

UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE

AUGMENTING FACETED BROWSING OF
THE ISIS BIBLIOGRAPHY OF THE HISTORY OF SCIENCE
WITH SMALL MULTIPLE TIME SERIES VISUALIZATIONS

A THESIS
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfilment of the requirements for the
Degree of
MASTER OF SCIENCE

BY
BHAVANA GOPALACHARY

Norman, Oklahoma

2021

AUGMENTING FACETED BROWSING OF
THE ISIS BIBLIOGRAPHY OF THE HISTORY OF SCIENCE
WITH SMALL MULTIPLE TIME SERIES VISUALIZATIONS

A THESIS APPROVED FOR THE
GALLOGLY COLLEGE OF ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr Chris Weaver

Dr Christan Grant

Dr Kash Barker

© Copyright by Bhavana Gopalachary 2021
All Rights Reserved.

ACKNOWLEDGEMENTS

I am very thankful to my thesis advisor Dr.Chris Weaver for giving me the opportunity to explore various things during my research, for always giving me the push in the right direction when needed and for his invaluable expertise throughout my research. I would also like to thank my committee members Dr.Christian Grant and Dr.Kash Barker for their time and effort throughout the thesis process.

I am deeply grateful to Dr.Stephen Weldon, for providing us with a wonderful platform like Isis Current Bibliography to work with and his constant support and expertise. I would also like to thank Julia Damerow for always being ready to help me with any questions I had. I would also like to thank my research group members Omkar Chekuri and Yonathan Hendrawan for always being interested and helpful in my research.

Most importantly, I am especially thankful to my parents and my sister for being such an amazing support system and encouraging me over the last few years.

Contents

1	Introduction	1
2	Background and Related work	4
2.1	Bibliographic Databases and Search Engines	4
2.2	Information Interfaces	5
2.3	User Interface Types	7
2.3.1	Web-based user interface	7
2.3.2	Graphical user interface	8
2.3.3	Faceted search user interface	9
2.4	Visualization	9
2.4.1	Small Multiple Visualizations	12
2.4.2	Time Series Visualizations	12
2.5	History of IsisCB and the IsisCB Explore Architecture	15
2.5.1	History of IsisCB	15
2.5.2	IsisCB Explore	17
2.5.3	Architecture	17
3	Discovery Study	21
3.1	User Study	21
3.2	Study Results and Analysis	23
4	Design and Implementation	26
4.1	Design Space	26
4.2	Design Specifications	28
4.2.1	Color	30
4.2.2	Interaction	31
4.3	Implementation	32
4.3.1	Initial Setup	33
4.3.2	Process/Procedure	33
4.3.3	Data Querying	34
4.3.4	Data Preparation	38
4.3.5	Visualization Development	40
4.3.6	Performance Analysis	44
5	Conclusion	49

List of Figures

1	Minard’s map of Napoleon’s 1912 march on Russia [27]	10
2	William Playfair’s <i>The Commercial and Political Atlas</i> , 1786 Source: https://commons.wikipedia.org/w/index.php?curid=42808916	11
3	Number of publications of an author over several years.	13
4	Nightingale’s Diagram of the Causes of Mortality[8]	14
5	The main page of the IsisCB Explore website	16
6	The IsisCB Architecture.	19
7	Participants Statements	23
8	Participants Accuracy Chart	25
9	The Authority page for Newton	27
10	Aigner’s example representation of Linear data	29
11	Aigner’s example representation of interval-based data	29
12	Color ranges considered	31
13	Stages of the implementation process	34
14	Screenshot of part of the Django code	35
15	Corresponding PostgreSQL code	36
16	Screenshot of an Authorities data model	37
17	Screenshot of a Citations data model	38
18	Screenshot of an ACRelations data model	39
19	Screenshot of the Query	40
20	Screenshot of part of the python script	40
21	The first version of our design	41
22	The second version of our design	41
23	Screenshot of Related Contributors facet with the tool-tip	42
24	The new IsisCB version after the addition of visualizations	43
25	The circles variation of our design	44
26	Snapshot of the large pop up for the Related Journals facet	45
27	Snapshot of the small pop up for one of the Related Authors and Con- tributors facet value	46
28	The updated stages of the implementation process	47

Abstract

Bibliographic data can be defined as content of a library catalog, union catalog or abstract and indexing service which connects users to the information they need [14]. A bibliographic record helps users find and access the resources they need. Many bibliographic records have key words or indexes that can be used to search these records. These records are often contained in bibliographic databases, or historically, on paper.

IsisCB is one such bibliographic database which started off on paper like most historical bibliographies. The website we are working on, IsisCB Explore, is a user interface that allows its users to access the large amounts of bibliographic data present in the IsisCB database. The main focus of this thesis is improving the usability of the IsisCB Explore user interface by adding small multiple time series visualizations.

The way the user interface is designed right now, it takes a lot of time and effort for a user to be able to perform tasks such as finding specific information, filtering information, and seeing and navigating through the data. All of these types of tasks allow users to find patterns in the data available on IsisCB. Most users of IsisCB Explore use the interface for analyses of these types. Currently, the user needs to go through different steps and pages to construct complex queries using a series of mouse clicks that map to these tasks. This makes the IsisCB Explore user interface complicated to use, especially on a regular basis or for someone with less familiarity. Through our work we want to make this process easier and more user friendly.

Visualization is one of the most common ways of exploring and analyzing data, especially when we have large amounts of data. By adding visualizations to the IsisCB user interface, we hope to make it possible for the user to analyze different subsets of the bibliographic data much faster and in a less error-prone way. This will result in improved user satisfaction and productivity for existing IsisCB users. By incorporating

the technique of small multiple time series visualizations in the IsisCB Explore interface, we aim to improve a user’s ability to find patterns from the bibliographic entries in the database over time by displaying them graphically.

Towards demonstrating this thesis, we utilized the raw data available on IsisCB Explore and we worked with experts on designing and implementing small multiple visualizations. A discovery user study was performed to understand the needs of a regular IsisCB user, and analyze our design against those needs. The focus of the user study was to help define the design space of the data to visualize and visual representations to potentially add. We wanted to see how interested and needed these visualizations are to our users in relation to the design space and time-oriented data from our discovery study.

From the process that we followed and from the results of our discovery user study, we conclude that small multiple visualizations provide an interactive way of understanding data and enable IsisCB Explore users to see different kinds of patterns in the IsisCB data that sometimes are not apparent by looking at data in textual. From these patterns, IsisCB users can gain meaningful insights to further their everyday work.

1 Introduction

Data is an important part of today's world. Data sets often contain a lot of information. It is crucial to store data properly and for this reason different types of data are stored in different ways. One example of huge amounts of data being stored in structured format is bibliographic data. Bibliographic data can be defined as organized collections of references to published literature like journal and newspaper articles, conference proceedings, reports, government and legal publications, patents, etc. A digitized version of this is called a bibliographic database. IsisCB Explore is such a digitized bibliographic database accessed through a website.

Bibliographic data is primarily nominal and categorical in nature. Because bibliographic data is a representation of a library catalogue, it consists of many data attributes (named and typed columns in a database). Most of the data attributes in the IsisCB database are either nominal, temporal or categorical, and even the temporal data is treated as nominal. Time data can be dates or years, but also historical periods, eras, etc. By doing so, the information content of these data attributes is absorbed or reduced. By showing data values as individual text items, it is hard to see the different patterns in these data attributes. For instance, the temporal data for each citation is only visible when one views a particular citation. This information is difficult to use for more than simply reading about that citation when it is in this form. For such nominal data to be analyzed and understood, it needs to be read through manually, or it needs to be transformed in some way. Reading through all citations and recording observations manually to get an overview is typically a tedious task.

Visualization can transform nominal data into something more easily accessible. Visualizations can provide an overview of the complete data, and interactive visualizations in particular can give users the ability to further explore the data. Transforming

data into visualizations reduces the amount of effort an individual will have to make to understand data. This is critical when it comes to a database like IsisCB, which has data with many nominal attributes and where the database is updated with hundreds of items, several times per year. The visualization additions we describe here make the user interface more accessible to non-expert IsisCB data users.

The Visualizations enable users to see patterns of multiple years within the data without having to perform a series of clicks to form their queries. In this thesis, we discuss how *the incorporation of small multiple time series visualizations improve the usability of the faceted search user interface called IsisCB Explore*. We designed our visualizations with expert help from the current editor of IsisCB, performed a discovery user study to help us understand the needs of IsisCB users, analyzed how well our design choices align with them, implemented the visualization designs and added them to the IsisCB Explore user interface. The steps in the development process that we followed to explore the thesis are described in the following chapters:

- Chapter 2 provides background on user interfaces and types of UIs in particular faceted search and graphical user interfaces, small multiple and time series visualizations, and the history and architecture of IsisCB and the IsisCB Explore website.
- Chapter 3 describes the discovery user study conducted to explore the domain and its areas of improvement, and the analysis and results of the study.
- Chapter 4 discusses the design space, design specifications and its implementation, the intricacies involved in the set-up of the augmented IsisCB Explore system, and problems faced during the process.
- Chapter 5 concludes with a discussion of the improvements made to the usability

of IsisCB Explore and how the additions to the user interface align with the needs of users.

2 Background and Related work

2.1 Bibliographic Databases and Search Engines

Bibliography is a formal way of defining collections of data. Most libraries have some form of bibliography manager to organize the large volume of resources they contain. Library websites are a common example of bibliography management. Bibliographic data is not only used by library personnel for inventory purposes or collection development, but can also be used to analyze the nature of the bibliographies. By performing analysis on bibliographic data, insights can be gained regarding different fields of research, different genres of resources, or the resource authors themselves.

Over the years, there have been a lot of bibliographic databases, tools and websites. Web of Science [6] is one of the largest sources of bibliographic information. It is now a subscription-based website with citations from many disciplines. Web of Science enables its users to perform analysis, generate reports, and example visualizations. It was formerly owned by the Institute for Scientific Information and is now maintained by Clarivate Analytics. Some other bibliographic databases that are part of Web of Science are the Chinese Science Citation Database, the Russian Science Citation Index, the Arts and Humanities Citation Index, to name a few.

Academic databases have a long history. Over time many bibliography stakeholders have shown interest in evaluating these databases, some from a content perspective, others from a usability perspective. Mongeon, et al. [24] performed a comparative analysis between Web of Science and Scopus, two major sources of bibliographic data. Another notable evaluation in the field of bibliographic databases and search engines was the evaluation in Gusenbauer, et al. [17] of Google Scholar, PubMed, and 26 other widely used academic search engines.

Bibliographic databases combined with analytical abilities can become a tool for research in any discipline. To be able to perform such analysis, it is important for the databases to have an interactive user interface. In this chapter, we focus on relevant types of user interfaces and some general principles of user interfaces, look at visualizations and their importance in analysis of data with a focus on small multiple and time series visualizations, and provide a brief history of IsisCB.

2.2 Information Interfaces

User interfaces are defined as the part of a system that handles interactions between humans and computers. The purpose of most user interfaces is to support ongoing interaction between a user and a computer. User interfaces are quite a common part of our everyday lives. Some general examples of user interfaces are those in cars, phones, exercise equipment and, various medical equipment, things that we don't often consider to be UIs. In this section, we discuss key general principles of user interfaces with particular focus on those that are relevant to IsisCB Explore and the ones that we choose to focus on in our discovery user study.

User interfaces have a long history. There are a number of principles proposed over time that define what a good interface is supposed to be. Some general principles that we are considering in our research have been proposed by Ngo [23] and evaluated for present day web interfaces by Bhaskar, et al. [5]. The proposed set of general principles to be noted while developing a good quality user interface that are relevant to our work are:

Clarity: The interface must be clear and should avoid any ambiguity. Everything must be made clear through language, flow, hierarchy, and metaphors for visual elements. The current IsisCB interface has all the data clearly shown in facet boxes where

each facet is an attribute type of the data records in IsisCB. Clearly showing attribute types in this way adheres to the Clarity principle.

Simplicity: The interface must be concise and not include so much information that it gets difficult to find the information needed. It should be simple to use and only have relevant information displayed. Without having too many facets, for all the different attributes of the data records in IsisCB, the user interface maintains its simplicity by displaying only some of the available attributes.

Familiarity: Even if someone uses an interface for the first time, certain elements must still be familiar so as to not confuse the user too much. A good way to achieve this would be to use real world metaphors while designing the system. The IsisCB Explore interface is a domain specific interface, but it still has some familiar elements, including a search box, page numbers at the bottom to navigate through search results and, links underlined upon mouse hover. These are elements found in most search engine UIs.

Responsiveness: The user interface must provide good feedback to the user without taking too much time. The feedback can be in any form: visual, textual, auditory, etc., but the system should let the user know their request has been processed. The IsisCB Explore interface is quite responsive and provides fast results typically less than a few seconds to most actions.

Aesthetics: It is important for the users to enjoy the time they spend working on the interface. The user interface must be pleasing to the eye. In terms of aesthetics, IsisCB Explore follows a color theme of blues, which is usually interpreted as a cool color that fosters a sense of calmness when looked at [20]. This makes it more likely for the user to work on the interface longer and enjoy the time that they spend on IsisCB Explore.

