CNI Fall 2016 Membership Meeting
12-13 December 2016

The Future of Finding at the University of Oxford

Catríona Cannon, Deputy Librarian, Bodleian Libraries
Christine Madsen, Co-Founder & Chief Innovation Officer, Athenaeum21
Outline:

1. Why are we doing this?
2. The Resource Discovery review
3. The Oxford Collections Visualization tool
<table>
<thead>
<tr>
<th>Collection</th>
<th>Collection Size</th>
<th>Digital metadata record</th>
<th>Digitised resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashmolean</td>
<td>1,000,000</td>
<td>330,000</td>
<td>60,000</td>
</tr>
<tr>
<td>Oxford Botanic Garden</td>
<td>30,000</td>
<td>7,500</td>
<td>7,500</td>
</tr>
<tr>
<td>Museum of the History of Science</td>
<td>45,000</td>
<td>40,000</td>
<td>8,000</td>
</tr>
<tr>
<td>Museum of Natural History</td>
<td>6,250,000</td>
<td>152,000</td>
<td>15,000</td>
</tr>
<tr>
<td>Pitt Rivers Museum</td>
<td>650,000</td>
<td>430,000</td>
<td>240,000</td>
</tr>
<tr>
<td>Bodleian Library Published (Print+Electronic)</td>
<td>12,000,000</td>
<td>12,000,000</td>
<td>1.5M</td>
</tr>
<tr>
<td>Bodleian Special Collections</td>
<td>18 kilometres</td>
<td>?</td>
<td>48%</td>
</tr>
</tbody>
</table>

Bodleian Special Collections
Collections

The Museum is an important centre for research and teaching, and is organised into Life Collections, Earth Collections, and Archival Collections.

Search collections online: The Museum aims to make its collections as accessible as possible. To search the Museum’s online databases select the subject link from the menu on the left, and then select databases.

We have recently launched a new unified Collections Online website. At present it holds the Lepidoptera (butterflies and moths), the minerals, and the archives. We will continue to migrate data from our existing online databases to the new site.

Scientific enquiries: every year hundreds of rocks, fossils and other finds are identified by our team of curators and scientists. If you have found anything of interest, bring it along on your next visit to the Museum.

Important note: the Museum does not condone any illegal action, including the collection of birds’ eggs, trapping live animals or unauthorised work in quarries.

Nuffield Research Placements
The Museum hosts Nuffield Research Placements in our Life and Earth collections in the fields of zoology, entomology and geology. Projects cover a wide range of disciplines from collections based curation and research to insect ecology and ecosystem services. Previous project titles include “Darwin’s Insects in the Hope Entomological Collections”, “Burcheifs African Insect Collection” and “Dung Beetle Ecology” All our Nuffield Research Placements take place through Science Oxford.
Life Collections

The Life Collections comprise the zoological collections and the Hope Entomological Collections.

The Zoological Collections comprise more than 250,000 specimens, including some 1,000 type specimens. Many extinct and endangered species are represented, including the most complete remains of a dodo in the world. Notable collections include those of Thomas Bell, William Burchell, and Charles Darwin.

Within the United Kingdom, the Hope Entomological Collections are second in size and importance to the national insect collection at the Natural History Museum, London. The collection houses over 25,000 arthropod types, and comprises over 5 million specimens. Of particular significance are:

- Hope-Westwood collections
- Extensive historic collections
- The Verrall-Collin collection of Diptera
- Comprehensive British collections
- The Pickard-Cambridge and Blackwall arachnid collections
- The Wytham Woods collections

To find out more, please search the entomological collections and the associated archives.

