

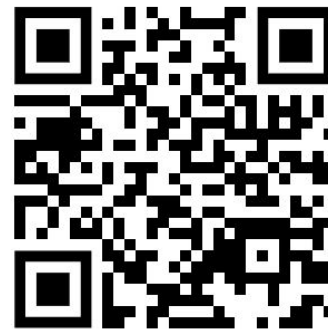
Why care about ARK identifiers?



- Because robust web links are rare – the average URL lifetime is 100 days
- ARKs can be “persistent” identifiers (PIDs), but we prefer “persistable”
- “Ten persistent myths about persistent identifiers”

<https://n2t.net/ark:/13030/c7gb1xh09>

The ARK (Archival Resource Key) identifier scheme was introduced in 2001.



How cultural heritage links break

GLAMs use DAMS,
DAMS use databases,
databases generate new URLs,

frequently.

For example, widespread URL change can occur

- when the database is reloaded,
- when the software is updated,
- when the content migrates between DAMS vendors

Venn of a 404



ARK anatomy

A labelled URL with a globally unique identity inside it



<https://n2t.net/ark:/12345/fk1234>

↑
makes ARK
actionable
(the resolver)

↑
core globally unique
identity (independent
of web and hostname)



N2T.net is a global “name” to “thing” resolver

Why not “ARKresolver.net” like
most other PID schemes?

Because ARKs are inclusive
and resolvers generalize
easily.

N2T keeps identifiers persistent,
forwarding them to the best
known web addresses

Any kind of name – ARK, DOI,
URN, Handle, PMID, PDB, Taxon,
GRID, arxiv, ISSN, ...

Partners with [EZID.cdlib.org](https://ezid.cdlib.org/),
[Identifiers.org](https://identifiers.org/), [Archive.org](https://archive.org/),
[YAMZ.net](https://yamz.net/) metadictionary

Any kind of thing – data, web
page, physical specimen, group,
vocabulary term, living being, ...

N2T is a global [ARK](#) resolver

**Also a meta-resolver for 900+
kinds of [compact Identifiers](#)**

ARK organizations

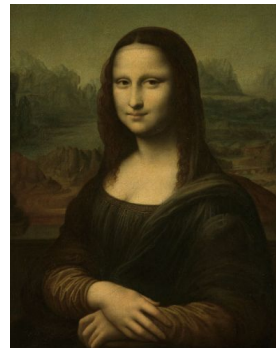
8.2 billion ARKs created by 1100+ institutions – libraries, archives, museums, publishers, data centers, educators, etc. For example,



Internet Archive
Bodleian Libraries
Berkeley Law Library
Bibliothèque Mazarine
New York Public Library
French National Archives
National Library of Austria
Library and Archives Canada

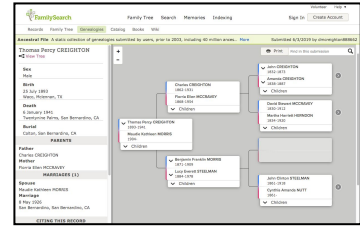
University of California Berkeley
Smithsonian National Museum
National Library of France
University of Chicago
Musée du Louvre
Family Search
British Library
Google

<https://n2t.net/ark:/53355/cl010066723> →



What are ARKs used for?

- genealogical records (8 billion [FamilySearch](#))
- publisher content (100 million [Portico](#))
- scientific datasets and records (22 million [INIST](#))
- scanned books and texts (30 million [Internet Archive](#))
- bibliographic records (15 million [BnF main catalog](#))
- museum specimens (15 million [Smithsonian Institution](#))
- public health documents (15 million [UCSF IDL](#))
- historical documents (21 million [CDL](#), 5 million [BnF Gallica](#))
- historical authors and scholars (4 million [SNAC](#))
- fine art museum collections (490,000 [Louvre](#))
- vocabulary terms (30,000 [Periodo](#), [YAMZ](#))



archive.org
ARKs: 13960



USMARC Code List for Languages

by [Network Development and MARC Standard office](#)

Publication date	1996
Collection	inlibrary ; printdisabled ; internetarchivebooks
Digitizing sponsor	Kahle/Austin Foundation
Contributor	Internet Archive
Language	English
Access-restricted-item	true
Addeddate	2023-03-08 20:13:15
Autocrop_version	0.0.14_books-20220331-0.2
Bookplateleaf	0002
Boxid	IA40872114
Camera	Sony Alpha-A6300 (Control)
Collection_set	printdisabled
External-identifier	um:lcp:usmarccodelistfo0000netw:epub:34d7b206-8305-40a5-9027-3cc1b010af2e um:lcp:usmarccodelistfo0000netw:lcpdf:ec98575a-5387-49cb-923f-3260f1adeadb
Foldoutcount	0
Identifier	usmarccodelistfo0000netw
Identifier-ark	ark:/13960/s2wj1b5txr4
Invoice	1652

History of “persistable” id schemes

- PURL (Persistent URL) – “URLs are fine if you *redirect* from purl.org”
- URN (Uniform Resource Name), DOI (Digital Object Identifier) & Handle
 - “URLs and domain names are bad, except for ours, and we redirect”
- Tim Berners-Lee – “cool URIs don’t break”
- ARK (Archival Resource Key) – “URLs are fine if managed well, but do tell us which of your URLs are meant for what kind of persistence”



