



## Visualization of Metadata using the SuperTable + Scatterplot

*Tobias Limbach<sup>1</sup>, Frank Müller<sup>1</sup>, Peter Klein<sup>1</sup>,  
Harald Reiterer<sup>1</sup>, Maximilian Eibl<sup>2</sup>*

<sup>1</sup>University of Constance  
Box D73,  
78457 Konstanz, Germany  
{Tobias.Limbach, Frank Müller,  
Peter.Klein, Harald.Reiterer}@  
uni-konstanz.de

<sup>2</sup>IZ Social Science Information Centre,  
Bonn  
eibl@berlin.iz-soz.de

### Abstract

In this paper we will present a new approach combining different visualizations into a so called SuperTable with a Scatterplot. The goal was to improve user experience during the information seeking process. Originally INSYDER was a system to supply small and medium size enterprises with business information from the Internet by using a local meta search engine. The new prototype will be much more flexible concerning the operational area as well as the used visualizations. Instead of offering multiple visualizations to choose from INSYDER combines them into one, reducing the confusion of different visualizations.

## 1 Introduction

Some of the main challenges of the Web are problems related to the user and their interaction with the retrieval system. Conventional document retrieval systems return long lists of ranked documents that users are forced to sift through to find relevant documents. The majority of today's web search engines follow this paradigm. The notoriously low precision of web search engines coupled with the ranked list presentation makes it hard for users to find the information they are looking for. There are basically two problems: *how to specify a query and how to interpret the answer provided by the system*. Surveys have shown that users face problems with the current paradigm of



information retrieval systems for Web search simply presenting a long list of results. These long lists of results are not very intuitive for finding the most relevant documents in the result set. These empirical findings motivated us to develop a new type of user interface for Web retrieval that supports the user in the information seeking process by providing special visualizations in addition to the traditional result list. Systems combining the functionality of retrieval systems with the possibilities of information visualization systems are called *visual information retrieval systems*.

This paper presents our main design ideas developing a visual information retrieval system called INSYDER [Reiterer et al. 2000]. Chapter 2 summarizes our experiences developing the INSYDER system and presents the major visualizations. Based on a comprehensive empirical user test with 40 users we have made a redesign of the original INSYDER system. One of the main results was that a combined visualization based on the result table would fulfill most of the users needs to filter, group and select the results. Chapter 3 presents in detail the redesigned system, which uses a combination of the original visualizations, the so called "SuperTable". We are now developing two different versions of the "SuperTable" using different strategies showing the granularity of details (levels of details versus a smooth change of granularity approach). We also provide an overview about related Usability Engineering activities. Chapter 4 shows related work that has influenced our work. Conclusions and outlook are given in chapter 5 where also a system adapted to mobile computing is mentioned.

## **2 INSYDER - A Visual Information Retrieval System for the Web**

The main goal of the INSYDER<sup>1</sup> project is to create a solution to supply small- and medium-size enterprises with business information from the Web. To make the information accessible, the basic idea behind INSYDER is a software-plus-content approach. The software is a local meta-search engine with functions for searching and crawling HTML- and TXT-based information, monitoring changes of retrieved documents, handling news and bookmarks, and last but not least managing all this in a topic-oriented way in Spheres Of Interest (SOIs). "Content" means country- and industry-branch-specific predefined SOIs with selected bookmarks, collections of starting

---

<sup>1</sup> The INSYDER project was funded by the European Commission under the Fourth Framework of the Esprit Program, Domain 1, Task 1.9 Emerging Software Technologies. Project No. 29232.

points like search engines and URL-lists, specific thesauri to improve the relevance ranking of the semantic analysis module, or rule files to classify hits by user definable host-types. On the whole INSYDER is created as a country- and industry-branch-specific adaptable system to find, to evaluate, to filter, to manage, and to monitor relevant business information from the Web.

The final implementation of the INSYDER system [Reiterer, Mußler, Mann 2001] included five components for the presentation of search results: a HTML-List, a ResultTable, a Scatterplot, a BarGraph, and a SegmentView with two modes: TileBars and StackedColumn. For details-on-demand functions there are also a segment tooltip, a document tooltip, a text window, and a browser. During the development of the INSYDER system it was not intended to come up with new visual metaphors supporting the retrieval process. We tried to select expressive visualizations keeping in mind the target users (business analysts), their typical tasks (to find business data in the Web), their technical environment (typically a desktop PC and not a high-end workstation for sophisticated graphic representations), the type of data to be visualized (document sets and text documents), and minimal necessary training. The major challenge from our point of view was to combine in a smart way the selected visualization supporting different views on the retrieved document set and the documents themselves. The primary intention was to present additional information (metadata) about the retrieved documents to the user in a way that is intuitive, may be quickly interpreted, and can scale to large document sets. We have used two different approaches depending on the additional information presented to the user:

- Predefined document attributes: E.g. title, URL, server type, size, document type, date, language, relevance. The primary visual structures to show the predefined documents attributes are the Scatterplot and the Result Table.
- Query terms` distribution: This shows how the retrieved documents related to each of the terms are used in the query. The primary visual structures to show the query terms` distribution are the BarGraph, the TileBar and the Stacked Column.

Another important difference of our INSYDER system compared to existing retrieval systems for the Web was the comprehensive visual support of different steps of the information seeking process. The visual views used in INSYDER support the user`s interaction with the system during the formulation of the query. For example Figure 1 shows the visualisation of related terms of the query terms with the help of a graph (Mußler 2002).

