

DHSI 2012: Online Tools for Literary Analysis
Instructors: Susan Brown and Stan Ruecker

Summaries by Susan Brown, Stan Ruecker,
Teresa Dobson, and Daniel Sondheim

Table of Contents

Bamboo DiRT [beta]	2
BreadBoard	3
Google Analytics	5
Timelines	6
Linked Data Prototypes (DBpedia; Scalar; Shared Canvas; SNAC; Yuma)	10
LORE: Literature Object Re-use and Exchange	14
HyperPo	16
JiTR: Just in Time Research collections management and analysis environment.....	17
Mandala Browser.....	18
ManyEyes Word Tree.....	20
Metadata Offer New Knowledge (MONK)	21
OJS Reading Tools.....	22
Old Bailey Datawarehousing Interface.....	24
Orlando: Degrees of Separation	25
Orlando Mandala.....	26
Orlando: Women’s Writing in the British Isles from the Beginnings to the Present.....	28
OVis	30
Pliny.....	34
TextArc.....	35
Voyant.....	36
Voyant Word Trends	37
Voyant Cirrus.....	38
Voyant Links	39
Voyant Bubblelines	40
Voyant Scatterplot.....	41
Wordhoard	42
Wordle	43
Zotero.....	44

Bamboo DiRT [beta]

Daniel Sondheim

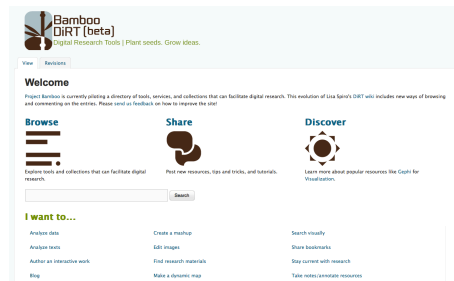
Bamboo DiRT [beta] provides links to a wide assortment of tools by different developers. Tools are organized by category; examples include annotation, brainstorming, data conversion, image editing, linguistic research, text mining, transcription, and visualization. Organization is also done by platform, cost, and level of use, as well as by tags. Users can share new tools of their own.

Example

Someone wishes to find applications that will enable collaborative editing of a document.

Getting Started

1. Go to <http://dirt.projectbamboo.org/>
2. Click on one of the categories under “I want to...”
3. Narrow your results by specifying the platform, cost, and other options provided at the top of the list.
4. Click on the title of a tool to see more information about it, or to comment on it.
5. Click on the URL provided in the tool’s description to use it.



Who has worked on creating it

Bamboo DiRT [beta] is based on Lisa Spiro’s DiRT wiki, and is developed and maintained by Project Bamboo and supported by the Andrew W. Mellon Foundation.

Where to read about it

A short “About” section can be found at <http://dirt.projectbamboo.org/about>

Availability

Bamboo DiRT [beta] is freely available online, at <http://dirt.projectbamboo.org/>.

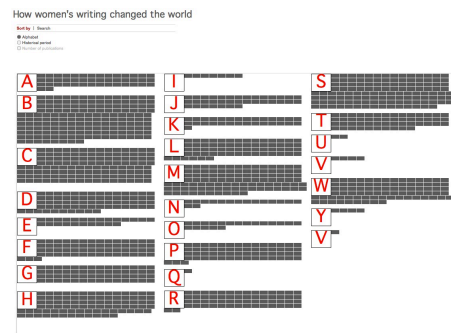
BreadBoard

Susan Brown

BreadBoard is an experimental interface for exploring the Orlando database. It allows the contents of the entire database to be viewed at once on a single screen, thereby allowing researchers to easily find and view authors in a way that shows a compact representation of their connections to other authors, places, literary references, organizations, and people. BreadBoard's representation can be sorted by the first letter of the author's last name, the historical period, or other search parameters.

Example

A literary scholar wishes to find out which people are most prominent in W. B. Yeats' social networks, and why.



Getting Started

Start

6. Go to the Orlando BreadBoard link at <http://medusa.arts.ualberta.ca/development/protected/prototype/breadboard/> [it is not currently live, but soon will be!]
7. Log in with your Orlando user name and password.
8. Wait for the page to load.
9. You will see a list of writers' names that are in the Orlando textbase.
10. These names are organized alphabetically by default.

Sort and Search

11. To sort by a writer's last name or historical period, click on "Alphabet" or "Historical Period".
12. To search, click on the "Search" option to the right of the "Sort by" option.
13. You can choose to limit the search results by scope and date.
14. To limit by scope, select or deselect "British Women Writers," "Male Writers," and/or "Other Women Writers."
15. To limit by date, select the desired date range from the drop down menu. Select only one date range.

View

16. Click on a writer's last name to see more information about that writer.
17. Wait for the page to load.
18. In the "Results Summary" tab, a summary of the writer's life and work appears below the writer's name.
19. The "Results Summary" tab shows a listing of all of the different links in the Orlando textbase entry for that writer.
20. These links are colour coded according to whether they relate to "Social Networks," "Organizations," "Literary Relations," or "Places." The colours correspond to the colours for each tab listed underneath the information about the writer.
21. Click on the "Social Networks," "Organizations," "Literary Relations," or "Places" tabs to see a tag cloud with large and small links.
22. The links that have the highest frequency within the entry are larger and bolder than entries that occur less frequently.
23. These tag clouds give an overview of the organizations, social networks (people), literary relations, and places with which the author was associated, based on how frequently these names, places, organizations, and works of literature are mentioned in the entry.
24. These tag clouds can be narrowed by deselecting or selecting the tags that appear on the left-hand side of each tab.
25. Click on a link to view its context in the writer's entry, which is extracted from the selected writer's entry in the Orlando textbase.
26. Wait for the page to load.

Who has worked on creating it

The Breadboard Interface was designed by Susan Brown, Milena Radzikowska, and Stan Ruecker. Programming by Jeffery Antoniuk and Mark Bieber. Funded by the Social Sciences and Humanities Research Council of Canada.

Where to read about it

A document describing Breadboard is available here: <http://entry.tapor.ca/?id=25>

Availability

Online, with account by arrangement with orlando@ualberta.ca.

Google Analytics

Daniel Sondheim

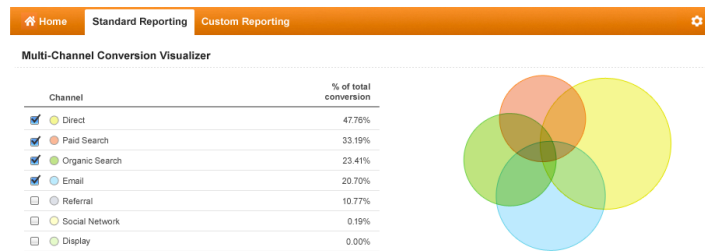
Google Analytics is a set of tools created and made available by Google. In general, these tools are designed to provide detailed statistics about website visitors. Various kinds of tools are available, including ones developed for content analytics, social analytics, mobile analytics, conversion analytics, and advertising analytics. There are also ones developed specifically for data analysis. Results may be displayed using a variety of methods, including a number of types of visualizations.

Example

The owners of a website wishes to determine the geographical locations of their website visitors.

Getting Started

1. Navigate to <http://www.google.com/analytics/index.html>
2. Click on “Create an account.”
3. Add tracking code to your website.
4. Learn about your website visitors using the various tools and visualizations available at the site.



Who has worked on creating it

Google Analytics is based on Urchin on Demand, originally developed by Urchin Software. Urchin Software has been owned by Google since 2005.

Where to read about it

Features of Google Analytics are described at <http://www.google.com/analytics/features/index.html>

Availability

Free and paid versions are available online at <http://www.google.com/analytics/index.html>.

Timelines

Susan Brown and Daniel Sondheim

ChronoZoom:

<http://eps.berkeley.edu/~saekow/chronozoom/projectinformation/index.html>

A browser-based tool that facilitates the study of “Big History,” including four historical epochs: Cosmic, Earth, Life, and Human history.

Getting Started: Click and drag the timeline left and right to pan, or zoom in and out using the mouse wheel. Click on the circles to view media. Click on the button that looks like a section of film in the upper left-hand corner of the ChronoZoom screen for a guided tour.

Advanced Tips: Click on the textual labels of the bars in the upper right-hand corner of the screen to zoom to the levels that correspond with the labels.

Who has worked on creating it: ChronoZoom was developed Underlying, Inc.

* * *

Chronos Timeline:

<http://hyperstudio.mit.edu/software/chronos-timeline/>

A browser-based timeline creation tool designed specifically for the humanities and social sciences community. It is currently in beta, and will be released as free software. It allows navigating by means of tags, and viewing of the timeline either vertically or horizontally.

Getting Started: Click and drag any of the three components of the timeline to pan at different rates. Click on the tags in the cloud at the right or within the stories to highlight points in the timeline based on their subject matter.

Advanced Tips: You can use your own data to create a timeline by clicking on “To try your own data set, click here” and entering a Google Spreadsheet URL. You must have “start,” “eventid,” and “title” fields in your spreadsheet.

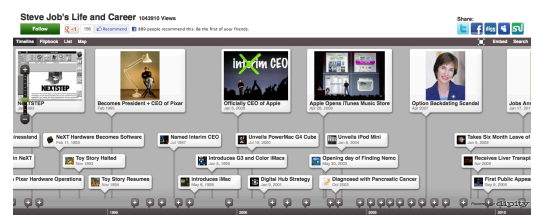
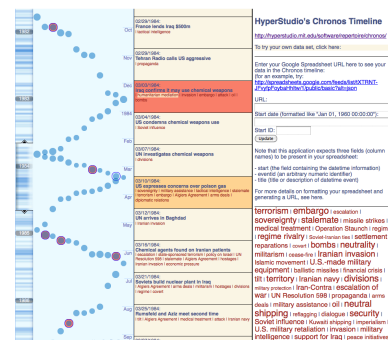
Who has worked on creating it: Chronos Timeline is developed by HyperStudio at the MIT School of Humanities, Arts and Social Sciences, and affiliated with the Comparative Media Studies Program.

* * *

Dipity: <http://www.dipity.com/>

Dipity is a browser-based timeline-creation tool. Users can create their own interactive timelines on any subject, and embed media of different types within them.

Getting Started: Click on “Create a Timeline” to create an account and create a timeline. To view a list of popular timelines, click on “Hot Topics.”



Advanced Tips: When viewing a timeline, zoom in and out by using the slider on the left, and pan forwards and backwards by using the scroll bar at the bottom. Click on the plus-sign tabs along the bottom of the screen to automatically zoom in and view additional elements at that point in time.

Who has worked on creating it: Dipity was developed by Roland Saekow and Walter Alvarez at the Earth and Planetary Science University of California, Berkeley.

* * *

Monk: <http://monkproject.org/>

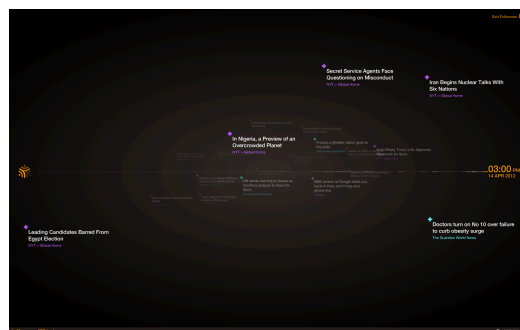
The workbench includes a timeline visualization tool that is included in some of the toolsets and can be used with the existing text collection without any downloads or installation. Please see the Monk entry in this coursepack for more information on Monk.

* * *

Voyage: <http://rssvoyage.com/>

Voyage is a browser-based RSS feed reader. Stories from various sources are organized chronologically.

Getting Started: Stories can be navigated by means of a mouse wheel or arrow keys. Feeds can be added and removed by opening the pane at the bottom of the screen (and registering as a user).



