

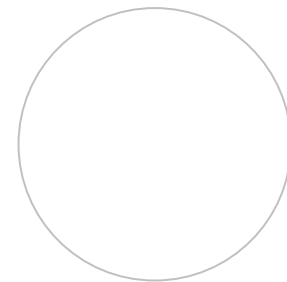
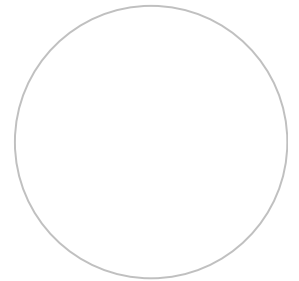
# IVEA: An Information Visualization Tool for Personalized Exploratory Document Collection Analysis

VinhTuan Thai, Siegfried Handschuh, Stefan Decker

# Outline



- Introduction
- IVEA- Visual Interface & Interactions
- Architecture
- Usability Study
- Recent Update & Future Work
- Related Work
- Conclusions



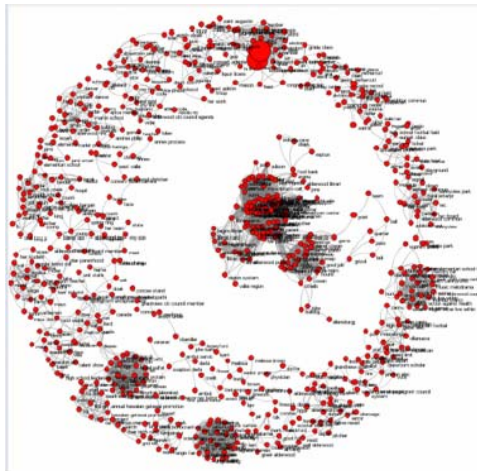
## ■ Motivation

- Knowledge work in many fields requires examining aspects of a document collection to attain meaningful understanding.
- a.k.a exploratory data analysis / information analysis
- Commonly carried out in various fields, e.g. science, intelligence and defense, or business
- Purpose:
  - to understand the distribution of topics
  - to identify trends and linkages between different entities

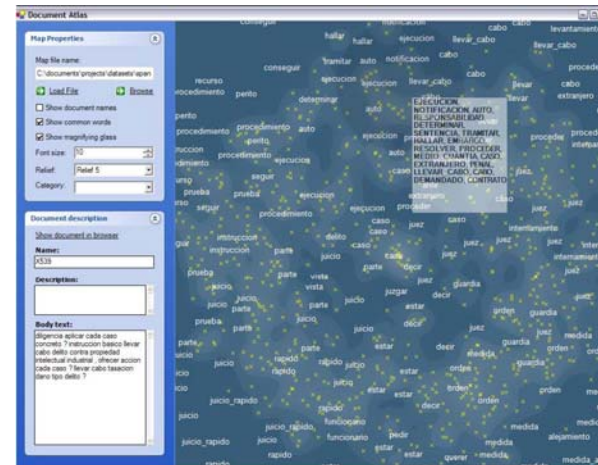
# Introduction



- Existing document corpus visualization tools
  - visualize linkages between main entities (e.g. topics, people, locations)
  - visualize clusters of documents



