DERI – DIGITAL ENTERPRISE RESEARCH INSTITUTE

TECHNICAL REPORT: VISUAL ABSTRACTION AND ORDERING IN FACETED BROWSING OF TEXT COLLECTIONS

VinhTuan Thai   Pierre-Yves Rouille
Siegfried Handschuh

DERI TECHNICAL REPORT 2010-07-19
AUGUST 2010

DERI – DIGITAL ENTERPRISE RESEARCH INSTITUTE

DERI Galway
IDA Business Park
Lower Dangan
Galway, Ireland
http://www.deri.ie/

DERI TECHNICAL REPORT
DERI TECHNICAL REPORT 2010-07-19, AUGUST 2010

TECHNICAL REPORT: VISUAL ABSTRACTION AND ORDERING IN FACETED BROWSING OF TEXT COLLECTIONS

VinhTuan Thai[1] Pierre-Yves Rouille[2] Siegfried Handschuh[3]

Abstract. Faceted navigation is a proven technique for exploration and discovery of a resource collection. In this paper, we report on a visual support toward the exploration of a collection of documents based on a set of entities of interest to users, in which faceted navigation is employed for the filtering process. Our approach can be used when metadata is not available and unlike other faceted browsing work, it treats documents as content-bearing items. We propose using a multi-dimensional visualization as an alternative to the linear listing of focus items. We describe how visual abstractions based on a combination of structural equivalence and conceptual structure can be used simultaneously to deal with a large number of items, as well as visual ordering based on the importance of facet values to support prioritized, cross-facet comparison of focus items. A user study was conducted and it showed that interfaces using the proposed approach can better support users in exploratory tasks and were also well-liked by the participants of the study, with the hybrid interface combining the multi-dimensional visualization with the linear listing receiving the most favorable ratings.

Keywords: Faceted browsing, text collections, visual exploration.

[1]DERI (Digital Enterprise Research Institute), National University of Ireland, Galway , IDA Business Park, Lower Dangan, Galway, Ireland. E-mail: vinhtuan.thai@deri.org

[2]Ecole Nationale Supérieure de Cognitique E-mail: py.rouille@gmail.com

[3]DERI (Digital Enterprise Research Institute), National University of Ireland, Galway , IDA Business Park, Lower Dangan, Galway, Ireland. E-mail: siegfried.handschuh@deri.org

Acknowledgements: This work has been supported by Science Foundation Ireland under Grant No.SFI/08/CE/I1380 (Líon-2) and in part by the EU under Grant No.FP7-ICT-2007-1-216048 (FAST project).

Preliminary results of this paper appeared in "Visual Abstraction and Ordering in Faceted Browsing of Text Collections" *Proceedings of the International Workshop on Intelligent Visual Interfaces for Text Analysis (IVITA), co-located with ACM IUI 2010, Hong Kong, February 07-10, 2010.*

Contents

1 Faceted Browsing of Text Collections

1.1 Overview

Faceted navigation is a proven technique for exploration and discovery of a resource collection [14]. Apart from being studied in many research works e.g. [41, 16, 7, 24], its usefulness is attested by the popular uptake in many commercial websites. At the core of faceted navigation is a set of flat or hierarchical facets, which are categories characterizing items in a large collection [14]. Each facet has one or more facet values and each item may be associated with a subset of these values [4]. As such, this navigation paradigm requires rich metadata expressing relationships between facet values and resources. In this multi-step resource seeking process, users start with some *initial constraint definition*, inspect the initial result set, and then continue with *orienteering* and *refinement* steps, and finish the process by *closely examining* the result set [13]. Within this process, user selections of facet values result in either conjunctive or disjunctive queries executed on the resource collection, depending on the nature of the facets and/or the design of the applications. Query previews [28] are often used to indicate how many items are associated with each of the facet values, and as such they help users avoid filtering by facet values that would return an empty result set [15], which is also known as the Poka-Yoke principle [34]. The matching items, or focus items [4], are then displayed as results of the filtering process.

We were interested in user experience with existing faceted user interfaces, therefore we invited nine persons who were familiar with faceted browsing to participate in contextual interviews. They were asked to demonstrate their recent use of a faceted user interface to achieve a real task of their own. They explained their interactions while we were observing. Afterwards, each of them discussed their experience with using these websites. Their feedback can be summarized as follows:

- The facets in these websites were highly appreciated as they were relevant and helpful to narrow down the search space. Some subjects were not happy with the default set of facets, as they did not reflect their most relevant criteria.

- Comparison of items across different facets was very important in many cases for making decisions. To avoid missing the best matches, users had to look at different combinations of facet values inherent in the focus items. Sorting by one facet at a time was therefore not considered effective.

- Facets were not equally important. Some facets were more important than others.

- Having to go through a long list of focus items was time-consuming. This is in line with results from a study on faceted user interface [22], which showed that while users spent equally much time on the query, the facets and the results on the first result page, they focused entirely on the items on the second and third result pages. This suggested that after choosing the facet values to narrow down their options, users still needed to spend a considerable amount of time to look for specific items.

It is worth noting that most of the websites used by the participants only allowed for the single-valued selection mode i.e. only one facet value could be selected at a time e.g. a brand or a price range (hence conjunctive queries used at each filtering step), only some, such as yelp.com and amazon.co.uk, also allowed for multi-valued selection mode e.g. on the "Categories" facet in yelp.com (and hence disjunctive queries used). In the former case, users needed to narrow down the search space in a stepwise manner by choosing only one value in a facet at a time. As such, if none of the items in the result set met their needs, they had to backtrack and explore different paths in order to compare items using a combination of facets that were of import to them to find the best matching items. While comparisons were important to users in making choices, not all sites using faceted navigation paradigm offered this feature. In the latter case, the multi-valued selection mode catered for vague criteria from users (e.g. multiple neighborhoods were considered acceptable for the location facet while a user was looking for a restaurant), and thus allowed for displaying a larger set of items matching one or more values of a facet. In this situation, users needed to look into the details of each item to figure out which values of a facet an item matched. In both cases, the users' feedback indicated that choosing an item matching their needs involved comparing items based on a combination of facets rather than just one.

1.2 Faceted Navigation on Text Collections

While none of the websites used by the participants dealt with documents, they shared a similar feature with most current faceted user interfaces for browsing a text collection, such as the one studied in [22], in terms of displaying the result set. Filtering interactions resulted in document items returned in a list and users still had to traverse through many result pages to select a particular document for further analysis. This similarity exists because even though documents are content-bearing items, their contents are completely ignored and only the (usually hand-crafted) metadata is utilized as with other kinds of resources. This has two setbacks:

- It is not often the case that metadata is available for a text collection to be explored, since the effort required to generate the clean metadata (e.g. to classify if a document belongs to one or more categories) is tremendous on large datasets. An example of this issue is highlighted in [29].
- Even when metadata is available, as in the case of carefully curated digital library collections, the existing paradigm, which is based only on metadata, however, falls short of being adequate to support users in the task of filtering for potentially relevant documents while exploring a text collection. Document contents are only taken into account in free text search to further narrow down the result set. This issue has also been highlighted in [3] after the authors conducted an experimental study on a visualization to support search interfaces and realized that "*users of digital libraries are most often interested in the content of the documents rather than their metadata*". The finding is intriguing and it reinforces the fact that while metadata is certainly useful for filtering and aggregation in general to obtain some basic analytical statistics and relationships, when documents are used, their contents are far richer in terms of useful information that can be used for the exploration activity.

Therefore, we believe that it would be useful to further support users, who are interested in a certain set of entities of interest, to explore a text collection whose metadata is not available. In this case, it is important to:

- *Show how relevant a document is to these entities*. This is due to the fact that unlike the binary relationships often seen between an item and a facet value (e.g. an electronic item and a brand), documents are content-bearing items and hence the relevance relationship is of importance to users in choosing which documents to focus on.
- *Use disjunctive queries instead of conjunctive queries*. When focus items are documents, users should be able to consider multiple facet values simultaneously since a document is usually relevant to a number of concepts at the same time. In addition, while exploring documents, the fact that some documents do not mention certain entities may be of interest to users (e.g. in the Enron email dataset, it may be interesting to identify the set of documents mentioning both "*leak*" and "*investigation*", and

the set of documents only mentioning "*leak*" and nothing else in the legal actitivy category). This feature is lacking in most faceted user interfaces since only a single facet value can be used at a time and thus facet values can only be ANDed together. Disjunctive queries can return a set of items that do not contain certain facet values and this could be of interest to users. This applies not only to documents but also to other digital resources when *slice-and-dice* operations are required [34], e.g. filtering for photos that satisfy the condition "*Australia and in 2004 and not portrait*" from a photo collection.

- *Show basic repository statistics*. This is to support comparisons between different groups of documents meeting certain criteria. In the example above, it may be useful to compare the sizes of those two sets of emails. The use of disjunctive queries enables comparisons directly on the user interface and also reduces the number of different steps that need to be carried out to achieve this goal.

