A Distant Symmetry

Final Report: U.S. Intelligence Community Tools

Amelia Acker & Katie Shilton

8-08-10

Contents

I. Introduction

Mass digitization of the collections of the great research libraries together with a generation of text conversion and editing projects have created a critical mass of data in digital form that is highly heterogeneous in level of markup, format, and language. This confluence of conversion activities and advances in technology offers humanities scholars an opportunity to examine questions that require scale and computational power but challenges them to deal with new tools and technologies as well as with sources of varying quality and perhaps unknown provenance.

As the necessity for a cyberinfrastructure to support these new forms of scholarship becomes apparent, digital humanities scholars have begun comparing themselves to analogous research in other fields. Borgman (2009), for example, compares digital humanities research practices to those in the relatively successful e-sciences. This project makes a similar comparison, but to a very different field: that of United States intelligence gathering. The intelligence community faces many of the same challenges as digital humanists, including processing vast corpuses of texts and translating and working in multiple languages. What can the digital humanities community learn from the make-up, tool building, and infrastructure of the intelligence community?

II. Research Questions

This project focuses on two questions designed to assess some of the similarities and differences between digital humanities and tool building in the intelligence communities. Our research asks:

1. What tools exist in the intelligence communities, and are they accessible to humanities scholars?
2. If a scholar can find a link to the tool, is there sufficient information for the investigator to understand its requirements as well as the relevance of the tool to the proposed research?

III. Methods

During the first half of the Intelligence Tools project (September – December 2009), we mapped the landscape of the intelligence community to aid our search for digital research tools. We also consulted with two key informants familiar with the intelligence and data mining communities to discern sources and keywords that might aid our search. We simultaneously performed a search of the intelligence literature to ground ourselves in the structure and infrastructure of the intelligence community. After creating a map of the web of agencies and federally-funded academic centers, we searched each entity website for shared digital tools.

Key informants

As we mapped the infrastructure of the US Intelligence community, we also spoke with two key informants who provided background and suggestions for places to look for digital research tools. Dr. Michael Welge, Research Scientist at University of Illinois/NCSA, and Dr. Douglas Maughan, Program Manager of the Cyber Security R&D Center at the Department of Homeland Security, both provided ideas for organizations building digital tools, as well as suggestions for language we should be using (e.g. "data analysis" instead of "data mining").

Douglas Maughan, in particular, clued us in to a problem that has been reinforced in our searching, and which we will discuss further below. Many intelligence tools are built on the "dark web" – places unconnected to the World Wide Web for security or secrecy purposes. Maughan also emphasized that sharing and reuse is not a priority in the intelligence community. Many agencies develop software either in-house or under private contracts, with little sharing or repurposing.

Literature search

Reference interviews with subject librarians at UCLA revealed some important suggestions for searching for digital technology within intelligence community. For instance, the phrase "digital tool(s)" is used infrequently to describe electronic resources in academic article databases and has yet to appear in the Library of Congress Authorities. "Intelligence service" is preferred over "intelligence community" and "data fusion" often produces artificial intelligence articles focused on "multisensor data fusion." Much of our time with reference librarians was spent defining and locating appropriate keywords, subject headings, and subject areas before beginning a thorough literature review. Initially, we searched social science subject specific databases, but after compiling a list of synonyms (or close matches) to "digital tools" we found that searching technology databases yielded a variety of sources as well.

As a result of our reference interviews and personal subject searches, we compiled a list of keywords, subjects, and descriptors to apply to database and general web searches. The terms we used (in combination with one another) were:

- analytics
- data analysis
- data fusion
- data integration
- data management
- data mining
- digital resources
- tool
- intelligence community
- intelligence service
- middleware
- open source
- software
- visual analytics
- visualization

Using our refined search list, we carried out a two-part literature review. First, we created a series of searches combining terms like data fusion or data analysis; tool or software, and an intelligence agency. For instance: "Defense Information Systems Agency" w/p "data fusion" + "tool" would retrieve any text with all three search terms within a paragraph from a full-text query of a subject-specific database. We used these search conventions in subject specific bibliographic guides and full-text search catalogs, including Columbia University's Information Resources Guide for the U.S. Intelligence Community; Lexis Nexis Congressional Advanced for relevant reports and testimonies featured in congressional daily record; and PAIS International for applicable conference proceedings, government publications, and public policy coverage related to intelligence in the U.S. We also searched the databases of three major government news sources that publish daily articles on government technology trends, including *Government Computer News*, *Washington Technology*, and *Federal Computer Week*. We looked at recent government publications, including the *2009 National Intelligence Strategy*, and found one of the salient suggestions was to integrate cyber expertise throughout agencies, allied intelligence services, industry, and academia. We also made use of the 'glossary' section of the *Guideline for Identifying an Information System as a National Security System* to further inform our subject-term lists (Barker, 2003).

The second part of the literature survey involved finding academic publications to provide insight into the intelligences service's attitudes towards tool building, dispersal, and sharing. We consulted *Covert and Overt: Recollecting and Connecting Intelligence Service and Information Science* (Williams & Lipetz, 2005); *Spying Blind: The CIA, the FBI, and the origins of 9/11* (Zegart, 2008); and Robert Williams' (2005) book chapter "The Information Science and Intelligence Service Literature: An Overview." A thorough citation analysis of these books' bibliographies listed further pertinent resources, including relevant government publications, newspaper articles, and recent testimonies before congress. A complete list of sources cited is provided in our bibliography.

Tool evaluation

Our literature and web searches finally revealed both evidence of "dark" or inaccessible tools, as well as links to accessible, shared tools. We discuss some of the implications of this in "Tool Findings," below. While federal agencies revealed few accessible tools, academic centers funded by the intelligence community boast many openly available tools. Our search also produced 41 accessible tools for evaluation in the second part of the project. A list and description of tools found is included in Appendix 1.

The project team evaluated each of the 41 tools according to the methods developed by Nguyen and Shilton (2008, 2009). The method is designed to evaluate the following questions:

- How easy is it to access DHC tools?
- How clear are the intentions and functions of DHC tools?

Based on these questions, we created two scales:

- Ease of access
- Clarity of use
- Sustainability

The scales respond to five variables: (1) identification of tool, (2) feature, display, and access, (3) clarity of description; (4) clarity of operation; and (5) long-term sustainability. To construct measurable scales, we divided the variables into distinct indicators that we could rank as poor, moderate, or excellent. The indicators and variables are listed in Table 1. More detail on the development of these indicators and variables can be found in Nguyen and Shilton (2008) and Shilton (2009).

Table 1: Indicators and variables

Variable	Component	Poor	Moderate	Excellent
Identification of tools	Word choice	Use of broader term	Use of narrower term	Use of the term "tool"
	Visibility on page	Buried within body of text	Moderately visible	Highly visible
Feature, Display, and Access	Tool placement within website	Buried under multiple pages (clicks)	2-click	1-click
	Downloading	Download link separated from tool description		Download link embedded in tool description
	Uploading	Link to upload dataset/resources separated from tool description		Link to upload dataset/resources embedded in tool description
Clarity of description	Function	Function of tool not stated	Function of tool difficult to understand	Function of tool clearly stated in an easy to understand manner
	User Group	Intended user groups not stated at all	User group difficult to understand	Intended user groups clearly stated (by subject, age, discipline, etc)
Clarity of operation	Preview	Tool can not be previewed in any manner	Tool can be previewed via screenshots only	Tool can be previewed via demos
	Technical requirements	Operating system requirements/limitations not provided	Operating system requirements/limitations are murky, hard to find, buried on	Clear and concise operating system requirements/limitations provided

			page	
	Technical requirements – additional software	The tool requires additional software HOWEVER does not provide clear statements about these requirements, nor does it provide direct links to the additional software, nor instructions on accessing and installing	[any 2 out of these 3] Clear descriptions on additional requirements Direct links to additional software Instructions on accessing and installing software	The tool does not require any additional software to run -or- The DHC provides clear statements on additional requirements, while providing direct links AND instructions on accessing and installing additional software requirements
	Instructions for download	No instructions are provided on how to download a tool	Instructions are either difficult to understand or not readily accessible	Clear and easy to understand instructions on how to download the tool are provided and readily accessible
	Instructions for data import or upload	No instructions are provided on how to connect data or resources to a tool	Instructions are either difficult to understand or not readily accessible	Clear and easy to understand instructions are provided and readily accessible
Sustainability	Age of the tool	0-3 years	4-6 years	7-10
	Versioning information	Version number		
	Evidence of ongoing support	No evidence	Evidence that a tool is discontinued	Evidence that a tool is active
	Open source?	Unclear	No	Yes
	Use of open-source development tools (SourceForge, etc.)	Unclear	No	Yes
	Responsibility for a tool	None identified	1 role identified	2 or more roles identified

