

MetaCrystal

Visualizing Authority and Ranking Effects

Anselm Spoerri

SCILS, Rutgers University



MetaCrystal

Visualize Overlap between Search Results

Documents Found by Multiple Engines
tend to be more likely to be Relevant

Saracevic & Kantor ('88) | Foltz & Dumais ('92) | Belkin et al. ('93)

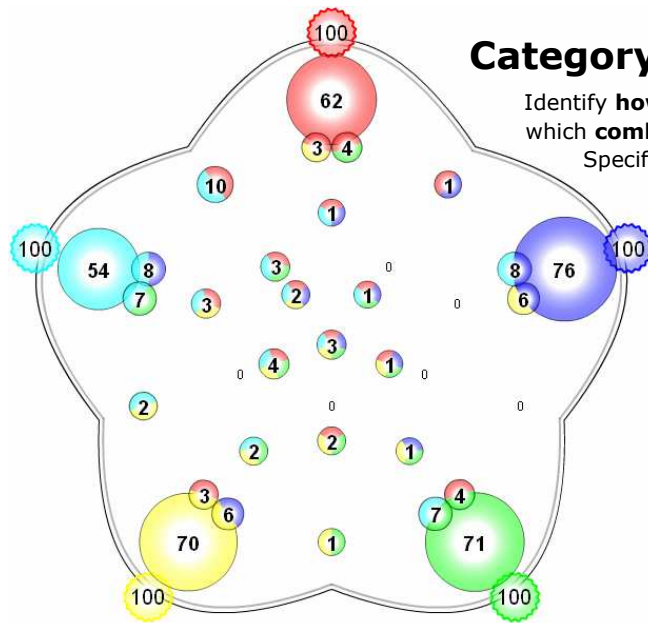
By Exploring the Data using MetaCrystal, will
Users be lead to Relevant Documents?

What is the **Signal** MetaCrystal makes visible?

How does Overlap between Results
Actually Correlate with Relevance?

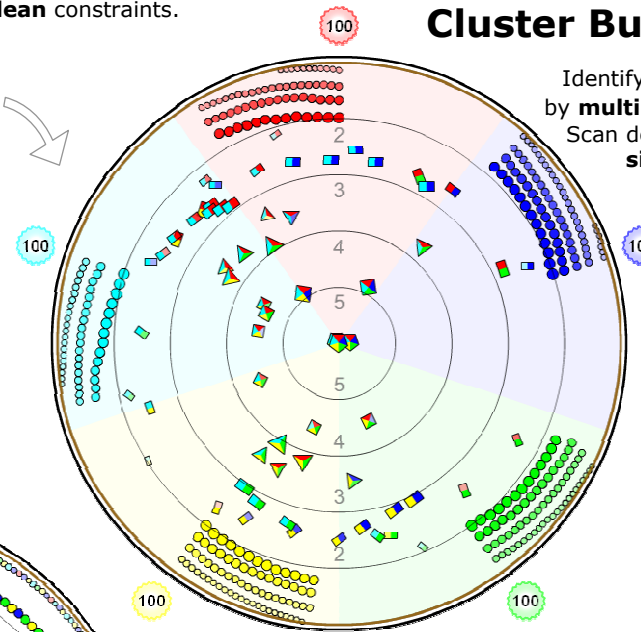


MetaCrystal Toolset



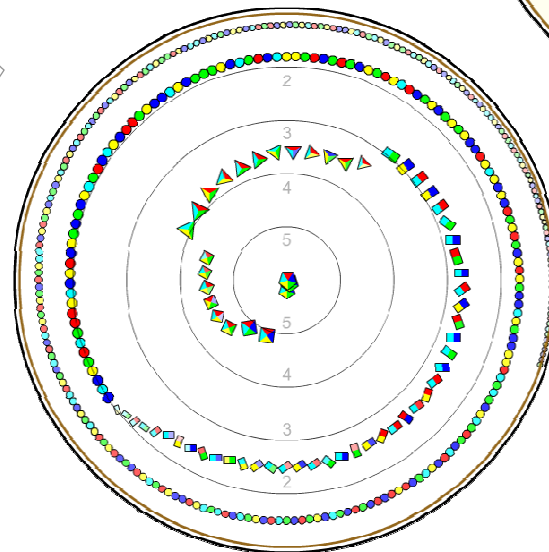
Category View

Identify **how many** docs found by which **combinations of engines**;
Specify **Boolean** constraints.



