

**Build Tomorrow's Library**

---

# <The Pitch>

Get your hands on the full bibliographic data about virtually all of the 12.7 million items in the Harvard Library collection, plus millions of images, plus some extremely anonymized usage data, and more. All of this can be put to use in the apps you'll learn to write that use the [API of Harvard Library Cloud, a new open metadata server](#). Library metadata is rich, complex, and very human, so let's dig in!

# The Harvard Library has a lot of stuff

- ~13 million bibliographic records
- ~4 million images cataloged
- ~5,000 finding aids for archival and document collections (representing ~1.8 million individual documents), including links to online content
- A usage metric

# All these items have metadata

- Metadata is data about the data
- Cataloging information - call numbers, subjects, holdings, formats, author, etc.
- Some objects are available online (e.g. full text, scanned images), but most are not
- Metadata is useful (just ask the NSA)

**Title:** The annotated Alice : Alice's adventures in Wonderland & Through the looking glass

**Title:** Alice's adventures in Wonderland

**Attribution:** by Lewis Carroll ; with illustrations by John Tenniel ; updated, with an introduction and notes by Martin Gardner.

**Author:** Carroll, Lewis, 1832-1898.

**Variant title:** Annotated Alice--the definitive edition

**Summary:** Forty years after Gardner's groundbreaking publication of the annotated version of Carroll's most famous work comes this new version, featuring fascinating insights, notes and newly discovered line drawings.

**Subject:** Carroll, Lewis, -- 1832-1898. -- Alice's adventures in Wonderland.;  
Carroll, Lewis, -- 1832-1898. -- Through the looking-glass.;  
Alice -- (Fictitious character from Carroll) -- Fiction.;  
Alice -- (Fictitious character from Carroll);  
Fantasy fiction, English.;  
Fantasy fiction, English -- History and criticism.;  
Alice (Fictitious character : Carroll) -- Fiction.;  
Alice (Fictitious character : Carroll)

**Other subject:** *Form/Genre: Dust jackets (Bindings) 2000.*;

**Published:** New York : Norton, c2000.

**Edition:** Definitive ed.

**Description:** xxviii, 312 p. : ill. ; 26 cm.

**Language:** English

**Author:** Gardner, Martin, 1914-;  
Tenniel, John, Sir, 1820-1914.;  
Carroll, Lewis, 1832-1898. Through the looking-glass.;  
Tenniel, John, 1820-1914.;  
Gardner, Martin, 1914-2010.

**Notes:** This edition combines the notes of Gardner's 1960, The annotated Alice with his 1990 update, More annotated Alice, as well as additional discoveries and updates drawn from Gardner's encyclopedic knowledge of the texts. It includes Tenniel's classic art with some recently discovered Tenniel pencil sketches.  
Includes bibliographical references.

**Contents:** Preface to the definitive edition of The Annotated Alice -- Introduction to The Annotated Alice -- Introduction to More Annotated Alice -- Alice's Adventures in Wonderland -- Through the Looking-Glass and What Alice Found There -- The Wasp in a Wig -- Original pencil sketches by Tenniel -- A note about Lewis Carroll societies -- Selected references -- Alice on the screen / by David Schaefer.

**Other title:** Through the looking-glass.;

**Identifier:** ISBN 0393048470; ISBN 9780393048476

**HOLLIS Number:** 008330208

**Creation Date:** c2000

**Source:** HVD ALEPH

This is library metadata

Components: 1 surrogate (with 1 image)

## Work

**Title:** Expanded Alice.  
**Classification:** 977 T 257 4 A1 (Harvard Fine Arts Library, Visual Collections - Slides and Digital Images)  
**Work Type:** print  
**Creator:** **Tenniel**, Sir John (1820-1914), England, artist  
**Date:** n.d.