We began the tool evaluation process by applying ratings to five tools simultaneously, and then comparing our rankings for inter-indexer consistency. We discussed our differences, came to agreement on definitions and ratings, and then simultaneously coded five additional tools. We achieved 90% inter-indexer consistency on this second set of tools. We then divided the remaining tools and each coded half.

Study limitations

The first stage of our work produced a series of questions that we reviewed with our supervisors at CLIR. Our first set of challenges had to do with the definition of tools. What was clearly defined in the digital humanities space (Nguyen & Shilton, 2008) seems less clear in the broader intelligence community. Specifically, we wondered if CLIR were interested in analyzing tools meant for purposes such as disaster preparedness, health analysis, or information assurance. Similarly, we debated whether we should investigate operating systems, programming languages, and other middleware built for use in the intelligence community.

After speaking with our supervisors at CLIR, we decided that such tools were all out of scope. While we have chosen to draw attention to the existence of such tools as we come across them, we will not add them to our list for analysis. We decided our efforts are best spent on tools that are data-ready (so, not tools designed to help build other tools, like middleware). Additionally, we should delineate tools that scholars might discover and that could logically be repurposed for humanities research goals. Tools that help researchers work with text at scale, visualize a corpus, or map data may be useful for a broad array of digital humanities research activities. Tools for disaster preparedness, health analysis and information assurance do not repurpose as clearly as tools for text mining, translation, visualization, and GIS. In addition, tools built by private organizations funded by In-Q-Tel, the venture capital arm of the intelligence community, do not appear in this study. Such tools might include such broadly familiar tools as Facebook, and would be a large and diverse (but possibly off topic) pool.

Similarly, we decided not to pursue tools beyond those easily available online. Such pursuit would make our target too broad. For the purposes of this project, we assume that because an average humanities researcher would not be able to access such tools, they should be excluded from our evaluation. However, we did make an effort to note any tools that might be available through networking or conversations with relevant academics or intelligence contacts. Such tools might be more easily available to researchers than classified or other dark tools.

IV. Literature Review Findings

Our review of intelligence literature yielded a number of insights applicable to our study. These included detailing the information technology needs of the intelligence community, helping to map the intelligence landscape, and explaining the relationship of the intelligence community with tool authorship and sharing.

Intelligence information technology needs

A review of information needs in the U.S. intelligence community by a researcher in library and information sciences (Marling, 2005) provided a concise list of needs to which intelligence agencies might address digital tools. While surely not comprehensive, this list provides a starting place to understand the sorts of tools the intelligence agencies might develop. These needs include systems for indexing and tagging; search engines for media as well as text and the deep web; language translation tools; speech recognition tools; real-time or breaking news aggregation; natural-language processing tools; entity extraction techniques; text mining; tools for summarizing and distilling large corpora; tools for link analysis and pattern recognition; and data visualization techniques (Marling, 2005). Many of these map to similar needs in the digital humanities community (Friedlander, 2009).

Mapping the intelligence landscape

Our literature search also allowed us to map the complex web of agencies and funded academic laboratories that make up the U.S. intelligence community. The mapping exercise allowed us to delineate places where digital tools of interest might be found, organize our search, and ensure that didn't overlook any potential tool sources. We constructed a map of intelligence agencies and affiliates by beginning with the organizational structure of the U.S. intelligence community. We used a helpful organizational chart provided by the documentary series *Frontline*, which can be found on the web at*:* http://www.pbs.org/wgbh/pages/frontline/darkside/etc/cia.html (Frontline, 2006). To this basic structure, we added National Laboratories, Department of Homeland Security (DHS) Centers of Excellence, Federally Funded Research and Development Centers (supported by the DHS, Department of Energy and Department of Defense), and other intelligence-funded units recommended to us by our key informants. Each of the nodes on this new organizational chart (Appendix 2) gave us a place to start searching for publicly-available tools developed by the intelligence community.

The relationship between intelligence work and tools

Our literature review also illuminated the relationship of the intelligence community to tools and software building, as well as to sharing and secrecy. It is clear that the intelligence community does not have the same culture of sharing, or publishing, as the academic community. Our project is premised on an academic community model, in which incentives exist for the sharing of publications, and increasingly digital tools and even digital data (Borgman, 2007). But the culture of U.S. intelligence publication and sharing is different in critical ways that clearly affect the dissemination and availability of digital tools.

Information sharing across agencies is an acknowledged problem in the intelligence community (Thompson, 2006). Secrecy, after all, is an important feature of intelligence. The U.S. intelligence service has strict rules about classification, information restriction, and clearance for resource access. Such hierarchies of information sharing necessitate role-based access control for many digital sharing tools. Designing and deploying tools utilizing state-of-the-art security and featuring complex access hierarchies requires both skill and customization (Thompson, 2006).

The culture of publishing within intelligence agencies has historically been centered upon the authorship and circulation of intelligence reports. Within agencies, such widely circulated intelligence reports are one of the few concrete and distinguishable ways that analysts can be promoted (Thompson, 2006). Until recently, raw information gathered and synthesized for such reports has not been networked between agencies (or even *within* agencies). Instead, collected intelligence has been stored in agency-specific 'air-gapped' data warehouses. And until the late 1990s, most intelligence hardware was based on cold-war information systems. Highly structured channels and walls that existed throughout organizations like the CIA and the FBI, coupled with legal and technological limitations between agencies, impeding elevant information exchanges.

Recent news sources we consulted clearly show a recent and developing trend towards changing this 'air-gapped' culture. Intelligence agencies are beginning to implement social networking technologies to foster information exchange among and within agencies (Beizer, 2009). But this shift in ethos is not broadly adopted. It appears that only officials and analysts at the 'top' and 'bottom' want to steer intelligence agencies toward Web 2.0, cloud computing and social intelligence technologies. Many other intelligence employees are highly skeptical of the new technologies. Employees known as "iron majors" (25-year career agents), who make up the bulk of the intelligence bureaus, may be wary of technical change. Experienced agents, who rose under the values of the traditional culture, are unsure if openness and collectivity outweigh the risks of co-locating vulnerable swaths of classified information.

Additionally, many new Web 2.0 and social networking technologies obscure authorship and hamper the ability to trace the provenance of editing and functions. When speaking in terms of the intelligence community's traditional publishing model, the utility of collaboration in most Web 2.0 technologies is both a poison and a cure for the conventions of the 'traditional' intelligence report. How can an agency assess an analyst's performance when producing a collaborative intelligence resource?