Consistency: A user interface must be consistent across the application allowing users to recognize usage capabilities. The same kind of action must result in the same kind of response whenever possible. IsisCB Explore's similarity to common search engines, as mentioned above, follows the consistency principle.

Efficiency: A great user interface must make the user productive by giving them shortcuts and good designs that make it easier to understand the underlying meaning of the data being displayed. The navigation from one point to other must be short when possible and involve a small number of mouse clicks. The current IsisCB Explore interface lacks this type of efficiency for seeing patterns in attribute values. We aim to contribute via this principle by incorporating small multiple designs into the user interface.

The above-mentioned set of user interface design principles are just some of the many factors to be considered for the design and development of a functional user interface.

2.3 User Interface Types

Over the long history of user interfaces there have been many styles of user interfaces such as menu-based interfaces[30], graphical interfaces[3], natural language interfaces[30], and so many more. Below is a background on some user interfaces that are relevant to the IsisCB Explore user interface in particular.

2.3.1 Web-based user interface

A Web-based user interface allows users to interact with content or software through a web browser. The data or content is downloaded from a web server to the web browser where the user can interact with it. Some well known web-based user interfaces are

online shopping sites like Amazon and eBay, social media sites like Facebook and Twitter, and online document sharing sites like OneDrive and Google Docs. Web-based user interfaces are ideal for allowing users to interact with large collections of data and find meaningful insights from them. An example for web-based user interfaces used for storing and analyzing data is the web-based user interface developed by Gasser et al [15] and his team as part of the PHENICX project.

IsisCB Explore is a web-based user interface being hosted on web using Django. The entire IsisCB data is available on the web to explore on the website.

2.3.2 Graphical user interface

A graphical user interface is an interface that accepts input in the form of mouse or keyboard and provide output in the form of graphics on a computer monitor. Graphical user interfaces are leading in the field of data analysis and data visualization. One of the major advantages of a graphical user interface is seeing information in the graphics itself. Visualizing and presenting the content of a user interface in the form of graphics enables users to better see and understand the underlying patterns in data.

One major style of design of graphical user interfaces is object oriented user interfaces. A graphical user interface that is developed by using objects and object-oriented programming is called an object-oriented graphical user interface. A GUI developed using Java components such as buttons, radio buttons, check-boxes, labels, etc. is an example of an object-oriented graphical user interface. In many imagined designs, user interfaces follow the principles of object-orientation.

IsisCB Explore is a graphical user interface in which the different graphical elements present in it, such as its timeline visualization that allows users to explore the different types of records available, and map visualization that gives users geographical information regarding a particular authority. Our addition of small multiples augments this

GUI by providing an overview of all time related data for each facet independently.

2.3.3 Faceted search user interface

A user interface that provides the result of a search query in a grouped manner is called a faceted search user interface. This type of user interfaces supports the faceted search technique, which involves combining traditional searching techniques along with a faceted navigation system, allowing users narrow down their search results. A prominent example of a faceted search user interface is the Amazon website. Most faceted search user interfaces organize their data using a faceted classification scheme that is tailored to their type of data. According to Tunkeleand [33], faceted search provides better information seeking support when compared to conventional search approaches. An example of a faceted interface is Facetedpedia [12], a faceted information retrieval system for information discovery and exploration in Wikipedia.

IsisCB Explore is a faceted search user interface with 12 facets in its interface, showing different types of information. Eight of the twelve facets are related attributes facets, with the remaining four being general information facets.

2.4 Visualization

Visualization is defined as "the communication of information using graphical representations" [36]. Visualizing data allows us to identify relationships between data by examining graphical patterns. This makes it easier to gain meaningful insights needed to improve the system being studied. Visualizations have been used for thousands of years. Early examples of visualizations include depictions of animal herds in cave paintings and Chinese trade maps. Some other notable visualizations early on in the history of visualization are Playfair's [26] charts and plots of commercial activity over time,

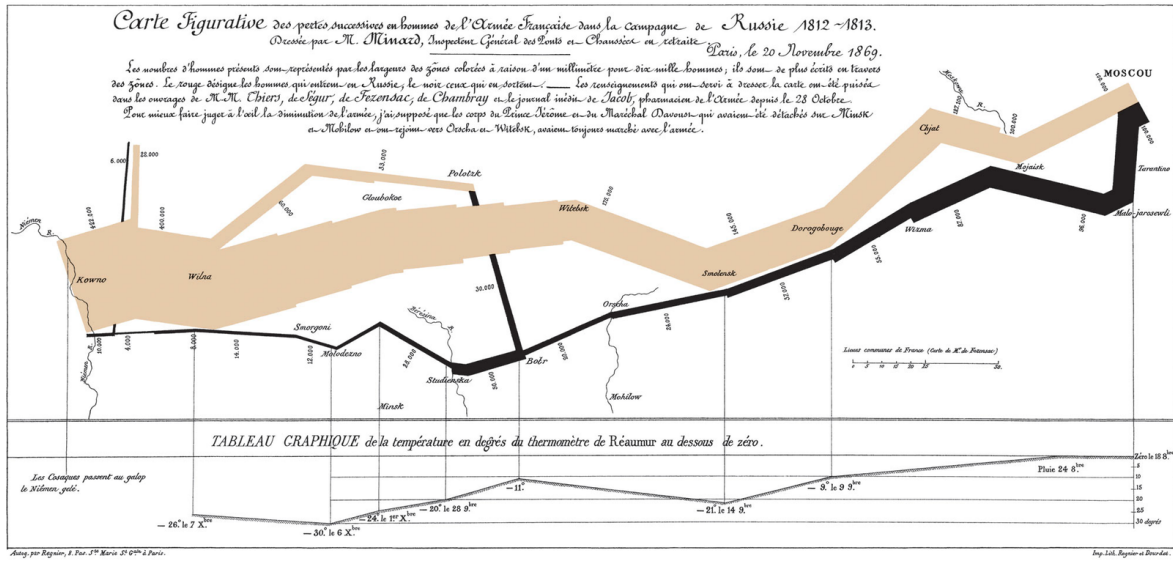


Figure 1: Minard's map of Napoleon's 1912 march on Russia [27]

Nightingale's rose charts [8] that show the causes of mortality during the Crimean War, the Cuban missile crisis illustration by Bertin [4] and Minard's map of Napoleon's 1912 March on Russia. These visualizations and many others are surveyed by Aigner [1].

Figure 1 shows Minard's map of Napoleon's 1912 March on Russia. The original map has descriptions in French. In this map, Minard shows at least six types of information regarding Napoleon's army during their entry and departure of Russia, including the size of the army, distance traveled, temperature, latitude and longitude, direction of travel, and location on specific dates [27].

Visualizations are often made up of multiple views, each of which contains data and controls for manipulating parameters that aid in the formation of different queries. Data points are visually encoded in views using visual channels such as position, color hue, texture, connection, containment, density, color saturation, shape, length, angle, slope, and area, each used to represent some aspect or attribute of the data [21]. Data points are mapped to visual elements appropriate to their attribute data types.

Historical visualizations are generally static, drawn on static media. More recently,

computer supported visualizations also allow for interaction, which lets users perform queries and interact with different parts of a visualization. The interactive elements can be the views themselves, to support navigation of elements and selection of data, or controls like sliders, buttons, and check-boxes, that allow users to interact with additional parameters of the visualizations, including parameters that affect the queries to produce data shown in the views. A user can perform actions like overview, filter, select, zoom using the interactive elements of a visualization.

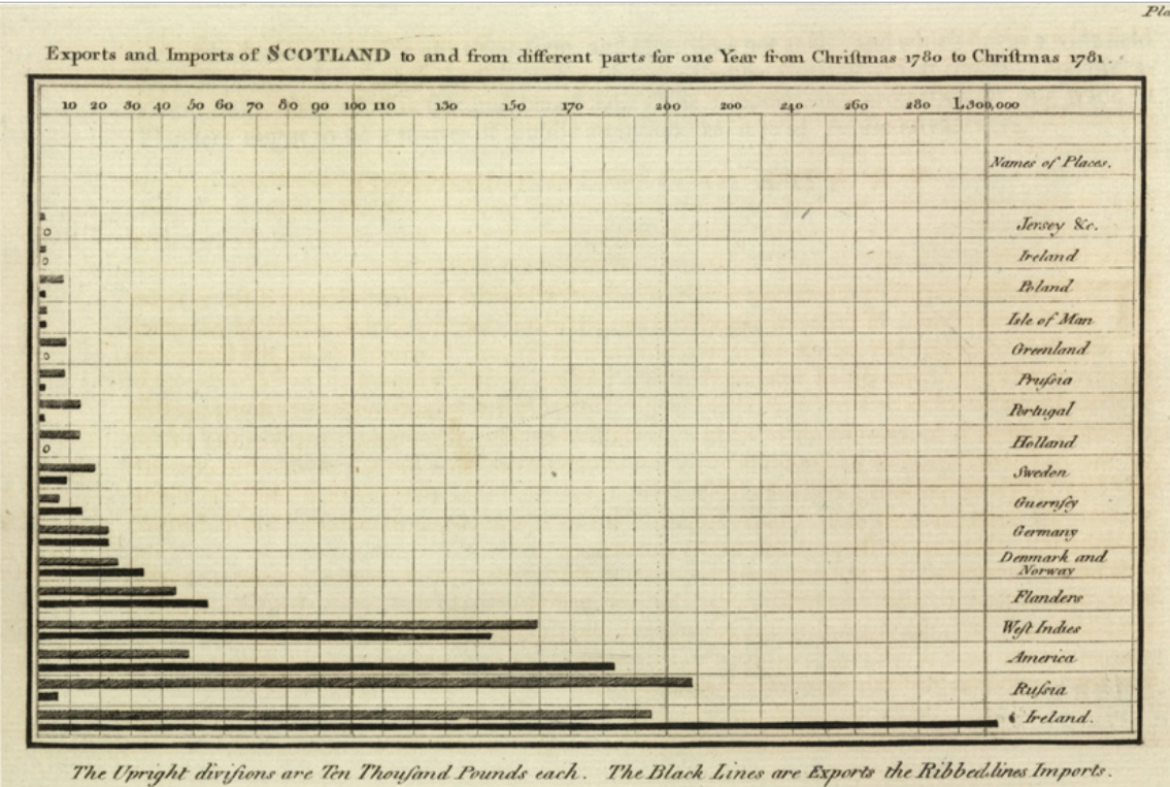


Figure 2: William Playfair’s *The Commercial and Political Atlas*, 1786
 Source: <https://commons.wikipedia.org/w/index.php?curid=42808916>

Research in visualization aims to develop new ways of exploring and analyzing data. Jigsaw [29], for example, is a visual analytics system for investigative analysis. It provides multiple coordinated views to help investigative analysts analyze large numbers of documents quickly and efficiently. Another visualization technique proposed to an-

alyze large amounts of data is the LifeLine [25] visualization environment. LifeLine is presented as a way to efficiently summarize personal data. Various interactive features are also provided to aid in any analysis needed. The hotels visualization is another notable example of using multiple coordinated views to represent historic data to find patterns and insights [38].

2.4.1 Small Multiple Visualizations

Small multiples are defined as ” a series of graphics, showing the same combination of variables, indexed by changes in other variables” [32]. Small multiples are basically series of charts but smaller in size and with minimal details. The concept of small multiples was popularized by Edward Tufte in 1983. According to Tufte [32] they are an excellent way to do a comparison analysis and are economical as well. Small multiples allow us to analyze a bigger or an overwhelming data set by breaking it down in to smaller sizes.

When we combine the concept of small multiples with time series data, we end up with an effective way of presenting data in small amounts of space. In a study by Brehmer [9], small multiples allow users to perform a set of specific tasks faster and more accurately than compared to other general visualization techniques, including animations. Although the experiment was specific to mobile based visualizations, it is likely to generalize well to small multiples on other devices like the IsisCB Explore window running on other displays like laptops.

2.4.2 Time Series Visualizations

Time is quite an important and interesting data attribute to visualize. Visualizing data over time can tell us a lot about the history of things represented in the data. Visualizations that illustrate data points at specific parts in time are called time series

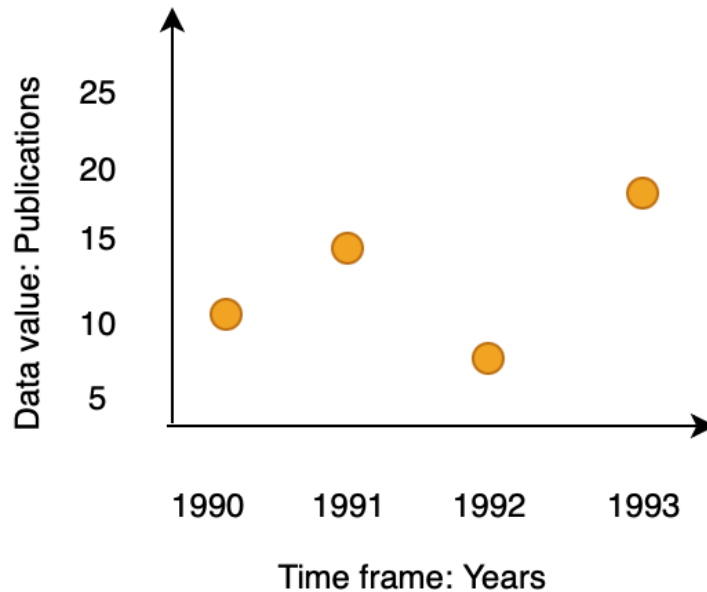


Figure 3: Number of publications of an author over several years.

visualizations. Each data point in the visualization corresponds to the value a the data point at that particular time. For instance, figure 3 shows a scatter plot with the number of publications by a particular author over a range of years.

