[7] Nielsen, J. & Landauer, T. A Mathematical Model of the Finding of Usability Problems. *Proceedings of INTERCHI '93*, ACM, 206-13, 1993.
[8] Pierotti, Deniese. Heuristic Evaluation: A System Checklist. In *Usability Analysis and Design*. Xerox Corporation, 1995.
[9] Department of Defense. *Defense Information Infrastructure Common Operating Environment (DIICOE) User Interface System Checklist.*

5 RECOMMENDATIONS

5.1 Text Mining Tools

As mentioned previously, we are considering three categories of tools – Text Mining, Investigative Analysis, and Bayesian tools that would help an analyst with his/her task. However, as the emphasis of this study was on Bayesian modeling tools, we have not had a chance to do an in-depth evaluation of text mining and investigative analysis tools. SRA, a regular user of vendor products, has gathered information from sources that do in-depth study of these tools.

For text mining tools, we relied on market research done by SRA. Table 28 lists well known tools and vendors in different categories of text mining.

Table 28. Text Mining Tools

Technique	Tool Name	Website
Categorization of documents	Autonomy	http://www.autonomy.com/Content/Technology/
	Convera	http://www.convera.com/Products/index.asp
	Entrieva	http://www.entrieva.com/entrieva/index.htm
	InXight	http://www.inxight.com/
	Mohomine	http://www.mohomine.com/
	Stratify	http://www.stratify.com/
	Verity	http://www.verity.com/products/index.html
Summarization	Copernic	http://www.copernic.com/en/products/summarizer/index.html
	InXight	http://www.inxight.com/

5.2 Social Network Analysis Tools

For social network analysis tools, the recommendation is based on the results from a detailed survey done for MITRE Corporation. In this report, i2's *Analyst Notebook Version 6* was the first recommendation and the next recommendation was Visual Analytics *VisuaLinks*. For further details regarding the criteria used, other tools evaluated and generation of results, refer to [3] (end of Section 4), included as Attachment 8 to this report.

5.3 Modeling Tools

5.3.1 Performance

From Section 4.5.6, it can be seen that *Hugin* ranks as the optimal Bayesian tool among all the ranked alternatives. The next best alternatives are *BayesiaLab* and *Netica*.

Most of *Hugin*'s high decision score can be attributed to the number of features it supported in the category of Input Manipulation and the fact that criteria Input Manipulation was given the maximum weight towards the final goal of choosing an optimal tool.

BayesiaLab has excellent design sense from a usability standpoint. Though there are some "ease of use" problems, they should be relatively easy to fix. Both *Hugin* and *BayesiaLab* personnel were very helpful and receptive to making changes and adding features in response to specific queries and customer requests (in fact the *BayesiaLab* folks have already made some changes/updates in response to queries from this study).

Netica and *BayesiaLab* representatives are already addressing some of the performance issues that led to their ranking below *Hugin* (see Table 12). With the addition of these planned extensions, it is likely that the performance rankings for *Netica* and *BayesiaLab* will improve.

SIAM is the only GOTS tool that we analyzed in detail, for reasons described in Section 4. One of the major drawbacks of *SIAM* is that it does not support Bayesian Inferencing. Other features that *SIAM* does not address: the learning of probabilities from a sample set and learning of the structure of the net from a database. These features are useful if the end user wants the tool to work with existing data. In Performance & Extensibility, *SIAM* ranked low because of the emphasis placed on the language used to develop the computational engine and the Graphical User Interface (GUI). The preferred language was C/C++ and *SIAM* was developed in Java. The lack of an Application Program Interface (API) for *SIAM* also contributed to its low score. All that said, *SIAM* does have performance features that will be useful for NASIC in the near term.

Based on the information gathered for *OCCAM*, it can be said that it seems to be addressing the domain specific needs of the customer. It also employs all the techniques that the customer is planning to apply to build a decision support system that helps model the influences impacting human behavior.