— Surrogate 1 of 1

Save this... ↕



Save this... ↕

**Classification:** 977 T 257 4 A1

**Work Type:** Slide

**Repository:** Harvard Fine Arts Library, Visual Collections - Slides and Digital Images  
1997.17420

**Record Identifier:** olvwork43816

So is this

Authors:	Amory, Harcourt, 1855-1925, collector.
Title:	Harcourt Amory collection of Lewis Carroll, 1843-1915. <b>&lt;&lt; More about this collection &gt;&gt;</b>
Finding aids:	Electronic finding aid available <a href="http://nrs.harvard.edu/urn-3:FHCL.Hough:h000171">http://nrs.harvard.edu/urn-3:FHCL.Hough:h000171</a>
Description:	3 volumes and 6 boxes (3 linear ft.)
Description:	Organized into the following series: MS Eng 718: Dodgson, The rectory umbrella; MS Eng 718.1: Dodgson, Mischmasch; bMS Eng 718.2: Dodgson, Hiawatha's photographing; bMS Eng 718.3: Dodgson,.; The only sister who would write to her brother; bMS Eng 718.4: Dodgson, Humpty Dumpty and Tweedledum; MS Eng 718.5: Dodgson, A series of autograph letters; bMS Eng 718.6: Tenniel, Studies for illustrations to Alice's adventures in wonderland; bMS Eng 718.7: Tenniel, Studies for illustrations to Through the looking-glass; bMS Eng 718.8: Tenniel, Drawings published in Punch; bMS Eng 718.9: Frost, Drawings for Rhyme? and reason? and A tangled tale; bMS Eng 718.10: Dodgson, Letters from Dodgson; bMS Eng 718.11: Dodgson, Compositions; bMS Eng 718.12: Dodgson, Letters to Dodgson; bMS Eng 718.13: Furniss, Illustrations to Sylvie and Bruno; bMS Eng 718.14: Photographs; bMS Eng 718.15: Amory, Papers concerning his collection of Lewis Carroll material; and bMS Eng 718.16: Harcourt Amory bookplate.
History notes:	Dodgson was an English author and mathematician, best known as Lewis Carroll. Sir John Tenniel was an English illustrator and satirical artist especially known for his work in Punch and his illustrations for the Lewis Carroll Alice books. Harcourt Amory was a collector of Lewis Carroll books and manuscripts.
Summary:	Contains manuscripts of Dodgson's prose and poetry with his illustrations including: drawings in pencil and pen and ink, including his designs of Humpty Dumpty and Tweedledum for Tenniel, and a sketchbook with drawings of girls. Includes letters by Dodgson on both personal and professional matters, including letters, 1844-1852, that he wrote to his sisters from school. Also includes: Tenniel's studies for his illustrations to Alice's Adventures in Wonderland and Through the Looking-Glass; drawings by others to illustrate Dodgson's work; portrait photographs of Dodgson and others, some taken by Dodgson; and notes, lists, and correspondence about Dodgson by Harcourt Amory.
Summary:	Includes: manuscripts of his childhood domestic magazines, The rectory umbrella and Mischmasch; drawings by Tenniel for Punch; drawings by Arthur Burdett Frost for Rhyme? and reason? and A tangled tale, both by Carroll; prints by Harry Furniss of Sylvie and Bruno; papers from Harcourt Amory concerning the collecting of Lewis Carroll materials; and designs and notes about the drawing by Sidney Lawton Smith of the Harcourt Amory bookplate.
Other forms:	Selected color digital images available; see finding aid.
Cite as:	Harcourt Amory Collection of Lewis Carroll (MS Eng 718-718.16). Houghton Library, Harvard University.
Subject:	Drawing, English -- 19th century. English literature -- 19th century. English poetry -- 19th century. Illustration of books -- 19th century. Amory, Harcourt, 1855-1925. Carroll, Lewis, 1832-1898. Tenniel, John, Sir, 1820-1914.
Form/Genre:	Drawings -- Great Britain -- 19th century. Bookplates

and this

# The metadata is locked up

---

All these metadata records are searchable through Harvard Library systems...by humans

How can we make it available for computers?