Enquiries regarding the collections should be sent to: life@oum.ox.ac.uk

Davies, K.C. and Hull, J. The Zoological Collections of the Oxford University Museum

For information about the Zoological Collections please download this free version. Complete book, 6630Kb

Smith, A. 1986. A History of the Hope Entomological Collections

For information about the Entomological Collections and associated academics please download this free version. Complete book, 9597Kb

For slower internet connections it may be easier to download the book in four separate parts:
- Part one, 6270Kb
- Part two, 1534Kb
- Part three, 3069Kb
- Part four, 488Kb

Stanley Bowstead and Thomas M. Eccles, 2012. Drawing Techniques for Publication Download

Children can discover the world of animals in the Learning zone

Children can discover the world of insects in the Learning zone
Life Collections: searchable databases

The zoological databases include the bird and mammal specimens, a small collection of Charles Darwin’s material, the extinct and endangered species represented in the collections, and the human remains.

Search the mammals database
Over 5,000 specimens representing 750 species, including British material
Search the birds database
Over 17,000 bird skins from around the world representing 5,000 species
Search Charles Darwin’s collection
Dried crustacea collected on the voyage of The Beagle
Search the extinct and endangered database
Around 2,000 specimens cross-referenced with the IUCN 2000 Red List
Search the human remains database
400 specimens representing the species of the genus Homo
How to use the databases
A guide to searching the databases

The entomological databases provide summary listings of all the entomological collections and archives. There are also databases of type specimens, each one representing a taxonomic order, and the Diptera databases of type material from the Bigot and Verrall-Colin collections.

Search collections
A list of entomological collections
Search archives
A list of entomological archives
Search Bigot-Marcquart Diptera
The Bigot-Marcquart collection of over 2,500 species
Search Verrall-Colin Diptera
The Verrall-Colin collection of approximately 1,000 species
Search Palearctic Diptera
The Palearctic Diptera collection of over 5,000 species
Search Pickard-Cambridge Arachnid collection
The named material from the O. Pickard-Cambridge Arachnid collection
Search Embioptera
A small collection of web-spinners
Search Odonata
An important collection of dragonflies and damselflies
Search Pteroptera
A small collection of lepidoptera
Search Strepsiptera
Includes the important collection of Sir S. S. Saunders
How to use the databases
A guide to searching the databases
Life Collections: mammals database

The mammal collections comprise nearly 5,000 specimens, representing more than 750 different species. Around 500 of these specimens are British, representing over 40 species native to the United Kingdom.

To search the Mammals database please enter a search term in one or more fields and press submit.

Scientific Name
Family
Order
Common Name: bat
Collector
Nature of Specimen
Country

How to use the databases
A guide to searching the databases
More about the database and collection
A description of the size and scope of the collection

Enquiries regarding the collections should be sent to: info@oum.ox.ac.uk
Life Collections: mammals database

Records returned where species Common name matches 'bat'
There are 214 records. Records 1 to 10.

Click on the red reference number to view full specimen details.

- 00595: Myrmecobius fasciatus: Numbat: Australia
- 01407: Vombatus ursinus: Common Wombat: Australia
- 03067: Myrmecobius fasciatus: Numbat: Australia
- 03124: Hipposideros diadem: Diadem Roundleaf Bat: Malaysia
- 03125: Rhinolophus ferrumequinum: Greater Horseshoe Bat: United Kingdom
- 03127: Cynopterus sphinx: Short-nosed Fruit Bat: not known
- 03130: Thropterus nigrescens: Swift Fruit Bat: not known
- 03131: Megaderma lyra: Greater False Vampire Bat: not known
- 03132: Scotophilus kuhlii: Lesser Asiatic Yellow Bat: not known
- 03138: Lavia frons: Yellow-winged Bat: not known

Back Next

Enter new search criteria

How to use the databases
A guide to searching the databases
More about the database and collection
A description of the size and scope of the collection

Enquiries regarding the collections should be sent to life@oum.ox.ac.uk
E-journal unavailability 4th and 7th December

Please note that due to server moves by our service provider, there will be no access to ejournals or ejournal articles via OU ejournals or SOLO during the following time periods:

- Sunday 4 December from 10:00 until 19:00 GMT
- Wednesday 7 December from 11:45 until 13:15 GMT

