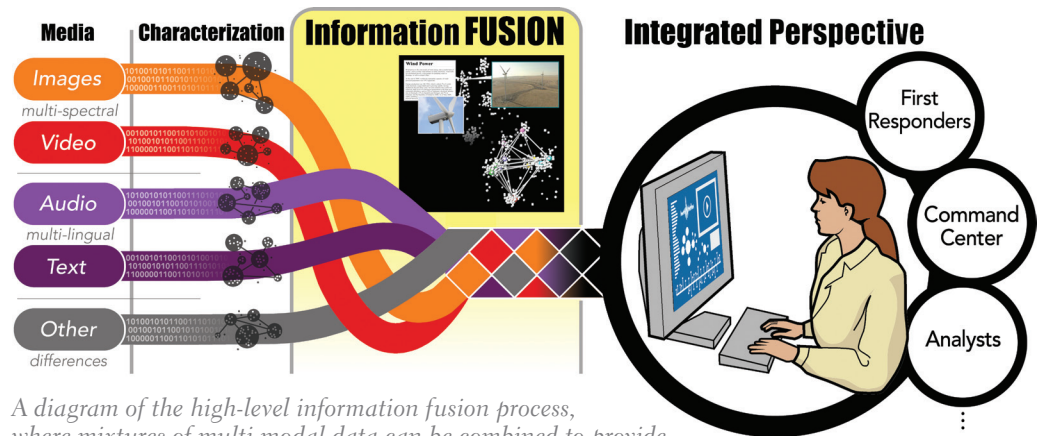


Equinox—Information Synthesis and Fusion:

A Method for Semantically Aligning Content for Text, Audio, Images, and More

A major challenge in analytics is creating the capability to process, organize, and analyze multi-modal data. In the past few years, we have seen an explosion of content: videos from YouTube™, images from Flickr™, and more news and blog outlets all vying for our attention. The amount of available digital information (more than can be read in lifetime) will only increase an analyst's load. Compounding the problem is the fact that most analytic tools are unimodal, only able to analyze and represent one type of data, causing analysts to use several tools to understand the content based on its modality. As the number of modalities included in an analysis increases, so do the tools; the analyst begins a juggling act of combining various analytic products from each tool. Unfortunately, this aggregation operation is not managed within a single process or tool but in the analyst's mind. Today's analytic environment and computational models have not addressed the need for an integrated analytic environment, especially at the semantic level.



A diagram of the high-level information fusion process, where mixtures of multi-modal data can be combined to provide an integrated perspective useful for insight and decision making

Equinox—Information Synthesis and Fusion

Over the last few years, the National Visualization and Analytics Center (NVAC) has been developing a framework called Equinox. This framework leverages various knowledge representations where content and context can be aligned or fused, thereby creating a new information space from which an integrated perspective can be formed. We are using semantic characterization technologies of the analytic and computational communities. These communities have focused on developing signatures that semantically represent their research content. Signatures, as we define them, are computational summaries based on specific features of a document (where a document represents an image, text, video, etc.). Semantic signatures are based on semantic features within the document and thus can help define a common basis in which information can be analyzed and represented. Also, many visual analytic systems are already poised to leverage this type of representation. Therefore, the methodologies this team has developed incorporate the semantic signatures of various information types and produce a fused signature where all data types could coexist.

InFusion: A Library for Content Alignment and Fusion Signatures

To create fusion signatures, NVAC's Equinox program has focused on researching and developing alignment strategies. The overall Equinox program has several components, including a visual analysis tool and a testbed that is centered on the core InFusion library, which houses the main algorithms for aligning various multimodal signature spaces into a common fused spaced. The InFusion library does not require complex ontologies or a large knowledge base; instead, it is based on specifying equivalence relationships using a set of content from the different data spaces being fused.

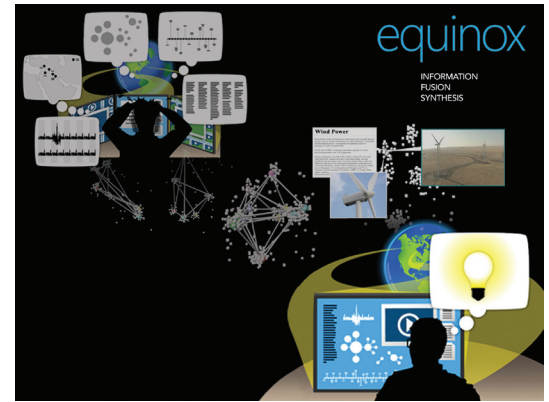
For example, consider combining a set of structured category records and a set of unstructured text documents. Our approach is to first establish a set of content equivalent or parallel documents that exist in both datasets (i.e., data types). These documents establish the semantic bridge between the different data spaces. Note these equivalent documents can be generated or be identified from the data sources a priori. Analysts then use the semantic signature algorithms to convert their respective datasets into semantic signatures. These signatures are then processed through InFusion, via the equivalent documents, resulting in a new signature set for both datasets. The new signatures are structurally equivalent to their predecessors and can be used in many of the visual metaphors and analytic processes as before.

A distinguishing feature of this process, in contrast to other fusion-based systems, is what it's not trying to do. It does not attempt to identify individual low-level semantic features and then align and reconcile them between spaces. Instead, it uses the co-occurrences of those low-level features and draws on statistical relations within each signature space as it's organized into the fused space. An important property of alignment is identifying when the content has significant contextual differences. InFusion can resolve the context issue by relying on the context as defined by the individual signature spaces. If the equivalence document pairs can be fused, then those portions of the space are deemed contextually similar. If, however, there is sufficient stress in aligning the equivalence document pairs, then they indicate a difference of context and are not aligned. At the heart of the fusion algorithm is the important process of balancing alignment vs. deformation for each equivalence document pair. As we execute the algorithm, we measure how much each space deforms as it attempts to meet the alignment criteria. The goal is to minimize the deformation of each signature space as we move toward an alignment of the equivalence document pairs. Part of our research is to determine how much deformation is acceptable versus how close of an alignment is necessary to achieve analytic value.

We have begun defining user testing scenarios and analytic tasks that will show the value of a combined fused environment. While we have fused datasets, it still requires a user study to quantify these issues of accuracy for deformation vs. alignment. To understand better how users deal with multimodal information, part of the user study will examine use of text, audio, images, and video in solving a scenario. We expect to have results with multimedia information in CY2011.

A Community Effort

For this effort to be successful, the domain communities must develop high-quality semantic signatures for the various data types. Over the past several years, we have seen huge leaps in bridging the semantic gap for imagery and video; however, much more is required before we are able to leverage true semantic signatures. While the semantic signatures may not be at the level



Equinox is a fusion-based system that enables users to align, test, and explore heterogeneous data spaces through the use of semantic signatures.

we would like, that does not mean a certain value cannot be achieved with what we have. For example, with some level of machine learning, some computer vision algorithms are able to identify certain scenery or objects. This provides us with the initial fodder to perform experiments, identify how much of a gap exists, and establish baselines that we can make improvements against.

Unifying the semantic landscape for heterogeneous data will continue to be a challenge, but with this initial research and university collaborations, researchers are overcoming the obstacles. In time, analysts will be able to focus on the problem and not just on how the pieces fit together.

Team Members

Pacific Northwest National Laboratory: Shawn Bohn, Grant Nakamura, Amanda White
Collaborators: Dr. Haesun Park, Jaegul Choo
(Georgia Institute of Technology)

Point of Contact

Shawn Bohn, Senior Research Scientist,
shawn.bohn@pnl.gov



Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965