

Thrust 3: Knowledge Representation & Management

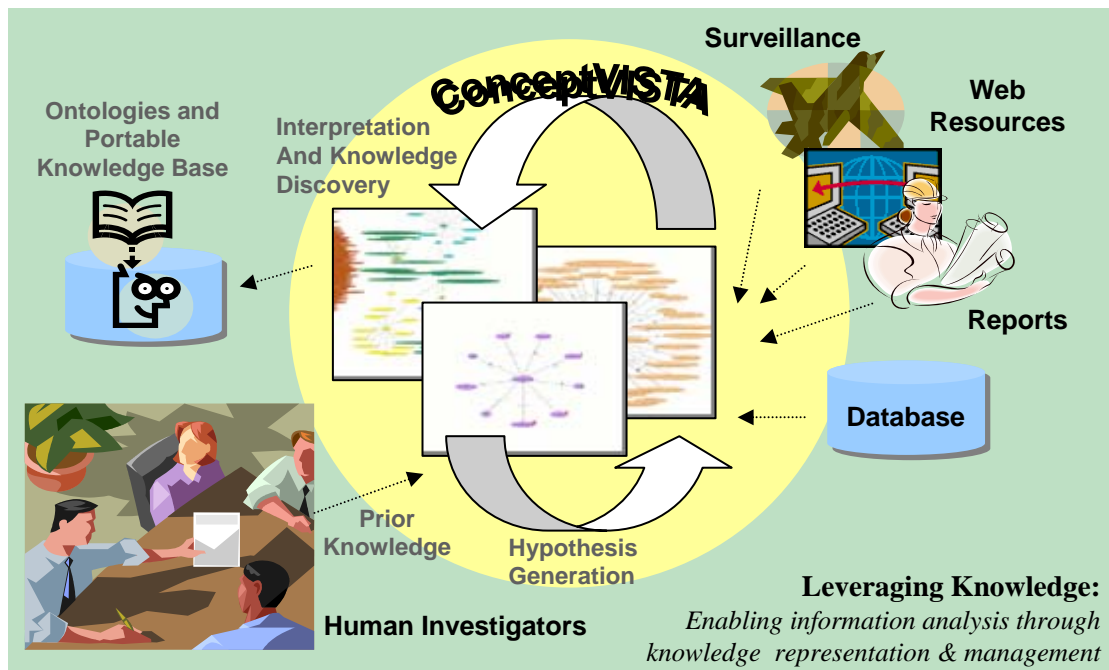
Benefit: *Relevant knowledge is hard to obtain. Thus, systems designed to support complex tasks such as threat analysis, situation assessment, and crisis management must include strategies to represent and manage both existing knowledge and that generated on-the-fly. NEVAC research in knowledge representation and management supports human decision making through an interactive environment where text documents, Web pages, ontologies, and data sets are encoded, visualized and coordinated to help human users explore, create, and share knowledge.*

Collaborator(s): *NVAC, Pacific Northwest National Laboratory.*

NEVAC is developing tools to help analysts and decision makers identify and manage complex concepts and relationships embedded in a range of information sources relevant to threat assessment and crisis management. Our main purpose is to support situations where human reasoning with complex information is critical. Reasoning is supported through tools that allow human analysts to interactively encode assumptions, hypotheses, conclusions, and other tacit knowledge as concepts and relations, and to visually connect them with documents, raw data, and shared common knowledge (encoded as ontologies) to assist high-level interpretation and sense making.

Our focus is on integrating technologies from three areas: **concept mapping and knowledge visualization**, which offers intuitive user interfaces for human analysts to encode tacit knowledge as concept maps and ontologies; **ontology management**, which ensures knowledge developed by various means (e.g. built manually or automatically harvested) can be queried, managed, and exchanged; and **text browsing / analyzing**, which provides visual linkages and coordination interfaces between knowledge visualization and other human-readable information from sources such as the Web.

Our main approach is to build an integrated software environment based on our knowledge visualization tools and relevant technologies developed by other NEVAC groups. We have extended ConceptVISTA, the main knowledge visualization platform developed by the GeoVISTA Center, to incorporate advanced text browsing tools (based on Java Desktop Integration Components, or JDIC), Web search engines, ontology reasoners (from Jena API), the Prefuse data visualization API, and Web map services. In addition, we have developed software interfaces between ConceptVISTA and other NEVAC tools such as FactXtractor and Improvise.



Funded by: *The Department of Homeland Security and the National Science Foundation.*

Early Development

Lab Prototype

Commercial Product

Jan 2008 – tools within this thrust cover a range from early development, through lab prototypes, to open source applications being distributed

Knowledge Representation and Management in ConceptVISTA

ConceptVISTA is a software platform for ontology development and knowledge representation. Recently, it has been re-engineered and extended to support integrated visual analysis, especially in the following areas:

Interactive concept mapping: In addition to basic concept map editing functionality, concepts and relations in a concept map can be sent to various search engines, with the results sent back and appended as new concepts. This means that human analysts can directly harness information from the Web to assist the encoding of tacit knowledge.

