs/2024/05/17/when-online-content-disappears/>

La durata media della vita di un sito web è 2 anni e 1 mese

Cosa archiviamo

Non possiamo archiviare tutto ma dovremmo preoccuparci il prima possibile, fin dalla formazione delle nostre risorse digitali, di come queste verranno archiviate.

Cosa archiviamo

*Di certo, la **tutela del web come oggetto culturale** non costituisce l'unica motivazione che conduce alla sua archiviazione: altre – e in determinati frangenti - più stringenti necessità possono essere rappresentate dalla **individuazione di quegli elementi in grado di garantire stabilità, oggettività e autorevolezza al contenuto giuridico e amministrativo dei documenti formati tramite esso e depositati in esso.***

Alberto Salarelli, *Luci e ombre del Web archiving*, 2025

Cosa archiviamo

- Siti e altre risorse web che hanno un interesse o un **valore "culturale"**.
- Siti e altre risorse web che hanno un **valore "documentale"** (non in alternativa alla conservazione dei documenti come da previsioni del CAD, ma come supporto probatorio).
- Siti e altre risorse web che le **tecnologie ci consentono di archiviare** (*Reborn web*).
- Siti e altre risorse web che **la normativa vigente** permette o impone di conservare.



Condivisi con me > UNO BIANCA

Tipo Persone Data modifica Sorgente

Nome	Proprietario	Data modifica	Dimensioni f	IP	Ordina
4 dicembre 25 biblioteca assemblea legislativa	maurizio.matrone	5 dic 2025	—		
4 gennaio 26 Mirare allo Stato	maurizio.matrone	13 gen	—		
8 gennaio 26 mirare allo stato - studenti	maurizio.matrone	18 gen	—		
12 dicembre 25 Mambo, mostra Nadalini	maurizio.matrone	13 dic 2025	—		
13 novembre 25 archivio di stato	maurizio.matrone	13 nov 2025	—		
19 novembre 25 memorie da una città ferita	maurizio.matrone	21 nov 2025	—		
20 novembre 25 biblioteca della assemblea legislativa	maurizio.matrone	27 nov 2025	—		
27 novembre 25 Biblioteca della assemblea legislativa	maurizio.matrone	27 nov 2025	—		
ARCHIVI pubblici e privati	maurizio.matrone	23 gen 2025	—		
ARTE	maurizio.matrone	25 giu 2024	—		
CINEMA	maurizio.matrone	25 giu 2024	—		
DOCUMENTARI	maurizio.matrone	25 giu 2024	—		
FOTOGRAFIE	maurizio.matrone	25 giu 2024	—		
FUMETTI	maurizio.matrone	25 giu 2024	—		

Limiti alla cattura, un esempio:
<https://www.unobianca.it/>

Archiviabilità dei siti web

Gli sviluppatori e i designer di siti web devono oggi tenere conto dei criteri di accessibilità, performance, SEO – Search Engine Optimization, compatibilità con gli standard del W3C e usabilità.

Tuttavia, tra le buone pratiche emergenti c'è quella di prendere in considerazione anche i criteri di archiviabilità dei siti web.

Per **archiviabilità** si intende l'insieme delle caratteristiche che i contenuti, la struttura, le funzionalità e le interfacce di un sito web devono possedere perché il sito stesso possa essere conservato e reso accessibile nel lungo periodo con gli attuali strumenti di web archiving.

- ↳ Linee guida per la realizzazione di siti web archiviabili
- ↳ Validatore



<https://www.bncf.firenze.sbn.it/biblioteca/archiviabilita-dei-siti-web/>

Il Web archiving in Italia e il servizio della Biblioteca nazionale centrale di Firenze

Il deposito legale delle risorse digitali: la situazione italiana

- In Italia il deposito legale dei documenti diffusi tramite rete informatica, compresi siti e altre risorse Web, NON È OBBLIGATORIO.
- Tuttavia, sulla base del mandato dell'art. 37 c. 2 del [D.P.R. 252/2006 - Regolamento recante norme in materia di deposito legale dei documenti di interesse culturale destinati all'uso pubblico](#), dal 2010 è attivo il servizio **MAGAZZINI DIGITALI**.

