Using Archival Resource Keys (ARKs) for Persistent Identification

Mark E. Phillips
University of North Texas Libraries

Texas Conference on Digital Libraries
June 5, 2008
ARK

Archival Resource Key

Is a naming scheme designed to facilitate the high quality and persistent identification of information objects.

A founding principle of the ARK is that persistence is purely a matter of service and is neither inherent in an object nor conferred on it by a particular naming syntax.
Archival Resource Key (ARK)

An ARK is a URL created to allow persistent, long-term access to information objects.

http://example.org/ark:/13030/654xz321/s3/f8.05v.tiff
\________________/  \_/  \_/  \_____/  \__________/ (replaceable)
| ARK Label |    |    | Qualifier (NMA-supported)
| Name Mapping Authority (NMA) |    |    |
| Name Assigning Authority Number (NAAN) |    |    | Name (NAA-assigned)
Name Mapping Authority (NMA)

http://example.org/

The name mapping authority is responsible for launching an ARK into the world as a url, otherwise it is no different from any other identifier scheme.

Name Mapping authorities can change over time, and probably will.

Access to content at multiple locations is possible.

http://tdl.org/ark:/12025/654xz321
http://digital.library.unt.edu/ark:/12025/654xz321
ARK Label

ark:/

Simple label that indicates that this is in fact an ARK Identifier. Makes searching for these quite easy in most search engines.
Name Assigning Authority Number (NAAN)

13030

Unique identifier for institutions creating ARKs.

Helps in the “globally unique” problem.

Will remain the same throughout the life of the identifier.
Name

654xz321

Minted by institution identified in the NAAN

The more opaque probably the better.

Preferred:

No sequences,
No more than two letters together
Qualifier

s3/f8.05v.tiff

Optional Qualifier for adding service points to identifiers

Do not have to be supported by other NMA's

Qualifiers don't need to be as opaque.
What should a persistent URL deliver?

Representation of the digital object?

Human readable display?

Metadata?

Service agreement?
Beautiful URLs

In many modern web frameworks the concept of "beautiful URLs" is very strong.

We should have this

http://example.com/news/article/131/

Instead of this

http://example.com/news.php?article=131
ARK in use

UNT Libraries is redesigning our digital library systems infrastructure and we are using ARKs as a key piece of the infrastructure.

ARks match up with the concept of “beautiful URLs” which is a common design practice for web applications.

ARKs provide an easy framework for building services.
http://digital.library.unt.edu/ark:/67531/metadc3211

Returns a human readable page like most systems
http://digital.library.unt.edu/ark:/67531/metadc3211?

Returns a machine and human readable metadata
record in the ERC format

erc:

who: Ezzell, Frank
what: The Ferris Wheel, Volume 4, No. 39, Saturday, June 12, 1897
when: 1897
where: http://digital.library.unt.edu/ark:/67531/metadc3211
Returns a service agreement about the ARK letting the user know how this identifier or object might change over time.

erc-support:
who: unt.edu
what: (:tba) Commitment statement pending.
when: 20071219
where: http://digital.library.unt.edu/ark/

Hyphens are ignored so this resolves to the same place.
Services

http://digital.library.unt.edu/ark:/67531/metadc3211/thumbnail

Thumbnail representation of the object, every object has some sort of thumbnail image.

http://digital.library.unt.edu/ark:/67531/metadc3211/square

Square representation of the object, every object has some sort of square image.

http://digital.library.unt.edu/ark:/67531/metadc3211/metadata

Listing of all metadata in our preferred format
All objects have quick access to many metadata schemes.
Objects

http://digital.library.unt.edu/ark:/67531/metadc3211/1/
  First fileSet in the object (probably a page turner)
http://digital.library.unt.edu/ark:/67531/metadc3211/1/thumbnail
  Thumbnail for first fileSet
http://digital.library.unt.edu/ark:/67531/metadc3211/1/square
  Square image for first fileSet
http://digital.library.unt.edu/ark:/67531/metadc3211/1/ocr
  OCR text for first fileSet
http://digital.library.unt.edu/ark:/67531/metadc3211/1/highRes
  Access to our high resolution image
http://digital.library.unt.edu/ark:/67531/metadc3211/1/medRes
  Access to the medium resolution image
Manifestations

Texas Reports:

One digital object
PDF version
HTML version
TXT version

http://digital.library.unt.edu/ark:/67531/metadc3211/m1/

PDF version
http://digital.library.unt.edu/ark:/67531/metadc3211/m2/

TXT version
http://digital.library.unt.edu/ark:/67531/metadc3211/m3/

HTML version
Books, Newspapers, Maps

http://digital.library.unt.edu/ark:/67531/metadc3211/1/zoom
  Zooming interface for maps and newspapers

http://digital.library.unt.edu/ark:/67531/metadc3211/1/zoom?q=Texas
  Highlights the word Texas on the page

http://digital.library.unt.edu/ark:/67531/metadc3211/1/boundingBoxes
  Bounding box file mapping words to locations on a page

http://digital.library.unt.edu/ark:/67531/metadc3211/1/bboxes?q=Texas
  JSON containing bounding box information for the location of the word Texas

Example: http://libbrandonbox.library.unt.edu/ark:/67531/metapth18837/1/zoom?q=cleaning
Rapid Development

If a designer knows just the ARK they have all of these services at their disposal.

ARK provides a framework to hang additional services as needed.