Cluster Bulls-Eye

Identify **docs** found by **multiple engines**;
Scan docs found by **single engine**



RankSpiral

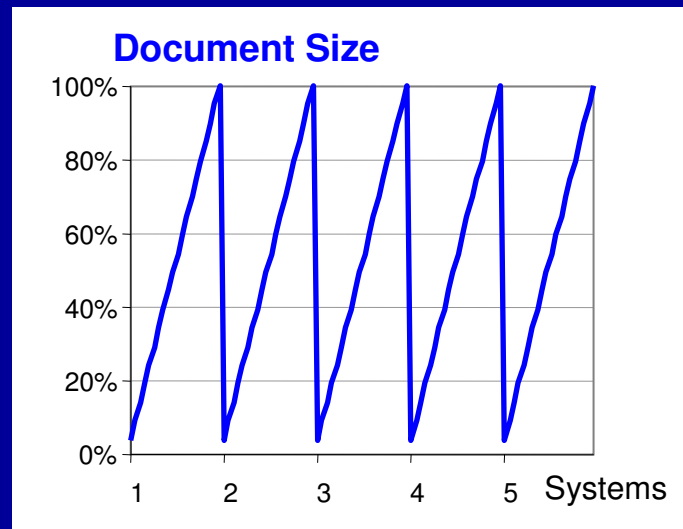
Identify **top docs** found by a **specific number** of engines;
Rapidly scan many docs titles.



MetaCrystal

Key Design Principles

- ⌘ **Map** Docs found by **Multiple Engines Toward Center**
- ⌘ **Map** Docs found by **Same Number of Engines** into the **Same Concentric Ring**
- ⌘ **Size** = Average of Rankings by Different Engines