Time series data can be encoded in visual channels such as position, color, density, color saturation, etc. It can also be shown via animation. Some common time series visualizations are line charts, bar charts, dot plots, and arc diagrams[37]. There are many different types of time series data, such as from weather forecast, stock trades, patient records, and public records. All of these types of data have some form of time variable associated with them, but it is the type of the other variable that determines the type of time series visualization that would represent this data well. Aigner, et al [2] expands on this in his work, in which he presents a taxonomy of techniques for visualizing time oriented data.

Figure 4 shows Nightingale’s rose charts[8]. In the figure, the red color represents

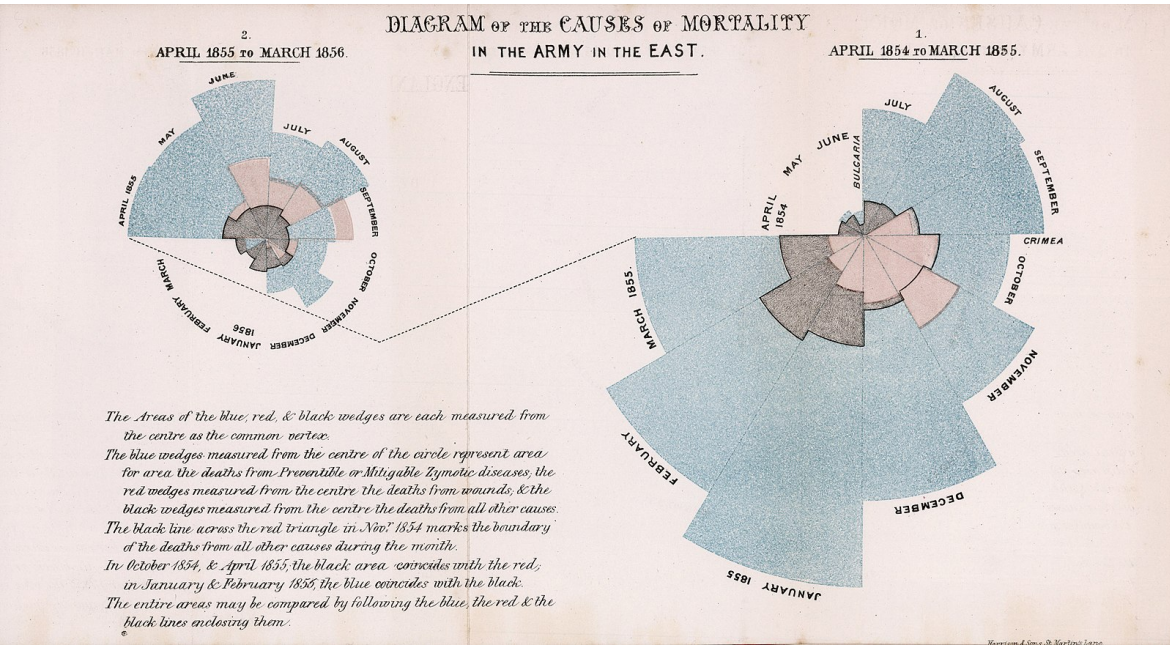


Figure 4: Nightingale’s Diagram of the Causes of Mortality[8]

deaths due to battle wounds, the black represents deaths due to accidents and other causes, and the blue color represents deaths caused by infections that could have been avoided with proper medical care and hygiene.

Over the years, there have been a variety of more complex techniques developed to visualize time-oriented data. ThemeRiver [18] represents the process of evolution of document collections over time. It is a good example of a complex technique visualization where the process of each dimension changing with time is represented. Mortiz and Fisher present a technique called DenseLine [22] to visualize many time series in a single chart. His technique allowed for large amounts of information to be available at a glance, but traded off on the clarity aspect. Dork, et al. [13] presented a technique to monitor events that happen on a large scale, such on social media like Twitter. Their technique deals with large amounts of data that is always evolving. All of these are in addition to the early historical visualizations mentioned earlier.

Time series visualizations are particularly important for data sets with a long his-

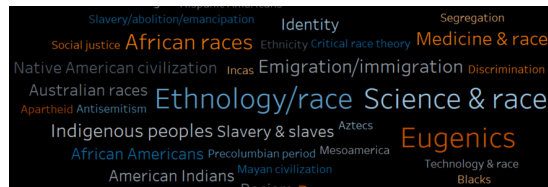
tory to them. For large data sets with hundreds of thousands of data points, time series visualizations provide a way of representing all the points within a single space. These visualizations provide clear insights into the different patterns and behavior of the data being visualized. For the IsisCB Explore data, the scalability of time visualizations technique to a large range of time is needed because it has publication records spanning over 100 years.

2.5 History of IsisCB and the IsisCB Explore Architecture

Isis Current Bibliography of the History of Science [39], IsisCB for short, is a bibliography that provides citations of published work from around the world. The users of IsisCB are primarily from the field of history of science and is the bibliography of the History of Science Society. The CB in IsisCB used to stand for "critical bibliography" but was later identified as "current bibliography" or "cumulative bibliography". The current editor of the Isis Current Bibliography is Dr. Stephen Weldon of the History of Science Department at the University of Oklahoma.

2.5.1 History of IsisCB

IsisCB was first published in 1913 by George Sarton, a Belgian scholar. Over 400,000+ works have been indexed and classified throughout the long history of IsisCB [40]. IsisCB first became digital under Dr. John Neu in the form of a database created for the history of science, technology, and medicine (HSTM). The digitized version of IsisCB that is being hosted as a web page is called IsisCB Explore. Neu was also responsible for the massive growth of IsisCB in terms of the number of records, as it was under him when IsisCB exceeded 4000 records annually, much larger than it had been prior to that point. During his term, he also reclassified the CB in IsisCB to



RACE AND SCIENCE WORD CLOUD

Today I was able to complete a small visualization that I've been working on in conjunction with the Race and the History of Science project announced in my June 9 [...] [Read more »](#)

DIGITAL SCHOLARSHIP SURVEY (UPDATES...)

March 5, 2021 | [admin](#) |

This past month, I have been doing more analysis of the survey. I am reporting on that analysis in the Spring HSS Newsletter (which is in publication as I write this). I'm now updating this short post from last year to highlight a few new files that I have created: The Google Drive folder where you can find the public [...] [Read more](#)

PANDEMIC IN SOUTH EAST ASIA

August 10, 2020 | [Stephen P. Weldon](#) |

The first bibliographical essay in our Pandemics series has arrived: Vivek Neelakantan's very interesting "History of Pandemics in Southeast Asia: A Return of National Anxieties?" and his bibliography. The article pioneers our experimental "early release and open review" process that allows us to post the first version of Neelakantan's article and bibliography while it is being reviewed. You can [...] [Read more](#)

RACE AND SCIENCE WORD CLOUD

July 15, 2020 | [Stephen P. Weldon](#) | [Leave a comment](#)

Today I was able to complete a small visualization that I've been working on in conjunction with the Race and the History of Science project announced in my June 9 post. I went through a list of keywords on this topic and found the most relevant ones. I then created a word cloud, which you can see in full below. [...] [Read more](#)

1 2 3 ... 20 »

SEARCH THE BIBLIOGRAPHY & EXPLORE HISTORY OF SCIENCE

Explore the IsisCB

SEARCH THE BLOG

Search ...

RECENT POSTS

- Digital Scholarship Survey (UPDATES...)
- Pandemics in South East Asia
- Race and Science Word Cloud
- Race and Racism in the History of Science
- Bibliographical Essays

TAG CLOUD

Alchemy Alfred P. Sloan Foundation Authorities Bibliography Biography Citation Commission on Bibliography and Documentation Digital Humanities Discovery service Double-key transcription Explore
FeaturedContent
 Functionality Geographical distribution HSS meeting International Congresses International Union for the History of Science and Technology interview IsisCB IsisCB Explore IsisCB Platform Development Isis Cumulative Bibliography Latin script Linked Open Data Neu-Whitrow Prize non-Western languages Open access TEI
THATCampHSS Unicode University of Oklahoma Libraries User comments V&A Zetoc

META

- Log in
- Entries feed
- Comments feed
- WordPress.org

SPONSOR: ALFRED P. SLOAN FOUNDATION



SPONSOR: HISTORY OF SCIENCE SOCIETY



SPONSOR: UNIVERSITY OF OKLAHOMA LIBRARIES



CONTACT INFORMATION

Stephen P. Weldon, Editor
 Isis Current Bibliography
 History of Science Department
 University of Oklahoma
 601 Elm Street
 Norman, OK 73019
 1-405-255-5187

[Privacy Policy](#)

Figure 5: The main page of the IsisCB Explore website

”current bibliography” from its previous ”critical bibliography”.

After the end of Dr. Neu’s tenure, his local database was no more functional due to the software and hardware used being outdated. When Dr. Weldon took over the bibliography in 2002 and moved to the University of Oklahoma, he along with Joy Harvey had to come up with a new system to publish as well as upload the citations. A new classification system with a few additional capabilities made possible by the advanced software was developed. This new system was both well-structured and easy to update.

2.5.2 IsisCB Explore

The IsisCB Explore is a digitized version of IsisCB. Currently the bibliography contains nearly 200,000 records, all related to different topics of interest to historians of science. It is generally used by people within the IsisCB community, mostly from the field of history of science, as well as people generally interested in bibliographies and bibliographic databases. The IsisCB Explore project is a collaborating between Dr.Stephen Weldon, History of Science department at University of Oklahoma and A Place Called Up Consulting as of 2021. The running of IsisCB Explore and its contents is managed by Dr.Weldon and his team.

This digitized version of IsisCB, IsisCB Explore is available online at <https://isiscb.org> as a Django web application. Figure 5 shows a screenshot of the main page of IsisCB Explore.

2.5.3 Architecture

The IsisCB Explore architecture, as seen in Figure 6, has 5 major components in it: a front-end Django layer, which is responsible for generating all the data visible on the IsisCB website; a Docker Django layer that performs tasks in the back-end based on the

messages received from the front-end; an Amazon SQS message queue that handles the back and forth communication between the front-end and the back-end; a PostgreSQL database that stores all the bibliographic data; and an elasticsearch component that is responsible for the indexing and fast retrieval of data records from the database.

For any task to be performed, the front-end Django component sends a message to the Docker contained Django component requesting a specific task to be performed. When it receives this message through the Amazon SQS, it takes a request to match data records from the PostgreSQL database and performs the requested query.

For instance, to generate an Authority page, one first searches for that particular authority, say "19th Century", in the search box, then from the search results finds "19th Century" and clicks on it. This triggers the process of generating the Authority page for the time period "19th Century". The process flow of generating the Authority page for the "19th Century" is as follows:

- **Process A** is getting a request for the Authority page of "19th Century" through a series of clicks performed on the web interface, and storing it in the message queue.
- **Process B** is the Docker container reading this message and starting the task.
- **Process C** is the Docker container querying the required result sets from the database.
- **Process D** is the Docker container performing the task of filtering and categorizing the data.
- **Process E** is displaying the data in the form of facets and generating the timeline visualization on the front-end Django component.

