UNIVERSITY OF CAPE TOWN

DEPARTMENT OF COMPUTER SCIENCE

CS Honours
Final Paper 2021

Title: The effectiveness of colour-icon pairs and skeuomorphic visualisations in the search results of digital libraries

Author: Dean Williams

Project Abbreviation: THETOOL

Supervisor(s): Hussein Suleman

Category	*Min*	*Max*	Chosen
Requirement Analysis and Design	*0*	*20*	0
Theoretical Analysis	*0*	*25*	0
Experiment Design and Execution	*0*	*20*	15
System Development and Implementation	*0*	*20*	15
Results, Findings and Conclusions	*10*	*20*	20
Aim Formulation and Background Work	*10*	*15*	10
Quality of Paper Writing and Presentation	*10*		10
Quality of Deliverables	*10*		10
Overall General Project Evaluation (*this section allowed only with motivation letter from supervisor*)	*0*	*10*	
Total marks		**80**	

The effectiveness of colour-icon pairs and skeuomorphic visualisations in the search results of digital libraries

Dean J. Williams
University of Cape Town
South Africa

Abstract

Digital libraries and academic search engines often have monolithic user interfaces and lack useful and sufficient visual features. This makes distinguishing between fields of academia very challenging and time consuming. In this project, an in depth analysis is done into the benefit of incorporating visualisations into the search results of a digital library. Also investigated, is the potential of improving user experience via an interactive book skeuomorphism. The search results mimic real life piles of books. Books are coloured and make use of iconography to best visually distinguish themselves from books of different genres. The results of this paper show that the visualisation greatly helped users find relevant results as well as improved the overall user experience.

Keywords

visualisation, digital library, digital libraries, library, academic search engine, subject classification, Dewey Decimal classification, DDC, skeuomorphism, research, ETD, NDLTD, electronic thesis, colour visualisation, iconography

1 Introduction

The invention and innovation of the Internet over the past 25 years has led to mass digitization of academic content. As a result, knowledge is almost entirely stored online in an electronic format.

Physical libraries are becoming less used over the years, as most academics move towards digital libraries as their go-to research source. However, while most digital libraries have extensive resources, they often lack sufficient user services [2]. While digital libraries provide many conveniences, the lack of sufficient user-centered services makes finding relevant publications a difficult and arduous task [2].

In order for society to progress, the efficient dissemination of knowledge across the world is paramount. As discussed by Bush, the ability to access and share such knowledge allows for progress in every facet of society [1]. One could say that the goal of all digital libraries, and those that work with them, is to provide humanity with quick and easy access to information, so that society may progress.

To help meet this ambitious (but necessary) goal, we be conducted an experiment related to enhanced user services for digital libraries. Our digital library of choice is the Networked digital library of theses and dissertations (NDLTD). The NDLTD is a network of many digital libraries across the world, with a database of metadata pertaining to over 6 million electronic theses and dissertations (ETDs). To access this vast amount of academic content, one must use the NDLTD global search – a simple search engine to interface with the networked digital library. Like many digital libraries, its contents are vast, but its user services are limited.

Currently, the NDLTD global search provides the following user services:

- The ability to query the metadata database using keyword queries.
- The ability to filter by date, source, tags and/or language.
- Tips on how to make "advanced" searches.
- Displaying of ETD metadata info as well as links to access the ETDs.

The vast amount of resources contained in the NDLTD make finding relevant publications a difficult task. Users often either have to make extremely specific searches or sift through hundreds of irrelevant ETDs, with no easy way of identifying what *is* relevant. The only way to do so is to read the title, description and/or tags of each individual ETD. Exacerbating this problem, ETD "tags" provided to the NDLTD come in a myriad of different formats (from Dewey Decimal, to hyphenated, to none at all). This inconsistency makes filtering by subject/tag an unfortunately incomplete and unreliable method of refining search results. It also means that there is a lack of consistency provided to the users, making identifying similar academic content much harder than it should be. In fact, the subject tags are so inconsistent that they are often left out from the search results entirely. A subject/tag string of 40 characters or more in length is ALWAYS discarded and NEVER visually shown to the user. This is especially detrimental since many NDLTD metadata-providers give subject tags as one, single combined string of many different subject tags. These composite-tags are usually over 40 characters long, and hence zero useful subject information is shown to the user.

Overall, the process of finding relevant publications using the NDLTD global search is far more difficult, time consuming and tedious than it should be. This research project aims to show potential solutions to these problems via enhanced user services. These services were thoroughly tested during development and were evaluated based on user feedback. Both numeric and qualitative feedback was collected to best determine the effectiveness of the features developed as well as best answer the research questions formulated. [1]

2 Background and related work

In a study by Theng, less than 50% of users found that digital libraries helped them achieve their goals quicker than searching traditional libraries [5]. Richardson et al states that the vast number of ETDs available makes it difficult for users to determine which dissertations to read [4]. Additionally, Buchanan et. al found that when information seeking moved from a strongly-defined goal into more uncertain areas, problems rapidly emerged [8]. Precise searches required careful selection of search criteria. This careful selection was only observed in a few users [8]. As discussed in the introduction, the NDLTD global search falls victim to this problem.
Sweetnam et. al. states that DLs are typically monolithic and difficult to navigate [3].

2.1 Visualisation

According to Richardson et. al., between 52% and 96% of academics are visual learners (depending on the field)[4]. It then comes as no surprise that many studies on user requirements, such as that of Sweetnam et. al.[3] and another study into what users want from digital libraries[11], state that one of the most requested features for digital libraries is a more visual based interface. One such visualisation that can be implemented is the visualisation of query results. Agosti et. al.[13] suggest that presenting search results in a more visual format should be more appealing and stimulating for novice users who access the system. They go on to suggest that users are more willing to learn and use visually interactive interfaces.
Studies have also found that visualisations allow for easier and more useful interaction[3] than that of text based interfaces [11]. Visual based interfaces can enable enhanced faceted searching of collections[3]. This has been identified[3] as an effective way of addressing the problem stated by Buchanan et. al[8].
Fox et al note that, with specific reference to the NDLTD, the structured display and visualisation of result sets should carry sufficient semantic information about the retrieved resources, which may be enhanced through the use of metadata[10]. Certain visualisations, such as map based[3] and/or force directed graphs, have been identified as highly useful forms of visualisations. However, these specific forms require a high computation time [4, 17].

2.2 Skeuomorphism

A "skeuomorph" is a visually displayed object resembling its real life counterpart. In many cases, users suggest that a DL should have features resembling a traditional library[3, 5, 11]. One highly requested, and important, feature of digital libraries – that is ubiquitous in traditional libraries – is that papers should be categorised based on their subject [11] and stored/displayed based on that category. Additionally, users want the ability to see *all* papers of the same subject[11].
Users also suggested the addition of humour to DL interfaces, perceived as "dry",

[1] Please note that due to last minute changes, some paragraphs are erroneously in future tense.

through the inclusion of appropriate animations[11].
Bainbridge developed a system to realistically visualise a book [15]. They state there to be many benefits in doing so, such as the fact that physical features of books help communicate ambient qualities of the document. They found that physical book models can increase the perceived value of a digital library system, since the experience of exploring digital library documents is often dull.

3 Research aim

From the background work done, we concluded that it was worth investigating the potential benefits of incorporating visualisation into the NDLTDs global search. However, due to the high computation times (and personal preference), we diverged from the typical visualisations. The benefit of incorporating skeuomorphism into the visualisation, mimicking a real life books (like Bainbridge).

3.1 Formulation

Background work has shown that a visual-based interface, alongside a visualisation, would prove highly effective. I hypothesize that the use of colour, imagery and skeuomorphism would greatly benefit users of the NDLTD global search. The plan is to group ETDs by genre, mimicking that of traditional libraries. Displaying ETDs as piles of library "books", with each book having colour schemes and imagery based on their genre, may allow for quick and easy identification of ETD subject matter. This may greatly aid in identification of relevant results, saving time that would have been otherwise spent tediously reading subject tags, titles and descriptions to yield the same effect.

3.2 Research questions

To determine the effectiveness of the system that will be developed, we shall compare it to the currently existing NDLTD global search's traditional, text-based, interface. To do so, we introduce the following three research questions (which will be referred to RQ1, RQ2 and RQ3)

3.2.1 RQ1: Genres

Will the grouping of ETDs by "genre" be a useful system of organisation?

3.2.2 RQ2: Visualisation

Will the visualisation be more useful for finding relevant ETDs than a text-based interface?

3.2.3 RQ3: Skeuomorphism

Will the skeuomorphic visualisation of search results as a "pile of books" provide a better user experience?