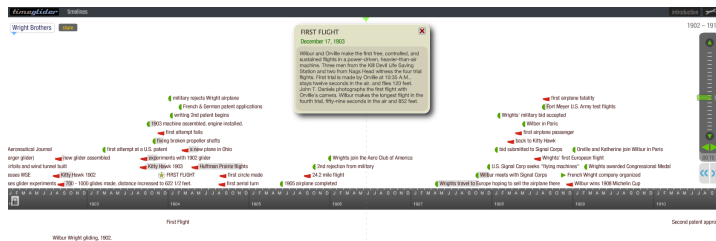
Advanced Tips: Voyage can be viewed full screen by clicking on the button in the upper right-hand corner of the screen.

Who has worked on creating it: Voyage was developed by Andy Biggs.

* * *

Timeglider: <http://timeglider.com/>

A browser-based application that allows users to create an interactive timeline with zooming and panning functionality. Events can be assigned “importance levels,” so that they can be made to show up at different zoom levels. Images can be included.



Getting Started: You can create a timeline without signing up at <http://timeglider.com/app/try.php>. To navigate timelines, click and drag the horizontal bar to pan; click and drag the vertical slider to zoom in and out. Click “GO TO...” to jump directly to a particular date.

Advanced Tips: Click on events in the timeline to see more information about them. Click on the wrench icon in the upper right-hand corner for further options.

Who has worked on creating it: Timeglider was developed by Mnemograph LLC.

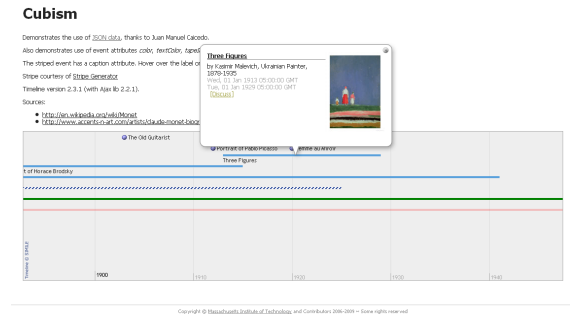
* * *

Timeline: <http://www.simile-widgets.org/timeline/>

An open-source web widget that allows you to create interactive horizontal timelines of temporal data. It allows for the inclusion of images as well as text, and can produce multiple temporal bands that can scroll independently of each other.

Getting Started: Follow the directions in the Timeline wiki at http://code.google.com/p/simile-widgets/wiki/Timeline_GettingStarted. If you do not have access to a server, you will need to download the local installation. Familiarity with javascript is desirable.

Advanced Tips: SIMILE offers other visualization Widgets at <http://www.simile-widgets.org/>



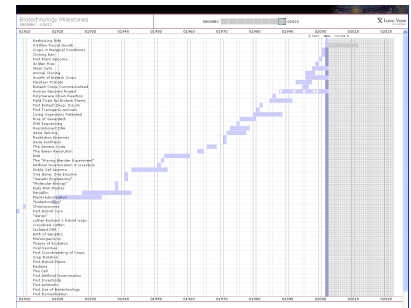
Who has worked on creating it: Timeline was developed by the Simile Project (Semantic Interoperability of Metadata and Information in unLike Environments) at MIT and is now part of Simile Widgets, an ongoing open-source project.

* * *

Longviewer: <http://redpuma.net/longview/tree/biotech-html/>

An open-source server side tool for showing a series of events over centuries and millennia, which has a number of interesting features including representing the future with a moving “now” point.

Getting Started: click on the Download link at <http://www.longnow.org/about/longview.php>



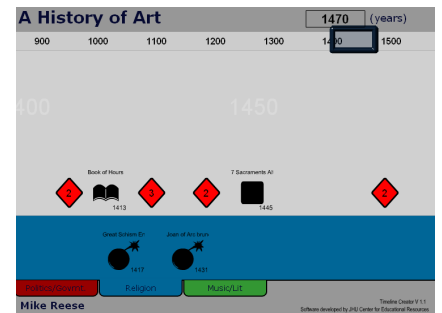
Who has worked on creating it: Stewart Brand, James Home, Dan Mosdale, Kurt Bollacker, Alexander Rose, with funding from The Long Now Foundation, co-founded by Brian Eno.

* * *

Timeline Creator: <http://timeline.cer.jhu.edu/community.htm>

A free Java application that allows those without multimedia development skills to develop an interactive timeline that can be published on the World Wide Web or presented directly from a PC. Can display up to six simultaneous timelines with descriptive text and media such as digital images, audio and video.

Getting Started: Follow the instructions on the “Download now!” tab.

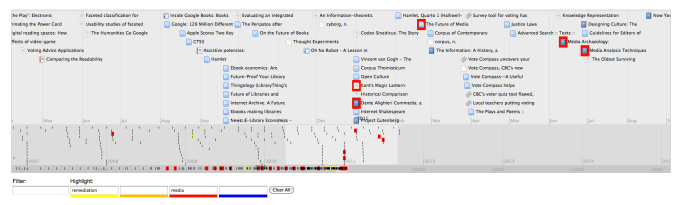


Who has worked on creating it: Center for Educational Resources, Johns Hopkins

* * *

Zotero Timeline:
<http://www.zotero.org/blog/visualize-your-zotero-collections-in-timelines/>

Timeline is a feature of Zotero, which is open source reference management software. Timeline



allows users to visualize their collections (or portions thereof) in a chronological fashion, with chosen keywords highlighted. Filters can also be employed in order to limit results.

Getting Started: First, you will need a copy of Zotero. You can download it from <http://www.zotero.org/>, either as an add-on for Firefox or as a standalone program. Once you have Zotero installed and have gathered a collection of references, select the references that you would like to visualize, and then click on the gear icon and select “Create timeline.”

Advanced Tips: Apply a filter by entering it into the appropriate input field. You can also highlight keywords of interest in four colours. You can specify what the bands in the timeline represent by choosing different options from the dropdown menus, allowing you to customize your visualization in different ways. Selecting an item in the timeline also selects it in your library.

Linked Data Prototypes (DBpedia; Scalar; Shared Canvas; SNAC; Yuma)

Susan Brown

“Linked Data is simply about using the Web to create typed links between data from different sources. These may be as diverse as databases maintained by two organisations in different geographical locations, or simply heterogeneous systems within one organisation that, historically, have not easily interoperated at the data level. Technically, Linked Data refers to data published on the Web in such a way that it is machine-readable, its meaning is explicitly defined, it is linked to other external data sets, and can in turn be linked to from external data sets.”

Christian Bizer, Tom Heath, and Tim Berners-Lee, “Linked Data - The Story So Far”:
tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf

Wikipedia: “linked data describes a method of publishing structured data so that it can be interlinked and become more useful”: http://en.wikipedia.org/wiki/Linked_Data

General resources on Linked Data:

Paul Miller, JISC-funded report Linked Data Horizon Scan, 2010: <http://linkeddata.jiscpress.org/>
 Tim Berners-Lee, Four Principles of Linked Data: <http://www.w3.org/DesignIssues/LinkedData.html>
 Time Berners-Lee, James Hendler, and Ora Lassila, “The Semantic Web,” *Scientific American* 2001:
<http://www.scientificamerican.com/article.cfm?id=the-semantic-web>

The following tools demonstrate the potential of linked data from diverse sources.

For a Linked Data tool that you can use with web materials of your choice, see [LORE: Literature Object Re-use and Exchange](#).

* * *

DBpedia: <http://dbpedia.org/>

This is a linked data set derived from the structured information embedded in Wikipedia articles (largely the content of the infoboxes on the top right of entries). It is one of the larger and more prominent linked data sets.

Getting started:

The [Faceted Wikipedia Search](#) site provides human search engine interface and browser at <http://dbpedia.neofonie.de/browse/>

The example queries give a sense of the power of SPARQL queries. Although you can put in a search term and then facet the results, it is most interesting to start from the left column and work your way down into the categories to produce results sets akin to those from the example queries. For instance, People who were artists who were born in London and active from 1800-1900.

The focus of DBpedia is to allow computers to interact with the dataset by means of SPARQL semantic web queries, for which there is a public endpoint at: <http://DBpedia.org/sparql>.

Who has worked on creating it: the [Wikipedia](#) community.

The screenshot shows the DBpedia website interface. On the left, there are search filters for 'Item type' (Person), 'Born in' (London), and 'Active from year' (1800-1900). The main content area displays search results for 'George Gordon Byron, 6th Baron Byron' and 'Michael Caine'. Each result includes a small portrait and a brief description of the person's life and work.

* * *

Scalar Authoring and Publishing Platform:<http://scalar.usc.edu/anvc/>

Produced by the Alliance for Networking Visual culture, the Scalar platform uses RDF to allow annotation of multimedia objects hosted elsewhere in the production of long-form scholarship. It supports collaborative authorship and rich commentary.

Getting started:

There are a number of example texts that use Scalar, including the digital companion to the book *The Nicest Kids in Town* by Matthew F. Delmont: <http://scalar.usc.edu/nehvectors/nicest-kids/introduction>

Jentery Sayers provides a description of and screenshots from his work with Scalar for various purposes on his blog at <http://www.jenterysayers.com/2012/scalar/>. He summarizes his conclusions as follows:

Scalar compellingly demonstrates why we are in an exciting moment for scholarly communication. Thus far, the platform has become a vehicle for me to compose, argue, speculate, and teach through new media, graphical expression, collaboration, context-sensitive design, metadata, and various ontologies. While I still have much to learn, at this juncture in my research I have realized the degree to which such platforms facilitate interpretation via dynamic content and computation, which need not be rendered positivist, determinist, or value-neutral in character.

As of the time of writing this, the beta release that would allow users to author their own materials in Scalar was not yet available, but was promised soon.

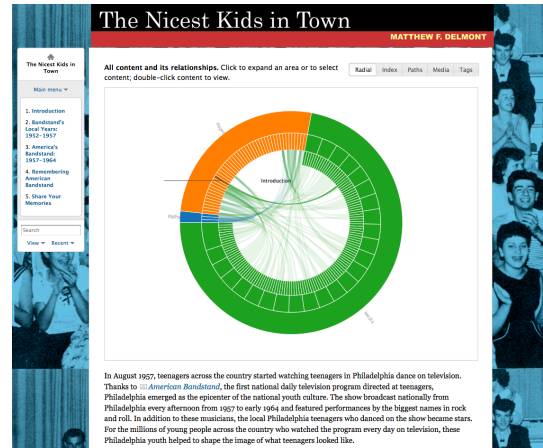
Who has worked on creating it: [The Alliance for Networking Visual Culture](#), with Tara McPherson as Lead PI and Wendy Chun, Brian Goldfarb, Nicholas Mirzoeff, and Joan Saab as co-PIs. The Core development team is Erik Loyer, Craig Dietrich, Tara McPherson, Steve F. Anderson, Jillian O'Connor, John Bell, Alex Louie and George Brower.

* * *

Shared Canvas: A Distributed Canvas Rendered from Linked Data Annotations: <http://www.shared-canvas.org>

This open-source project is producing a data model and set of web services designed to allow distributed resources to be combined to produce a composite object. The impetus comes from medieval studies in which portions of manuscripts often exist in various locations and heavy annotation of the visual object in various media is required. The data model is based on the emergent standard for annotations being developed by the [Open Annotation Collaboration](#).

More information about the project can be found at: <http://www.shared-canvas.org/slides/> including an [overview](#) from June 2011.



Yuma needs to be installed in a host environment to work, and the test sites do not allow you to upload your own objects for annotation.

Further information is available at: <http://ercim-news.ercim.eu/en86/special/yuma--crowd-sourced-metadata-enrichment-for-online-collections>

Getting started:

Go to the test site page at <http://dme.ait.ac.at/annotation/>. There are links here to screencasts of the tool in action, instructions on usage, and to test sites for the the four different media types. Note that some of the test items are very slow to load, and not all seem to work at all times.

Who has worked on creating it: Joachim Jung, Rainer Simon and Bernhard Haslhofer and others at the Austrian Institute of Technology and Cornell University.