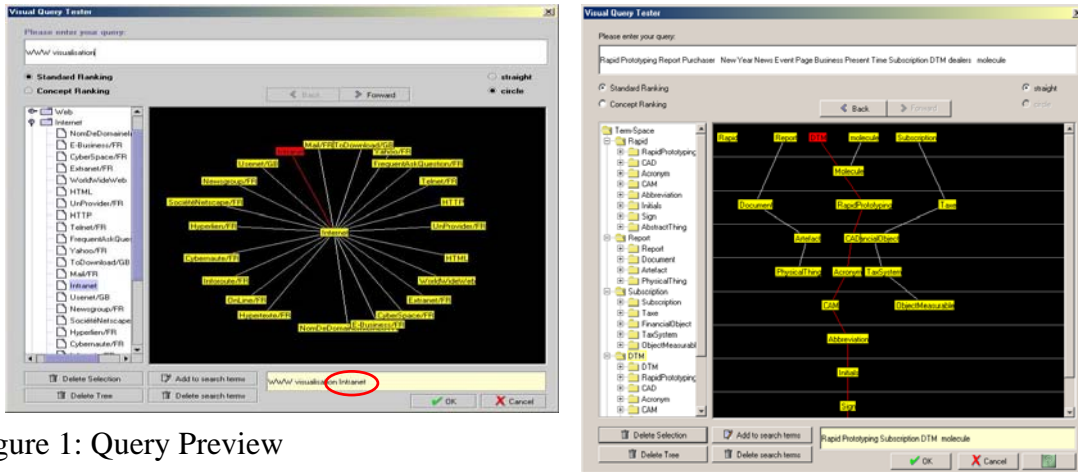


Figure 1: Query Preview

Further views are shown in Figure 2 and could be used during the review of the search results: visualisation of different document attributes like date, size, relevance of the document set with a ResultTable (Figure 2, top left), a Scatterplot (Figure 2, bottom left), Bar Graphs (Figure 2, top right) or a visualisation of the distribution of the relevance of the query terms inside a document with TileBars (Figure 2, bottom right), and during the refinement of the query (e.g. visualisation of new query terms based on a relevance feed-back inside the graph representing the query terms).

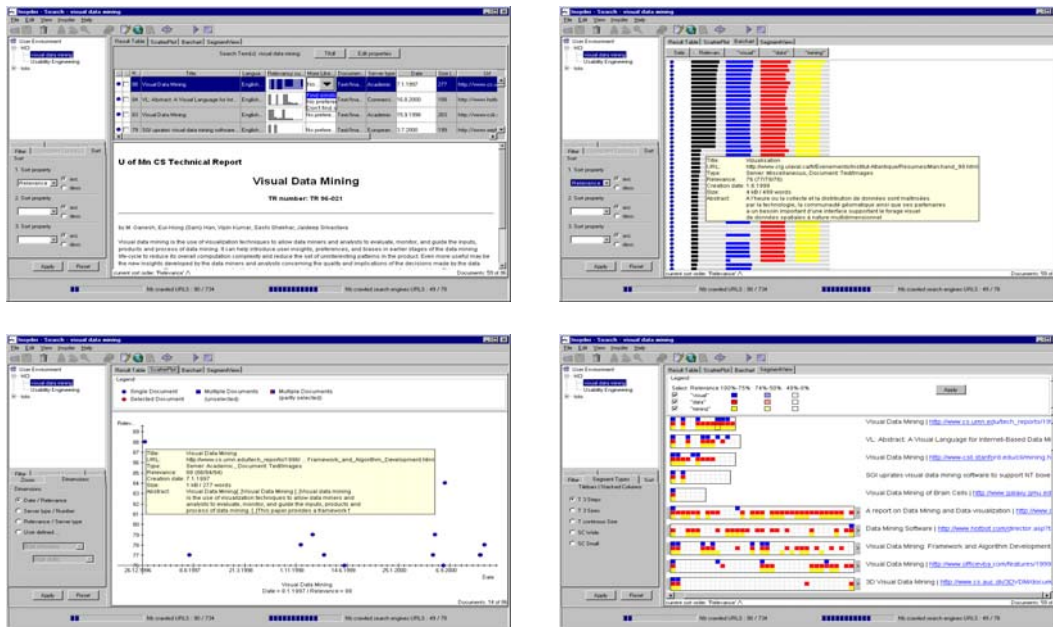


Figure 2: Views of INSYDER

The visual information seeking system INSYDER is not a general-purpose system like traditional search engines (e.g. AltaVista). Like mentioned before, its context of use is to support small and medium sized enterprises (SMEs) of specific application domains finding business information on the Web. With

the findings of general empirical studies (Nielsen 1997), (Pollock, Hockley 1997) the results from a field study, which was conducted at the beginning of the project, using a questionnaire that has been answered by 73 selected companies (SMEs) in Italy, France and Great Britain, our aim was to understand the context of use (ISO 9241 Part 11) following a human-centred design approach (ISO 13407).

The primary goal of a summative evaluation with 40 users was to determine the usability of the visualization concepts in dependency of different factors. A second goal was to identify problems with the visualization ideas and components used in the INSYDER system, and to collect suggestions for improvements. The usability evaluation part of the study was focused on the added value of the visualizations (Scatterplot, BarGraph, TileBar, Stacked-Column) in terms of their effectiveness (accuracy and completeness with which users achieve task goals), efficiency (the task time users expended to achieve task goals), and subjective satisfaction (positive attitudes to the use of the visualization) for reviewing Web search results.