The above requirements, together with what we can learn from the user feedback on existing faceted user interfaces, beg the question if properties of documents can be visually displayed in such a way that users can make a better informed decision faster than having to traverse all pages of results and looking at one document after another. The key limitation of existing work is the lack of an aggregated view showing how each document relates to each of the entities. While certain issues were raised about faceted user interfaces in general [14], they tend to focus on the display of a large number of facets, e.g. which facets to show (adaptively) when there are many. The issues identified here regarding focus items representation for faceted browsing of a text collection have largely been left untouched. We argue that user experience can be improved if the following design desiderata are met:

- Each focus item should have a compact representation expressing its correspondences to facet values.
- Users should be able to perform cross-comparisons of focus items over different values of one or more facets.
- The display can be visually abstracted to deal with a large amount of focus items.
- Users should be able to interactively reorder facets based on their preferences, resulting in different displays of focus items.

As such, we propose using a multi-dimensional visualization, one dimension for each facet value, as an alternative to the linear result listing paradigm. In the rest of the paper, we describe the proposed solution, the user study findings and discuss on future work.

2 Proposed Approach

2.1 Context

Our proposed approach toward faceted browsing of a text collection is developed within the context of a filtering mechanism to support exploration of a document collection in a personalized manner. The design of the initial prototype IVEA [38] is, to a certain extent, based on Shneiderman's visual information-seeking mantra "*Overview first, zoom and filter, then details-on-demand*" [31]. It employs multiple coordinated views to guide users through this workflow. The core element in IVEA is a simple user-defined ontology encapsulating their sphere of interest. Various statistics about the relationships between documents and

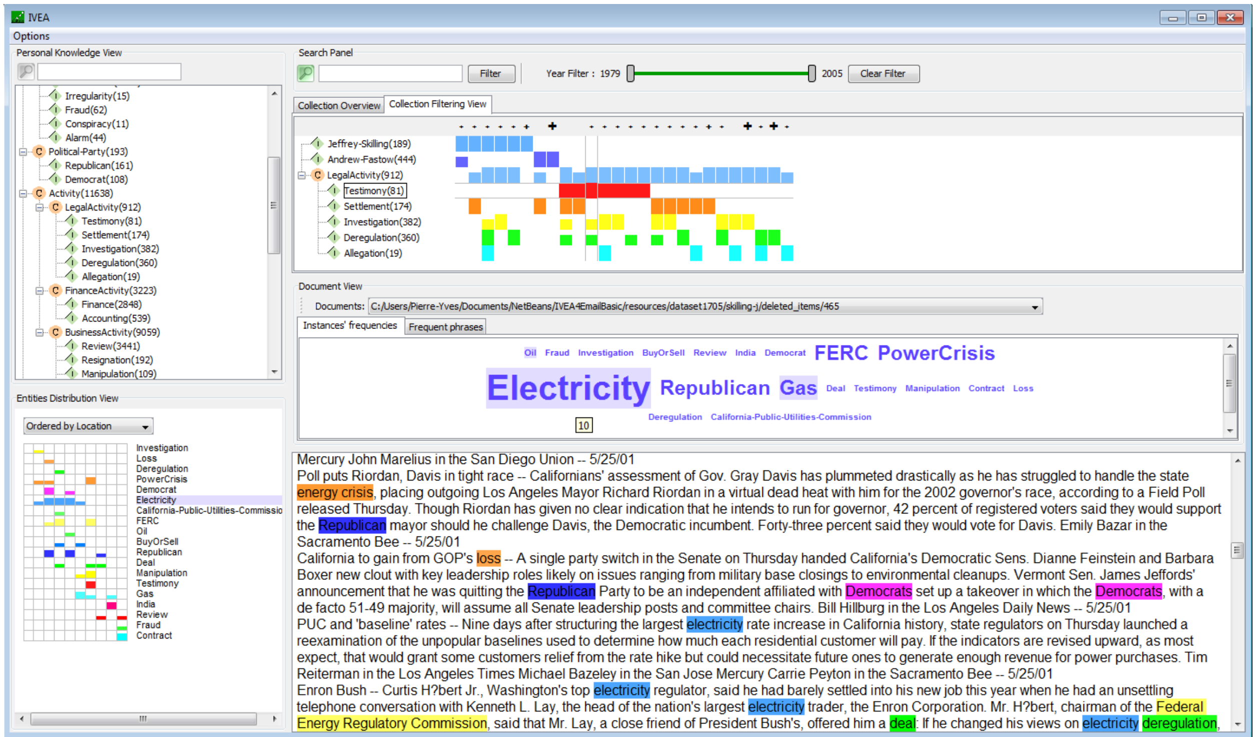

Figure 1: Current version of the IVEA prototype

entities in the ontology are visually presented to facilitate the exploration, as shown in Fig.1. Brushing technique is used to highlight relevant information across the coordinated views. A screencast[1] is available at http://resources.smile.deri.ie/projects/ivea/screencast/ .

2.2 Initial Design

To show the correspondence between documents and entities of interest to users, we proposed in [39] the initial idea on a matrix-based multi-dimensional visualization, which was inspired by FOCUS [33] and TableLens [30]. This representation, in effect, blends visualization with faceted browsing in that users can select entities from hierarchical facets and then documents relevant to one or more of the selected entities are displayed on the visualization. The main differences to other faceted browsing work are that our approach focuses on enabling users to explore a document collection which does not have associated metadata and disjunctive queries are used given the characteristics of documents as content-bearing items. The relationships between documents and entities of interest to users are automatically obtained based on analysis of the documents' contents instead of metadata.

Since hierarchical relationships between entities are taken into account, selecting a class will result in the automatic inclusion of all of its direct instances and recursively, all of its subclasses. Thus, facet selection for filtering can be done at different levels of granularity and multiple facet values can be selected in a single operation. In the matrix, rows represent selected entities, columns represent documents containing at least one of those entities, and each cell shows the relevance value (TF-IDF based score) of a document with

[1] The ontology entity search feature (in the upper-left corner of Fig.1) was an additional feature after the screencast was made.

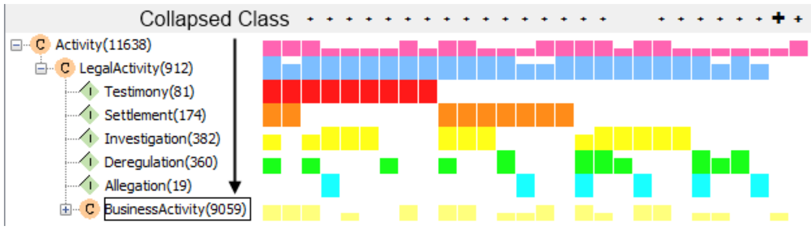

Figure 2: Semantic Zooming

respect to an entity via its height. Here each entity can be linked to a user-defined set of associating terms. As such, abbreviations, linguistic variations (plural/singular forms, verb tenses, etc.), conceptually related terms and synonyms can be taken into account when necessary[2].

To decide a cell's height, we use k-means clustering to identify three clusters of relevance values, and the maximal values of the three clusters are used as thresholds. The vertical part of the cross-hair highlighter helps to focus on which entities a document contains and its horizontal part helps to show the distribution of an entity in a collection. Entity de-selection is made as easy as entity selection, i.e. users can de-select facet values by right-clicking on an entity and the whole respective row is removed from the visualization. Keyword search is also available to further restrict the matrix to display a particular subset of the result set. If the timestamps of documents can be extracted, they can also be used as facet values. Pivoting is also supported when users click on a specific entity that is contained within the document in focus.

Although it meets two of the design desiderata, the visualization provides no visual abstraction to cater for a large number of documents and no interactive ordering of facet values to enable users to easily compare items across different facet values (entities). Next we present the proposed solution. Here, documents are information items and concepts/entities of interest to users are facet values.

3 Visual Abstraction of Documents

Relationships between documents and a set of concepts can be represented by a bipartite graph $G = (D, C, E)$ whose vertices belong to two disjoint sets D representing documents and C representing concepts. If $d_i \in D$ is relevant to $c_j \in C$, then there is an edge $(d_i, c_j) \in E$ connecting them, whose weight is the relevance of d_i with respect to c_j.

Given the typically available screen resolution, too much data is confusing and limiting information manipulation. Therefore, it is important to collapse visual information when desired so that it takes up less screen space and users can focus on what is being shown more effectively. This need is equivalent to interactively collapsing and expanding vertices in the set of documents and the set of concepts so that the view is collapsed and expanded accordingly. In this respect, two mechanisms are provided: semantic zooming and document grouping, which can be employed simultaneously to achieve a highly flexible visual abstraction when the number of documents makes it challenging to digest and explore.