LORE: Literature Object Re-use and Exchange

<http://itee.uq.edu.au/~eresearch/projects/aus-e-lit/lore.php>

Susan Brown

LORE is an extension to the Firefox web browser that allows you to annotate objects that are elsewhere on the web. It employs the W3C's RDF ([Resource Description Framework](#)) model for data interchange for the [Semantic Web](#). It emerges from the [AUS-e-LIT](#) project's work towards infrastructure for study of Australian literature and print culture, but can be used with any web object.

As outlined on the site, LORE allows you to:

- *Author Compound-objects:*
 - **Gather** web resources and organise and tag them, describe links between them and make notes about them via the graphical editor.
 - **Discover** compound objects that others have created about related topics and web resources.
 - **Publish** the information that scholars have assembled as compound objects so that others can find them and reuse the information.
 - **Communicate** the information using automatically generated documents, slideshows and diagrams.
- *Annotate web resources:*
 - **Contribute** comments, notes and additional information by annotating web resources.
 - **Search and browse** annotations contributed by others.
 - **Discuss** topics and resources by replying to annotations created by others.

Getting started:

Documentation:

Two pdf overviews of LORE: [Using Annotations](#); [Using Compound Objects](#)

Useful video demonstrations/tutorials: <http://itee.uq.edu.au/~eresearch/projects/aus-e-lit/demos.php>

If you don't use it already, download and install Firefox from the [Mozilla](#) site. If you use Firefox, ensure you are using version 3.5 or later.

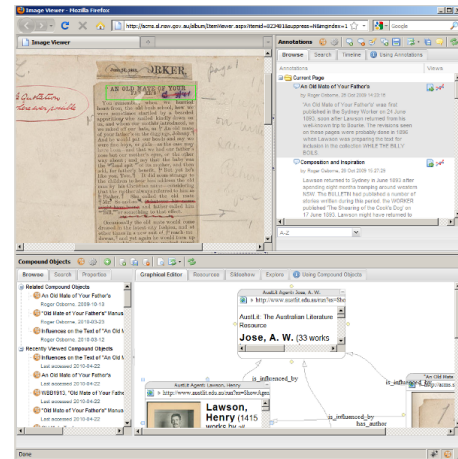
Go to <https://addons.mozilla.org/en-US/firefox/addon/146333/> to install the add-on.

In order to create annotations, you need to create a user account and sign in.

LORE installs a small gold button on the lower right corner of the browser window, near the resizing button. Clicking on the button opens two additional panes, the Annotation Pane and the Resource Map pane, in your browser. These can be independently closed and resized depending on your activities. Clicking the gold button again closes both panes. Help is available in tabs within these panes.

Advanced tips: To set the name associated with your annotations or other preferences, in Firefox go to the Tools → Add-ons menu item to get to your Add-ons Manager, then click on Extensions (not Add-ons) to get to your preferences. You can also get these settings by using the gear icon in the LORE annotation pane.

Who has worked on creating it: Lead researchers are Jane Hunter, Kerry Kilner, and Richard Osborne, and Anna Gerber, Damien Ayers, Andrew Hyland and Chris Davoren have worked on the coding. The Aus-e-Lit project is funded by the National Collaborative Research Infrastructure Strategy (NCRIS) Platforms for Collaboration, through the National eResearch Architecture Taskforce (NeAT), and by the University of Queensland.



Google N-gram viewer

<http://books.google.com/ngrams/>

Susan Brown

The Google N-gram viewer allows you to search words or sequences of words from a subset of the Google Books corpus in seven languages. It provides a visualization of the frequency of the character string occurrence over time.

Getting started:

Go to the site and type a word or words.

Separate words or strings that you wish to compare with commas

The results are in the form of a graph and links to searches on the books in which the words occur.

Note that the data is the result of mass digitization and contains OCR errors, metadata errors and other inconsistencies.

Documentation: <http://books.google.com/ngrams/info>

Further reading:

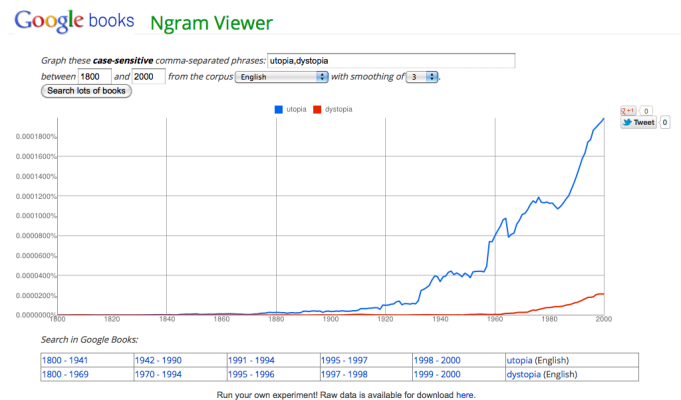
The flagship article: Jean-Baptiste Michel, Yuan Kui Shen, Aviva P. Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak and Erez Lieberman Aiden. "Quantitative Analysis of Culture Using Millions of Digitized Books," *Science*, December 2010.

<http://www.sciencemag.org.subzero.lib.uoguelph.ca/content/early/2010/12/15/science.1199644>

[The Culturomics site](#)

["The Jargon of the Novel. Computed,"](#) *New York Times*, July 29, 2011

Who has worked on creating it: Google, and Harvard University's Cultural Observatory, directed by Erez Lieberman Aiden and Jean-Baptiste Michel.



HyperPo

Teresa Dobson

HyperPo is an online tool described on the HyperPo 6.0 site as a “user-friendly text exploration and analysis program” designed to “close the gap, created by traditional text analysis tools, between an original text and data generated about that text. It accomplishes this by creating hypertextual links between different representations of the text.” HyperPo allows one to input for analysis text documents from on the web or from a local drive. Some of its key features are as follows: it creates frequency, KWIC, co-occurrence and distribution lists; it colour codes words to reflect repetition information; it allows for the simultaneous comparison of data from multiple texts; and it allows for the use of external resources such as Google and ARTFL.

Example

Here the text of Milton's *Areopagitica* has been submitted to HyperPo. The full text displays in the top left pane. Colour-coded (blue) words are those that appear most frequently. A word frequency list appears to the right. The word “licensing” appears in the KWIC display at bottom of the screen. (Scrolling within the bottom pane would also display a series of tables and graphs showing the frequency, distribution and collocations of *licensing*.)

The screenshot shows the HyperPo interface with three panes. The top-left pane displays the text of Milton's *Areopagitica* with blue highlighting on frequently occurring words. The top-right pane shows a word frequency list with columns for the word, its count, and its percentage of the total text. The bottom pane shows a KWIC display for the word "licensing", with a table of surrounding text snippets.

Word	Count	Percentage
up	12	10.00%
one	12	10.00%
first	12	10.00%
many	12	10.00%
some	12	10.00%
new	12	10.00%
her	12	10.00%
much	12	10.00%
god	12	10.00%
hath	12	10.00%
new	12	10.00%
been	12	10.00%
do	12	10.00%
's	12	10.00%
cannot	12	10.00%
only	12	10.00%
things	12	10.00%
great	12	10.00%
experience	12	10.00%

Context
this authentic Spanish policy of licensing books, if I have said among us, for all that licensing can do! Yet this is they do, but must be licensed what they may say, and Canada westward, while our Spanish licensing gaps the English press never unless it were approved and licensed under the hands of two and perfection. Suffer not these licensing contributors to stand at every get into their custody the licensing of that which they say from all quarters to twenty licensing larger? Had anyone written and field, we do injuriously, by licensing and prohibiting, to mislead her should be referred to the licensing of some more sober workmasters and the original of book licensing roused up and drawn as

Getting Started

1. Visit <<http://hyperpo.org/>>.
2. Decide upon a source text (you may select an existing website, paste in text, or upload a local file).
3. Decide whether you wish to perform the linguistic analysis and select “submit” at the bottom of the screen.
4. If you complete the linguistic analysis, the results will display in a table.
5. Select “continue.” The text will display in a three-pane screen as above: full text, frequencies, and information.

Advanced Tips

- Scrolling down in any of the three panes will reveal options for display. For example, you may reveal markup in the text pane or choose whether to enable the colour-coding feature for word frequencies. In the frequencies pane you may view options for sorting and displaying results (e.g., whether word forms are lemmatized or lexical, whether function words are displayed, and so on). In the information pane you may specify words to match, increase or decrease the amount of context to display, and so on.

Who has worked on creating it

HyperPo was initially designed and written by Stéfan Sinclair, who notes that the program is now moving into an “open-source” phase where it is hoped that collaboration will bring about more improvements and features.

Where to read about it

Sinclair, Stéfan. “Computer-Assisted Reading: Reconceiving Text Analysis.” *Literary and Linguistic Computing* 18 (2): 175-184, 2003.

Availability

HyperPo is available online here: <http://hyperpo.org/>

JiTR: Just in Time Research collections management and analysis environment

Susan Brown

JiTR provides a means of compiling, tracking, sharing, editing, and annotating online research collections. It makes it possible for researchers to create “mashups” or combinations of streams of texts from different sources for purposes of scholarly editing, sociolinguistic study, and literary, historical, or conceptual analysis. It is currently an alpha prototype.

Example

Someone wishes to compile, sort, and annotate a collection of online documents connected with their research, and to share it with collaborators and research assistants.

Getting Started

Start

1. Go to the JiTR link at <http://ra.tapor.ualberta.ca/~jitr/>
2. Log in with your username and password.
3. Wait for the page to load.
4. You can either view existing collections or create a new collection.

Create a Collection

5. To create a new collection click on "New Collection" on the home page.
6. Choose a name and a description for you collection and click "Create".
7. On the home page you will see your new collection.
8. Click on your collection to add a new text.
9. To add a text click "Add a text" or "Upload XML" under "Import Text" on the right hand column.
10. To add an XML file go to "Upload XML" and choose your file then click "Upload".
11. To add text, HTML or XML choose "Add Text".
12. Here you can give your file a name, tags, description and text format.
13. You then paste your text into the Source text textbox and press "Create"

Analyzing Texts

14. To analyze your text click on the collection from the home page
15. Under the text you want analyzed click "Analyze" and Voyant Tools will open to analyze your text

Advanced Tips

- Use “tags” to organize and interlink your collections.
- Use “notes” to annotate your documents. Notes are linked to the document and also stored in your Notes collection.
- [if implemented in time] Use the JavaStart XML editor to begin creating a textual edition as part of your JiTR collection.

Who has worked on creating it

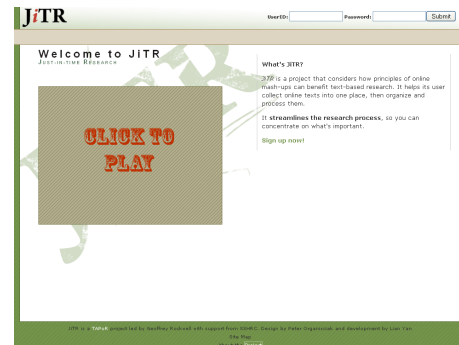
JiTR has is being developed by a team led by Geoffrey Rockwell (first at McMaster University and now at the University of Alberta) in connection with the TAPoR Project, involving Susan Brown, Stan Ruecker, Stéfan Sinclair, Peter Organisciak, Kamal Ranaweera, James Chartrand, Lian Yan, and Shawn Day.

Where to read about it

At <http://tada.mcmaster.ca/Main/MashTexts>.

Availability

Accounts available by arrangement with the development team.



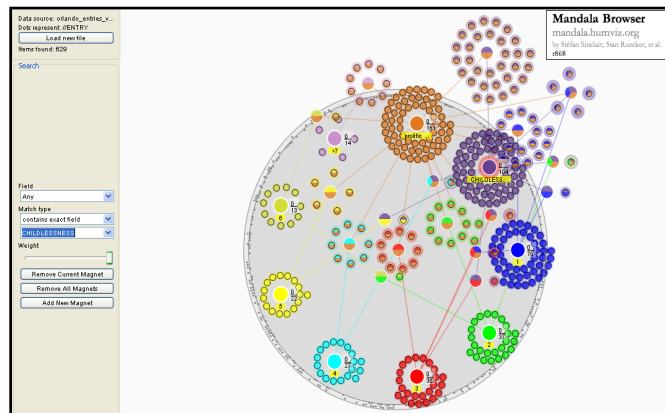
Mandala Browser

Stan Ruecker

It allows you to do fairly sophisticated visual browsing of any XML document or collection of documents. Parts of documents or entire documents appear as dots around the periphery. You create colourful magnets and assign values to them. These magnets draw the dots into the centre space. There is a text viewer off to the right, so you can see what is inside any given dot. You can also export what you find for later study.

