What can linked data do for archives?

Gloria Gonzalez, Elizabeth Russey Roke, & Mark A. Matienzo
What is Linked Data?
STRUCTURED, MACHINE
ACTIONABLE DATA
with an open license
for sharing & reuse
What's a triple?

object

predicate

subject
What's a triple?

Susan Sontag → created → email
What’s a triple?

http://link.library.ucla.edu/resource/rde_eBRgadA/

Susan Sontag

http://bibfra.me/vocab/lite/creator

is the creator of

http://link.library.ucla.edu/resource/l4lhHaSqGGE/

email
Read by humans

Read by machines

#### Restrictions on Access

Series contains files that have been restricted until December 2044.

#### Processing Information

The original structure, naming, and contents of the material have been maintained. Sontag migrated files from one computer to the next; accordingly, duplication and variation in files are common. Files and folders with the same or similar titles may not have exactly the same contents; it is also not uncommon for segments of text to be repeated across files of different names, as many of the file instances are drafts and often build upon one another.

#### Scope and Content

This series represents the contents of three of Sontag's computer hard drives, which were acquired as a part of the 2012 addition to the Sontag archive. These materials touch upon several of the previous series created for the collection as a whole, specifically: correspondence, works by Susan Sontag, public activity, political activity, Bosnia involvement, PEN involvement, material about Susan Sontag, personal materials, and photographs.

Included are Sontag's email correspondence, both personal and professional in nature; photographic material related to family, public activity, travel, film productions, and Sarajevo; and personal material. The personal material includes agendas, calendars and itineraries; notes; financial, medical and travel documents; lists of vocabulary, favorite works, and films; lists of hotels, restaurants, museums, and other addresses; arranged geographically; music files; and reproductions of various genres of artwork. Also included are speeches written for and notes to Annie Liebowitz, and biographies and personal statements written for Sontag's son, David Rieff.

Material related to Sontag's works includes notes, research, and finished and unfinished drafts for Sontag's works of fiction and non-fiction, essays, op-ed pieces, plays,
[2 hard drives containing Susan Sontag's email correspondence: purchased from her son David Rieff through his representative, The Wylie Agency]@en;
schema:name "[2 hard drives containing Susan Sontag's email correspondence: purchased from her son David Rieff through his representative, The Wylie Agency]"@en.

Sontag, Susan"@en;
bf:attribute "UCLA Library."@en;
cc:license <https://creativecommons.org/licenses/by/4.0/>;
og:title "Sontag, Susan (Person) – UCLA Library"@en;
dcterms:title "Sontag, Susan (Person) – UCLA Library"@en;
schema:name "Sontag, Susan"@en.
Archivists can create local connections with Linked Data
Local Connections

http://link.library.ucla.edu/portal/Oral-history-interview-with-Larry-Margolis--.../Y99pmK32Eb0/

BF:Item,
\[\text{rdfs:label} \text{"Oral history interview with Larry Margolis ...}, 1989"@en, \]
\[\text{"Oral history interview with Larry Margolis ...}, \text{by Carole Hicke"}@en;\]
\[\text{bf:contributor} \text{http://link.library.ucla.edu/resource/F5C2jrRjG-I/}, \]
\[\text{http://link.library.ucla.edu/resource/FTDmMcxD1vs/}, \]
\[\text{http://link.library.ucla.edu/resource/sxDkIEFw_Rc/}; \]
\[\text{bf:creator} \text{http://link.library.ucla.edu/resource/U_qeqFp0wiQ/};\]
\[\text{schema:name} \text{"Oral history interview with Larry Margolis ..."@en}, \]
\[\text{"Oral history interview with Larry Margolis ...}, 1989"@en, \]
\[\text{"Oral history interview with Larry Margolis ...}, \text{by Carole Hicke"}@en;\]
\[\text{schema:potentialAction} \text{[a schema:BorrowAction;}, \]
\[\text{schema:lender} \text{http://link.library.ucla.edu/#ucla}; \]
\[\text{schema:target} \text{https://catalog.library.ucla.edu/vwebv/holdingsInfo?bibId=4267159}, \]
\[\text{[a schema:RegisterAction;}, \]
\[\text{schema:object} \text{[a schema:LibrarySystem;}, \]
\[\text{schema:id} \text{http://link.library.ucla.edu/}; \]
\[\text{schema:name} \text{"UCLA Library"@en}; \]
\[\text{schema:url} \text{https://www.library.ucla.edu/}]; \]
\[\text{schema:target} \text{[a schema:EntryPoint;}, \]
\[\text{schema:actionPlatform} \text{schema:AndroidPlatform,}, \]
\[\text{schema:DesktopPlatform,}, \]
\[\text{schema:IOSPlatform,}, \]
\[\text{schema:WindowsPlatform}; \]
\[\text{schema:urlTemplate} \]
\[\text{https://speccoll.library.ucla.edu/logon?_ga=2.143646645.1412163549.1507837147-867728627.1505506163}; \]

