

# NanoPort: A Web Portal for Nanoscale Science and Technology

Michael Chau, Hsinchun Chen, Jialun Qin, Yilu Zhou, Wai-Ki Sung, Yongchi Chen,  
Yi Qin, Daniel McDonald, Ann Lally, Matthew Landon

Artificial Intelligence Lab

Department of Management Information Systems

The University of Arizona

Tucson, Arizona 85721, USA

1-520-621-2748

{mchau, hchen, qin, yilu, wai-ki}@bpa.arizona.edu, mark@ai.bpa.arizona.edu,

{yiqin, dmm, alally}@bpa.arizona.edu, mlandon@u.arizona.edu

## 1 INTRODUCTION

Areas related to nanotechnology, or nanoscale science and engineering (NSSE), have experienced tremendous growth over the past few years. While there are a large variety of useful resources available on the Web, such information are usually distributed and difficult to locate, resulting in the problem of information overload. To address the problem, we developed the NanoPort system, an integrated Web portal aiming to provide a one-stop shopping service to satisfy the information needs of researchers and practitioners in the field of NSSE [1]. We believe that the approaches taken also can be applied to other domains.

## 2 NANOPORT SYSTEM FEATURES

We will demonstrate the NanoPort system prototype, both the system as a whole as well as the individual components. The NanoPort system has the following features:

### Vertical Searching

As NSSE is a multi-disciplinary domain, relevant Web pages are distributed over the Web. It is difficult to identify from the Web a set of high-quality, relevant pages. We will demonstrate the system's capabilities in searching Web pages only in the domain, as well as the backend algorithms that collect and filter NSSE-related pages.

### Meta-searching

In addition to the self-maintained search index, NanoPort connects with different search engines and combines the search results. These search engines include general-purpose search engines (e.g., Google and AltaVista), vertical search engines (e.g., nanotechnology.com), academic search engines (e.g., Scirus, MedLine), and news search engines (e.g., ABCNews, USAToday).

### Noun Phrasing

In a traditional ranked list display, a user has to read the search results one by one to identify the desired pages. This can be a very time-consuming and mentally exhausting process. The NanoPort system employs the Arizona Noun Phraser (AZNP) to extract the key phrases from the search results in a session [3]. These phrases are used to create a

folder-structure that categorizes search results based on key concepts, making it much easier to understand the results.

### Self-organizing Topic Map

In addition to the folder structure, a user can also choose to generate a 2-dimensional topic map based on the self-organizing map (SOM) algorithm [2]. The automatically generated map contains of a few regions, each representing a key concept in the set of documents. More important concepts occupy larger regions, and similar concepts are grouped in a neighborhood. The documents are categorized into different regions for easier concept identification.

### Automatic Summarization

Another component of NanoPort is the AI Summarizer, which can summarize any Web page from the search result into a few sentences. Utilizing both segmentation and summarization methods, the AI Summarizer tries to extract the most important and representative sentences from a document. Such summarization can help the user decide quickly whether a page is interesting or not.

## 3 TECHNICAL SPECIFICATION

The system was mainly implemented in Java, Java Servlet, and JSP. The AZNP and SOM were implemented in C.

## 4 ACKNOWLEDGMENTS

This research is supported in part by the National Science Foundation under grant numbers CTS-0204375, IIS-9800696, and IIS-9817473.

## REFERENCES

- [1] Chau, M., Chen, H., Qin, J., Zhou, Y., Qin, Y., Sung, W. K., McDonald, D. Comparison of Two Approaches to Building a Vertical Search Tool: A Case Study in the Nanotechnology Domain. In Proceedings of JCDL'02, Portland, OR, USA, July 2002.
- [2] Orwig, R., Chen, H., and Nunamaker, J. A Graphical, Self-Organizing Approach to Classifying Electronic Meeting Output, JASIS, 48(2) (1997), 157-170.
- [3] Tolle, K. M. and Chen, H. Comparing Noun Phrasing Techniques for Use with Medical Digital Library Tools. JASIS, 51(4) (2000), 352-370.