Substantial evidence shows that before 2001, the U.S. intelligence community had neither the intention nor the infrastructure to share, exchange, pool, and network intelligence information (Zegart, 2007b). Since then, the government has established centers and initiatives to promote intelligence-sharing and coordination efforts (Bain, 2009a). Many of these efforts have relied on independent contractors to build digital tools for the community. In fact, the intelligence community may see itself primarily as a software consumer, not a software builder. For example, federal regulations dictate that government bodies must give preference to commercial software, including open source (Jackson, 2009b). There are ongoing internal debates as to whether it is within agency mandates (as well as wise and economical) to develop software in-house (Jackson, 2009a; Welsh, 2009). News articles track the many private firms contracting with intelligence agencies to provide information technology and tool-building services. Examples include translation services (Wakeman, 2008) and Intellipedia (Beizer, 2008b), a classified wiki encyclopedia for the intelligence community. The CIA also has an investment arm, In-Q-Tel, specifically devoted to funding private technology development (Wakeman, 2008).

Though these tools may be under development in the private sector, it remains to be seen how and if these tools will be used, and if agents and analysts alike will trust these tools. The publishing model of Web 2.0, blogs, and wikis is quite opposed to the traditional intelligence community's organizational concept of intelligence reporting, because it is based upon intelligence sharing. As many scholars have pointed out, the adoption of these coordinated technologies will depend on "bottom-up cultural transformation as well as top-down policy changes" (Zegart, 2007a).

V. Tool Evaluation Findings

Evidence of inaccessible and "dark" tools

As our term implies, evidence of many "dark" or inaccessible tools was shadowy at best. It is clear that the U.S. intelligence community is building many digital tools that might be of interest to the humanities community, but most of these tools remain inaccessible. Few federal intelligence agencies seem to make their tools openly available. Whether this is an oversight (due to institutional cultures that value secrecy over sharing) or purposeful (classified tools like A-Space, the "Facebook for Analysts" contain classified material are not available to the public) (Bain, 2008) remains unclear.

The largest evidence of this tool-building activity is the existence of Forge.mil, a system that "enables the collaborative development and use of open source and DoD community source software" (http://www.disa.mil/forge/.) Forge.mil is quite obviously modeled on SourceForge.net, a repository of code for open source tools ("SourceForge.net," 2009). However, access to Forge.mil is restricted to employees of federal agencies with the right levels of clearance.

We found another source of evidence of inaccessible tools in newspaper articles detailing private contracts to develop tools. Featured largely in news sources like *Federal Computer Week* and *Washington Technology*, these articles point to a variety of digital tools for information processing under development (Beizer, 2008b, 2008a; Jackson, 2009c; Lipowicz, 2008, 2007, 2006; Wakeman, 2008). A partial list of inaccessible, closed, or dark tools is included in Appendix 3. This is provided for purposes of illustrating the range of tools under development by the U.S. intelligence community, and in no way is meant to be comprehensive.

Accessible tools evaluation

We completed evaluation of 41 intelligence community tools on the variables of identification; feature, display and access; clarity of description; clarity of operation, and sustainability. Our analysis yielded both qualitative and quantitative observations.

Our first qualitative observation is that the evaluation tool developed by Nguyen and Shilton seems to work well for tools developed outside of the humanities community. Neither team member experienced problems applying the evaluation instrument to intelligence community tools. One area we feel needs further examination is the relationship between a tool's sustainability and the nature of its stewardship. Does it matter if one person or a group of team members sustain the tools? Do the roles and responsibilities of stewards shift over time? Shilton's original report on sustainability was within a Digital Humanities Center context, and assumed that a team was more likely than an individual to be able to sustain a tool over time. However, several of the best (openly available) intelligence community tools seem to be the scholarly work of single researchers. Perhaps when a tool is part of a researcher's academic oeuvre, and counts as part of their scholarship, it is more likely to be updated and supported over time. On the other hand, are tools built by graduate students or postdoctoral scholars abandoned when these researchers move to other institutions?

We also noticed that a number of the tools we evaluated were part of larger toolkits. It was difficult to evaluate the toolkit components as single tools, because they were often sparsely described and difficult to comprehend without exploring the entire suite. Individual tool descriptions would frequently rely on familiarity with the utilities of other tools in the kit. Consequently, some tools that are part of broader toolkits may suffer on clarity of description and clarity of operations rankings. Our evaluation methods may need to be adapted slightly for tools that are part of larger suites, because they are currently not designed to rate descriptions of many tools at once.

Our evaluation also turned up an expectation on the part of many tool builders of advanced software literacy. Working knowledge of languages such as java is required to use many of the tools we evaluated. Many of the tools buried descriptive and operations information in "ReadMe" text files included in the tool download package. Humanists looking for tools of interest may need to dig for descriptive information, and need to possess a certain amount of software acumen to find this information.

Data analysis

Once coded using our scheme, the 41 intelligence tools ranged in total score from 38 to 18 points. We calculated the mean and standard deviation for the tools' total scores (x=28, sd=5). We then used the standard deviation to analyze the overall distribution of the tools and construct three tool groups (lowest quartile, middle two quartiles, and top quartile). Tools scoring 33 points or above were categorized within the highest-scoring group. Those scoring between 23 and 32 points were grouped within the middle-scoring set. Those scoring 22 points or less were categorized within the lowest-scoring group. The highest-scoring group comprised 11 tools, the middle group 22 tools, and the bottom group 8 tools. Organizing the tools into three distinct sets allowed us to average individual variable scores within each set. This provided a comparison of the major differences among the groups. Table 2 below provides an overview of which tools fell within which groups.

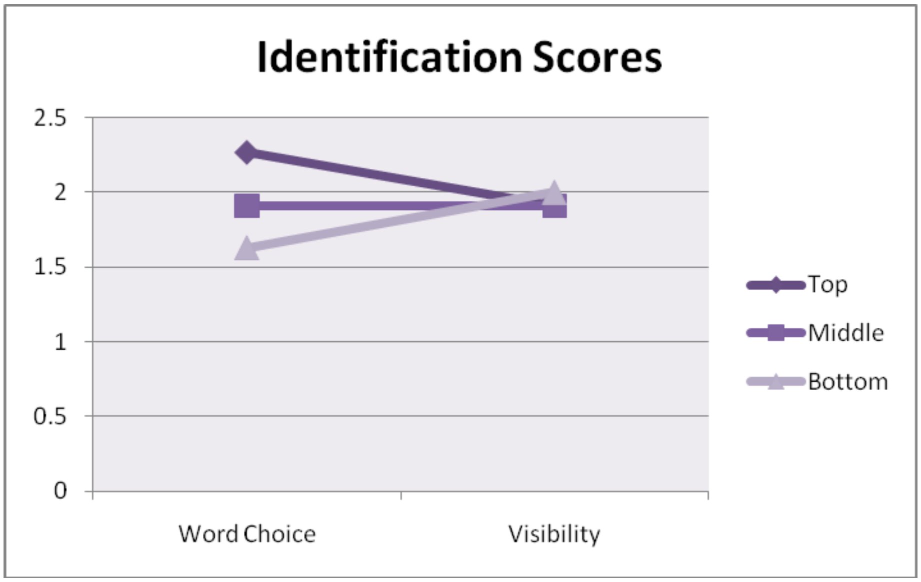
Table 2: Tool groupings

Top Group	Score
GeoViz Toolkit	38
Food and Agriculture Criticality Assessment Tool (FASCAT)	37
GeoVISTA Studio	36
Semantic Role Labeler (SRL)	35
Bioinformatics Resource Manager Software	35
Improvise	35
Unstructured Information Management Architecture (UIMA)	34
Color Brewer	34
ESTAT (Exploratory Spatial-Temporal Analysis Toolkit	34
pySNoW	33
Venn Diagram Plotter	33