# Why open the metadata?

- Custom online catalog viewer (e.g., Stacklife)
- Recommendation engine
- Specialized search engine
- Library analytics
- Integrate with Web sources (e.g. NPR, Wikipedia)

and especially, things we haven't thought of yet

# How can we make the data available?

## Complications

- Multiple systems, multiple formats
- Copyright restrictions
- Difficult for end users to manage large data sets

## Approach

- A single platform aggregates and normalizes data from multiple sources
- A public API provides access

# The LibraryCloud Item API

<http://api.lib.harvard.edu/v2/items>

Get all the items

# URL, deconstructed

<http://api.lib.harvard.edu/v2/items>

v2 - Version number. If there are breaking changes, we increment the version

items - The type of thing that's available

# Searching for donuts

<http://api.lib.harvard.edu/v2/items?q=donuts>

# Response formats

## Schemas

MODS (Complex, default) <http://api.lib.harvard.edu/v2/items?q=donuts>

Dublin Core (Simple) <http://api.lib.harvard.edu/v2/items.dc?q=donuts>

## Data Representations

JSON <http://api.lib.harvard.edu/v2/items.json?q=donuts>

XML (default) <http://api.lib.harvard.edu/v2/items.xml?q=donuts>

# Dublin Core + JSON (.dc.json)

```
"item": {
  "title": [
    "Donuts: an American passion",
    "American passion"
  ],
  "contributor": "Edge, John T.",
  "type": "Text",
  "publisher": "G. P. Putnam's Sons",
  "language": "eng",
  "format": [
    "print",
    "176 p. : ill. ; 20 cm."
  ],
  "description": [
    "\nIn Donuts, Edge walks us though the donut's inception as Dutch fare, the Salvation Army's wartime donuts, the invention of the donut machine, the 1950s donut-shop craze, the Krispy Kreme revolution, the appropriation by other ethnicities, and the fanatical chefs that take donuts to a new art form. Nothing encourages our sweet-tooth cravings like the donut. It is honest. It is satisfying. It is a national symbol that has survived the low carb-diet dogma and the death of the local donut shop, and it is making a comeback into the hearts of Americans.\n" -- publisher website (October 2006).",
    "Becoming american -- Of eels and donuts and New England -- Beyond lent -- California dalliance -- New traditionalists -- The future.",
    "John T. Edge."
  ],
  "subject": [
    "Doughnuts",
    "TX770.D67 E44 2006",
    "641.8/653"
  ],
  "identifier": [
    "isbn: 0399153586",
    "isbn: 9780399153587",
    "lccn: 2005058682",
    "010010858"
  ]
}
```

# Dublin Core + XML (.dc.xml)

```
<item>
  <dc:title>Donuts: an American passion</dc:title>
  <dc:title>American passion</dc:title>
  <dc:contributor>Edge, John T.</dc:contributor>
  <dc:type>Text</dc:type>
  <dc:publisher>G. P. Putnam's Sons</dc:publisher>
  <dc:language>eng</dc:language>
  <dc:format>print</dc:format>
  <dc:format>176 p. : ill. ; 20 cm.</dc:format>
  <dc:description>"In Donuts, Edge walks us through the donut's inception as Dutch fare, the Salvation Army's wartime donuts, the invention of the donut machine, the 1950s donut-shop craze, the Krispy Kreme revolution, the appropriation by other ethnicities, and the fanatical chefs that take donuts to a new art form. Nothing encourages our sweet-tooth cravings like the donut. It is honest. It is satisfying. It is a national symbol that has survived the low carb-diet dogma and the death of the local donut shop, and it is making a comeback into the hearts of Americans." -- publisher website (October 2006).</dc:description>
  <dc:description>Becoming american -- Of eels and donuts and New England -- Beyond lent -- California dalliance -- New traditionalists -- The future.</dc:description>
  <dc:description>John T. Edge.</dc:description>
  <dc:subject>Doughnuts</dc:subject>
  <dc:subject>TX770.D67 E44 2006</dc:subject>
  <dc:subject>641.8/653</dc:subject>
  <dc:identifier>isbn: 0399153586</dc:identifier>
  <dc:identifier>isbn: 9780399153587</dc:identifier>
  <dc:identifier>lccn: 2005058682</dc:identifier>
  <dc:identifier>010010858</dc:identifier>
</item>
```