Ontology management: ConceptVISTA supports complex ontology queries through a visual query editor and offers customized ontology matching algorithms based on selected concept similarity measures. This is particularly useful when exploring large ontologies because it allows analysts to focus on relevant parts of a huge concept map.

Web page builder: ConceptVISTA incorporates advanced text content extraction and semantic matching tools based on WordNet and derived from our *Information extraction and contextualization* thrust. These analyze text document contents (including Web pages), identify key words in the text that semantically match selected concepts in a concept map, and insert “tags” or new styles into the documents to highlight the keywords to support human analysts.

Finally, ConceptVISTA can be used to visualize and analyze results obtained from automated knowledge extraction tools such as FactXtractor. The knowledge visualization component of ConceptVISTA has also been coordinated with data visualization tools using the *Improvise* coordination architecture, part of our *Visual-Computational Methods* thrust.

Integrated Environment for Visual Analysis

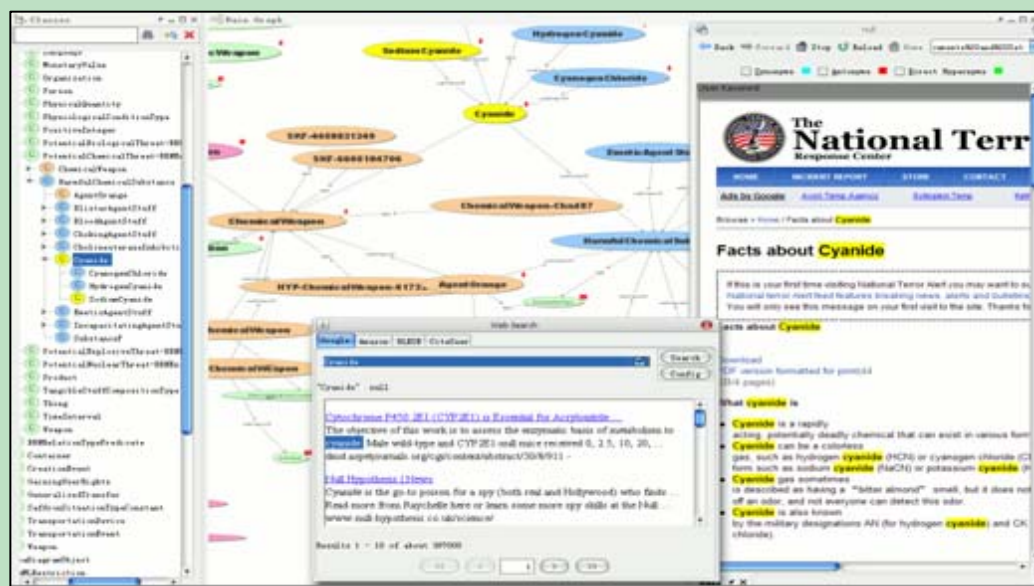
Our ultimate goal is to leverage work in all four thrusts to develop an integrated environment for visual analytics. A key challenge is to clarify the inter-connections between data and knowledge, and to accurately define their roles during the course of visual reasoning. We are examining this issue from theoretical, system design, and practical levels.

At the theoretical level, we are focused on defining a conceptual foundation for visual reasoning. From the perspective of this work, we are particularly interested in theories of human problem solving, especially how knowledge is used and generated in different tasks.

Our immediate system design focus is to create an open coordination framework that incorporates both tabular data and graph-based knowledge representation. Preliminary work has focused on building direct mapping between data elements and knowledge entities. We will next extend this approach to support fully-functioning cross-queries between tabular databases and graph-based knowledge bases or ontologies.

At the practical level, it is necessary to develop realistic user cases and user evaluation scenarios to align the software design with mission needs. It is important to evaluate different data-knowledge coordination patterns under different problem solving situations (as identified at the theoretical level), and to study effectiveness of the software design in real world threat analysis and crisis management scenarios.

In summary, we believe the representation and management of knowledge is fundamental to visual analytics; sensemaking must be grounded in existing and evolving knowledge. The approach we are taking integrates advances in representation and management of knowledge with those in information extraction, visual-computational analysis, and group work.



Linking knowledge with data: This example illustrates a typical use scenario for the tools described—grounding an online query in existing knowledge. The concept map shows the query result for “cyanide” in OpenCyc, one of the most comprehensive online knowledge sources. Analysts can activate various Web Search engines to find additional resources, and the keyword “cyanide” can be highlighted in the Web browser to facilitate human reading.