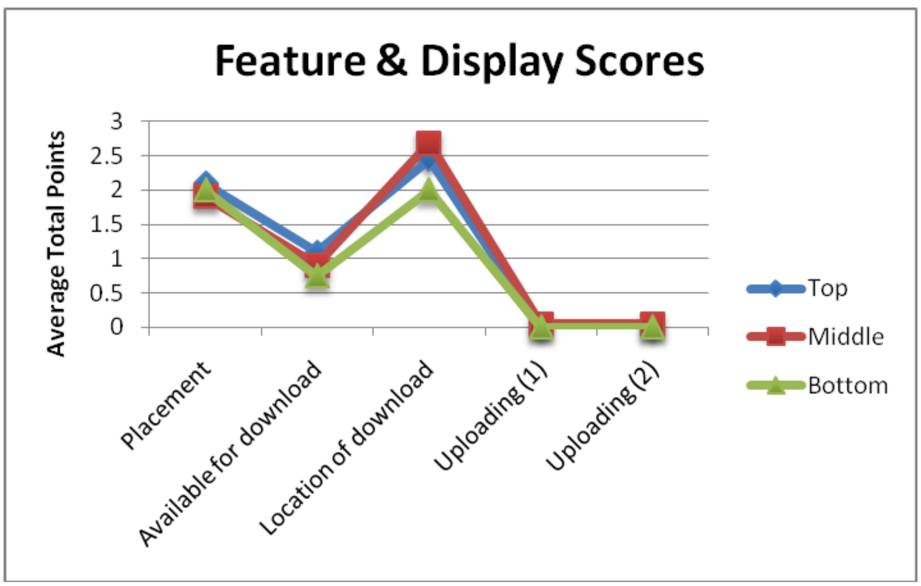
Middle Group	Score
PHARAOH	31
G-Ex Portal	30
Pandemic Visualization (PanViz) Toolkit	29
Tiburon	29
Visual Integration for Bayesian Evaluation (VIBE)	29
Learning Based Java (LBJ)	29
DAnTE	29
LBJ-based Coreference Package	28
Synthetic Syndromic Surveillance Data Creation Toolkit (SYDOVAT)	28
LBJ Chunker	27
Relational Feature Extraction Language (FEX)	27
LBJ Part of Speech Tagger	26
SNoW-based Part of Speech Tagger	26
Global Arrays Toolkit	26
SNoW-based Shallow Parser	25
SVM Technique for Evaluating Proteotypic Peptides (STEPP)	25
Electronic Mass Casualty Assessment and Planning Scenarios	25
SNoW Learning Architecture	25
LBJ Named Entity Tagger	24
ConceptVista	24
Visual Inquiry Toolkit	24
CoRanker	23

Bottom Group	Score
Jigsaw	22
Listed First-Order Probabilistic Inference, versions 1.0a and 1.0b	21
National Bio-Surveillance Integration System (NBIS)	21
Middleware for Data Instensive Computing (MeDICi) Integration Framework	21
ReWrite Decoder	20
Car Detection Software	19
MPQA Opinion Corpus, versions 1.2 and 2.0	18
Name Entity Tagger	18

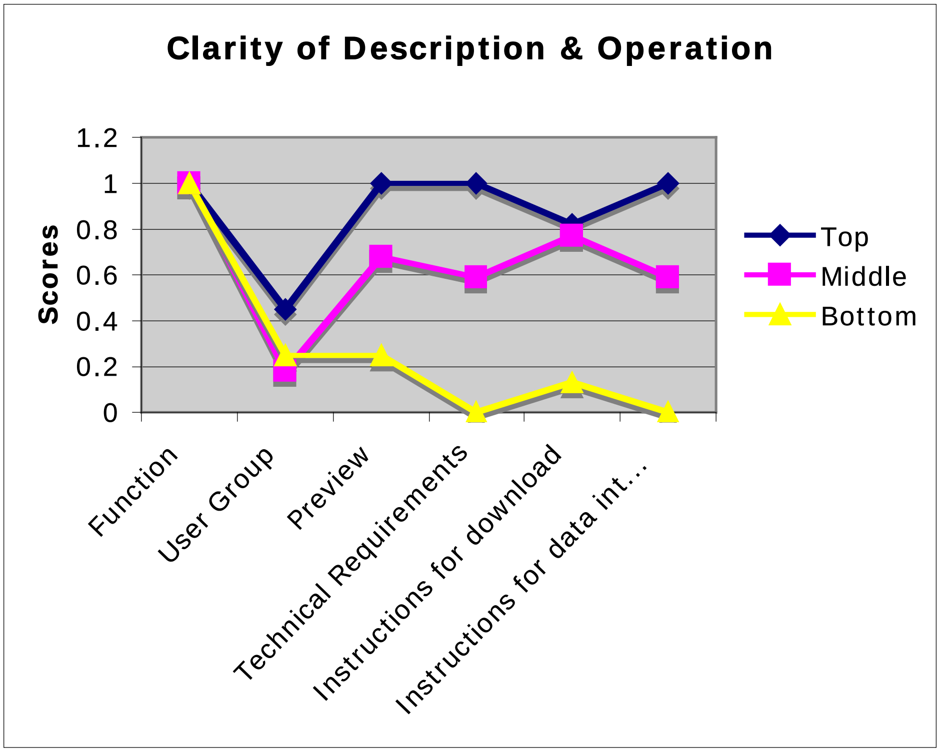
When broken down by category, the differences between top, middle, and bottom-ranking tools become apparent.



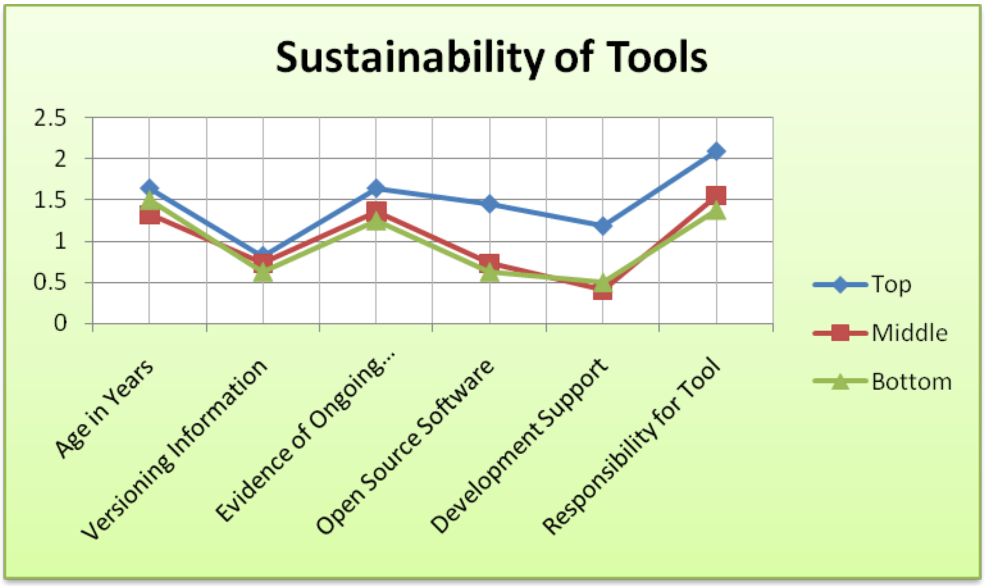
Identification scores: The visibility of a tool is almost identical across the top, middle, and low range of tools. It is word choice where a distinguishable difference in quality can be seen. This points to the wide variety of terms by which these tools are identified. *Software*, *kits*, and *applications* are all employed as descriptors. This finding is quite similar to the findings of the initial humanities report – there is little existing agreement on tool terminology.



Feature and display scores: The closeness of feature and display scores show that these were not a distinguishing feature between top, middle, and bottom-scoring tools. The fact that all tools were relatively well-placed, and tool download pages were easy to locate, suggests basic web design literacy among tool developers and intelligence site web hosts. Interestingly, no tools surveyed accepted data upload. This is similar to our findings in the initial humanities tool survey.