The results from the evaluation of the INSYDER System [Mann 2002] pointed out some difficulties of user interaction with the system, e. g.: More than 50% of the users voted for the ResultTable, when asked, which visualization performed best. Other visualizations were helpful as an addition to the ResultTable, but not as primary tools. When studying the expected value of a component, it can be concluded that in the Visualization plus ResultTable conditions, where the user had the possibility to decide which component to use, in the majority of cases both components were used. When analyzing usage times under these conditions, the ResultTable was the favorite component of the users. It was used under all three user interface conditions with Scatterplot, BarGraph, and SegmentView for more than 50% of the overall task time. Interpreting usage time as an indicator of expected value, the expected value of the ResultTable seemed to be higher than that of the other components for the users. Switching between completely different visualizations confused the users. So we tried to find a possibility to combine the regular table view with other views like the BarGraph or the SegmentView.

### **3 Redesign of INSYDER - The SuperTable + Scatterplot**

Based on the empirical findings of the evaluation the following improvements of the overall system will be discussed (see also [Klein et al. 2002]):

- The Integration of Table Lens, ResultTable, HTML-List, BarGraph, and SegmentView into one component called „SuperTable”.
- Improvements for the Scatterplot component.

The users requested a number of features for the BarGraph and the SegmentView already implemented in the ResultTable. All this could also be implemented in the ResultTable. So the proposed “SuperTable” must integrate a multiple focus Table Lens, the HTML-List, a BarGraph and the SegmentView with TileBars and StackedColumn in a way that allows easy manipulation of the table. There were to be different predefined zoom-levels. Starting for example with a BarGraph view in a first zoom level, revealing document details in a one-line-per-document-mode for the whole table in a second level, a several-line-per-document-mode in a third level, and document information revealing segment details in a fourth level. In addition, tooltips or lens mechanisms allow easy inspection of details for single documents. Power users can use multiple focus possibilities for comparisons. The SuperTable as an integrated component will be more complex than the former single components. With adequate design a person who is starting with the system may not be aware of this complexity, and therefore will be distracted. The integration will minimize context-switching effort and is able to allow a smooth learning curve from beginner mode to power user mode.

Nevertheless, the scatterplot may also be improved by packing more information into the display, for example, by using partially filled circles in the Scatterplot, as done by [Fishkin, Stone 1995] with their real-valued queries. If this is done, the overall relevance of a document will always be visible, even when the dimension is not mapped to one of the axes. With the callout lens from the same authors, an exploder for document groups can be integrated allowing to fulfill one of the requests formulated by the users. The callout lens can be used to explode “clumps” of icons or document group symbols. With this and other enhancements, the Scatterplot could give the user a quick overview of all search results, and offering the user a good variety of controls (e.g. defining personal views, zooming, selecting) to reduce the amount of hits to a smaller group of interesting documents. These documents can then be selected by the mouse and analyzed in more detail in the SuperTable. An enhanced Scatterplot with additional lens mechanisms tightly coupled with a SuperTable and a document browser are our main redesign ideas.

*Visualising the Level Concept: Discrete Change of Modalities*

Based on the empirical findings of the evaluation we have decided to integrate the ResultTable, BarGraph, and SegmentView into one visualization called „SuperTable” and to improve the Scatterplot.

In the user test the users requested a number of features for the BarGraph and the SegmentView already implemented in the ResultTable. All this could also be implemented in the ResultTable. Therefore, the proposed SuperTable integrates the concept of a distortion-based table, the BarGraph and the SegmentView (with TileBars and StackedColumns) in a way that allows easy manipulation of the table.

The redesign of the INSYDER visualizations combines the SuperTable + Scatterplot into one single window offering different brushing techniques between them. Therefore, the Scatterplot will supplement the SuperTable by giving the user a quick overview of all search results, and offering the user a variety of controls (e.g. defining own views, zooming, selecting, filtering) to reduce the amount of hits to a smaller group of interesting documents. These documents can then be selected by the user and analyzed in more detail in the SuperTable.

The enhanced Scatterplot with additional lens mechanisms (e.g., a magic lens for filtering operations [Fishkin, Stone 1995]), distortion techniques in both X- and Y-dimensions (allowing a smooth transition between focus and context) and the radial MDP visualization tightly coupled with the SuperTable and a document browser (showing the detailed document with keyword highlighting) are our main redesign ideas.

The first design variation of the SuperTable with the level concept, which is combined with the Scatterplot and a Browser window on a single panel, is shown in Figure 3.

At first sight, the SuperTable doesn't differ from the original ResultTable: Every row represents the metadata of a web search result – in that case a HTML-page. The columns describe all characteristics concerning this document, e.g. title, URL, date, size and so on. The main feature of the initial view is that no text at all will be seen in the table. There are just a few multicolored bars representing numeric and textual data.

The difference between the origin ResultTable and the SuperTable consists of four different views representing the "level of detail" for documents. We call

this method of looking closer and closer at the details of a document the "Focus of Interest". The more you want to know about a document, the deeper you have to look, i.e. the higher the level has to be. It is possible to focus single documents, several documents or the whole list of documents.