\[\text{Data from Civil rights -- California -- UCLA Library}\]
Archivists can connect with communities and across networks with linked data.
Network Connections

Civil rights -- California

Focus: Civil rights
Sub Focus: California

Shared in Network

Currently shared across 45 libraries

- http://link.afplis.org/resource/xgkGnlqwq9Q/
- http://link.library.ucla.edu/resource/xgkGnlqwq9Q/
- http://uofme.library.link/resource/xgkGnlqwq9Q/
- http://link.lib.umanitoba.ca/resource/xgkGnlqwq9Q/
- http://link.sandiego.edu/resource/xgkGnlqwq9Q/
Dallas Public Library’s summer 2018 partnership with KERA public TV

http://learn.kera.org/greatamericanread/
Archivists can define archives as places with linked data.
## Archives as places

### Statement Collection #17
**Entity**
**Attributes**
- rdf:type
- schema:url
- Name
- http://bibfra.me/vocab/lite
- schema:memberOf
- http://library.link/vocab/bib
- schema:address
- schema:location
- schema:telephone
- schema:openingHours
- schema:openingHours
- schema:openingHours

### Statement Collection #18
**Attributes**
- rdf:type
- schema:streetAddress
- schema:addressLocality
- schema:addressRegion
- schema:postalCode
- schema:addressCountry

### Statement Collection #19
**Attributes**
- rdf:type
- schema:geo

### Statement Collection #20
**Attributes**
- rdf:type
- schema:latitude
- schema:longitude

---

```json
{  "type": "Library",  "address": {"streetAddress": "A1713 Charles E. Young Research Library",  "addressLocality": "Los Angeles",  "addressRegion": "CA",  "postalCode": "90095-1575",  "addressCountry": "US"},  "geo": {"type": "Point",  "coordinates": ["34.0749691", "-118.441466"]}
```
Schema.org

- Created in 2011 by Bing, Google, and Yahoo to address structured data format proliferation for search engines
- Provides single schema across range of topics: people, creative works, places, etc. (598 types, 862 properties)
- Used on 1+ billion web pages & many popular websites
- Provides extension mechanism for both Schema.org-hosted and external extensions
- Expressible as JSON-LD, RDFa, and Microdata
About
From "Democratic Motherism" in The world Will Follow Joy: Turning Madness into …

My Books
To learn more about each book, just click on the cover. Alice Walker encourages you to to …

New Books
Wonders Never Ceasing – Queen Sugar and The Star of the Sea; Children in Cages: …

See results only from alicewalkersgarden.com

Spring Poem
We Can Know the World is Changing When the African Woman Freely Speaks Her Mind, …

Book List
Pema Chodron and Alice Walker in Conversation (video) Collected Interviews …

Democratic Womanism
We Can Know the World is Changing When the African Woman Freely Speaks Her Mind, …

Alice Walker - Wikipedia
https://en.wikipedia.org/wiki/Alice_Walker

Children: Rebecca Walker
Occupation: Novelist, short story writer, poet, …
Notable works: The Color Purple

Overview Contents Early life Writing career Activism and political criti

Alice Walker is an American novelist, short story writer, poet, and activist. She wrote the novel The Color Purple, for which she won the National Book Award for hardcover fiction, and the Pulitzer Prize for Fiction. She also wrote the novels Meridian and The Third Life of Grange Copeland, among other works. An avowed feminist, Walker coined the term "womanist" to mean "A black feminist or feminist of color" in 1983.

See more on en.wikipedia.org · Text under CC-BY-SA license

Alice Walker - Author, Activist, Women's Rights Activist ...
https://www.biography.com/people/alice-walker-9521939

Watch video · Alice Walker is an African-American, Pulitzer Prize-winning novelist known for The Color Purple. Learn more at Biography.com.
# Book
Canonical URL: [http://schema.org/Book](http://schema.org/Book)

**CreativeWork > Book**

A book.

