Tech Mining to Identify Emerging Technology Opportunities

July, 2009

Alan Porter
Director of Research and Development
aporter@searchtech.com
&
Technology Policy & Assessment Center
Georgia Tech

Search Technology

AGENDA

1. Introduction
   a. Open Innovation
   b. Tech Mining
2. Tech Mining Tales
   a. “Research Profiling” Country study: Brazil
   b. Biomaterials Patenting: Technology Opportunities Analyses
   c. NanoBioSensors: Innovation Path Mapping

3. Could these tools work for you?

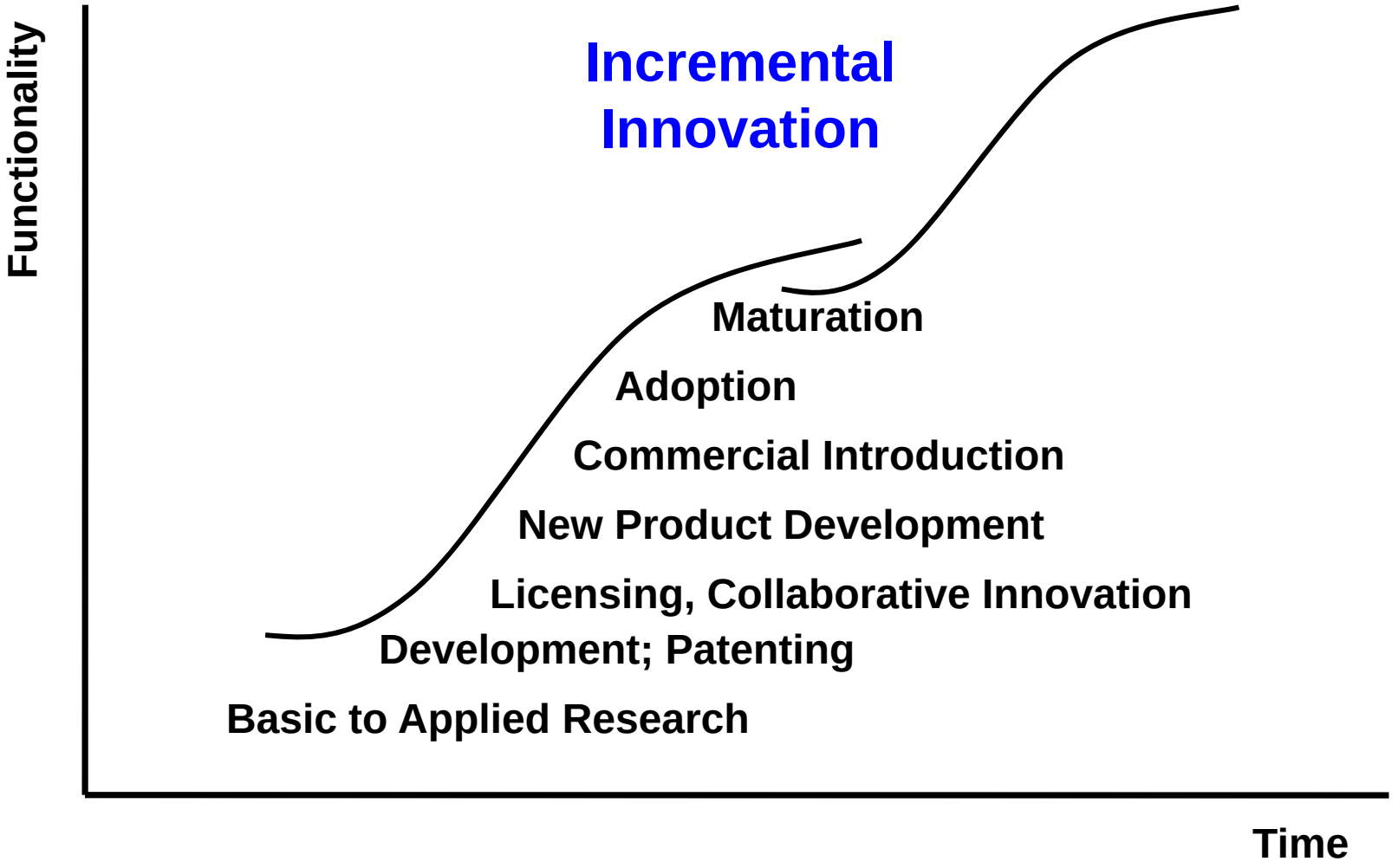
1. Introduction

- Open Innovation
- Tech Mining

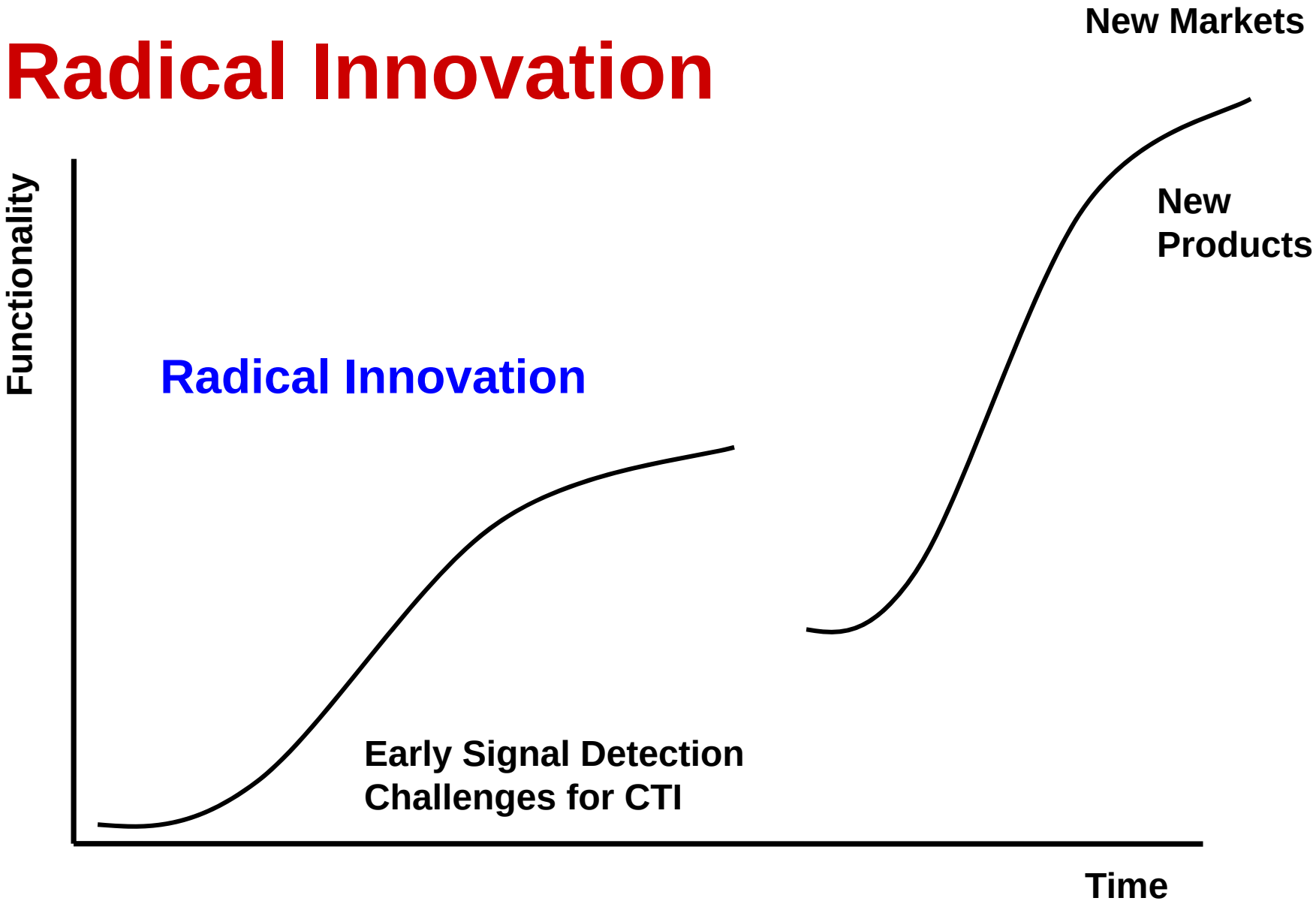
Types of Innovation