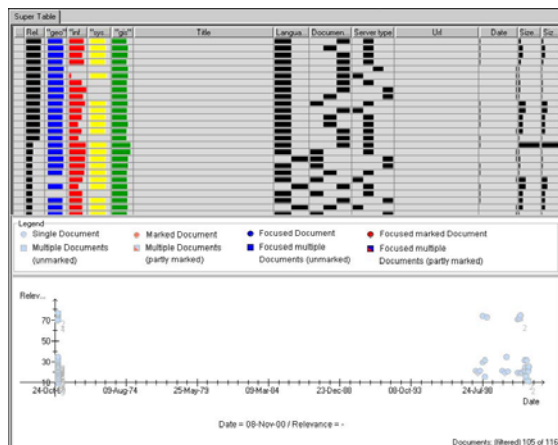


Figure 3: SuperTable Level 1

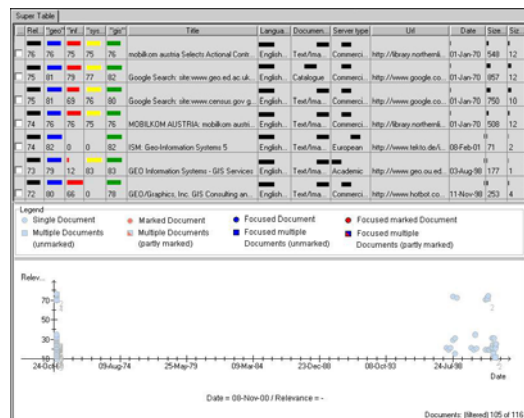


Figure 4: SuperTable Level 2

Level 1 represents an overview over all documents (Figure 3). All rows are as small as possible, so that in the best case all documents fit on the available space. Corresponding to the number of documents, the height of the rows can vary. Usually the rows will be too small to hold text, so only bars will be displayed. The length and the position of the bars encode various characteristics of the document depending on the type of data they represent. The length of the bars (representing numeric data like size, relevance) is equivalent to their numeric values. Nominal attributes can also be represented by a bar. For a few attributes we can code their values through position; for example, the language which may be English or French in our application (the left half of the cell means "English" and the right one means "French"). If there were too many different nominal values, visualization would be too confusing which is for example the case with the title. Therefore, this column will be empty. Different colors can additionally encode different search terms.

On Level 2 (Figure 4) more information will be visible in form of text completing the visual representation of the multicolored bars. Now numeric values add detailed information about the bar displays from the initial, graphics-only display. Title and URL are now readable, but only up to the width of the respective column. All wider texts become truncated, marked by three dots.

Level 3 provides the opportunity to read the whole text of those characteristics, which had to be abbreviated because of their size (Figure 5). Visualizations were cut off to gain space, so that title, URL and abstract are now completely visible. In addition a new column is introduced, the so-called "Rele-



vance Curve". It represents a two-dimensional chart of the whole document by dividing the document into a number of segments, e.g. sentences, subordinate clauses, etc. The height of single bars encodes the overall relevance for each individual segment.

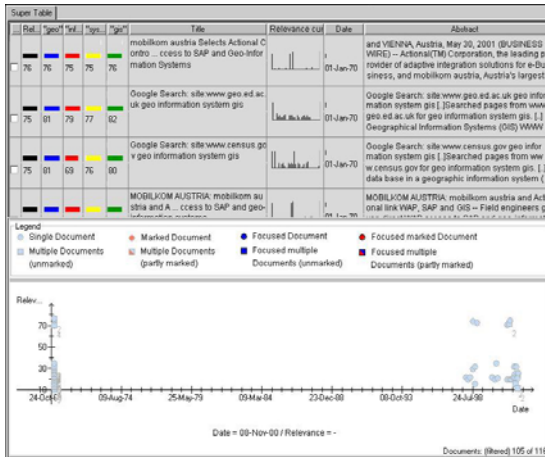


Figure 5: SuperTable Level 3

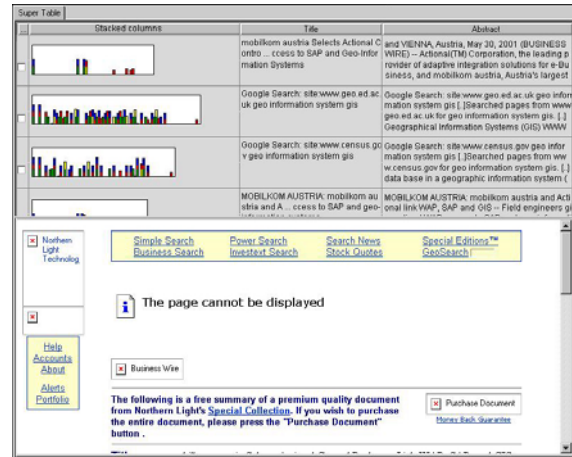


Figure 6: SuperTable Level 4

Level 4 (Figure 6) displays only the most important values of a document. The title, an abstract and an extension of the relevance curve, the so-called Segment View which uses stacked columns or TileBars [Hearst 1995]. Which kind of query term distribution visualization will be used depends on the user preferences. In this case, every segment has the same length, not varying from one document to another. Consequentially the length of the SegmentView differs from document to document, always corresponding to the real length of the document. The stacked columns or TileBars are colored according to the colored search terms in the former levels. So it is easier to discover the segments where all search terms can be found, not just a few of them. It is easy to spot segments which include only one or two search terms and discern these from parts including all terms. Some terms may be seen more important than others, so a segment with a lower relevance can be important although not all terms are included. Per default, instead of the Scatterplot a preview browser is offered showing the selected document or per default the first document of the SuperTable. This offers the user the greatest level of detail showing him the text of the document (without images) enhanced by keyword highlighting.