Storyline

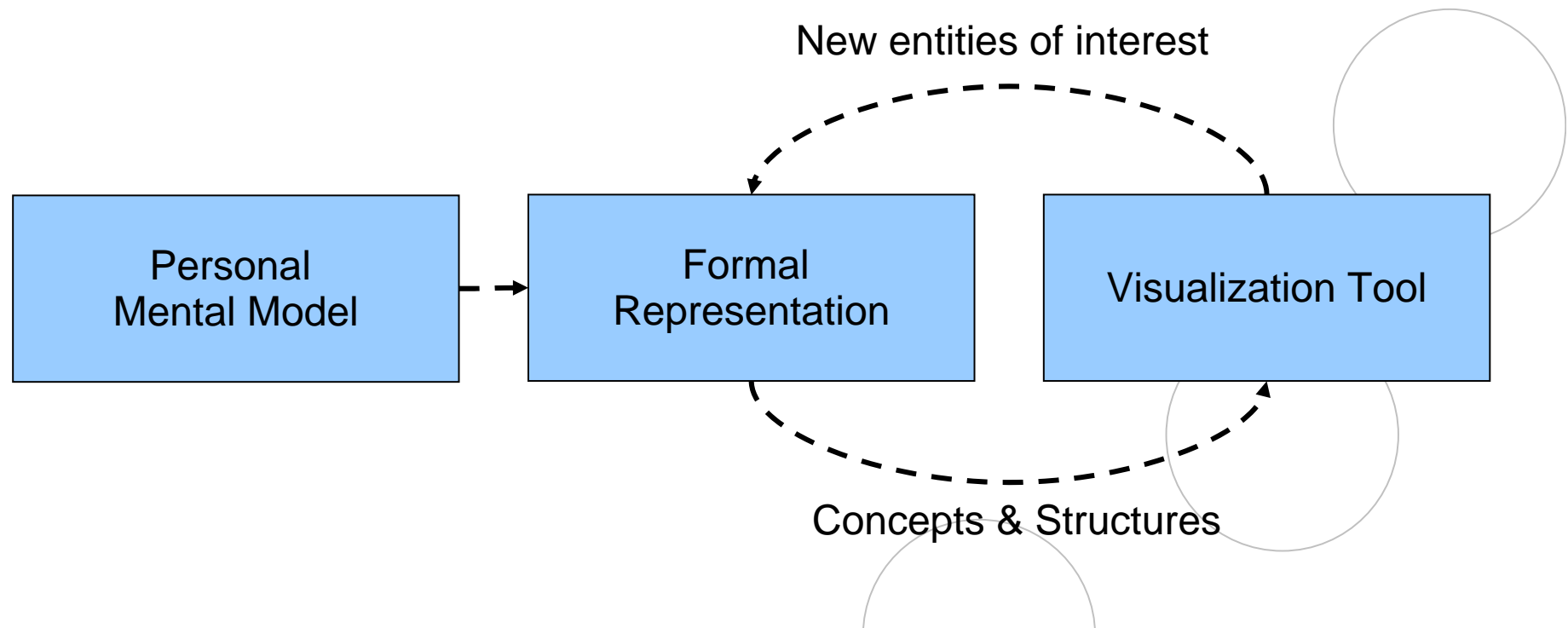


Text Garden

- findings presented are independent of the users' interests

## ■ Goal

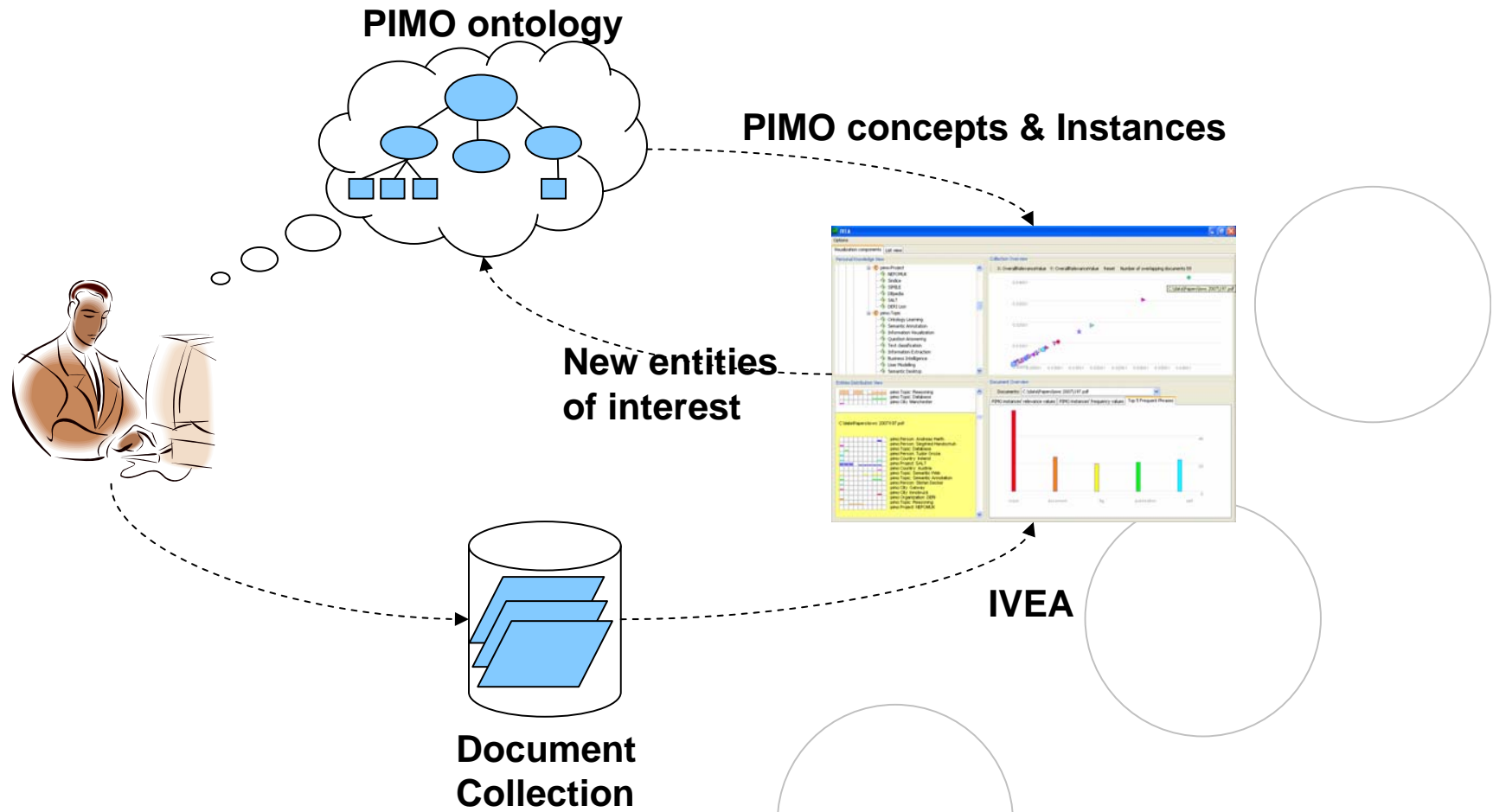
- An approach toward **personalized visual exploration** of a text collection



# Introduction



## Proposed approach



## ■ Proposed approach

### □ Advantages:

- aligned with the **users' interests**
- **user-controlled**
- flexibility to explore at **different levels of detail**
- dynamic **ontology enrichment**
  - **better representation** of the users' interest
  - **more personalized** to the users
  - benefits **other PIMO-based applications**

## ■ Design

### □ Factors to be taken into consideration:

- the nature of the **task**
- the type of **data**
- the target **users**
- their typically available **work environment**

### □ Shneiderman's visual information-seeking mantra

- “Overview first, zoom and filter, then details-on-demand”



