*Paper accepted for the **2005 Intelligence Analysis Conference**, May 2005, Washington, DC.*

Information Triage with TRIST

David Jonker, William Wright, David Schroh, Pascale Proulx, Brian Cort
Oculus Info Inc.

{david.jonker, bill.wright, david.schroh, pascale.proulx, brian.cort}@oculusinfo.com

Keywords: Multi-INT/fusion, All Source Intelligence, Novel Intelligence from Massive Data, Search and Retrieval

Abstract

TRIST ("The Rapid Information Scanning Tool") is the information retrieval and triage component for the analytical environment called nSpace. TRIST uses Human Information Interaction (HII) techniques to interact with massive data in order to quickly uncover the relevant, novel and unexpected. TRIST provides query planning, rapid scanning over thousands of search results in one display, and includes multiple linked dimensions for result characterization and correlation. It also forms a cohesive platform for integrating computational linguistic capabilities such as entity extraction, document clustering and other new techniques. Analysts work with TRIST to triage their massive data and to extract information into the Sandbox evidence marshalling environment. Initial experiments with TRIST show that analyst work product quality is increased, in half the time, while reading double the documents.

1 Introduction

As part of the Novel Intelligence in Massive Data (NIMD) research program [ARDA, 2002], new interactive, information visualization techniques are being investigated which tightly couple massive data, software agents and the analyst's exploration task. A break-through in finding novel intelligence is believed possible if all the components are combined in a system of systems. Progress has been made towards an integrated cognitive space where analysts will see, and interact with, more information, more quickly, with more comprehension. This space is called "nSpace" and is the combination of the multi-dimensional linked views found in TRIST with the visible and flexible cognitive mechanisms of the Sandbox.

Figure 1. Workflows supported by nSpace.

This paper focuses on TRIST. First, the results of a cognitive task analysis are presented. This is followed by a discussion of related information visualization work. Then there is a technical discussion of TRIST concepts and capabilities. Finally, results are reviewed from an experiment conducted at NIST that measured the performance impacts of TRIST.

2 Analysis Cognitive Task Analysis

Cognitive task analysis (CTA) is a layered system-based framework used to optimize the match between a task's cognitive demands and the design solution [Schraagen, 2000]. CTA techniques include a variety of systematic observation, interview and analysis methods, all aimed at providing insight into the mental processes underlying complex human tasks.

In preparation for investigating nSpace concepts, two CTA studies were completed: structured interviews and a review of analyst activity as logged by the NIMD "Glass Box" [Cowley et al, 2004]. The observations from these CTA studies were used to refine the performance objectives for TRIST and potential task metrics.

2.1 Speaking with Analysts - Observations

Structured interviews were conducted with fourteen analysts who work with a variety of sources and on a range of short/long term, narrow/broad focus subjects [Wright and Kapler, 2004]. The interviews were wide ranging and touched on many topics: IR, Analysis Methods, Tools, Work Products, Objectives, etc. With respect to information retrieval, it is clear that executing queries, sequentially scanning results, opening and reading documents is a common, time consuming task for analysts. The following are illustrative excerpts on massive data and IR:

> "**You start with 30,000 hits** … which you need to understand and put in some order. And I need to see documents and in nine different ways.

> "I have about 1,000 messages or reports a day. About 300 are relevant. And maybe 30 go to my log. I cut and paste fragments into my log. **Doing updates takes most of my time**.

> "I need to see the searches. **I need to jump back to any level** in the searches. … And I can't refine searches now …and there are no nested searches …

"With Google, **I read and read and read, then I get brain dead.** After so many pages, you get exhausted … after the first 45, I clear it. You get worn out.

"More than **95% of what we deal with is noise**. Try data analysis with > 95% noise in the data, and with varying degrees of reliability of the relevant data.

"I need to **see relations between concepts and entities**. I need to pull other than through keywords."

2.2 NIMD Glass Box Activity Analysis

The NIMD "Glass Box" (GB) provides an experimental laboratory to examine analyst work by collecting data about what analysts do and how they do it [Cowley et al, 2004]. Initial analysis of NIMD GB data allowed a characterization of open source analyst activity, including the observed IR process, as shown in Figure 2 and Figure 3 [Wright, Kapler, 2004]. Much GB analyst time is spent on IR. On average, for the six month period examined, the analysts spent 3.2 hours per day in the GB of which 82 minutes were spent using Internet Explorer, Acrobat and Media Player (IR tools) and 72 minutes were spent using MS Word and Windows Explorer (evidence marshalling tools).