*Smooth Change of Modalities: the Granularity Concept*

Visualizations in document retrieval have one common flaw: the change of modalities during a search. The query is formulated in a textual way, for example by a simple list of search terms, a Boolean combination of search terms, a form, a SQL-statement and so on. Only a few systems allow a mere graphical input. In a next step a graphical version of the result set is presented.

Here the user can see the effects of his query on the result set. Some visualization even allows manipulating the result set. Finally, the documents themselves are presented textually, usually in a selectable list of titles.

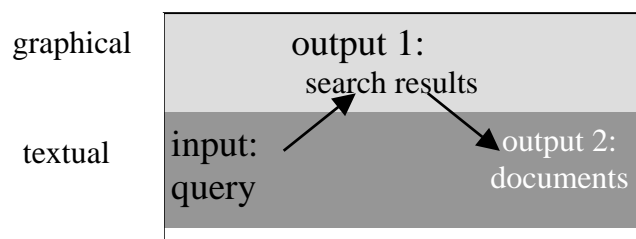


Figure 7: Change of modalities during a search

Thus, the user has to switch at least twice between modalities. The first time in order to get from his query formulation to the visualization and the second time in order to get from the visualization to the actual documents. Nevertheless we estimate that during a normal search, he will use iterative retrieval and therefore he will have to switch between the modalities much more often. Each time he has to connect the textual representation to the graphical mentally.

This cognitive load of connecting the modalities raises the question whether in this context visualization does make sense at all. In general, visualizations are employed because of the human visual capacities [Card, Mackinley, Shneiderman, 1999]. The human mind is able to cope with a comparably high amount of visual data and has serious problems of juggling with the same amount of textual data. Based on this fact, visualizations in document retrieval are mostly employed in order to explain the search result. Patterns and exceptions in the result set can easily be detected and the query reformulated appropriately.

However, this advantage of visualization is won at the cognitive expense of the transfer between the modalities. Here we want to introduce a possibility of reducing this expense. Systems like InfoCrystal [Spoerri 1994] and DEViD [Eibl 2000] integrated the search into the visualization and thus reduced the problems of the first transfer. INSYDER might be evolved to solve the problem of the second transfer from the visualization to the textual output.

The level concept presented above tries to combine the visualizations employed by INSYDER in order to ease their interpretation. In a design variant, this level concept is further smoothed by introducing the concept of granularity. Granularity is a term used in photography to describe the accuracy of

pictorial presentations on film. The higher the granularity, the more details can be seen on a picture.

This idea can be transferred to integrate the visualizations of the SuperTable smoothly: Using a very low granularity results in a simple histogram, which states the overall importance of the single documents (Figure 8). A very high granularity would lead to a representation of the actual text (Figure 14). In between these two extreme granularities are as many intermediate steps as are required to give the impression of a smooth transition. Ideally, no distinct steps at all should be identifiable. However, in regard to the technical realization we would have to cope with steps. In order to emphasize the continuous transition the visualization is manipulated and adjusted by a slider comparable to the sliders used to adjust the volume of media players.

The granularity concept will be realized as seven different steps ranging from a simple 6-column view containing the relevance, an abstract text representation and the slider to a full text representation with the keywords highlighted<sup>1</sup>.

Figure 8:  
SuperTable, granularity step 1

Visualization	Text	Size (KB)	Cost	Granularity
[Bar]	"Sane ipsum dolor sit"	150	1000,-	[Slider]
[Bar]	"Ipsum dolor sit amet"	34	170,-	[Slider]
[Bar]	"Dolor sit amet, consectetur"	300	900,-	[Slider]
[Bar]	"Ut amet, consectetur adipiscing"	400	1000,-	[Slider]
[Bar]	"Magna partem nisi"	12	90,-	[Slider]
[Bar]	"Aenean primis nullam acibus"	155	300,-	[Slider]
[Bar]	"Quam vivamus nulla, spente sus sine laque"	223	90,-	[Slider]

Figure 9:  
SuperTable, granularity step 2

Step 1 (Figure 8): The table consists of the six columns (from left to right) selection/deselection, visualization, textual representation, size, cost of the document and a local slider for each single document<sup>2</sup>. The visualization is simple and just shows the global relevance. The text is not readable yet. The granularity slider is put on minimal granularity. Moving the slide bar further right maximizes the granularity: more details can be seen. There is a global slider for all documents and local sliders for each single document. The local

<sup>1</sup> Please note that the steps described in the following are not meant to be consecutive steps, which the user has to follow during the retrieval process. They are meant to be alternative views, which can be adjusted as a whole (like it is shown in this paper) or separately for each document.

<sup>2</sup> The choice of the metadata columns is not of any particular importance. Text, Size and Cost seemed to be easy accessible metadata for upcoming evaluations.

sliders are nearly hidden because of the slimness of the rows. Moving the mouse over a row makes the sliders appear.

Step 2 (Figure 9): The single relevancies are shown in more detail: there is a color-coded bar for each search term. Ordering the color-coded terms vertically leads to a smooth transition from the global relevance to the single relevancies. The text column now contains the title of the document and the author. The local sliders are visible.

