



Foresight and Understanding from Scientific Exposition (FUSE)

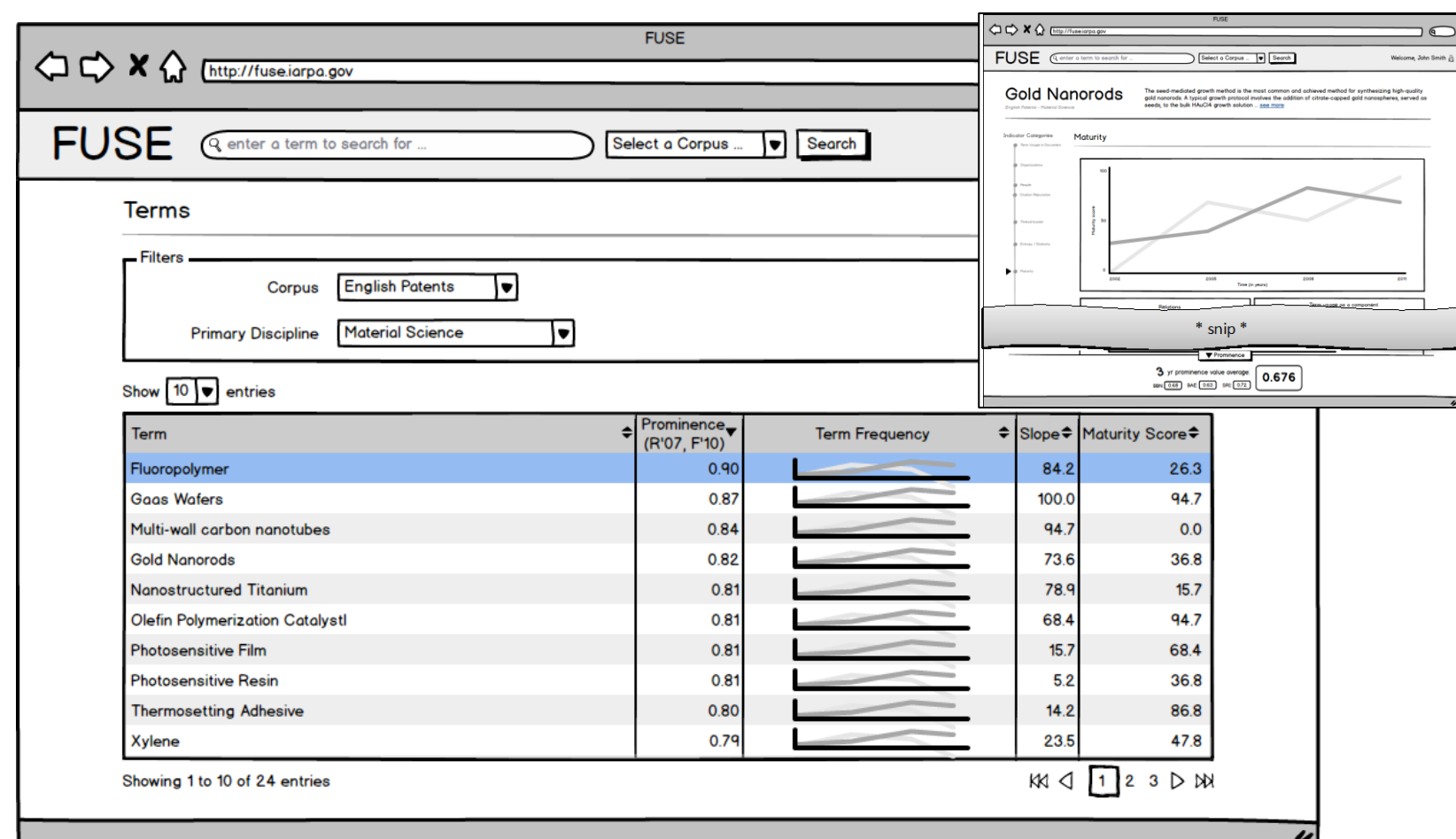


Predicting Technical Emergence from Scientific and Patent Literature

Program Manager: Dr. Dewey Murdick; E-mail: dewey.murdick@iarpa.gov

