

Application of Scientometric Methods to Identify Emerging Technologies

Presented at:
Quantitative Methods in Defense and
National Security (QMDSS 2012)

George Mason University

Robert K. Abercrombie, Ph.D.

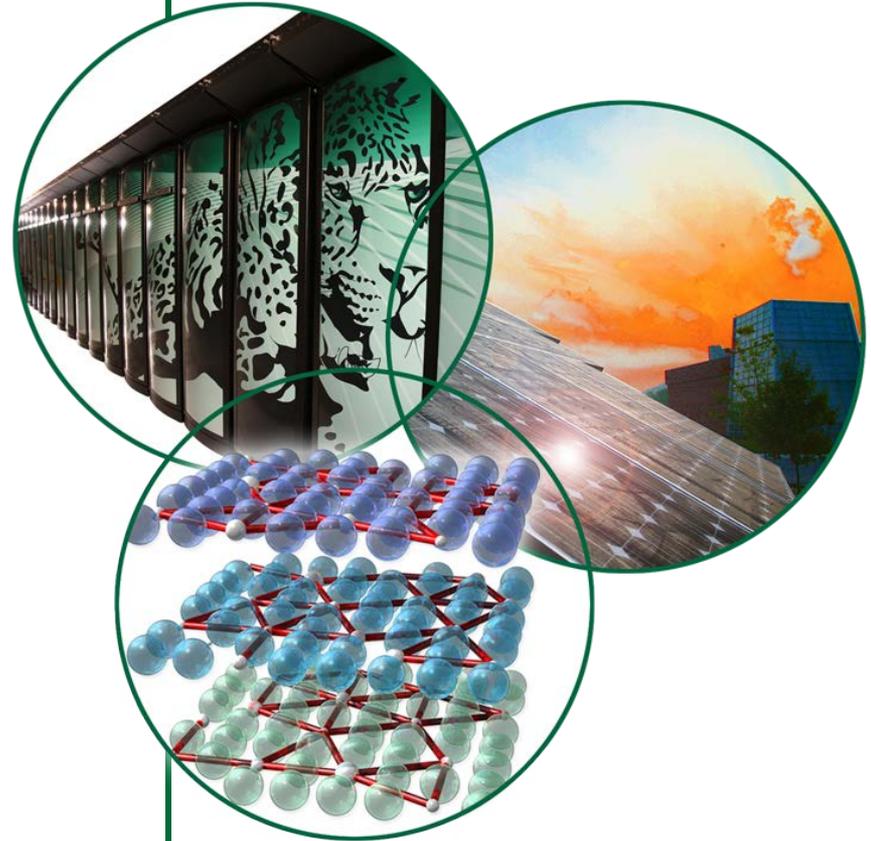
April 30, 2012

Authors:

Robert K. Abercrombie (ORNL) and
Bob G. Schlicher (ORNL)