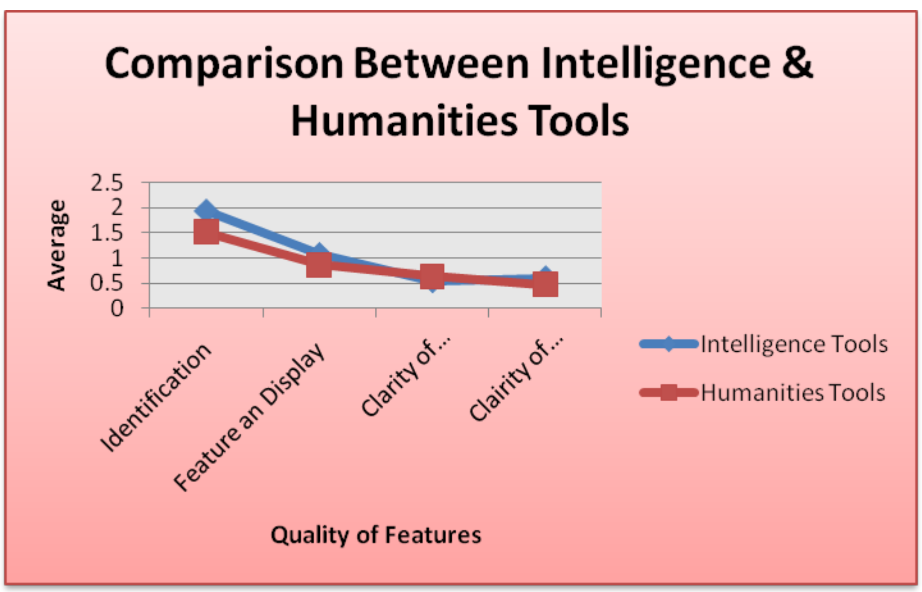
Clarity of description and operation scores: We combined the most notable features from clarity of description and operation variables. (We excluded variables on which tools did not vary significantly: additional software and additional technical requirements). The remaining variables illustrated here are clarity of function, user group, preview of tool, technical requirements, instructions for download, and data interaction. The notable distinction between top tools and the middle/bottom groups were availability of previews, enumeration of necessary technical requirements, and clear and easily available instructions for data interaction. Several tools suffered in these categories by obscuring this kind of information in 'ReadMe' text files embedded within download packages, rather than making directions directly available on the web.



Sustainability scores: The major differences in tool sustainability were in the categories of evidence of ongoing support, the use of open source protocols and development support, and demarcation of responsibility for the tools. Top scoring tools used open source protocols and community development support tools such as Sourceforge. Top tools also benefitted from clearly demarcated tool creators, distributors, and stewards.

Comparisons with humanities tools

When averaged, intelligence tools scored higher on the variables of identification (1.94 vs. 1.51), feature and display (1.07 vs. 0.86), and clarity of operation (0.60 vs. 0.47) than humanities tools.[1]



However, humanities tools scored slightly higher on clarity of description (0.63 vs. 0.54) when all scores were averaged. So while intelligence tools might be said to be slightly more findable and usable, these differences are slight. Intelligence tools produced by the academic community on the whole suffer from some of the same problems of those produced by humanities researchers: low visibility of, and obscured access to, downloading and uploading features; and problems with discerning and acting upon a tool's intended use.

[1] Sustainability was the only ranking that could not be easily compared, because fewer sustainability metrics were used in the intelligence tools evaluation.

VII. Next steps

Next steps for this project could go in a variety of directions. Now that the evaluation tool has been used on two separate sets of tools, it may be worth thinking about how the tool might be used in other fields. Outreach to communities who might be interested in this evaluation tool for their own disciplines would be a worthwhile next step. One way that we plan to accomplish this is to publicize the evaluation tool in an academic journal article. The workshop that accompanies this project may also provide a space to brainstorm outreach efforts.

We also suggest that academic libraries might find the evaluation tool of interest as they become repositories for tools on their campuses. Libraries have a history of focus on findability and sustainability. We have yet to search libraries for tools to evaluate, but libraries seem like a natural home for tools for scholarship, whether from the digital humanities or the intelligence community.

Another research avenue would be to fine-tune portions of the evaluation instrument to appraise discipline-specific features of digital tools for further studies. This would involve tailoring sections of the evaluation schema to fit other genres of tools (e.g. GIS tools, e-social science or e-science tools) while maintaining general indicators for broad comparison across genres. For example, the high correlation between all three groups of tools in the *feature and display* scores may indicate that alternative display characteristics should be considered for intelligence tools. Similarly, display characteristics might be evaluated slightly differently for tools emerging from the e-science or GIS communities. Additional evaluation instrument sections specifically tailored to different areas of scholarship, their needs, and the nature of their tools, might enhance the ability of the evaluation instrument to cross disciplines.

A different direction for further research might focus on expanding cooperation between intelligence and digital humanities toolmakers. Efforts to bring together intelligence and humanities research communities, through invited workshops or mini-conferences, would allow for increased cross-pollination and perhaps illumination of dark tools that would be useful for humanities projects. Another step would be to re-evaluate existing humanities tools using new criteria. Now that findability and understandability of these tools has been evaluated, metrics such as purpose, usability, kinds of data supported, infrastructure dependencies, and use cases would be helpful to humanities researchers. Complimentary work on evaluation categories for tools is underway at arts-humanities.net (http://www.arts-humanities.net/), and collaboration with this group might be worthwhile.

VIII. Conclusions

Our literature and tool search have made it clear that the task of locating, culling, deciphering, and using online tools built for the intelligence community would be a considerable challenge for humanities scholars. The lack of visibility for many of the tools—a mix of poorly described resources, unavailability of tools to the public, and purposeful classification—proved to be a challenge. It took significant time and research to locate a sufficient sample of tools to evaluate, and the available sample was limited to tools produced by academic institutions. In addition, the evaluation instrument made apparent tool builder's pervasive assumptions that intelligence tool users could employ a high degree of software literacy and navigate tool sites with little or no instruction. While there are certainly tools developed by the intelligence community that may be of use to humanities scholars, the present state of open access tools and their identification, access, clarity and sustainability make interdisciplinary sharing difficult to envision. We recommend that scholars and research institutions working on tool building, coordination and funding efforts in the future to consider these vital qualities of access and sharing across disciplines.

VIII. Bibliography

Bain, B. (2008, September 3). A-Space set to launch this month. *Federal Computer Week*. Retrieved from http://fcw.com/articles/2008/09/03/aspace-set-to-launch-this-month.aspx

Beizer, D. (2008a, February 14). CSC unveils location awareness tool. *Washington Technology*. Retrieved from http://washingtontechnology.com/articles/2008/02/14/csc-unveils-location-awareness-tool.aspx?sc_lang=en

Beizer, D. (2008b, June 26). A key to understanding data. *Washington Technology*. Retrieved from http://washingtontechnology.com/articles/2008/06/26/a-key-to-understanding-data.aspx?sc_lang=en

Borgman, C. L.The Digital Future is Now: A Call to Action for the Humanities. *Digital Humanities Quarterly*.

Borgman, C. L. (2007). *Scholarship in the digital age: information, infrastructure, and the internet*. Cambridge, MA and London: The MIT Press.

Friedlander, A. (2009). Asking Questions and Building a Research Agenda for Digital Scholarship. In *Working Together or Apart: Promoting the Next Generation of Digital Scholarship* (pp. 1-15). Washington, D.C.: Council on Library and Information Resources.