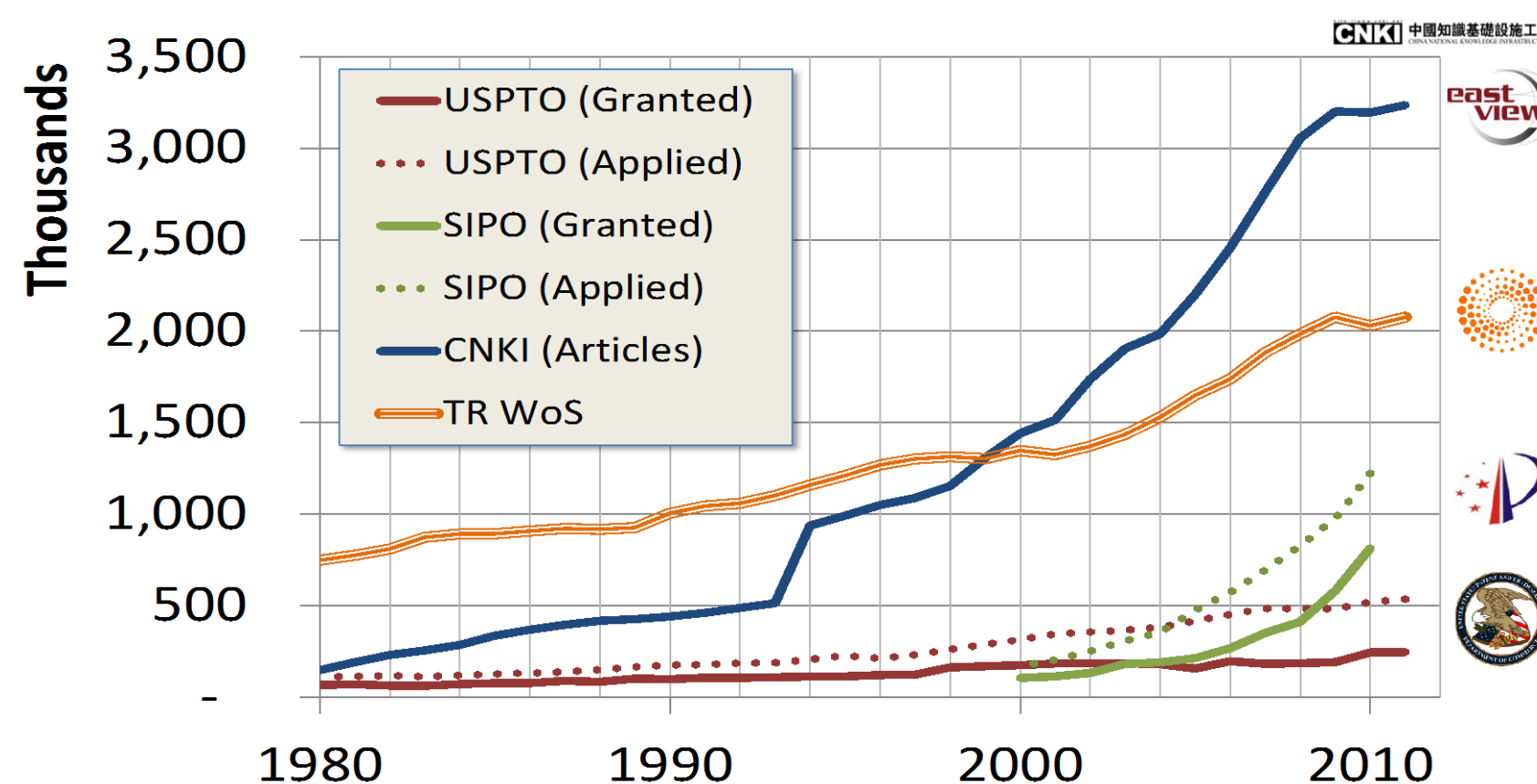
Goal

Mitigate technology surprise by the reliable and early detection of emerging scientific and technological concepts across multiple disciplines through automated analysis of the primary scientific literature