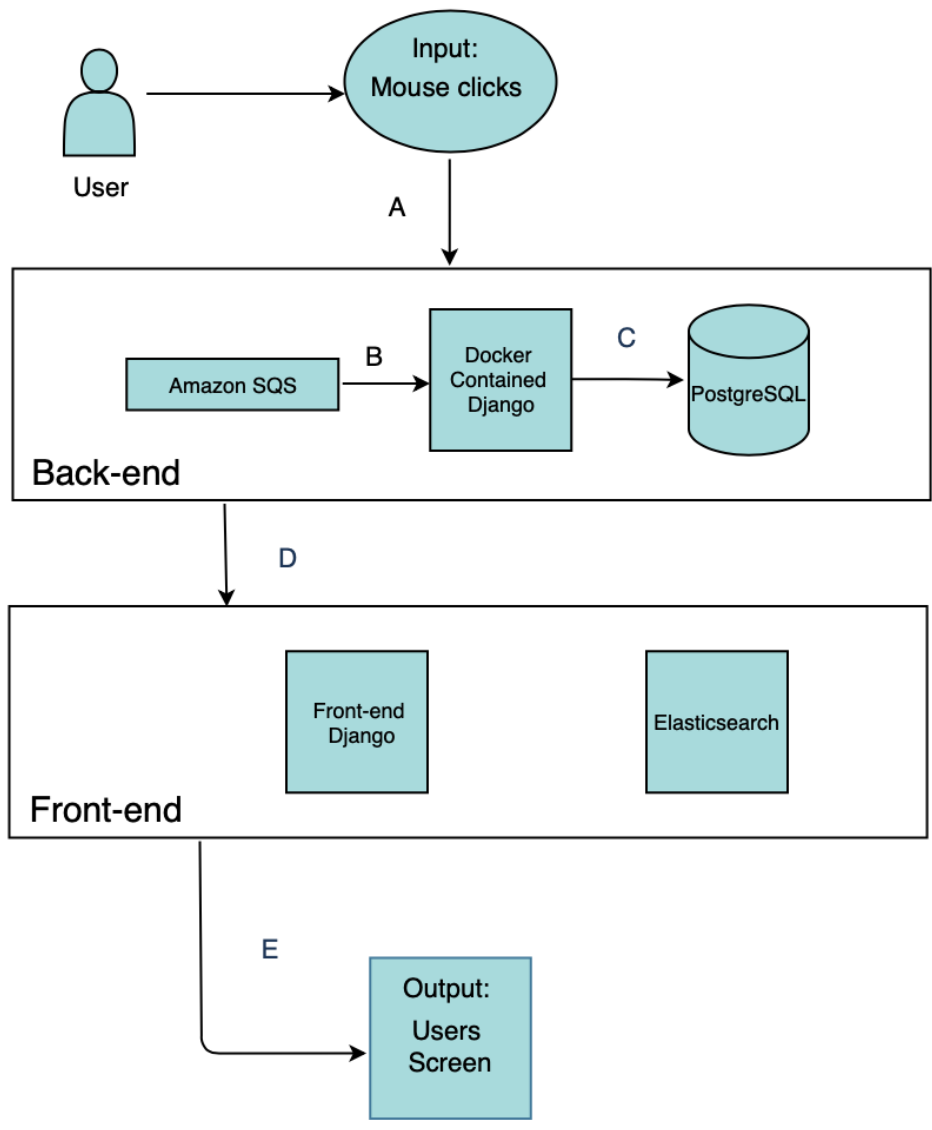


Figure 6: The IsisCB Architecture.

With the use of the above mentioned visualization techniques, we aim to add new analysis capabilities to the current IsisCB Explore system. Using the IsisCB architecture, the current IsisCB Explore website displays the result set as a set of faceted boxes and allows the users to interact with the facets to support navigation, as shown in the process flow in Figure 6. In our work we add small multiple visualizations to each of the facet values using the data provided in the result set. These visualizations provide a compact view of all the underlying time series data within the facets. By doing this, we predict that the usability and overall user experience of IsisCB Explore users will be improved by giving users more ways to interact with the system visually.

3 Discovery Study

To anticipate the effectiveness of our design additions to the IsisCB Explore interface, we performed a discovery user study. Our aim for this study is to understand how people utilize the existing faceted browsing interface in IsisCB Explore, identify improvements to its current design, and assess the potential effectiveness of our new data visualization features as a means of expanding its capabilities. The entire session will be conducted via Zoom with each participant using the IsisCB Explore web interface and answering questions via survey instruments (in Qualtrics).

The IsisCB community has a few hundred people who use the website. Our pool of potential participants were people who are more directly involved with IsisCB rather than just being an occasional user. A list of potential participants was taken from Dr Weldon and they were about 25 in number. The participants for the study were recruited by email.

3.1 User Study

We designed our discovery user study with the guidance of the current editor of IsisCB Dr Stephen Weldon. The tasks were designed with the focus on finding the parts of the user interface that most users usually visit and use. By finding these areas we wanted to see how closely do our additions align with the needs of a regular IsisCB Explore users.

The discovery user study was conducted via Zoom. The users were screen recorded while performing tasks using the IsisCB Explore interface. Our pool of users were people with prior knowledge of IsisCB Explore or prior experience of IsisCB Explore. There were a total of 7 participants a mix of graduate students, staff of IsisCB and researchers who are part of the board of directors of IsisCB. All the participants work

with IsisCB regularly either as developers or content managers or as other members of the community.

At the beginning, the participants were asked to complete one knowledge assessment survey and a demographic plus IsisCB usage survey. These surveys took about 15 minutes. Through these surveys we wanted to know what kinds of things a typical IsisCB user does with the IsisCB Explore interface and what new features they'd like to see added to it. Next, the participants were asked to complete a set of tasks. These tasks were focused on having the participants interactively explore IsisCB to gain insights about user usage. These tasks were designed to be similar to the tasks an IsisCB user would perform on a typical day. There are 4 tasks starting off with the participants being asked to explore the IsisCB Explore interface for a few minutes and then to choose a citation and write down that information. Through this task, we want to make any non-regular IsisCB users more comfortable with the user study. Then, the participants are directed through a series of steps to the Authority page and are asked to answer a few questions. Then we slowly transition in to questions related to time series data and then encourage the participants to use the single timeline visualization to answer some more questions. All these tasks took about 30 minutes. The participants were then asked to complete a post-study feedback survey. In this survey, we collect participants feedback for each of the tasks. This survey took about 10 minutes on average.

Through out this whole process, we focus on having the user use the IsisCB Explore interface, in particular the various facet boxes in the Authority page as a tool to find data. By watching this process, we want to monitor and record the series of steps a participant takes to perform these tasks so as to find area of the interface where improvements were possible. We also wanted to validate our choices of design space

that is the Authority page and our choice of visualizing time-oriented data. From the study, we wanted to see if the needs we were trying to address were actually needs of a regular IsisCB Explore user and if our choice of using visualizations as a means of improving the users experience were of some interest to the users.

3.2 Study Results and Analysis

We analyzed the information gathered through the discovery study by taking in to account their answers for the IsisCB focused survey and by answering two sets of questions regarding the participants behaviour and actions for the tasks survey. From the IsisCB focused survey, we gathered data regarding the types of temporal data the participants generally use, the parts of the web interface they use and the facet boxes they find the most important and use the most. The answers to these were used to validate our our choice of working with the Authority page and in particular time series data within the facet boxes and our approach of using visualizations as the means of improving the interface. For the tasks survey, through the first set of questions, we

	Statement 1: "I frequently use the Authority page as part of my work."	Statement 2: "The Bar Chart made finding information easier."	Statement 3: "Decade is an important time scale in my use of IsisCB Explore."
Participant 1	Agree	Agree	Agree
Participant 2	Agree	Agree	Neutral
Participant 3	Agree	Agree	Agree
Participant 4	Disagree	Agree	Disagree
Participant 5	Neutral	Agree	Agree
Participant 6	Agree	Agree	Agree
Participant 7	Agree	Agree	Agree

Figure 7: Participants Statements

wanted to access the participants familiarity of the IsisCB Explore interface. From the second set of questions we wanted to see which parts of the user interface did the participants use most or struggle with while finding the answers to the questions. By combining these results, we wanted to see how the participants familiarity level with the IsisCB Explore interface was related to areas of the interface they are using. This lets us know if the participants highly familiar with the IsisCB Explore interface and participants with less familiarity are using the same parts of the interface while performing queries. This allows us to find most used parts of IsisCB Explore and see how these map with the additions we made to the user interface. Our choices of design space and data were evaluated based on the participants response to the following statements:

- Statement 1: I frequently use the Authority Page as part of my work.
- Statement 2: The Bar Chart made finding information easier.
- Statement 3: Decade is an important time scale in my use of IsisCB Explore.

From figure 7, we can see that only 14 percent of our participants disagreed with Statement 1 regarding our choice of design space, the rest all either agreed or were neutral. 100 percent of our participants agreed with our choice of using visualizations as a better means of presenting the data to the users. Similar to Statement 1, only 14 percent of our participants disagreed with Statement 3 regarding our choice of data to be time-oriented data. Through these statements and the participants level of agreement to them, we were successful in validating our choices.

We also used accuracy as a way of determining the usefulness of visualizations in improving the users ability to access and understand the data in IsisCB Explore interface. From the bar chart in Figure 8, we can see that 42 percent of our participants

failed to answer questions in Task 3 right, the rest were successful. 71 percent of our participants answered the questions in Task 4 right.

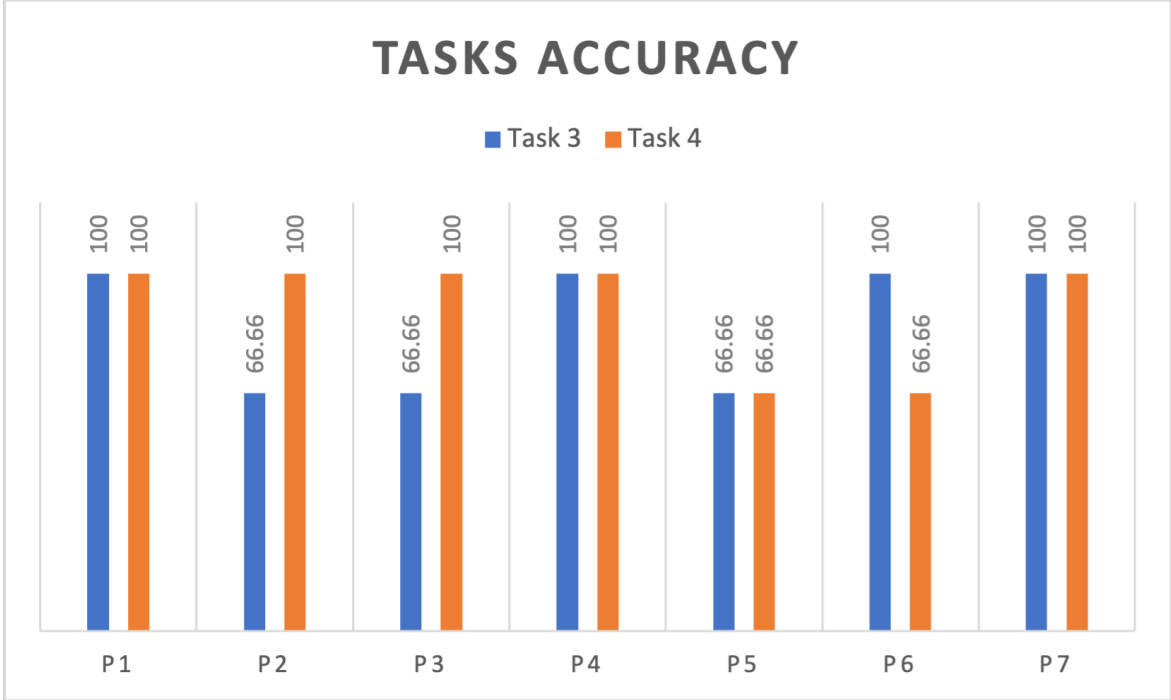


Figure 8: Participants Accuracy Chart

From the results of the discovery study, we conclude that the area of the IsisCB Explore user interface namely the Authority Page with all the facet boxes is one of the most used parts of the IsisCB Explore user interface. Most users also felt time oriented data to be an important attribute and were especially interested in decades and centuries. Based on these conclusions, we anticipate our design additions of small multiple time series visualizations to the IsisCB Explore interface to be an important new feature that aligns with the needs of IsisCB Explore users.

4 Design and Implementation

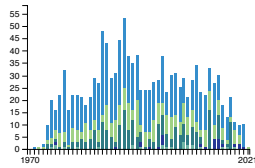
To display the abundant amount of data available in IsisCB Explore, we designed a small multiple time series visualization with some interactive features to explore all the different attributes. In this chapter, we will take a closer look at the design space and the various considerations we made during the process of designing our visualizations. We will also see the implementation process, the various concepts we used for implementing the visualizations, our final design and some efficiency issues faced and their solutions.

The data we choose to focus on, the years and publication counts within the time periods available in the IsisCB Explore database, are a mix of quantitative and categorical data. According to McKinlay [21], the most appropriate way to display this kind of quantitative data is through position or length. But due to the lack of space available we choose to go ahead with color hue as our way of representing the data since color is the best way to describe categorical data which our data comes under as well. In this chapter, we will take a closer look at the design space and the design specifications as well.

4.1 Design Space

The Authority page which is a faceted user interface has about 12 facets which have information about different attributes for a specific Authority record. In simple terms, an authority record is information about a Citation. An authority record can either be a topic or a person or a concept or any of the other attributes that every Citation has. The Authority page allows IsisCB Explore users to further analyze attributes related to a specific attribute. Figure 9 shows a screenshot of the Authority page for Isaac Newton.

Publication dates of IsisCB bibliographic records
(click to expand)



Person

IsisCB ID: CBA000073903

Newton, Isaac

Show all 1255 citations to this authority. Show 33 citations as author, etc.
Show 1244 citations as subject or category.