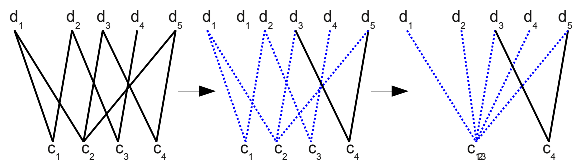

Figure 3: Semantic Zooming Bigraphs

3.1 Semantic Zooming

Hierarchy is an important organizational paradigm, which helps users abstract key concepts from groups of similar items and structure their reasoning [3]. The semantic zooming feature is based on the hierarchical relationships among entities (facet values) and hence helps avoid cluttered interfaces by providing different levels of abstraction. The hierarchy attached with the matrix allows users to dynamically drill down or roll up to achieve views at different conceptual levels of detail, as indicated by the collapsible glyph in Fig.2, to focus on particular subgroups. This, in effect, is the aggregation of vertices in the set of concepts C, hence replaces edges that connect documents to instances of a class with edges that connect documents to that class only. For instance, the left bigraph in Fig.3 shows the relationships between 5 documents and 4 concepts. Assuming that the concepts c_1, c_2, c_3 are instances of a class, they therefore can be grouped into a single concept c_{123} representing that class. Any document that contains any or all of the three concepts c_1, c_2, c_3 is considered as containing the concept c_{123}. Thus, the resulting bigraph on the right has fewer edges, as shown in Fig.3, whereby all documents now connect to the concept c_{123} instead of to the three concepts individually (edges connecting the involved documents and concepts are shown with dashed lines). In general, the levels of semantic zooming correspond to the number of hierarchical relationships between entities defined in the ontology.

3.2 Documents Grouping

While semantic zooming can abstract away a lot of details, the number of documents in a relatively large collection can still be too much to be effectively displayed on a limited screen space. Here the documents grouping feature provides further abstraction based on the notion of Structural Equivalence of individuals in social networks [25], defined as below.

Definition 3.1 *Objects a, b of a category C are structurally equivalent if a relates to every object x of C in exactly the same way as b does [25].*

This notion is used to partition objects in a set into classes of structurally equivalent objects, which leads to the ability to derive a reduced set of categories in which belonging objects are considered equivalent. When these objects are individuals in a social network, the set of derived categories represents “maximal relationally homogeneous groups” [25].

Here, we treat documents and concepts as objects and adapt the Structural Equivalence notion as per below.

[2]This is important, since in many cases (e.g. business report filings), documents’ authors adopt a rhetorical writing style by choosing different wordings, and use them as a semantic camouflage with the intended purpose of influencing readers into accepting misleading interpretations of the information being presented.

Definition 3.2 *Given a set of concepts* $C = \{c_1, .., c_n\}$*, the set of structurally equivalent documents with respect to* C *consists of documents that contain all elements of* C*.*

In other words, documents d_i and d_j in D are structurally equivalent with respect to C if there exist edges (d_i, c_k) and $(d_j, c_k) \in E$, for $k = (1, .., n)$. As such, given a set of selected entities C , we can identify a set of structurally equivalent documents with respect to this set and treat them as a group. This, however, has two limitations:

- The requirement for documents to be structurally equivalent is strict in that they need to contain *all* elements of a given concept set, therefore the number of documents satisfying this requirement will not likely be large.
- Only one group can be identified given a set of entities, which is not significantly helpful in dealing with a large collection.

Thus, our approach is to consider the powerset of C (excluding the empty set) and find groups of structurally equivalent documents with respect to each of those subsets. For example, there are four concepts c_1, c_2, c_3, c_4 , and five documents relevant to them in such a way that is represented by the left bigraph in Fig.4. Here the documents d_1, d_3, d_5 contain c_2, c_4 and not any other concepts. Therefore, these three documents can be put together into a group d_{135} as shown in the right bigraph.

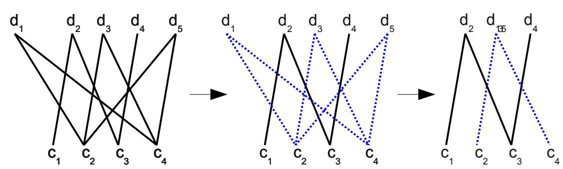

Figure 4: Document Grouping Bigraphs

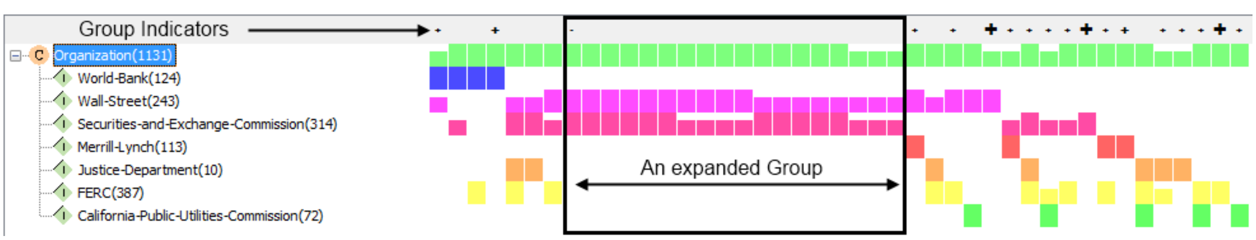

Figure 5: Document Grouping

It is worth noting that while collapsing and expanding facets with semantic zooming are already widely employed in many facet browsing websites, the advantage of our approach is that documents grouping can be used **simultaneously** with semantic zooming to achieve visual abstraction on two different dimensions of concepts and documents at the same time, as can be seen on Fig.2 and Fig.5. As such, this approach enables even more flexibility with regard to the levels of granularity at which information is viewed and manipulated. As in Fig.5, although the screen space is limited, the visualization can still cope with a large set of documents. Here, the filtering process results in 1131 relevant documents. However, since 18 structurally equivalent groups are identified, only 26 columns need to be shown, as only one (randomly chosen) document of a group is initially displayed on the matrix, while other documents that do not belong to any groups

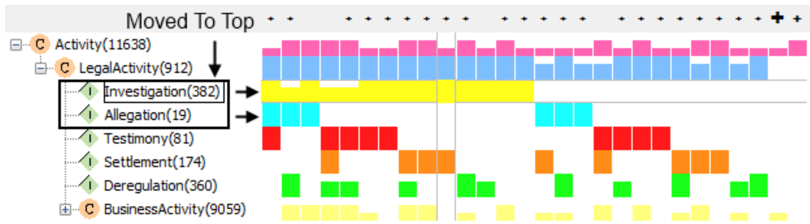

Figure 6: Facets Reordering

are still displayed as a regular column each. In fact, if there are n selected facet values, only a maximum of $2^n - 1$ columns are needed for the initial display. The column of the representative document of a group has a '+' sign on top, which is a visual cue to indicate that there are more documents containing exactly the same set of entities. We also use k-means clustering on different group sizes to find three clusters of sizes and use the maximal values of the three clusters as thresholds. Thus, the size of the '+' sign can indicate the relative size of a group. Hovering over a '+' sign will pop-up the exact number of documents in that group. This visual cue overcomes the need to show numeric values, which require varying spaces and can distort the consistent layout of the matrix. Clicking on this representative column will make visible all documents in a group and its visual cue changes to '-' as shown in Fig.5. Clicking again on the representative column will hide other documents in that group. This "focus+context" interaction simplifies the comprehension of the visual display of a large number of documents without users having to examine a matrix containing a large number of columns, since the initial display does not depend on the actual number of focus items (documents).

4 Visual, Interactive Ordering based on Facet Values

As previously mentioned, not all facets are equally important. For each user and/or for each task, there is an order of importance of facet values accordingly. Therefore, we consider facet values that are placed on top to be more important than those below them. Thus, documents and groups of documents are reordered based on their correspondences with facet values. As in Fig.2, in the facet "LegalActivity", the value "Testimony" has the highest position, therefore documents and groups of documents that are relevant to "Testimony" are moved to the left and those that do not are moved to the right. Within these two groups, they are subsequently ordered by their relevances to the second value, "Settlement". This ordering is done similarly until the last facet value. This ordering can be efficiently achieved using a bitmap of a document or a group of documents, which is constructed by assigning a 1 bit if a document/group of documents is relevant to a facet value, a 0 bit otherwise, and the first facet value corresponds to the highest order bit. For instance, the left-most document in Fig.6 corresponds to the value "11111000", highest among the derived bitmap values.

Furthermore, users can interactively reorder facet values while exploring a text collection. As shown in Fig.6, in the facet "LegalActivity", if the values "Investigation" and "Allegation" are considered more important, they can be moved (via drag-and-drop) on top. The documents' order is changed accordingly as a result. We believe that this visual ordering based on facet values enables users to easily compare focus items, in this case documents/groups of documents, across facet values in a meaningful, prioritized way.

5 User study

We conducted a user study to test the null hypothesis that *there is no difference on users' performance and preferences on interfaces using the matrix-based representation and interfaces using the linear listing paradigm.*

5.1 Method

5.1.1 Participants

18 people (9 females, 9 males) participated in the study, two of whom participated in pilot sessions. Six of them were in the age range of [18,25], nine in [26,40], and three in [41,60]. In terms of occupations, the subjects were accountants, administrative assistants, project administrators, managers, programmers, undergraduate and graduate students in History, Business, Physics, Industrial Engineering and Information Technology/Computer Science. All were familiar with web search and had needs to seek for information from documents (with frequencies ranging from a few times a week to hourly). They had all used commercial websites employing the faceted browsing paradigm. None had any prior experience with the IVEA prototype.

5.1.2 Materials

We used a set of 38180 emails from the Enron email dataset (available at http://www.cs.cmu.edu/∼enron/) as the test collection in the study. These emails varied in length, from relatively short to long ones. Many of them were internal disseminations of news articles together with further discussion. A set of pre-defined facets and facet values were used for the study. The study was conducted on a 15" laptop with a 1440x900 screen resolution.