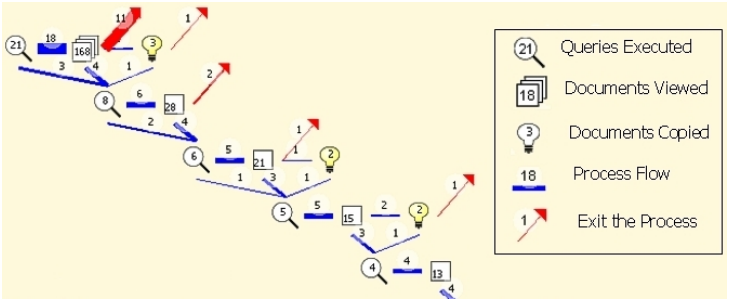

Figure 2: Observed Analyst Information Retrieval Behavior.

countryA + countryB + countryC + countryD + XYZ weapons + proliferation
countryA + missiles + countryB + countryC + countryD
countryA + missiles + countryB + countryC + countryD + XYZ weapons
countryA + missiles + delivery of XYZ warheads

Figure 3 – Query Refinement. Observed Analyst Query Trails.

The well known analysis "bath tub" curve [Rose, 1996], with most analyst time spent in IR and reporting and less time using analytic methods, was replicated in the analysis of GB data. Once again, a system that would yield a measurable, order of magnitude, increase in productivity in IR is needed, so that more time can be spent on analytic methods.

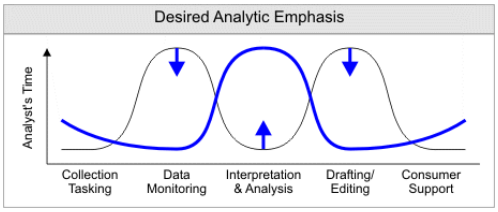

Figure 4. Need for New Analytical Emphasis

2.3 Conclusions from Observations - What Do Analysts Need?

Executing queries, sequentially scanning results, opening and reading documents is a common workflow. Queries are often iteratively refined, can become quite complex, or be freshly developed and established as new thoughts are followed like trails. Results are scanned for when written, source, buzzwords, keywords, corroborating evidence, new items, trustworthy baseline document, summaries, relevance, etc. The nature of the data is varied and voluminous. People feel overwhelmed working now with just hundreds and thousands of items such as observations, reports and events, but if analysts were able to work with hundreds of thousands and millions of items, they would. Keeping track of sources and queries is time consuming. Fatigue and cognitive strain are factors. Analysts need an IR system that will increase their productivity in a 'triage' workflow without removing information on which human judgments can be accurately and quickly made.

Analyst work is not sequential, and moves back and forth, across multiple tasks, at a moment's notice. There is a need for an integrated approach for supporting analysts. An integrated environment should provide a common visual vocabulary for analytic work as well as creating a mixed-initiative platform for the whole analysis workflow. Analysts need a system that can easily integrate new/different IR technologies. There is an opportunity for a test bench approach. Not every method performs the same in the context of all tasks. Analysts need a way to determine which tools and methods are most effective for the task at hand. Finally, information seeking is only one part of the full work process, and must be connected with sense-making.

3 Related Work

Information visualization techniques amplify cognition by increasing human mental resources, reducing search times within displays, improving recognition of patterns, increasing inference making and increasing monitoring scope. There is a significant body of work on information visualization techniques for IR [Card et al, 1999] [Hearst,