3.3 Methodology overview

The goal of this project is to determine the answers to the research questions. To do so, we developed a suitable system. After its completion, we conducted an experiment to obtain user feedback. Upon analysing said feedback, we discuss and determine the answers to each research question.

3.4 Expected results

The usage of skeuomorphism – in the form of search results imitating a pile of library books – as a means to aid user experience, is a very uncommon form of visualisation. Skeuomorphism is omnipresent in modern user interfaces, so we believe it has merits in its inclusion into digital libraries, which are often monolithic and thus antiquated in their design. In summary, we expect to find that:

- Users overwhelmingly agree that grouping-by-genre is useful.
- Users mostly agree that the visualisation is useful.
- Most users will enjoy the inclusion of skeuomorphism.

4 System design and development

To ensure the comparison between this system and the NDLTD global search was fair, we kept the latter's core functionality intact. The system developed, which shall be referred to as the "visualisation system" from now on, consists of a search bar, a filtering system and a way to access/display ETD metadata content. Any searches/queries the user makes interacts with identical metadata databases.

4.1 Metadata Storage

As previously mentioned, the database used to store ETD metadata is identical to the one used by the NDLTD. Said database is a SOLR, index of ETD metadata, spanning hundreds of digital libraries world wide.

4.1.1 Metadata format

Metadata is based on the Dublin Core standard (DC).

4.1.2 Limitations

Unfortunately, many NDTLD metadata providers (digital libraries that share ETD meta data with the NDLTD) give significantly less useful metadata than others. Some providers share their metadata without subject fields. Some leave out ETD descriptions/abstracts entirely. When subject data is provided, it is almost never in a controlled vocabulary, making machine readability difficult.

4.1.3 Solutions

A solution, to the *inconsistency problem*, that was implemented, was restricting the metadata "pool". This pool consists of *only* metadata provided by the British Library. This pool was also restricted such that only papers from the year 2000 onwards were included. The was further restricted so that only ETDs with non-null subject fields were shown. Additionally, choosing the British Library ensured that the majority of ETDs were in English.

All of this ensured that we were working with a collection of consistently formatted metadata. This pool contained almost three-hundred-thousand ETDs, more than enough for suitable testing. This was mostly due to the sheer size of the British Library's metadata contribution[2] as well as their relatively high consistency in metadata formatting.

Restricting queries to this pool was as simple as appending the following string to each SOLR query sent to the database:

```
subject:[* TO *] AND source_set_names:"Ethos UK"
AND date:[2000-01-01T00:00:00Z TO NOW]
```

This process is obviously abstracted from the user. While this solution is simple, it is not necessarily the best solution in terms of query speed. Adding this extra query criteria increases search/query time significantly. Ideally, the system would query a database that consists only of the restricted pool, keeping the search results fast while also restricted. This, however, would require creating an entirely new SOLR index, requiring the indexing process to be repeated. This is, unfortunately, unfeasible due to personal lack of SOLR experience as well as time constraints. This, however, should not be *too* much of a detriment to the experiment as only very broad searches (giving significantly more results) will have noticeable loading times. For the purposes of this experiment, we placed a maximum limit as to how many ETDs were retrieved per search.

4.1.4 Metadata retrieval

As previously mentioned, metadata was retrieved by querying the same SOLR metadata index that the NDLTD used. When a searches, a query is sent to the server-side database (with the previously mentioned restriction criteria) and retrieved by the client as JSON data. This JSON data is processed via JavaScript into manageable lists containing the metadata. These lists may be easily managed client-side. Making up for the, sometimes, slow retrieval speeds, these lists provided extremely fast filtering and navigation capabilities.

4.2 Classification and grouping of genres

Subjects were given in DDC, so ETDs were very easily be grouped and displayed by genre.

Client-side dictionaries[3] allow for quick conversions of Dewey Decimal codes into their full English subject names. The subject metadata stored in DDC is fed

[2]Ethos, or the electronic thesis online service, is one of the NDLTDs largest metadata providers
[3]These dictionaries were generated via a self developed python script that takes machine readable text data from wikipedia and output a JavaScript readable dictionary

through the dictionary, returning both the full genre name (eg. Physics) as well as its sub-genre (eg. Classical mechanics). It is worth noting that DDC's inception was almost a hundred years ago, and while it is being updated annually, most consider its structure to be quite outdated. So in an attempt to account for this, we merge many genres together (eg. Language and Linguistics are classified as two separate genres in DDC, but we merge both into the single subject "Language and Linguistics"). I have manually edited the dictionary so that it contains 52 genres mapped from the original 100. However, the just under 1000 sub-genres have been left with their original names.

4.3 Visualisation design

Now that a genre-dictionary has been established, we may use it to quickly and easily provide visual distinctions for each ETD genre. We shall now begin developing an efficient design scheme to procedurally visualise the 52 different genres.

4.3.1 Colour-icon pairs

To visually distinguish between different genres, we apply the following principles.

- Each genre must have a colour scheme to distinguish themselves.
- Each genre must have an associated image-icon, so that it may be easily visually-identified without having to read any textual descriptions.

In accordance with these principles, unique and visually distinguishable colour schemes were manually identified and taken from the colour-scheme assisting website, colorhunt[20]. Subject colours were saved in a key-value dictionary, with the subject name being the key and the chosen colour scheme being the value. Colour schemes are made up of 4 hexadecimal colour strings. These strings were mostly chosen with subject name in mind (for example, plant/nature related subjects like biology or earth sciences were given green colour schemes), however, they were primarily chosen so as to provide a noticeable visual distinction between different subjects. This was done relatively quickly, a simple process of saving 52 different colour scheme strings. I will note that, for ideal results, colour schemes should be picked more carefully, by people with extensive design and colour-theory experience.

Additionally, icons were retrieved via the free icon website: flaticon[21]. Each of the 52 subject icons was individually picked so as to best visually convey their subject matter. Subject icons are stored and saved as their respective genre name, so that once again they may very easily and quickly be retrieved via the dictionary.

To ensure that the colour-icon pairs effectively conveyed their subject matter, I went through each pair and determined if I was able to identify the exact name of the subject without reading any text. Some icons were then replaced to better visually convey their genre. See examples below.

Figure 1: example icons: astronomy; biology; chemical engineering; commerce and travel; earth sciences; engineering; entertainment; history; language and linguistics; logic; medicine and health; physics; and religion.

Combined with colour schemes, these color-icon pairs shall provide not only visually distinguishing features, but also quick identification and recall of subject matter.

4.3.2 Book Skeuomorphism

To best approach the procedural generation of books, I made careful study of what book spines looked like[23] and what necessary information they conveyed. I determined the following:

- Books should have eye-catching and readable titles.
- Books should show the author's name.
- Books should vary in size.
- Books should vary in colour.
- Books should contain imagery.
- Books should contain "stripes" or lines to aid with colour scheme and visually distinguishing features.
- Most importantly, books should *look* like books. They should appear 3D with rounded curves and appropriate shading.

Incorporating both the identified visually-essential features of a book, as well as the colour-icon pairs discussed previously, yielded a pleasant looking and satisfying book visualisation (see example books below).

Figure 2: example entertainment book

Figure 3: example zoology book

Shown above are two ETD search results with completely different subject matter. Both, however, are retrieved when searching the metadata database for "cricket". Looking at the two, it is quickly and easily apparent which of the two is related to the sport, and which of the two is related to the insect.

Both examples were generated via the visualisation system automatically. Their colour-icon pairs were generated and assigned via the dictionaries. Colour-scheme strings were ordered so that the two highest contrast colours, were the colours of the title-background and the title-text. This way, the titles would always pop out and be readable on a web-browser. A list of unimportant words that could be present in titles (such as words like "of" or "and") was made. Whenever a title word matched a word in the unimportant list, it was made to have a smaller font-size. This way, titles mimic the eye-catching nature of standard book titles. A serif font, Georgia, was chosen as the title font, due to its formal looking nature.

A rather surprisingly complicated process was the resizing of fonts so that the titles fit their books. Since books have varying heights as well as varying lengths of titles, there was no one-size fits all for the book fonts. A quick solution was to set each font to a large size, then shrink it with JavaScript until the full title fits (i.e. its HTML offsetHeight was less than the other's). This was paired alongside a scheme that made books, with larger titles, bigger than books with smaller titles. To finish the title scheme, an algorithm is run to procedurally shorten book titles (removing parenthesis and quotations as well as splitting compound titles into two and picking the "best" one).

4.3.3 Book animations

To help tie in the skeuomorphic visualisation as well as provide humour to the system, a falling and bouncing animation was developed. When a user is shown search results, the books fall from a point above the screen and bounce when they hit the floor, landing in two separate piles of five. The books have slight x-coordinate variations (randomly generated before falling) so as to give the appearance of a more realistic looking pile.