Birth and Death Dates: 1642-1727

Authority URI: <http://data.isiscb.org/isis/authority/CBA000073903>

Related places (List view)



Related Authors & Contributors (See the top 100)

- 1 Newton, Isaac (33 citations)
- 1 Westfall, Richard S. (22 citations)
- 1 Hall, A. Rupert (20 citations)
- 1 Ducheyne, Steffen (20 citations)
- 1 Shapiro, Alan E. (19 citations)
- 1 Erlichson, Herman (18 citations)
- 1 Cohen, I. Bernard (17 citations)

Related Journals (See the top 100)

- 1 Notes and Records: The Royal Society Journal of the History of Science (50 citations)
- 1 Studies in History and Philosophy of Science (45 citations)
- 1 Archive for History of Exact Sciences (33 citations)
- 1 Isis: International Review Devoted to the History of Science and Its Cultural Influences (21 citations)
- 1 Historia Mathematica (21 citations)
- 1 Annals of Science: The History of Science and Technology (21 citations)
- 1 British Journal for the History of Science (19 citations)

Related Publisher (See the top 100)

- 1 Cambridge University Press (27 citations)
- 1 Oxford University Press (15 citations)
- 1 Kluwer Academic (11 citations)
- 1 University of Chicago Press (6 citations)
- 1 Springer (6 citations)
- 1 Green Lion Press (5 citations)
- 1 Akademie-Verlag (5 citations)

Datasets

- 1 Isis Bibliography of the History of Science (1253 citations)
- 1 Technology & Culture Bibliography (2 citations)

Related Concepts (See the top 100)

- 1 Science (184 citations)
- 1 Physics (173 citations)
- 1 Mechanics (168 citations)
- 1 Optics (108 citations)
- 1 Mathematics (84 citations)
- 1 Astronomy (82 citations)
- 1 Science and religion (71 citations)

Related Categories (See the top 100)

- 1 The profession; the person in the subject (449 citations)
- 1 comments on specific subjects (441 citations)
- 1 Seventeenth century (393 citations)
- 1 Science (176 citations)
- 1 Mechanics (124 citations)
- 1 Physics; physical sciences, general works -- 17th century (102 citations)
- 1 Eighteenth century (83 citations)

Related People (See the top 100)

- 1 Newton, Isaac (1244 citations)
- 1 Leibniz, Gottfried Wilhelm von (60 citations)
- 1 Descartes, René (36 citations)
- 1 Galilei, Galileo (34 citations)
- 1 Locke, John (27 citations)
- 1 Huygens, Christiaan (25 citations)
- 1 Boyle, Robert (23 citations)

Related Institutions (See all 14)

- 1 Royal Society of London (12 citations)
- 1 Cambridge University (7 citations)
- 1 Jesuits (Society of Jesus) (4 citations)
- 1 Oxford (1 citations)
- 1 Shanghai Lida Polytechnic Institute (1 citations)
- 1 Deutsche Gesellschaft für Geschichte der Medizin, Naturwissenschaft, und Technik (1 citations)
- 1 Groningen, Academie (1 citations)

Outside Links (Show all)

- 1 Linked Data: <http://d-nb.info/gnd/118587544/about/html> (DNB) <http://www.biodiversitylibrary.org/creator/23134> (BHL)
- 1 Stanford Encyclopedia
- 1 Wikipedia
- 1 Archive Grid
- 1 Consortium of History of Science
- 1 SNAC
- 1 H-Net
- 1 Twitter

Related Times (See all 16)

- 1 17th century (715 citations)
- 1 18th century (319 citations)
- 1 19th century (40 citations)
- 1 16th century (18 citations)
- 1 Early modern (16 citations)
- 1 20th century (12 citations)
- 1 Enlightenment (7 citations)



Comments

Be the first to comment!

Log in or register to comment

Figure 9: The Authority page for Newton

The relevant attribute facets are: Related Authors and Contributors, Related Journals, Related Publisher, Related Concepts, Related Categories, Related People, Related Institutions, Related Times, Related Places. These are the facets that will have the visualizations added to. For every facet, the top 7 attribute values are displayed in the facet. A large pop up allows the users to access the rest of the values by clicking on “see all” option available at the top right corner of the facet box right next to the name of the facet. Each value has a link to access all the citations that are have mentioned or are related to the particular attribute value right next to the attribute value. On clicking this, the user will be redirected to a set of search results of all the citations related to that particular attribute.

The data that we visualized is the data within each of these facets. For every attribute value in every facet, a visualization showing the timeline of publications will be added. For the visualizations to be added with the small space available, we needed a design that is not only compact enough to fit in the space but also informative enough that it will capture all the data and display it clearly within that space. To overcome this concern, a small multiple time series visualization was designed.

4.2 Design Specifications

A small multiple visualization was designed to represent the behavior of each of the facet attributes over a few decades. Design aspects defined by Aigner [1] where he draws from Frank [34] and Goralwalla et al [16] were some of our major considerations in our design process along with Javed et al’s work on multiple time-series [35]. They used principal orthogonal aspects such as the scale, scope, arrangement and viewpoint to characterize different types of time. The data we are representing is interval-based data that is discrete and linear. Our data progresses from the past to the future in an

ordered time domain, which means that events occur one after the other. Figures 10 and 11 show recreated images of Aigner’s example representations of linear data and interval-based data. Tominski et al [31] developed two visual arrangement of axes – TimeWheel and the MultiComb, where one of the advantages of his methods are the ability to move around the axes representing the other variables in a time series plot. He states in his paper that the relation between time and its dependent attribute can be efficiently analyzed when both the axes are parallel to each other either vertically or horizontally. He also further says that when axes are perpendicular to each other it gets difficult to analyze it.

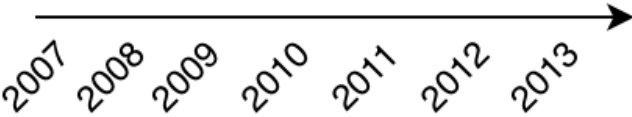


Figure 10: Aigner’s example representation of Linear data

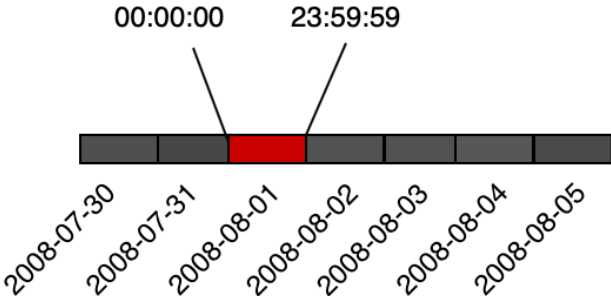


Figure 11: Aigner’s example representation of interval-based data

Byron and Wattenberg [10] evaluated stacked graphs as a way of representing temporal data. They concluded that when legibility and aesthetics are concerned, for large

amounts of time series data, stacked graphs present an alternative to the usual line graphs and bar charts. Stacked graphs are a shared space technique for different time series data with same dimensions. Each facet in the IsisCB Explore Authority page has a fixed horizontal width range of about 290-300px whereas the height is adjusted according to the number of values for each of the facet attributes. This presents an issue of space. Borrowing from the concept of stacked graphs and small multiples [32], we came up with a stacked horizontal bars design. Small multiples split the different time series in to separate graphs, this helps us overcome the space problem while keeping the visualization legible and aesthetic. Stacked graphs allow for comparison across different time series, though our design is not exactly stacked, bars representing time series data for different attributes are stacked on top of each other, with the facet attribute value being the only thing separating them.

4.2.1 Color

Drawing from the above mentioned works, a horizontal stacked bar chart with time as its x-axis and a range of colors that fill each bar in the stacked bar chart was designed. Instead of using the y-axis to represent the other variable, these range of colors were used to represent the publication counts for each of the facet attribute values. The colors are filled right on top of the length of the x-axis which represents a particular decade, thereby making the time variable and its dependent variable which in our case is the publications count, in parallel to each other. Our main consideration while deciding on a color scheme was for the users to be able to clearly differentiate between colors representing different ranges of values.

Our color range starts from White which represents the count 0 to a dark brownish red which represents the max publication count range. The colors range from pale yellow to orange to red. The lighter colors represent a smaller count, and the darker

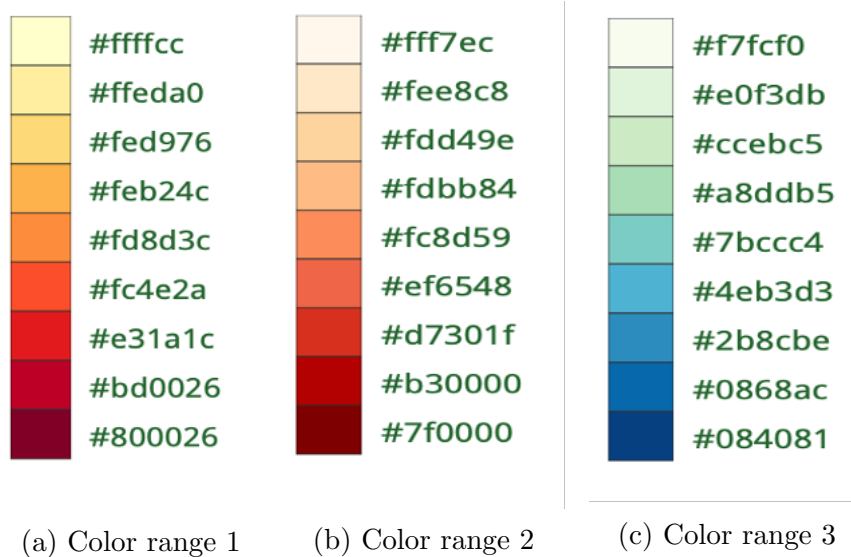


Figure 12: Color ranges considered

colors represent a higher count. Figure 12a shows the color range along with its codes. Figures 12b and 12c show some of the other color ranges we tried as well. The color ranges in Figure 12b did not provide a clear enough variation for counts that fall into the middle of the range and the color range in Figure 12c failed to do the same at the very beginning of the color range. Though the colors in Figure 12c went well with the entire theme of the IsisCB Explore user interface aesthetically speaking, the colors at the lower end of the color spectrum were very similar to white which caused some confusion. The color range in Figure 12a had enough of a variation throughout the color spectrum and was clear and visible, an important consideration especially for small multiple visualizations because of their compact size.

4.2.2 Interaction

The starting point for any user interface is Shneiderman’s [28] mantra and a part of Shneiderman’s mantra is details-on-demand. Though not a complex form of interaction, a tool-tip gives the opportunity of implementing the concept of details-on-demand.

For a visualization design such as ours, where availability of space is a constraint, adding axis labels to clearly show the information being represented becomes difficult. Adding a tool-tip to our visualization, that clearly shows the starting year of the decade and the count of publications within that decade, provides that necessary information within the given constraints. For every partition in the bar graph, the tool-tip gives these details when the mouse hovers over it.

Because the facets show only the top 7 values of an attribute within the facet box, another variation of the details-on-demand mantra, a pop-up that shows all the values up to 100 along with the small multiple visualization is also provided. On clicking the “see all” button right next to the label of the facet box, this pop-up can be accessed. Once done, clicking the close or the cross at the bottom right corner and top right corner respectively, closes the pop-up.

Similarly, another pop up that allows the user to zoom the visualization is also provided. This pop-up can be accessed by clicking on a small square box right next to the citation counts displayed for every value of the facet attribute. As with the other pop-up, clicking the close at the bottom right corner or the cross at the top right corner, closes the pop-up. This form of interaction can be mapped to the one of the seven tasks of Shneiderman’s. Though our design does not necessarily implement every one of Shneiderman’s principles, at every stage of designing, developing and implementing the interactions, we heavily relied on his principles and teachings.

4.3 Implementation

The IsisCB Explore interface is a Django Web Application. It is majorly implemented in Python along with some components being implemented in JavaScript, HTML and CSS. All the data related to IsisCB is stored in a Postgres database and is accessed

using Django ORM. Django ORM allows interaction with any database just like you would with a generic database such as SQL. In this chapter, we will discuss how the system is setup, the data is prepared and accessed, the visualizations are built and the efficiency of the system once the visualizations have been added.

4.3.1 Initial Setup

The digitized version of IsisCB, IsisCB Explore has been in production for the past 6 years and over this period it has evolved into a complex system with a lot of dependencies. To get started on this, quite a few set up steps had to be followed. The Django web application was developed on Python version 2.6 and has over 20 python modules that are used to access the various other technologies needed for the system to work. The next step was to create and populate the database using Postgres. Once the database was ready, for the application to be able to access the data records on a new system, all the data records, the 200,000 citation records and 150,000 authority records had to be re-indexed. This was done using elasticsearch. Once this was done, the data and the system are ready to use. To run this web application locally every time, the celery worker, elasticsearch and the Django server have to be started. Once these three are running, the IsisCB Explore website can be accessed locally at the address <http://127.0.0.1:8000/>.

4.3.2 Process/Procedure

There are 3 main stages in the implementation of our design: Data Querying, Data Preparation and Visualization Development. In the Data Querying stage, the required data is accessed from the database using Django ORM queries, then the queried data is filtered and transformed into a easily workable python dictionary and finally this dictionary is sent to the front-end where the visualizations are developed using D3js.