Example

You might be interested in the question: do women writers who have children write less than women writers without children? To address this question, you could load the Orlando collection, and make each dot represent a woman writer. Create magnets 0-9 representing numbers of children. You see immediately that there are diminishing numbers of women writers for each number of children, but there are some writers for every number. However, maybe the ones with more children wrote less. So you create new magnets dealing with measures of prolixity, by looking at the author summary tag, which summarizes a writer's achievements, for the words "prolific" or "numerous" or "many." It appears that by these measures, having more children equates to being more prolific. Now the third step is to access the documents, either by export or by logging into Orlando, to pursue further details.



Getting Started

To set up Mandala:

1. Go to <http://tapor-dev.mcmaster.ca/~humviz/mandala/dev/standalone/>.
2. Click on the highest version number.
3. A ZIP file should download to your computer.
4. Unzip the file to create the mandala folder.
5. Doubleclick the .jar file to run the mandala browser.

To use Mandala:

1. Figure out what tag in your xml file will be used as a dot in the mandala.
2. Copy your xml file into the mandala directory.
3. Copy the file name.
4. Doubleclick the jar file to start the mandala.
5. Paste the file name into the "Open" field top left.
6. Type your dot tag into the "Dots Represent" field top left.
7. Push the "Load" button. Dots should appear around the periphery. Large files may take a few minutes to load.
8. An empty magnet is preloaded. Set up the criteria for it and watch the dots gather to it. For a free-text search on all XML fields, type the search text in the box and hit return. To narrow the search, first select an XML field from the picklist above the search box. You can also choose to see a type-ahead response while typing by clicking on the checkbox called "Search as you type."
9. Create the next magnet, and fill in its criteria.
10. Iterate as necessary.
11. Read items in the panel on the right by clicking them or lassoing them in the main circle.

12. If you click on a given magnet, you may be unable to see information that you want. If this is the case, then click on the Display palette and on the "Fields always displayed" dropdown and select more boxes to display. If you return to the magnet, you will now be able to see the extra information in the reading pane on the right.
13. If you would like to see the finer details of your Mandala visualization, you may zoom in / out and adjust the screen with the Display palette. Click on the Display palette and you may zoom in / out with the '+' and '-' signs. Another way to zoom is with a mousewheel / two-finger scroll. You may also navigate the screen with the ←, →, ↓, ↑ arrows. If you'd like more control, you can click on the hand tool to drag the screen around.
14. We may also export our Mandala visualization by clicking Export on the bottom-right side of the window. If you choose Text, you can save a copy of the text displayed in the reader panel. If you choose Screenshot, you can save a visual copy of the Mandala visualization.

Advanced Tips

- You can use regular expressions to set up magnet values, allowing you to have several tags, for instance, attracted to a single magnet.
- You can also search on tag attributes.

Who has worked on creating it

Mandala was developed with funding from SSHRC by Stefan Sinclair, Anthony Sapp, Matt Patey (McMaster University), Stan Ruecker, Oksana Cheypesh, Constanza Pacher, Rhiannon Gainor (University of Alberta), Sandra Gabriele (York University)

Where to read about it

Brown, Susan, Patricia Clements, Isobel Grundy, Stan Ruecker, Jeffery Antoniuk, Sharon Balazs, Stéfan Sinclair and Matt Patey. "Thinking Beyond the Text: Using the Mandala Browser to Explore Orlando." Paper presented at the Society for Digital Humanities/ Société pour l'étude des médias interactifs annual conference at the 2008 Congress of the Social Sciences and Humanities, University of British Columbia. June 2-3, 2008.

Gainor, Rhiannon, Stéfan Sinclair, Stan Ruecker, Matt Patey, and Sandra Gabriele. [Forthcoming]. "A Mandala Browser User Study: Visualizing XML Versions of Shakespeare's Plays." *Visible Language* 43(1). 2009.

Ruecker, Stan. "Experimental Interfaces Involving Visual Grouping During Browsing." *Partnership: the Canadian Journal of Library and Information Practice and Research*. 1(1). 2006.

Availability

Web site at www.humviz.org.

Download: <http://tapor-dev.mcmaster.ca/~humviz/mandala/dev/standalone/>

See a demo video at www.ualberta.ca/~sruecker/mandala_demo.mov

ManyEyes Word Tree

Stan Ruecker

With this robust online site from researchers at IBM, you can upload your own data and choose from an ever-expanding menu of ways to see it visualized. All of the standard charts are available, but there are also some that are more fun, like Word Clouds and Wordles, and the one that we recommend as being usefully exploratory, called Word Tree. It is a form of visual concordance that shows a document with the repeated words enlarged, as they are in a Word Cloud, but left within the sequence of the document.

Example

On the Word Tree page, they show the speech by Martin Luther King, Jr. where you can quickly see the resonating phrase “I have a dream” and all of the variations Dr. King used to complete the sentence.



Getting Started

1. Note first of all that ManyEyes is a public site, so any data that you post will be visible to the world.
2. Visit manyeyes.alphaworks.ibm.com/manyeyes/page/Word_Tree.html to read what it is about
3. IBM has the whole thing behind an account login, so you need to go and create an account at www.ibm.com/developerworks/dwwi/jsp/WSLogin.jsp. You can also get there from the link on the left nav bar called “Upload a data set”
4. Once you are logged in, you can upload a text document by pasting it into the box. There are also optional fields where you can describe your document.
5. Enter a search term
6. Traverse the results. Try another search term. And so on.

Advanced Tips

- There are radio buttons that let you reverse the order, so the word you are looking for will appear at the end of phrases rather than at their beginning.
- By default, clicking will zoom in on a word, but you can change that setting so that clicking will highlight.

Who has worked on creating it

ManyEyes was created by Martin Wattenberg, Fernanda Viegas, and others at IBM (<http://manyeyes.alphaworks.ibm.com/manyeyes/page/About.html>).

Where to read about it

Danis, Catalina M., Fernanda B. Viégas, Martin Wattenberg, Jesse Kriss. “Your Place or Mine? Visualization as a Community Component.” CHI 2008.

A more complete list is available here: <http://www.bewitched.com/manyeyes.html>

Availability

The list of visualization options can be found at:

http://manyeyes.alphaworks.ibm.com/manyeyes/page/Visualization_Options.html.

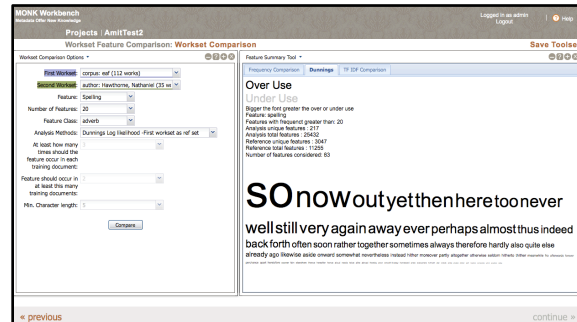
Metadata Offer New Knowledge (MONK)

Stan Ruecker

MONK is a workbench of tools that have been organized into toolsets. Working with the relatively large set of primary documents that have been pre-ingested into MONK, you have the opportunity to use sophisticated algorithms to cluster documents, automatically search for similar documents, study representative vocabulary, and so on. In clustering, for example, you can be used to cluster texts around some concepts that are of interest to you. Then you can see what features the system thinks each of the items in those clusters have in common, and you can use those features to see what other documents the system thinks belong to your clusters.

Example

If you were interested in themes in sermons, you might find several sermons in the MONK collection and create clusters of sermons that represent the key themes. Examples might include grace/works, salvation/damnation, and political commentary. Run the workbench, choose the toolset that deals with Clustering, then look at what features have been found to represent each cluster. Look also at documents that appear to be similar based on those features.



Getting Started

1. Visit www.monkproject.org
2. Start the workbench.
3. Choose an existing Project, such as AmitTest2
4. Choose the toolset called Workset Feature Comparison
5. Select a reference workset and a second workset that will be compared against the first to identify its characteristic vocabulary. The algorithm for this process is called Dunning's Log Likelihood ratio. It finds words that appear both more and less frequently in the second workset than in the first one. For example, you might look at what adverbs characterize the writing of Nathaniel Hawthorne compared with other authors from his period.
6. Look at the results, clicking on any of the words in the cloud to call up a concordance view where you can see where they occur in the workset.

Advanced Tips

- MONK has a variety of fairly complex toolsets. Try creating your own worksets, consisting of documents that interest you, and using different toolsets on them.
- Each toolset also has a variety of settings. Try changing the settings to get variations on the results.

Who has worked on creating it

MONK was developed with funding from the Mellon Foundation by more than three dozen researchers at seven universities. A complete list of contributors is available at www.monkproject.org

Where to read about it

Tanya Clement, "What to do, what to do, what to do with 3182 paragraphs: the nora project and repetition in Gertrude Stein's *The Making of Americans*," presented at the Chicago Colloquium on Digital Humanities and Computer Science, University of Chicago, November 2006.

Availability

Online at www.monkproject.org.

OJS Reading Tools

Teresa Dobson

OJS reading tools, designed by the Public Knowledge Project under the direction of John Willinsky, come in discipline-based sets with PKP's Open Journal Systems, which editors then select to appear with their online articles. The tools are designed "to support reading academic literature in online settings for a wider range of readers than has traditionally been the case" (Siemens, Willinsky, Blake, et al). OJS reading tools include features such as the following: view author biography, capture citation, view metadata, notify colleague, email author, add comment. As well, readers may access related readings such as books, book reviews, dissertations, journal articles, literary criticism, and articles in popular media sources. Many of these tools enable readers to input terms for a Boolean search and then offer a selection of databases from which that information might be harvested.

Example

This example from the *Journal of the Association for the Study of Australian Literature* shows the abstract of an article by David McCooney, "Poetry and Public Speech: Three Traces." To the right of the abstract are the OJS reading tools.

Getting Started

1. Locate an OJS journal in your area. A list of possible selections is attached.
2. To view the example provided above, visit <http://www.nla.gov.au/openpublish/index.php/jasal>.
3. Click on the abstract of an article in which you are interested.
4. The tools will appear to the right of the abstract.
5. Click on the tool you wish to use.
6. A new browser window will appear either with the information you have requested or with a search form that will enable you to either search using provided terms or input new terms.

Advanced Tips

- Some OJS journals do not make use of the reading tools. For example, *MHRA Working Papers in the Humanities* and *English Studies in Canada* are OJS journals; however, their editors have not elected to have the reading tools appear with their articles.
- When you are using OJS reading tools, keep track of the browser window in which they first appear. If you search, minimize rather than close the window, and later try to use another reading tool, you may get the impression the system is not responding to your request. In fact, you'll find the tool you're trying to use is being displayed in the browser window that appeared when you first used the OJS reading tools.
- If you are a journal editor wishing to move your journal online or better manage online publication, consider employing OJS. OJS is not a hosting service; rather, it is an open source content management system specifically designed for journals. You may download it free of charge and then arrange your own hosting.

Who has worked on creating it

Open Journal Systems was developed, and continues to be improved, by the Public Knowledge Project under the direction of John Willinsky (Stanford; UBC).

Where to read about it

Siemens, Ray, John Willinsky, Analisa Blake, Greg Newton, Karin Armstrong, and Lindsay Colahan. *A Study of Professional Reading Tools for Computing Humanists*. Victoria: Electronic Textual Cultures Laboratory. Available: http://etcl-dev.uvic.ca/public/pkp_report/

Availability

For information about the Public Knowledge Project and the open source (free) journal management and publishing software OJS, visit <<http://pkp.sfu.ca/>>. Below is a selected list of OJS humanities journals (note that not all of these journals use reading tools).