Each books falling velocity, V_f, is based off the $1D$ equations of motion. This simulates a life-like gravitational fall. Similarly, when the books hit their respective "floors", their V_f is negated and scaled down slightly, allowing for the equations

of motion, and gravitational acceleration, to naturally create a "bounce" effect.

As a side note, at one point certain atypical interactions with the web-application (such as resizing the view-port during the bounce animation) caused extremely high bounce-velocities. Books would, humorously, fly off the screen at supersonic speeds. The simple fix was setting a maximum, terminal, velocity.

4.4 User Interface

As mentioned before, the user interface was designed to maintain the core-functionality of the NDLTD global search. As such, we shall only discuss new/modified user interface features. One key difference, is that due to the humorous nature of the book visualisation, the user interface has been designed to match this theme. Overall, the user interface has a rather jovial nature to it.

4.4.1 Subject Filters

When a user makes a search, a list of subject filters is generated and displayed to the left side of the screen. These filters mimic the appearance of the books that they may be used to filter. They have the same shading and curvature as the $3D$ books. These little book-looking filters consist of a genre-name and colour-icon pair. The filters are ordered by frequency in the query results, with the most common genres displayed at the top and the least common at the bottom. This way, users may more easily and effectively filter search results, since the more frequently occurring genres are likely to be the most relevant.

Users may interact with the filters in two different ways:

- Users may click the body of a filter, toggling all other filters on or off.
- Users may click the icon of a filter, toggling that specific filter on or off.

To make this as user friendly as possible, when hovering over a filter, it moves slightly and is highlighted in white. Additionally, hovering over the icon turns the icon into a cross or a tick depending on its toggled state. Hovering over either shows a tooltip, explaining what the respective user action will do. In addition, if a filter is toggled off, it is made transparent to show that it is inactive. By default, all filters are active.

Toggling a filter off will ensure that any ETDs of that genre will be removed from the displayed search results. This way, a user may easily filter out/in genres that they want, ensuring only relevant academia is shown.

With further development, these filters could be made storable/saveable so that a user may retrieve previous filter configurations to help streamline the research process. However, this has, regretfully, not been implemented due to time constraints.

4.4.2 Navigation buttons

To navigate between piles of books, 4 navigation buttons are provided to the user. Back and forward buttons allow users to go backwards and forwards through piles. The back to beginning and go to end buttons allow the users to navigate to the first and last piles respectively. Each time one of these buttons is clicked, new books are shown on the screen, following the standard falling and bouncing animation developed.

4.4.3 Search results

The center pane of the screen contains the search results. In other words, the piles of books retrieved from the database. To aid with user experience, hovering over a book highlights it with a white background and makes it move slightly to the left, similar to hovering over a filter. Similar to clicking a filter, clicking a book makes all other books transparent, bringing focus to the book that was clicked. Clicking a different book makes the new book opaque and the previously clicked book transparent like the rest. Clicking away removes the transparency effect giving focus back to the piles of books.

4.4.4 Book information

When a book is clicked, its information is shown on the right side of the screen. This information includes the ETD's shortened title (with font and colour mimicking that of the book displayed) as well as the ETD's full title, its author and source, its subject, its description and the date it was created. All fonts (except for the book titles) are open-sans. This is due to its readability, especially in small sizes. A link to where the user may access the original ETD is displayed under the author and source in a large and noticeable font. Clicking said link opens a new tab in the user's browser. In this new tab will be the website containing the electronic/pdf copy of the ETD. This way users will not lose their progress on the visualisation system upon accessing the ETDs.

4.4.5 Tooltips

Another feature included to help aid usability and user experience, is the addition of tooltips. Tooltips simply provide textual information when hovering over certain user interface objects (such as the filters, books and navigation buttons). Tooltips may be toggled off or on at any point using the inconspicuous Tooltips checkbox. The tooltips checkbox is half transparent by default. Only upon hovering over does it become fully opaque. This is done to reduce "noise".

5 Experimental design

Since the primary goal of this research paper is to answer the three research questions, we design the experiment to best test them. Since the developed system's effectiveness is subject to user feedback, we will be conducting user surveys. Users will be given a list of tasks to perform, with both the NDLTD and the visualisation system. Upon completion of the tasks, they will complete a questionnaire. This process will be streamlined through the usage of limesurvey.

5.1 Participants

Since the purpose of this paper is to determine the effectiveness of incorporating a visualisation into academic search engines, the target market are academics. Participants will have the requirement that they are currently studying towards or have achieved a degree. We shall attempt to gather many Honours students (or above) as the primary participants. For a decent enough sample size, we would like to gather at least 30 full responses.

5.2 Tasks and instructions

Users will be tasked to perform instructions so that they have sufficient learned information about the system to best answer the questionnaires. To ensure that users make sufficient usage of both the NDLTD global search and the visualisation system, they will be tasked with the same instructions for each system. However, to present the participants with the benefits of the visualisation system, we will task participants in performing a scenario that showcases said benefits. A good example is the previously mentioned case where a user searches "cricket". Making such a search will yield 2 vastly different types of academic content: that which relates to the sport and that which relates to the insect. Using the NDLTD global search to distinguish between the two types is quite difficult. One has to read the title, description and tags of each search result just to determine whether its about sport or the insect. This is really time consuming for something that should be made much easier. Thus, to show this to the user in a fair and controlled way, they shall perform the same instructions with each system.

Users will begin by searching for ETDs on the sport, cricket, using the NDLTD global search. They will be asked to identify search results that are sport related and note how easy/difficult that was. They will then be tasked to filter the search results so that they are confident no insect related search results will be displayed. They will then switch over to the visualisation system, performing the exact same task of finding cricket (sport) related ETDs. Ideally, the users will notice that it was far easier to perform this filtering functionality with the visualisation system.

Upon completion of both these tasks, the users will be asked to "mess around" with the filters/visualisation of the developed system. They will be given some tasks, such as "Attempt to filter-out all subjects except a subject of your choice", so that they may get a general feel of the system and the experience it provides.

Users will have experience using academic search engines and digital libraries, not only because they are all academic students, but because they will have used a separately developed system (a user-generated list system by Hugh Bedford, hereafter referred to as "the user-list system") prior to using both the global search and the visualisation system. This should help reduce the learning affect that could arise, that would cause potential biases towards either system. If anything, users should have a slight bias against the visualisation system as

it massively deviates from the norm in terms of traditional text-based research systems. To account for this, users should be given sufficient time to play with the visualisation before answering the questionnaire. Any users who fails to do so will be noted when analysing results.

5.3 Questionnaires

In order to answer the research questions, we require user feedback. We require both numeric feedback as well as qualitative feedback. Additionally, so that we may best interpret said feedback, we ask for some very basic, but non-identifying, participant information. This way, feedback may remain anonymous, while also shedding light into different types of users. Participants will be explicitly told that any feedback they provide shall be completely anonymous, ensuring that they will provide honest feedback.

5.3.1 Participant information

Participants will be asked the following three, non-identifiable, personal questions.

- Which faculty are you in?
- Which discipline or field do you currently study/work in?
- What qualification are you currently working towards?

These questions are designed to provide us with useful information in analysing results, while still keeping the participants entirely anonymous.

5.3.2 USE Questionnaire: Usefulness, Satisfaction, and Ease of use

A well documented and widely used questionnaire to quantitatively analyse usability of software systems is the USE Questionnaire[22]. This questionnaire consists of 29 questions, all designed to best test the usability of a system. It provides numeric feedback, scoring the system's usefulness, user satisfaction and ease of use. Participants may provide a low score representing a strong disagreement and a high score representing the opposite.
Sufficient answers to these questions should provide *almost* enough data to analyse, so that we may best answer our research questions.

5.3.3 Designed questions

Key-word being *almost*. In addition to the 29 USE Questionnaire questions, we ask participants for their stance on the following two statements

- The visualization was useful.
- The book genres were useful.

These questions are a modification of a question, from the original questionnaire, "it is useful". This way if a participant, for any reason, feels that the visualisation was not useful, we may at least see their stance on the usage of book genres (and vise versa). Ideally, we would like positive responses to both statements. The visualisation system is tied to the book skeuomorphism, so answers to the first statement shall be indicative of both.

Both the use questionnaire and the specifically designed questions will be scored by the user on a scale of 0 to 10. Zero being a strong disagreement, 5 being a neutral stance and 10 being a strong agreement.

5.3.4 Long answer questions

We asked participants 4 long answer questions:

- What feature did you enjoy the most? Why?
- What feature did you enjoy the least? Why?
- If you could change one thing about this system, what would it be and why?
- Any general comments about this system?

These questions will help us identify reasoning behind their numeric responses.