Usage: Between 10,000 and 50,000 domains

<table>
<thead>
<tr>
<th>Property</th>
<th>Expected Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Properties from Book</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abridged</td>
<td>Boolean</td>
<td>Indicates whether the book is an abridged edition.</td>
</tr>
<tr>
<td>bookFormat</td>
<td>BookFormatType</td>
<td>The format of the book.</td>
</tr>
<tr>
<td>illustrator</td>
<td>Person</td>
<td>The illustrator of the book.</td>
</tr>
<tr>
<td>numberOfPages</td>
<td>Integer</td>
<td>The number of pages in the book.</td>
</tr>
<tr>
<td><strong>Properties from CreativeWork</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>about</td>
<td>Thing</td>
<td>The subject matter of the content.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inverse property: subjectOf.</td>
</tr>
<tr>
<td>accessMode</td>
<td>Text</td>
<td>The human sensory perceptual system or cognitive faculty through which a person may process or perceive information. Expected values include: auditory, tactile, textual, visual, colorDependent, chartOnVisual, chemOnVisual, diagramOnVisual, mathOnVisual, musicOnVisual, textOnVisual.</td>
</tr>
<tr>
<td>accessModeSufficient</td>
<td>Text</td>
<td>A list of single or combined accessModes that are sufficient to understand all the intellectual content of a resource. Expected values include: auditory, tactile, textual, visual.</td>
</tr>
<tr>
<td>accessibilityAPI</td>
<td>Text</td>
<td>Indicates that the resource is compatible with the referenced accessibility API (<a href="https://www.w3.org/2011/06/webvitals/#accessibilityapi">WebSchemas wiki lists possible values</a>).</td>
</tr>
<tr>
<td>accessibilityControl</td>
<td>Text</td>
<td>Identifies input methods that are sufficient to fully control the described resource (<a href="https://www.w3.org/2011/06/webvitals/#accessibilitycontrol">WebSchemas wiki lists possible values</a>).</td>
</tr>
<tr>
<td>accessibilityFeature</td>
<td>Text</td>
<td>Content features of the resource, such as accessible media, alternatives and supported enhancements for accessibility (<a href="https://www.w3.org/2011/06/webvitals/#accessibilityfeature">WebSchemas wiki lists possible values</a>).</td>
</tr>
<tr>
<td>accessibilityHazard</td>
<td>Text</td>
<td>A characteristic of the described resource that is physiologically dangerous to some users. Related to WCAG 2.0 guideline 2.3 (<a href="https://www.w3.org/2011/06/webvitals/#accessibilityhazard">WebSchemas wiki lists possible values</a>).</td>
</tr>
<tr>
<td>accessibilityRelatedLink</td>
<td>Text</td>
<td>A human-readable summary of specific accessibility features or guidelines (<a href="https://www.w3.org/2011/06/webvitals/#accessibilityrelatedlink">WebSchemas wiki lists possible values</a>).</td>
</tr>
</tbody>
</table>
Example RDFa Schema.org Markup: Book

...<div property="breadcrumb">
  <a href="/category/books.html">Books</a> >
  <a href="/category/books-literature.html">Literature & Fiction</a> >
  <a href="/category/books-classics">Classics</a>
</div>

<div property="mainEntity" typeof="Book">
  <img property="image" src="/catcher-in-the-rye-book-cover.jpg"
       alt="cover art: red horse, city in background"/>
  <span property="name">The Catcher in the Rye</span> -
  by <a property="author" href="/author/jd_salinger.html">J.D. Salinger</a>
  <div property="aggregateRating" typeof="AggregateRating">
    <span property="ratingValue">4</span> stars -
    <span property="reviewCount">3077</span> reviews
  </div>
  <div property="offers" typeof="Offer">
    Price: $<span property="price">6.99</span>
    <meta property="priceCurrency" content="USD"/>
    <link property="availability" href="/http://schema.org/InStock">In Stock
  </div>
<br>
Product details
  <span property="numberOfPages">224</span> pages
  Publisher: <span property="publisher">Little, Brown, and Company</span> -
  <meta property="datePublished" content="1991-05-01">May 1, 1991
</div>
Schema for Finding Aids?

Descriptive Summary

Creator: Carson, Ciaran, 1948-
Title: Ciaran Carson papers, circa 1970-2010
Call Number: Manuscript Collection No. 746
Extent: 34.5 linear feet (67 boxes) and 2 oversized papers boxes and 5 oversized papers folders (OP)
Abstract: Personal and literary papers of Irish poet Ciaran Carson including correspondence, literary notebooks, literary manuscripts, and collected printed material.
Language: Materials in English with some items in Irish.