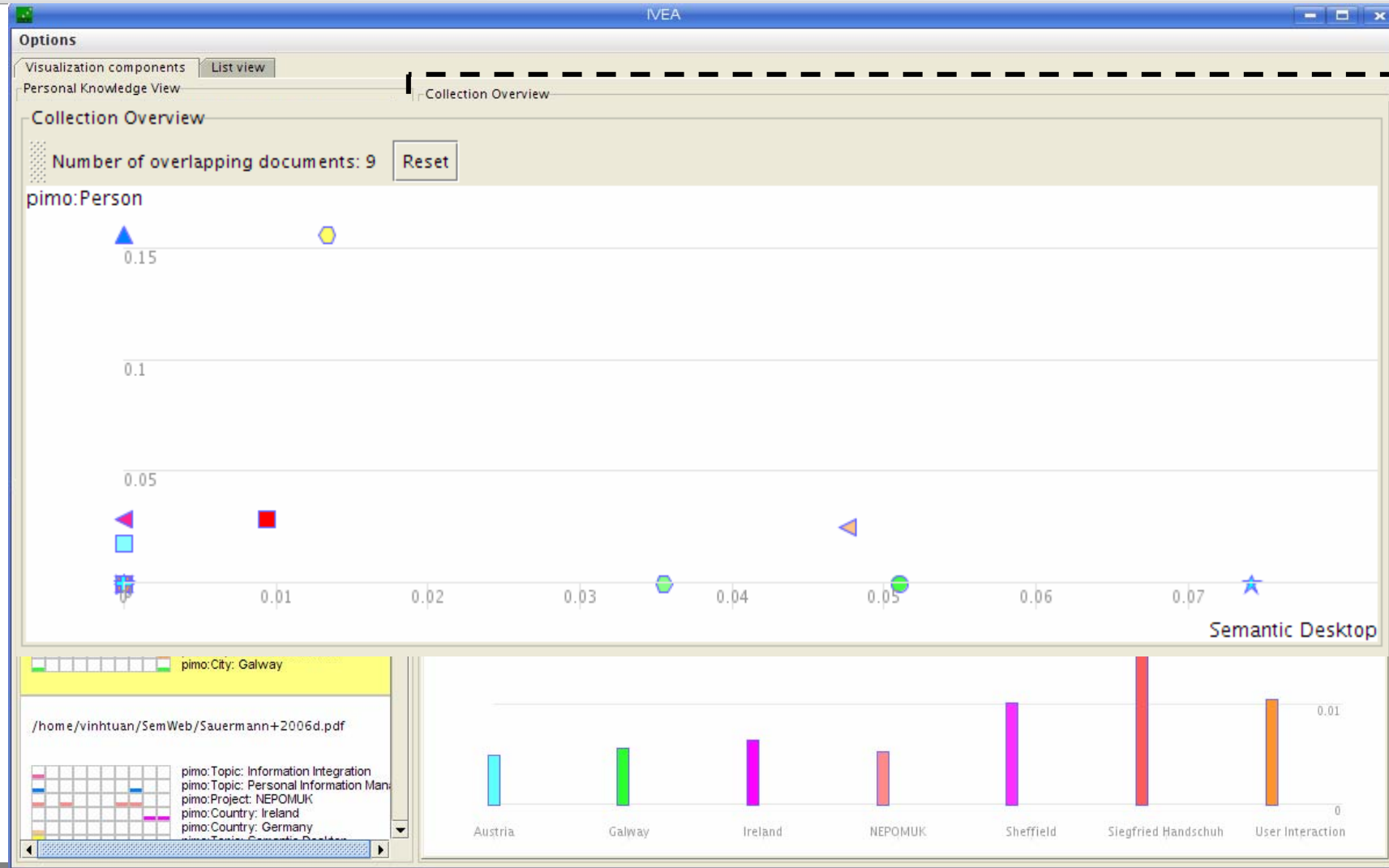
# IVEA- Visual Interface



The screenshot displays the IVEA- Visual Interface with several overlapping windows:

- Entities Distribution View:** Shows a grid visualization and a list of entities for the document `/home/vinh Tuan/SemWeb/795.pdf`. The entities are: `pimo:City: Sheffield`, `pimo:Person: Siegfried Handschuh`, `pimo:Project: NEPOMUK`, `pimo:Country: Ireland`, `pimo:Country: Austria`, and `pimo:Topic: User Interaction`.
- Document View:** Shows a list of documents and a bar chart for 'conference' with a red bar.
- Options:** A panel with 'Visualization components' and 'Personal Knowledge View'.
- Docum:** A window showing a bar chart for 'semantic web' with a cyan bar.
- Fragment Detail:** A blue box showing details for 'Fragment number: 1' and 'Frequency: 1'. The contents are: **Recipes for Semantic Web Dog Food – The ESWC and ISWC Metadata Projects** Knud Møller<sup>1</sup>, Tom Heath, Siegfried Handschuh, and John Domingue<sup>1</sup> Digital Enterprise Research Institute, National University of Ireland, Galway knud.moeller@deri.org, siegfried.handschuh@deri.org<sup>2</sup> Knowledge Media Institute and Centre for Research in Computing, The Open University, Milton Keynes, UK {t.heath, j.b.domingue}@open.ac.uk Abstract. ...