SOLO Live Help

You can also text us on: 02030 953910
Monday-Friday, 9am-5pm

About SOLO

Coverage  
Feedback  
Getting Help  
SOLO Service Updates and Development

Other Catalogues and Services

Chinese Catalogues  
Oxford Research Archive (CRA)  
Special Collections catalogues

Additional Information

Library passwords  
Book recommendations

Managing your references

RefWorks, Endnote, Zotero, Mendeley, Covels, Papers and BibTex

SOLO video tutorials

- Searching  
- Reserving a book  
- Renewing your books  
- Saving searches  
- Signing on  
- Requesting closed stack material  
- Using the e-Shelf  
- Searching SOLO by date
Title: Legal status of Diaoyutai (Senkaku) Islands / 統家修士
Author: 統家修士
Publisher Details: 台北: 蘭汀大鷹法學會
Publication Date: 1990
Format: 畫冊1件(276頁): 25公分
Language: Chinese
Source: Allegro Chinese
Type: text
Location Information: Please follow the link on the right hand side of the screen for item information
Senkaku Islands

books
current & rare

Journal articles
e & print

manuscripts

theses

digitized card catalogues

digitized printed catalogues

online educational resources

WebLearn
iTunesU

data
business
economics

researchers

China Centre

Chinese language

Japanese studies

International relations

law

archaeology

Museum of the History of Science

collection management tool

history

geography

earth sciences

maps

Allegro

SOLO

LibGuides

digitized collections

art collections

interpretation of collections

tailored reading lists

librarians & curators
Outline:

1. Why are we doing this?
2. The Resource Discovery review
3. The Oxford Collections Visualization tool
The Resource Discovery Review

1. What we did
2. What we found out
3. How we tried to solve 1 key issue
Resource Discovery: What We Did

106 interviews, 18 site visits, 3 literature reviews

- 38 Interviews with users of collections around Oxford
- 30 Interviews with collection ‘experts’ (representing all collections)
- 22 Interviews with external institutions (including 11 site visits)
- 16 Interviews with vendors/suppliers
- 7 Site visits to Oxford libraries and museums to observe researchers
- 3 Literature reviews

---

Analysis across all data
Final Report Available Online

Resource Discovery @ The University of Oxford 2015

Analysis & Recommendations by Christine Madsen & Megan Hurst, Athenaeum21 Consulting
Research by

Christine Madsen  Megan Hurst  Simon McLeish
Iain Emsley  Ray Stacey  Masha Garibyan
Alfie Abdul-Rahman  Saiful Khan
Search behaviour & skills at Oxford are:

- about **training people to become expert in their field**, rather than finding things in the collections

- very discipline-specific. This is **not a matter of Google vs the Catalogue**

- while there is evidence that students need to learn how to search, **discovery is not as simple as ‘novice’ vs ‘expert’**. Experts in their fields use some of the the same discovery tools and techniques as young students when they switch to a new discipline or are looking for something specific

- as much about knowing **who to ask** as it is about looking for something specific

- still a **very ‘analogue’ process** for many collections and disciplines
What We Found

Search behaviour & skills at Oxford are not:

• tied to a specific discovery tool – **Google is heavily used, but by no means universal**

• correlated to a specific discipline – ‘**digital natives’ are universally bad at searching regardless of their field**

Outside of Oxford

• **Vendors** are eager to partner with their customers because they **don’t know what to do next**

• Academic and cultural institutions outside of Oxford are largely in the same place: not satisfied with their current discovery tools and **looking for what is next**
3 Recommended Areas for Investment

1. Map the landscape of things
2. Map the network of people
3. Save people time
3. Save People Time

- **Getting existing metadata out** to the places where researchers work: Wikipedia/Wikidata, Google, Google Scholar, subject-specific repositories like arXiv and PubMed, publishers like JSTOR

- **Facilitate citation chaining.** Citation chaining is ubiquitous in all areas of research across all disciplines and levels of expertise. Users use cited references as authoritative points of departure for finding more resources on a topic.
2. Map the Network of People

- **Visualization of the network of experts and research.** A graph of the professional networks at Oxford would facilitate discovery and navigation within and between fields. (NB: a number of institutions have attempted this with varying degrees of success. Care should be taken to learn lessons from others’ successes and failures.)
1. Map the Landscape of Things

Orienting the users to the corpora of collections (digital and non-digital) by:

1. **Visualizing the scope of collections at Oxford**: providing an interactive diagram that represents the range of collections at Oxford (using collection-level metadata).