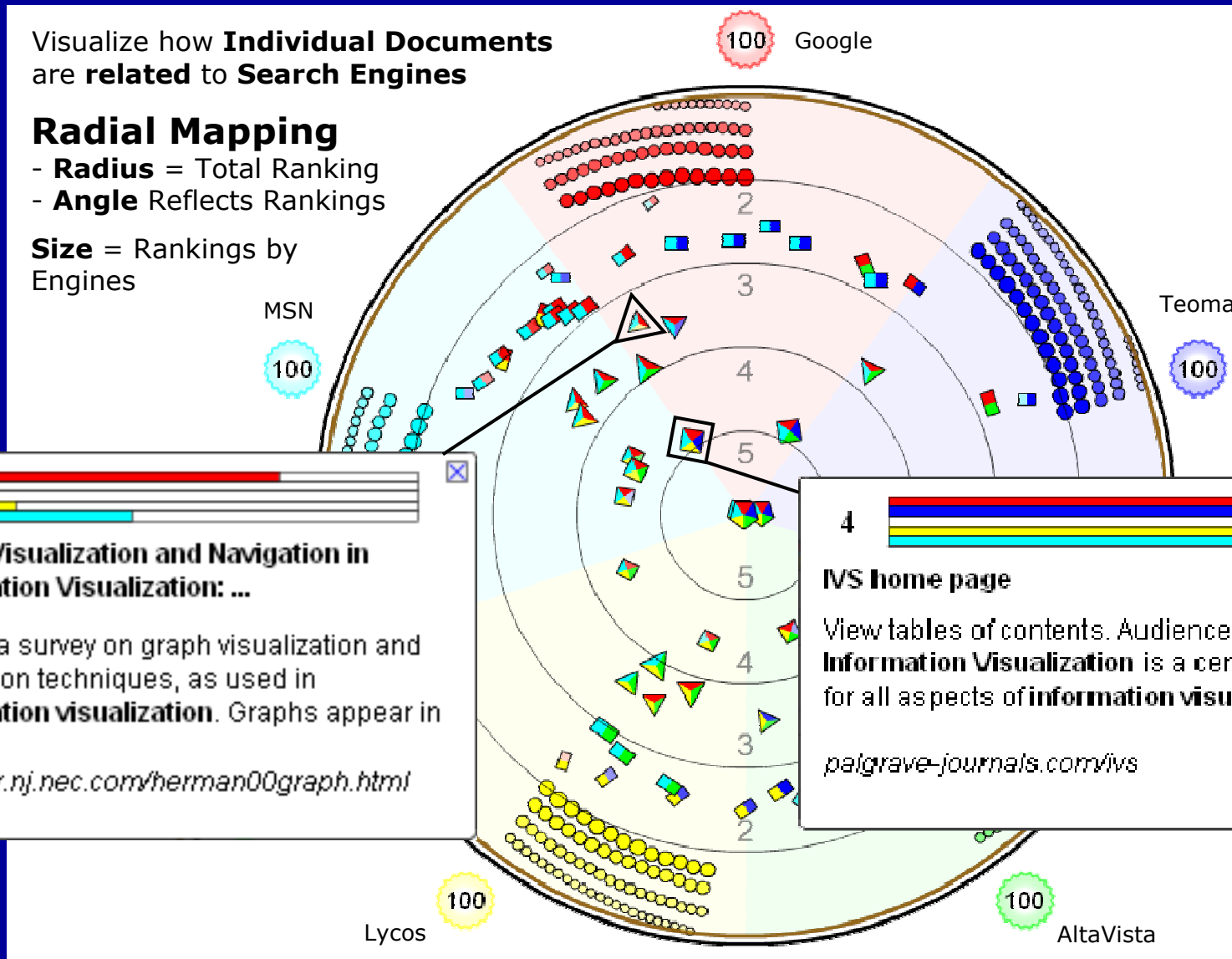
MetaCrystal – Cluster Bulls-Eye

Visualize how **Individual Documents** are **related** to **Search Engines**

Radial Mapping

- **Radius** = Total Ranking
- **Angle** Reflects Rankings

Size = Rankings by Engines



23

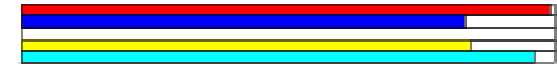


Graph Visualization and Navigation in Information Visualization: ...

This is a survey on graph visualization and navigation techniques, as used in **information visualization**. Graphs appear in

citeseer.nj.nec.com/herman00graph.html

4



IVS home page

View tables of contents. Audience.
Information Visualization is a central forum for all aspects of **information visualization**

paigrave-journals.com/ivs



Overlap between Search Results Correlated with Relevance?

Major **Data Fusion** methods use both **Voting** and **Merging** and makes use of these phenomena:

- **Authority Effect**: the more systems that find a document, the greater the probability of its relevance
- **Ranking Effect**: the higher up document in ranked list and the more systems that find it, the greater the probability of its relevance

No Systematic Analysis of Structure of Overlap for Large Text Databases as used in TREC



Overlap between Search Results Correlated with Relevance?

Method

- Use **Ad-hoc track** data for **TREC 3, 6, 7, 8**
- Systems search the **SAME Database**
- **Automatic Short Runs**
- **50 Topics** and **1,000 Documents per topic** \approx 50,000 documents
- Retrieval systems can submit multiple runs
- \approx **Select Best Run** based **Mean Average Precision**