Figure 13 shows the implementation process pipeline.

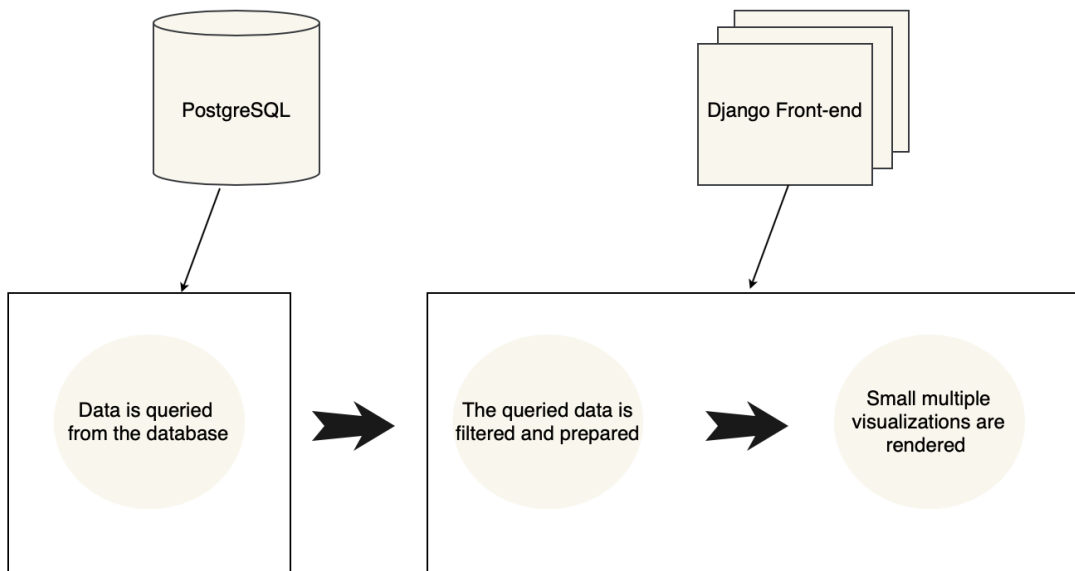


Figure 13: Stages of the implementation process

The specifics of each stage is discussed in the below sections.

4.3.3 Data Querying

The data is stored in a Postgres database named isiscbnew. This database was created manually and has over 200,000 citation records and 150,000 authority records. The data that we need, the years and citations belong to the isiscbnew table and to get the same data columns but for the attributes specific to each facet box, we use the isiscbnew table. The data is retrieved from these tables using Django ORM. Django ORM uses the concept of Object-Relational Mapping, which in a sense creates a virtual object database that can be used from within Python. This virtual database allows for querying and manipulation of data just as a regular SQL database. Because

Python is an Object-Oriented Language all of the different Object-Oriented Concepts like Inheritance, Abstraction, Class, Objects etc are all can be used to operate this virtual database.

Before writing the queries, a data model needs to be created. A model is in a sense all the information about your data, the fields and its behavior stored in a single place. Generally, one data model maps to a single database table. Every model is a Python class, and every attribute of a model represents a database field. Once the data models with fields are created, Django ORM queries can be written. These queries are written in Python and the query result set values are stored in objects.

```
class CachedTimeline(models.Model):
    """
    Caches timeline data generated for authorities.
    """
    created_at = models.DateTimeField(auto_now_add=True)
    authority_id = models.CharField(max_length=255, blank=True, null=True)
    complete = models.BooleanField(default=False)
    recalculate = models.BooleanField(default=False)
```

Figure 14: Screenshot of part of the Django code

Figure 14 shows an example data model. From the figure it can be inferred that the database table that this model is representing is a Cached TimeLine table with attributes createdat, authorityid, complete and recalculate. When this data model is run, it will create a database table named Cached TimeLine with attributes createdat, authorityid, complete and recalculate is automatically added. Figure 15 shows the corresponding PostgrgreSQL query that is generally used to create the Cached TimeLine table.

Once the data model is created, Django automatically gives a database abstraction API that allows for creation, retrieval, updating and deletion of objects. The data we

```

CREATE TABLE CachedTimeline(
    created_at DATE NOT NULL,
    authority_id VARCHAR(30) NOT NULL,
    complete BOOLEAN NOT NULL,
    recalculate BOOLEAN NOT NULL
)

```

Figure 15: Corresponding PostgreSQL code

need can be accessed through two tables, isisdatacitation and isisdataauthority. To get the data we need from these tables for the different facets in an Authority page, we need to retrieve every citation that has an AC relation with the authority record currently being displayed on the web page. For a query like this, we take all the data records from the Authorities table and the Citations table that are related to each other through a type of relation called ACRelation. The IsisCB database has 3 types of relations: AARelation, ACRelation and CCRelation. To query the data records from the facets in an Authority page, we need Authorities and Citations in those facets to be related through an ACRelation. ACRelation has three types in it: ACRelation type controlled(ACREL), ACRelation type broad controlled(ACRELB), ACRelation authority type controlled(AUTHBR). Then using the acrelation type and a keyword that every facet has, data records from each of those facets can be retrieved.

The keywords for each facets are:

- Related Authorities and Contributors: ACRELB PR
- Related Journals: ACREL PE
- Related Publisher: ACREL PU

- Related Concepts: AUTHBR CO
- Related Categories: ACREL CA
- Related People: AUTHBR PE
- Related Institutions: AUTHBR IN
- Related Places: AUTHBR GE
- Related Times: AUTHBR TI

Figure 16, 17 and 18 show the data models of an Authority, a Citation and the data model of an ACRelation. The Authority and Citation data models have different properties associated with them as well. These properties allow access to other data models.

```
class Authority(ReferencedEntity, CuratedMixin):
    ID_PREFIX = 'CBA'

    class Meta(object):
        verbose_name_plural = 'authority records'
        verbose_name = 'authority record'

    history = HistoricalRecords()

    name = models.CharField(max_length=1000, db_index=True, help_text=help_text("""
Name, title, or other main term for the authority as will be displayed.
"""))

    name_for_sort = models.CharField(max_length=2000, db_index=True, blank=True, null=True)
    """ASCII-normalized name."""

    created_on_stored = models.DateTimeField(blank=True, null=True)
    # CHECK: Had to add on_delete so chose cascade -> JD: deleting users shouldn't delete authority
    created_by_stored = models.ForeignKey(User, null=True, blank=True,
        help_text=help_text("""
The user who created this object."""), related_name="creator_of_object", on_delete=models.SET_NULL)
```

Figure 16: Screenshot of an Authorities data model

Figure 19 shows the screenshot of the query used to filter facet data records. We then filter out the records that have their type as Contributors and remove any du-

```
class Citation(ReferencedEntity, CuratedMixin):
    """
    A bibliographic record.
    """
    ID_PREFIX = 'CBB'

    history = HistoricalRecords()

    # Allowing blank values is not ideal, but many existing records lack titles.
    title = models.CharField(max_length=2000, blank=True,
                             help_text="The name to be used to identify the"
                             " resource. For reviews that traditionally have no"
                             " title, this should be added as something like"
                             " '[Review of Title (Year) by Author]'."))

    title_for_sort = models.CharField(max_length=2000, blank=True, null=True,
                                      db_index=True)
    """ASCII-normalized title."""

    title_for_display = models.CharField(max_length=2000, blank=True, null=True)

    additional_titles = models.TextField(blank=True, null=True,
                                         help_text="Additional titles (not"
                                         " delimited, free text).")
    book_series = models.CharField(max_length=255, blank=True, null=True,
                                   help_text="Used for books, and potentially"
                                   " other works in a series.")

    created_native = models.DateTimeField(blank=True, null=True)
    # CHECK: Had to add on_delete so chose cascade -> JD: deleting user shouldn't delete citation
    created_by_native = models.ForeignKey(User, null=True, blank=True,
                                         help_text=help_text("""
                                         The user who created this object."""), related_name="creator_of", on_delete=models.SET_NULL)

    # CHECK: Had to add on_delete so chose cascade -> JD: deleting subtype shouldn't delete citation
    subtype = models.ForeignKey('CitationSubtype', blank=True, null=True, on_delete=models.SET_NULL)
```

Figure 17: Screenshot of a Citations data model

plicate values. Every facet attribute has a type name that is used to filter them. For instance, the Related Authors and Contributors facet has its type name as PERSONALRESPONS. Once we get all the citation records in the Related Authors and Contributors facet, we filter out just the publication year column.

4.3.4 Data Preparation

Once we have the data records we need, the data still needs to be worked on for us to actually build visualizations using it. Because we want to build a time series plot for

```
class ACRelation(ReferencedEntity, CuratedMixin):
    """
    A relation between a :class:`.Authority` and a :class:`.Citation`.

    For example, between a paper and its author(s), a book and its place of
    publication, etc.
    """
    ID_PREFIX = 'ACR'

    class Meta(object):
        verbose_name = 'authority-citation relationship'
        verbose_name_plural = 'authority-citation relationships'

    history = HistoricalRecords()

    # CHECK: Had to add on_delete so chose cascade -> JD: ACRelations shouldn't be deleted when citations/authorities are
    citation = models.ForeignKey('Citation', blank=True, null=True, on_delete=models.SET_NULL)

    # CHECK: Had to add on_delete so chose cascade -> JD: ACRelations shouldn't be deleted when citations/authorities are
    authority = models.ForeignKey('Authority', blank=True, null=True, on_delete=models.SET_NULL)

    name = models.CharField(max_length=255, blank=True)
    description = models.TextField(blank=True)

    # TODO: conditional validation on these fields.
    # Allowed values depend on the value of the Type.Broad,controlled
    # if Type.Broad.controlled = 'HasPersonalResponsibilityFor'
    AUTHOR = 'AU'
    EDITOR = 'ED'
    ADVISOR = 'AD'
    CONTRIBUTOR = 'CO'
    TRANSLATOR = 'TR'
    # if Type.Broad.controlled = 'ProvidesSubjectContentAbout'
    SUBJECT = 'SU'
    CATEGORY = 'CA'
    # if Type.Broad.controlled = 'IsInstitutionalHostOf'
    PUBLISHER = 'PU'
    SCHOOL = 'SC'
    INSTITUTION = 'IN'
    MEETING = 'ME'
    # if Type.Broad.controlled = 'IsPublicationHostOf'
    PERIODICAL = 'PE'
    BOOK_SERIES = 'BS'
    COMMITTEE_MEMBER = 'CM'
```

Figure 18: Screenshot of an ACRelations data model

citation count over the decades, we need decades, and the citation counts in for these decades. Using some in-built functions in the Numpy library of python, we create a dictionary with the decade number as type and citation count as the key. This dictionary is returned at the end of the method.

A python method is written for every facet and the data is directly queried from the database. Figure 20 shows the initial python method for the Related Times facet.

```
c2=Citation.objects.filter(arelacion__authority_id=authority_id)
.filter(arelacion__authority_id=other_authority_id, arelacion__type_controlled=relation_type)
```

Figure 19: Screenshot of the Query

```
@register.filter
def get_times(auth1, auth2):
    #code
    c2=Citation.objects.filter(arelacion__authority_id=auth2).filter(arelacion__authority_id=auth1, arelacion__authority_type_controlled=
    #, arelacion__public=True, arelacion__authority_public=True).distinct()
    years = []
    for c in c2.all():
        if c.publication_date:
            years.append(c.publication_date.year)
    counts,bins = numpy.histogram(years , bins=6 , range=(1970,2030))
    bins = bins.astype('int32')
    z = dict(zip(bins , counts))
    new1 = [{"type": f, "count": k} for f,k in z.items()]
    return new1
```

Figure 20: Screenshot of part of the python script

4.3.5 Visualization Development

The visualizations have been developed using D3js. D3js is a JavaScript library that allows for transforming data into visualizations using HTML, CSS and SVG [7]. The python method that returns the prepared data is called within the div element of the facet using built-in template tags in Django. These built-in tags act as the general for or with loops. The returned data is iterated over and sent into the JavaScript function for every facet attribute value. This is done for every facet in the Authority page.

Before deciding on and developing the final design, a few initial designs were explored. Figure 21 shows the first version of the small multiple visualization. In this version, we tried using color saturation as our means of visual encoding. As visible in the figure, this did not work as this design lacked legibility.

In figure 22 the second version of our design can be seen. We realized a deeper problem during this version. In the beginning we were considering citation records from 1900 to 2020 as we wanted to cover a long period of time. Just like our final design, we used white for decades with 0 publications. As you can see, this design had scaling issues and also showed us zero records, till the later part of the 20th Century.