Mockup of a demo user interface under development that will allow navigation of term lists and exploration of indicator values for selected terms

There is a flood of scientific and technical information that provides an opportunity to be used as a source of technical emergence



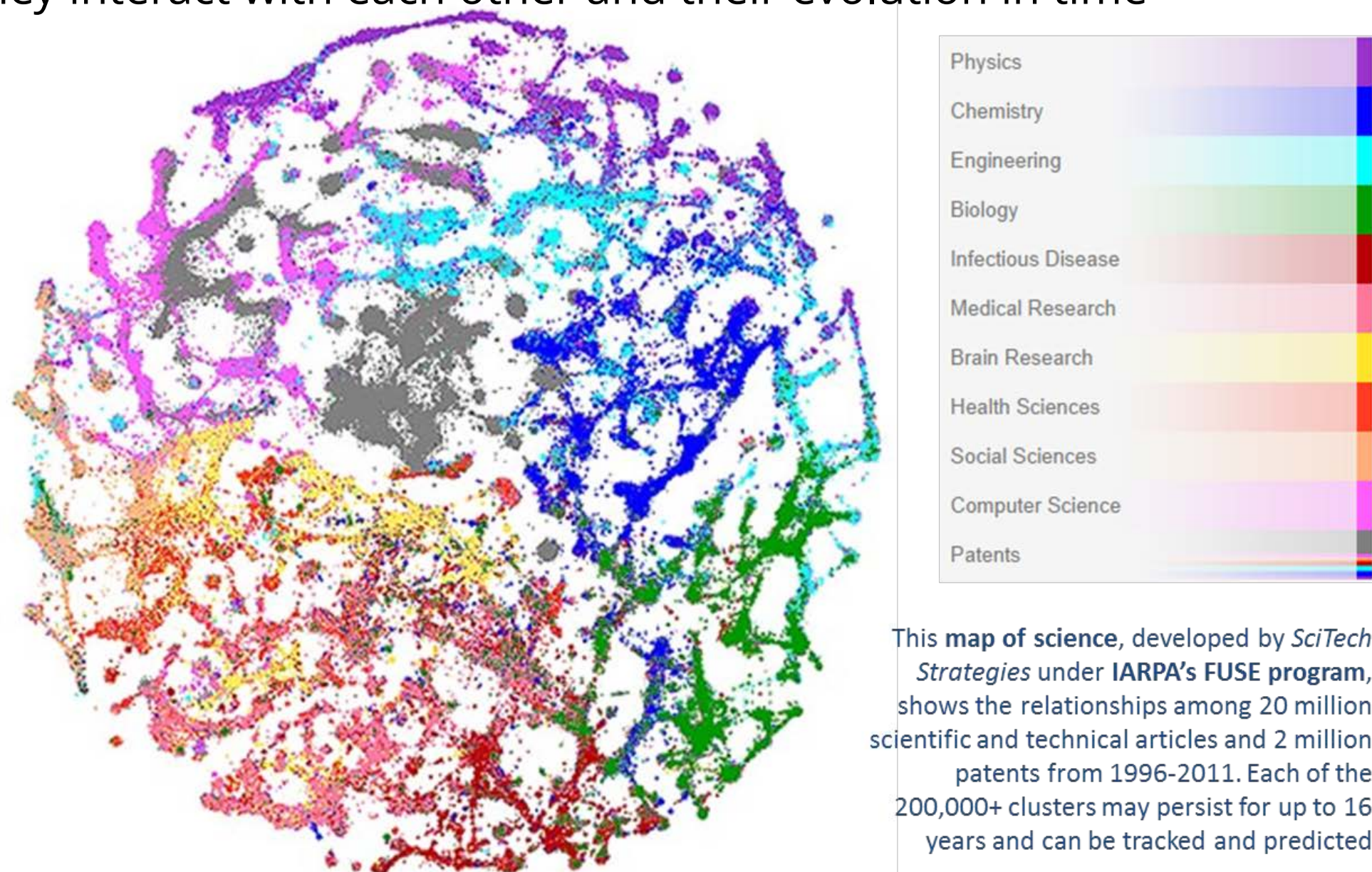
Roughly 600k scientific articles and patents in English and Chinese – millions of pages of text per month.