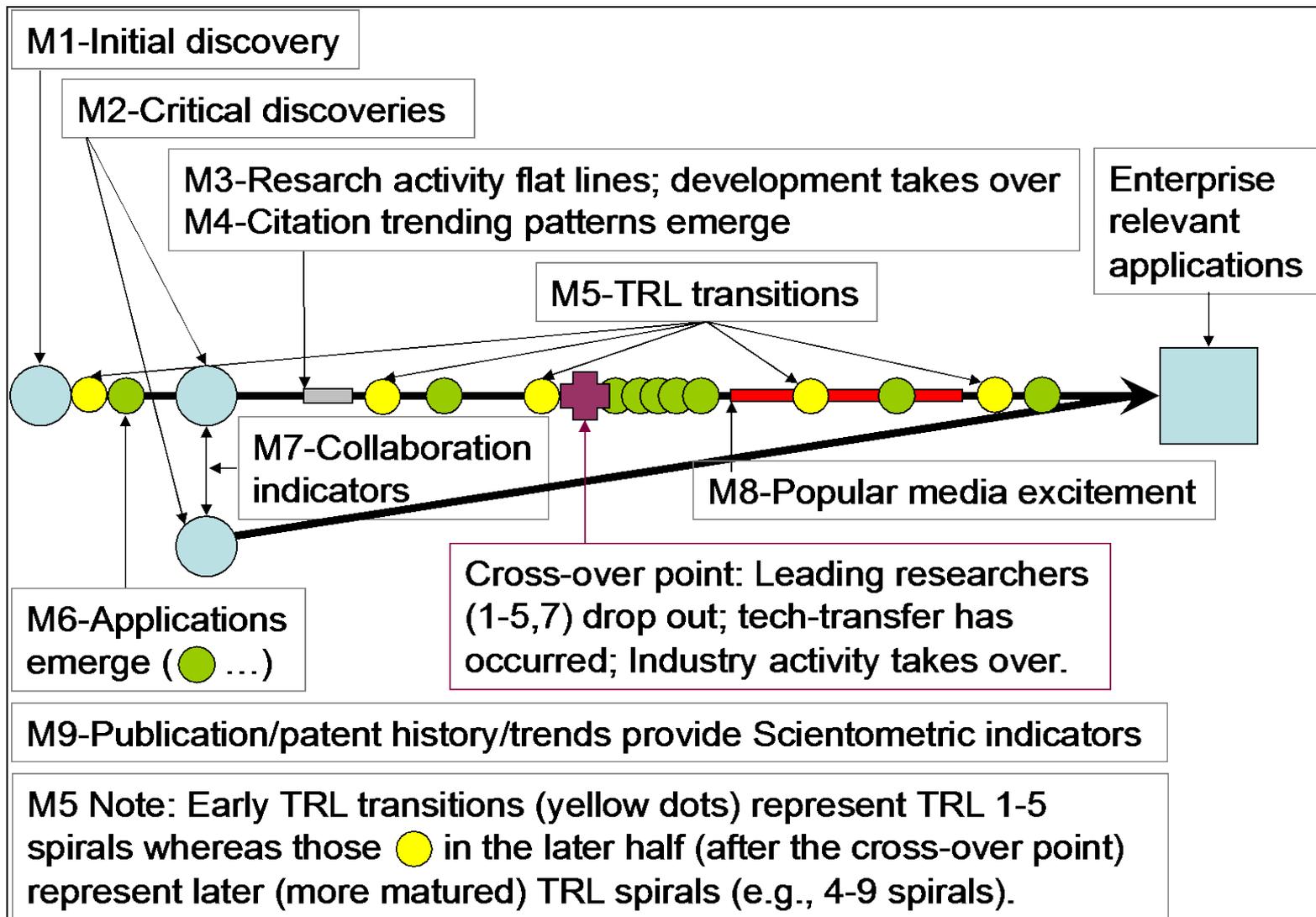
The submitted manuscript has been authored by a contractor of the U.S. Government under contract DE-AC05-00OR22725. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.



Agenda For Today's Presentation

- Purpose
 - Present findings of one subject domain, the Simple Network Management Protocol (SNMP), as applied to ORNL's Technology Evolution Model (TEM)
- Explanation of one possible systematic flow through TEM
- Stepwise Results of the TEM Milestones using SNMP as an example
 - Milestone 1: Initial Discovery
 - Milestone 2: Critical Discoveries
 - Milestone 3: R&D Activities Flat Lines
 - Milestone 4: Citations Activities Flat Lines - Corollary to Milestone 3
 - Milestone 5: Technology Readiness Level (TRLs)
 - Milestone 6: Commercial Applications emerge
 - Milestone 7: Collaborations indicators
 - Milestone 8: Sentiment and Excitement
 - Milestone 9: Publication/patent history/trends
- Conclusions and Future Directions

Hypothesis: General of technology evolution from initial discovery to enterprise relevant application



