



Topics in Sandia Cyber Research

David H. Rogers

**Manager, Dept 1424
Data Analysis and Visualization Department
Sandia National Labs**

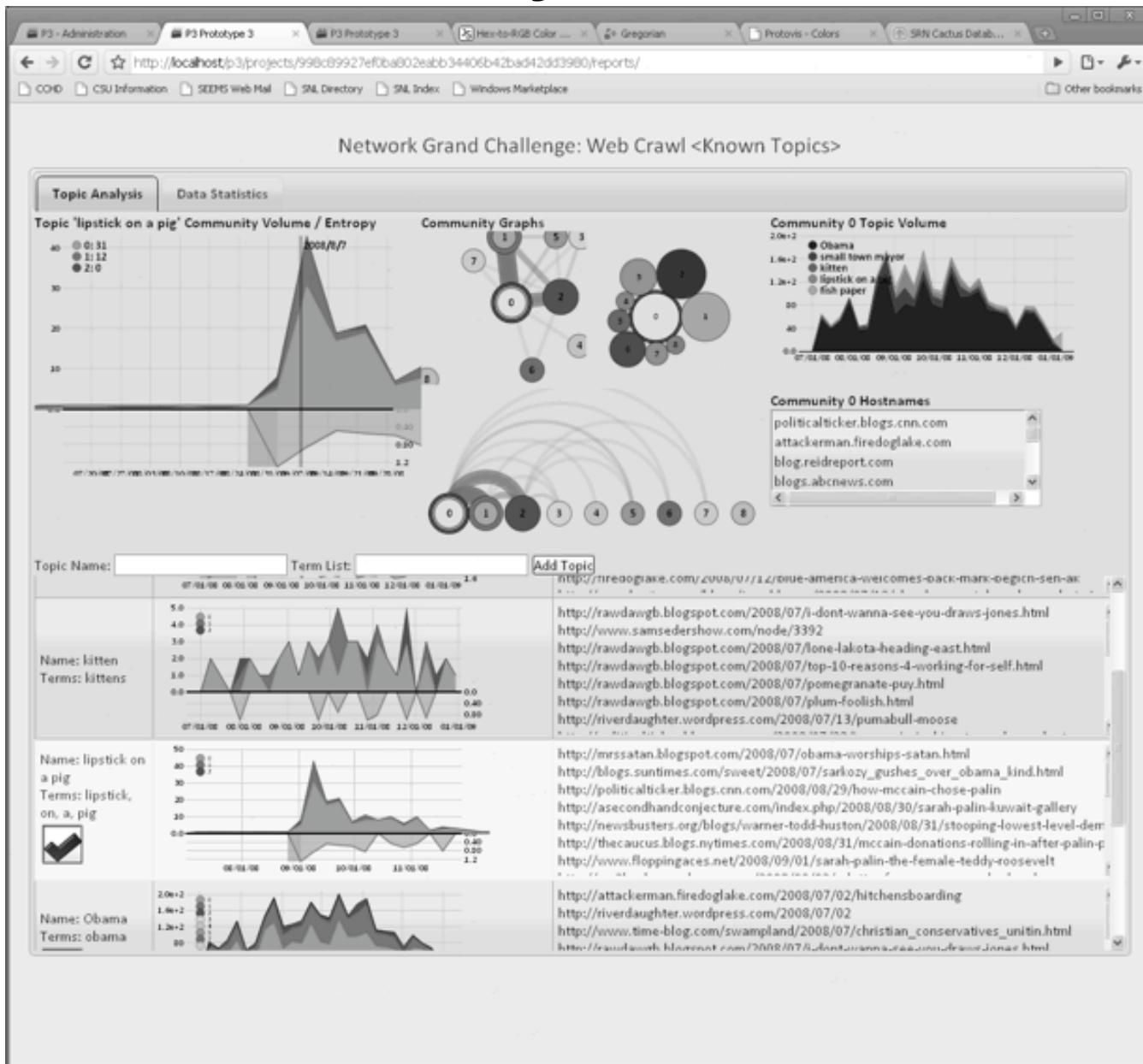


Sandia is Applying Historic Strengths to Cyber Problems

- **Strengths**
 - Parallel Algorithms and High Performance Computing
 - Scalable Visualization and Data Analysis
 - Range of Deep Subject Matter Expertise
- **Examples**
 - Sentiment Analysis of Blogs
 - Multi-lingual Doc Analysis on Red Storm
 - Integrated Packet, Topic and Traffic Analysis of Cyber Data
- **Strategic vision for investment**