*BC Studies**Black Theology: An International Journal**Buddhist Studies Review**Bulletin of the Indo-Pacific Prehistory Association**Comparative Islamic Studies**Cosmos and History: the Journal of natural and Social Philosophy**Culture Society and Praxis**Dialogue & Discourse**Doomsday: Journal of the Thomas Lovell Beddoes Society**ESC: English Studies in Canada**Gender and Language**Glossator: Practice and Theory of the Commentary**Humanities Research Group Working Papers**International Journal of Mormon Studies**JASAL: the Journal of the Association for the Study of Australian Literature**Journal for the Orthodox Center for the Advancement of Biblical Studies**Journal of Lao Studies**Journal of Natal and Zulu History**Journal of the International Association of Buddhist Studies**Journal of the International Institute for Tibetan and Asian Studies**KRITIKE: An Online Journal of Philosophy**Latin American Theatre Review**Law and History Review**Left History**MHRA Working Papers in the Humanities**Museum Anthropology Review**Newfoundland and Labrador Studies**Nordic Journal of English Studies**Performance and Spirituality**Philippine Studies**Public History Review**Public Writing**Queensland Archaeological Research**Research Review of the Institute of African Studies**Reviews in Australian Studies**S: Journal of the Jan van Eyck Circle for Lacanian Ideology Critique**Semantics and Pragmatics**Spontaneous Generations: A Journal for the History and Philosophy of Science**Studies in Canadian Literature/Études en littérature canadienne**Sydney Journal**Sydney Studies in English**The Heritage Journal**The Journal of World Christianity**The Looking Glass: New perspectives in children's literature**Theology: An International Journal**Transformative Works and Cultures**University of Toronto Art Journal**Velox: Critical Approaches to Contemporary Film**Wagadu**Women in Judaism: A Multidisciplinary Journal*

Old Bailey Datawarehousing Interface

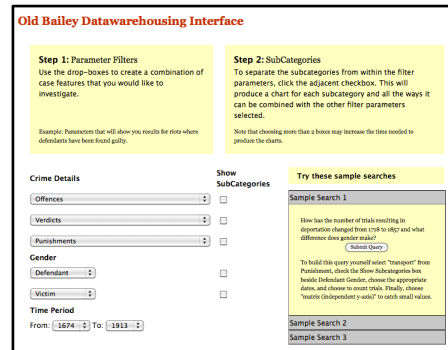
The Old Bailey Online is a database of 197,745 XML-encoded records of criminal trials held at London's central criminal court. The Old Bailey Database Interface provides a way for users to create customized visualizations of the Old Bailey site.

Example

You may be interested in when the verdict of guilt by reason of insanity became popular in London.

Getting Started

1. Point your browser to <http://ra.tapor.ualberta.ca/~digging2data/>.
2. Choose data that you would like to visualize by selecting options from the dropdown menus on the left.
3. Click on the "show subcategories" boxes to the right of the dropdown menus if you would like to create separate visualizations for the subcategories of the options shown in the dropdown menus.
4. You can also filter by words. Type a word into the "Filter Words" box in order to show only results that include that word in the text. There is a radio button that allows you to search for the exact forms of words as well as one to search for multiple words using "and/or."
5. "Track Words" allows for the comparison of instances of particular words as they occur in different trials.
6. Dropdown menus in the "Output" section allow you to select the way in which the results are visualized. Depending on the parameters selected above, certain options will produce more readable results than others.



Advanced Tips

- Read the text under each of the sample search questions in order to learn how to construct that kind of query.

Who has worked on creating it

The Old Bailey Data Warehouse Interface was programmed by John Simpson under the direction of Joerg Sanders and Geoffrey Rockwell as part of the Criminal Intent project. Interface design by Milena Radzikowska and implementation by John Simpson.

Where to read about it

At <http://criminalintent.org/>.

Availability

Online at <http://ra.tapor.ualberta.ca/~digging2data/>.

Orlando: Degrees of Separation

Stan Ruecker

Orlando contains a relatively large corpus, currently consisting of details about the life and writing careers of roughly 1000 British women writers, amounting to 6.8 million words with 2.2 million semantic tags for everything from paragraphs to politics, plots, or relations with publishers. The 6 degrees of connection project is an early experimental attempt to leverage this textbase for use by researchers who are interested in studying the network of connections between writers.

Example

To trace a connection between Ella Baker and Frances Harper through organizational affiliations, chose the two names using the controls on the left, then select the radio button for “organizations”. The proximity choice and exclusions list underneath the “Linked by” choices are there to reduce the number of connections that may not prove to be very useful. Most British women writers, for instance, spent some time in London, so connecting them by places other than London is likely to yield more interesting results.

Getting Started

1. Log in to the Orlando DoS link (you will need to obtain a userid and password by contacting Susan Brown).
2. Choose two writers.
3. Decide the settings you want to use for type of link, proximity, exclusions, and restrictions.
4. Examine the results, using the links provided to read the relevant passages.
5. Try varying the settings. Exclusions in particular may need some adjustment as you work.

Advanced Tips

- The Restrictions setting allows you to specify that the chain of connections should begin or end or both with the specified people.
- The default of close proximity typically works best, but for cases where there are fewer connections, it may be worthwhile loosening the restriction.

Who has worked on creating it

The Orlando Project was led by Patricia Clements, Isobel Grundy, and Susan Brown at the University of Alberta and the University of Guelph. The 6 degrees research team also includes Blair Nonnecke (Guelph), Claire Warwick (UCL), and Stan Ruecker (U of A).

Where to read about it

Brown, Susan, Jeffery Antoniuk, Sharon Farnel, Isobel Grundy, Stan Ruecker, Matt Patey, Stéfan Sinclair, and Milena Radzikowska. “Making Sense of Literary History: The Dense Associative Web of Orlando.” Paper presented at the third Digital Humanities and Computer Science (DHCS) Colloquium. Nov 1–3, 2008. Chicago: University of Chicago, 2008.

Some theory about connections is available in:

Barabási, Albert-László. *Linked: The New Science of Networks*. Cambridge, MA: Perseus Publishing, 2002.

Availability

Online, with account by arrangement with Susan Brown.

Orlando Mandala

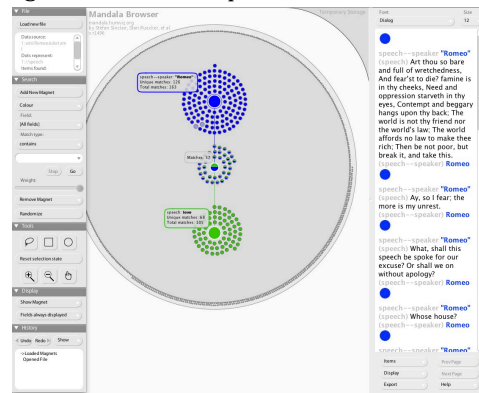
Orlando Mandala allows you to engage with Orlando Project materials in new ways, by looking at the ways patterns work across the whole set of author entries. Entries appear as dots around the periphery. You create colourful magnets and assign values to them, including values based on Orlando's semantic markup. These magnets draw the dots into the centre space. The text viewer to the right, allows you to see the matches for your magnets in context.

Example

You might be interested in the question: do women writers who have children write less than women writers without children? To address this question, you could load the Orlando collection, and make each dot represent a woman writer. Create magnets 0-9 representing numbers of children. You see immediately that there are diminishing numbers of women writers for each number of children, but there are some writers for every number. However, maybe the ones with more children wrote less. So you create new magnets dealing with measures of prolixity, by looking at the author summary tag, which summarizes a writer's achievements, for the words "prolific" or "numerous" or "many." It appears that by these measures, having more children equates to being more prolific. Now the third step is to access the documents, either by export or by logging into Orlando, to pursue further details.

Getting Started

1. Click on the appropriate icon to run Orlando Mandala.
2. Click on the Open File button
3. Select the Orlando XML file from the directory in which the program resides (it will have "orlando" in its name).
4. Using the "Dots Represent" picklist select //ENTRY
5. Click Load.
6. An empty magnet is preloaded. Set up the criteria for it and watch the dots gather to it. For a free-text search on all XML fields, type the search text in the box and hit return. To narrow the search, first select an XML field from the picklist above the search box. You can also choose to see a type-ahead response while typing by clicking on the checkbox called "Search as you type."
7. Create the next magnet, and fill in its criteria.
8. Iterate as desired.
9. Read items in the panel on the right by clicking them or lassoing them in the main circle.
10. Drag items to temporary storage if you find you do not want them in your results set.
11. Export items from the reading panel for further reading.



Advanced Tips

- You can use regular expressions to set up magnet values, allowing you to have several tags, for instance, attracted to a single magnet.
- You can also search on tag attributes.

Who has worked on creating it

Mandala was developed with funding from SSHRC by Stefan Sinclair, Anthony Sapp, Matt Patey (McMaster University), Stan Ruecker, Oksana Cheypesh, Constanza Pacher, Rhiannon Gainor (University of Alberta), and Sandra Gabriele (York University). Orlando Mandala was developed in partnership with Susan Brown and Jeffery Antoniuk. Mark Bieber has worked on the web version of Orlando Mandala that works through the Orlando API.

Where to read about it

Susan Brown, Stan Ruecker, Jeffery Antoniuk, Sharon Farnel, Matt Gooding, Stéfán Sinclair, Matt Patey, Sandra Gabriele. "Reading *Orlando* with the Mandala Browser: A Case Study in Algorithmic Criticism via Experimental Visualization." *Digital Studies / Le champ numérique* 2.1 (2010).

Brown, Susan, Patricia Clements, Isobel Grundy, Stan Ruecker, Jeffery Antoniuk, Sharon Balazs, Stéfán Sinclair and Matt Patey. "Thinking Beyond the Text: Using the Mandala Browser to Explore Orlando." Paper presented at the Society for Digital Humanities/ Société pour l'étude des médias interactifs annual conference at the 2008 Congress of the Social Sciences and Humanities, University of British Columbia. June 2-3, 2008.

Gainor, Rhiannon, Stéfán Sinclair, Stan Ruecker, Matt Patey, and Sandra Gabriele. [Forthcoming]. "A Mandala Browser User Study: Visualizing XML Versions of Shakespeare's Plays." *Visible Language* 43(1). 2009.

Ruecker, Stan. "Experimental Interfaces Involving Visual Grouping During Browsing." *Partnership: the Canadian Journal of Library and Information Practice and Research*. 1(1). 2006.

Availability

By arrangement.

Orlando: Women's Writing in the British Isles from the Beginnings to the Present

Stan Ruecker and Susan Brown

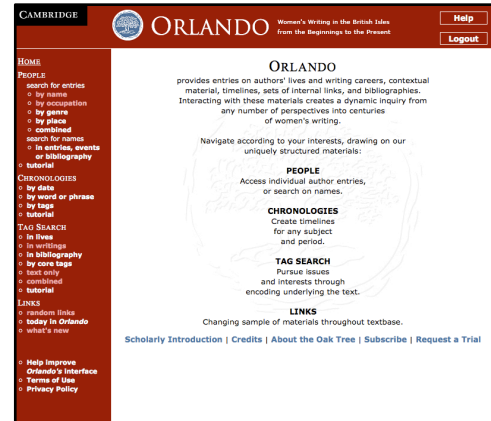
Orlando provides literary biographies of about a thousand British women writers, detailing their lives and their writing careers. It also contains some biographies of women writers who are not British, some men who were writers, and roughly 30,000 contextualizing historical events.

Example

Someone with an interest in the history of women writing ghost stories might begin by going to the link on the left nav bar called "Chronologies: by tags," choosing "ghost stories" from the picklist called "or by genre," and pushing the GO button. The search will produce 39 hits in chronological order from March 1796 to August 2008.

For a different set of results, try clicking on the left nav bar "People: by genre", picking "ghost stories" from the picklist called "by genre" and selecting Results Type: "excerpts" in the options box on the lower right. This search produces 95 results from writers' entries.