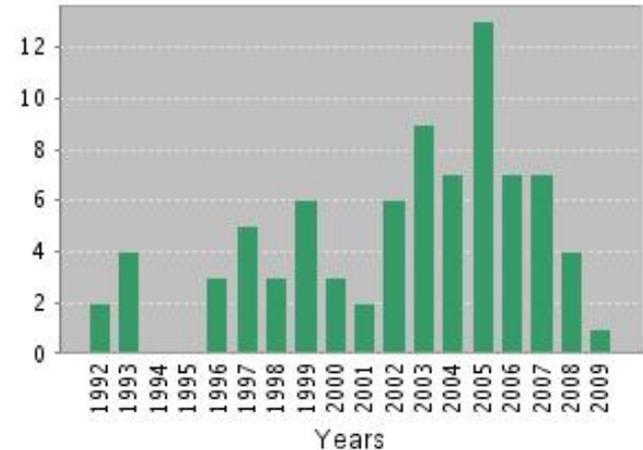
Testing of the Technology Evaluation Model (TEM)

- **Subject Domain of Study**
 - Simple Network Management Protocol – SNMP is a standard operations and maintenance Internet protocol
 - SNMP is a component of the Internet Protocol Suite defined by IETF
 - The history of SNMP illustrates the process of TEM and tests the hypothesis milestones identified in general terms
- **Collection of Dataset from Searchable On-line Data Networks**
 - Scholarly Data (Publications and Citations)
 - Patent Data
 - Web News Data Sources
 - Business/Product Starts

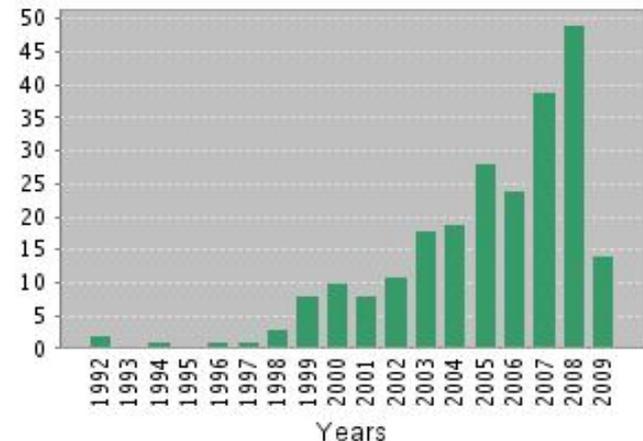
Peer-reviewed scholarly publications and citations per year

- Dataset - scholarly on-line searches identify the following trends from 1992 through the beginning of 2009.
 - Only complete yearly data were analyzed for trends (i.e., 2009 was excluded from trend analyses).
- Top Figure identifies scholarly publications per year.
- Bottom Figure identifies the numbers of citation per year.

Published Items in Each Year

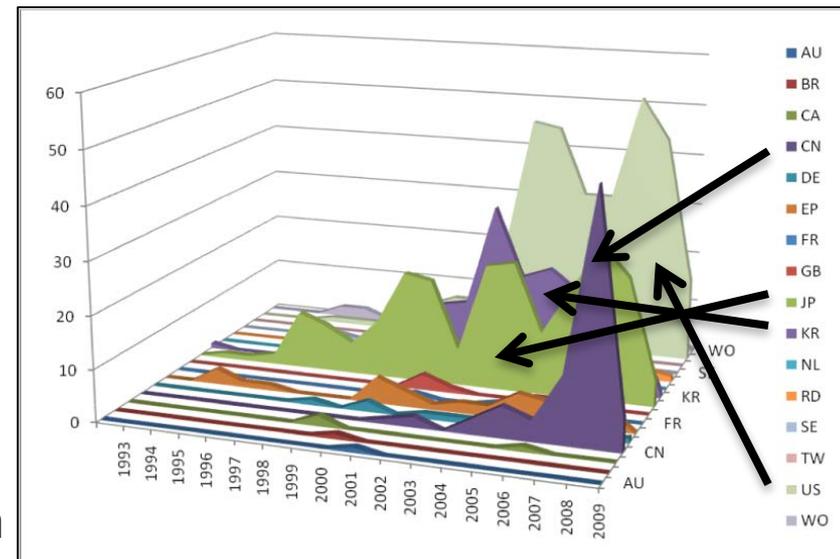
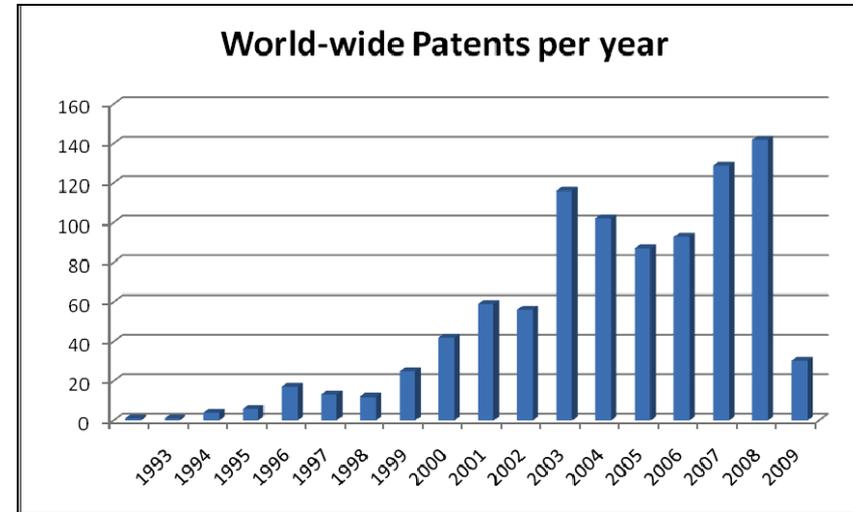