2. **Cross-Collection Search**: Overlaying this diagram with the ability to search within or across collections. In such a context, researchers will understand not only what they are finding, but what they might be missing.
<table>
<thead>
<tr>
<th>Institution</th>
<th>Collection Size</th>
<th>Digital metadata record</th>
<th>Digitised resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ashmolean</td>
<td>1,000,000</td>
<td>330,000</td>
<td>60,000</td>
</tr>
<tr>
<td>Oxford Botanic Garden</td>
<td>30,000</td>
<td>7,500</td>
<td>7,500</td>
</tr>
<tr>
<td>Museum of the History of Science</td>
<td>45,000</td>
<td>40,000</td>
<td>8,000</td>
</tr>
<tr>
<td>Museum of Natural History</td>
<td>6,250,000</td>
<td>152,000</td>
<td>15,000</td>
</tr>
<tr>
<td>Pitt Rivers Museum</td>
<td>650,000</td>
<td>430,000</td>
<td>240,000</td>
</tr>
<tr>
<td>Bodleian Library Published (Print+Electronic)</td>
<td>12,000,000</td>
<td>12,000,000</td>
<td>1.5M</td>
</tr>
<tr>
<td>Bodleian Special Collections</td>
<td>18 kilometres</td>
<td>?</td>
<td>48%</td>
</tr>
</tbody>
</table>
Everything Available to Researchers
50-100M+ items

Catalogued 13M+ items

Digital Surrogates 250k+ items
(excludes “born-digital” items and eBooks)
Everything Available to Researchers
50-100M+ items

Digital Surrogates
250k+ items
(excludes “born-digital” items and eBooks)

Catalogued
13M+ items
“It's clear that reasonably diligent students are strikingly not sophisticated in their searching. Students search in one place, and if they don't find anything on the first try, they think it doesn’t exist.”
The Challenge

• We know what can be done with lots of metadata...
http://spatialinformationdesignlab.org/project_sites/library/catalog.html
Periods & Sectors from ~4000 until 2020

Overview visualization along the time periods and cultural heritage sectors. Selecting a time span shows the most common keywords, places, persons and organizations for these periods.
Deutsche digitale Bibliothek visualisiert – Urban Complexity Lab, Potsdam
A network of the most prolific persons and organizations from 1826 to 1850
Map the Landscape of Things

The Challenge

• We know what can be done with lots of metadata...

• But what can you do without it?
Create a map of all of the museum and library collections at Oxford

• Interactive visual navigation of Oxford’s collections

• Regardless of whether they have metadata

Build something that is beautiful, easy to use and transports the user into the world of Oxford’s collections

• Like a Google maps of Oxford collections
How does this project relate to other efforts like linked data?

- Assumes users have little or no prior knowledge of collections
- Is NOT dependent upon item-level metadata
- Will display collections that have no metadata
- “What does Oxford have that is relevant to my research?”
- Dependent upon item-level metadata
- Will only ever include ’catalogued’ things
- “How is this thing I have found related to other things in the collections”

A Supplement – Not a Substitute
The Project: Target Audience

“what collections are at Oxford that might be useful for my research?”

• Many incoming researchers are spending 1-2 years just finding their primary resources

“I know these things exist, why aren’t I getting any search results?”

• Most researchers have no idea the scale of what is catalogued electronically or not
The Project: Deliverables

1. A **working demonstrator**, that is, an elegant, visualization design that leverages real data (a subset of the available data) with minimum intervention. The prototype will be user-tested and will help to illustrate the feasibility and scalability of the approach.

