Brief Bio and (PR)^2: Problems & Pitches – Rants & Raves by *Luis M. A. Bettencourt*

Dear invited participant of the NSF Workshops on Knowledge Management and Visualization Tools in Support of Discovery at the National Science Foundation, Arlington, VA on March 10 & 11, 2008, see http://vw.slis.indiana.edu/cdi2008/workshop1.html.

In preparation for the workshop we ask you to provide a brief bio and your input to the questions below. Your input will be available to all participants and people which are "interested (but cannot attend)" before the meeting to complement the introduction of participants and to structure the workshop more effectively.

Thank you for your time.

Self Introduction

Please introduce yourself by providing a
* photo of yourself
* brief biography of about 250 words

Luís M. A Bettencourt [Technical Staff Member, T-7, LANL] obtained his PhD from Imperial College, University of London, in 1996 for work on critical phenomena in the early Universe, and associated mathematical techniques of Statistical Physics, Field Theory and Non-linear Dynamics. He held postdoctoral positions at the University of Heidelberg, Germany, as a Director's Fellow in the Theoretical Division at LANL, and at the Center for Theoretical Physics at MIT. In 2000 he was awarded the distinguished Slansky Fellowship at Los Alamos National Laboratory for excellence in interdisciplinary research. He has been a Staff Member at LANL since the spring of 2003, first at the Computer and Computational Sciences Division (CCS), and since September 2005 in the Theoretical Division (T-7: Mathematical Modeling and Analysis). He is also External Research Faculty at the Santa Fe Institute [effective July 2007], and a Research Professor at the Department of Mathematics and Statistics and the School of Human Evolution and Social Change at Arizona State University.

Dr. Bettencourt carries research in the structure and dynamics of several complex systems, with an emphasis on dynamical problems in biology and society. Currently he works on real time epidemiological estimation, social networks of human communication, distributed sensor networks, information processing in neural systems and urban organization and dynamics. He maintains many diverse multidisciplinary collaborations at LANL, and beyond at the Santa Fe Institute, ASU, Princeton University, Harvard, MIT, SAMSI, and UC Davis. He is a member of advisory committees for international conferences and referees for journals in physics, mathematics and computer science, and for international fellowship programs. In the last three years he has co-advised four PhD thesis, in epidemiology, complex networks, computational neuroscience and statistical mechanics. He is also a consultant for the Office Science and Technology Information of the US Department of Energy on the subject of scientific innovation and discovery

* up to five major publications (they are linked)

Growth, innovation, scaling, and the pace of life in cities

Identification of functional information subgraphs in complex networks

Towards Real Time Epidemiology: Data Assimilation, Modeling and Anomaly Detection of Health Surveillance Data Streams

Separating the Wheat from the Chaff: Practical Anomaly Detection Schemes in Ecological Applications of Distributed Sensor Networks

The power of a good idea: Quantitative Modeling of the spread of ideas from Epidemiological models
- link to your home page and
http://math.lanl.gov/~lmbett/

- links to data or software you serve (if applicable).

N/A I use web of science data, data from delphion on patents, and some exaples of streaming ecological data, such as:
http://sevilleta.lanl.gov/warming/graph/day/airtmp/

General Questions

1) What is (are) your main interest(s) in attending the workshop?

Learn from other participants about state of the art electronic (streaming) data, its management and visualization. Come of with new concepts where these opportunities can be harnessed to do fundamental new science. I.e. I am not after just tool building, it needs to generate new scientific opportunities that are transformational in the NSF's lingo.

2) What information/knowledge management needs do you have?
Explain your 'dream tool' for scientific discovery and innovation.

I'd like to be able to integrate streaming data with statistical and dynamical models that can : 1) characterize statistical regularities in data streams, 2) characterize statistical exceptions (anomalies) in such data, 3) make predictions with quantified uncertainty about future data, and 4) archive and visualize past data seamlessly. I imagine applying this to scientific and technological data, e.g. to records of publications, patents, and possibly usage (downloads).

Building such environment would be tantamount to being able to predict (at least in the short term) scientific and technological trends.

Something like this
http://emm.jrc.it/NewsBrief/clusteredition/en/latest.html
but for scientific and technological data (say new papers and downloads/citations), then coupled with predictive dynamical and statistical methods that automatically analyze the data for analomalies and make predictions of future events (we'd have to discuss wuat those would be) ….

3) What is the most insightful visualization of static or dynamic phenomena you know?
 [Ideally this visualization led to a major discovery/innovation. Examples could come from science, art, or any other field of human endeavor. Note that we plan to use this visualization on your name card.]

I have to go with an oldie: the structure of DNA. Because it implies how biological (genetic) information is coded for.

I also like similarity trees but their meaning is ambiguous usually and they don't reveal functional information.

4) What would you like to learn / achieve at the workshop?

Build a group of researchers engaged in a collaboration to analyze and predict the course of science and technology, from online streaming data.

*** Please send the completed document by <u>February 20th, 2008</u>***
*** to Katy Borner <katy@indiana.edu> and Elisha Hardy <efhardy@indiana.edu>***