

# Interactive Poster: Name That Cluster – Text vs. Graphics

James Abello†

Hans-Jörg Schulz‡

Benoit Gaudin+

Christian Tominski‡

DIMACS, Rutgers University

University of Rostock

University College Dublin

University of Rostock

## ABSTRACT

Given a user query, search engines generally return a very sizeable collection of possible answers. Clustering has been proposed as a tool to partition the possible answer set into more manageable subsets of related results. There is no current agreement on the preferred mode of presentation of these clusters. Currently, most search engines display the set of results in an almost purely textual form. However, relatively recently we have witnessed some timid attempts to use some graphical representations. To elucidate when and why text appears to outperform graphics for certain fundamental clustering related tasks, this work presents a preliminary user study with three interfaces to display flat clusters of user queries.

**Keywords:** applications of information visualization, graph visualization, user studies.

## 1. THE SETUP

When deciding on how to present a partition of clustered search engine results to the user, there are basically three general choices: a pure textual interface or a graphical interface – or, since each of these two choices has its Pros and Cons, maybe even a combination of both with some graphical and some textual elements. The ongoing user study [2] is an attempt to provide some directions to support this design decision. Other works like [4] aim to incorporate user feedback to help determine the criteria influencing clustering quality, but our emphasis is not on any clustering technique but rather on the presentation mode and associated interaction methods. In that respect, our work is more related to [2] since it considers graphical techniques to present textual clustering. However, our study also includes pure textual representation and clustering related tasks.

Since we wanted to mimic as close as possible the scenario of a user’s query web search without encumbering the study with web page content descriptions or extra URL clicks we settled for presenting the user with sets of related web search queries instead of web page descriptions. Namely, after choosing an input query from a scrolling query list the user’s task is to *Select*, *Explore*, and finally *Grade and Name* a set of query clusters offered as related to the input query.

In the course of the study, we collect the user’s response as well as the time used to explore, rate and name the clusters. This allows us to correlate the different interface types with the measured performance, cluster sizes and user responses. An interface could then be considered better, when it outperforms another interface in terms of completion time or completeness of the answers (less unnamed or ungraded clusters).

## 2. INTERFACE DESCRIPTIONS

Even though, different approaches and techniques have been used to implement the different interfaces, close attention was paid to make the interfaces as functionally similar as possible regarding the used color schemes or the evaluation.

† abello@dimacs.rutgers.edu

+ benoit.gaudin@ucd.ie

‡ {hjschulz, ct}@informatik.uni-rostock.de

The **textual interface** (Figure 1) presents clusters of queries as a scrollable list. It is implemented as a DHTML/JavaScript web page. The clustered queries associated with an input query are displayed in rows on the screen, together with the form fields needed to evaluate each cluster.

The **graphical interface** (Figure 2) represents queries as round, marble-like items scattered across a 2-dimensional canvas (call them graphical query-items). Elements of the same cluster are positioned close to each other and a polygonal frame is drawn around each cluster. The graphical query-items are annotated with textual labels using a level-of-detail approach that shows more labels depending on the level of zooming being applied to a cluster. This interface has been implemented in JAVA/OpenGL and it reuses some of the mechanisms offered by the CGV-platform [1]. The described graphical setup is in our view a logical extension of the 1-dimensional, list-like display used in the textual interface.

We also offer a third **hybrid interface** (Figure 3) that incorporates in our view the “best” of the textual and the graphical interfaces. Concretely, the clusters are now presented both in textual and graphical form. These views are linked with each other in a coordinated way using a Model-View-Controller Pattern offered in the CGV platform. Thus, navigational changes from inter- or intra-cluster exploration are reflected by both views in a coordinated way. This way, the interface indicates both: the current position of the item the user is currently viewing in the textual representation, and corresponding visual information on the cluster appearing in the graphical component. For the latter, the density of a cluster is visible at a glance.

## 3. HOW TO PARTICIPATE

There are two ways to help our effort to compare the described interfaces:

- There will be an opportunity to test the three interfaces by joining our user study at the poster presentation during the InfoVis conference.
- We are now working on extensions of our experiment set up to a web-based platform. This will enable us to reach out to a larger and broader set of participants. The URL of this web user study will be made available at the InfoVis conference.