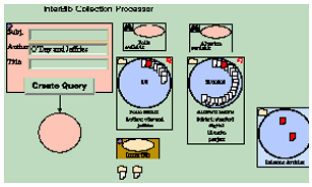

Figure 5. DLITE

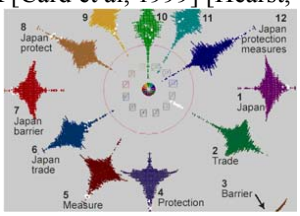

Figure 6. Sparkler

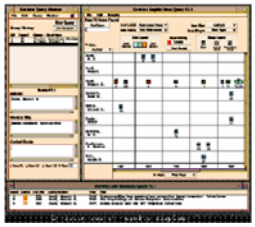

Figure 7. Envision

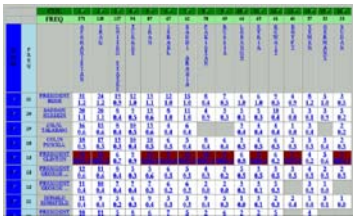

Figure 8. PathFinder

1999] [Geroimenko, Chen, 2003]. Approaches have been investigated to help users formulate queries, select information sources, understand results, adjust their search strategy and keep track of their progress. Nevertheless, the top three of ten suggested challenges in digital libraries were visual information retrieval, visual information exploration and visual information organization [Chen, 2001].

Several conceptual models explain user information seeking behavior including information foraging, sense making [Pirolli, Card, 1999] and "berry-picking" [Bates, 1989]. Three kinds of information seeking tasks are often discussed: monitoring a well known topic over time, following a planned search to achieve a particular goal and exploring a topic in an undirected fashion. Information seeking goals change as partial results satisfy needs. The common core is an adaptive, highly interactive process.

User interfaces for commercial search engine products such as Google, ClearForest, RetrievalWare typically show search results as a 1-D list of references in order of their computed relevance. About ten to twenty references are shown at a time. RetrievalWare also provides a 2-D table view withsearch results categorized in each dimension [Convera, 2005]. Recent work from Microsoft uses a list view but also includes selectable filters [Dumais et al, 2003]. These approaches often assume that the computed relevance will provide one or two high utility references that will satisfy the user's need.

ThemeScape moves beyond the 1-D list allowing thousands of documents to be visually, succinctly described, navigated and accessed [Wise et al, 1995]. *DLITE* [Cousins, 1997] is a graphical query system that uses iconic representations of queries and results, but relatively small amount of results can be shown. *Sparkler* [Havre, 2001] provides comparison of concept occurrence in a result set using visual profiles. *Envision* [Nowell, 1996] and *Search Result Explorer* [Andrews, 2001] group search results by displaying documents in a tabular 2-D display according to their metadata (e.g. author, date). Icons encode meta-data. The *PathFinder* co-occurrence matrix uses a 2-D grid display [Presearch, 2005]. Pathfinder operates on thousands of documents and supports the broader analytical workflow by providing twenty or so separate functions. Each function provides utility and is loosely integrated into a whole workflow but requires the analyst to transition among many separate single-purpose displays.

4 TRIST Performance Objectives

The overall objective of TRIST is to increase combined human and system productivity in IR in order to free analyst time for analytic methods. Most of the information retrieval (IR) workflow is supported in an integrated interface. Analysts are able to formulate, efficiently refine, organize and execute queries. Selecting results reveals metadata, content, contexts, as well as entities contained in those results. Exploration of results can start in any dimension or with any entities. Relevant, interesting, unexpected results can be isolated. Those results can efficiently be skimmed or read in the integrated document viewer. Nuggets that could contribute to the analysis are saved and organized in the Sandbox analysis space. The flow between analysis and IR is unrestricted and fluid, so that lines of thoughts can be followed freely without losing work or context.

Query comparison is also an objective. The visible comparison of multiple results, queries and IR methods should improve performance by closing the feedback loop. Analysts will be able to see the effect of query changes and search tools via quick visual indication of what is common, unique and new in result sets. Highlighting what is different is the first step to highlighting what is unusual.

5 TRIST Capabilities

5.1 Overview

TRIST is divided into panes for queries, dimension setup, dimension viewing and a separate entity dimension area. In the query pane, in the upper left of Figure 9, analysts formulate, refine, organize and execute concurrent, multiple queries. Query re-use, shortcuts and organization tools make for more efficient management of queries on multiple search engine systems. Analyzing results in the center screen visually feeds back directly into query reformulation on the left where queries are discarded and recalled, and previous result sets easily retrieved and reviewed. Favorite queries, for long standing assignments, can be browsed, selected and executed.

