MetaCrystal: Visual Interface for Meta Searching

Anselm Spoerri

School of Communication, Information & Library Studies, Rutgers University

4 Huntington Street, New Brunswick, NJ 08901, USA

aspoerri@scils.rutgers.edu

ABSTRACT

MetaCrystal visualizes the degree of overlap between the top results returned by different search engines. Linked overview tools support rapid exploration, facilitate advanced filtering operations and guide users toward relevant information. The direct manipulation interface enables users to iteratively compose and edit meta searches. MetaCrystal addresses the problem of the effective fusion of different search engine results by helping users control and gain insight into how to combine and filter them.

Categories and Subject Descriptor

H.3.3 Information Storage and Retrieval: Information Search and Retrieval - information filtering. H.5.2 Information Interfaces and Presentation (e.g., HCI): User Interfaces - graphical user interfaces (GUI).

INTRODUCTION

Users can choose from a multitude of Internet search engines, which tend to return different results for the same query [5]. Meta search engines address this limitation by combining the results from different engines. The automatic and effective fusion of different search engine results can be difficult [1]. While meta search interfaces exist that visually organize the retrieved documents [3, 4], none of them provide users with an overview of the precise overlap between the engines. Meta searching could benefit from such a visual interface, because: a) it is difficult to predict the quality of coverage for single search engines, which tend to index less than 20% of the Internet [5]; b) the overlap between different engines is difficult to predict; c) some engines are more effective than others, depending on the search domain; d) users may prefer or trust some engines more than others. This suggests that active user involvement can make a difference when it comes to deciding how to combine search engine results. While items found by multiple engines tend be more relevant [2], users also want to examine the items only found by their most trusted engines. MetaCrystal can address this range of requirements.

METACRYSTAL

Users can iteratively compose and edit meta searches that visualize the precise overlap between up to five search engines. Implemented in Flash, its linked overview tools support rapid exploration, enable advanced filtering operations and guide users toward relevant information. The *Category View* displays the number of items retrieved by different search engine combinations. Modeled on the InfoCrystal layout [6], the number of items retrieved by all engines is shown at the center. Shape (size), color, orientation and proximity coding are used to visually organize the different search engine combinations. The *Cluster Bulls-Eye View* displays *all* retrieved items; items with similar ranking scores for the different engines are placed in close proximity and items found by multiple engines cluster toward the center. This view helps users identify items retrieved by specific engine combinations and at the same time scan the top items found by a single engine. Shape, color and orientation coding show which engines retrieved an item. Search engines commonly display their results as ranked lists, which can only show a limited number of items. The *RankSpiral View* overcomes this limitation by placing *all* items sequentially along a spiral based on their total ranking scores. MetaCrystal enables users to perform complex filtering operations visually and interact with a *filtered results* view. "Details on demand" provides users with an immediate sense of an item's content and how the different engines contributed to its total ranking. Users can apply different weights to the engines to create their own ranking functions that reflect their engine preferences. MetaCrystal and its overview tools help users control and gain insight into how to combine and filter the top items retrieved by different search engines. The next step is to conduct a formal evaluation of MetaCrystal.

REFERENCES

1. Callan. J. (2000). Distributed information retrieval. In Croft W.B. (Ed.), Advances in Information Retrieval. (pp. 127-150). Kluwer Academic Publishers

2. Foltz, P. and Dumais, St. (1992) Personalized information delivery: An analysis of information-filtering methods. Communications of the ACM, 35, 12:51-60.

3. Grokker – www.groxis.com

4. Kartoo – www.kartoo.com

5. Lawrence, S., & Giles, C.L. (1999). Accessibility of information on the Web. Nature, 400, 107–109.

6. Spoerri, A. (1993) InfoCrystal: A Visual Tool for Information Retrieval. Proc. IEEE Visualization. 150-157.

MetaCrystal: Visual Interface for Meta Searching

Anselm Spoerri, SCILS, Rutgers University, aspoerri@scils.rutgers.edu

Search

information visualization — **Found**
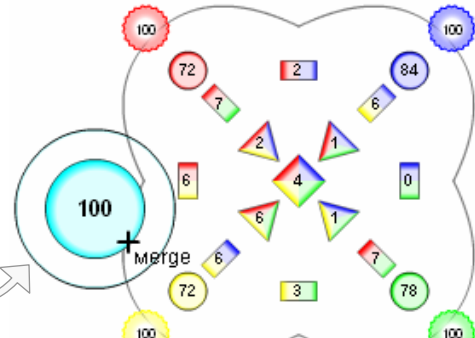
Google	100
Teoma	100
AltaVista	100
Lycos	100
MSN	100