TREC 3 **19** systems **928,709** documents found

TREC 6 **24** systems **1,192,557** documents found

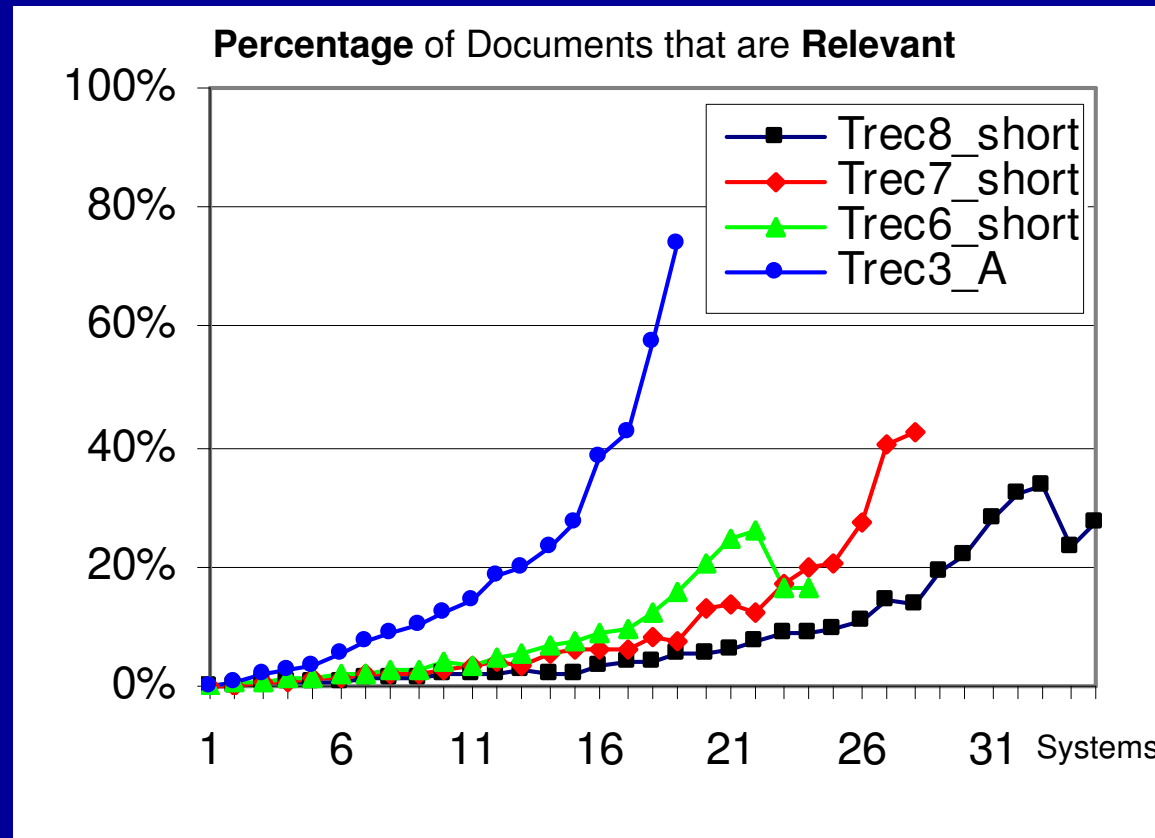
TREC 7 **28** systems **1,327,166** documents found

TREC 8 **35** systems **1,723,929** documents found

- **Compute Average** by summing over all 50 topics and divide by 50



How does Overlap Correlate with Relevance?