For the fullest results set, click in the left nav bar on "Tag Search: in writings", select "Genre" in the first box called "Search", then click on "select an attribute of this tag" right below the search boxes, select the radio button for "Genre Name" and select "Ghost Story" from the picklist. This produces 102 excerpts from writers' entries, plus the timeline of 39 events on a separate tab.



Getting Started

1. Log in to the Orlando project through the databases available at your university, or through your own subscription from Cambridge University Press.
2. Choose an entry point from the list in the nav panel on the left.
3. Set up your search criteria using the fields available. These differ for each of the entry points.
4. Narrow your search using the additional palettes on the right (displayed in a lighter shade of blue).
5. Pursue the links involved in the search results.

Advanced Tips

- The most sophisticated searching of Orlando is found under the "Tag Search: combined" link in the left nav panel.
- Some of the tutorials from the left nav panel include an optional link back to automatically fill in the search criteria, making it easy to see how the examples were constructed.

Who has worked on creating it

The Orlando Project was led by Patricia Clements, Isobel Grundy, and Susan Brown at the University of Alberta and the University of Guelph.

Where to read about it

Brown, Susan, Isobel Grundy, Patricia Clements, Renée Elio, Sharon Balazs, Rebecca Cameron. "Intertextual Encoding in the Writing of Women's Literary History." *Computers and the Humanities* 38 (2004): 191-206.

A more complete list is available here: www.ualberta.ca/ORLANDO/publications.htm

Availability

By online subscription from Cambridge University Press. Trial user accounts are also available in connection with the Orlando user study by emailing orlando@ualberta.ca.

OVis

Susan Brown

OrlandoVision (or OVis) is an experimental, interactive link-node graph designed for exploring the Orlando textbase. It enables the user to visualize the writers with entries in the Orlando textbase and the more than 30,000 people associated with them. OVis represents the connections between people as nodes and the relationship between them as edges, and enables the user to read about the context of the link through the text reading pane. The links or edges are colour-coded, which corresponds to the colouring of the particular semantic tag that indicates the context in which the association is mentioned, such as <politics>, <textualfeatures>, or literary <responses>. The tool also has a multi-functional search capacity which enables users to navigate large swaths of data at once. Users can learn about the connections between individuals in the Fabian Society, for instance, by searching for "Fabian" and tracing the links between the nodes (people) who appear on the graph.

Example

Someone wishes to look at all people connected to the "poetess" movement who were alive during the early Victorian period.

Getting Started

Open and Load

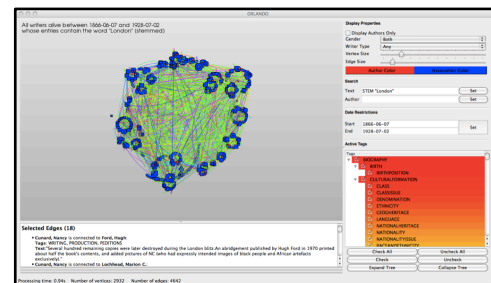
1. Double-click the Ovis icon on the desktop.
2. Under "File," select "Open Data."
3. Navigate through the folders by clicking on "Desktop," "CWRC," "ovis," and "DATA" until you find a file called "orlando_entries_all." Double-click it.
4. Wait briefly for the dataset to load.

Navigate

5. In order to navigate within the graph itself, you will need a conventional mouse.
6. To zoom into and out of the graph, use the scroll wheel, or right click and drag.
7. To move the entire graph, click the scroll wheel, hold, and drag into desired position.
8. To highlight (or select) a node or edge, left click on a node or edge.
9. To highlight multiple nodes or edges, press down the shift key and left click additional nodes and edges, or left click and drag a rectangular lasso.
10. Once a node is highlighted, the person's name will appear in the text pane at the bottom of the screen.
11. Once an edge is highlighted, the text extracted from the Orlando database will describe the context of the link in the text pane.
12. If you select multiple edges, the text extracted from the Orlando database will describe all the links you have selected. This text will appear in the scrollable text pane. The most recent links you have selected will appear at the top of the text pane.
13. To deselect nodes and edges, left click any empty space in the graph.

Search

14. In OVis, you can search using any of the following search options: "Text", "Author" (writer), or "Date Range".
15. To use the "Text" search feature, click on "Set" next to "Text" on the right-hand side of the screen.
16. Click "Add" in the window that pops up.
17. Double-Click in the empty box under "Search Term."
18. Type a search term in the text box.
19. To enter multiple search terms, click "Add" and enter another search term.
20. Click the box under "And/Or" to select "And" or "Or".
21. Click the box under "Stem" to select whether to stem your search term and include all variations of that term.



22. Select whether you want your search to be "broad" (showing all writers with entries in the Orlando database that mention the search term AND all other people to whom they are connected) or "narrow" (showing only the entries in the Orlando database that mention the search term).
23. Click "Remove" to remove the bottom search term.
24. Click "Cancel" to cancel search.
25. Click "OK" to set your search.
26. To use the "Author" (writer) search feature, click on "Set" next to "Author" on the right-hand side of the screen.
27. Click "Add" in the window that pops up.
28. Double-Click in the empty box under "Search Term."
29. Type an author name in the text box.
30. To use the "Date Range" search feature, click on "Set" under "Date Range" on the right-hand side of the screen.
31. A window will pop up.
32. Choose one of the options under "Historical Dates" or "Monarch Dates". DO NOT select both of these options.
33. Alternatively, enter a Start year into the text box below "Start" and "Year (1+)", select a Month and Day, and enter a End year into the text box below "End" and "Year (1+)", and select a Month and Day.
34. Click "OK."

Limit Results

35. Limit your results in OVis by selecting or deselecting the following options from the right-hand side of the screen below "Display Properties": "Visible Nodes", "Writer Sex" and "Writer Type".
36. Results may also be limited by selecting or deselecting Tags, on the lower right-hand side of the screen below "Active Tags".
37. To select all tags, click the "Check All" button.
38. To deselect all tags, click the "Uncheck All" button.
39. To select or deselect a particular tag, click the box to check or uncheck that tag.
40. Additional limiters are available through the "Search" function.

Modify Layout

41. To change the size of nodes and links (edges), use the sliders on the top right of the screen.
42. To change the node colour, click "writers with entries" or "other people". Select colour from the colour wheel. Click "OK".
43. To change the node style, click on the "View" menu and navigate to "Set Node Style", which has more customization options.
44. To change the background colour, click on the "View" menu and navigate to "Set Background", which has more options.
45. To change the text that appears in the upper left hand corner of the screen, click on the "View" menu and navigate to "Set Annotation". A text box will appear. Write in your desired annotation.
46. To eliminate the text in the upper left hand corner of the screen, select or deselect "Show Annotation" in the "View" menu.
47. To see different ways to display the graph, click on the "View" menu, navigate to "Set Layout Strategy," and select one of the other options. (**For a more detailed exploration of the Layout Strategies, see "Features" below.)

Save and Reload

48. To save a copy of your OVis Session, select "Save Session As" from the "Session" menu.
49. Write in the name of your session in the textbox that appears.
50. Select a folder in which to save your session.
51. Select the "Save" button.
52. To save subsequent changes to your session, select "Save Session" in the "Session" menu. This will save any subsequent changes you make to the session under your original session name.
53. To revert to your saved session once you have made changes to your graph that you wish to remain unsaved, select "Reload Session" in the "Session" menu.

54. Once you are finished with your session for a time, quit OVis.
55. To reload your previous session, double-click the OVis icon.
56. Under "Session," select "Open Session."
57. Navigate through the folders until you find the file you saved. Double-click it.
Take a Screenshot
58. To take a screenshot of your graph, select File from the menu and select Screenshot.
59. Save your screenshot where you want it.
60. To make the resolution of your screenshot higher, select "view" and then "Set Snapshot Magnification". Select desired magnification (from 1-10).
61. This screenshot feature will only take a picture of the graph itself, without the OVis frame.
62. To take a screen shot of the entire screen, press CTRL+SHIFT+4 on a Mac Keyboard, or the PrtSc button on a PC.

Advanced Tips

Layout Strategies

- The layout strategies available in Ovis can be divided into two categories: Force-Directed and Non Force-Directed.

Force-Directed

- In the case of the Orlando dataset, Force-Directed layouts will generally prove to be more useful, at the cost of requiring more time to produce results.
- Force-Directed layouts use a physics based approach to arrange the nodes of a graph. The general goal of a force-directed layout is to attempt to ensure that the edges, i.e. the connections between nodes, are similar in length, and that they do not cross. However, given the structure of the graph this may not always be possible.
- The best way to understand a force-directed layout is to think of the edges of a graph as though they were springs, and the nodes as though they were electrically charged particles. The nodes repel each other, while the springs attract. The interplay of forces continues until eventually the system reaches a point of relative equilibrium. At this point the graph is drawn using these positions.
- For the most part, the force-directed layouts described below tend to represent the nodes with their most edges between them as being closer together in the graph. This is because the more edges between nodes, the stronger the attractive force between the nodes. This is particularly true of the "Simple 2D" and "Force-Directed" Graphs, which may prove the most useful for the Orlando dataset, at the cost of requiring more time to produce results.
- For a more detailed explanation, please see: http://en.wikipedia.org/wiki/Force-based_algorithms_%28graph_drawing%29

Simple 2D

- This layout uses the Fruchterman-Reingold algorithm (<https://nwb.slis.indiana.edu/community/?n=VisualizeData.Fruchterman-Rheingold>). Like the ForceDirected algorithm, it is fairly slow. A unique feature of this algorithm is that it uses an overall "temperature" in order to guarantee that the system will reach an equilibrium. This algorithm tends to produce graphs which are more dispersed than the ForceDirected algorithm.

Force-Directed

- This layout uses the Kamada-Kawai algorithm. It is a slow, force-directed layout algorithm and tends to produce graphs which have more clusters than the Simple2D algorithm.

Clustering2D, Community2D, and Fast2D

- These three layout algorithms are all fairly similar. While Clustering2D tends to produce a graph with more crossed edges, Community2D and Fast2D produce almost identical results. These algorithms create graphs much more quickly than the previous two layout algorithms, because they use a density grid which provides a random initial rough layout of the graph.

Non Force-Directed

- Non Force-Directed layouts tend to be more random than Force-Directed layouts. This is particularly true of the "Random," and Circular layouts, which both place nodes within the boundaries of a graph in a random pattern.

Random

- Random is an extremely fast layout algorithm. As might be expected, this layout randomly places nodes within the boundaries of the graph. It is therefore not very helpful for displaying meaningful connections; however, due its speed it could be useful on slower computers and/or when working with a large number of nodes.

Circular

- This is a fast layout algorithm which arranges nodes around the circumference of a circle. Since there is no logic applied to the positioning of the nodes, it is not very useful for displaying meaningful connections.

Span Tree

- This is a fast layout algorithm which arranges nodes into a series of circular spanning trees. A spanning tree can be defined as the minimum number of edges that connect all the nodes of a graph. In the case of the Orlando dataset, we are usually working with a spanning forest, which consists of multiple, separate spanning trees. This layout is best suited to sparse graphs, wherein the number of edges is about the same as the number of nodes. For more information on spanning trees please see: http://en.wikipedia.org/wiki/Spanning_tree

Who has worked on creating it

Susan Brown guided the prototype design from a literary historical perspective, Michael Bauer directed the technical work, Jennifer Berberich and Jonathan Cable did the coding of the first prototype, and Patrick D. Emond coded the second prototype. Jana Smith Elford did significant testing and drafted the documentation. Work on OrlandoVision has been generously supported by the Sharcnet High Performance Computing Consortium, which provided programming resources, by the Social Sciences and Humanities Research Council of Canada, and by the University of Alberta's Faculty of Arts.

Where to read about it

Smith Elford, Jana, Susan Brown, Michael Bauer, Jennifer Berberich and Jonathan Cable. "Elevating Influence": Victorian Literary History by Graphs." *The Victorian's Institute Journal Digital Annex*. 1 (1). 2011.