Citations in Each Year



Number of patents per year by country

- Dataset - 904 patents from 1992-2008.
- Preliminary charts were developed, which provided a foundation to quickly interpret the remaining data from other sources.
- Top Figure identifies worldwide patent data presented in its raw form as patents per year.
- Bottom Figure presents the numbers of patents per year and illustrates emerging trends with respect to different countries over time.
 - For example, the US, China, S. Korea and Japan exhibit patent spurts indicative of both +/- correlations



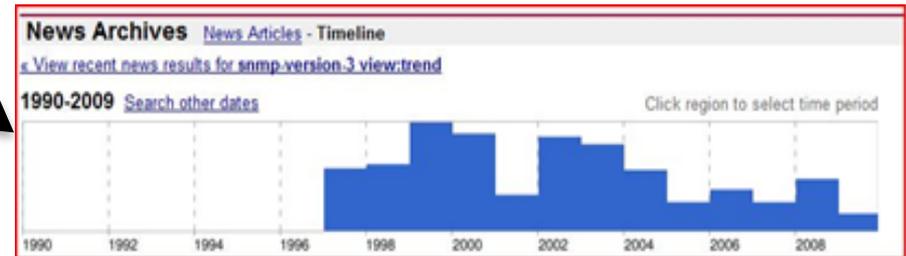
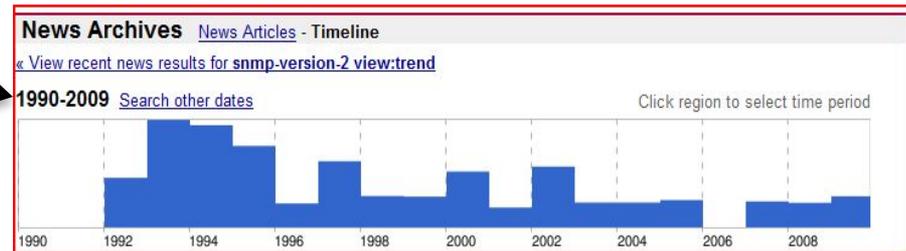
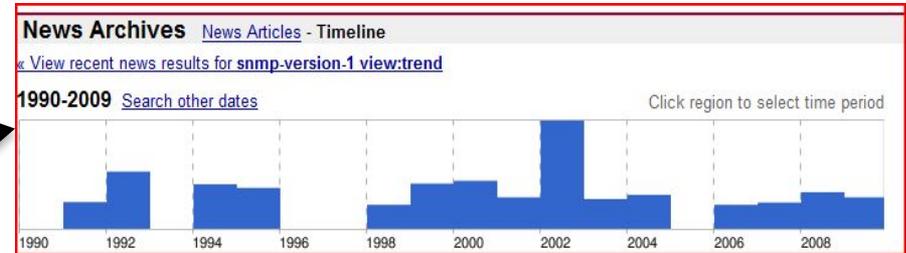
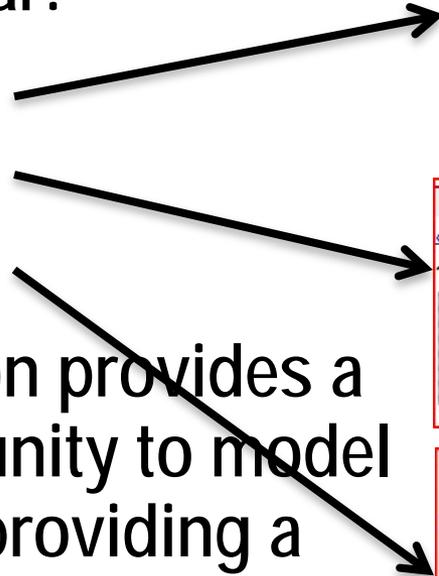
Web News Data Sources for each version of SNMP