Since the web experiment is expected to draw a large number of participants, we plan to apply classical statistical analysis to this “vox populi”. Such a statistical approach applied to results obtained from a large number of participants could make our results robust to “unusual behaviors” of just a few participants.

## REFERENCES

- [1] J. Abello et al, CGV – Coordinated Graph Visualization, Interactive Poster, InfoVis 2007
- [2] M. Ghoniem et al, A Comparison of the Readability of Graphs Using Node-Link and Matrix-Based Representations, InfoVis 2004
- [3] J.E. Swan II, Experimental Design and Analysis for Human-Subject Visualization Experiments, IEEE InfoVis 2006 Tutorial.
- [4] C. Ware, et al, Cognitive measurements of graph aesthetics, Information Visualization, Vol.1 No.2, 103-110, 2002.

Step 1	Step 2: inspect the contents of the group after having clicked the "Select This Group!" button.	Step 3: answer part A and B then click the "I'm finished evaluating this group!" button.
Select This Group!	<pre> * Acacia+Chemical+Formula * What+is+Acacia+Used+in+%3f * chemical+formula+for+acacia * acacia+fiber What+is+the+Ingredient+Acacia+%3f what+is+the+chemical+formula+for+acacia acacia+powder what+is+gum+acacia define+acacia acacia+nilotica </pre>	<p>Part A: how would you describe the relevance of this group according to the phrase selected in Step 0?</p> <p><input type="radio"/> fantastic <input type="radio"/> good <input checked="" type="radio"/> o.k. <input type="radio"/> poor <input type="radio"/> disaster</p> <p><input type="checkbox"/> Too hard to rate.</p> <p>Part B: fill in keywords that you think best describe the contents of the group.</p> <p>acacia gum</p> <p><input type="checkbox"/> Too hard to describe.</p> <p style="text-align: center;">I'm finished evaluating this group!</p>

Figure 1: The textual interface.

STEP 1 and 2

acacia+furniture  
acacia+wood+furniture  
acacia+wood  
pictures+of+acacia  
where+does+acacia+wood+come+from+%3f  
what+is+acacia+wood  
definition+of+acacia  
what+does+acacia+wood+look+like

acacia+trees  
acacia+tree  
acacia  
eucalyptus  
acacia+species  
acacia+fraternity

STEP 3

Part A: how would you describe the relevance of this group according to the phrase selected in Step 0?

fantastic  good  o.k.  poor  disaster

Too hard to rate.

Part B: fill in keywords that you think best describe the contents of the group.

acacia wood

Too hard to describe.

I'm finished evaluating this group!

Figure 2: The graphical interface with one cluster already evaluated (gray) and one currently being in the process of evaluation (yellow).

STEP 1 and 2

- acacia+plant
- acacia+flowers
- acacia+plant+information
- acacia+group
- pictures+of+plant+acacia
- acacia+plants
- australian+acacia
- what+is+a+acacia
- acacia+africa
- acacia+pycnantha

11 queries:

- acacia+plant
- wattle
- acacia+flowers
- acacia+plant+information
- acacia+group
- pictures+of+plant+acacia
- acacia+plants
- australian+acacia
- what+is+a+acacia
- acacia+africa
- acacia+pycnantha

14 queries:

- gum+acacia
- acacia+gum
- malindi+acacia
- acacia+chemical+formula
- what+is+acacia+used+in+%3f
- acacia+fraternity
- acacia+nilotica
- acacia+fraternity

giraffe

acacia+farnesiana

acacia+plant

acacia+wood

acacia+tree

baobab

acacia+flower

acacia+dealbata

Figure 3: Hybrid interface showing textual and graphical components. The flat cluster list is visible on the left; the graphical representation is at the right. Both are linked, so that the selection of an item in either one of them is reflected in both views – the highlighted item in the tree view corresponds to the highlighted item in the graphical cluster. User evaluations are performed through an interface identical to the one presented in Figure 2.