5.1.3 Design

The study used a within-subjects design. The independent variable was the interface type and the dependent variable was the participants' performance and their subjective ratings. In this study, we implemented four different interfaces to display the result set returned from faceted filtering:

- **Simple linear listing (UI1)**: In this baseline interface, as shown in Fig.7, items were displayed linearly, ordered by their relevance to the set of selected entities (facet values). Each result item had an email subject, a set of entities that it contained among those used for filtering and a short text snippet. This interface resembled current faceted browsing websites in terms of showing which facet values an item is associated with for a multi-valued facet.
- **Linear listing with grouping (UI2)**: As shown in Fig.8, this interface was similar to UI1 in that it also employed the linear listing paradigm that users were already familiar with. However, instead of showing the names of the entities contained in each result item, small squared icons were used as their visual representations. Entities' names could be known from the legend or the tooltip texts. Each square was filled with colors based on their relevance values, in the same fashion as the cells of the matrix representation. In addition, items were also ordered and grouped in a similar fashion to our proposed approach in Section 3 and 4 in that items containing exactly the same set of entities were displayed continuously and ordered by their binary signatures. This interface served as a strong baseline.

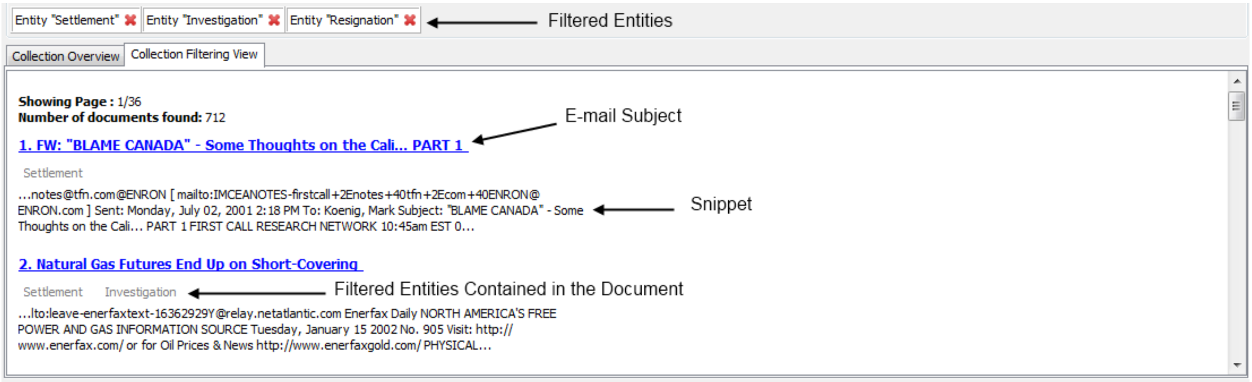

Figure 7: UI1

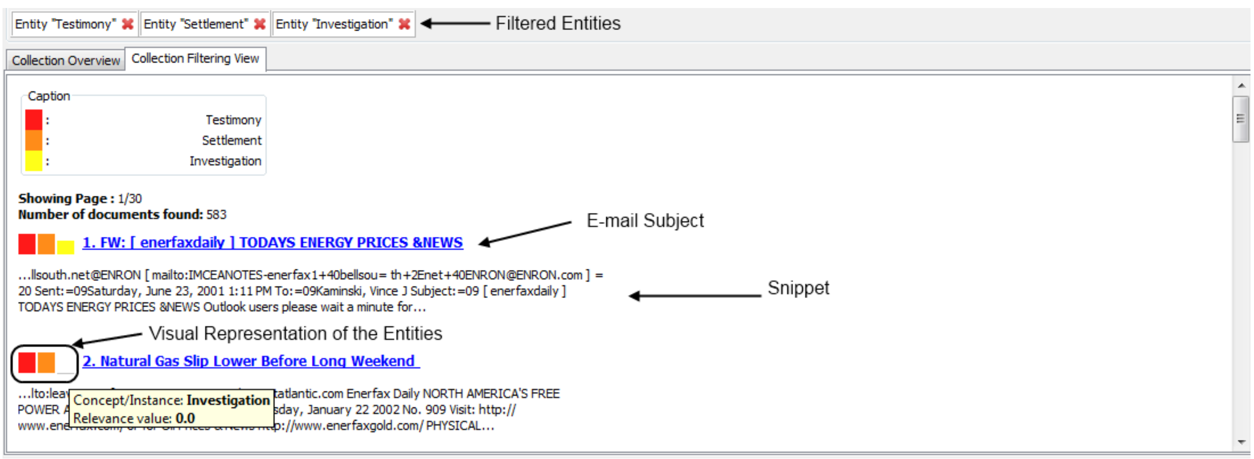

Figure 8: UI2

- **Matrix-based view (UI3)**: This interface employed the proposed approach in Section 3 and 4.
- **Matrix-based view with linear listing (UI4)**: As shown in Fig.9, this hybrid interface was a combination of UI3 and a slightly-modified version of UI2. When users clicked on the representative column of a group of structurally equivalent items, they were not expanded as columns in the matrix. Instead, this group of documents were shown in a linear listing view in which items were ordered by relevance with respect to the set of entities shown in the representative column. The representative column was visually highlighted in the matrix to indicate which combination of entities were contained in that group of items. The visual indicator for each document was made more compact as non-relevant entities were removed, which resulted in a more focused representation while the correspondence between a document and the whole set of selected entities was still visually available in a larger context in the matrix.

The above four interfaces were considered to cover a range of variations from the one that resembled the current linear listing paradigm, to linear listing enhanced with icon representations, which was also grouped and ordered, to the proposed matrix-based representation and a hybrid version of matrix-based and linear listing view. During the study, each implemented interface was referred to by a neutral, color-based name

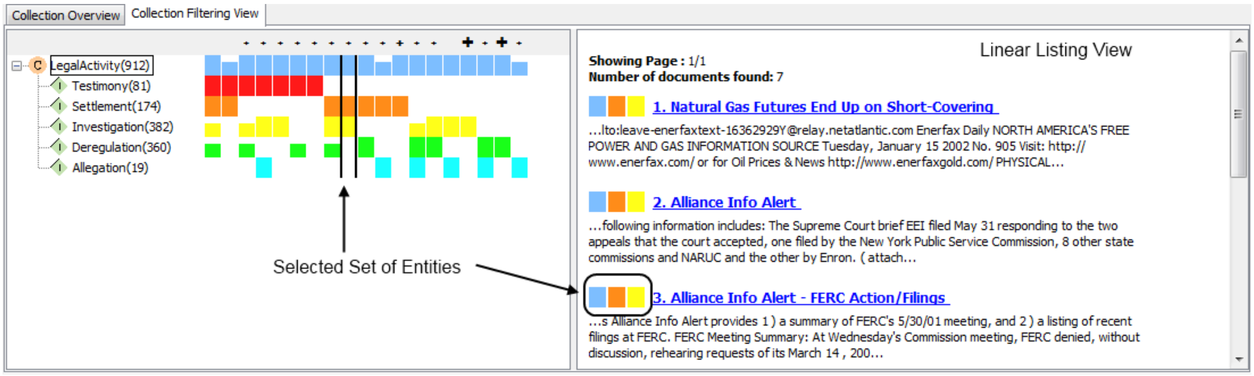

Figure 9: UI4

e.g. “FacetExplorer-Green” for UI4.

5.1.4 Procedure

The subjects were given a short briefing about the study. They were also introduced to and shown video screencasts of all four different interfaces and asked to carry out a number of sample tasks on each of them. They were not informed of which ones employed our proposed approach. Then they were asked to perform four different sets of tasks, one for each interface. To eliminate learning effects, the order of the interfaces was counterbalanced. The assignments of the four different sets of tasks to the four interfaces were also counterbalanced using Latin Square as suggested in [18].

Each set of tasks consisted of the below three parts:

- The first task was related to basic repository statistics, i.e. asking for the number of documents meeting some criteria in two subtasks.
- The second task asked the subjects to filter for (i.e. identify) specific documents meeting some criteria including a time period.
- The last task was an optional task involving the subjects freely filtering for documents of interest.

The first two tasks were preceded with short excerpts from news articles from the New York Times, which were related to the Enron event and hence the dataset, to set the contexts. The contexts were helpful to provide the subjects with some topically interesting background information and hence they could relate to the situation and be motivated to carry out the tasks [21]. In addition, we designed the first two tasks such that they required the same amount of efforts (the same number of steps) to find out the answers i.e. all subtasks involved the same number of different entities, however their associated Boolean operators were varied to avoid learning effects. While the first two tasks enabled us to compare users’ performances with respect to obtaining basic repository statistics and identifying the relevant set of documents meeting certain criteria, the third task was an opportunity for the subjects to freely experiment with various features of the interfaces being evaluated and hence we could solicit their comments. As such, we took into consideration the task completion time and accuracy in the first two tasks and sought qualitative results from the third one. One of the four sets of task is available in the Appendix. The participants did not have to write down answers as doing so may introduce large variations in terms of time. They only pointed to them on the screen and verbally informed the facilitator, the accuracy and task completion time were extracted in post-hoc analysis based on the screen and voice capture .

