Big Data and Open Source Information at EUCOM

Dr. Regan Damron
EUCOM J2-IE, Strategic Foresight
damron_regan@bah.com
regan.w.damron.ctr@mail.mil

Prepared for presentation at the 8th Annual Strategic Multi-Layer Assessment Conference, 28 October 2014



Big Data: Premises

- When is data "big"?
 - When its structural characteristics (e.g., Laney's "volume, velocity, & variety") become issues themselves (storage, algorithmic efficiency, unstructured data management, etc.)
- Such data requires a paradigm shift
 - From data sharing to access sharing
 - Private sector leads this with APIs and SDKs
 - Intel is not very good at this; we must improve
 - We can learn using open-source information



Big Data: Premises (Cont.)

- Big data (and associated TTPs) is generally best for:
 - Automation of simple and/or formulaic tasks
 - Augmentation of analyst capabilities
- Big data is not (currently) good for:
 - Automation of complex, adaptive, and/or high-stakes tasks
 - Replacement of human analysts/human reasoning
 - ...but it's getting better.



Big Data: Dangers

- Big data can be dangerous because:
 - Use of "normal" statistical techniques without compensating for data "bigness" can yield erroneous results
 - With a large enough dataset, anything can be shown to be related to anything (statistical significance)
 - Results can be highly sensitive to variable definition (operationalization)
 - Results must be interpreted (substantive significance);
 this requires:
 - Knowledge of context and contextually-appropriate theory
 - Training to recognize and avoid fallacious reasoning
 - There is no substitute for clear thinking.



EUCOM Strategic Foresight



Knowledge Commons

- Big data, open source repository
 - Aggregator and reader for content
 - Provides situational awareness on topics without forcing user to read all content
 - Object Based Production semantically categorizes content against topics being tracked
- Environment sensor/scanner
 - Discerns trends and events from noise
 - Alerts users to surprises and developments
 - Open source indicators and warnings
- Analysis to rapidly understand topics
 - Extracted entities used to reduce user need to read source content
 - Temporal and faceted filtering
 - Highly visual aggregate analytics



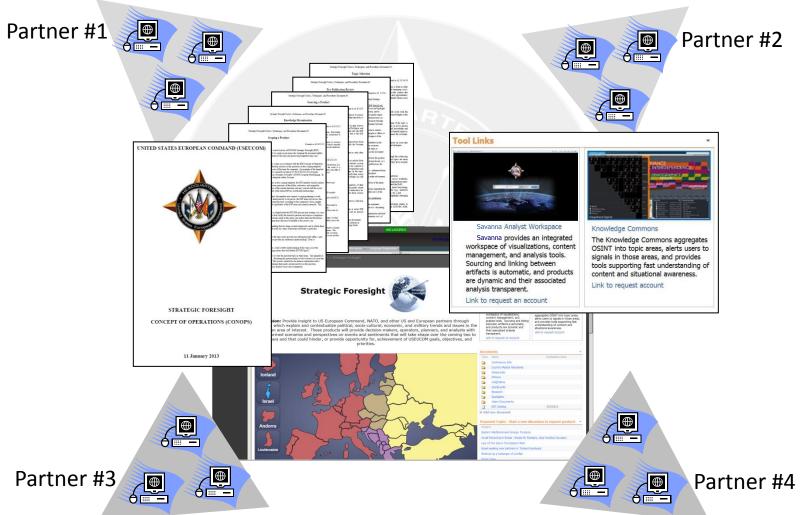
Savanna

- Analyst workspace focusing on information synthesis and production
- Focuses on the relationships between items
- Explicitly represents analysts' thoughts in mind maps
 - Enables review, collaboration, and oversight
 - Links conclusions to contributing factors
 if factors change affected conclusions are obvious
- All data is "live" changes propagate to all products
- > Highly visual and interactive





The ESF Vision: A Community of Practice





A EUCOM-Wide Effort

- In the last month, ECJ2 has been leading an effort to systematically:
 - Understand, codify, and verify process/doctrine regarding the use of open-source information
 - Requires getting legal findings on differences between OSINT and "open source research"
 - Assess the potential for collaboration/coproduction with NATO allies and partner nations in an open-source environment
 - Engage in a dialog to convince others in the Intel establishment that open-source info has intrinsic value—it's more than just another input to "all-source analysis"



Questions?

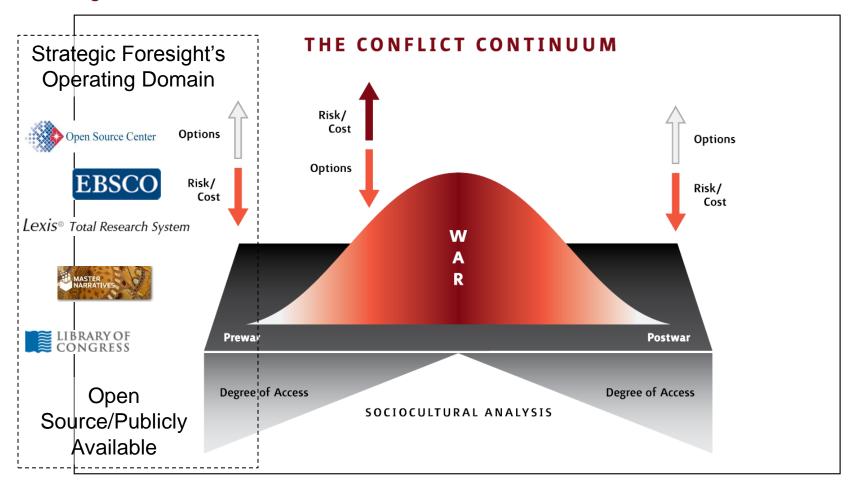


Backup Slides



The Need: Why Strategic Foresight?