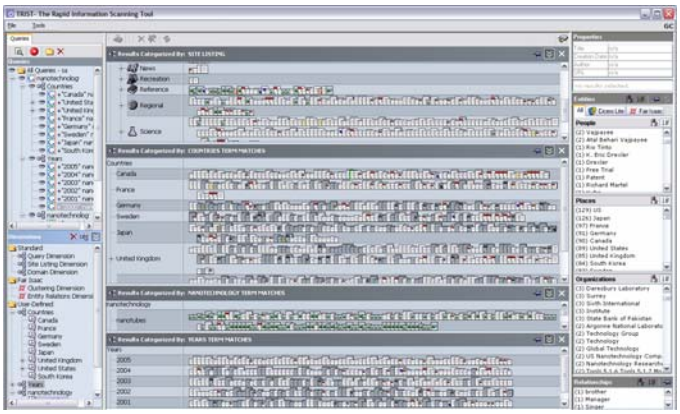

Figure 9. TRIST Layout. The center shows four dimensions: web site categories, countries, technology and year.

The center Results View provides multi-dimensional characterization that allows for fast, user controlled parallel processing of metadata. Results can be organized by dimensions such as source, country, technology, date. etc. Tailored dimensions and their associated "bins" can be selected or determined by the analyst in the lower left.

Metadata is encoded in the document icons (e.g. size, type, read, annotated, duplicate). TRIST displays hundreds of documents per dimension. The rich intuitive

iconic encodings maximize display density, making it possible to scan an order of magnitude more results.

Due to the compact visual representation of results, multiple dimensions can be seen simultaneously. Seeing the same information in each dimension reveals different aspects of content or context of the information. Deeper insight into larger amounts of information is provided by seeing information from many perspectives all at once. Uninteresting results can be identified and ignored, and unusual observations can be isolated. Entity extraction results, e.g. people, places, organizations, are placed on the right. TRIST uses a modular approach and a variety of extraction systems can be incorporated (e.g. CiceroLite from LCC and Fair Isaac IE technology). Entities that characterize the result sets can be used during scanning to identify the distribution of themes. Relationships can be seen in context.

5.2 Comparison and Difference Visualization

Sets of documents can be compared at a glance with a difference visualization, which quickly identifies what is new or unread. Grayed icons indicate duplicates. Colored icons indicate documents already opened. Difference visualization provides the feedback necessary to know whether a slightly different search method yields better or worse results. This judgment can be made without having to open a single document. Thus the difference visualization, or comparative analysis, closes the feedback loop and yields improvement in performance when trying to refine a query, or compare the results from multiple search technologies.

5.3 Defining Dimensions

In the lower left pane, dimensions and their associated "bins" can be selected or determined by the analyst. Any ontology can be used to create custom dimensions. For example, a Technology dimension is made by selecting senses (e.g. nanotechnology, aerospace, etc. with any number of sub-categories) from an ontology, such as WordNet or the Library of Congress Classification System. The analyst may also type terms to define dimensions manually. Dimensions are saved and are thus re-usable across tasks. Tailored dimensions make TRIST applicable to any task domain and reveals information of most interest to individual analysts. Watch lists can be set up to monitor and flag topics of interest.

5.4 Linked Selection

With linked selection viewing [Eick, 1995], selecting a subset of documents in one dimension, for example country, will highlight all those documents as they occur in the other dimensions, such as source and year. The analyst can quickly see trends and exceptions. For example, all information on one country is all published from the same source in recent years. Each dimension defines a selection context and selection highlighting occurs simultaneously across all views as shown in Figures 10 and 11. With linked selection, multi-dimensional display and careful iconic visuals, an analyst can quickly and correctly identify which documents, in result sets of hundreds, contain relevant or useful information without having to open and read them.