# IVEA- Interactions



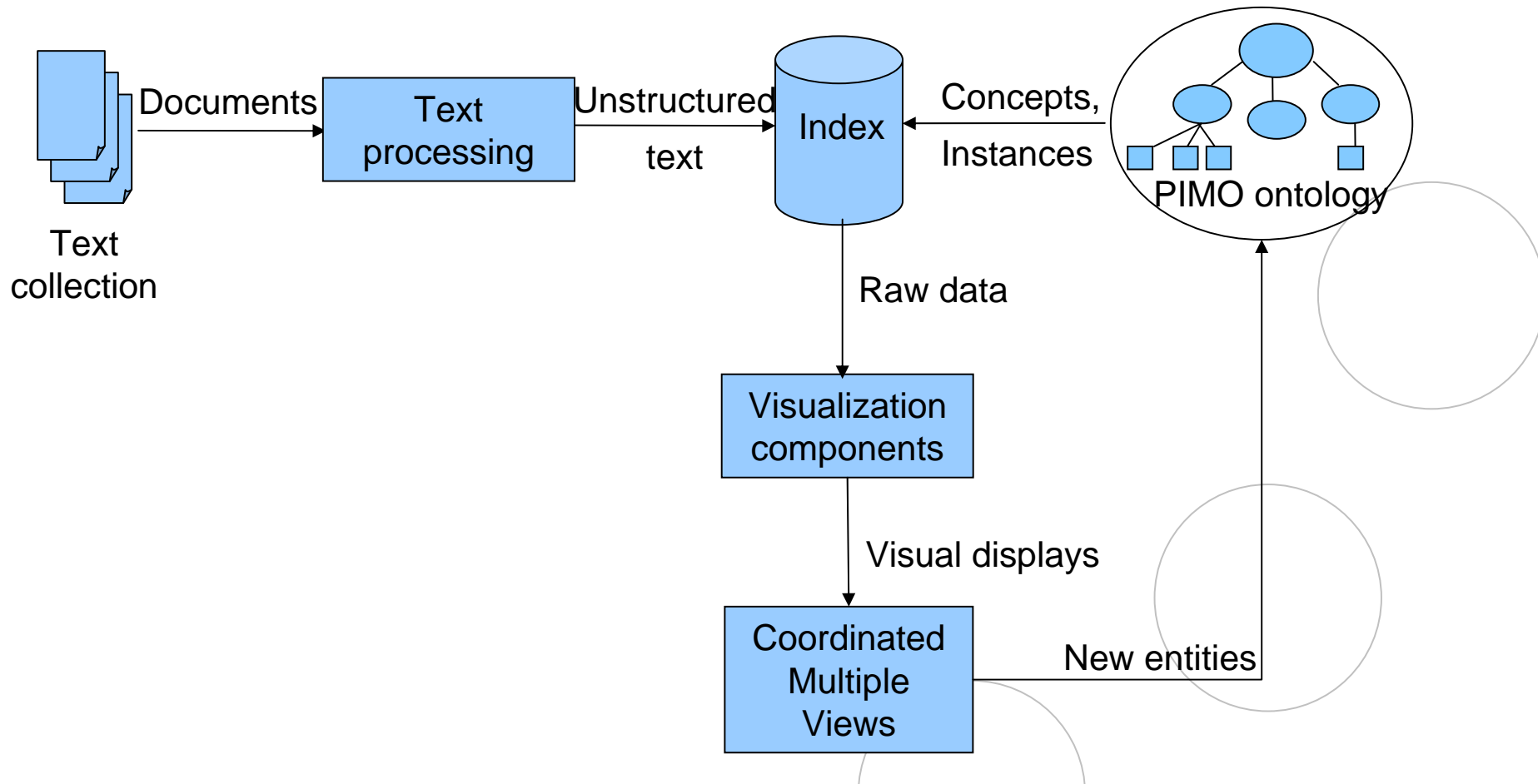
# IVEA- Interactions



The screenshot displays the IVEA application interface with several key sections:

- Options:** Includes a 'Visualization components' tab and a 'List view' button.
- Personal Knowledge View:** A hierarchical tree structure showing ontologies and concepts. A red dashed arrow points to 'pimo:conference'.
- Collection Overview:** Shows 'Number of overlapping documents: 9' with a 'Reset' button. Below is a scatter plot for 'pimo:Person' with various data points.
- Entities Distribution View:** Displays two document-specific entity distributions:
  - Document 1: /home/vinh Tuan/SemWeb/795.pdf. Entities include pimo:City: Sheffield, pimo:Person: Siegfried Handschuh, pimo:Project: NEPOMUK, pimo:Country: Ireland, pimo:Country: Austria, pimo:Topic: User Interaction, and pimo:City: Galway.
  - Document 2: /home/vinh Tuan/SemWeb/Sauermann+2006d.pdf. Entities include pimo:Topic: Information Integration, pimo:Topic: Personal Information Man, pimo:Project: NEPOMUK, pimo:Country: Ireland, pimo:Country: Germany, and pimo:Topic: Semantic Desktop.
- Document View:** Shows 'Documents: /home/vinh Tuan/SemWeb/795.pdf'. It contains three sub-views:
  - 'PIMO instances' relevance values BarChart' (not fully visible).
  - 'PIMO instances' frequency values BarChart' showing bars for 'conference', 'datum', 'eswc', 'iswc', and 'semantic web'.
  - 'Top 5 Frequent Phrases' (not fully visible).