Frontline. (2006, June 20). The dark side: the changing intelligence community. Retrieved August 28, 2009, from http://www.pbs.org/wgbh/pages/frontline/darkside/etc/cia.html

Jackson, J. (2009a, September 16). DOD rethinks buy versus build software quandary -- Washington Technology. *Washington Technology*. Retrieved from http://washingtontechnology.com/articles/2009/09/15/dod-rethinking-buy-versus-build-software.aspx?sc_lang=en

Jackson, J. (2009b, October 30). DOD open-source memo could change software landscape. *Washington Technology*. Retrieved from http://washingtontechnology.com/Articles/2009/10/28/DoD-OSS-II.aspx?sc_lang=en&Page=1

Jackson, J. (2009c, November 20). A-Space melds social media and intelligence gathering. *Government Computer News*. Retrieved from http://gcn.com/Articles/2009/11/30/A-Space-DIA-intell-sharing-wiki.aspx?sc_lang=en&Page=1

Lipowicz, A. (2006, September 8). DHS plans major data fusion project. *Washington Technology*. Retrieved from http://washingtontechnology.com/articles/2006/09/08/dhs-plans-major-data-fusion-project.aspx?sc_lang=en

Lipowicz, A. (2007, August 14). Data mining program near rock bottom. *Washington Technology*. Retrieved from http://washingtontechnology.com/articles/2007/08/14/data-mining-program-near-rock-bottom.aspx?sc_lang=en

Lipowicz, A. (2008, February 4). ICE rolls out new data tool. *Washington Technology*. Retrieved from http://washingtontechnology.com/articles/2008/02/04/ice-rolls-out-new-data-tool.aspx?sc_lang=en

Marling, G. L. (2005). Technology for open source government information and business intelligence. In R. V. Williams & B. Lipetz (Eds.), *Covert and Overt: Recollecting and Conecting Intelligence Service and Information Science*, ASIS&T Monograph Series (pp. 129-145). Lanham, MD: The Scarecrow Press, Inc.

Nguyen, L., & Shilton, K. (2008). Tools for humanists. In D. Zorich (Ed.), *A Survey of Digital Humanities Centers in the United States*. Washington, D.C.: Council on Library and Information Resources.

SourceForge.net. (2009, November 11). *Wikipedia.org*. Retrieved December 6, 2009, from http://en.wikipedia.org/wiki/SourceForge

Thompson, C. (2006, December 3). Open-Source Spying. *The New York Times*. Retrieved from

http://www.nytimes.com/2006/12/03/magazine/03intelligence.html?_r=1
Wakeman, N. (2008, July 22). Translation services firm gets boost from In-Q-Tel. *Washington Technology*. Retrieved from http://washingtontechnology.com/articles/2008/07/22/translation-services-firm-gets-boost-from-inqtel.aspx?sc_lang=en
Welsh, W. (2009, September 30). DOD move to develop software stirs up storm. *Washington Technology*. Retrieved from http://washingtontechnology.com/articles/2009/09/23/software-development-storm.aspx?sc_lang=en
Zegart, A. (2007a July 8). Our Clueless Intelligence System. The Washington Post. Retrieved November 18, 2009 from http://www.washingtonpost.com/wp-dyn/content/article/2007/07/06/AR2007070602004_pf.html.
Zegart, A. (2007b). Spying Blind: The CIA, the FBI, and the Origins of 9/11. Princeton, N.J.: Princeton University Press.

APPENDIX 1

Tool Name	Website	University/Org	Affiliated Center/Lab	Funding Sources	Description
Unstructured Information Management Architecture (UIMA)	http://incubator.apache.org/uima/	Apache		Private donations to the Apache Software Foundation	A software system that analyze large volumes of unstructured information in order to discover knowledge that is relevant to an end user. An example UIM application might ingest plain text and identify entities, such as persons, places, organizations; or relations, such as works-for or located-at.
MPQA Opinion Corpus, versions 1.2 and 2.0	http://www.cs.pitt.edu/mpqa/databaserelease/	University of Pittsburgh, Cornell University, University of Utah		ARDA	A system that processes documents and automatically identifies subjective sentences as well as various aspects of subjectivity within sentences, including agents who are sources of opinion, direct subjective expressions and speech events, and sentiment expressions.
Jigsaw	http://www.cc.gatech.edu/gvu/ii/jigsaw/	Georgia Tech	Information Interfaces Laboratory	Department of Homeland Security, NSF	A visual analytics system to explore, analyze, and make sense of document collections. Jigsaw provides visualizations that portray different aspects of the documents, focused on presenting identifiable important entities (people, places, organizations, etc.) and their direct or indirect connections.

Car Detection Software	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=SPARSE_PARTS	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	Purpose unclear from web description
CoRanker	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=CORANKER	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	An algorithm for cross-lingual Named Entity discovery in bilingual weakly temporally aligned corpora.
LBJ Chunker	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=FLBJCHUNK	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	A chunker or ("shallow parser") is a program that partitions plain text into sequences of semantically related words.
LBJ Named Entity Tagger	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=FLBJNE	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	This is a state of the art NER tagger that tags plain text with named entitites (people / organizations / locations / miscellaneous). It uses gazetteers extracted from Wikipedia, word class model derived from unlabeled text and expressive non-local features.
LBJ Part of Speech Tagger	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=FLBJPOS	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	Identifies and tags parts of speech in a text corpus.

LBJ-based Coreference Package	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=FLBJCOREF	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	The coreference package finds matches among text in a corpus. It include gender and number match, WordNet relations including synonym, hypernym, and antonym, and ACE entity types (e.g. semantic classes such as person, organization, and geopolitical entity).
Learning Based Java (LBJ)	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=LBJ	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	Learning Based Java is a modeling language for the rapid development of software systems with one or more learned functions, designed for use with the JavaTM programming language. LBJ offers a convenient, declarative syntax for classifier and constraint definition directly in terms of the objects in the programmer's application.

Listed First-Order Probabilistic Inference, versions 1.0a and 1.0b	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=FOPI	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF, ARDA	This implementation of a Lifted First-order Probabilistic Inference algorithm allows the user to define a First-order probabilistic model and query it. Regular probabilistic inference algorithms operate on the propositional level (as opposed to first-order level when there is a notion of multiple objects).This tool seeks to provide a lifted first-order probabilistic inference algorithm that performs at a first-order level even during inference itself, and not only at the specification stage.
pySNoW	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=PYSNOW	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	pySnow is a minimal python interface to the SNoW - Sparse Network of Winnows learning architecture.
Relational Feature Extraction Language (FEX)	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=FEX	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	FEX is a tool for extracting features from text. These features can be used to generate examples for use with Machine Learning software such as SNoW.

Semantic Role Labeler (SRL)	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=SRL	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	Semantic Role Labeler is a machine-learning tool that analyzes for a shallow semantic information of a given sentence. The tool is capable of outputing verb-argument structure following the notation defined by the Propbank project.
SNoW Learning Architecture, versions 3.1.8 and 3.2.1	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=SNOW	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	The SNoW (Sparse Network of Winnows) learning architecture is a multi-class classifier that is specifically tailored for large scale learning tasks and fpr domains in which the potential number of features taking part in decisions is very large, but may be unknown a priori. It learns a sparse network of linear functions in which the targets concepts (class labels) are represented as linear functions over a common feature space.
SNoW-based Part of Speech Tagger	http://l2r.cs.uiuc.edu/~cogcomp/asoftware.php?skey=POS	University of Illinois Urbana-Champaign	Cognitive Computation Group	NSF	An older version of a part of speech tagger; replaced by the LBJ part of speech tagger.

PHARAOH	http://www.isi.edu/publications/licensed-sw/pharaoh/	University of Southern California	Information Sciences Institute	DARPA, Department of Energy, Department of Homeland Security	Pharaoh is a machine translation decoder released to the research community to aid research in statistical machine translation.
ReWrite Decoder	http://www.isi.edu/publications/licensed-sw/rewrite-decoder/	University of Southern California	Information Sciences Institute	DARPA, Department of Energy, Department of Homeland Security	The ISI ReWrite Decoder is a program that translates from one natural languge into another using statistical machine translation.
Tiburon	http://www.isi.edu/publications/licensed-sw/tiburon/	University of Southern California	Information Sciences Institute	DARPA, Department of Energy, Department of Homeland Security	Tiburon is a tree transducer package designed to handle weighted regular tree grammars, context-free grammars, and both tree-to-tree and tree-to-string transducers, and can perform composition, intersection, application, determinization, inside/outside training, pruning, and returning k-most likely trees.