- Histograms for each version of SNMP by year:

- Version 1
- Version 2
- Version 3

- This information provides a unique opportunity to model the networks, providing a perspective about the:

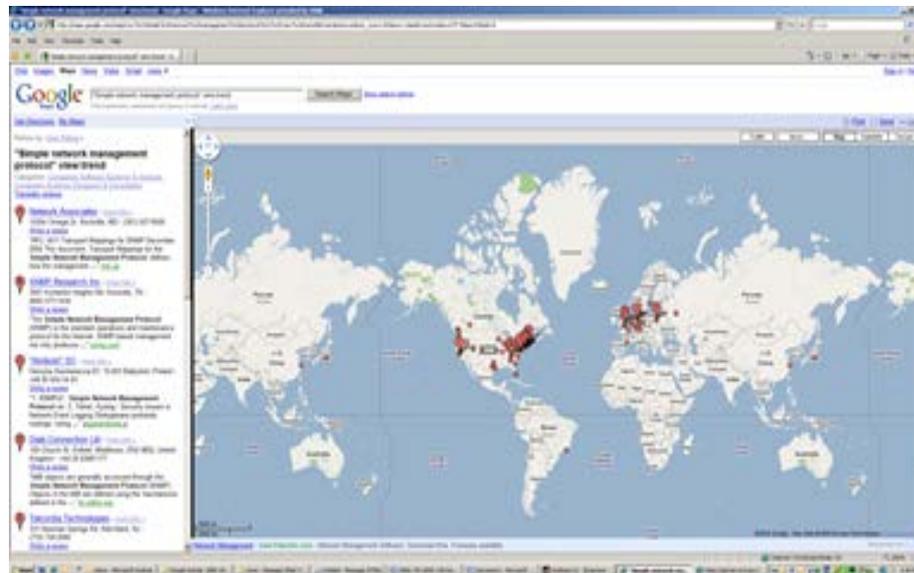
- persistence of sentiment and
- excitement (i.e., popularity) associated with a news worthy (possibly hyped) activity over a number of years.



Note: Comparison of Panels - knowing the maximum values [tallest bar] for Panels 1-3 [top to bottom] to be 16, 5 and 16 respectively)

Worldwide Distribution of Companies Associated with SNMP product applications offerings

- Search criteria from the Web News feed acquisition was expanded to “Simple Network Management Protocol” and resulted in a richer all-inclusive dataset of 333 entries.
- Figure identifies the initial search and the interactive display from which the refinement began.
- Further research (culling) of each reference site, narrowed the dataset to 56 distinct entities.
- These were identified as either:
 - new company starts, or
 - new products from established companies.