2. A **plan for gathering summary data** to provide complete coverage of Oxford’s collections

3. A **detailed plan and budget to develop the project into a second phase** with a broader partnership
provide an immediate visual guide to:

• **what exists at Oxford** and where it is held;
• **relative sizes** of different types of collections;
• which collections are findable/searchable **electronically**;
• which collections are catalogued in **print indices**; and
• which are **not yet catalogued**.
The Prototype: How was it built?

Data from two sources:

• Existing **electronic metadata** from museums and libraries through batch exports
  • Data will be analysed and summarised to provide the highest level of navigation

• Collection **summary data** for collections that do not have item-level electronic metadata through interviews and spreadsheets
  • Every collection curator or manager at the University has a good sense of what is in their collection. Much of this summary level data already exists in spreadsheets.
The User Journey

• Users **will not start from a search term**

• Journey **may end** in a detailed metadata record, digital object, or contact information for a collection manager

• The visual navigation should **uncover the unexpected** -- that the Bodleian has art or that the Pitt Rivers has manuscripts
Carte Figurative des pertes successives en hommes de l'Armée française dans la campagne de Russie 1812-1813.

Paris, le 20 Novembre 1869.

Les nombres d'hommes présents au début, représentés par les lignes des courbes ici tracées, sont multipliés en carré 10 000 pour 100 000 hommes ; ils sont plus évidents en termes des pertes. Le tracé montre les hommes qui entrèrent en Russie, le noir ceux qui en sortirent, sans compter l'erreur dans le nombre des morts. Nous voyons au départ 30 000 hommes, 40 000 entrèrent, 90 000 en sortirent. Les guerres ont été faits grâce à l'armée de la vie. Il est dommage que les corps de l'armée aient été dispersés, mais nous avons été traités par nous. Nous avons été traités par nous.

TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au-dessus de zéro.

Les températures sont en grises.

-16° à 7° A.
-5° à 14° A.
-19° à 9° A.

Paris 24 déc.

One of the possible spatialization schemes at the top level

Research Data

Images

Specimen

3D Objects

intersection areas or new category

intersection areas

intersection areas, or new category

intersection areas or new category

intersections

TEXT

The current zoomed-in window

Prof Min Chen
search and information panel

supporting **keywords search** as well as structured search according to predefined attributes (e.g., subjects, formats, materials, dates, hierarchy, etc.)

map-based visual interface

supporting explorative interactions (e.g., navigation, zoom, virtual tours, spatial warp, etc.)

The actual visual encoding is to be designed in the pilot project.

mouse-activated snippet, showing detailed information about a collection or sub-collection.
What Have we Accomplished

• First version of a demonstrator
  • But it doesn’t quite do what we want it to do
• Much better understanding of the problem
  • Scale of the metadata issue
  • Problems with visualization
• New design ideas
• More data analysis
Welcome to the Oxford Collection Map

Below you’ll see a map of the hundreds of collections held in the various institutional collections across Oxford. Try filtering the map by Type of Thing, University, Department, or Subject. If you prefer, you can also simply noodle around by subject, type of thing or institution, if you like.

Legend

A collection
Some metadata is online
Some digitised materials
A collection, but we aren’t sure of these

Choose a filter... then more options will come up here

Clear

What am I looking at?

Each collection is represented as a circle or a square. A circle means we know how big it is, and a square means we don’t.

If we know how many things in the collection have a digital metadata record, we show proportion as a piece of pie.
Welcome to the Oxford Collection Map

Below you'll see a map of the hundreds of collections held in the various institutional collections across Oxford. Try filtering the map by Type of Thing, University Department, or Subject. If you prefer, you can also simply noodle around by subject, type of thing or institution, if you like.