Global Arrays Toolkit	http://www.pnl.gov/science/highlights/highlight.asp?groupid=853&curpage=2&id=654	Pacific Northwest National Laboratory		Department of Energy	The GA provides high-level interfaces for writing parallel programs, it supports users who do large-scale computational research and greatly simplifies the development of programs for unconventional architectures on distributed-memory computers.
Bioinformatics Resource Manager Software	http://www.sysbio.org/research/bsi/bioanalytics/bioinformatics.stm	Pacific Northwest National Laboratory		Department of Energy	BRM is a middleware architecture to link scientific applications and heterogeneous data source; the framework can be used to automate multidisciplinary data mining processes used for high-throughput molecular profiling.
Middleware for Data Instensive Computing (MeDICi)/ MeDICi Integration Framework (MIF), versions 2.1.2 and 3.0.1	http://medici.pnl.gov.download_mif	Pacific Northwest National Laboratory		Department of Energy	MeDICi is an evolving middleware platform for building complex, high performance analytical applications that comprise a pipeline of software components, each of which performs some analysis on incoming data and passes on results to the next step in the pipeline.

DAnTE	http://omics.pnl.gov/software/DAnTE.php	Pacific Northwest National Laboratory		Department of Energy	DAnTE allows users to perform various downstream data analysis, normalization, data reduction, and hypothesis testing steps; it includes a graphical environment for visualizing the data during the processing.
SVM Technique for Evaluating Proteotypic Peptides (STEPP)	http://omics.pnl.gov/software/STEPP.php	Pacific Northwest National Laboratory		Department of Energy	STEPP software contains an implementation of a trained Support Vector Machine that can compute a score representing how "proteotypic" a peptide is by LC-MS. The program can read a protein file, perform an in-silico digestion, and compute the observability score for each tryptic or partially tryptic peptide.
Venn Diagram Plotter	http://omics.pnl.gov/software/VennDiagramPlotter.php	Pacific Northwest National Laboratory		Department of Energy	This program will draw correctly proportioned and positioned Venn diagrams, supporting both 2-circle and 3-circle; it includes a graphical user interface and diagrams can be displayed, colors customized, and diagrams can be copied to the clipboard or saved to disk.

Visual Integration for Bayesian Evaluation (VIBE)	http://omics.pnl.gov/software/VIBE.php	Pacific Northwest National Laboratory		Department of Energy	VIBE software is a visualization tool that allows users to observe classification accuracies at the class level and evaluate classification accuracies on any subset of available data types based on the posterior probability models defined for the individual and integrated data.
Color Brewer	http://www.personal.psu.edu/cab38/ColorBrewer/ColorBrewer_intro.html	Pennsylvania State University	GeoVista Center, North-East Visualization and Analytics Center, Pacific Northwest National Laboratory	Department of Homeland Security	Color Brewer is an online diagnostic tool designed to help users select appropriate color schemes for maps, geospatial visualizations and other graphics.
ConceptVista	http://www.geovista.psu.edu/ConceptVISTA/index.jsp	Pennsylvania State University	GeoVista Center, North-East Visualization and Analytics Center, Pacific Northwest National Laboratory	Department of Homeland Security	ConceptVISTA is an ontology creation and visualization tool that allows users to define and link concepts and resources pertaining to a conceptual domain.
ESTAT (Exploratory Spatial-Temporal Analysis Toolkit	http://www.geovista.psu.edu/ESTAT/index.html	Pennsylvania State University	GeoVista Center, North-East Visualization and Analytics Center, Pacific Northwest National Laboratory	Department of Homeland Security	ESTAT toolkit provides user-friendly, open-source software designed to support exploratory geographic visualization

G-Ex Portal	http://www.geovista.psu.edu/G-EXPortal/index.html	Pennsylvania State University	GeoVista Center, North-East Visualization and Analytics Center, Pacific Northwest National Laboratory	Department of Homeland Security	The G-EX Portal is a web-based geocollaboration tool for the dissemination of geovisualization software and tools, learning artifacts supporting the use of these tools, and analysis artifacts generated during application of these tools.
GeoVISTA Studio	http://www.geovistastudio.psu.edu/jsp/index.jsp	Pennsylvania State University	GeoVista Center, North-East Visualization and Analytics Center, Pacific Northwest National Laboratory	Department of Homeland Security	GeoVISTA Studio is an open software development environment that allows users to quickly build applications for geocomputation and geographic visualization.
GeoViz Toolkit	http://www.geovista.psu.edu/geoviztoolkit/index.html	Pennsylvania State University	GeoVista Center, North-East Visualization and Analytics Center, Pacific Northwest National Laboratory	Department of Homeland Security	GeoViz Toolkit is an application to support analysis of multivariate relationships in geographic space.
Improvise	http://www.cs.ou.edu/~weaver/improvise/index.html	University of Oklahoma	Center for Spatial Analysis		Improvise is a fully-implemented Java software architecture and user interface that enables users to build and browse highly-coordinated visualizations interactively.

Visual Inquiry Toolkit	http://www.geovista.psu.edu/VIT/index.html	Pennsylvania State University	GeoVista Center, North-East Visualization and Analytics Center, Pacific Northwest National Laboratory	Department of Homeland Security	The Visual Inquiry Toolkit provides a visual programming environment for spatial data analysis; it integrates visual, computational, and cartographic methods for incremental searching patterns.
Pandemic Visualization (PanViz) Toolkit	http://pixel.ecn.purdue.edu:8080/~rmacieje/PanViz/	Purdue University	Purdue University Regional Visualization and Analytics Center	Department of Homeland Security	PanViz is a suite of visual analytic tools for analyzing the spread of pandemic influenza.
Synthetic Syndromic Surveillance Data Creation Toolkit (SYDOVAT)	http://pixel.ecn.purdue.edu:8080/~rmacieje/SYDOVAT-web/index.html	Purdue University	Purdue University Regional Visualization and Analytics Center	Department of Homeland Security	SYDOVAT provides researchers with a range of synthetic data sets from emergency departments that can be used in the evaluation of algorithms and methods dealing with multivariate, spatiotemporal data exploration, analysis and visualization.
Food and Agriculture Criticality Assessment Tool (FAS-CAT)	http://fazd.tamu.edu/products/information-analysis-systems/food-and-agriculture-criticality-assessment-tool-fascat.html/	Texas A&M University & Kansas State University	**National Center for Foreign Animal and Zoonotic Disease Defense**	Department of Homeland Security	FAS-CAT is an advanced Excel application that helps identify critical assets in the food and agriculture sector and provide reporting mechanisms to Homeland Security.

National Bio-Surveillance Integration System (NBIS)	http://fazd.tamu.edu/products/information-analysis-systems/national-bio-surveillance-integration-system-nbis.html/	Texas A&M University & Kansas State University	**National Center for Foreign Animal and Zoonotic Disease Defense**	Department of Homeland Security	NBIS is a dynamic display integration system that displays data from a variety of sources in a single window to provide increased situational awareness of biological and disease-related events.
Electronic Mass Casualty Assessment and Planning Scenarios	http://www.pacercenter.org/pages/about_emcaps.aspx	Johns Hopkins University	National Center for the Study of Preparedness & Catastrophic Event Response	Department of Homeland Security	EMCAPS software allows users to model disaster scenarios for planning and education.