# Architecture

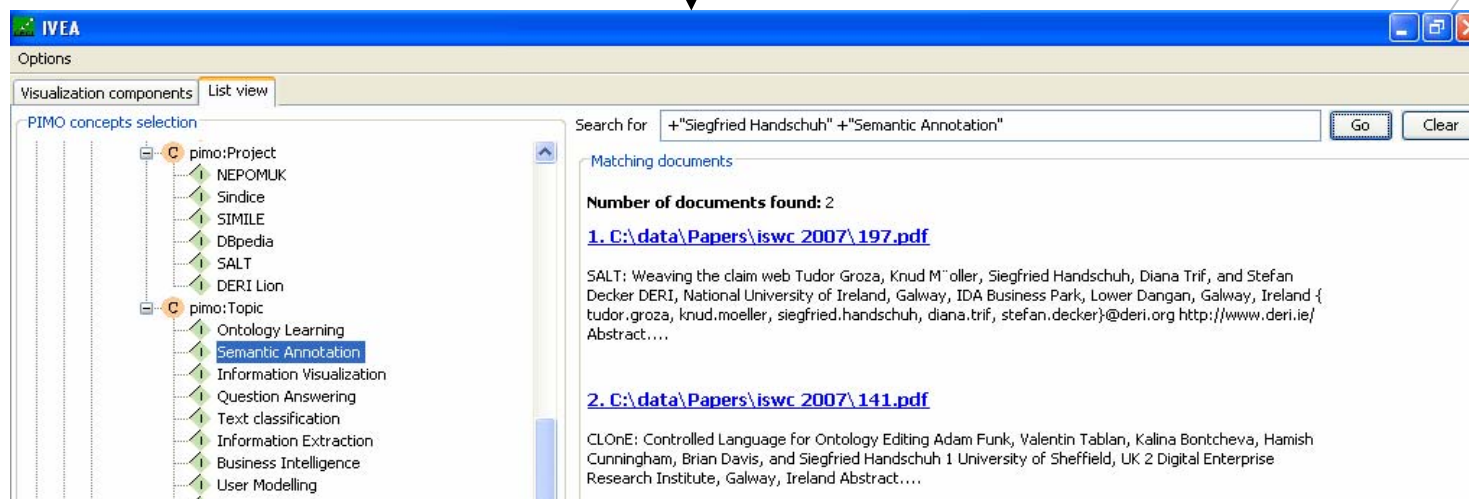


# Usability Study



## ■ Purpose

- identify potential usability problems
- get initial indications about the users' performances by using IVEA versus by using the baseline interface



## ■ Description

- Subjects:
  - 6 researchers
- Assumption:
  - same sphere of interest encoded in a predefined PIMO ontology
- Test collection:
  - 62 research papers
- Tasks:
  - 4 simple tasks and 4 complex tasks

## ■ Description

### □ Measures:

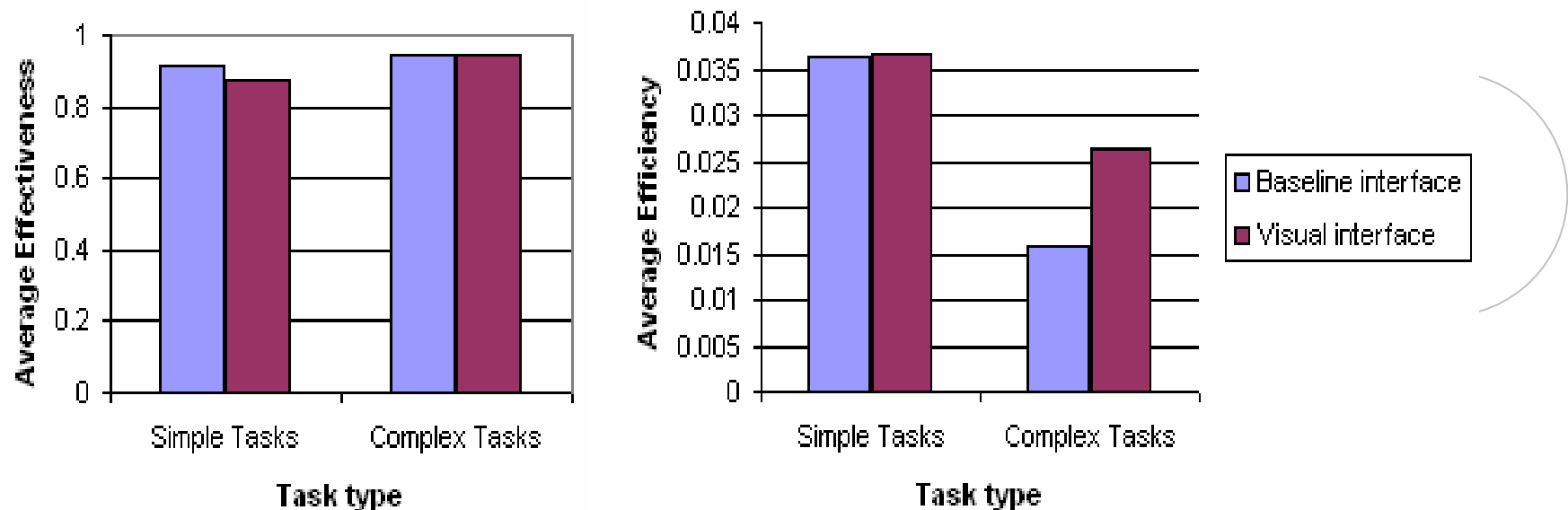
- **Effectiveness**: the accuracy and completeness
- **Temporal efficiency**: effectiveness/ task completion time
- **Users' satisfaction**:
  - 12 Likert-scale questions
  - 1 open-ended question seeking suggestions on improving the design.