Analysis of the Combined Datasets

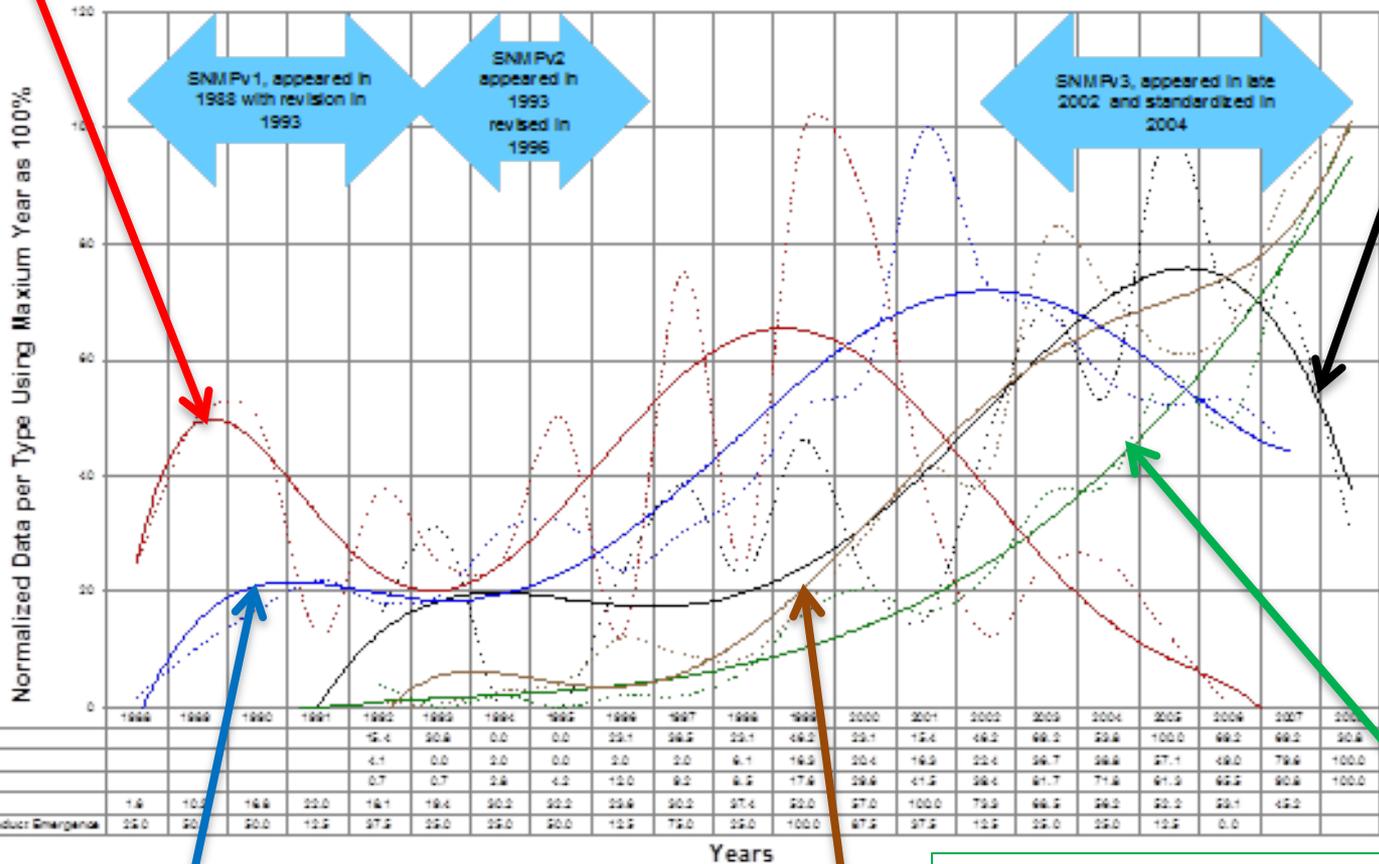
- Combination of the datasets from the previously defined four distinct and separate on-line sources become the collective dataset used to quickly (inexpensively) analyze the temporal flow of the SNMP subject domain
 - from initial discovery (via original scientific and conference literature),
 - through critical discoveries (via original scientific, conference literature and patents),
 - transitioning through some or all of the various TRLs, and
 - ultimately on to enterprise relevant application, while tracking news media.