In order to compare the subjects' performance when the Latin Square design was used, it was important to take the tasks' characteristics into consideration while comparing them. While the number of entities involved were equal across tasks, the sizes of the result sets were not, and that was something we could not control. To answer the structured tasks, the subjects had to process result sets of different sizes and as larger result sets may require more time, we should not attribute that to the interfaces themselves. In addition, the accuracy of the answers should also be considered, since participants who made mistakes should be considered as having lower performance score. Therefore, we employed a metric similar to the *qualified search speed* proposed in [18], called *relevant speed*, which reflected the number of focus items the subjects were able to process per minute, weighted by the accuracy of their answers.

$$\text{Relevant Speed} = \frac{\text{Result Set Size}}{\text{Task Completion Time}} \times \frac{\text{Number Of Correct Answers Found}}{\text{Number Of Total Correct Answers}} \quad (1)$$

Once the subjects finished the tasks on all interfaces, they were asked to give their subjective ratings on each of the interfaces, based upon the following adjectives: "*Easy to use*", "*Easy to browse*", "*Stimulating*", "*Satisfying*", "*Overwhelming*", "*Flexible*", "*Self-descriptive*", "*Organized*","*Tedious*". The ratings were on a Likert scale from 1 (Strongly disagree) to 9 (Strongly agree), with 5 being neutral. Similar to the study conducted in [41], we used a wide range to have a more sensitive testing instrument. In addition, the reason the subjects were asked to give these ratings once they finished all tasks was to give them a chance to interact with all interfaces, and therefore they could calibrate their ratings in an informed manner. Previous work showed that the order of interfaces could have an effect on the subjective ratings (e.g. in [41]) and hence we want to avoid this effect. Upon completion of tasks on all four interfaces, the subjects were asked to answer an overall questionnaire. The findings are presented next.

5.2 Results and Discussion

Before we discuss the findings, it is worth noting a few challenges we faced in setting up the tasks for this experimental study. As faceted navigation is usually used to support exploratory activities, constructing tasks for evaluation is known as a challenging issue [21]. Among the desired characteristics proposed in [21], exploratory tasks should indicate ambiguity in information needs, but also suggest a knowledge acquisition, comparison or discovery task. As such, a combination of structured and unstructured tasks are usually used in experimental studies on faceted browsing (e.g. [41]), with "known item" search tasks also used in some cases (e.g [22]). While it may be better to evaluate a tool supporting exploratory search in a longitudinal study, our evaluation effort here was not on the tool (IVEA) itself as a whole, but specifically on its way of presenting items from the filtering step. Therefore, we designed the filtering/comparison tasks in such a way that they could be carried out on all four interfaces. Although the prototype could provide more features than reflected in the structured tasks, these features were emphasized in the introduction and the unstructured tasks instead. In a controlled user study, the type of pre-formulated query tasks used was considered acceptable for comparing different ways of presenting the result set since they helped to control the variations potentially caused by many elements of an experimental study and to focus on a particular interface component instead of the whole tool [18, 15].

In this discussion, for clarity we refer to the interfaces using their abbreviations e.g. UI1 instead of the neutral names used in the study.

5.2.1 Quantitative results

When the subjects carried out the tasks, six of them gave up with UI1 on all tasks as they found it too tedious to continue. The rest of the participants tried but also gave up on at least one task with UI1. Therefore, we did not have sufficient data on the subjects' performance on UI1 to include it in further analysis.

We analyzed the dependent variable (relevant speed, as defined in Eq.1) using 3x2 (3 interfaces x 2 task types) repeated-measures ANOVA. The results showed that:

- The **main effect** of **interface type** was **significant**, $F(2,30) = 4.55, p < .05, \eta = .23$, hence the null hypothesis was rejected. Pairwise comparisons with Bonferroni adjustment showed that the relevant speed on UI2 ($M = 909, SD = 130$) was significantly less than that on UI3 ($M = 1599, SD = 191$) and UI4 ($M = 1476, SD = 170$), both $p's < .05$. There was no significant difference between the relevant speeds on UI3 and UI4. The reason that the mean relevant speeds seemed to be high was because groupings (in all three interfaces) and ordering (in UI3 and UI4) allowed the subjects to skip many of the irrelevant groups of focus items. Thus, it was not the case that the participants looked into every single focus item to decide if it matched or not, but only at the visual encodings of representative items of the groups. As such, even when there were thousands of documents to be processed, only a small number of them needed to be considered instead. This reflected the advantage of using grouping and ordering.
- The **main effect** of **task type** was also **significant**, $F(1,15) = 17.99, p < .05, \eta = .55$. Pairwise comparisons with Bonferroni adjustment showed that the relevant speed on task 1 was significantly less than that on task 2, $p < .05$. This may be attributed to the fact that in each task set, task 1 was always before task 2 on all interfaces, so the subjects became more familiar with the interfaces as they carried out the tasks and faster on task 2 as a result. In addition, in task 2, there were always two subtasks that used the same set of entities but different Boolean operators attached (for the purpose of comparison), hence the participants did not have to form new queries (and hence needed not process a new result set), they just needed to find relevant answers from the same set of results.
- The **interaction effect** was **not significant**, $p > .05$.

Furthermore, the mean values of the participants' ratings on the degree to which they agreed with the adjectives used to describe the four interfaces are shown in Fig.10. In general, UI3 and UI4 received more positive ratings than UI1 and UI2 in almost all measures. Below we discuss the ratings for each adjective based on ANOVA analysis and post-hoc tests:

- "*Easy to browse*": The effect of interface type was significant, F(3,60) = 18.29, $p < .001$. Tukey HSD post-hoc tests showed that it was significantly less easy to browse on UI1 than on the other three interfaces, $p < .001$. (No other pairwise differences were significant.)
- "*Easy to use*": The effect of interface type was significant, F(3,60) = 5.44, $p < .05$. While the significance value for homogeneity of variances in this case had $p < .05$, the Welch and Brown-Forsythe tests were both significant with $p < .05$, therefore we could reject the null hypothesis that there was no difference in the mean ratings with respect to the "*Easy to use*" aspect. Games-Howell post-hoc tests showed that UI1 was significantly less easy to use than UI3 and UI4, $p < .05$.
- "*Flexible*": The effect of interface type was significant, F(3,60) = 18.12, $p < .001$. Tukey HSD post-hoc tests showed that UI1 was significantly less flexible than the other three interfaces, $p < .001$.

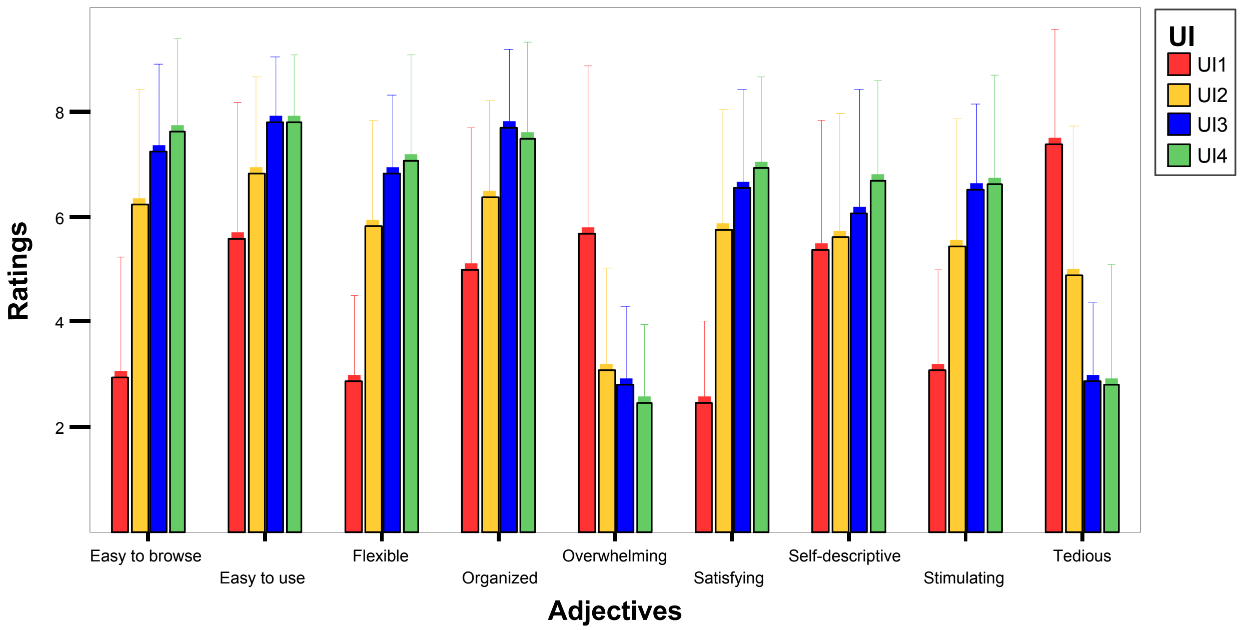

Figure 10: Mean subjective ratings.