6 Results and analysis

In total, 42 participants were gathered of which 30 provided full responses and only 28 of these were deemed legitimate"Hi, bye, hello world. U19 provided long answers suggesting they did not use the visualisation system at all, just the global search. They mention features like advanced search tips and unhelpful subject tags. As a result, both users responses were discarded". See the appendix for pie charts showing participant composition:

6.1 Qualitative analysis

Long answers were analysed using open coding. They were read one by one, in order of response ID, and whenever a new concept/idea was expressed by a user, a new unique code was created. After going through each user, the users were analysed again and tagged with any matching codes. See figures 4 and 5 for a full list of codes and tagged frequency.

Table showing the openly assigned positive codes, as well as their description and frequency

CODE	Description	Frequency
'VISUAL'	the visualisation or its features	18
'SUBJCT'	the usage of book genres/subjects	18
'ENJOYM'	praising the system's enjoyment/fun	15
'FILTER'	the filtering system	14
'HLPFUL'	the system or a feature was helpful/useful	14
'RELVNT'	relevant results/academic content	11
'COLOUR'	the usage of colour	11
'COLCOD'	the usage of colour coding	8
'EASUSE'	the system or a feature was easy to use	8
'VISEFF'	the visual effects/animations of the visualisation	6
'EASIER'	the system or a feature was easier to use	6
'ICONSV'	the usage of icons	6
'INTERA'	features were interactable	3
'SAVTIM'	explicit mention of saving time	2
'TOLTIP'	the tooltips	1
'MATURE'	the childlike nature of the system	1

Figure 4: positive codes in order of frequency

Table showing the openly assigned negative codes, as well as their description and frequency

CODE	Description	Frequency
'FILFET'	a feature of the developed filter	7
'USRFRN'	the user-friendliness of the system	4
'INTERF'	the system's user interface	4
'MATURE'	the childlike nature of the system	4
'FEWRES'	a lack of visible results on screen	3
'FILUSE'	difficulty using the filter/did not notice functionality	3
'LOADIN'	the loading times for searches made	2
'VISEFF'	the visual effects/animations of the visualisation	2
'FILLIM'	the filters were limited in variety	2
'FILVIS'	the visualisation of the filter	1
'COLOUR'	the usage of colour	1
'TOLTIP'	the tooltips	1
'FILREC'	difficulty recovering previous filter configs	1

Figure 5: negative codes in order of frequency

The experimental features developed (SUBJCT and VISUAL) were the most positively tagged codes, and hence the most enjoyed features of the system.
The next highest was ENJOYM, with 15 tagged users, showing that the majority of participants had an enjoyable user experience.
11 users were assigned the RELVNT code, showing that they found the system to provide more relevant results than the NDLTD's global search.
6 users specifically mentioned enjoying the visual effects of the system's book skeuomorphism (like the falling animation).

6.1.1 Quotes to support RQ1:

U7 — "this is a useful system for finding resources"
U18 — "I enjoyed how easy it was to filter what category I wanted"
U24 —"the original ETD search for cricket is illustrative of a common issue which [filtering by subject area] solves"
U38 — "visual differentiation of subjects is extremely helpful" — "it helps me to skim read much faster" — "it saves me time having to work out what is relevant and what isn't".

6.1.2 Quotes to support RQ2:

U12 — “The colour-coding of papers by subject [...] made it far easier to filter out irrelevant results”
U13 — “Greatly helped with finding topics”
U16 — “the genres and the colours [...] make navigation easier”
U27 — “visual labelling really helped a lot to refine results!”
U29 — “the visualisations [...] made it a lot easier to find a certain topic”
U30 — “[the icons] really help hammer home the visual relationship I have with these subjects”

6.1.3 Quotes to support RQ3:

U3 — “visual effects made it more pleasant” — “seems gucci”
U6 — “rather splendid”
U13 — “Great visualization, easy to use and read”
U16 — “flawless” — “the bouncing animation is quite a nice touch”
U18 — “amazing” — “perfection”
U22 — “the icons of the different categories made the experience pleasant”.
U23 — “fun and interactive”
U24 — “very efficient and fun to use”
U27 — “visually, the whole set up was entertaining” — “I really enjoy[ed] the visual graphics and the thought of the whole page” — “I wish there were more easily accessible examples out there like this”
U30 — “fantastic to use”
U31 — “this stuff is the bomb-diggity” — “visual aspect is both crisp and clean” — “it breaks up the monotony of viewing academic journals” — “Add more colours. Apple red would be appreciated.”
U37 — “the visualisation of the books falling down [...] was fun”.
U38 — “looks great!” — “its fun to use”.
U39 — “I really enjoyed the book system” — “and how [it] looked visually”. — “good stuff”

6.1.4 Negative feedback and constructive criticism

The most common negative codes were for features of the specific filtering system developed, as well as user friendliness and the user interface.

A few users found the system to be too “childish”. This was a particular problem for U42, stating “it felt a little childish as an ETD search engine”. They state that they personally prefer the visualisation system to the NDLTD, but that they feel it needs to “look a little more sophisticated”.

Both U22 and U41 did not enjoy the visual effects. U41 states that there is “too much noise and clutter”, with reference to the animations. While U41 was a big fan of the usage of colour and visualisations to differentiate subject matter, stating “the function behind it is strong”, they felt the system to be visually overwhelming, stating “it needs to be represented in a more user friendly way”.

Some users state the user-interface-layout to be their least favourite feature. U9 in particular was heavily influenced by this in their numeric responses.

While mention of loading times was to be expected, it was only seen in the responses of U4 and U12.

7 users felt that there were features of the filtering system that needs improvement (tagged with FILFET).
3 users had difficulty using the filters (tagged with FILUSE). 2 of which were identified to completely miss the dual toggling capabilities of the filtering system. U3 stated they did not like that “you cannot just select two categories you want to scroll through you have to manually deselect all others”. This is of course not true, the filters have dual toggling capabilities, but is indicative that the filters’ functionality is not 100% obvious at first glance. U12 stated “I would make it such that it was possible to have multiple filters active simultaneously” in response to what they would change about the system, indicating again that the filter’s functionality needs to be made more obvious on first glance.
U27, another user tagged with FILUSE, stated “with out written instruction [a] user would have to play around a bit with how to reset [the filters]” and that they “wouldn’t have known [how to do so] unless told”. They suggest that “maybe it could be mentioned with the tooltips”. This was a serious blunder on my part. The tooltips *do* mention and explain the filters’ functionality, but in an attempt to make the system more user friendly, the tooltips were made as concise as possible (see system design and development section). I was not aware of the fact that “toggle” is not a well known word, and thus the concise tooltips contained confusing jargon for many users.
This is unfortunate since once a user knows how to use the filtering system, its capabilities are very strong and easy to use. For example, U20 states “The ability to double click and reset filters, As a user this feature is quick and easy to use” — in response to the feature they found most enjoyable.
U20 really enjoyed the ease of use of the filters and their dual-toggling capabilities, but state “some icons are difficult to see and recognise easily and some colouring of subjects are similar”, referring to the appearance of the filters.
U29 wanted the ability to search for specific filters. U37 wanted the ability to save/go back to previous filter configurations, stating “I did not like that if I misclicked a filter I could not go back to what the filters were before”. As discussed previously, this is a feature that I would have liked to have added, but could not due to time constraints. U29 and U38 wanted more in-depth filters, not limited to the 52 subjects currently available. U38 states “I feel as though the usefulness of tags cap out” and that “more specific tags might be nice to have”.
U6, U12 and U24 all stated that they would like to be able to see more than 10 results on the screen at once.

6.2 Numeric feedback

To reiterate, there were 28 legitimate responses to the surveys. Users were shown a statement and asked to rate their thoughts on the statement from 0-10.
From the use questionnaire, all ratings provided for each question have been used, except for questions 4, 5, 13 and 29, which many users found to be too vague and provided highly spread answers. See the Use questionnaire [22] for the specific questions used.

Figure 6 is a graph containing the cumulative user ratings for each section of the use questionnaire. Each bar is measured as the percentage frequency of occurrence for that rating in that specific section.

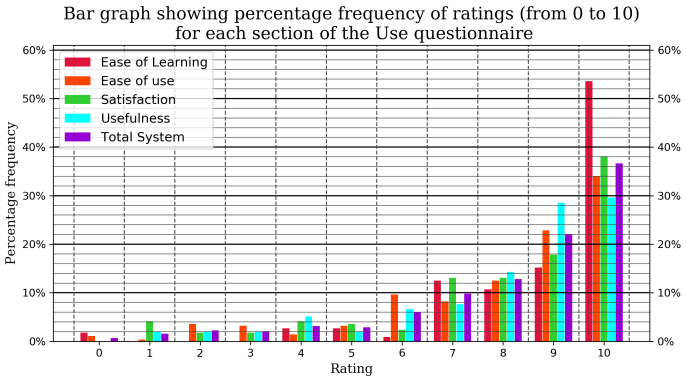

Figure 6: percentage frequency of ratings per section.