Appendix 3: Inaccessible Intelligence Tools
Amelia Acker & Katie Shilton 12-14-09

Tool	URL	Description	Source	Accessible?
VACCINE	http://www.cerias.purdue.edu/site/projects/de	broad suites of tools to deal with heterogeneous information at scale with emphasis on visual analytics and presentation	CERIAS	No
Location Object Field Tracking (LOFT) technology		Designed to help organizations visualize people, vehicles and infrastructure anywhere on Earth to better protect and account for mobile assets.	Computer Science Corporation (CSC), a private company	No
The Risk Analysis Workbench (RAW)	http://create.usc.edu/research/risk_a	RAW is a software platform that enables policy/decision-makers and risk analysts to share computing tools, models, data, analysis and results, and supports each in their unique roles and needs in risk-sensitive assessments and planning.	CREATE @ USC	no
Analysis, Dissemination, Visualization, Insight and Semantic Enhancement (ADVISE)		Large DHS data mining program - may or may not still be in operation	Department of Homeland Security	No
Intelligence and Information Fusion (I2F) system		The system will use advanced computer processes for collecting, tagging, classifying and organizing data to gather and analyze information about potential terrorists.	Department of Homeland Security	No

Appendix 3: Inaccessible Intelligence Tools
Amelia Acker & Katie Shilton 12-14-09

Tool	URL	Description	Source	Accessible?
VICTIMS database	http://www.victimsidproject.org/login.aspx	VICTIMS is national database of unidentified human remains cases created as part of a research project within the FBI Laboratory. The goal of VICTIMS is to identify these individuals by providing a federally-sponsored, user-friendly website that anyone will be able to search.	FBI	No
IARPA	http://www.iarpa.gov/index.html	Many calls for tool development (IARPA Knowledge Discovery Dissemination Program, IARPA Socio-cultural Content in Language Program) but no links to tools	IARPA	No
ICE Pattern Analysis and Information Collection (ICEPIC)	http://www.ice.gov/pi/news/factshe	ICEPIC looks for unexpected new links between individuals and groups among databases collected by law enforcement agencies and supplemented by information from commercial databases. The system looks for patterns and links between databases, similar to what data mining programs do, but it is not a data mining program in the privacy impact assessment.	Immigration and Customs Enforcement	No
Intellipedia		A classified Wikipedia for the intelligence community	Intellipedia is a private company	No

Appendix 3: Inaccessible Intelligence Tools
Amelia Acker & Katie Shilton 12-14-09

Tool	URL	Description	Source	Accessible?
Counterproliferation Analysis and Planning System (CAPS)	https://www-gs.llnl.gov/operationalplan	A planning tool to conduct missions against facilities that support the production of WMDs. CAPS may be accessed through any of the major classified military networks.	Lawrence Livermore National Laboratory	No
End-to-End Exploitation (E3)	https://www-gs.llnl.gov/e3.html	The program allows data from a range of different sensor types, such as imagers and point detectors, to be ingested into an Oracle database. Once stored, the data can be searched by time, location, and features of the data such as confidence metrics or chemical name. Includes visualization tool.	Lawrence Livermore National Laboratory	No
Homeland Defense Operational Planning System (HOPS)	https://www-gs.llnl.gov/operationalplan	An operational planning tool accessible to authorized users who plan homeland defense exercises.	Lawrence Livermore National Laboratory	No
Story Tracking			Lead investigator is Stuart Rose, and he presented on the technology at the IEEE VAST in 2008	
Lingotek	http://www.lingotek.com	Translation software - privately developed, funded by In-Q-Tel	Lingotek (funded by In-Q-Tel, intelligence communtiy are clients)	Demo
California Herberia -- Spatio-temporal and georeferenced objects			North-East Visualization and Analytics Center	
GeoDeliberator			North-East Visualization and Analytics Center	

Appendix 3: Inaccessible Intelligence Tools

Amelia Acker & Katie Shilton 12-14-09

Tool	URL	Description	Source	Accessible?
ReliefWeb		mapping program that also integrates text sources, in this case, to do a project for humanitarian aide agencies	North-East Visualization and Analytics Center	
A-Space		Facebook for spies - an online collaboration space for Analysts	Office of the Director of National Intelligence	No
In-Spire	http://in-spire.pnl.gov/	N-SPIRE™, powerful information visualization software developed by Pacific Northwest National Laboratory, can give people the ability to see something different in the data they already have.	Pacific Northwest National Lab	Licenseable
Starlight Visualization Technologies	http://starlight.pnl.gov/	The Starlight Information Visualization System graphically depicts information, dramatically accelerating and improving human ability to derive meaningful knowledge from increasingly large and complex information resources. Visit the Overview section to learn more.	Pacific Northwest National Lab	For purchase

Appendix 3: Inaccessible Intelligence Tools
Amelia Acker & Katie Shilton 12-14-09

Tool	URL	Description	Source	Accessible?
Fused Analytic Desktop Environment (FADE)		FADE is an integration of five Lab-developed software technologies suited for exploration of unstructured and semi-structured textual data. The five software tools include Analyst Driven Knowledge Enhancement and Analysis (AKEA), Concept Based Clustering/Frame of reference Visualization (CBC/ForVIZ), Collaborative Analytical Toolbox (CAT), IN-SPIRE, and Universal Parsing Agent (UPA).	Pacific Northwest National Laboratories (Sold to Metatomix)	Not yet publicly available
Just-In-Time Command and Control Center (JITC3)	http://parvac.washington.edu/projects/view.php?shortname=jitc3	The JITC3 visualization system uses GIS data to generate a cohesive geo-spatial representation of an urban environment that can be navigated and manipulated using Augmented Reality.	PARVAC	no
First Look			PNNL, National Visualization and Analytics Center (NVAC)	No
Linked Animal-Human Health Visual Analytics (LAHVA)	http://pixel.ecn.purdue.edu:8080/~rmacieje/flier	LAHVA application allows users to visually search the data for clusters in statistical model view and spatio-temporal views.	Purdue University Regional Visualization and Analtics Center.	No
CCICADA			Rutgers?	No

Appendix 3: Inaccessible Intelligence Tools
Amelia Acker & Katie Shilton 12-14-09

Tool	URL	Description	Source	Accessible?
Dynamic Preparedness System	http://fazd.tamu.edu/products/inform	The DPS provides emergency response managers access to FAZD-based modeling, decision-support, and situational awareness tools during training or during an actual FAZD event. It will also be used to analyze information in the aftermath of an outbreak to develop lessons learned and best practices.	The National Center for Foreign Animal and Zoonotic Disease Defense (FAZD)	No
Food and Agriculture System Transportation (FASTRANS) Model of Interstate Movement of Livestock (with the National Center for Food Protection and Defense)	http://fazd.tamu.edu/products/inform	The DHS has provided special funding to the National Center for Food Protection and Defense (NCFPD) and the FAZD Center to acquire the data and to build a national transportation model that will be input to multiple epidemiologic modeling efforts. The initial effort will focus on beef, dairy, and swine, but we are planning follow on efforts for other commodities. This will provide for the first time a quantitative estimate of what is probably one of the most important factors in the spread of foreign animal or zoonotic disease through the interstate movement of large numbers of animals over long distances.	The National Center for Foreign Animal and Zoonotic Disease Defense (FAZD)	No (under development)

Appendix 3: Inaccessible Intelligence Tools
Amelia Acker & Katie Shilton 12-14-09

Tool	URL	Description	Source	Accessible?
Models and Databases to Assess Consequences of Prevention and Intervention	http://fazd.tamu.edu/products/inform	The FAZD Center has worked with multiple partners including several national laboratories to develop a suite of molecular analytical tools that has provided valuable and often unanticipated insight into select agent disease pathways, and is now being employed for the study of other important agents including Avian Influenza.	The National Center for Foreign Animal and Zoonotic Disease Defense (FAZD)	No