Approach

Automatically digest English & Chinese scientific literature and patents and discover patterns of emergence for concepts that will likely emerge in several years.



A visualization of the thousands of technical areas in science, how they interact with each other and their evolution in time



This map of science, developed by SciTech Strategies under IARPA's FUSE program, shows the relationships among 20 million scientific and technical articles and 2 million patents from 1996-2011. Each of the 200,000+ clusters may persist for up to 16 years and can be tracked and predicted

Develop, implement and test:

- Theories of technical emergence.
- Technical term extraction.
- Novel feature development (e.g., community detection, topic thread persistence, citing sentence sentiment, technical term classification)
- Indicators of technical emergence.
- Indicator fusion and forecasting models.
- Forecast evidence explanation.

| Technical Terms | Scientific Lit Indicators | | | | | |
|----------------------------------|---------------------------|--------------------------|----------------------|------------------------|------------------------|-------------------|
| | Venue Impact | Term Family Venue Impact | Historical Citations | Average Author H-Index | Average Research Level | Doc Cluster Count |
| DNA microarray | | | | | | |
| Generative model | | | | | | |
| Graphene sheets | | | | | | |
| Heteronuclear magnetic resonance | | | | | | |
| Microquasars | | | | | | |
| Nanosensors | | | | | | |
| Neurexins | | | | | | |
| Qubit | | | | | | |
| Silent transgene | | | | | | |
| Topological insulator | | | | | | |

Scientific Indicators track emerging areas of research and assess maturity, community dynamics, and combine to estimate future prominence.

Program schedule: July 2011 to August 2015.

Evaluation

Score accuracy of 2-5 year forecasts of future prominence of technical terms and score clarity of evidence. Nominated terms represent 1%-20% of a complete term list. Currently Chinese patents tests show:

- 41-45% of nominated terms are correct for 3 year forecasts (precision)
- 53-60% of terms are correctly nominated (recall).
- >90% passed evidence clarity evaluation (S&TI analyst proxy).

All current teams are expected to meet forecasting nomination and evidence clarity targets for English and Chinese data sets.

Potential Impact

Provide an analytic force multiplier to maintain technical vigilance across all disciplines and multiple languages. Aid strategic planning and reshape the way we do technical horizon scanning.

Notable Media Attention:

- Wired** - "Is Innovation Predictable?" 11/25/13 and "The Shape of the Future: From TIPS to FUSE," 12/3/13
- Nature** - "Text-mining offers clues to success," 5/20/14
- The Washington Post** - "What's the next big tech trend? This federal agency thinks it can predict the answer" 9/21/14.

Over 90 journal articles published: See goo.gl/u2kum

Research Teams

Research Teams

BAE SYSTEMS
1790 Analytics
Brandeis University
New York University
Rensselaer Polytechnic Institute

SRI International
Intelligent Information Services Co. (IISC)
University of California, Irvine
University of Illinois, Urbana-Champaign
University of Michigan

Raytheon BBN Technologies
SciTech Strategies
University of Mass, Amherst

Test & Evaluation

MITRE
Booz | Allen | Hamilton
Avian Engineering, LLC
Tarragon Consulting Corp.

OAK RIDGE National Laboratory

Thomson Reuters Ila
IEEE East View
...

ALSO, see the ForeST Program Poster to understand how FUSE synergizes with ForeST to create a combined technology forecasting solution