In general, frequencies of ratings are in ascending order, with 10s and 9s being by far the two most frequently given. There were very few ratings less than 5/10. The Total System score (i.e. the cumulative ratings for all 27 questions) shows a consistent upward trend. 6% of ratings are 3 or below, 12% are between 4 and 6, 23% are 7s and 8s, and 59% are 9s and 10s. This gives a median system rating of 9/10. Thus, the vast majority of users were very happy with the system developed, and only a few ratings given showed otherwise.

6.3 Combined analysis

For each user, the table (figure 15 in the appendix) shows their average ratings for Usefulness (**U**), Ease of Use (**EoU**), Ease of Learning (**EoL**) and Satisfaction (**S**). Then, it shows the, equally weighted, mean of these four sections, as the total system score (**Tot**). After that, the table also shows the ratings for each of the designed questions (**Vis** and **Gen**). The last numeric column shows the rating for question 26 (**Fun**). Then, the last two columns show the **Positive Codes** and **Negative Codes** assigned through open coding.
Below the user ratings, the mean (**Men**), median (**Med**), mean absolute deviation from the median (**MAD**), inter quartile range (**IQR**) and standard deviation (**StD**) are shown for each numeric column.

6.3.1 System colour gradient

The table is numerically colour-coded (see the below key) and is almost entirely made up of lush greens. This shows that the majority of users were extremely happy with the system developed. The only notable exceptions (outliers) are U4, U8, U9, U41 and U42, which we will discuss in detail in a later subsection.

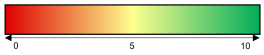

Figure 7: ratings gradient-key

6.3.2 Variability

While we want to measure the variability of ratings in our dataset, we cannot blindly use standard deviation to measure variability, since most sections in the dataset are not normally distributed[24, 25] (see figure 6, all sections have upward trends). Thus, to more accurately determine the variability of ratings, we use the mean absolute deviation (MAD), taken from the median. Since MAD is calculated using absolute differences, it is less affected by outliers. This more robust measurement of variability shows how far, on average, the typical datapoints are (on either side of the median).

6.3.3 Outlier analysis

There are 5 outliers which are very visible in the table as the only users with yellow/orange cells in the Tot column. In terms of their Tot ratings, U4, U8, U9, U41 and U42 were 2.8, 2.8, 5.1, 3.9 and 2.5 mean absolute deviations from the median. This is enough to declare them all outliers as they are far outside the typical user ratings.

6.3.4 Not outliers?

We note that U18 and U31 provided entirely 10s for every question in the Use questionnaire. While this may appear as outlier behaviour, their Tot ratings are actually within 0.74 mean absolute deviations from the median, suggesting that their ratings are quite typical. Both users provided very in depth and well thought out long answer feedback and both users said that they would not change a thing about the system.

6.3.5 Usefulness

The typical U ratings are between 7.5/10 and 10/10.
Only 1 typical (non-outlier) user fell outside this range (U7).

In addition, the typical Vis rating is between 7.32/10 and 10/10. Only 2 typical users fell outside this range (U22 and U32).

The typical Gen rating is between 7.9/10 and 10/10, with only one typical user (U20) falling below this range.
The mean scores for each of these 3 columns mentioned are 8, 8.3 and 8.7, suggesting that even amongst outliers, the system and the features developed were on average very useful.

6.3.6 Ease of use

The typical EoU rating is between 6.9/10 and 9.9/10.
Only 2 typical users fell below this range (U12 and U29).

6.3.7 Ease of learning

The typical EoL rating is between 8.4/10 and 10/10.
Only 3 typical users fell below this (U7, U29 and U32). Since this MAD is the lowest of the 4 sections, and thus the most consistently highly rated section.

6.3.8 Satisfaction

The typical S rating is between 7.5 and 10/10. The number of typical users falling outside this range was only 2 (U3 and U7). This section has some of the highest number of 9s and 10s, but also some of the highest number of 7s. S has the highest IQR and MAD, so it had the highest spread of ratings.

Even amongst the outliers, the mean satisfaction score is 8/10, which indicates a a satisfying user experience.

6.3.9 Fun!

The typical user rating for Fun is between 7.5/10 and 10/10. Zero typical users fall outside this range and zero provide a fun score of 6 or less. Thus numerically speaking, the system is very fun to use in terms of a research tool.

Even amongst the outliers, there was only 1 user that rated the system below a 5/10 (U41, who is currently not studying so perhaps they did not have much to gauge the fun relative to other research tools).

6.3.10 Overall total system rating

The median Tot rating is 9/10, with a MAD of 1.35/10. We see lush greens throughout this column, with exclusion of outliers. The typical users' lowest Tot score is 7.5/10 (U29) and the highest is 10/10 (U18 and U31). Overall, users were very happy with the system developed.
Even amongst the outliers, there were only 2 users with a score below 5/10. The mean Tot rating is 8.2/10, showing that the average rating for the system (overall) was very high.

6.3.11 Constructive feedback

Careful attention will be made to any typical users that provided sectional ratings significantly lower than their Tot ratings, to find problem areas and constructive feedback. To do this, a table (figure 8, shown below) is generated. This table shows only the 23 typical users.

Table showing the difference between total system ratings and sectional ratings for each typical user

ID	U	EoU	EoL	S	Tot	Vis	Gen	Fun
3	0	0.2	0.5	-0.83	8	0	0	-1
6	0.1	0	-0.15	0.1	9.9	0.1	0.1	0.1
7	-0.74	0.9	0.4	-0.43	7.6	0.4	0.4	0.4
11	-0.01	-0.2	0.7	-0.3	8.3	-0.3	-0.3	-0.3
12	0.91	-2.5	1.2	0.37	8.8	1.2	1.2	1.2
13	0.03	-0.1	0.35	-0.23	9.4	0.6	0.6	-0.4
16	-0.74	0.4	0.4	0.07	9.6	-0.6	-0.6	0.4
18	0	0	0	0	10	0	0	0
20	-0.14	-0.7	1	0	9	1	-2	1
21	-0.24	0	-0.1	0.4	9.1	-0.1	-1.1	-0.1
22	-0.09	-1	1.2	-0.13	8.8	-3.8	1.2	1.2
23	-0.36	0	0.5	-0.33	9.5	-0.5	-0.5	-0.5
24	-0.66	0.2	0.2	0.2	9.8	-0.8	0.2	0.2
27	-0.11	-0.2	0.1	0.1	9.4	0.6	0.6	0.6
28	-0.43	0	1	-0.5	8	1	1	0
29	0.79	-1	0.25	0	7.5	2.5	0.5	0.5
30	0.29	-0.8	0.25	0.33	9	1	1	1
31	0	0	0	0	10	0	0	0
32	-0.23	0.2	-0.05	0.03	7.8	-0.8	0.2	-0.8
35	-0.13	-0.1	0.3	-0.2	9.7	0.3	-0.7	-0.7
37	-0.46	-0.2	0.4	0.4	9.6	0.4	0.4	0.4
38	-0.86	1	1	-1	9	1	1	1
39	-0.31	-0.6	0.4	0.4	9.6	0.4	0.4	0.4
Men	-0.15	-0.2	0.428	-0.07	9.017	0.157	0.157	0.2

Figure 8: the differences between sectional ratings and Tot ratings for each typical user

The above table has identical columns and headings to Figure 15, but numbers shown are the difference between a user's Tot rating and their average numeric rating for that section/question. The table is also colour coded. Pink cells indicate a section that the user rated less than their Tot rating and blue cells indicate the opposite.

The majority of users with very pink cells were tagged with FILFET (U3, U12, U20, U29, U37 and U38) suggesting that many of the users problems were with the filtering system developed. Some others were tagged with specific features of the visualisation like COLOUR or ICONSV (Refer back to section 6.1.4 for more).

6.4 Outlier removal

For the rest of the numeric analysis, the sample studied excludes outliers. This sample contains 23 participants, all of which were deemed to be representative of the typical users and typical user ratings. This has been done to provide a better representation of the typical user results, without the outliers affecting the means too much, so that problem areas may be better identified.

Since the number of ratings at or below 6/10 are very minimal, their cumulative frequencies are combined into one group, labeled ≤ 6 (referred to as low ratings). See the below pie charts, showing rating compositions.