# MODS + JSON (.json)

```
"items": {  
  "mods": {  
    "version": 3.4,  
    "titleInfo": [  
      {  
        "title": "Donuts",  
        "subTitle": "an American passion"  
      },  
      {  
        "type": "alternative",  
        "title": "American passion"  
      }  
    ],  
    "name": {  
      "type": "personal",  
      "namePart": "Edge, John T."  
    },  
    "typeOfResource": "text",  
    "originInfo": {  
      "place": [  
        {  
          "placeTerm": {  
            "authority": "marccountry",  
            "type": "code",  
            "$": "nyu"  
          }  
        }  
      ],  
    }  
  }  
}
```

# MODS + XML (.xml)

```
<item>
  <mods:mods version="3.4">
    <mods:titleInfo>
      <mods:title>Donuts</mods:title>
      <mods:subTitle>an American passion</mods:subTitle>
    </mods:titleInfo>
    <mods:titleInfo type="alternative">
      <mods:title>American passion</mods:title>
    </mods:titleInfo>
    <mods:name type="personal">
      <mods:namePart>Edge, John T.</mods:namePart>
    </mods:name>
    <mods:typeOfResource>text</mods:typeOfResource>
    <mods:originInfo>
      <mods:place>
        <mods:placeTerm authority="marccountry" type="code">nyu</mods:placeTerm>
      </mods:place>
      <mods:place>
        <mods:placeTerm type="text">New York</mods:placeTerm>
      </mods:place>
      <mods:publisher>G. P. Putnam's Sons</mods:publisher>
      <mods:datelssued>c2006</mods:datelssued>
      <mods:datelssued encoding="marc">2006</mods:datelssued>
      <mods:datelssued encoding="marc">2006</mods:datelssued>
      <mods:issuance>monographic</mods:issuance>
    </mods:originInfo>
    <mods:language>
      <mods:languageTerm authority="iso639-2b" type="code">eng</mods:languageTerm>
```

# Searching a specific field

<http://api.lib.harvard.edu/v2/items?title=donuts>

# Paging

## Page 1 (20 items)

<http://api.lib.harvard.edu/v2/items?title=donuts&limit=20>

## Page 2

<http://api.lib.harvard.edu/v2/items?title=donuts&limit=20&start=20>

# Faceting

<http://api.lib.harvard.edu/v2/items.json?q=donuts&facets=subject,languageCode>

The screenshot shows the Amazon search interface for the query "donuts". At the top, the Amazon Prime logo is on the left, and a search bar contains "donuts". Below the search bar, there are navigation options: "Departments" (with a dropdown arrow), "Fire & Kindle" (with a dropdown arrow), and "For Jeffrey". A secondary navigation bar includes "Books", "Advanced Search", "New Releases", "Best Sellers", and "Th".

The main content area displays "1-12 of 1,831 results for Books: 'donuts'". Under the heading "Show results for", there is a list of facets for "Books":

- Any Category
- Breakfast (12)
- Pastry Baking (21)
- Cake Baking (10)
- Confectionary Desserts (11)
- Baking (48)
- Cookbooks, Food & Wine (283)
- Bread Baking (5)
- Travel Dining Reference (12)
- Desserts (57)
- Gastronomy History (5)
- Mystery, Thriller & Suspense (95)
- Children's Books (177)

A "+ See more" link is located at the bottom of the list. To the right of the facets, there are "Related Searches: baked donuts" and "Book Format: Hardcover | Kin". Below these are two book covers: "Donuts" by The Donut Shop and "Hand-Ferred Doughnuts" by The Donut Shop.