- "*Organized*": The effect of interface type was significant, $F(3{,}60) = 6$, $p < .05$. Games-Howell post-hoc tests showed that UI1 was significantly less organized than UI3 and UI4, $p < .05$.
- "*Overwhelming*": The effect of interface type was significant, $F(3{,}60) = 7.66$, $p < .001$. Games-Howell post-hoc tests showed that UI1 was significantly **more** overwhelming than the other three interfaces, $p < .05$.
- "*Satisfying*": The effect of interface type was significant, $F(3{,}60) = 18.87$, $p < .001$. Tukey HSD post-hoc tests showed that UI1 was significantly less satisfying than the other three interfaces, $p < .001$.
- "*Self-descriptive*": The effect of interface type was **not** significant, $p > .05$.
- "*Stimulating*": The effect of interface type was significant, $F(3{,}60) = 10.44$, $p < .001$. Tukey HSD post-hoc tests showed that UI1 was significantly less stimulating to use than the other three interfaces, $p < .05$.
- "*Tedious*": The effect of interface type was significant, $F(3{,}60) = 14.45$, $p < .001$. Games-Howell post-hoc tests showed that UI1 was significantly **more** tedious than the other three interfaces, $p < .05$.

In summary, it was faster and more accurate for the subjects to filter/explore with the interfaces that employed the matrix-based representation (UI3 and UI4) than the linear listing paradigm enhanced with visual encodings and grouping (UI2). The simple baseline, linear listing only interface (UI1) was perceived to be too tedious to be useful for this kind of activity (which was also well reflected in the subjective ratings). Both the interfaces that used the matrix-based representation were well-liked by the subjects, as they were considered to be easy to use, to browse, flexible, organized, satisfying and stimulating. While the strong baseline interface (UI2) received lower ratings on these aspects on average, the difference was not significant in comparison with UI3 and UI4 . The pure baseline (UI1) was also considered to be overwhelming. There was no difference in terms of self-descriptiveness for all interfaces.

5.2.2 Qualitative results

The subjects' answers to the overall questionnaire (listed in the Appendix) on the four different interfaces are described below. The finer-grained, elaborated details from the participants were useful as more could be understood about their preferences.

The participants' answers indicated that they thought the tasks were equally difficult. When asked which interface was the easiest to use, nine subjects chose UI4, one chose UI3, four of them thought UI3 and UI4 were equally so and only two chose UI2. Furthermore, six subjects chose UI4 as helped them learn the most about the collection, two chose UI3, five of them thought UI3 and UI4 equally did, one chose UI2, while one chose UI2/UI4 because of the listing, and one did not know. All participants thought the use of the matrix-based representation changed the way they used to explore a collection, with positive and encouraging comments such as "*I think that the ability to look at the nodes [document-entity relations] is really really good. This is often the way I wanna look at things but it's not easy to do [with existing interfaces].*" The participants' feedback for questions on interfaces' features are summarized in Table 1 via an example for each type of responses, ordered by the non-exclusive **Count** figures indicating how many subjects mentioned similar comments.

Description	**Example**	**Count**
The document grouping feature helped	"It's faster, it's far easier, you can see pretty much all documents, which ones are relevant and which ones are not."	15
The facet reordering feature helped	"If you were looking for something that is in the document and something that is not in the document, you just put the one that you want at the top."	13
Scrolling through lists is tedious	"UI2 type of grouping is not as helpful [as Matrix grouping in UI3 and UI4] because you have to scroll down the list. I would give up after a while."	8
The matrix provides clear Boolean statements	"It made it much easier to answer the type of queries you ask. Mentions X, Y but not Z for instance. And also I think it doesn't require a lot of experiences in logic queries. You just see what's there and what's not there."	5
The matrix speeds up search	"Yes, I think it helped me to fasten my way of finding something, especially with a huge collection of documents, so I wouldn't mind having something like this on Google website."	4

Table 1: Comments on Interfaces' Features

The subjects also commented on their experience when browsing/ filtering freely without any guided tasks in order to gather documents potentially containing relevant pieces of information. Most were really positive in the case they were enjoying exploring the collection, such as "*Easy enough to get the information out to get an overview, especially with UI3 and UI4*" or "*It was fine, but it's hard to start. It allows for a lot of trials and errors, adding and removing concepts, when you try something and then you get rid of it, the* *overall result will be more satisfying.*" on UI4. Another subject was surprised about her findings when trying to find documents mentioning a combination of some entities in particular: "*There were certain categories where I thought there would be overlap but there wasn't*." Some responses were more reserved when some of the commenters did not know much about the collection (the Enron event) e.g. "*The documents were not interesting for me, if it was something that I would need, then I would probably very much enjoy it.*", and did not define the set of entities of interest themselves. As a result, not all participants performed the unstructured tasks on every interface. In addition, we observed that:

- Asking the subjects to carry out *unstructured* tasks in a *controlled* experimental study was troublesome, as the nature of the unstructured tasks were contradictory to that of the settings. The subjects were not in a state of mind that would normally encourage them to freely explore a text collection.
- With our evaluation focusing only on the visual representation to support exploring the result set returned from the filtering step, it was not easy for the subjects to resist using the prototype as a whole.

Furthermore, the subjects' overall rankings of the four interfaces are shown in Table 2. It is clear that UI4 was most liked, followed by UI3. UI1 was disliked by all participants. The rankings were in fact surprising to us, as prior to the user study, we believed that UI3 had the advantage of not requiring users to process more information. But as it turned out, the additional list view in UI4 was perceived as usefully providing more information. Further feedback on their justification for the rankings is summarized in Table 3.

Rank	**UI1**	**UI2**	**UI3**	**UI4**
1		12.5%	6.3%	81.3%
2		6.3%	75%	18.8%
3		81.3%	18.8%	
4	100%			

Table 2: Overall Rankings

Description	**Example**	**Count**
Disliked UI1	"With UI1 it's nearly impossible when you have a lot of results."	16
Preferred Matrix views over List views	"Overall, it's incredibly powerful [...] I've never seen that kind of visualization before, and it's really good."	9
Preferred UI4 over UI3	"Compared to UI3, the additional list view is really helpful. "	9
Preferred UI2 over Matrix	"Because the way it is ordered make the user think more about what she wants."	2
Preferred UI3 over UI4	"You don't need to click on the column to go to the documents."	2
Preferred List views over Matrix views	"It gives you directly the list of items."	1

Table 3: Comments on Interface Comparison

In terms of other usability aspects, the subjects were also asked if they think it would make a difference if the matrix only used one color. The responses were unanimous, e.g. "*Yes, it would be less effective*", "*It would be more difficult without colors*". Finally, certain usability issues were raised by the participants as described in Table 4. Since their frequency was low, we believe that with proper documentation/training, the subjects would be able to overcome these issues once they get more familiar with a new visual representation.

Description	**Example**	**Count**
The matrix's + sign is unclear	"I don't really get why you have bigger or smaller + signs. It makes it more difficult to click if it's small. It's a bit tricky."	3
Unclear that a column is a group	"It's easy to forget that a column represent a group of documents. While interacting with the interface I kinda forgot that that was a group of documents."	1
The matrix gets confusing	"It could get confusing for instance if you have 2 white spaces. Or if the relevance is really less, you might mistake it for a white space."	1

Table 4: Comments on Usability Issues

6 Related Work

Since our work focuses on faceted browsing of text collections, it is related to both faceted browsing in general as well as faceted browsing of text collections. Reviews and thorough analysis of research work and commercial applications employing the faceted navigation paradigm can already be found in two excellent sources, interested readers are referred to [15] and [34] for more details. In this section, we highlight some recent work that have not been covered in the above two publications and other visualizations that are related to ours.

A number of faceted navigation systems support exploratory tasks by visualizing the number of items matching the concepts (used as facet values) such as RAVE (using bar charts) [42], Elastic Lists (using list items' sizes) [35] and VisGets (using tag cloud color intensities) [7] . Meanwhile, in IVEA, we use the query previews [28]. The main difference between the three above systems and our work is the result set presentation. All three systems mentioned above employ the linear listing paradigm on the result set once certain concepts have been selected. In this approach, there are no indicators of which concepts an item match and how relevant it is to each of them.