5.5 Clustering

Document clustering, using a context vector method from Fair Isaac [Caid, 1997], is an example of a computed dimension. An example is shown in Figure 12. Document similarity is scored and the document is assigned to the nearest cluster if it scores at least the threshold for that cluster. With unsupervised clustering (UC), results dragged into that dimension are clustered automatically on the fly and categories representative of the clusters are created. No prior knowledge of categories

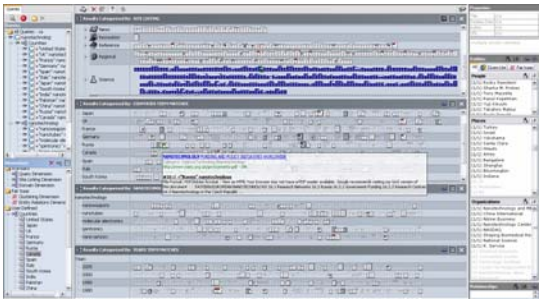

Figure 10. Linked Views – Selected dimension is blue. Occurrences in other dimensions are white.

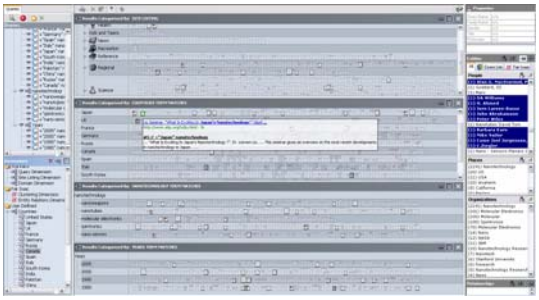

Figure 11. Linked Views – Selected entities are blue. Occurrences in other dimensions are white

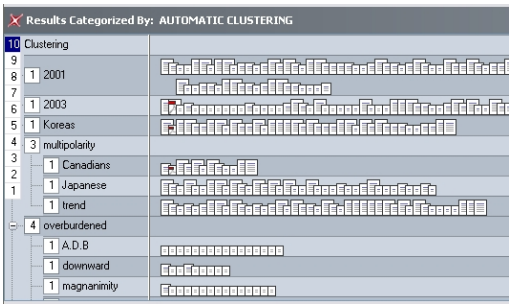

Figure 12. Clustering.

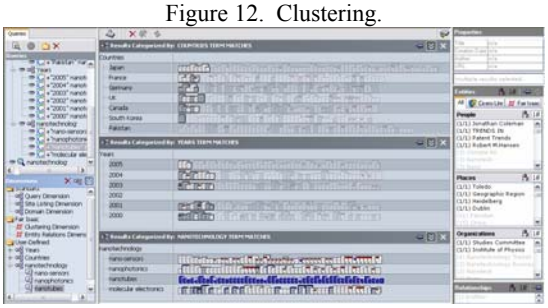

Figure 13. Trend Analysis.

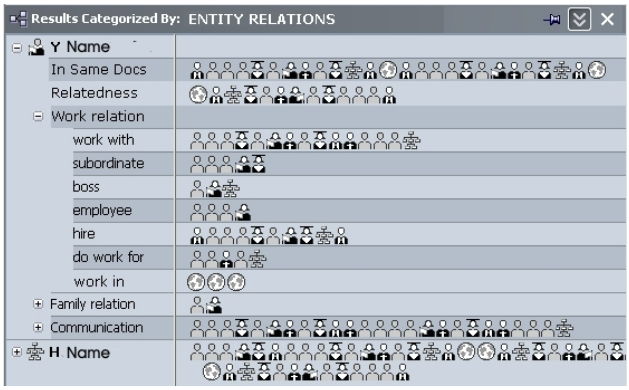

Figure 14. Entity Relations.

is necessary. The analyst can reclassify all or a subset of the results on the fly, building quick understanding of content by chunking it in many different ways. Unusual results or specific results can be isolated. Lightly supervised clustering, trained with keywords and sample documents, can also be integrated within the TRIST framework.

5.6 Trend Analysis and Massive Data

"Hot spots" and "indications" of trends can be detected across a broad spectrum of topics which can then be further investigated. Searching for the unexpected requires working very broad, parallel, guided searches and very large result sets involving many thousands of documents to understand trends or new developments in the issue space. Tailored dimensions can be used to create bins and sub-bins to build a picture of interesting subject areas. By adding dimensions, for example, for time and for countries, trends are visible using the linked views to highlight, for example, amount of activity in diverse technology areas over time and/or by country. In Figure 14, in the bottom pane, one technology type has been selected (in blue), and then using linked selection with sort-by-selection, trends can be seen in the country and time dimension views just above.