Sample might be useful to NASIC/BPB if they plan to employ multiple artificial intelligence techniques (in addition to Bayesian Nets).

Although Metrica's PSYOP DSS tool is not applicable as one of the ranked alternatives (does not use Bayesian methods), the feature of building a database

that contains data supporting PSYOP objectives (and helping the user based on this information) *is* a very useful technique. This technique can be applied to Bayesian tools to help the user determine the a-priori probabilities.

5.3.2 Ease of Use

Where *Hugin* and *BayesiaLab* were the top modeling tools from a performance perspective, *Netica* and *Analytica* stood out from those two in usability. *Netica* rated most "user friendly" in terms of user support, efficiency of entry, error management, and supported visualizations. *Analytica* was a close second in all these categories, with the *Hugin* and *BayesiaLab* ranking significantly lower in ease of use.

SRA continues to receive information from all these companies. As we get more information, we will update our AHP analysis to ensure we have accurately modeled them and get the correct relative ranking. As we get results, we will provide them to the customer. However, from both a performance and a usability perspective, it would be far preferable to validate results from this study against a representative sample problem or scenario.

5.3.3 Cost

Table 29 summarizes the cost per single license, quantity licenses, and site license of the top five modeling tools. As some of the tools prices are listed in Euros, refer to http://www.x-rates.com/calculator.html for an equivalent price in Dollars.

Table 29. Cost Summary.

Tool	Pricing Structure
Hugin	(GUI and API Priced) 1 License – 6300 Euros 2 Licenses – 8390 Euros 5 Licenses – 10,465 Euros 10 Licenses – 12,550 Euros Site License – 16,785 Euros
BayesiaLab	(GUI and API Priced) 1 License – 3450 Euros 5 Licenses – 10,350 Euros 10 Licenses – 17,250 Euros 50 Licenses – 43,125 Euros
Netica	GUI - $585 per License, API - $685 Note 1: Site licenses available for 5 times as much. Note 2: API embedded is $20 to $175 depending on the volume.

Tool	Pricing Structure
Analytica	Professional - $1295 Enterprise - $2495 Analytical Decision Engine - $6000 Note 1: Website (Attachment 3) gives comparison of features of these different versions) Note 2: 15% discount for 10+ Licenses Note 3: 25% discount for 50+ Licenses
SIAM	All US gov't agencies hold a license to use *SIAM* for gov't purposes (just need federal ID)

Complete cost information for the rest of the evaluated tools can be found in the *"General Features"* worksheet in Attachment 3.

6 CONCLUSION

6.1 NASIC's Investment in Analytical Capabilities

Many viable text/data mining options are available to NASIC/BPB. SRA is very familiar with NetOwl, and has used it for a variety of applications. However, other options were presented in Section 5.

From discussions with experts in the field of Social Network Analysis, we were referred to a Mitre Corporation report (Attachment 9) that ranked i2's Analyst Notebook Version 6 at the top of the list of these capabilities. The next best capability recommendation was Visual Analytics VisuaLinks.

Based on the results of this study, NASIC/BPB has at least five viable options to meet mission modeling requirements: *Hugin*, *BayesiaLab*, *Netica*, *Analytica*, and *SIAM*. Unfortunately, the tool that ranked best in terms of performance (*Hugin*) is also rated, relative to the other tools, more difficult to use. *BayesiaLab*, another highly ranked tool in terms of performance, is also ranked lower in "ease of use." *Netica* and *Analytica*, while easier to use, did rank lower in performance. The lack of a clear "winner" is further complicated by the fact that some extensions to the performance capabilities of both *Netica* and *BayesiaLab* are planned in the next release of these tools. None of these options is out of the running in terms of cost/price, and (of the five top-rated performance options) *SIAM* has the advantage of being available for licensing free of charge to government customers.

On balance, especially if ease of use is a critical factor for NASIC and the customer is willing to wait for the performance improvements (Table 12) planned