# Example

Query the API and display some results

<http://localhost/~jlicht/cs50.html>

Another example

<http://hlswebtest.law.harvard.edu/dev/david/googleHollis/googleHollis.html>

# Add Usage Data

- Usage data is a clue to how valuable a resource is within a community
  - Checkouts, reserves, etc.
- Privacy concerns
- Convert into a stackscore between 1 and 100

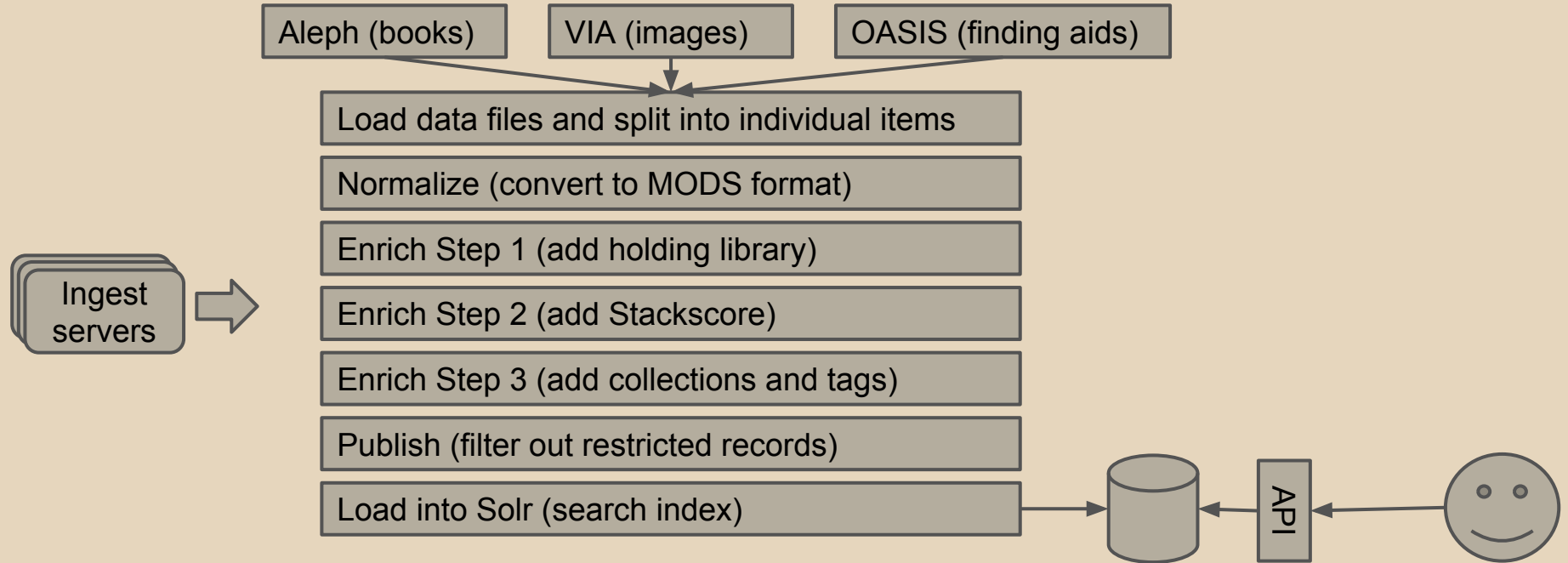
# Stacklife demo

---

<http://library.harvard.edu>



# How does this all work?



# Documentation

---

<https://wiki.harvard.edu/confluence/display/LibraryStaffDoc/Library+Cloud>

# Questions?

Jeffrey Licht  
Pod Consulting  
[jeff@podconsulting.com](mailto:jeff@podconsulting.com)