Figure 21: The first version of our design

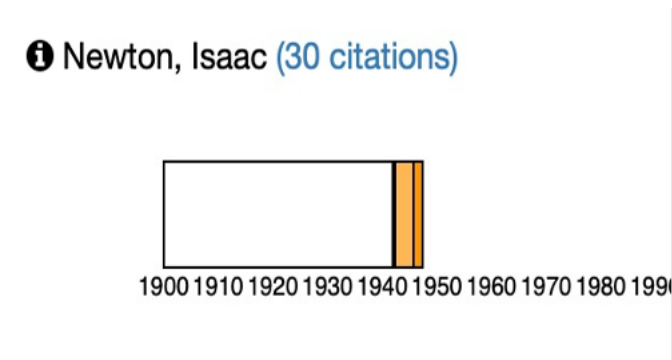


Figure 22: The second version of our design

After some analysis, it was clear that IsisCB Explore had almost no records before the year 1970. This changed our scales of time to start from 1970 to end with 2030 as we wanted to include the year 2021 as well. Figure 23 shows our final design.

In one of our design variations, we used circles placed horizontally with each circle representing a decade. Figure 25 shows this variation. These circles were equally spaced with the same color encoding as our final design, yellow and orange hue to dark red with the color white representing the value 0. A tool-tip that gives information of the data encoded was also added to this design. This information was available on demand using a mouse hover.

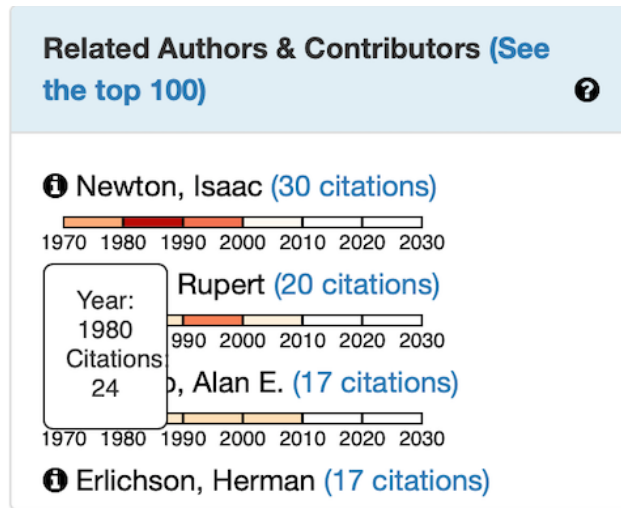


Figure 23: Screenshot of Related Contributors facet with the tool-tip

Our final design as seen in Figure 24, shows data from 1970 to 2030. Each facet has the small multiple designs in it for the top 7 attribute values. We used a color spectrum that goes from white for 0 publications to a yellow and orange hue for citations in the middle to a dark red to nearly black for bigger values.

All the interactions, the tool-tip and both the pop-ups were also added using the same JavaScript function and HTML. Figure 23 shows a closer look of the Related Authors and Contributors facet from our final design with the tool-tip.

Figure 26 shows a snapshot of the large pop-up for the Related Journals facet. In this figure we can see our visualization design added to all the values of this facet. This pop can be used to analyze citations further down the list if needed.

Figure 27 shows a snapshot of the small pop up for one of the Related Authors and Contributors facet value Rupert A Hall. The functionality of the small pop up is to allow users to be able to better distinguish the color variations.

[Home](#)
[About](#)
[Blog](#)
[Help](#)
[Feedback](#)
[Log In](#)

Explore the IsisCB...

Person IsisCB ID: CBA000073903

Newton, Isaac

Show all 1111 citations to this authority. Show 30 citations as author, etc. Show 1102 citations as subject or category.

Birth to death dates: 1642-1727

Authority URI: <http://localhost:8000/isis/authority/CBA000073903>

Publication dates of IsisCB bibliographic records
(click to expand)

⚠ Updated timeline is loading. Please wait.
The server may be processing several jobs...

Related Authors & Contributors (See the top 100)

- 1 Newton, Isaac (30 citations)
- 1 Hall, A. Rupert (20 citations)
- 1 Shapiro, Alan E. (17 citations)
- 1 Erlichson, Herman (17 citations)

Related Journals (See the top 100)

- 1 Studies in History and Philosophy of Science (41 citations)
- 1 Notes and Records of the Royal Society (41 citations)
- 1 Archive for History of Exact Sciences (31 citations)

Related Publisher (See the top 100)

- 1 Cambridge University Press (23 citations)
- 1 Oxford University Press (11 citations)
- 1 Kluwer Academic (11 citations)

Datasets

- 1 Isis Bibliography of the History of Science (2 citations)

Outside Links (Show all)

- Q Stanford Encyclopedia
- Q Wikipedia
- Q Archive Grid
- Q Consortium of History of Science
- Q SNAC
- Q H-Net
- Q Twitter

Related Places (See all 26)

- 1 British Isles (40 citations)
- 1 Great Britain (35 citations)
- 1 England (20 citations)
- 1 France (15 citations)

Related Times (See all 14)

- 1 17th century (637 citations)
- 1 18th century (247 citations)
- 1 19th century (33 citations)
- 1 16th century (13 citations)

Related People (See the top 100)

- 1 Newton, Isaac (1102 citations)
- 1 Leibniz, Gottfried Wilhelm von (54 citations)
- 1 Descartes, René (31 citations)

Related Concepts (See the top 100)

- 1 Science (184 citations)
- 1 Mechanics (156 citations)
- 1 Physics (130 citations)
- 1 Mathematics (103 citations)

Related Categories (See the top 100)

- 1 comments on specific subjects (441 citations)
- 1 Seventeenth century (395 citations)
- 1 Science (176 citations)

Related Institutions (See all 11)

- 1 Royal Society of London (9 citations)
- 1 Cambridge University (7 citations)
- 1 Jesuits (2 citations)
- 1 Shanghai Lida Polytechnic Institute (1 citations)



Comments

Be the first to comment!

Log in or register to comment

Figure 24: The new IsisCB version after the addition of visualizations

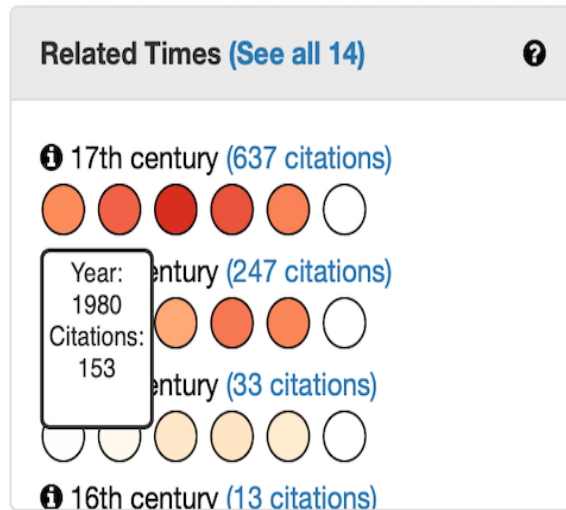


Figure 25: The circles variation of our design

4.3.6 Performance Analysis

To analyze our version of the IsisCB Explore user interface we go back to the set of general principles of a good quality user interface proposed by David Chek Ling Ngo [23]. In particular, we considered responsiveness, efficiency, familiarity and aesthetics.

According to Bhaskar et al [5], responsiveness is one of the general principles of a good user interface and it goes along with another one of his principles: Efficiency. A responsive system allows its users to perform tasks with efficiency. As visualizations were being added to each of the facets one by one, there was a noticeable degradation in the responsiveness of the website. By doing some basic bench-marking using lighthouse on Google Chrome, we found that the speed index of the Authority page was 35 seconds once all the facets had visualizations added to them. The reason for this was because we were accessing the data directly from the database and the counts of publications data was being calculated for every decade every time the page was being loaded. This was taking quite a bit of time and nothing was being displayed till all the calculations were done and all the visualizations were rendered.

The screenshot displays the IsisCB Explore interface with a 'Related Journals' pop-up window. The pop-up lists the following journals:

- History of Science (18 citations)
- Centaurus: International Magazine of the History of Mathematics, Science, and Technology (18 citations)
- Isis: International Review Devoted to the History of Science and Its Cultural Influences (17 citations)
- American Journal of Physics (17 citations)
- Journal of the History of Ideas (16 citations)
- British Journal for the History of Science (16 citations)
- Philosophia Naturalis (13 citations)
- Archives Internationales d'Histoire des Sciences (13 citations)

The background interface includes several facets:

- Related Authors & Contributors (See the top 100):** Newton, Isaac (30 citations), Hall, A. Rupert (20 citations), Shapiro, Alan E. (17 citations), Erlichson, Herman (17 citations).
- Related Journals (See the top 100):** Studies in History and Philosophy of Science (41 citations), Notes and Records of the Royal Society (41 citations), Archive for History of Exact Sciences (21 citations).
- Related Publisher (See the top 100):** Cambridge University Press (23 citations), Oxford University Press (11 citations), Kluwer Academic (11 citations).
- Related Categories (See the top 100):** comments on specific subjects (441 citations), Seventeenth century (395 citations), Science (176 citations).
- Related Institutions (See all 11):** Royal Society of London (9 citations), Cambridge University (7 citations), Jesuits (2 citations), Shanghai Lida Polytechnic Institute (1).
- Related Places (See all 26):** British Isles (40 citations), Great Britain (35 citations), England (20 citations), France (15 citations).
- Related Times (See all 14):** 17th century (637 citations), 18th century (247 citations), 19th century (33 citations), 16th century (13 citations).

Additional elements include 'Outside Links' (Stanford Encyclopedia, Wikipedia, etc.), 'Datasets' (Isis Bibliography of the History of Science), and a 'Comments' section at the bottom.

Figure 26: Snapshot of the large pop up for the Related Journals facet

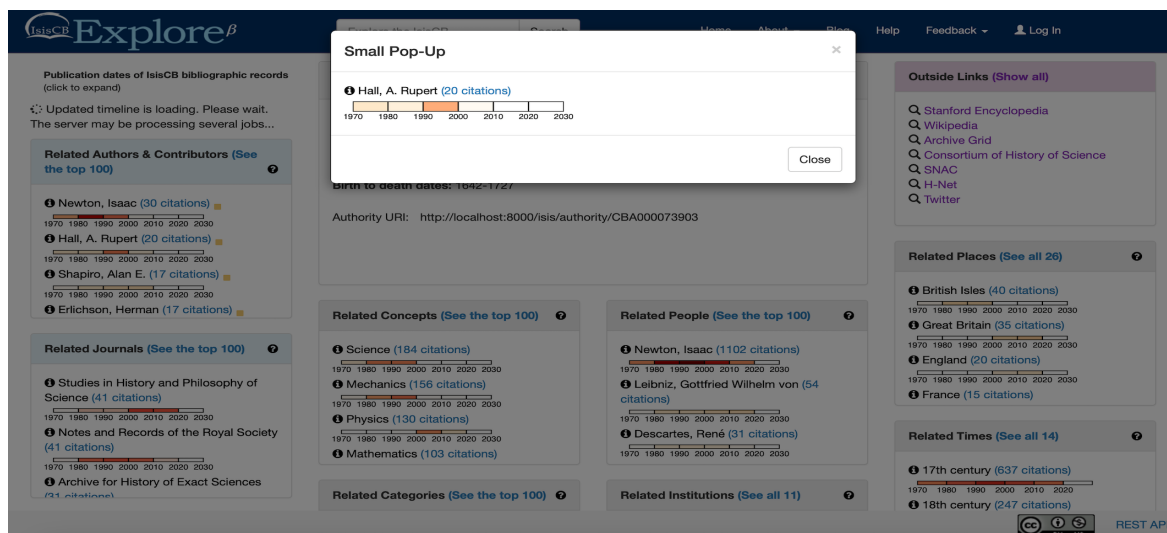


Figure 27: Snapshot of the small pop up for one of the Related Authors and Contributors facet value

Chang et al [11] created temporary tables with only the specific data that was required as a way to render the visualizations to deal with the large amounts of data produced when a wire transaction occurs. Drawing from this, we created a new table in the database, because the root cause of the issue was that the calculations were taking so long every time the page was reloaded. This way, the first time a particular page was loaded, the data calculation that are needed for the visualizations were performed and stored in a new table in the database. So, the next time this page was loaded, there was no need for the values to be again queried from the isiscbnew database and calculated as the data was already available in a new table in the database.

We also used Django URLs. Django uses the URL pattern to map the URL pattern to its Python method. This python method is the one that queries the data from the database and calculates the required data for the visualizations. Django quickly complies these methods and the visualizations are rendered as and when the data is calculated for each facet. Though it still takes some time for all the visualizations to be rendered, it won't feel very long anymore as we can see visualizations being added

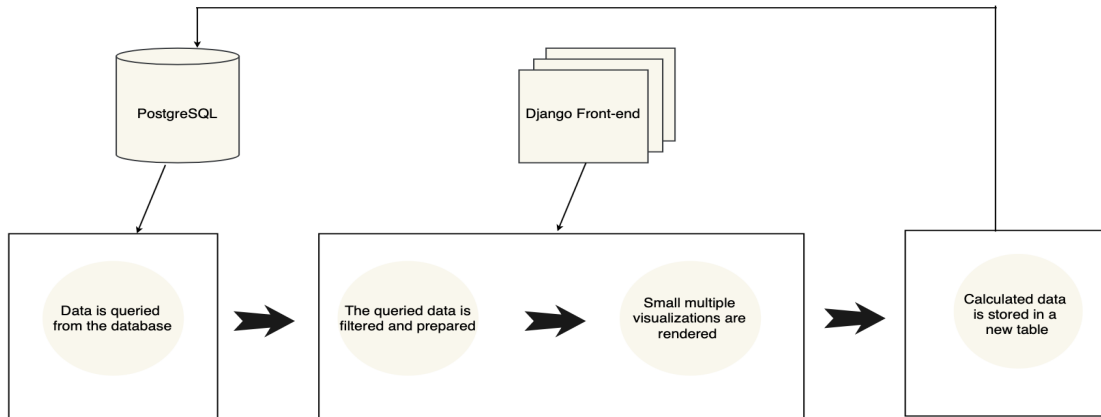


Figure 28: The updated stages of the implementation process

one by one to the facets. This is not a solution but just an alternative way to make the problem more acceptable. Once these changes have been made, the pipeline of the visualization development looks like in Figure 28. These changes improved the speed index of the Authority page to 6 seconds.