5.7 Entity Relations

One of the dimensions specializes in exploring and scanning entity relations and makes use of Fair Isaac computationally extracted entity relations. Information level contact chaining exploration is possible. For example, the analyst might be curious about a person mentioned in a few relevant documents. Dragging that person into this dimension might reveal connections to organizations not necessarily mentioned in the examined documents. The analyst may then follow her train of thought further and create a dimension for some of those organizations to quickly see who else is mentioned as working there and how they are related to the interesting person.

5.8 Integrated Document Reader

Promising documents can be skimmed, read, and annotated in the reader. A summary of the entities found in the document can be scanned, and points of interest in the document are indexed so the analyst can quickly jump there. Entities and query terms are highlighted.

5.9 Collecting Nuggets in the Sandbox

Whole documents and relevant fragments judged as relevant can simply be dragged into the Sandbox analysis space. The Sandbox is tightly integrated with TRIST and supports organizing information and thoughts freely [Wright et al, 2005]. Any relevant information saved in the Sandbox keeps its link to the documents and the queries it came from automatically. Selecting evidence in the Sandbox will highlight its source in all TRIST dimensions, quickly reminding the analyst about information on the source and facilitating credibility assessment of the evidence.

5.10 Supporting Analysis Science

TRIST, and nSpace, has been designed for integrating and evaluating new analysis science and technologies. It provides side-by-side comparison of alternative technologies and an integrated workspace to perform evaluations in a whole workflow context. Making performance visible will help improve IR technologies and provide insight about impact of tools and methods. TRIST now provides access to 'Google Web Search, 'Google Image Search', 'NCBI PubMed Search', Altarum/Sarnoff Ant Hill Investigations (based on queries derived from user modeling), and the ability to search a corpus of documents compiled by Fair Isaac and a related database of all the entities contains in those documents.

6 Technical Architecture

nSpace - TRIST has a multi-tier architecture for scalability and ease of deployment. Web Services standards are implemented to encapsulate the services in each tier and to provide scalability, modularity and data processing functionality. An Activity and Knowledge Base is maintained and accessed through the Application Services layer. Other background application processing, such as Search execution, can be offloaded to the Application Services layer to reduce load on the client.

7 Pilot Evaluation at NIST

In a two day experiment at NIST, six analysts used a pilot version of TRIST. Results showed that analyst work product quality increased, in half the time, while reading double the number of documents [Scholtz et al, 2004].

Aggregate quality rankings of the reports produced showed that some analysts with less domain expertise were able to outperform the baseline domain expert while using TRIST. Correlation was found between report score and two of the primary metrics: unique documents viewed and number of queries executed. Analysts who viewed more documents, and who executed more queries, tended to produce reports that were ranked higher.

Rank	Analyst Report
1	Report B
2	Report F
3	Report Expert
4	Report A
5	Report C
6	Report D
7	Report E

Table 1. Rank of Reports.

Analyst	Time	Searches	Time/Search	Docs. Read	Time/Doc.
A	220	9	24.4 mins	126	1.7 mins
B	254	26	9.8	284	0.9
C	209	9	23.2	162	1.3
D	258	4	64.5	159	1.6
E	148	3	49.3	165	0.9
F	200	13	15.4	182	1.1
Expert	540	40	13.5	88	6.0

Table 2: TRIST improves IR productivity.

Analysts with TRIST were compared to the baseline expert analyst who performed the same task using Google and MS Word. As shown in Table 2, TRIST users A to F spent less time on the task and read more documents. The analysts provided qualitative feedback about the concepts behind TRIST during the "hot-wash" discussion period:

> Searching *seemed* faster compared with Google. Searches happened asynchronously, so analysts were able to perform several activities at once.
>
> Much more information was able to be reviewed with less fatigue. The relentless list after list experience was avoided. It was easier and faster to identify relevant results.
>
> Analyst confidence in the analysis increased as more information was reviewed in less time.
>
> The Query Planning view helped with task decomposition and report building during the IR phase. Visible query trails saved time and contributed to organizing the work.
>
> TRIST was able to highlight items that may not have been apparent in Google.
>
> Analysts agreed their products were of higher quality with TRIST than with Google.