Administrative Information

Restrictions on Access
Collection stored off-site. Researchers must contact the Rose Library in advance to access this collection.

Terms Governing Use and Reproduction
All requests subject to limitations noted in departmental policies on reproduction. No special restrictions apply.
<rdf:Description rdf:about="http://pid.emory.edu/ark:/25593/8z3n4#collection">
  <schema:keywords xml:lang="en">Robinson, Tim, 1935-</schema:keywords>
  <schema:keywords xml:lang="en">Poets, Irish--20th century.</schema:keywords>
  <schema:relatedLink rdf:resource="http://pid.emory.edu/ark:/25593/8zkmf"/>
  <schema:about rdf:resource="http://viaf.org/viaf/120096893"/>
  <schema:name xml:lang="en">Ciaran Carson papers, circa 1970-2010</schema:name>
  <schema:keywords xml:lang="en">Gallagher, Tess.</schema:keywords>
  <schema:about rdf:resource="http://viaf.org/viaf/112404721"/>
  <schema:keywords xml:lang="en">Songs, Irish.</schema:keywords>
  <schema:about rdf:resource="http://viaf.org/viaf/49280158"/>
  <schema:about rdf:resource="http://viaf.org/viaf/113957087"/>
  <schema:keywords xml:lang="en">Ormsby, Frank, 1947-</schema:keywords>
  <schema:keywords xml:lang="en">Northern Ireland.</schema:keywords>
  <schema:keywords xml:lang="en">Poems.</schema:keywords>
  <schema:keywords xml:lang="en">English literature--Irish authors--20th century.</schema:keywords>
  <schema:keywords xml:lang="en">Poet.</schema:keywords>
  <schema:keywords xml:lang="en">Montague, John.</schema:keywords>
  <schema:about rdf:resource="http://viaf.org/viaf/24613408"/>
  <schema:keywords xml:lang="en">Arts, Irish.</schema:keywords>
  <schema:keywords xml:lang="en">Irish poetry--20th century.</schema:keywords>
  <schema:keywords xml:lang="en">Carson, Ciaran, 1948-</schema:keywords>
  <schema:about rdf:resource="http://viaf.org/viaf/129928623"/>
  <schema:description xml:lang="en">Personal and literary papers of Irish poet Ciaran Carson including correspondence, literary notebooks, literary manuscripts, and collected printed material.</schema:description>
  <rdf:type rdf:resource="http://purl.org/archival/vocab/arch#Collection"/>
  <schema:about rdf:resource="http://sws.geonames.org/2641364"/>
  <schema:about rdf:resource="http://viaf.org/viaf/109557338"/>
  <schema:about rdf:resource="http://viaf.org/viaf/79088985"/>
  <schema:keywords xml:lang="en">Heaney, Seamus, 1939-</schema:keywords>
  <schema:about rdf:resource="http://sws.geonames.org/2655984"/>
  <rdf:type rdf:resource="http://schema.org/CreativeWork"/>
  <schema:keywords xml:lang="en">Proulx, Annie.</schema:keywords>
</rdf:Description>
Example RDFa Schema.org Markup: Finding Aid

Data from: Wikipedia · IMDb · Goodreads · Famousbirthdays · Amazon · Freebase

Text under CC-BY-SA license

Suggest an edit
Archives and Linked Data Interest Group
Goals

- Evaluate existing models
- Investigate minimally viable ways to publish linked data about archives
- Development of mappings from archival description standards to linked data models
- Production of archival description as linked data directly from archives management systems
Schema Architypes

Archive

Woodson Research Center

By appointment

name

accessConditions

holdingArchive

location

archiveHeld

Archive that holds a Collection

Collection (bib.schema.org)

0.5

LF

value

unitCode

materialExtent

accessConditions

Stored offsite.