By making these changes, the system became faster than before with fewer lines of code and the user experience of the new version of IsisCB Explore was much better than before. This way the responsiveness and efficiency of the user interface was analyzed and concerns regarding it were addressed.

Analyzing our design according to the aesthetic principle, the factor we focused on was the color encoding we choose for our design. Though these were not consideration taken in to account while deciding on our color scheme initially, we are using this for our analysis. As mentioned previously, the current aesthetic of IsisCB Explore, Blue is considered to be a cool color. Cool colors induce a sense of calmness in the person looking at it and this makes for good user experience. According to an experiment on effect of colors in a user interface design, it was found that warmer colors such as

oranges and reds tend to foster responsiveness and trust within people [19]. Our design has a color scale that goes from pale yellowish orange to red and black. This aspect of our design aligns with the aesthetic principle where the user should have a positive experience while looking at the user interface.

The familiarity principle states that elements of the user interface must be familiar to the user. The IsisCB Explore interface had this to begin with and the addition of our small multiples within the facet boxes stayed within the bounds of this principle. By making sure no major changes were made while incorporating our design with the existing design of the IsisCB Explore user interface, we retained the familiarity of the IsisCB Explore user interface.

Based on our analysis, we maintain these principles of a good user interface in our version of IsisCB Explore. A user study evaluating this interface with regular IsisCB users would be a next step.

5 Conclusion

In this thesis, we proposed the addition of new designs to improve the usability of the bibliographic website IsisCB Explore. Our design provides a way to summarize the large amounts of time series data and present them in the form of visualizations. These visualizations in turn allow for improved user experience while performing analyses. The design was developed keeping in mind the space constraints of the IsisCB Explore user interface. We drew on the concept of small multiple visualizations and combined it with well-known techniques of visualizing time series data to develop a design that would fit well in the existing faceted user interface layout.

We anticipate our design additions to greatly improve the user experience by allowing our users to interact with the data at a higher level of overview than before. The different interactive abilities of the new design allows users to get summarized time related data of all the publications available in the IsisCB database. This would in-turn be helpful for users for various things like research on bibliographic data or for finding citations from a specific period of time. Our discovery study provided motivation for the additions and how they align with the areas of the IsisCB Explore interface that users visit and work with most often. Through this study, we were able to identify promising choices of design to integrate in to the authority page. Our choice of visualizing time-oriented data focused in particular on the decade time scale. We analyzed our additions to the IsisCB Explore user interface based on principles of good user interface design. We predict that our design additions will improve the user experience of IsisCB Explore by allowing users to gain more insights in to time phenomena in data than was practical before. This thesis contributes:

- An enhanced version of IsisCB Explore with new visualizations along with interaction capabilities implemented to support deeper analysis.

- A discovery study of the existing version of IsisCB Explore interface.
- Increased understanding of how visualization and in particular small multiples can be used in bibliographic user interface systems.

Some of the future additions that can be done to this are:

- A usability evaluation of the new version of IsisCB Explore along with a comparative analysis of the results from this study and the discovery study.
- Adding visualizations with more interactive abilities, like showing a yearly publications bar chart in one of the small pop-ups.
- Exploring more design variations for the existing visualization, such as line charts or heat maps instead of bar charts.

We anticipate that with the addition of new visualization designs and corresponding interactive capabilities, IsisCB Explore can transform from being a relatively typical model bibliographic website into an analysis tool that can be used by people from various backgrounds to better understand the bibliography of History of Science and its allied fields.

Slowly but surely, visualization has become one of the most common ways of presenting and analyzing data. Because of its ability to be understandable to people of many backgrounds, it has aided in eliminating the domain barrier that sometimes crops up when people from different fields are collaborating. In simple terms, visualizations have made data analysis available to all. We present our work as a step towards doing this in the bibliographic and in particular the IsisCB community.

References

- [1] Wolfgang Aigner et al. *Visualization of Time-Oriented Data*. London: Springer-Verlag, 2011.
- [2] Wolfgang Aigner et al. “Visualizing Time-Oriented Data—A Systematic View”. In: *Computers & Graphics* 31 (2007), pp. 401–409. DOI: 10.1016/j.cag.2007.01.030.
- [3] Paul S Barth. “An object-oriented approach to graphical interfaces”. In: *ACM Transactions on Graphics (TOG)* 5.2 (1986), pp. 142–172.
- [4] J Bertin. “Semiology of graphics: Diagrams, networks, maps.” In: University of Wisconsin Press., 1983.
- [5] N. Bhaskar et al. “General Principles of User Interface Design and Websites”. In: *International Journal of Software Engineering (IJSE)* (2011).
- [6] Caroline Birkle et al. “Web of Science as a data source for research on scientific and scholarly activity”. In: *Quantitative Science Studies* (2020). DOI: https://doi.org/10.1162/qss_a_00018.
- [7] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. “D³ Data-Driven Documents”. In: *IEEE Transactions on Visualization and Computer Graphics* (2011). DOI: 10.1109/TVCG.2011.185.
- [8] Lee Brasseur. “Florence Nightingale’s Visual Rhetoric in the Rose Diagrams”. In: *Technical Communication Quarterly* 14 (Apr. 2005), pp. 161–182. DOI: 10.1207/s15427625tcq1402_3.
- [9] Matthew Brehmer et al. “A Comparative Evaluation of Animation and Small Multiples for Trend Visualization on Mobile Phones”. In: *IEEE Transactions*

- on *Visualization and Computer Graphics* (2020). DOI: 10.1109/tvcg.2019.2934397.
- [10] Lee Byron and Martin Wattenberg. “Stacked Graphs – Geometry Aesthetics”. In: *IEEE Transactions on Visualization and Computer Graphics* (2008). DOI: 10.1109/TVCG.2008.166.
- [11] Remco Chang et al. “Scalable and Interactive Visual Analysis of Financial Wire Transactions for Fraud Detection”. In: *Information Visualization 7.1* (2008), pp. 63–76. DOI: 10.1057/palgrave.ivs.9500172.
- [12] Li Chengkai et al. “Facetedpedia: Dynamic Generation of Query-Dependent Faceted Interfaces for Wikipedia”. In: *Knowledge and Information Systems* (2010). DOI: 10.1145/1772690.1772757.
- [13] Marian Dörk et al. “A Visual Backchannel for Large-Scale Events”. In: *IEEE Transactions on Visualization and Computer Graphics (Proceedings of Visualization / Information Visualization 2010)* 16.6 (2010), pp. 1129–1138. DOI: 10.1109/TVCG.2010.129.
- [14] Nancy Fallgren. “Users and Uses of Bibliographic Data Background Paper for the Working Group on the Future of Bibliographic Control”. In: College of Information Studies, University of Maryland, Apr. 2021.
- [15] Martin Gasser et al. “Classical Music on the Web - User Interfaces and Data Representations”. In: *16th International Society for Music Information Retrieval Conference* (2015).
- [16] Iqbal Goralwalla A., M. Tamer Ozsu, and Szafron Duane. “An Object-Oriented Framework for Temporal Data Models.” In: Springer, 1998.

- [17] M Gusenbauer and NR Haddaway. *Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources*. 2020. DOI: 10.1002/jrsm.1378.
- [18] S. G. Havre, Whitney Hetzler, and Nowell PaulT. “Themeriver: Visualizing Thematic Changes in Large Document Collections”. In: *Visualization and Computer Graphics, IEEE Transactions* (2002). DOI: 10.1109/2945.981848.
- [19] Florian Hawlitschek et al. “Colors and Trust: The Influence of User Interface Design on Trust and Reciprocity”. In: *2016 49th Hawaii International Conference on System Sciences (HICSS)* (2016), pp. 590–599.
- [20] Sevinc Kurt and Kelechi Osueke. “The Effects of Color on the Moods of College Students”. In: *SAGE Open* 4 (Feb. 2014). DOI: 10.1177/2158244014525423.
- [21] Jock Mackinlay. “Applying a Theory of Graphical Presentation to the Graphic Design of User Interfaces”. In: *Proceedings of the 1st Annual ACM SIGGRAPH Symposium on User Interface Software*. Association for Computing Machinery, 1988. DOI: 10.1145/62402.62431.
- [22] Dominik Moritz and Danyel Fisher. *Visualizing a Million Time Series with the Density Line Chart*. 2018. arXiv: 1808.06019 [cs.HC].
- [23] David Check Ling No, Lian Seng Teo, and John G. Byrne. “Evaluating Interface Esthetics”. In: *Knowledge and Information Systems* (2002). DOI: <http://dx.doi.org/10.1007/s10115-002-8193-6>.
- [24] Mongeon Philippe and Paul-Hus Adèle. *The journal coverage of Web of Science and Scopus: a comparative analysis*. 2016. DOI: 10.1007/s11192-015-1765-5.
- [25] Catherine Plaisant et al. “LifeLines: Visualizing Personal Histories”. In: *Association for Computing Machinery* (1996). DOI: 10.1145/238386.238493.

- [26] W. Playfair. “The commercial and political atlas: Representing, by means of stained copper-plate charts, the progress of the commerce, revenues, expenditure and debts of England during the whole of the Eighteenth Century”. In: (1789).
- [27] Sandra Rendgen. *The Minard System. The Complete Statistical Graphics of Charles-Joseph Minard*. New York, Princeton Architectural Press, 2018. ISBN: 9781616896331.
- [28] B. Shneiderman. “The eyes have it: a task by data type taxonomy for information visualizations”. In: *Proceedings 1996 IEEE Symposium on Visual Languages*. 1996. DOI: 10.1109/VL.1996.545307.
- [29] John Stasko et al. “Jigsaw: Supporting Investigative Analysis through Interactive Visualization”. In: *Visual Analytics Science and Technology* (2007). DOI: 10.1109/VAST.2007.4389006.
- [30] Craig W Thompson et al. “Building Usable Menu-Based Natural Language Interfaces To Databases.” In: *VLDB*. 1983, pp. 43–55.
- [31] Christian Tominski, James Abello, and Schumann H. “Axes-based visualizations with radial layouts”. In: Association for Computing Machinery, 2004. DOI: 10.1145/967900.968153.
- [32] Edward R Tufte. “The Visual Display of Quantitative Information”. In: *Graphics Press Cheshire* (1983).
- [33] Daniel Tunkelang. “Faceted Search”. In: *Synthesis Lectures on Information Concepts, Retrieval, and Services* (2009). DOI: 10.1007/978-0-387-39940-9_159.
- [34] Frank A. U. *Different Types of “Times” in GIS*. In Egenhofer, M. J. and Golledge, R. G., editors, *Spatial and Temporal Reasoning in Geographic Information Systems*. Oxford University Press New York NY USA, 1998.

- [35] Javed Waqas, Bryan McDonnel, and Niklas Elmqvist. “Graphical Perception of Multiple Time Series”. In: *IEEE Transactions on Visualization and Computer Graphics (Proceedings of Visualization / Information Visualization 2010)* (2010). DOI: 10.1109/TVCG.2010.162.
- [36] Matthew O. Ward, Georges Grinstein, and Daniel Keim. *Interactive Data Visualization: Foundations, Techniques, and Applications*. 2nd. CRC Press, 2015.
- [37] Martin Wattenberg. “Arc Diagrams: Visualizing Structure in Strings”. In: *Proceedings of the IEEE Symposium on Information Visualization (InfoVis)*. Boston, MA: IEEE, 2002, pp. 110–116. DOI: 10.1109/INFVIS.2002.1173155.
- [38] Chris Weaver et al. “Visual Exploration and Analysis of Historic Hotel Visits”. In: *Information Visualization 6.1* (2007), pp. 89–103. DOI: 10.1057/palgrave.ivs.9500145.
- [39] Stephen P. Weldon. “Bibliography Is Social: Organizing Knowledge in the Isis Bibliography from Sarton to the Early Twenty-First Century.” In: (2013). DOI: 10.1086/673273.
- [40] Stephen P. Weldon. “The Isis Bibliography from Its Origins to the Present Day: One Hundred Years of Evolution of a Classification System”. In: *Circumscribere: International Journal for the History of Science* (2009).