Figure. The Conflict Continuum





From "Left of Bang: The value of socio-cultural analysis in today's environment," <u>National Security Challenges: Insights from Social, Neurobiological, and Complexity Sciences</u>, White Volume, Jul 2012.

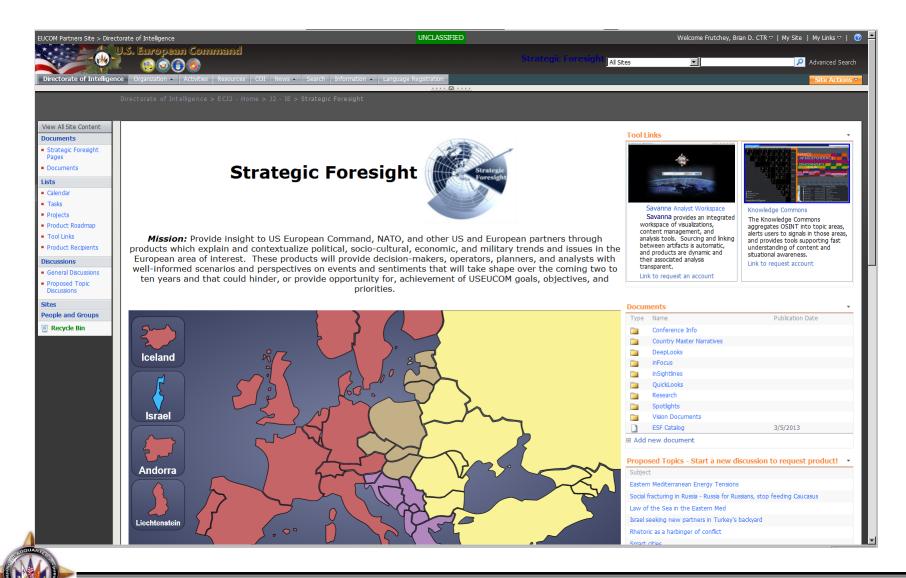
Strategic Foresight Publications



... at U.S. European Command

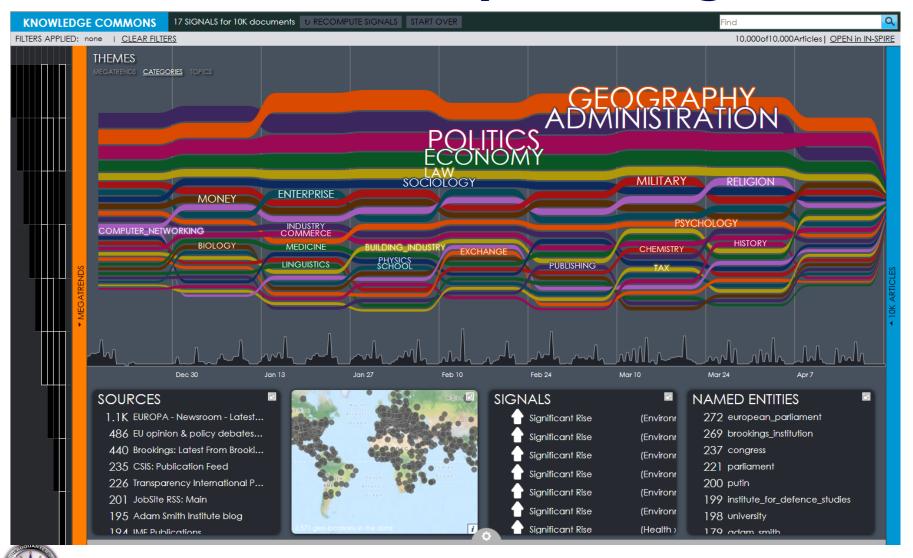


SharePoint Portal

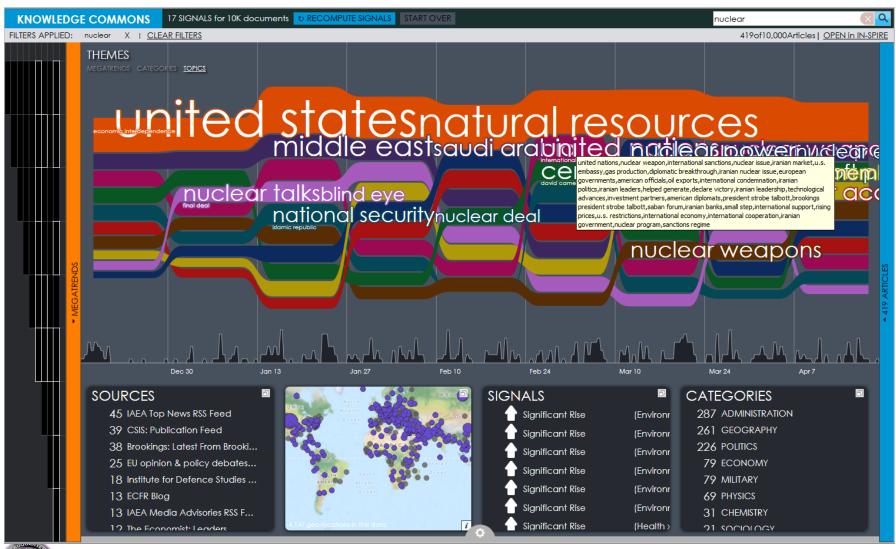


United States **European Command**

Theme River Analysis - Categories



Theme River Analysis - Topics





Analytic Process Management