Search New

Views

Category View groups and shows the number of documents retrieved by different combinations of search engines. Shape (size), color, orientation and proximity coding are used to visualize the different combinations.
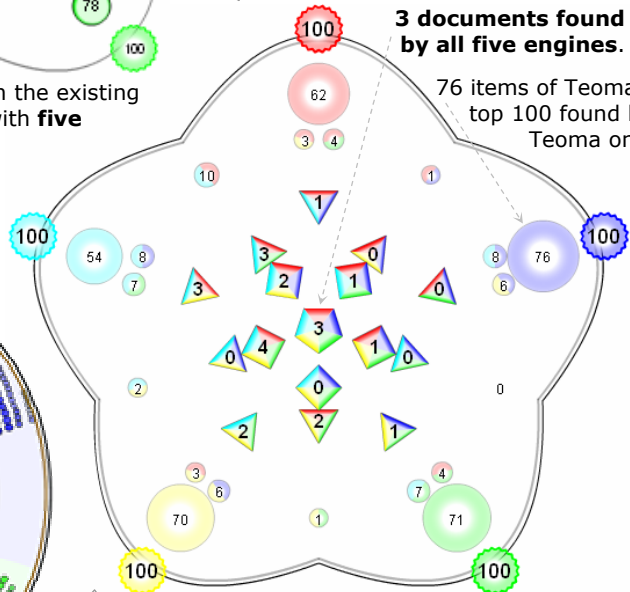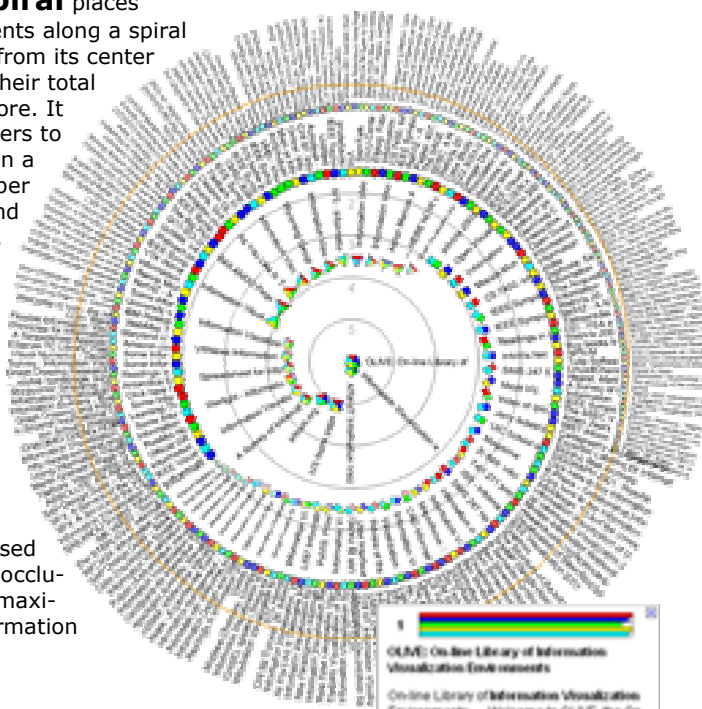
Drag MSN top 100 results & **merge** with the existing crystal to create a new one with **five intersecting result sets**.

3 documents found by all five engines.

76 items of Teoma's top 100 found by Teoma only

Cluster Bulls-Eye shows *all* documents. Those found by multiple engines cluster toward the center and users can scan top documents found by single engine. Shape, color and orientation coding show how many and which engines retrieved a document. Documents found by the same number of engines cluster in the same ring. The total ranking score increases toward the center.

Filtering applied to select, for example, only documents retrieved by **at least three engines**. Size coding used for category icons not selected.
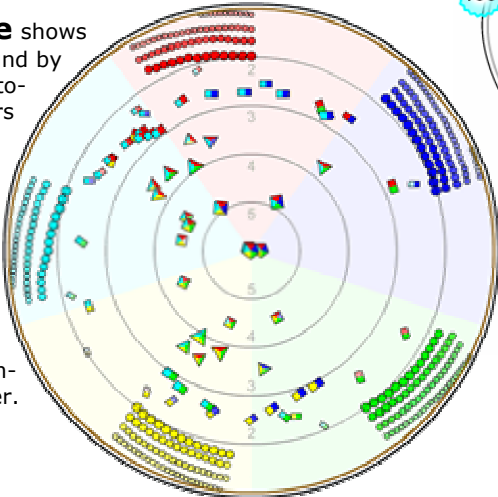
RankSpiral places *all* documents along a spiral and away from its center based on their total ranking score. It enables users to rapidly scan a large number of items and their titles. This tool makes it easy to identify the top items found by a specific number of engines. The spiral mapping used minimizes occlusions and maximizes information density.

Filtered Results

1. OLIVE: On-line Library of **Information**
2. **Information Visualization** at PNNL
3. AT and T **Information Visualization**
4. IVS home page
5. infovis.org
6. A Survey of **Information Visualization**
7. **Information Visualization** via Hyperbolic

Filtered Results View shows how the different search engines contributed to the total ranking score of the selected items.

MetaCrystal toolset helps users **visualize** the degree of **overlap** and **similarity** between the top results returned by different search engines. It addresses the problem of the **effective fusion** of different search engine results by enabling users to **visually compose** and refine **meta searches**. Users can perform **advanced filtering** operations visually and apply different weights to the search engines to create their own ranking functions. A hierarchy of aggregation is employed to provide users with **quick insights** into how to combine and filter the top results by different search engines.

Details-on-Demand gives users an immediate sense of a document's content and ranking scores.

OLIVE: On-line Library of Information Visualization Environments

On-line Library of **Information Visualization** Environments! ... Welcome to OLIVE, the On-line Library of **Information Visualization** Environments! ... Description: Includes Temporal, 1-D, 2-D, 3-D, Multi-D, Tree, Network, and workspace environments. Many links...

 cital.umd.edu/Olive