Normalized data trend "fitting" analysis

Application Specific Product Emergence – red line

Academic Articles – black line

Normalized Data (dashed lines) with Polynomial Trend "Fitting" Analysis (bolder solid lines)



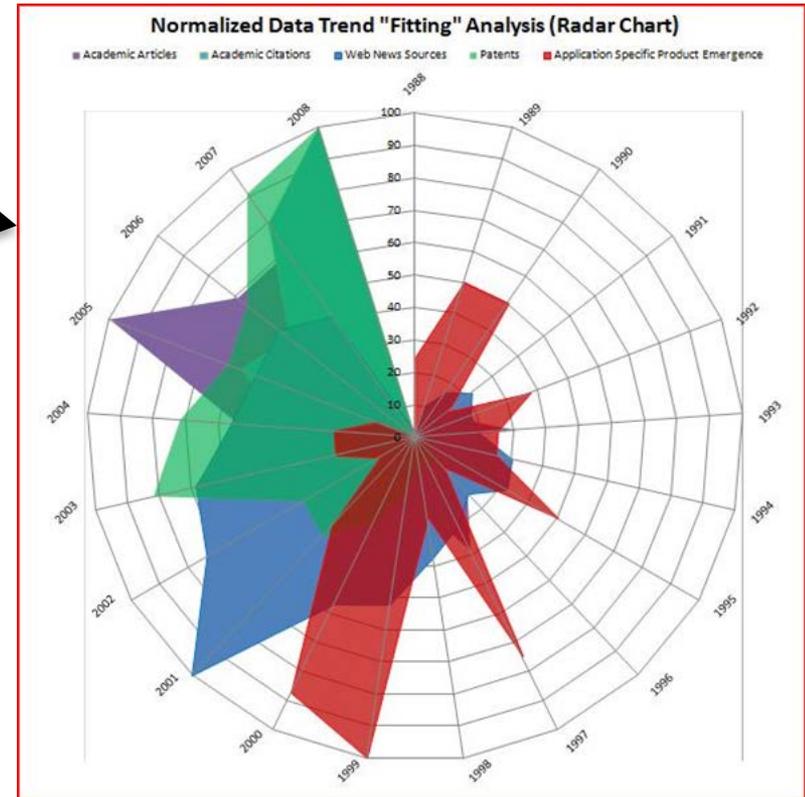
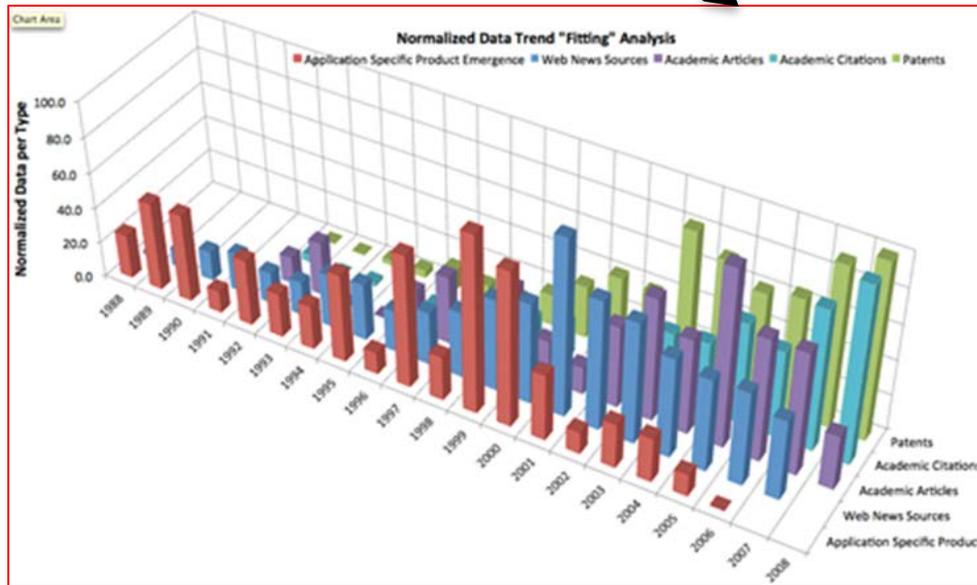
Web News Sources – blue line