### □ Notes:

- not a formal summative study, but still provides **initial suggestive indications**
- **advanced functionalities** (e.g. ontology enrichment) only available in **IVEA**, not in the baseline interface

## ■ Findings

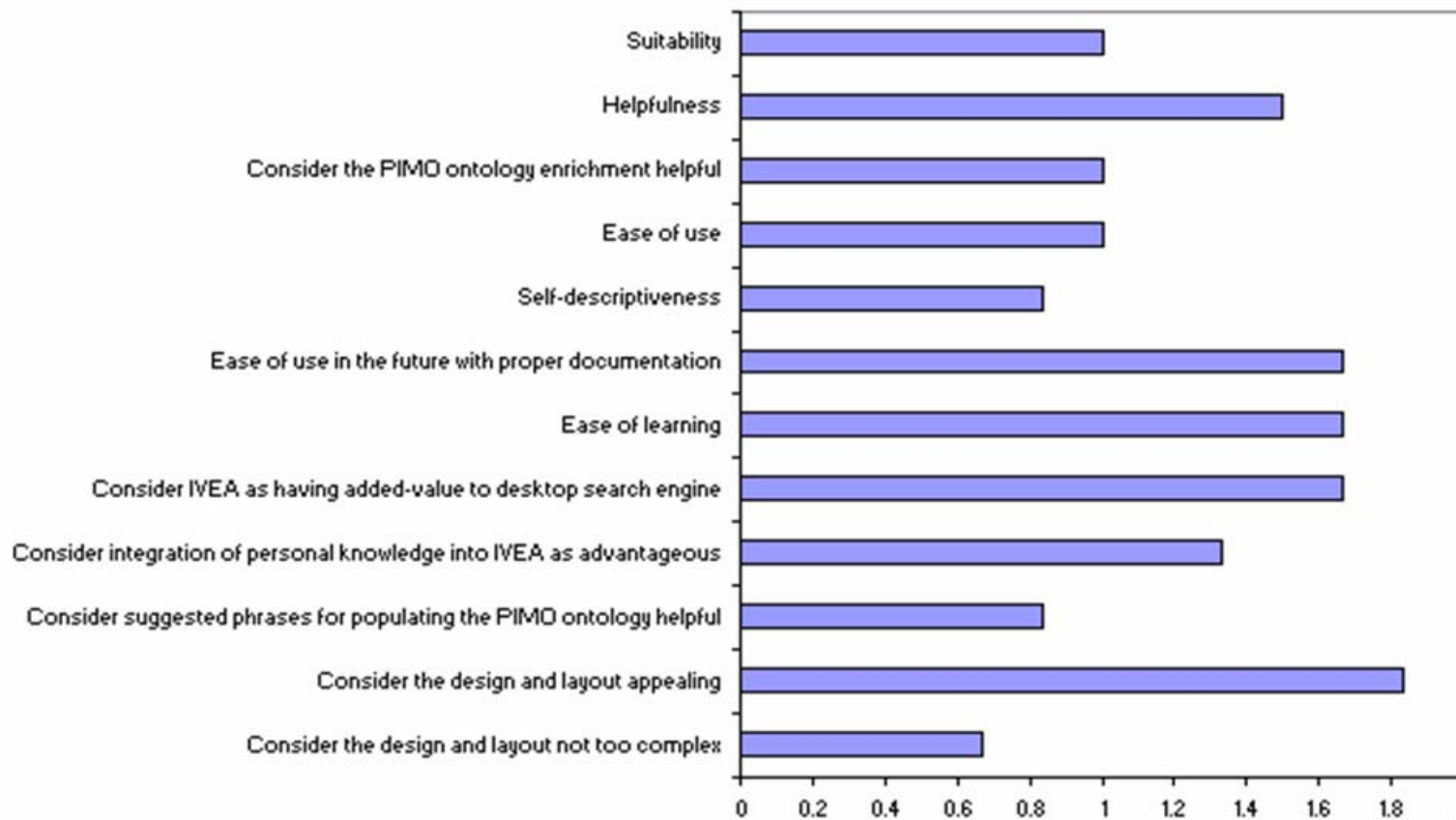
### □ Participants' performances





## ■ Findings

### □ Participants' satisfaction



## ■ Findings

### □ Participants' suggestions

- **Collection Overview** at more than 2 dimensions.
- **Highlight** the most relevant document.
- Display the **context** surrounding **PIMO instances** for TileBars' cells.
- **Group** the visual items **by class**.