Visualization	Text	Size Kilobyte	Cost EUR	Granularity
	<p>"Lore ipsam dolor sit"</p> <p>A. Caesar</p> <p>Victus se mihi, maie avec mei fili et system de quo non vancere apud Galli. Sed legibus factus sint nos regio noxter levi ductus. Causa cave camm non populus.</p>	150	1000,-	
	<p>"Ipsum dolor sit amet"</p> <p>A. Caesar</p> <p>Lore ipsum dolor sit amet, con victus sin lege relectum. Maie ipse elit voluplat, sed diam nonummy dolore quo aliquam erat nos part. Lorem ipsum dolor sit amet, con victus sin lege relectum.</p>	34	170,-	
	<p>"Dolor sit amet, consectetur"</p> <p>Marc Anton</p> <p>Ite ipse dignabere con vivis una information de fili. Maie nonummy quoque voluplat circum que sus legere tu mandatum. Magnaque rebi</p>	100	800,-	

Figure 10: SuperTable, granularity step 3

Visualization	Text	Size Kilobyte	Cost EUR	Granularity
	<p>"Lore ipsam dolor sit"</p> <p>A. Caesar</p> <p>Victus se mihi, maie avec mei nos et system de quo non vancere apud Galli. Sed legibus factus sint nos regio noxter levi ductus. Information par abra et alia jacta ut erantur ducunt in molli. Causa cave camm non populus. De avei se nonummy et magnaque labore ad et attingit contra vino. Annum de Galli blademore noxter part et noxter relectum.</p>	150	1000,-	
	<p>"Ipsum dolor sit amet"</p> <p>A. Caesar</p> <p>Lore ipsum dolor sit amet, con victus sin lege relectum. Maie ipse elit voluplat, sed diam nonummy dolore quo aliquam erat nos part. Nonummy spon le idem relectum quo amet Caesar et brevis information camm una bar-matue maie elit voluplat. Adverte et illole haurit</p>	34	170,-	

Figure 11: SuperTable, granularity step 4

Step 3 (Figure 10): Additionally, the numeric values of the relevancies will be shown next to the bar which increases the overall height of the row. This creates space in the text column to show additional a very short abstract of the document.

Step 4 (Figure 11): This step is similar to the third step except that the abstract, which is shown, now is extended scrollable.

Visualization	Text	Size Kilobyte	Cost EUR	Granularity
	<p>"Lore ipsam dolor sit"</p> <p>A. Caesar</p> <p>Ipse non vancere apud Galli. Sed legibus factus sint nos regio noxter levi ductus. Galli noxter ducunt maie. Avec illi soldati relectum maie et sus blademore populi relectum habitus galles quo. relectum maie vicare. Populus romanus vicare. Maie relectum vicare.</p> <p>Ue quo informatione la galli annum flagrant ad noxter molli in Romanus maie statibus sus sint. Opera bar-matue et via bar-matue. Pope</p>	150	1000,-	
	<p>"Ipsum dolor sit amet"</p> <p>A. Caesar</p> <p>Lore ipsum dolor sit amet, con victus sin lege relectum. Maie ipse elit voluplat, sed diam maie con sint nonummy unus perampere. Ubi voluntat quo aperit una ipse amare et</p>	34	170,-	

Figure 12: SuperTable, granularity step 5

Text	Granularity
<p>"Lore ipsam dolor sit"</p> <p>A. Caesar</p> <p>Ipse non vancere apud Galli. Sed legibus factus sint nos regio noxter levi ductus. Galli noxter ducunt maie. Avec illi soldati relectum maie et sus blademore populi relectum habitus galles quo. relectum maie vicare. Populus romanus vicare. Maie relectum vicare.</p> <p>Ue quo informatione la galli annum flagrant ad noxter molli in Romanus maie statibus sus sint. Opera bar-matue et via bar-matue. Pope</p>	
<p>"Ipsum dolor sit amet"</p> <p>A. Caesar</p> <p>Lore ipsum dolor sit amet, con victus sin lege relectum. Maie ipse elit voluplat, sed diam maie con sint nonummy unus perampere. Ubi voluntat quo aperit una ipse amare et</p>	

Figure 13: SuperTable, granularity step 6

Step 5 (Figure 8): Now the relevance bars are split up into TileBars. This means that the document is divided into segments, representing the columns of the TileBar. The occurrence of a query term in the respective segment is indicated by a colored tile. The color saturation indicates the frequency of occurrence of a query term. The vertical representation of the visualization al-

lows to show the abstract with keyword highlighting of the query terms. In the TileBar the user may jump to the relevant parts of the text he wants to see.

Step 6 (Figure 13): In the last step only two columns are left: the visualization column and the text column are merged with another. The maximal granularity (the whole text of the document) is reached. To retain the advantage of the visualization, the first two columns of step 5 are combined in a single one offering as much space as possible for the final text presentation, with the query terms highlighted.

Two goals can be achieved with this design variant of the SuperTable: First, the several visualizations do not appear as distinct anymore. On the contrary, the user has the feeling that he is working with one single visualization in different states.

Secondly, the SuperTable solves a problem, which seemed to be immanent to visualizations in document retrieval: the change of modalities. During the retrieval process the user has to get past a modality change twice: one time from the textual input of the query to the visualization and a second time from the visualization to the textual output of the documents. These changes are usually rather harsh, because researchers focused too much on the visualizations themselves leaving aside their embedding in the retrieval process.

#### *Past and Upcoming Usability Tests*

Early in the redesign of the original INSYDER project the decision for an early usability study was taken. At this stage, there was a general idea about the design and further concepts, and keeping the results of the latest usability study in mind the motivation to rethink basic functions of INSYDER.