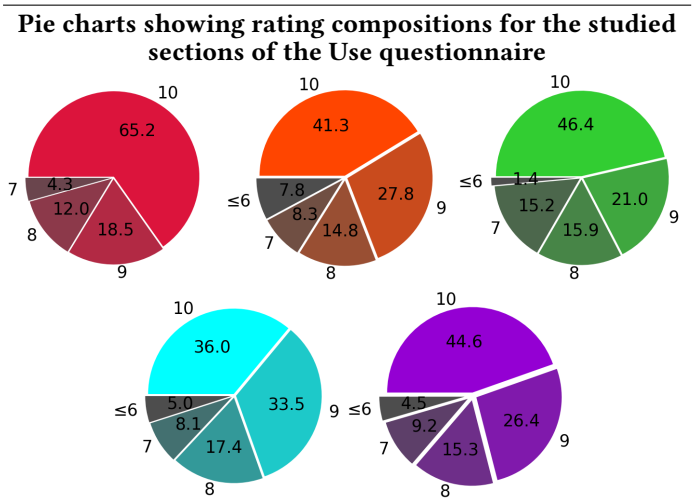

Figure 9: sectional rating composition. The EoL, EoU, S, U and Total system ratings are crimson, orange, green, aqua and violet respectively, matching that of figure 6.

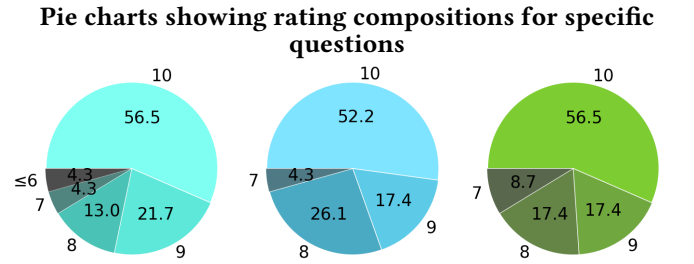

Figure 10: ratings compositions for three specific questions. Those being Vis, Gen and Fun (from left to right).

Looking at figure 9, we see again that ratings are in ascending order.

6.4.1 Fantastic results

The highest percentage frequency of 10/10s is for EoL (crimson), with a percentage frequency of 65.2%. Thus the majority of users found the system to be extremely easy to learn. In fact, 95.7% of ratings provided for all questions in EoL are 8/10 or above, with only 4.3% of ratings being 7/10 and 0% of ratings less than that. Thus 100% of users agreed the system was easy to learn, none indicating otherwise.

Looking at the EoU chart (orange), 83.9% of ratings are at or above 8/10. Thus, the vast majority of users found the system to be easy to use. Additionally, 41.3% of ratings are 10/10, showing that many users found the system to be extremely easy to use.

In the S chart (green), 46.4% of ratings were 10/10, indicating many users were fully satisfied by the system. In fact, there were more 10/10 ratings than 8/10 and 9/10 ratings combined. Additionally, 83.4% of ratings were at or above 8/10, showing that overall, the vast majority of users were satisfied by the system. Only 1.4% of ratings were low, for a total of 98.6% of ratings being at least a 7/10.

Looking at the U chart (aqua), the percentage frequency of 10/10s and 9/10s are almost equally high (36% and 33.5% respectively). While this chart shows the highest frequency of 9/10s (as well as the lowest frequency of 10/10s) of any chart, the percentage frequency of scores 8/10 is 86.9%. Thus, the vast majority of users found the overall system to be useful.

Now looking at the total system ratings chart (violet), we see that the largest portion (44.6%) of ratings was 10/10. The second largest portion of ratings was 9/10 (26.4%), and the third largest portion was 8/10 (15.3%). Only 9.2% of ratings were a 7/10, and a minuscule 4.5% of ratings were low (≤ 6). The vast majority of ratings provided were very high (70% of ratings were 9/10 or above). Thus, the overall system was very highly evaluated. We note that the violet and green charts (total system scores and satisfaction scores) look very similar, with the only notable exception being that the S chart has the highest number of 7/10 ratings of any section.

Now, looking at figure 10, we see that these three questions were very highly evaluated, with each having a median score of 10/10. We see that 91% of users found the visualisation to be at least an 8/10, 95.7% found the book genres to at least an 8/10, and 91.3% found the system's Fun to be at least an 8/10. Thus, with specific mention to our research questions, we see that the visualisation was useful, the book genres were useful and the system had a great user experience.

6.4.2 Low ratings

There were very few low ratings provided. The highest number of low ratings was for ease of use (EoU), with 7.8% of ratings being ≤ 6. In the qualitative analysis, we mentioned that many users missed the easy to use filtering system capabilities because they were not obvious enough at first glance. We see this effect showing in the pie chart (orange).

6.5 Relative results

We now compare the full numeric results of the use questionnaire (i.e. all 27 questions), with each question being measured relative to each other.

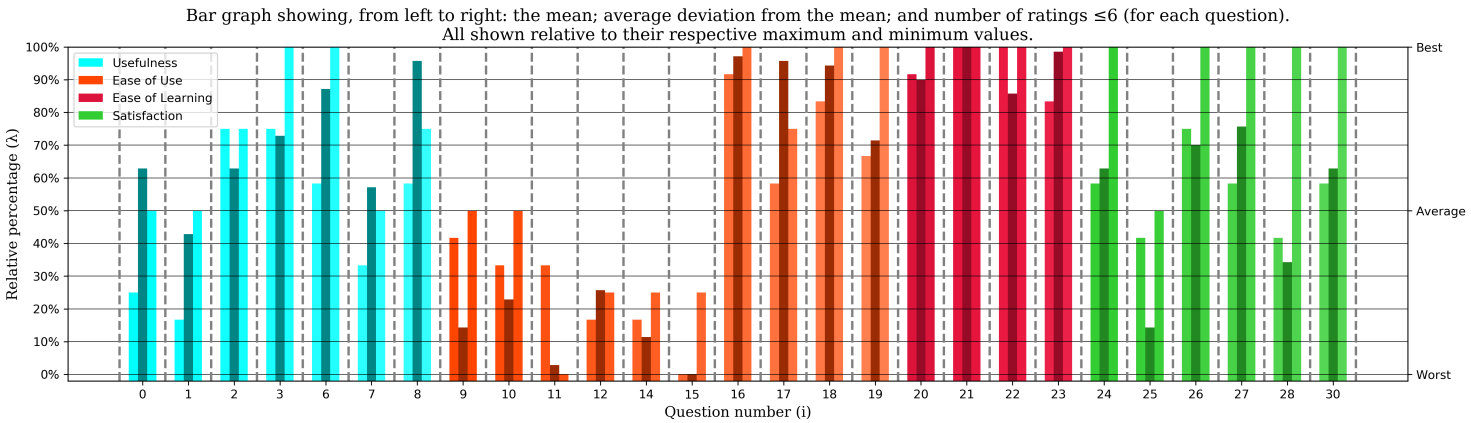

Figure 11: best and worst features of the system

To generate the above graph: for each question i of the Use questionnaire, we calculated the mean-rating $\overline{x}$, the mean-average deviation σ, and the number of low ratings ω. We then performed the following calculations to achieve the bar heights (λ).

$$\lambda_{i,j} = \begin{cases} \dfrac{\overline{x}_i - \overline{x}_{min}}{\Delta\overline{x}} & j == 0 \\ 1 - \dfrac{\sigma_i - \sigma_{min}}{\Delta\sigma} & j == 1 \\ 1 - \dfrac{\omega_i - \omega_{min}}{\Delta\omega} & j == 2 \end{cases}$$

Where $\overline{x}\epsilon[8.3, 9.5]$, $\sigma\epsilon[0.67, 1.37]$, $\omega\epsilon[0, 4]$ and $\Delta \equiv$ range operator.

Now we have a graph with bars measuring features from best and worst. Bars represent $\overline{x}$, σ and ω in order. The tallest $\overline{x}$-bar represents the highest rated question. The tallest σ-bar represents the most consistently-scored (least spread) question. The tallest ω-bar represents the question with the least number of problems (low ratings). So in all cases, tallest is best and shortest is the worst.

The most consistently answered questions were q8 (Usefulness); q16, q17 and q18 (EoU); q21 and q23 (EoL).

6.5.1 Best features

We see that the questions with the highest $\overline{x}$ bars are questions 21 and 22. This shows that, on average, the system's best feature was that it was easy to learn and remember.
The questions in the top 10% were q16 and q20. Thus, the system was very consistent (no bugs) and quick to learn.
The top 25% contains both q2 and q3 (our designed questions). This indicates that both the visualisation and the book genres are among the most useful and best features of the system. Also in the top 25% of means was q26, showing that the system's level of fun was one of its best features.

6.5.2 Areas of improvement

The lowest ω bar for usefulness was q7, which means that, the system did not meet the needs of a few users. We see that the section with the lowest ω-bars is EoU. The question with the lowest ω-bar is q11. As such, the user friendliness is a specific area that could use improvement. We see that the question with the lowest $\overline{x}$ bar is q15, which means, relatively, users felt they might struggle to use the system without written instructions.