Availability

By arrangement.

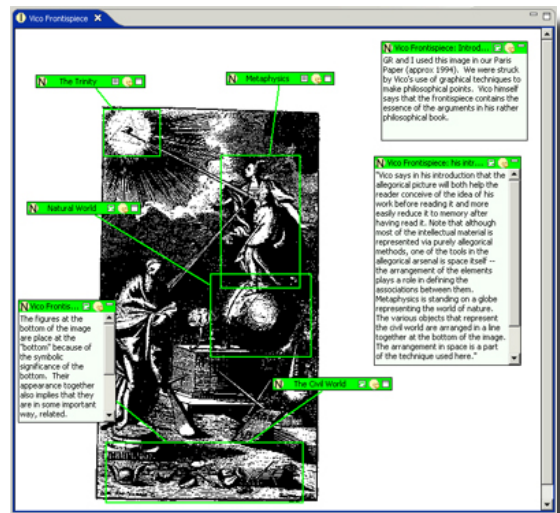
Pliny

Teresa Dobson

Pliny is free downloadable software that facilitates note-taking and annotation. It allows readers to integrate their initial notes into a representation of an evolving personal interpretation and can be used with both digital and non-digital materials, such as websites, images, and pdf files. According to the Pliny website, this application is designed “assuming that you might well accumulate many notes (perhaps even thousands) over time, and that your note collection will then provide a base platform upon which you might develop some sort of interpretation of the materials they refer to. Thus, Pliny has components in it that go beyond annotation facilities to help you manage the notes and organise them, even if you have thousands of them to work through.”

Example

Several examples of Pliny “in action” are provided on the screenshots page of the Pliny website. The one duplicated here shows a user working with an image – the frontispiece of Vico's *New Science*. Some of the annotations are shown closed and some are shown open. Pliny comes equipped with a wide range of tools for finding and bookmarking things both online and on your computer, for making notes about them and links between them, and for importing and exporting text and images. It comes with its own built-in web browser.



Getting Started

1. Visit < <http://pliny.cch.kcl.ac.uk/> > to read about and download Pliny.
2. Detailed instructions for downloading Pliny for Windows or Macintosh are linked from the welcome page.
3. Pliny is an “Eclipse” application. You may find it is unlike other software with which you’ve worked, and therefore it is useful to read the help section and review the tutorials before.
4. If you’ve started the standalone version, the welcome page within the application obscures the program. Close it to begin.
5. Choose icons from the list across the top to launch the browser, find items, and make notes about them.

Advanced Tips

- The containment view provides a powerful way to visually represent which items refer to others. It can be used to help track citations, for example.
- Pliny is still under development and may contain bugs. According to the developer, Pliny has been tested most and is less buggy in a Windows environment.

Who has worked on creating it

Pliny is a project of John Bradley, Center for Computing in the Humanities, King's College London.

Where to read about it

Bradley, John and Paul Vetch. Supporting Annotation as a Scholarly Tool--Experiences From the Online Chopin Variorum Edition. *Literary & Linguistic Computing* (22:2) June 2007, 225-241.

Availability

Pliny is available for download here: <http://pliny.cch.kcl.ac.uk/>

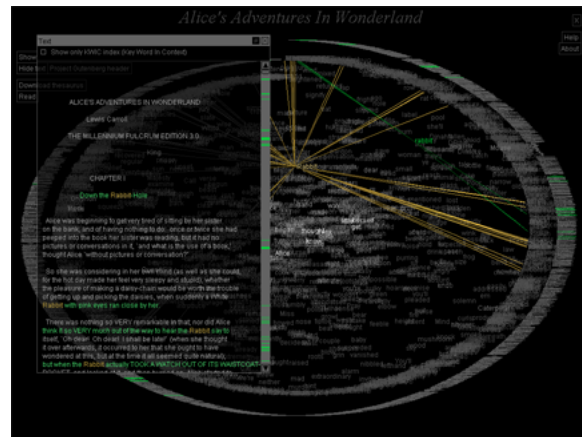
TextArc

Teresa Dobson

TextArc, created by interaction designer W. Bradford Paley, is described on the TextArc “Applications” page as “a Structuralist text analysis tool to show distribution of words in texts that have no ‘meta-data’ descriptions, such as a table of content or an index.” This text visualization tool draws the text line-by-line (and then again word by word) in a minute font around the outside of the screen. The arc provides a sense of global text structure. Words that appear more than once are drawn within the arc in their average position, and mousing over these words displays a series of lines radiating out to places where the word occurs in the text. A pop-up window offers full-text or KWIC displays. There are also concordance and association features, as well as a nifty reader that traces an electronic reading of the text, looping around the screen from word to word as it does so, and at the same time displaying sequential lines of text at the bottom of the screen.

Example

There are a number of screen shots on the TextArc website (see Still Screen Images). This visualization of *Alice in Wonderland* is one of the stills in the TextArc collection. In this example, the word “Rabbit” is highlighted and lines radiate outward showing where the word appears in the text. The overlay full-text window provides the keyword in context.



Getting Started

1. Visit <http://www.textarc.org/> to read about TextArc and experiment with the application.
2. *Hamlet* and *Alice and Wonderland* are offered for the purpose of such experimentation.
3. There is also a portal enabling access to texts at *Project Gutenberg*. (Note, however, that *Project Gutenberg* texts contain a great deal of metadata that will be reflected in your visualization.)

Advanced Tips

- The text on the “Applications” page of the TextArc site gives an overview of how this visualization tool might assist you in interpreting a text.

Who has worked on creating it

TextArc is a personal project of interaction designer, W. Bradford Paley, who, beyond his independent design work is an occasional lecturer at Columbia University. Credits are listed here: <http://www.textarc.org/ThoseInvolved.html>.

Where to read about it

Emerson, Lori. "Digital Poetry as Reflexive Embodiment." *Cybertext Yearbook*, Markku Eskelinen and Raine Koskimaa (eds.), 88-106. Saarijärvi: University of Jyväskylä, 2003. Available:

<http://epc.buffalo.edu/authors/goldsmith/emerson.pdf>

Paley, Bradford W. "TextArc: Showing Word Frequency and Distribution in Text." Poster presented at IEEE Symposium on Information Visualization 2002 (InfoVis 2002), October 28-29, 2002.

Available: http://www.textarc.org/appearances/InfoVis02/InfoVis02_TextArc.pdf

Mirapaul, Matthew. "Secrets of Digital Creativity Revealed in Miniatures." *New York Times*, 16

September 2002. Available: <http://www.nytimes.com/2002/09/16/arts/design/16ARTS.html>

Availability

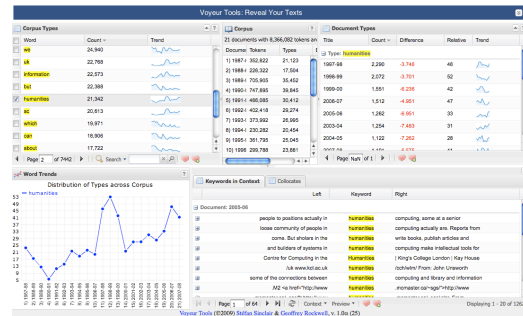
TextArc is available here: <http://www.textarc.org>

Voyant

Voyant is currently a beta release by Stéfan Sinclair and Geoffrey Rockwell. Voyant is the next generation in a series of text analysis tools that include HyperPo and Taporware. It provides tables and graphs related to word use across a single document or a collection. Voyant adds, among other things, the ability to handle much larger files than the previous tools could.

Example

Someone has a collection consisting of the entire archive of the Humanist list. It takes only a few clicks of the mouse to see, for instance, that the word “information” saw a peak usage in 1998/9 and that this peak corresponds to a similar one for the word “humanities.”



Getting Started

1. Go to <http://voyeurtools.org/>
2. Provide a URL or paste in some text.
3. Explore the various options for viewing the text analysis results.
 - a. For example, read the text in the "Summary" box, and click on interesting links.
 - b. Click on the "gear" button at the top of the box and choose the Taporware stop word list from the dropdown menu, and then click "OK." Note how common words disappear from the results.
 - c. You might also try clicking on a word in the Cirrus word cloud in the upper left. Then click on a point in the graph that appears under "Word Trends" on the right.
 - d. Also note that clicking on the title bars for "Words in Documents" and "Words in the Entire Corpus" reveals new tools.
4. Note that the column headers have explanatory rollover text, and there are also help icons scattered around the interface.

Advanced Tips

- It is important to remember that this is a beta release, which means that there will be bugs. Don't trust everything you see without stopping to consider that it might be a mistake of some kind. As Stéfan says, “expect the unexpected.”

Who has worked on creating it

Voyant is the product of Stéfan Sinclair and Geoffrey Rockwell.

Where to read about it

No papers have been published at this point, although the journal *Literary and Linguistic Computing* has published many articles that show the kind of work that is possible with a tool like Voyant. Here is one example:

Fink, Peter. “The Evolution of Order in the Chapter Lengths of Trollope's Novels.” *Literary & Linguistic Computing* (21:3) Sep 2006, 275-282.

Availability

Freely available online at <http://voyeurtools.org/> or else <http://voyeur.hermeneuti.ca/>

Voyant Word Trends

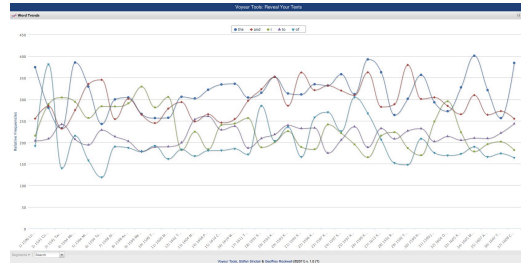
Term Frequencies Chart shows how terms are distributed across document(s) in a corpus (documents are shown in the order in which they were added). Once you arrive to Term Frequencies Chart, insert / upload your content and let the tool perform its analysis. Every charted line represents one word common throughout the entire corpus. If you hover over specific points it will give you specific information about that word in a specific document.

Example

Find the distribution of the word “love” across a corpus.

Getting Started

- Go to <http://voyeurtools.org/tool/TypeFrequenciesChart/> to load a document. Alternatively, use a preloaded corpus, and then skip ahead to step 4.
- Provide a URL of a document you would like to visualize, or paste in some text. You can also upload your own document or open one of the included corpora by clicking the appropriate buttons.
- Click “Reveal.”
- When you analyze a corpus with Term Frequencies Grid, you will initially have common words at the top of the chart with colour codes. You will see lines within the graph which are coloured accordingly to those words. If you click on one of the terms at the top, it will omit that term from the graph.
- When we hover over the segment points, we can see the frequency of that term in that segment. If you click on the point, Voyant will open a new window with detailed information of that segment and term within its Document KWICs tool.



Advanced Tips

- If you would like to see less or more segments on the chart, simply click on “Segments” at the bottom left of the chart to choose the desired segments.

Who has worked on creating it

Voyant is the product of Stéfán Sinclair and Geoffrey Rockwell.

Where to read about it

No papers have been published at this point, although the journal *Literary and Linguistic Computing* has published many articles that show the kind of work that is possible with a tool like Voyant. Here is one example:

Fink, Peter. “The Evolution of Order in the Chapter Lengths of Trollope's Novels.” *Literary & Linguistic Computing* (21:3) Sep 2006, 275-282.

Availability

Freely available online at <http://voyeurtools.org/tool/TypeFrequenciesChart/>

Voyant Bubblelines

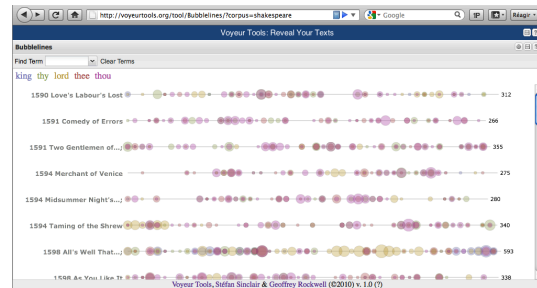
Bubblelines is a visualization tool that helps to understand patterns of word repetition in one or more documents. Each document is represented as a horizontal line and each search term is represented as a bubble – the bubble represents the frequency of the term in the corresponding segment of text (the text is divided into segments of equal length). The larger the bubble, the more frequent the term.