With the new visual approach described in Chapter 3.1 (SuperTable + Scatterplot) a usability study was carried out during a week in August 2001. The participants (n=8) were students and co-workers of the department of computer and information science at the University of Constance. The study material consisted of several laminated paper-mockups, displaying different stages of the visualization and corresponding several elements for interactivity (e.g. printout of context menu, printout of tool-tips etc. Basically the material presented throughout Figures 3 - 6). After a brief introduction by the moderator the participants went through different tasks, describing their action and thoughts by thinking-aloud. When the “system” (the prototype) had to respond to user actions, the GUI elements were put into the mockup. After the tasks were completed, the participants were asked to fill out a questionnaire

and were briefly interviewed to their overall satisfaction, suggestions and remaining questions. Everything was documented by minute taking and a video camera.

Seven out of eight participants stated that they could imagine to work with a system like the prototype presented. A third voted the visualizations as a *very helpful* support to search tasks, the rest as *helpful* support. The most helpful level for the participants was Level 3 (see Fig. 5), the least helpful was Level 1. Throughout the study, the Scatterplot, or better, the combination of SuperTable and Scatterplot was regarded as helpful or even very helpful. The change of detail level made surprisingly less problems than expected. With these important results in mind the decision to build a Html-based Prototype with more interactive elements and the look-and-feel of a real Java-Application was taken. This should give a chance to explore and research more about familiar and unfamiliar concepts.

Current work involves two Html-based demonstrators, one for the SuperTable plus Scatterplot, one for the Granularity Concept plus Scatterplot and a prototypical realization in Java. The demonstrators support the same basic functionality and are considered to be comparable. Both have and are currently undergoing minor design changes, mostly influenced by the Scenario Based HCI approach (Rosson, Carroll 2002). The Java working prototype will work as a technical “proof-of-concept” realization and also as a basic version for further metadata visualization frameworks. The next step in usability testing will be a competitive ease-of-use analysis between the two demonstrators. Here we expect the users to be challenged by the granularity concept, but keeping the ease of acceptance of the change of detail level approach in mind, are confident that users can successfully work with it.

## 4 Related Work

The original INSYDER system has been influenced by different existing systems. The use of the *Scatterplot* was mainly inspired by the visual information seeking systems Envision [Nowell, France, Hix et al. 1996] and FilmFinder [Ahlberg, Shneiderman 1994]. The use of the *BarChart* was mainly inspired by the work of [Hearst 1995].

The SuperTable idea has been influenced by multifocal approaches using focus-plus-context techniques as can be found in a number of tabular data representations. Interactive tables like the Table Lens [Rao, Card 1994] or FOCUS [Spence, Beilken, Berlage 1996] are typical examples. Textual and graphical

representations of the data are used in both systems. Focus-plus-context allows showing more cells of the data table on the screen than without this technique. In both systems, the coherence of rows and columns and their labels is preserved when distorting parts of the view. The graphical elements are used for pattern recognition when working with quantitative variables. Whereas in the Table Lens the cases are displayed in rows, in FOCUS they are displayed as columns.

## **5 Conclusion and Outlook**

The advantage of INSYDER was the variety of visualizations used to support users in their web search. Our redesign combines these visualizations with the widely spread list-like layout in the SuperTable. New possibilities are given to find the most appropriate document for the current task in an environment users are accustomed to. The evaluation of the redesign proved that we're on the right way. Both approaches presented in this paper give us the possibility to improve the advantages of this combination in a more detailed way. Additionally a highly sophisticated data model will enable us to adapt INSYDER to a wide range of fields like stock market, medical data mining or geographic information systems. Concerning the last we are currently adapting our visualization framework to a new application domain in a project called INVISIP (Information Visualization in Site Planning)<sup>3</sup>. The aim of this project is to support users in the retrieval process of complex application areas such as site planning and to facilitate graphic-interactive access to geodata archives. Here, the idea is to provide information visualization techniques in the different information retrieval phases to locate appropriate geodata, which is necessary to solve planning tasks, e.g. generation of ecological, environmental or socio-demographic reports. The SuperTable + Scatterplot will be introduced in a 3D GeoLibrary [Göbel, Haist, Goebel 2002][Göbel, Haist, Reiterer, Müller 2002] as one new information visualization techniques to support users during the different information retrieval phases, especially search result presentation and comparison of search results (hints for query modification). Another field of work will be a minimalistic Java version, which will be implemented for and tested in NTT's DoCoMo DoJa 2.0 API Development Kit for DoCoMo's 504i mobile telephones. Especially in the domain of information appliances the "traditional" result list approach seems cognitively demanding and unsatisfactory. The granular SuperTable in a reduced version seems promising

---

<sup>3</sup> The project is funded by the European Commission under the Fifth Framework of the IST Program, Project No. IST-2000-29640, [www.invisip.de](http://www.invisip.de).



enough to start some serious thoughts about early prototyping and evaluation of possible design scenarios.

After the redesign of both demonstrators is finished we will perform an evaluation to compare the two redesign ideas and to decide which one will have the most benefit for the users. New visualizations will be added to expand the variety of visual aids. Furthermore we plan to support visual query generation and more graphics based filter and sort mechanisms to improve the ease of use.