7 Conclusions

Overall, the system was extremely highly evaluated, with a mean total-system score of 9/10 excluding outliers (8.2/10 including outliers). Users showed a very high level of satisfaction and fun. Users overwhelmingly found the system to be easy to learn and useful. The majority of users found the system to be easy to use, with only a select few finding otherwise.

7.1 Research questions

7.1.1 The book genres were useful

The system had a mean genre usefulness of 8.7/10, which was the highest of all columns in figure 15. It had the lowest standard deviation too, showing conclusive evidence that even amongst outliers, the book genres were useful.

After outlier exclusion:
All users (23 out of 23) found the book genres to be useful, with 22 (96%) of the users finding it to be very useful. Users mentioned how the grouping of ETDs into genres, and how they were distinguished/displayed visually, massively helped them to find information in specific areas of academia. Overall, users consistently found the usefulness of the book genres to be very high, with median and mean ratings of 10/10 and 9.2/10 respectively. We conclude that the first research question has been answered with a resounding yes.

7.1.2 The visualisation was useful

The visualisation had a mean usefulness of 8.3/10, which is another excellent result for the experiment.

After outlier exclusion:
All except one (96%) of the users found the visualisation, specifically, to be be useful. 21 (91%) of the participants found the visualisation to be very useful. There was an overwhelming amount of long answer feedback provided to support this (see users coded with VISUAL and HLPFUL in figure 4). Many users made very clear that they loved the system and that the visualisation massively helped them find relevant ETDs. Thus we conclude our second research question to be a resounding yes too.

87% of ratings for the overall system's usefulness (U section of questionnaire) were 8/10 or above, showing that overall, users found the system to be very useful. With that said, the book genres and the visualisation stood out as the most useful features of the system (shown by the blue mean cells in figure 8 and the highest $\overline{x}$-bars in figure 11). So again, we have conclusive evidence that the first two research questions are a resounding yes.

7.1.3 The skeuomorphism provided an enjoyable experience

The system's mean fun rating was 8.5/10, suggesting a great user experience, even amongst the outliers. The mean satisfaction score was also high (8/10).

After outlier exclusion:
All 23 users found the system to be fun to use. An overwhelming 91% of users found the system to be very fun to use. Many users explicitly mentioned their enjoyment and fun they experienced using features of the book skeuomorphism, like the interactive books and their animations/visual effects. Many users really enjoyed the usages of colours and icons attached to the books. Many users found that the system allowed for a more enjoyable experience performing research/finding ETDs. Out of the 20 users that provided long answer feedback, 15 mentioned the visualisation positively and made explicit mention of a high level of fun/enjoyment.

There was only one user who did not enjoy the visual effects. This user, however,

still rated the system's fun to be a 10/10. So overall, the system was extremely fun to use and had a great user experience.

7.2 Problem areas

The worst feature of the system was that it needed written instructions to be used. The reason behind this has been identified as the filtering system's functionality not being obvious at first glance.

A few users found the system to not be user friendly. The varying degrees to which visualisations aid different learning types, has been identified as the primary reason for this. Some users did not appreciate the falling animations and the wide variety of colours. The majority, however, did find the system to be user friendly (see figure 11 showing a much higher $\overline{x}$-bar than the other two bars for q11), suggesting that user friendliness was tied to some specific features that a few users had grievances with, not the overall system.

7.3 Future work

The filtering system needs to be improved in many ways if it is to be implemented as a publicly accessible ETD website. Users have suggested many improvements such as allowing for saving of filter configurations, being able to search for specific filters and being able to see more specific sub-genre filters (for example Computer Vision). The system has many features designed to improve perceived filter-functionality (see system design section), the dual-toggling capabilities of the filtering system were still missed by a few users. In the future, instead of tooltips, hovering over a filter or filter icon should show a "glimpse" into the future, which would show what happens when said filter is clicked. This could be done by either making filters that are about to be turned on or off "jiggle" slightly (like in IOS app deletion), or making them less/more transparent. Perhaps a better solution would be giving users the option to drag and drop the skeuomorphic book-looking filters between two different filter columns (one being active filters and one being inactive filters). This could provide a very satisfyingly-interactive and easy to use filtering system.

The ability to toggle on/off the usage of specific features like: colour, icons, skeuomorphism and/or visual effects — may help accommodate users with specific grievances. Many users only had an issue with one feature, while loving the other 3. All had specific favourites amongst the 4, so blanketly removing any of these features would not be wise. However, giving users the option to turn a specific feature "off" would solve the problem nicely. This could be done relatively easily in an options menu, making the visualisation more accessible and user friendly to all users.
Another feature that should be included in this settings menu, should be a slider to modify the pile size (number of search results on the screen). This has been identified as a feature that would aid with user friendliness.

Since the system is an experimental prototype, it was only designed to incorporate 52 different genres. With a controlled vocabulary and sizeable enough dictionary, it would be possible to allow for even more genres. This could be done successfully by using a pseudo-random algorithm to make unique colour-scheme variations of the parent genre. This would massively help with sub-genre differentiation (as suggested by U38), vastly improving its real world (non-experimental) usefulness.

We were unable to survey any art students. This is unfortunate since the benefits of a visualisation would show more significantly in that specific field (high percentage of visual based learners).
We were unable to survey enough participants to achieve any meaningful correlations between faculties/degrees of study and user ratings. In the future, a higher sample size which includes art students should be surveyed and analysed.

7.4 Final words

Overall, this paper has shown that incorporating visualisations into the search results of digital libraries has a large amount of merit. To implement this successfully as a non-experimental (publicly accessible) system, an accurate (and controlled) subject classification needs to be used. Users really enjoyed how easy it was to visually distinguish between different genres, but without controlled vocabulary, differentiation between genres and sub genres is impossible.

References

[1] Bush, Vannevar. "As We May Think." The Atlantic Monthly, July 1945. https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881, preserved at https://perma.cc/RG3J-GSWG.
[2] Suleman, Hussein Fox, Edward. (2008). The Open Archives Initiative. Journal of Library Administration. 35. 125-145. 10.1300/J111v35n01_08.
[3] Sweetnam M.S. et al. (2012) User Needs for Enhanced Engagement with Cultural Heritage Collections. In: Zaphiris P., Buchanan G., Rasmussen E., Loizides F. (eds) Theory and Practice of Digital Libraries. TPDL 2012. Lecture Notes in Computer Science, vol 7489. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-33290-6_8
[4] Richardson, W.R., Srinivasan, V. Fox, E.A. Knowledge discovery in digital libraries of electronic theses and dissertations: an NDLTD case study. Int J Digit Libr 9, 163–171 (2008). https://doi.org/10.1007/s00799-008-0046-9
[5] Yin Leng Theng, Norliza Mohd-Nasir, and Harold Thimbleby. 2000. Purpose and usability of digital libraries. In Proceedings of the fifth ACM conference on Digital libraries (DL '00). Association for Computing Machinery, New York, NY, USA, 238–239. DOI:https://doi.org/10.1145/336597.336674
[6] Castelli D., Pagano P. (2002) OpenDLib: A Digital Library Service System. In: Agosti M., Thanos C. (eds) Research and Advanced Technology for Digital Libraries. ECDL 2002. Lecture Notes in Computer Science, vol 2458. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45747-X_22
[7] Fox, E.A., McMillan, G., Suleman, H., Gonçalves, M., Luo, M.,(2004). Networked Digital Library of Theses and Dissertations(NDLTD). In J. Andrews D. Law (Eds.), Digital Libraries: Policy, Planning and Practice (pp.M167T187). Hants, England: Ashgate.
[8] Buchanan G., Cunningham S.J., Blandford A., Rimmer J., Warwick C. (2005) Information Seeking by Humanities Scholars. In: Rauber A., Christodoulakis S., Tjoa A.M. (eds) Research and Advanced Technology for Digital Libraries. ECDL 2005. Lecture Notes in Computer Science, vol 3652. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11551362_20
[9] Suleman, Hussein and Fox, Edward A. and Madalli, Devika (2003) Design and Implementation of Networked Digital Libraries: Best Practices, Proceedings of DRTC Workshop on Digital Libraries: Theory and Practice, March 2003, Bangalore, India.
[10] Fox, Edward A. and Suleman, Hussein and Gaur, Ramesh C. and Madalli, Devika P. (2005) Design Architecture: An Introduction and Overview, Design and Usability of Digital Libraries: Case Studies in the Asia Pacific, 22-37, Information Science Publishing.
[11] Kani-Zabihi, E., Ghinea, G. and Chen, S.Y. (2006), "Digital libraries: what do users want?", Online Information Review, Vol. 30 No. 4, pp. 395-412. https://doi.org/10.1108/14684520610686292
[12] Webley, Lawrence and Chipeperekwa, Tatenda and Suleman, Hussein (2011) Creating a National Electronic Thesis and Dissertation Portal in South Africa, Proceedings of 14th International Symposium on Electronic Theses and Dissertations, 13-17 September 2011, Cape Town, South Africa.
[13] Maristella Agosti, Franco Crivellari, Giorgio Maria Di Nunzio, and Silvia Gabrielli. 2010. Understanding user requirements and preferences for a digital library Web portal. Int. J. Digit. Libr. 11, 4 (December 2010), 225–238. DOI:https://doi.org/10.1007/s00799-011-0075-7
[14] Agosti, Maristella Orio, Nicola. (2012). User Requirements for Effective Access to Digital Archives of Manuscripts. Journal of Multimedia. 7. 10.4304/jmm.7.2.217-222.
[15] Yi Chun, Chu Bainbridge, David Jones, Matt Witten, Ian. (2004). Realistic books: A bizarre homage to an obsolete medium?. 10.1145/996350.996372.
[16] https://scholar.google.com/intl/en/scholar/citations.html
[17] https://en.wikipedia.org/wiki/Force-directed_graph_drawingDisadvantages
[18] https://en.wikipedia.org/wiki/List_of_academic_fields
[19] https://en.wikipedia.org/wiki/List_of_Dewey_Decimal_classes
[20] https://colorhunt.co/palettes/popular
[21] https://www.flaticon.com/
[22] https://garyperlman.com/quest/quest.cgi?form=USE
[23] https://law.marquette.edu/facultyblog/2020/04/national-poetry-month-create-book-spine-poetry/
[24] https://core.ac.uk/download/pdf/206095228.pdf
[25] https://stats.stackexchange.com/questions/81986/mean-absolute-deviation-vs-standard-deviation