# Recent update



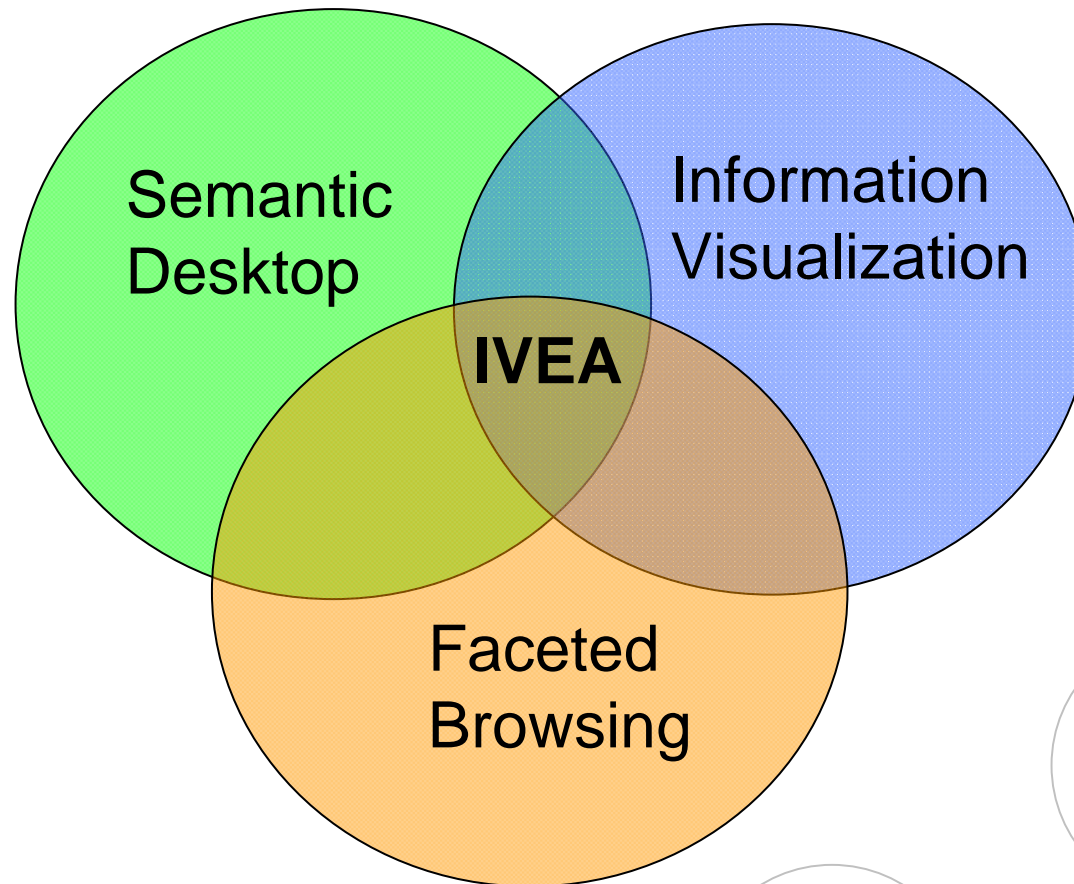
The screenshot shows a software window titled "IVEA <2>". It features a "Personal Knowledge View" on the left with a tree structure containing nodes like "pimo:SocialEvent", "pimo:Meeting", "pimo:Task", "pimo:Project", "NEPOMUK", and "DBpedia". The main area displays a "Collection Overview" with a list of items: "pimo:Person", "Siegfried Handschuh", "Tudor Groza", "NEPOMUK", and "Semantic Annotation". Below this is a document viewer showing a PDF document titled "SALT: Weaving the claim web" by Tudor Groza, Knud Möller, Siegfried Handschuh, Diana Trif, and Stefan Decker. The document content includes contact information for DERI, National University of Ireland, Galway. A metadata panel at the bottom left of the document viewer lists properties such as "pimo:City: Innsbruck", "pimo:Organization: DERI", "pimo:Person: Siegfried Handschuh", "pimo:Project: NEPOMUK", "pimo:Person: Tudor Groza", "pimo:Country: Ireland", "pimo:Project: SALT", "pimo:Country: Austria", and "pimo:City: Galway".

# Future Work



- Improve functionalities based on [user feedback](#)
- [Co-reference resolution](#) of PIMO instances (e.g. “European Central Bank”, “ECB”, “the bank”)
- [Ambiguity](#)
- [Scalability](#)
- Carry out a [formal summative evaluation](#) to gauge its usability