ContentLandscape [35], ResultMaps [3], FacetLens [24] and DocuBrowse [11] are recent applications that visualize result set distribution using treemaps or treemap-inspired representations. The ContentLandscape application aims at supporting resource analysis by allowing users to see the coverage of the selected resource set (e.g. if all product groups are represented by the result set) as well as to split the result set to up to three dimensions and compare various measures [35]. This application does not target at documents as resources and hence uses conjunctive queries as often seen in other applications. ResultMaps, while not being a faceted browser, is relevant to our work because it uses a treemap to represent hierarchical metadata and employs the brushing technique to link the categories that a result list item matches [3]. It is interesting that the experimental study on ResultMaps provided useful findings: (1) while the mean task time and accuracy were better in the interface with ResultMaps, there was no statistically significant effect by the interface type on both measures and (2) users of digital libraries, who are among those that have the need to explore a document collection, actually are most often interested in the content of the documents rather than their metadata, which could partially explain (1). FacetLens [24] is also a faceted browser which employs a more text-oriented variation of the treemap paradigm to show the result set, augmented by a timeline showing basic statistics. FacetLens shares certain visual features with its predecessor FacetMap [32], for which certain concern about the usage of screen estate was raised in [14]. Besides, FacetLens is also only based on metadata and hence has the same limitation that researchers working on ResultMaps had come to realize when documents were not treated as content-bearing resources even though documents' contents matter to most users. DocuBrowse is a recent addition to faceted browsing research, which targets at usage in an enterprise setting [11]. It makes use of metadata and the document collection file structure (organizational hierarchy) to support search and browsing. In this enterprise context, DocuBrowse also focuses on file types as genres (spreadsheet, presentation slides, etc.) and these genres, among others, can be used as facet values. While DocuBrowse maintains the organization structure in the visualization of results, its approach is different from ours in that it relies on the documents' metadata and organization hierarchy information. The document contents are only made use of for recommendation purposes instead in DocuBrowse.

It is also worth noting Camelis [9] and Microsoft Pivot[3]. Camelis is a faceted browser that allows flexible query formulations via various navigation modes. This feature is advantageous when different Boolean operators can be associated with facet values, e.g. filtering for photos that satisfy the condition "*Australia and not portrait*" from a photo collection. Most faceted browsers do not provide this feature despite it being an important one. However, while Camelis targets at supporting agile browsing of a document collection, it does not treat documents as content-bearing resources and there is also no grouping or specific ordering available for the result set. Microsoft Pivot, on the other hand, is a recent faceted browsing application which allows for grouping and ordering of items. This application, however, only works with item collections that have a set of associated clean metadata.

The use of a matrix-based representation to show the correspondence between concepts and documents was proposed earlier in the TileBars paradigm to show distribution of query terms within full text documents [12]. The TileBars approach, however, uses a separate visual representation for each document and its purpose is different from ours. DocBlocks [2] is a recent work that is also based on TileBars to support explorations of US congressional bills. This work is valuable in that it visualizes the thematic contents of documents based on their structures. However, due to its focus on visualizing the internal structure of each document (which the IVEA prototype also focuses on in a different manner via a reordered variation of the TileBars paradigm [37]), DocBlocks does not provide the grouping and ordering of documents at the collection level.

In addition, there are a number of other interesting approaches to visualize the correspondence between concepts and documents. DocuBurst uses a radial, space-filling layout approach to visualize the frequency of words in a document that match a hierarchical language structure such as WordNet [5]. While this representation is intuitive, visually appealing and helpful for content analysis of a single (usually long) document, it could potentially pose a challenge to users if many documents are to be compared at the same time. AutoFocus is an ontology-based visualization tool which displays search results for documents on the desktop as clusters of populated concepts [10]. While allowing for the exploration of a document collection based on the extracted metadata and search keywords, AutoFocus does not focus on the contents within documents. DocCube [26] is based on OLAP principles and employs 2D and 3D scatter plots to show the number of documents belonging to several categories, which are grouped into dimensions. DocCube,

[3]http://www.getpivot.com/

however, does not allow for relative comparison between documents to see how they are relevant to a set of concepts or categories. It is also limited to displaying at most 3 dimensions at a time. Compus [8] is an application to visualize structured document corpora encoded in XML. In Compus, each document is represented by a vertical bar, which comprises of color-coded elements. While comparisons between documents can be made to a certain extent, it seems to be quite difficult to use since Compus lacks a mechanism to deal with the scalability issue. SPIRE/Themescapes [40] is a tool for document clustering via a document landscape on which the strength and interrelations of different themes are represented. In an analogous manner, VIBE [27] supports users in getting an overview of a document collection relatively to their interests. In VIBE, the positioning of documents on a 2-dimensional display is based on the degrees of overlapping between keywords that users define for their points of interest (POI) and documents' contents. However, no intra-document details are available in either SPIRE/Themescapes or VIBE. Scalability issue is also left untouched in VIBE as there is no information on how it can cater for a relatively large document collection and set of POI.

Since by nature our work is also concerned with showing a collection of items having a set of attributes, it is worth noting other early work on multi-dimensional visualization in the literature. Of importance are fundamental ones such as: parallel coordinates [17], relational table [36, 30], iconic displays [12], circle segments [1], star coordinates [19] and dimensional stacking [23]. A good overview of multi-dimensional visualization approaches can be found in [20]. Finally, FeatureLens [6] is an interesting information visualization tool for exploring text collections. However, it focuses on presenting the relative locations and supports of text patterns mined from text collections. While the FeatureLens's interface is very intuitive, it does not allow for visual exploration at different levels of detail based on a hierarchical structure of concepts.

7 Conclusions and Outlook

In this paper, we have proposed a novel approach to support users in exploring a text collection by blending a multi-dimensional visualization with the faceted navigation paradigm. Our analysis of the application area has resulted in further understanding of users' needs, e.g. the ability to treat documents as content-bearing information resources and the ability to carry out prioritized comparisons between them across facet values. Our proposed approach is based on a matrix representation which is analogous to the already familiar spreadsheet paradigm. Its visual simplicity makes it usable as it can convey the correspondence between a document and a set of concepts. The semantic zooming, document grouping and facet reordering features, which can be used simultaneously, enable the tool to deal with a large amount of documents as well as to provide users with a substantial amount of flexibility in terms of choosing the levels of abstraction and priorities of the chosen facet values. We conducted a user study on a variety of interface types and the results showed that those that used the matrix representation were more capable of supporting users in exploratory tasks and were also well-liked by the participants of the study, with the hybrid interface receiving the most favorable ratings.

In future work, we plan to focus on studying more ways to support users to visually analyze and navigate within lengthy documents. Documents are information-rich resources and while this is valuable, it also inevitably poses significant challenges. TileBars, DocuBurst, DocBlocks are a few examples of effective and visually engaging approaches. It is our interest to further research in this direction.

References

[1] Mihael Ankerst, Daniel A. Keim, and Hans-Peter Kriegel. Circle Segments: A Technique for Visually Exploring Large Multidimensional Data Sets. In *Proceedings of Visualization 96, Hot Topic Session*, San Francisco, CA, 1996.

[2] Yannick Assogba, Irene Ros, and Matt McKeon. DocBlocks: communication-minded visualization of topics in U.S. congressional bills. In *CHI EA '10: Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, pages 4117–4122, New York, NY, USA, 2010. ACM.

[3] Edward Clarkson, Krishna Desai, and James Foley. ResultMaps: Visualization for Search Interfaces. *IEEE Transactions on Visualization and Computer Graphics*, 15:1057–1064, 2009.

[4] Edward C. Clarkson, Shamkant B. Navathe, and James D. Foley. Generalized formal models for faceted user interfaces. In *JCDL '09: Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*, pages 125–134, New York, NY, USA, 2009. ACM.

[5] Christopher Collins, Sheelagh Carpendale, and Gerald Penn. DocuBurst: Visualizing Document Content using Language Structure. *Computer Graphics Forum*, 28(3):1039–1046, 2009.

[6] Anthony Don, Elena Zheleva, Machon Gregory, Sureyya Tarkan, Loretta Auvil, Tanya Clement, Ben Shneiderman, and Catherine Plaisant. Discovering interesting usage patterns in text collections: integrating text mining with visualization. In *CIKM '07: Proceedings of the 16th ACM Conference on information and Knowledge Management*, pages 213–222, New York, NY, USA, 2007. ACM.

[7] Marian Dörk, Sheelagh Carpendale, Christopher Collins, and Carey Williamson. VisGets: Coordinated Visualizations for Web-based Information Exploration and Discovery. *IEEE Transactions on Visualization and Computer Graphics*, 14:1205–1212, 2008.

[8] Jean-Daniel Fekete and Nicole Dufournaud. Compus: visualization and analysis of structured documents for understanding social life in the 16th century. In *DL '00: Proceedings of the fifth ACM conference on Digital libraries*, pages 47–55, New York, NY, USA, 2000. ACM.

[9] Sébastien Ferré. Agile Browsing of a Document Collection with Dynamic Taxonomies. In *DEXA '08: Proceedings of the 2008 19th International Conference on Database and Expert Systems Application*, pages 377–381, Washington, DC, USA, 2008. IEEE Computer Society.

[10] Christiaan Fluit. AutoFocus: Semantic Search for the Desktop. In *IV'05: Proceedings of the Ninth International Conference on Information Visualisation*, pages 480–487, Washington, DC, USA, 2005. IEEE Computer Society.

[11] Andreas Girgensohn, Frank Shipman, Francine Chen, and Lynn Wilcox. DocuBrowse: faceted searching, browsing, and recommendations in an enterprise context. In *IUI '10: Proceeding of the 14th international conference on Intelligent user interfaces*, pages 189–198, New York, NY, USA, 2010. ACM.

[12] Marti A. Hearst. TileBars: visualization of term distribution information in full text information access. In *CHI '95: Proceedings of the SIGCHI Conference on Human factors in computing systems*, pages 59–66, New York, NY, USA, 1995. ACM.