8 Appendix

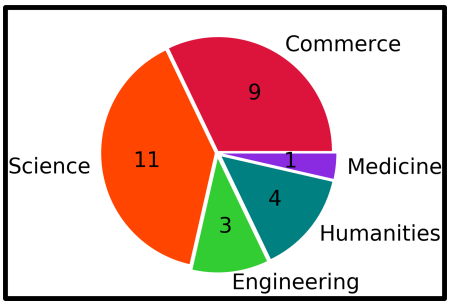

Figure 12: faculty composition

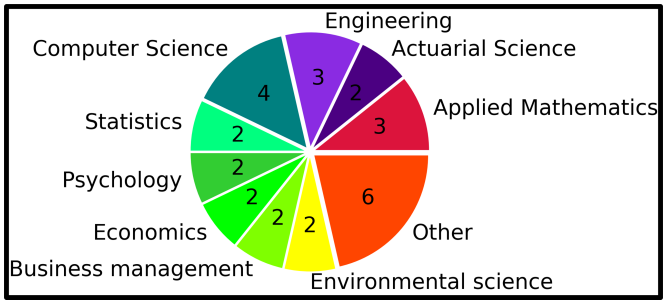

Figure 13: discipline composition

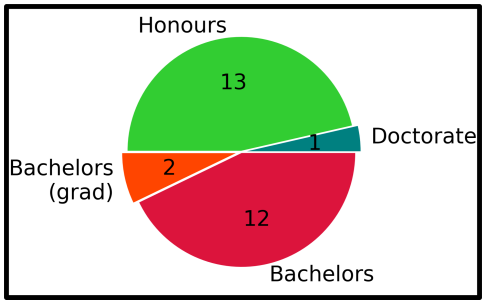

Figure 14: degree of study composition

The other (orange in figure 13) disciplines are: Information security, History, Medicine, Biology, English Literature and Finance.

Combined table showing numeric ratings alongside open-codes for each participant

ID	U	EoU	EoL	S	Tot	Vis	Gen	Fun	Positive Codes	Negative Codes
3	8	8.2	8.5	7.2	8	8	8	7	VISUAL, VISEFF, ENJOYM	FILUSE, FILFET
4	5.9	5.5	4.8	4.8	5.2	6	6	5	FILTER, RELVNT, HLPFUL, SUBJCT	LOADIN, USRFRN
6	10	9.9	9.8	10	9.9	10	10	10	FILTER, RELVNT, HLPFUL, SUBJCT, ENJOYM	FEWRES
7	6.9	8.5	8	7.2	7.6	8	8	8	FILTER, RELVNT, HLPFUL, SUBJCT	TOLTIP
8	5	4.5	7	4.2	5.2	3	7	5	FILTER	USRFRN
9	3.1	1.4	2	1.8	2.1	4	7	5	VISUAL, VISEFF, INTERA	INTERF
11	8.3	8.1	9	8	8.3	8	8	8		
12	9.7	6.3	10	9.2	8.8	10	10	10	VISUAL, COLCOD, FILTER, RELVNT, EASUSE, EASIER, SUBJCT, COLOUR, HLPFUL	LOADIN, FEWRES, FILUSE, FILFET
13	9.4	9.3	9.8	9.2	9.4	10	10	9	VISUAL, EASUSE, HLPFUL, RELVNT, SUBJCT, ENJOYM	
16	8.9	10	10	9.7	9.6	9	9	10	VISUAL, VISEFF, COLCOD, ENJOYM, SUBJCT, COLOUR, EASIER, EASUSE, HLPFUL	
18	10	10	10	10	10	10	10	10	VISUAL, VISEFF, EASIER, EASUSE, FILTER, RELVNT, SUBJCT, ENJOYM, HLPFUL	
20	8.9	8.3	10	9	9	10	7	10	FILTER, SUBJCT, EASUSE, SAVTIM	FILVIS, FILFET
21	8.9	9.1	9	9.5	9.1	9	8	9		
22	8.7	7.8	10	8.7	8.8	5	10	10	VISUAL, FILTER, RELVNT, ICONSV, SUBJCT, ENJOYM, INTERA	COLOUR, VISEFF
23	9.1	9.5	10	9.2	9.5	9	9	9	VISUAL, FILTER, ENJOYM, INTERA	
24	9.1	10	10	10	9.8	9	10	10	FILTER, SUBJCT, EASIER, HLPFUL, EASUSE, ENJOYM	FEWRES
27	9.3	9.2	9.5	9.5	9.4	10	10	10	VISUAL, SUBJCT, RELVNT, ENJOYM, VISEFF, HLPFUL, COLOUR, COLCOD, ICONSV, EASIER	FILUSE, MATURE, FILFET
28	7.6	8	9	7.5	8	9	9	8	VISUAL, COLOUR	
29	8.3	6.5	7.8	7.5	7.5	10	8	8	VISUAL, COLOUR, COLCOD, EASUSE, EASIER, SUBJCT, RELVNT	FILLIM, INTERF, MATURE, USRFRN, FILFET
30	9.3	8.2	9.3	9.3	9	10	10	10	ICONSV, HLPFUL, SUBJCT, ENJOYM, COLOUR, COLCOD, VISUAL	MATURE
31	10	10	10	10	10	10	10	10	COLOUR, HLPFUL, VISUAL, ENJOYM, EASUSE	
32	7.6	8	7.8	7.8	7.8	7	8	7		
35	9.6	9.6	10	9.5	9.7	10	9	9	ENJOYM	
37	9.1	9.4	10	10	9.6	10	10	10	VISEFF, VISUAL, ENJOYM, COLOUR, SUBJCT, ICONSV, FILTER, TOLTIP	FILREC, FILFET
38	8.1	10	10	8	9	10	10	10	VISUAL, HLPFUL, SUBJCT, SAVTIM, RELVNT, ENJOYM, MATURE	FILLIM, FILFET
39	9.3	9	10	10	9.6	10	10	10	ENJOYM, VISUAL, SUBJCT, FILTER, COLOUR, ICONSV, COLCOD	
41	2.7	3.9	7	1.7	3.8	1	9	1	FILTER, SUBJCT, HLPFUL, VISUAL, ICONSV, COLOUR, COLCOD	VISEFF, INTERF, USRFRN
42	4.3	5.6	7	5.7	5.6	6	4	9	VISUAL, FILTER, COLOUR, COLCOD, SUBJCT, HLPFUL, RELVNT	INTERF, MATURE
Men	8	8	8.8	8	8.2	8.3	8.7	8.5		
Med	8.9	8.4	9.6	9.1	9	9	9	9		
MAD	1.4	1.5	1.2	1.6	1.4	1.7	1.1	1.5		
IQR	1.5	1.6	1.8	2	1.7	2	2	2		
StD	2	2.2	1.9	2.4	2	2.5	1.5	2.2		

Figure 15: showing average user ratings, mean ratings, median ratings, mean absolute deviation, the inter-quartile range and standard deviation