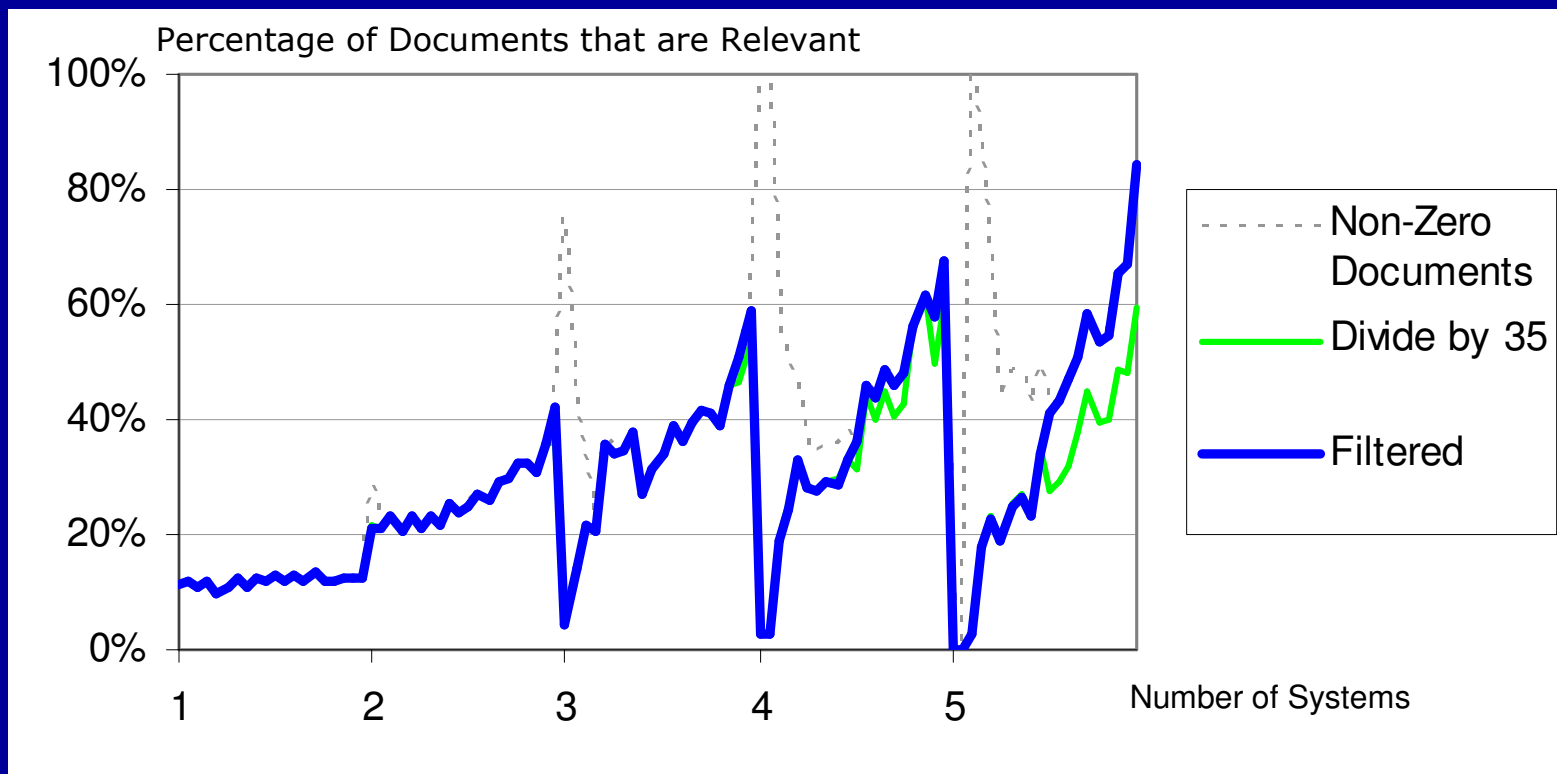


To Appear in ASIST 2005

⌘ Authority Effect



TREC 8 – Impact of Average Rank Position?



Submitted to InfoVis 2005

Compute overlap structure between top 50 search results of 35 random groupings of 5 retrieval systems for 50 topics.

Ranking Effect



How does Overlap between Search Results Correlate with Relevance?

Authority Effect – the more systems that find a document, the greater the probability that it is relevant

Ranking Effect – the higher up a document in a ranked list and the more systems that find it, the greater the probability of its relevance

- ⌘ Validates MetaCrystal's Design Approach
- ⌘ MetaCrystal Visualizes Authority & Ranking Effects
- ⌘ MetaCrystal can Guide User's Exploration Toward Relevant Documents