[13] Marti A. Hearst. Clustering versus faceted categories for information exploration. *Communications of the ACM*, 49(4):59–61, 2006.

[14] Marti A. Hearst. UIs for Faceted Navigation: Recent Advances and Remaining Open Problems. In *HCIR '08: Proceedings of the Second Workshop on Human-Computer Interaction and Information Retrieval*, October 2008.

[15] Marti A. Hearst. *Search User Interfaces*. Cambridge University Press, New York, NY, USA, 2009.

[16] David Huynh, David Karger, and Rob Miller. Exhibit: Lightweight Structured Data Publishing. In *WWW '07: Proceedings of the 16th International World Wide Web Conference*, Banff, Alberta, Canada, 2007. ACM.

[17] Alfred Inselberg and Bernard Dimsdale. Parallel coordinates: a tool for visualizing multi-dimensional geometry. In *VIS '90: Proceedings of the 1st Conference on Visualization '90*, pages 361–378, Los Alamitos, CA, USA, 1990. IEEE Computer Society Press.

[18] Mika Käki and Anne Aula. Controlling the complexity in comparing search user interfaces via user studies. *Information Processing and Management*, 44(1):82–91, 2008.

[19] Eser Kandogan. Star Coordinates: A Multi-dimensional Visualization Technique with Uniform Treatment of Dimensions. In *Proceedings of IEEE Information Visualization, Hot Topics*, pages 4–8, 2000.

[20] Daniel A. Keim. Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics*, 08(1):1–8, 2002.

[21] Bill Kules and Robert Capra. Creating exploratory tasks for a faceted search interface. In *HCIR '08: Proceedings of the Second Workshop on Human-Computer Interaction*, 2008.

[22] Bill Kules, Robert Capra, Matthew Banta, and Tito Sierra. What do exploratory searchers look at in a faceted search interface? In *JCDL '09: Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*, pages 313–322, New York, NY, USA, 2009. ACM.

[23] Jeffrey LeBlanc, Matthew O. Ward, and Norman Wittels. Exploring N-dimensional databases. In *VIS '90: Proceedings of the 1st conference on Visualization '90*, pages 230–237, Los Alamitos, CA, USA, 1990. IEEE Computer Society Press.

[24] Bongshin Lee, Greg Smith, George G. Robertson, Mary Czerwinski, and Desney S. Tan. FacetLens: exposing trends and relationships to support sensemaking within faceted datasets. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 1293–1302, New York, NY, USA, 2009. ACM.

[25] Francois Lorrain and Harrison C. White. Structural equivalence of individuals in social networks. *Journal of Mathematical Sociology*, 1:49–80, 1971.

[26] Josiane Mothe, Claude Chrisment, Bernard Dousset, and Joel Alaux. DocCube: multi-dimensional visualisation and exploration of large document sets. *Journal of the American Society for Information Science and Technology*, 54(7):650–659, 2003.

[27] Kai A. Olsen, Robert R. Korfhage, Kenneth M. Sochats, Michael B. Spring, and James G. Williams. Visualization of a document collection: the VIBE system. *Information Processing and Management*, 29(1):69–81, 1993.

[28] Catherine Plaisant, Tom Bruns, Ben Shneiderman, and Khoa Doan. Query previews in networked information systems: the case of EOSDIS. In *CHI '97: Extended abstracts on Human factors in computing systems*, pages 202–203, New York, NY, USA, 1997. ACM.

[29] Catherine Plaisant, Jean-Daniel Fekete, and Georges Grinstein. Promoting Insight-Based Evaluation of Visualizations: From Contest to Benchmark Repository. *IEEE Transactions on Visualization and Computer Graphics*, 14(1):120–134, 2008.

[30] Ramana Rao and Stuart K. Card. The table lens: merging graphical and symbolic representations in an interactive focus + context visualization for tabular information. In *CHI '94: Proceedings of the SIGCHI Conference on Human factors in computing systems*, pages 318–322, New York, NY, USA, 1994. ACM.

[31] Ben Shneiderman. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *IEEE Visual Languages*, pages 336–343, 1996.

[32] Greg Smith, Mary Czerwinski, Brian Meyers, Daniel Robbins, George Robertson, and Desney S. Tan. FacetMap: A Scalable Search and Browse Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):797–804, 2006.

[33] Michael Spenke, Christian Beilken, and Thomas Berlage. FOCUS: the interactive table for product comparison and selection. In *UIST '96: Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 41–50, New York, NY, USA, 1996. ACM.

[34] Moritz Stefaner, Sebastian Férré, Saverio Perugini, Jonathan Koren, and Yi Zhang. *Dynamic Taxonomies and Faceted Search: Theory, Practice, and Experience*, volume 25 of *The Information Retrieval Series*, chapter 4 - User Interface Design, pages 75–112. Springer, 2009.

[35] Moritz Stefaner, Thomas Urban, and Marc Seefelder. Elastic Lists for Facet Browsing and Resource Analysis in the Enterprise. In *Proceedings of the International Workshop on Database and Expert Systems Applications*, pages 397–401, Los Alamitos, CA, USA, 2008. IEEE Computer Society.

[36] Chris Stolte, Diane Tang, and Pat Hanrahan. Polaris: A System for Query, Analysis, and Visualization of Multidimensional Relational Databases. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):52–65, 2002.

[37] VinhTuan Thai and Siegfried Handschuh. Enhanced Navigation and Focus on TileBars with Barycenter Heuristic-based Reordering. In *AVI '10: Proceedings of the International Conference on Advanced Visual Interfaces*, pages 349–352, New York, NY, USA, 2010. ACM.

[38] VinhTuan Thai, Siegfried Handschuh, and Stefan Decker. IVEA: An Information Visualization Tool for Personalized Exploratory Document Collection Analysis. In *ESWC '08: Proceedings of the 5th European Semantic Web Conference*, pages 139–153. Springer, 2008.

[39] VinhTuan Thai, Siegfried Handschuh, and Stefan Decker. Tight coupling of personal interests with multi-dimensional visualization for exploration and analysis of text collections. In *IV '08: Proceedings of the 12th International Conference on Information Visualisation*, pages 221–226, Washington, DC, USA, 2008. IEEE Computer Society.

[40] J. A. Wise, J. J. Thomas, K. Pennock, D. Lantrip, M. Pottier, A. Schur, and V. Crow. Visualizing the non-visual: spatial analysis and interaction with information from text documents. In *INFOVIS '95: Proceedings of the 1995 IEEE Symposium on Information Visualization*, page 51, Washington, DC, USA, 1995. IEEE Computer Society.

[41] Ka-Ping Yee, Kirsten Swearingen, Kevin Li, and Marti Hearst. Faceted metadata for image search and browsing. In *CHI '03: Proc. of the SIGCHI Conf. on Human factors in computing systems*, pages 401–408, New York, NY, USA, 2003. ACM.

[42] Junliang Zhang and Gary Marchionini. Coupling browse and search in highly interactive user interfaces: a study of the relation browser++. In *JCDL '04: Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*, pages 384–384, New York, NY, USA, 2004. ACM.

APPENDIX

Tasks

Below is one of the four sets of tasks used in the user study:

1. "*...Even as the celebrations unfolded, accountants and trading experts at the company's Houston headquarters were desperately working to contain a financial disaster, one that threatened – and ultimately would destroy – everything Enron had become. A handful of executives were struggling to sound the alarm, but with Enron's confidence in its destiny, the warnings went unheeded...*"

In this context, you may want to focus on emails that mentioned accounting and Kenneth L. Lay. Among them,

a. How many emails mentioned loss ?

b. Among those that did not mention loss, how many emails in 1999-2002 mentioned alarm.

2. "*On Aug. 14, stunning the market, Jeff Skilling announced he was resigning after just six months as chief executive, citing undisclosed personal reasons. He left assuring investors that the finances of Enron had never been better.*"

a. You may want to find out more about internal communications surrounding this event. Identify emails in 2000-2004 which mentioned Jeff Skilling, resignation, and the role Chief Executive Officer.

   Among them,

b. Identify emails that mentioned investigation

c. Identify emails that did not mention investigation.

3. Filter for emails of interest based on either of the two contexts above.

Overall questionnaire

- Do you think the tasks are equally difficult? (apart from the different result set sizes).
- Which interface is the easiest to use?
- Which interface helps you learn the most about the document collection while exploring?
- Did the matrix-based presentation of result items change the way you used to explore a collection of text documents?
- Can you describe an example where the matrix-based presentation of result items helped / hindered, frustrated your exploration process?
- Do you think the grouping of documents containing exactly the same set of entities helped or hindered exploring a large number of relevant documents? (both groupings in list views and in matrix views)
- Do you think the ability to abstract the view on the matrix simultaneously on two dimensions (concepts and documents) helped or hindered the document exploring process?
- Do you think that the interactive ordering in a matrix-based presentation of result items helped or hindered while you are trying to have a clearer idea of the result set? Can you describe an example ?
- Would it make a difference if the matrix only used one color ?
- What was your experience when browsing/ filtering freely without any guided tasks in order to gather documents potentially containing relevant pieces of information?
- Overall, how would you rank the interfaces? Please provide comments on your ranking.