The large volumes of data from open and sensitive sources require today’s analysts to perform a difficult balancing act. Analysts must be able to extract pertinent facts from the data flow while not discarding other important information. The categorization and organization of data and searches for information are often dependent upon the individual analyst and the implied semantics of the data organization may be hard to convey between different users. Sharing this data with additional analysts is possible via networked drives or web-enabled folders but the organizational structure of the data may be unclear to others.

ISX has developed a visualization and organization tool to assist today’s analyst. The Semantic Navigator will display, navigate and visualize the discovered entities and relationships contained in the source data. Entity extraction technology categorizes incoming data and ontologies provide information organization. Ontology hierarchy navigation and ontology element partial phrase matching are used to build queries. Query results display the relationships of the entities extracted from the documents and allows the user to navigate between data nodes. Collaboration and information sharing is achieved by packaging the product as a tool in the Groove collaboration framework.

ISX is applying Semantic Web (Berners-Lee et al. 2001) technology to address the visualization and organizational needs of today’s analyst. As the amount of raw data available to analysts increase, the task of managing and exploiting the knowledge inside of this information space becomes increasingly difficult. The Semantic Web aims to ease these difficulties by providing an ontologically grounded and machine-accessible index of content, thereby dramatically improving the ability to leverage the knowledge contained in the information space. Ontologies allow the expression of the entities and relationships found in the information, which resides in the analyst workspace. The DARPA Agent Markup Language (DAML) (DAML 2005) captures these ontologies and enables the categorization and organization of the information in an analyst workspace. The Semantic Object Web (Kettler et al. 2003) extends the idea of the Semantic Web into an object-centric index of the information space, where actual bits of information are encapsulated in the Semantic Web as Knowledge Objects.

The Semantic Navigator leverages the Semantic Object Web to remove the barriers between heterogeneous information spaces and provide users with a seamless mechanism for exploiting knowledge across traditional boundaries. The tool resides in a Groove (Groove 2005) workspace and shares data between users in the workspace.

The Semantic Navigator uses Inxight’s Smart Discovery® (Inxight 2005) entity extraction technology to discover entities and simple relationships as data enters the workspace. Documents, user actions, and discussion threads from the workspace are processed and this information is stored in a Groove-base knowledge base (GBKB) based upon the Groove record set engine. The GBKB supports queries and query results are viewed via a hyperbolic tree representation. The representation of the ontology categories in the hyperbolic tree preserves
the underlying organization of the information. The tool aids the user in the tasks of categorization and organization of data; secure storage and sharing of the information; automatic resynchronization of information; queries and information retrieval; and information visualization.

2. Information Organization

Information gathered by the Semantic Navigator is stored in the GBKB. The GBKB will ingest the semantic content of the workspace and transform that data into entities to be stored in the Semantic Object Web. The data ingest occurs as information is added to the workspace and may include documents added to the workspace; discussion threads entered into the workspace; and designated process information.

Data ingestion modules control the flow of information into the GBKB. These modules are responsible for ingesting information from the other tools in the space and indexing that information into the GBKB. These modules read the events and initiate the content conversion process. The process conversion module converts information about actions that take place in the information space. The semantic conversion module converts the actual semantic content of an information space.

3. Knowledge Discovery and Information Exploration

The Semantic Navigator provides a familiar Boolean keyword search and query capability, that is enhanced by the categories and relationships. All search results are returned in a hierarchical manner, organized into categories of objects according to the ontology(ies). This allows the user to rapidly scan the different classes of objects and selectively analyze those classes pertaining to their intended search.

The Semantic Navigator facilitates the discovery of previously unknown information with a graphical navigation view that allows users to navigate the GBKB via a single unified index for the entire information space. Information contained in the GBKB, including query results, may be viewed via a hyperbolic tree view or a traditional tree-like folder structure. The hyperbolic tree is part of the Inxight tool suite. Ontologies define information in a graph structure that maps well to a hyperbolic tree view and eliminate the boundaries induced by the packaging of information into various artifacts. As individual elements are displayed in the view, the relationships between data elements are preserved and the navigation is allowed between the data nodes. (Figure 1)

The Semantic Navigator provides several methods for managing the history of a user’s exploration. These include the highlighting of a traversed path and a list of the nodes visited. Users may use these features to reverse their traversal or directly jump to previously viewed nodes. The user may also annotate nodes with textual or audio notes during the traversal. Users may also save their explored paths as an independent file and retrieve it for later use.

4. Information Sharing

The Semantic Navigator supports two modes of information sharing. The Groove environment provides a secure collaboration environment for Groove-enabled tools. Tools share documents, messages and other content via this method. Pointers to the source material are included with the ingested content and supported with drill-down into the source information. The second mode of information sharing allows multiple users to join together and view the same path of information exploration.

References