QuantitativeValue

ArchiveComponent

Collection in an Archive

hasPart

isPartOf

Series II: General

CreativeWork

A component (e.g., a series) of type CreativeWork in an archive

Existing Scheme.org type/property

Proposed Schema Architypes type/property
# Mapping description to Schema.org (1)

| DACS Element | | IAD/G Element | | Schema.org | | AS element | | AtoM element (to DACS) | | Notes |
|--------------|---|----------------|-------------|---------------|-------------|--------------------------|-----------------------------|-----------------------------|
| 2.1 Reference code | 3.1.1 Reference code(s) | 3.1.1 Identifier | R | schema/identifier | identifier | identifier | ok without this; URI should suffice; call number for books maps to schema/sku (?) |
| 2.2 Name and location of repository | R | | | | schema/holdingArchive | Repository name; Telephone number, email (should be linked to the repository record) | Name and location of repository (link to IBDIAH record) |
| 2.3 Title | 3.1.2 Title | | R | schema/name | Title | Title | |
| 2.4 Date | 3.1.3 Date(s) | | R | schema/dateCreated | Date/Expression / Start / End, when Label = "Broadcast/Created/Publication" and Type = "Inclusive/Single" | Date / Start / End / Type | will need to use date ranges within property |
| 2.5 Extent | 3.1.5 Extent and medium of the unit of description | | R | schema/materialExtent; schema/unitTextExtent; schema/unitValue | Extent statement (combination of 6 data elements: portion, extent number*, extent type*, container summary, physical details, dimensions) | Extent | see schema/isPartOf, schema/hasPart |
| 2.6 Name of creator(s) | 3.2.1 Name of creator(s) | O | schema/creator | Name of creator(s) | Name of creator(s) | |
| 2.7 Administrative/Biographical history | 3.2.2 Administrative/Biographical history | O | schema/Person/Organization: schema/description | Agent bio note (can be attached to collection or agent record, but my preference might be to ONLY to include it in schema.org when it's attached to the agent record) | (history is pulled in on save from the related authority record) | |
| 3.1 Scope and content | 3.3.1 Scope and content | O | schema/description | Scope and content note (like all notes in Advance, this can be repeated); abstract, note if no scope and content note | Scope and content | |
| 3.2 System of arrangement | 3.3.4 System of arrangement | O | schema/arrangement | Arrangement note | System of arrangement | not appropriate for search engine results | |
| 4.1 Conditions governing access | 3.4.1 Conditions of access | O | schema/AccessAndUse | Conditions governing access note | Conditions governing access | |
| 4.2 Physical access | 3.4.4 Physical characteristics and technical requirements | O | | Physical characteristics and technical requirements note | Physical access | ignore; not necessary for schema | |
| 4.3 Technical access | 3.4.4 Physical characteristics and technical requirements | O | | Technical access | ignore; not necessary for schema | | |
Mapping description to Schema.org (2)

Elements of DACS and ISAD(G) mapped to both existing Schema.org properties/types and proposed properties/types from the Architypes Proposal.
This collection contains items documenting the Houston Folk music scene from the 1970s-1980s.

During the 1960s-1980s, Houston, TX had a vibrant folk scene. Local musicians and those from other parts of Texas and the U.S. socialized and played at a variety of music establishments around the city, including Anderson Fair Retail Restaurant, Sand Mountain Coffeehouse, Liberty Hall, The Old Quarter, Theodore's and Corky's. This scene spawned many musicians, Townes Van Zandt, Guy Clark, Robert Earl Keen, Nanci Griffith, Lynn Langham, Wrecks Bell, Eric Taylor, John Vandiver, Danny Everitt, Vince Bell, Richard Dobson, Don Sanders, Wheatfield/St. Elmo's Fire, Dogtooth Violet, Bill Cade, Lyle Lovett, and many more.

This material is open for research.

Permission to publish from this material must be obtained from the specific copyright owner.

Albums donated by Brendan Doss, III. Townes Van Zandt reel donated by Craig Calvert.
What's Next?
What can we do with this?
Thank You!

Resources

- Archives and Linked Data Interest Group: https://archival.github.io/
- W3C Schema Architypes Community Group: https://www.w3.org/community/architypes/

Acknowledgements

The Archives and Linked Data Interest Group:

- Scott Carlson (Rice University)
- Mark Custer (Yale University)
- Patrick Galligan (Rockefeller Archives Center)
- Dan Gillean (Artefactual Systems)
- Gloria Gonzalez (Zepheira)
- Maggie Hughes (UCLA)
- Mark Matienzo (Stanford University)
- Dave Mayo (Harvard University)
- Laney McGlohon (LYRASIS/ArchivesSpace)
- Evelyn McLellan (Artefactual Systems)
- Katy Rawdon (Temple University)
- Elizabeth Russey Roke (Emory University)
- Ruth Kitchin Tillman (Penn State)