Il deposito legale delle risorse digitali: la situazione italiana



Magazzini Digitali è il servizio nazionale di conservazione e accesso a lungo termine alle pubblicazioni digitali:

- tesi di dottorato
- ebook (deposito volontario attualmente sospeso)
- ejournal Open Access (deposito volontario attualmente sospeso)
- siti e altre risorse web (Web e Social media archiving)

Il deposito legale delle risorse digitali: la situazione italiana

- I siti e i contenuti web sono soggetti alla normativa relativa al Diritto d'autore, che ne impedisce la copia non autorizzata.
- La normativa sul deposito legale digitale, tra le altre cose, dovrebbe regolare le eccezioni al Diritto d'autore per finalità storiche e culturali per quanto riguarda i siti web e le altre risorse digitali.
- In assenza di tale normativa, è possibile che le istituzioni depositarie del deposito legale trovino degli accordi con i produttori dei contenuti che regolino “depositi volontari” con opzioni *opt-in* o *opt-out*.
- **La BNCf porta avanti il servizio nazionale di Web archiving chiedendo un'adesione volontaria ai gestori dei siti (opt-in), che hanno anche facoltà di decidere se rendere le copie archiviate liberamente accessibili sul Web o soltanto dalla rete interna della Biblioteca.**

BIBLIOTECA RISORSE ATTIVITÀ AMMINISTRAZIONE

Biblioteca Nazionale Centrale di Firenze

Home > Servizi > Web archiving

Web archiving

La Biblioteca Nazionale Centrale di Firenze, nell'ambito del servizio Magazzini Digitali, raccoglie, conserva e rende accessibili in maniera permanente contenuti web di interesse per la cultura e la storia italiana.

Il progetto

I siti web e i documenti in essi contenuti sono considerati oggetti digitali "effimeri", chiunque, infatti, navigando sul Web ha fatto esperienza dei cosiddetti "link rotti", vedendosi restituire l'errore 404. È però altrettanto indubbio che essi siano una fonte ormai imprescindibile per la storia e la cultura contemporanea.

Per questo motivo, nel 2018 la BNCF, nell'ambito del più ampio servizio di conservazione e accesso a lungo termine alle pubblicazioni digitali, ha avviato un **programma di Web archiving**, in maniera analogica e complementare a quanto già fanno le grandi istituzioni della memoria del resto del mondo.

Il progetto

Come aderire
Requisiti tecnici per la raccolta
Archiviabilità dei siti web
Accesso alle collezioni
Link utili
Contributi online in italiano
Contatti

HOME EXPLORE LEARN MORE CONTACT US

The leading web archiving service for collecting and accessing cultural heritage on the web Built at the Internet Archive

Explore >> National Central Library of Florence

National Central Library of Florence

Archive-It Partner Since: gen. 2018
Organization Type: National Institutions
Organization URL: <http://www.bncf.firenze.sbn.it>
Description: Biblioteca Nazionale Centrale di Firenze

Narrow Your Results

Collections Sites Search Page Test

Enter search terms here

Sort By: Collection Name (A-Z) | Collection Name (Z-A)

Page 1 of 1 (11 Total Results)

There are no further ways to narrow your results.

The following results were found for the term(s): **MAB**

- 1 Sites were found.
- Additional results for **MAB** may be found by searching within the page text.

Sort By: Best Match | Title (A-Z) | Title (Z-A) | URL (A-Z) | URL (Z-A)

Title: MAB Italia
URL: <http://www.mab-italia.org/>

Description: MAB è l'acronimo con cui AIS (Associazione Italiana Biblioteche), ANAI (Associazione Nazionale Archivistica Italiana) e ICOM Italia (International Council of Museum - Comitato Nazionale Italiano), nella primavera del 2011 (svolgendo un'iniziativa precedente avviata dalle loro Sezioni parimenti), hanno dato vita a un coordinamento permanente per esplorare le prospettive di convergenza tra i musei e gli istituti in cui operano i professionisti degli archivi, delle biblioteche, dei musei.