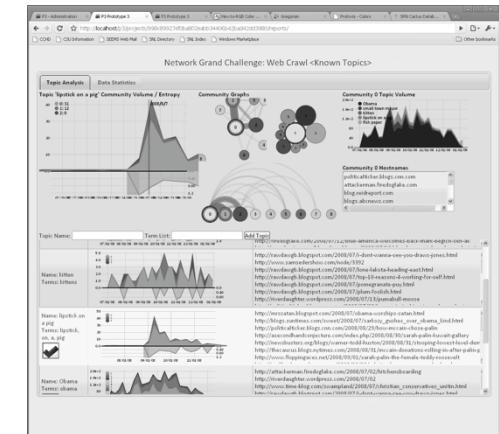
Sentiment Analysis of Web Traffic





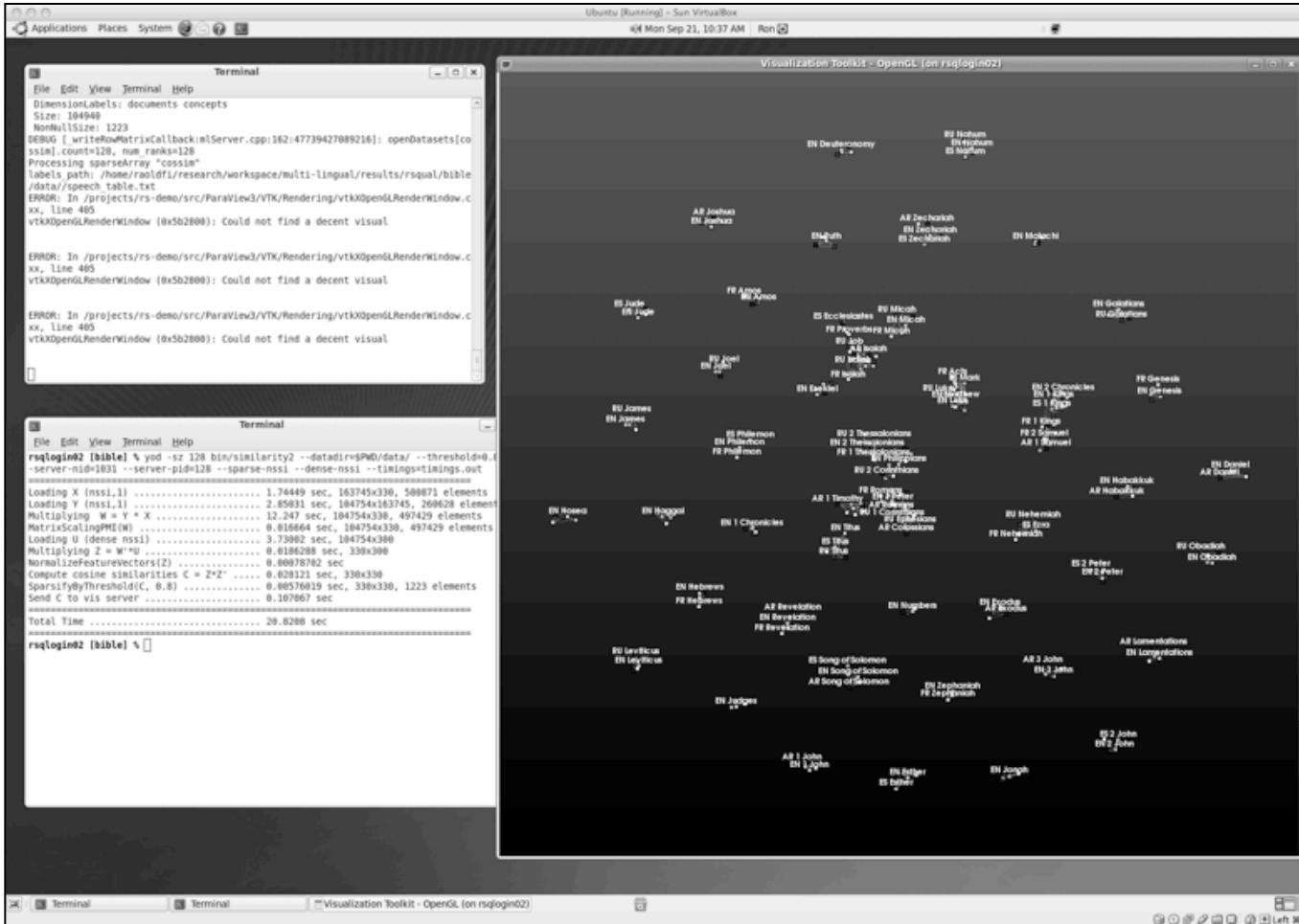
Sentiment Analysis of Blogs

- **Investment: Networks Grand Challenge, a three year, LDRD-funded project**
 - **Problem: can we characterize and predict the evolution of web-based memes (ideas), and understand which ones will ‘take off’**
 - **Result: Yes. Prototyped integrated browser-based analysis UI with a query-driven server analyzing large data from a web crawl**





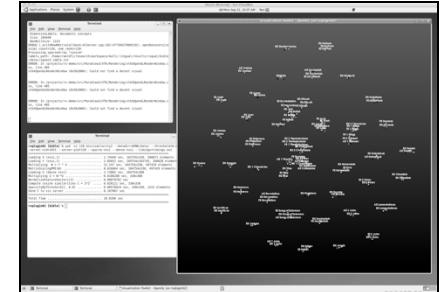
Multi-lingual Text Analysis on Red Storm





Multilingual Text Analysis on Red Storm

- **Question: can we do document triage – quick assessment of relative importance of massive data – on large collections of multi-lingual documents.**
 - Can we use the power of computation to cluster important documents by topic, without translating them?
- **Example: Using scalable tensor methods on distributed memory computers,**
- **Answer: Yes, given an appropriate training set, millions of documents can be analyzed with scalable tensor methods, and clustered for further investigation.**





Integrated Packet, Topic, and Traffic Analysis of Cyber Data





Integrated Packet, Topic and Traffic Analysis of Cyber Data

- **Question: can we use High Performance Computing to help analyze cyber packet information**
- **Example: integrates packet analysis, topic analysis and traffic analysis into a single application.**
 - Promotes integrated analysis of all levels of cyber data, using appropriate algorithms at each level.
- **Answer: Yes, but this application of technology requires a large investment in UI and interaction design – algorithms aren't enough**





Current Open Research Summary

- **Sandia's existing expertise in scalable algorithms and high performance computing is being applied to analysis of web, digital document collections, and cyber data**
 - An array of methods – from distributed memory algorithms to massively multithreaded algorithms – can serve as a set of tools to analyze extremely large, complex data
- **Future investments will be aligned with Sandia's strategic vision for computing, analysis, and customer-focused projects.**



Strategic Vision for Investment

- Leverage existing expertise and technology across a range of R&D at the Lab
 - Make strategic investments in core technology areas, and partner with external institutions and industries as appropriate
- Lab Directed Research and Development (LDRD) in Cyber
 - Focusing investment in strategic, distinguishing areas of technology for Sandia
 - First call, FY10, second call coming out Q1 FY11
- Cyber Engineering Research Institute (CERI)
 - Vision to bring together internal subject matter experts, researchers, and external partners to address problems of national importance