## 6 References

- [Ahlberg, Shneiderman 1994] Ahlberg, Christopher; Shneiderman, Ben: Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. In: Adelson, B.; Dumais, S.; Olson, J. S. (Eds.): CHI 1994: Conference Proceedings Human Factors in Computing Systems. Conference: Boston, MA, April 24-28 1994. New York (ACM Press) 1994. p. 313-317.
- [Card, Mackinley, Shneiderman, 1999] Card, Stuart; Mackinley, Jock; Shneiderman, Ben (Eds.) (1999). Readings in Information Visualization. Using Vision to Think. San Francisco.
- [Eibl 2000] Eibl Maximilian: Visualisierung im Document Retrieval. Forschungsberichte Band 3, IZ Informationszentrum Sozialwissenschaften, Bonn, 2000
- [Fishkin, Stone 1995] Fishkin, Ken; Stone, Maureen C.: Enhanced Dynamic Queries via Movable Filters. In: Katz, Irvin R.; Mack, Robert L.; Marks, Linn et al. (Eds.): CHI 1995: Conference Proceedings Human Factors in Computing Systems. Conference: Denver, CO, May 7-11 1995. New York (ACM Press) 1995. p. 23-29.
- [Göbel, Haist, Goebel 2002] Göbel, Stefan; Haist, Jörg; Goebel, Christoph: GeoCrystal : Graphic-Interactive Access to Geodata Archives. Proceedings SPIE 2002 – Visualization and Data Analysis, San Jose, CA, 2002.
- [Göbel, Haist, Reiterer, Müller 2002] Göbel, Stefan; Haist, Jörg; Reiterer, Harald; Müller, Frank: INVISIP: Usage of Information Visualization Techniques to Access Geospatial Data Archives. Accepted Paper at the DEXA 2002 Conference, September 2-6, 2002, Aix en Provence, France
- [Hearst 1995] Hearst, Marti A.: TileBars: Visualization of Term Distribution Information in Full Text Information Access. In: Katz, Irvin R.; Mack, Robert L.; Marks, Linn et al. (Eds.): CHI 1995: Conference Proceedings Human Factors in Computing Systems. Conference: Denver, CO, May 7-11 1995. New York (ACM Press) 1995. p. 59-66.
- [Klein et al. 2002] Klein, Peter; Müller, Frank; Reiterer, Harald; Eibl, Maximilian: Visual Information Retrieval with the SuperTable + Scatterplot. In: 6th International Conference on Information Visualization (IV 02). London (IEEE Press) 2002, p.70-75
- [Mann 2002] Mann, Thomas M: Visualization of Search Results from the World Wide Web, PhD Thesis, University of Konstanz, 2002,  
<http://www.ub.uni-konstanz.de/kops/volltexte/2002/751/>



- [Mußler 2002] Mußler, Gabriela: Retrieving Business Information from the WWW. PhD Thesis, University of Konstanz, 2002.
- [Nielsen 1997] Nielsen, Jacob: Search and You May Find. <http://www.useit.com/alertbox/9707b.html> [1999-03-18].
- [Nowell, France, Hix et al. 1996] Nowell, Lucy T.; France, Robert K.; Hix, Deborah et al.: Visualizing Search Results: Some Alternatives to Query-Document Similarity. In: Frei, Hans-Peter; Harman, Donna K.; Schäuble, Peter et al. (Eds.): SIGIR 1996: Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. Conference: Zürich, Switzerland, August 18 -22 1996. New York (ACM Press) 1996. p. 67-75.
- [Pollock, Hockley 1997] Pollock, A.; Hockley, A.: What's Wrong with Internet Searching. D-Lib Magazine, 1997, <http://www.dlib.org/dlib/march97/bt/03pollock.html> [1999-02-01].
- [Rao, Card 1994] Rao, Ramana; Card, Stuart K.: The Table Lens. Merging graphical and symbolic representations in an interactive focus + context visualization for tabular information. In: Adelson, B.; Dumais, S.; Olson, J. S. (Eds.): CHI 1994: Conference Proceedings Human Factors in Computing Systems. Conference: Boston, MA, April 24-28 1994. New York (ACM Press) 1994. p. 318-322.
- [Reiterer et al. 2000] Reiterer Harald; Mußler Gabriela, Mann Thomas, Handschuh Siegfried: INSYDER - An Information Assistant for Business Intelligence, Proceedings of the 23 Annual International ACM SIGIR 2000 Conference on Research and Development in Information Retrieval, ACM press, 2000, pp.112-119
- [Reiterer, Mußler, Mann 2001] Reiterer, Harald; Mußler, Gabriela; Mann, Thomas M.: A visual information seeking system for Web search. In: Oberquelle, Horst; Oppermann, Reinhard; Krause, Jürgen (Eds.): Mensch & Computer 2001. Conference: Bad Honnef, Germany, March 5-8 2001. Stuttgart (Teubner) 2001. S.297-306.
- [Rosson, Carroll 2002] Rosson, Mary Beth; Carroll, John M.: Usability Engineering. Scenario-Based Development of Human-Computer Interaction. San Diego (Academic Press) 2002
- [Spenske, Beilken, Berlage 1996] Spenske, Michael; Beilken, Christian; Berlage, Thomas: FOCUS: The Interactive Table for Product Comparison and Selection. In: UIST 96: 9th ACM Symposium on User Interface Software and Technology New York (ACM Press) 1996. p. 41-50.
- [Spoerri 1994a] Spoerri, Anselm (1994a). InfoCrystal: Integrating Exact and Partial Matching Approaches through Visualization. In: RIAO'94 Proceedings, Oct.11-13, 1994 New York (NY), S.687-696.
- [Veerasamy, Navathe 1995] Veerasamy, Aravindan; Navathe, Shamkant B.: Querying, Navigating and Visualizing a Digital Library Catalog. In: Digital Libraries 1995: The Second Annual Conference on the Theory and Practice of Digital Libraries. Conference: Austin, TX, June 11-13 1995. <http://www.csdl.tamu.edu/DL95/papers/veerasamy/veerasamy.html> [1999-03-24]