Legend

- A collection
- Some metadata is online
- Some digitised materials
- A collection, but we aren't sure of size

Choose a filter... then more options will come up here Clear
ASHMOLEAN

Heberden Coin Room: Objects
Casts, Electrotypes and Forgeries
Collection size not known
No metadata online
Ancient Near East

This collection mostly contains tablets, seals and terracottas about archaeology and art history dated about 1200BC-500AD. Materials are either from or about Near East.

60,000 things

http://www.ashmolean.org/research/
Ancient Near East

This collection mostly contains **tablets**, **seals** and **terracottas** about **archaeology** and **art history** dated about **1200BC-500AD**. Materials are either from or about **Near East**.

60,000 things

STAFF COMMENTS ON THIS COLLECTION

http://www.ashmolean.org/research/
Ancient Near East

This collection mostly contains tablets, seals and terracottas about archaeology and art history dated about 1200BC-500AD. Materials are either from or about Near East.

60,000 things

http://www.ashmolean.org/research/
Ancient Near East

This collection mostly contains **tablets, seals and terracottas** about **archaeology** and **art history** dated about **1200BC-500AD**. Materials are either from or about **Near East**.

60,000 things

STAFF COMMENTS ON THIS COLLECTION

http://www.ashmolean.org/research/

**STATS**
- 60,000 things
- 28% have **metadata online**
- 1% of those have a **digitised version**
The Issues

• Requires too much interaction before understanding
• Does not display enough information
• Relies too much on text
• Users get lost

Next Steps

• Analysis of data
• More design options
Iceberg chart showing the % of the collection catalogued (above the line) and uncatalogued (below the line). Top chart shows categorical breakdown by division/academic department, bottom chart shows type of material.
Snake Oil?

Scientific evidence for popular health supplements showing tangible human health benefits when taken orally by an adult with a healthy diet.

EVIDENCE

STRONG

- fish oil
- green tea
- probiotics
- Vitamin D

MID

- Folic Acid
- cocoa
- cinnamon

PROMISING

- Aloe vera
- ginger root
- echinacea
- Vitamin B6
- Vitamin C

CONFLICTING

- arginine
- black tea
- garlic
- goji (wolfberry)

ILLUSORY

- caffeine
- green tea
- L-arginine
- magnesium
- omega-3

EVIDENCE IN PRACTICE

- copper
- lavender
- magnesium
- omega-3
- probiotics

See Data

Source: PubMed, Cochrane.org

Large human blind placebo-controlled trials only.
Snake Oil Supplements?
Scientific evidence for popular health supplements
Showing tangible human health benefits when taken orally by an adult with a healthy diet (Sep 11th 2015)
turmeric (curcumin)
peptic ulcer, ulcerative colitis
Click here to view our notes
Immortality
Biographies of the famously long-lived examined for clues to longevity

LIFE
Upbringing
Schooling
University
Post-graduate
Hardship
Lived in a city
Lived in the countryside
Lived by the sea

HEALTH
Vegetarian
Regular exercise
Sport
Outdoors
Overall active life
Regular sleep
Major illness

VICES
Alcohol early life
Alcohol mid life
Alcohol late life
Alcohol very late life
Cigarettes early life
Cigarettes mid life
Cigarettes late life
Cigarettes very late life
Cigars
Soft drugs
Hard drugs
Prescription drugs
Additory

RELATIONSHIPS
Sociability
Sex life
Marriages
Quality

RATING
0 1 2 3 4

CELEBRITY / AGE AT DEATH

RECIPE
Live in a city
Sleep well
Quit smoking
Married
“If I get dropped into the middle of the landscape, I can deduce where I am and navigate my way out, whereas my students will latch on to the first tree that looks interesting.”
Lessons Learned

• There is a place for visualisation and a place for search

• No way around creating metadata

• Whatever you do is going to take a lot of work
  • Off the shelf packages don’t seem so terrible

• Data visualization vs info visualization
  • Tolerance for abstraction
  • Data analysis versus simple visualisation

• If you want something to be design-led, hold off on data
Next Steps
Questions and discussion