Captured 15 times between June 30, 2023 and October 16, 2025
Videos: 3796 Videos Captured

Nel 2018 la BNCF ha attivato un servizio di [Web archiving](https://wayback.archive-it.org/11345/*) e archivia periodicamente quasi 600 siti web liberamente accessibili su [Archive-it](https://wayback.archive-it.org/11345/*).

https://wayback.archive-it.org/11345/* / <http://www.mab-italia.org/>

Giornata della Memoria 2026. “La favola di Natale” di Giovannino Guareschi

13.11.2025

Sospensione informazioni +39 055 24919 96-97

Dal 27 novembre Codice Fiscale: 80020550481

Tesi di dottorato

BNCF

Chi siamo

Apertura e orario dei servizi

Dove siamo

Privacy Policy

Cookie Policy

Dichiarazione di accessibilità

[Sito archiviato](#)

Amministrazione

URP

Contatti

Amministrazione trasparente

Area riservata

Attività, eventi, incontri

Iscriviti alla nostra newsletter per non perderti le ultime novità.

Nome

E-mail

☐ *Dichiaro di aver letto e compreso la Privacy Policy

[ISCRIVIMI](#)

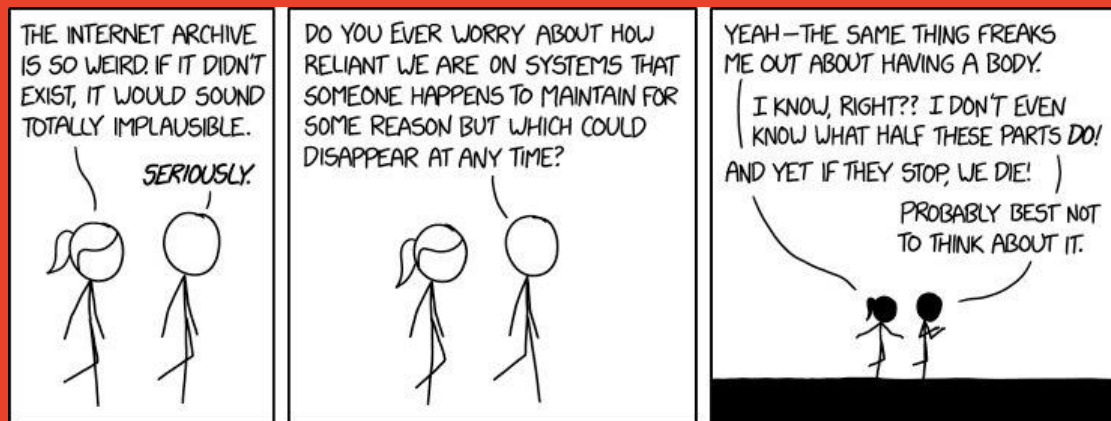
Le [copie archiviate](#) del [sito di BNCF](#)

Il Web archiving come sfida collaborativa

La **conservazione delle risorse web** è una sfida inedita che può essere affrontata solo con un **modello di governance completamente nuovo**.

Il Web archiving come sfida collaborativa

- Individuare le istituzioni pubbliche deputate alle attività di Web archiving **abbandonando le tradizionali categorie di “risorsa bibliografica”, “archivistica” o “museale”** e ridistribuendo gli ambiti di competenza in base alla tipologia di informazione veicolata più che alla forma delle risorse o a chi le ha prodotte.
- **Coinvolgere le "comunità"**, intese sia come gruppi portatori di una storia o di un interesse comune, per individuare e preservare le “memorie particolari”, sia come comunità di esperti / istituzioni private (es. Internet Archive) e attivisti (es. [Archive Team](#)) che insieme lavorano per salvaguardare l’infrastruttura cooperativa per eccellenza: il Web.



Fonte:
<https://xkcd.com/>

Se dopo questo incontro, ormai consapevoli dell'importanza dell'archiviazione delle risorse web, vorrete usufruire del servizio della BNCF, ne saremo felici. Tuttavia, il *team* dedicato a queste attività è davvero ridotto quindi vi chiediamo di essere pazienti: risponderemo a tutti!

bnc-fi.magazzinidigitali@cultura.gov.it