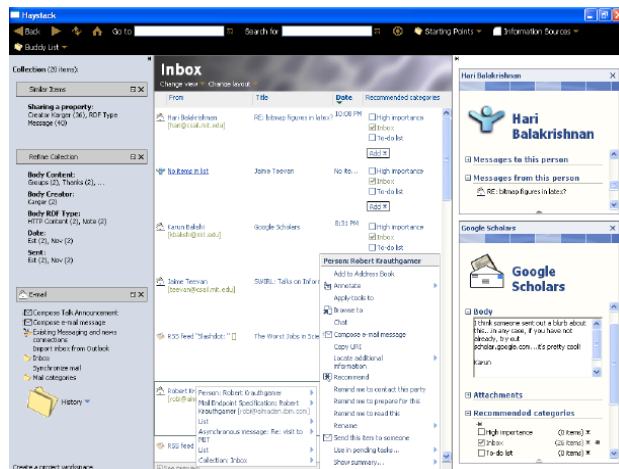
# Related Work



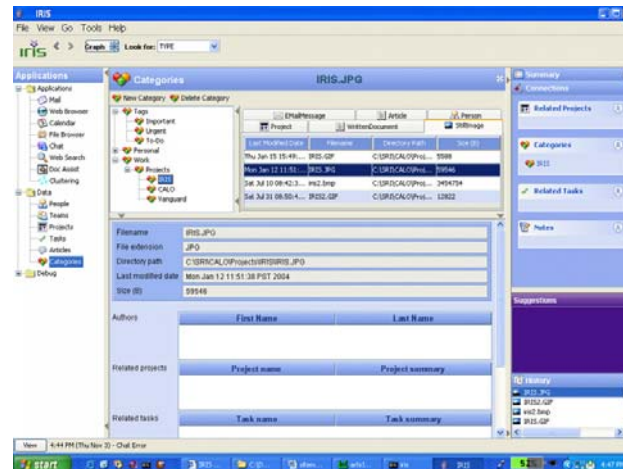
# Related Work



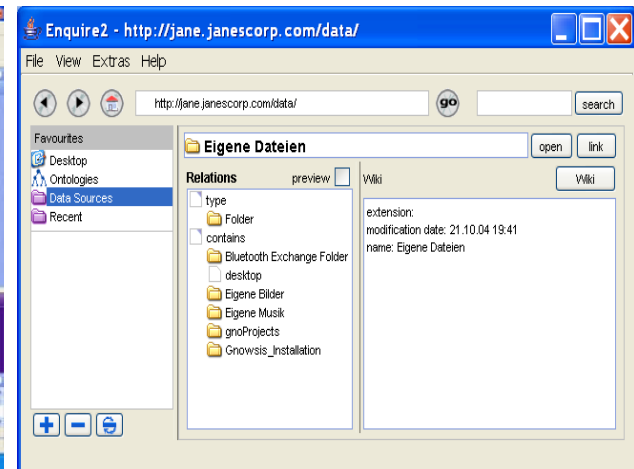
## ■ Semantic Desktop



Haystack



IRIS



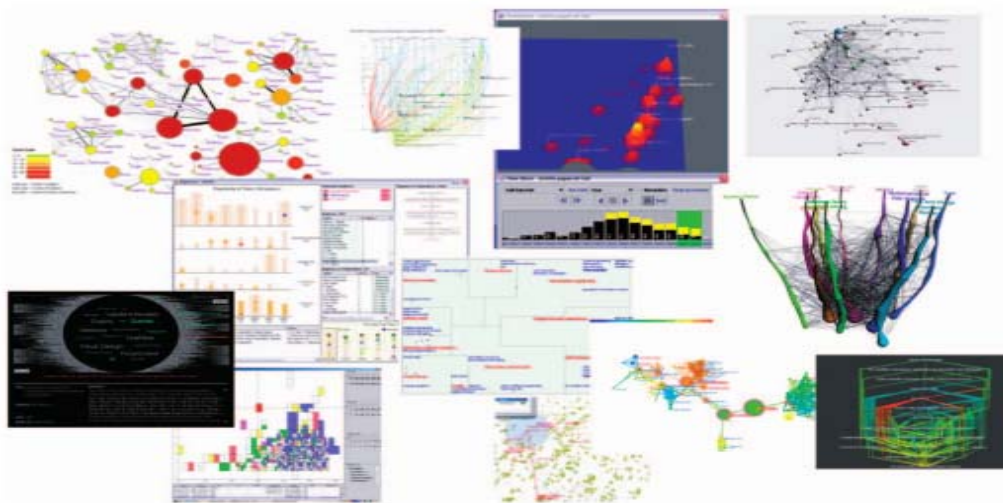
Gnowsisi

# Related Work



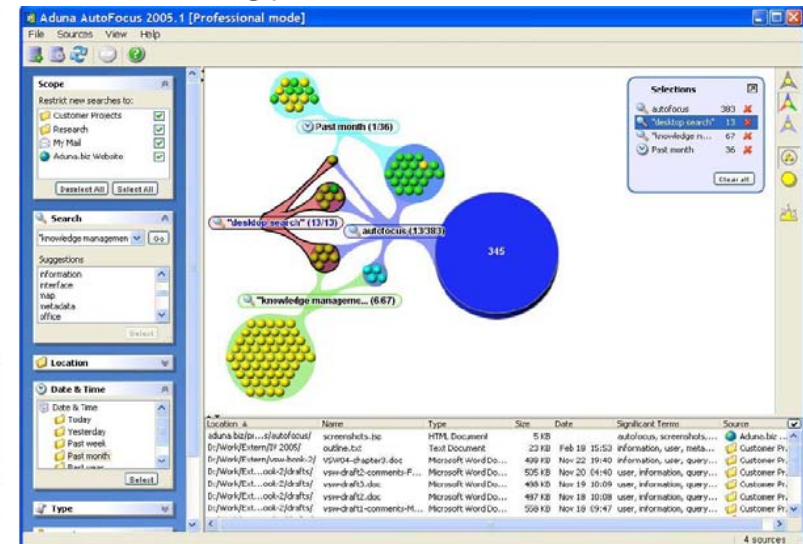
## Information Visualization

### Exploratory Visualization



InfoVis 2004 Contest

### Ontology-based Visualization



AutoFocus

# Related Work



## Faceted Browsing

The screenshot shows a search interface for Nobel Prize winners from 1901 to 2004. The search results are filtered by 'United States of America' and 'chemistry'. The interface displays various facets for refining the search, including Gender (female: 6, male: 272), Country (Australia: 1, Austria: 2, Belgium: 2, Canada: 2, China: 2, Egypt: 1, Federal Republic of Germany: 4, France: 1, Israel: 1, Japan: 3, more...), Affiliation (all: United States of America), Prize (chemistry: 53, economics: 41, medicine: 93, peace: 16, physics: 80), and Year (1943-2004). The main results are grouped by prize, showing portraits and names of winners like Ahmed Zewail, Alan G. MacDiarmid, Alan Heeger, Bruce Merrifield, A. Michael Spence, Clive W.J. Granger, Daniel Kahneman, and Daniel L. McFadden.

Flamenco

The screenshot shows a search interface for breakfast cereal characters. The search results are filtered by 'General Mills' and 'Post'. The interface displays various facets for refining the search, including Brand (General Mills, Kellogg, Nabisco, Post, Quaker Oats, Ralston), Decade (1940, 1950, 1960, 1970, 1990, 2000), and Country (Great Britain). The main results show character images and names like Bugs Bunny, Jimmy Cricket, Fred Flintstone, Alpha-Bits Sailor Euy, So-Hi, Sugar Bear, Mickey Mouse, and Honeycomb Kid.

Exhibit



# Conclusions



- **IVEA**- a tool for personalized visual exploration of text collections
  - integrates the **PIMO** ontology representing the **users' interests**
  - allows for **dynamic exploration** at different levels of detail
  - has the potential to support **complex exploration tasks**
  - enables **ontology enrichment** on-the-fly
  - benefits other PIMO-based applications on **Semantic Desktop** environments

Thank you