Patents = brown line

Academic Citations – green line

Alternate Views of normalized data with trend fitting

- Radar Chart
- Bar Chart



Color Coded as previous slide

Conclusions

- Overall impressions of these datasets reveal that the initial product offerings were a result of support from the IETF standards committee.
- Once companies started offering specific products, scholarly publications began appearing in 1992 with patents appearing within a year (IP protected). These growth trends continued with numerous patents, scholarly publications and citations appearing through 2008.
- We have described an innovative approach to track the emergence and evolution of relevant enterprise applications.
- Future research will continue to address the dynamic nature of these collective networks and websites as well as considering:
 - (1) social networks (multiple connections between people),
 - (2) interdependent technologies emergence/evolution,
 - (3) organizational stakes, and
 - 4) funding sources at the enterprise, state/federal and international levels.

TEM Future Directions - Milestones

- Continue to investigate how the TEM might be strengthened to generalize to other technologies.
- One concern regarding the TEM Milestones is specificity.
 - Continue to more concretely define the various milestones well enough to be falsifiable and more convincingly argued:
 - How measurable are the interdependencies between different datasets or events)
 - How does the analysis strongly show a meaningful repeatable pattern in the data?
 - » A significant body of cognitive bias literature shows that people are excellent at finding patterns in noise, but can machines do this effectively?

TEM Future Directions – Temporal, Domain Specific and Visualization Issues

- Questions to be addressed:
 - Are there characteristic lag times and frequencies associated with scholarly articles, citations and patents?
 - Are such measures of merit domain specific?
 - Is there an efficient way to backtrack articles being cited by patents?
 - What is the effect of different periods of economic times on these datasets?
 - One obvious next step is to explore the effect from the dot.com demise on the same or similar datasets (i.e., technologies assessed in light of the TEM). We would expect to see damping effects.
 - Another question would be the factors that decompose each of the identified entities into sub-relations (a deep dive) to examine the emergence of roles, identity, and organizational structure associated with the R2 analysis within the subject domain of interest.
 - What are the best ways to visualize these relations towards identifying new knowledge, patterns and predictors?

TEM General Future Directions

- The value proposition of the TEM is inherent in:
 - investigating the activities within each milestone,
 - investigating comparisons of disparate data, and
 - articulating the transitioning between milestones.
- Extend the model
 - outside the current linear approach toward making TEM more specific and
 - generally more applicable especially because innovation itself is a much more complex process than is described by a linear path.
- More specific criteria need to be established for the TEM so
 - one can decide what fits and what doesn't in terms of predicting what might or might not happen next.

References

- Patents

- R. K. Abercrombie, "System and Method for Normalizing Data for Analysis of Search Results," U.S. Patent Pending (February 2, 2011).

- Journal Articles

- R. K. Abercrombie, A. W. Udoeyop, and B. G. Schlicher, "A Study of Scientometric Methods to Identify Emerging Technologies via Modeling of Milestones," *Journal of Scientometrics*, Volume 91, Number 2, pp. 327-342, Springer, Akadémiai Kiadó, Budapest, January 11, 2012 (*Online First*), May 2012 (Journal issue), ([DOI 10.1077/s11192-011-0614-4](https://doi.org/10.1077/s11192-011-0614-4)).

- Conference Proceedings

- R. K. Abercrombie and A. W. Udoeyop, "A Study of Scientometric Methods to Identify Emerging Technologies," Proceedings of 13th Conference of the International Society for Scientometrics and Informetrics ([ISSI 2011 Conference](#)), Durban, South Africa, July 4-7, 2011, pp. 1-12 ([PDF](#)).

Oak Ridge National Laboratory: Meeting the challenges of the 21st century



Robert K. Abercrombie, Ph.D.
Email: abercrombier@ornl.gov
Phone: (865) 241-6537

<http://www.ornl.gov/~abe>