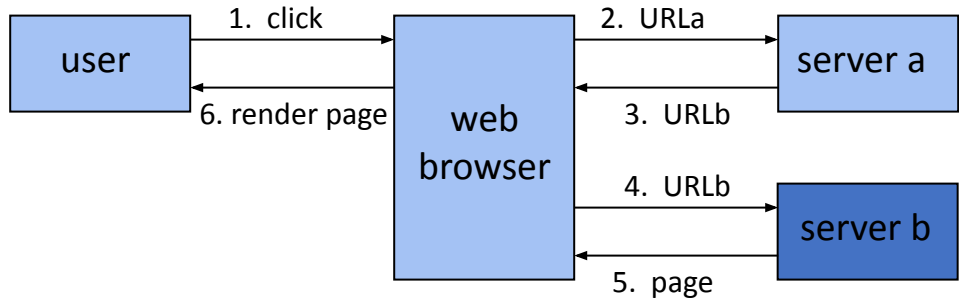
PID schemes – pessimist view

Helps with major causes of broken links?	PURL	Handle	URN	DOI	ARK
Prevents fire, war, flood, attack, bankruptcy, ...	No	No	No	No	No
Prevents human or service provider error	No	No	No	No	No
Guarantees your links, or fixes them for you	No	No	No	No	No
Best practices guard against copy/paste errors	No	No	No	No	Yes
Global resolver downtime less than 1 day per year	No	No	No	No	Yes
Identity independence from lost domain/server name	No	No	Yes	No	Yes

Web access – direct



Web access – indirect



HTTP
redirect

Example: archive.example.org/photo123 → photos.example.org/vault/123

A redirect is like forwarding a (request) message to a new address

PID schemes – optimist view

Features and costs	PURL	Handle	URN	DOI	ARK
Decentralized resolution	No	No	No	No	Yes
Inferenceable syntax (variants, containment)	No	No	No	No	Yes
Flexible metadata by design, including none	No	No	No	No	Yes
Inflections (...?info) and content negotiation	No	No	No	No	Yes
Nuanced persistence statements by design	No	No	No	No	Yes
Path extensions during resolution (suffix passthrough)	Yes	No	Yes?	No	Yes
<i>Free, non-paywalled, in unlimited numbers</i>	Yes	No	Yes	No	Yes

PID schemes – ecosystem view

Identifiers in an Internet context	PURL	Handle	URN	DOI	ARK
Appear in Data Citation Index, HathiTrust, Wikipedia, Wikidata, Internet Archive, ORCID profiles	Yes	Yes	Yes	Yes	Yes
Major adoption by most academic publishers outside the global South	No	No	No	Yes	No
Free (subsidized) account and admin interface for one-off use, e.g., purl.org, zenodo.org, archive.org	Yes?	No?	No?	Yes	Yes?
IETF standard URI, validated by web browsers	No	No	Yes	No	No
Replicated global resolver architecture	No	Yes	No	No	No

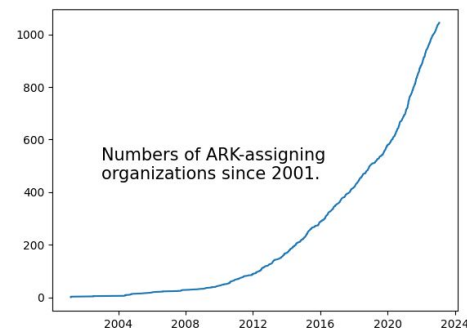
Summary: ARK benefits

ARKs can serve as persistable identifiers with metadata

- found in the Data Citation Index, HathiTrust, Wikipedia, Wikidata, Internet Archive, ORCID profiles, etc.

In contrast to other id schemes, ARKs have

- no fees, no limits, no walled gardens (decentralized)
- very flexible metadata, including none
- can be assigned to anything digital, physical, or conceptual



Smithsonian ARKs: 65665

The Smithsonian Libraries & The Smithsonian Institution

- ARKs for collection metadata & multimedia objects
- Started in 2015
- By 2020 over 15 million ARKs and counting....

“ARKs are a perfect fit for our [Smithsonian] collections”

- Project size
- Cost
- Ease of implementation
- Permanence



Courtesy of the Smithsonian Libraries.
Alexandre, Arsène. Noé dans son arche.
Combet et Cie, 1902.

Smithsonian ARK record and image examples



Scientific specimens from the National Museum of Natural History
<http://n2t.net/ark:/65665/381440f27-3f74-4eb9-ac11-b4d633a7da3d>



Cultural artifacts from the National Museum of American History
<http://n2t.net/ark:/65665/ng49ca746b2-42dc-704b-e053-15f76fa0b4fa>



Sculpture from the Freer Gallery of Art & Arthur M. Sackler Gallery
<http://n2t.net/ark:/65665/ye3080ce305-a705-49cc-a70d-99aff8cb65da>



Photographs from the National Museum of African American History and Culture
<http://n2t.net/ark:/65665/fd5ad97cb86-caaf-4209-8fde-98d70f52f072>

Paintings from the Smithsonian American Art Museum
<http://n2t.net/ark:/65665/vk7a466371d-0413-451f-bd76-ca0becc46f94>



Wrap up

Questions?