8 Conclusion

TRIST and the Sandbox are the first of two components of nSpace, a fluid, flexible medium of analysis and expression. Quick understanding using multiple dimensions, visible correlations and linked selection allows efficient triage of massive data. Results show a measurable, significant increase in analyst productivity particularly when dealing with massive data. These tools will continue to evolve as a framework for supporting the whole analytic workflow and integrating new agent technology. Collaborations with analysts and technology partners are essential for success.

New challenges include working at the information and cognition level, exploring new cognitive interactions and new information visualizations that interact with the analyst as the analyst interacts with the information space. Further productivity increases are expected as TRIST moves to a more information-focused analysis.

Acknowledgement

This study was supported and monitored by the Advanced Research and Development Activity (ARDA) and the National Geospatial-Intelligence Agency (NGA) under Contract Number NMA401-02-C0032. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official Department of Defense position, policy, or decision. The authors wish to thank the ARDA NIMD Program, and all ARDA staff for their support and encouragement.

9 References

Andrews, K., C. G¨utl, J. Moser, V. Sabol, and W. Lackner, Search Result Visualisation with xFIND, In Proc. UIDIS 2001, pages 50–58, Zurich, Switzerland, 2001.

ARDA, Novel Intelligence From Massive Data, NIMD, http://www.ic-arda.org/Novel_Intelligence/, 2002.

Caid, W. and Pu Oing, System and Method of Context Vector Generation + Retrieval, US patent 5,619,709, 1997.

Cousins, S., A. Paepcke, T. Winograd, E. Bier, and K. Pier, The Digital Library Integrated Task Environment (DLITE), Proc. of the 2nd ACM International Conf. on Digital Libraries, 1997.

Cowley PJ, Greitzer FL, and Hampson E, Glass Box: Progress and Plans, Technical Report for NIMD, 2004.

Bates, M., Design of Browsing and Berry Picking Techniques for the On-Line Search Interface, Online Review, v13, 1989.

Chen, C., Visual Spatial Thinking in Digital Libraries – Top Ten Problems, Conf. on Digital Libraries, 2001.

Dumais, S., E. Cutrell, J. Cadiz, G. Jancke, R. Sarin, D. Robbins, Stuff I've Seen: A System for Personal Information Retrieval and Re-Use, ACM SIGIR'03, Toronto, 2003.

Eick, Stephen and G. Wills, High Interaction Graphics, European Journal of Operations Research, Mar. 1995.

Geroimenko, V. and C. Chen, Visualizing the Semantic Web, Springer-Verlag, London, 2003.

Havre, S., E. Hetzler, K. Perrine, E. Jurrus, and N. Miller, Interactive Visualization of Multiple Query Results, Proceedings of the IEEE Information Visualization Symposium, 2001.

Hearst, M., User Interfaces and Visualization, Chapter 10 in Ricardo Baeza-Yates, Berthier Ribeiro-Neto, Modern Information Retrieval, Addison-Wesley-Longman, 1999.

Nowell, L., R. France, D. Hix, L. Heath, and E. Fox, Visualizing Search Results: Some Alternatives to Query-Document Similarity, In Proc. SIGIR'96, pages 67–75, Zurich, 1996.

Pirolli, Peter, Stuart Card (1999). Information Foraging. Psychology Review Vol. 106, No. 4. (pp.643-675)

Presearch, http://www.presearch-inc.com/pf.htm, 2005.

Rose, Russ, Chair P1000 Committee, P1000 Report, 1996.

Scholtz, J. et al, Pilot Evaluation of TRIST, Internal NIST Technical Report, February, 2004.

Schraagen, J.M.C. et al, State of the Art Review of Cognitive Task Analysis Techniques, in Cognitive Task Analysis, NATO RTO Technical Report 24, 2000.

Wise, J., J. Thomas, K. Bennock, M. Pottier, A. Schur, and V. Crow, Visualizing the Non-visual: Spatial Analysis and Interaction with Information from Documents, Proc. InfoVis 95.

Wright, W. and Kapler, T, Speaking with Analysts – Observations of Current Practices with Massive Data, submitted to Journal of Intelligence Community Research and Development, 2004.

Wright, W., D. Schroh, P. Proulx, B. Cort and D. Jonker, Advances in nSpace – The Sandbox for Analysis, Poster, Conference on Intelligence Analysis, 2005.