Example

You wish to find how often a character's name is mentioned throughout a story.

Getting Started

1. Go to <http://voyeurtools.org/tool/Bubblelines/> to load a document. Alternatively, use a preloaded corpus, and then skip ahead to step 4.
2. Provide a URL of a document you would like to visualize, or paste in some text. You can also upload your own document or open one of the included corpora by clicking the appropriate buttons.
3. Click "Reveal."
4. Hovering over a bubble, or set of bubbles, will cause a box to appear that displays the frequency counts for that segment of text.
5. Similarly, hovering over the number at the end of the line will cause a box to appear that summarizes the frequency for the entire document.
6. When Bubblelines first loads a corpus, you may see terms that have been pre-selected and included in the URL or embedded page. If no terms are specified, Bubblelines automatically fetches the five most frequent terms and displays bubbles based on those. You can remove the default terms by clicking on the "Clear Terms" button.
7. You can add additional terms to be displayed using the "Find Term" box. Note that available terms will appear as you type and you can pick an item from the list to have it added.



Advanced Tips

- In addition to adding and removing terms, you can toggle the display of the terms that have been loaded. To do so simply click on the term (active terms are underlined).

Who has worked on creating it

Voyant is the product of Stéfán Sinclair and Geoffrey Rockwell.

Where to read about it

No papers have been published at this point, although the journal *Literary and Linguistic Computing* has published many articles that show the kind of work that is possible with a tool like Voyant. Here is one example:

Fink, Peter. "The Evolution of Order in the Chapter Lengths of Trollope's Novels." *Literary & Linguistic Computing* (21:3) Sep 2006, 275-282.

Availability

Freely available online at <http://voyeurtools.org/tool/Bubblelines/>

Voyant Scatterplot

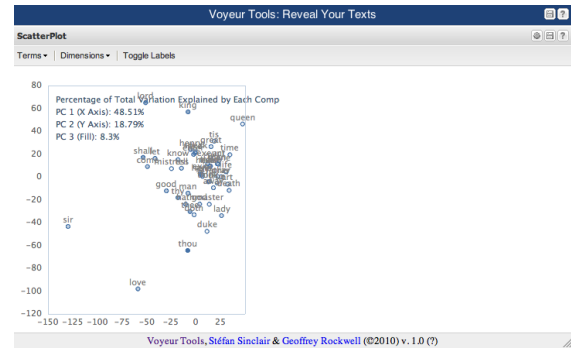
ScatterPlot creates a scatter plot graph of terms, spaced by their variation from one another. Once you arrive to ScatterPlot, insert / upload your content and let the tool perform its analysis. You may hover over these dots and click on them for more information.

Example

You wish to view variance between the words “king” and “queen” in a play.

Getting Started

1. Go to <http://voyeurtools.org/tool/ScatterPlot/> to load a document. Alternatively, use a preloaded corpus, and then skip ahead to step 4.
2. Provide a URL of a document you would like to visualize, or paste in some text. You can also upload your own document or open one of the included corpora by clicking the appropriate buttons.
3. Click "Reveal."
4. When you first load ScatterPlot, you will see a variety of terms plotted on a graph. If you hover over the terms, you will see their variation explained by each component on the x and y axis.
5. If you click on any of the terms, it will bring you to the Document KWICs tool for further analysis.



Advanced Tips

- ScatterPlot offers options for changing the plot. The terms button allows you choose how many terms should be displayed. The dimensions button lets you switch between a two or three dimensional graph. Toggle labels simply removes or adds labels for the terms on the graph. You can also add terms by using the input box.
- Click on the "gear" button at the top of the box and choose the Taporware stop word list from the dropdown menu, and then click "OK." Note how common words disappear from the results.

Who has worked on creating it

Voyant is the product of Stéfan Sinclair and Geoffrey Rockwell.

Where to read about it

No papers have been published at this point, although the journal *Literary and Linguistic Computing* has published many articles that show the kind of work that is possible with a tool like Voyant. Here is one example:

Fink, Peter. "The Evolution of Order in the Chapter Lengths of Trollope's Novels." *Literary & Linguistic Computing* (21:3) Sep 2006, 275-282.

Availability

Freely available online at <http://voyeurtools.org/tool/ScatterPlot/>

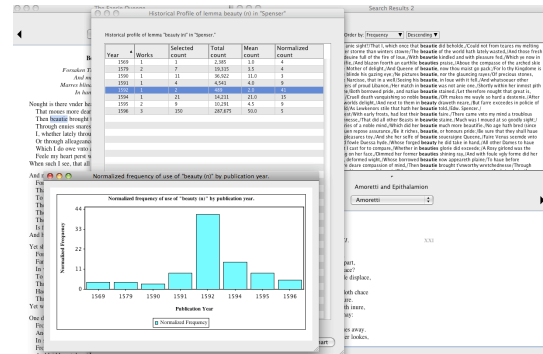
Wordhoard

Stan Ruecker

Wordhoard is a philological tool. In the words of its creator Martin Mueller, Wordhoard is all about “going from the **word** here to the **words** there.” Using Wordhoard, anyone interested in Early Greek Epic, Chaucer, Shakespeare, and Spenser can quickly access information about the use of words, lemmas, or parts of speech as they have been used across all of the poems, all of the plays, or any subset of interest.

Example

In Spenser’s *The Faerie Queene*, beauty is a recurring theme. One way to examine it is to start looking through the reading view, then selecting the word “beautie” by clicking on it and hitting command-F (for Find Word). This calls up a new window showing the lemmatized search for beauty, which yields 205 results, not just in the *Faerie Queene*, but across all the works by Spenser. They can be grouped in various ways, and examined individually using a keyword in context view. Under the “Analysis” menu, the option “Track Word Form over Time” allows the user to see a table of use of the noun “beauty” by year. A chart is available by clicking a button at the bottom right of the screen. The chart shows the danger of normalized counts, suggesting that Spenser used “beauty” a lot in 1592. The raw counts show that, on the contrary, he just used it twice, but the total word count for that year was only 489 words. More interesting is the following year, where the normalized count is still high (at 15 words per 10,000) but the total word count is more than 14,000 words.



Getting Started

1. Wordhoard is a Java web start, which means it doesn’t run through your browser. Instead, you have to download it and run it locally on your machine.
2. Go to <http://wordhoard.northwestern.edu/userman/index.html> and click the link called “Download and run Wordhoard.”
3. Choose the Main Public Site to get copies of the files. Users at Northwestern also have access to two other collections. It is also possible to add new files once they have been properly prepared.
4. Select options from the menus. Under the “Analysis” menu, for instance, you can select “Create Word Form List” which will give you a summary table of all the words in a given work, or under the “Find” menu you can perform a “Find Words”.

Advanced Tips

- There are a variety of analysis methods available in Wordhoard, described here: <http://wordhoard.northwestern.edu/userman/analysis-intro.html>

Who has worked on creating it

Wordhoard is the brainchild of Dr. Martin Mueller at Northwestern University.

Where to read about it

<http://wordhoard.northwestern.edu/userman/whatiswordhoard.html>. Anyone interested in philological tools in general and Wordhoard in particular is well advised to read through the rich list of materials available here. Stéfan Sinclair provides an encomium: <http://stefansinclair.name/node/54>

Availability

Freely available online at <http://wordhoard.northwestern.edu/userman/index.html>

Wordle

Teresa Dobson

Wordle, created by Jonathan Feinberg at IBM Research, is described as "a toy for generating 'word clouds' from text that you provide." This online site allows you to upload your own data or input the URL of a website. Wordle generates a word cloud of your source text in which greater prominence is given to words that appear more frequently. You may generate multiple word clouds of the same source text in different colours and fonts using either a "randomize" option or by using the formatting menu to modify language, font, layout, and colour. This is not a robust research tool; however, there are multiple classroom applications at any level of education.

Example

The image at right is a Wordle of the above paragraph.



Getting Started

1. Note first of all that Wordle is a public site, so any data that you post will be visible to the world.
2. Visit <http://www.wordle.net/> to read about Wordle and to create your own.
3. Select "create" to get started and follow the directions.
4. Once you have created a wordle you may modify it using the formatting menu or generate a different wordle of the same text using the "randomize" button. (Note, however, that after clicking the "randomize" button you will not be able to retrieve your previous Wordle unless you've saved it to the public gallery.)

Advanced Tips

- You may use screen capture to save your Wordle to your local drive for your own uses.
- Images created by Wordle.net are fall under the Creative Commons license. They may be used anywhere; however, they should be attributed to <http://www.wordle.net/>.

Who has worked on creating it

Wordle is a personal project of Jonathan Feinberg at IBM Research. Credits are listed here: <http://www.wordle.net/credits/>.

Where to read about it

Clement, T., Plaisant, C. and Vuillemot, R. (November 2008). The Story of One: Humanity scholarship with visualization and text analysis. HCIL Research Report. Available: <http://hcil.cs.umd.edu/trs/2008-33/2008-33.pdf>

Wattenberg, M. and Viegas, F. (2008). "Tag clouds and the case for vernacular visualization." *Interactions*, 15.4: 49-52.

Availability

Wordle is available here: <http://www.wordle.net>

There is also a Wordle available on the ManyEyes site.

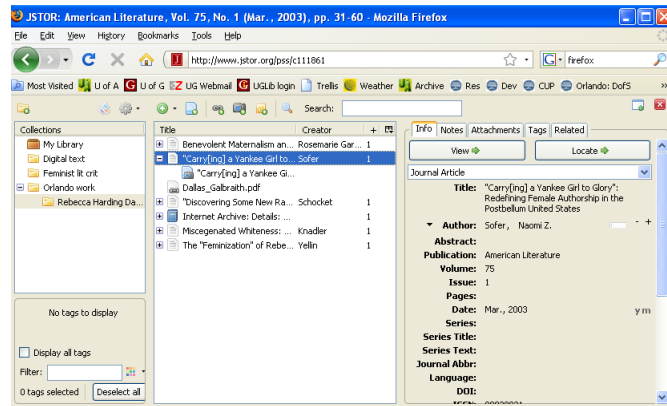
Zotero

Teresa Dobson

Zotero is a Firefox web browser extension and standalone application that allows you to collect, manage, cite and annotate research resources. The next release (now in beta) includes social networking features designed to support organizations and communities of scholars.

Example

Someone who has been using proprietary bibliographical management software wishes to move to a more web-friendly and integrated environment, with a view to being able to manage citations more effectively and share them more easily with collaborators and students.



Getting Started

1. Download Zotero at www.zotero.org.
2. To install, read the documentation at <http://www.zotero.org/support/installation>.
3. If you are using Zotero in Firefox, click the Zotero button in the lower right-hand corner of the browser window.
4. Click the left-most button with the icon of a folder and a plus sign to begin a new collection.
5. Add items to your collection in a range of formats: web pages, citations from catalogs, manual citations, sets of articles found through JSTOR or other databases. Try using it to download pdfs of scholarly articles as well as web pages.
6. Drag files from your computer into Zotero so that you can group them with your collections.

Advanced Tips

- Once you have some sites in place, try the Create TimeLine tool available under the Gear icon on the horizontal menu of the Zotero window.
- Download the Word or Open Office plugins to try the citation and bibliography generation features.
- You can try downloading large datasets from RefWorks, Endnote, or other bibliographical software to see how it behaves in Zotero. There are easy-to-follow instructions in the Zotero discussion forums.
- Try social networking features by creating and joining groups.

Who has worked on creating it

Centre for History and New Media, and George Mason University.

Where to read about it

<http://www.zotero.org/>

Daniel J. Cohen, "Zotero: Social and Semantic Computing for Historical Scholarship." *Perspectives*, American Historical Association, May 2007.

<http://www.historians.org/perspectives/issues/2007/0705/0705tec2.cfm>

Availability

Free download at <http://www.zotero.org/>