- **Incremental Innovation**
- **Radical Innovation**
- **Open Innovation**

A “Linear” view of Innovation Processes

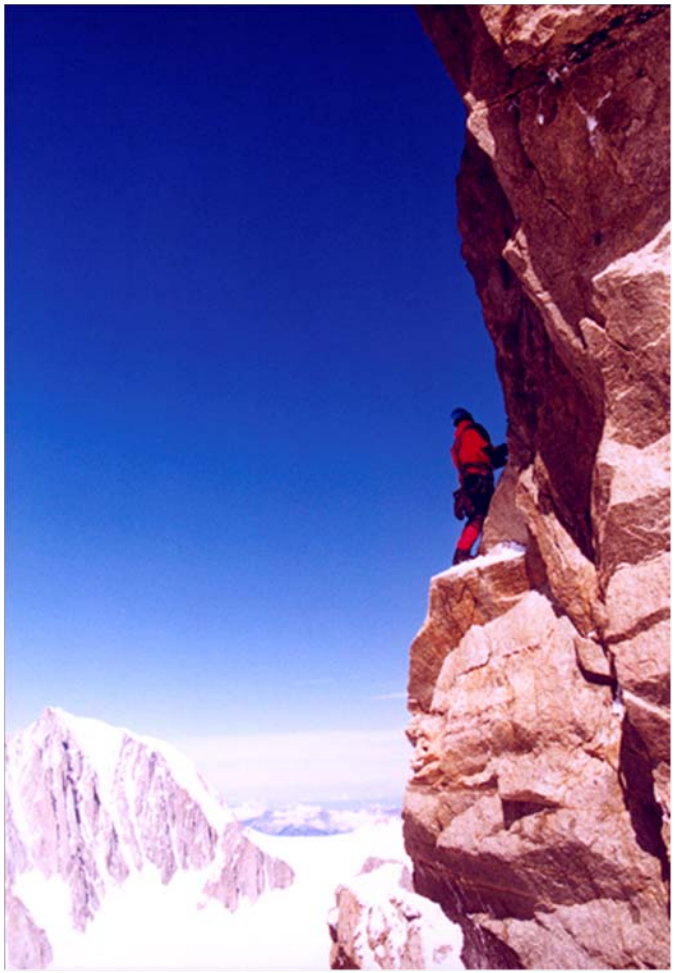


Radical Innovation



Radical Innovation: A Tougher Challenge

- Discontinuous change – implies unfamiliarity
- Increasingly science-based technologies (challenge to predict breakthroughs)
- Need to address socio-technical systems – contextual factors less stationary



Internal Path: R&D Management

- Project selection & evaluation
- Concurrent engineering
- Life cycle design & Sustainability analyses [external issues important too!]

External Path: Technology Acquisition

- Intelligence
- Assessment
- Multiple paths
  - purchase: integrated systems or piecewise
  - customer/supplier collaboration to innovate
  - technology licensing
  - joint ventures

Sources of Innovation

- **Internal Research and Development** **$$$$**
  - **Requires fixed assets and labor**
- **Mergers, Acquisitions** **$$$**
  - **Requires capital**
- **License Intellectual Property** **$$**
  - **Requires the ability to absorb invention**
- **Leverage “Knowledge Economy”** **$**
  - **Requires the ability to leverage external ideas**

Open Innovation: Leveraging the Knowledge Economy

- **“Most innovations fail.”**
- **“And companies that don’t innovate die.”**

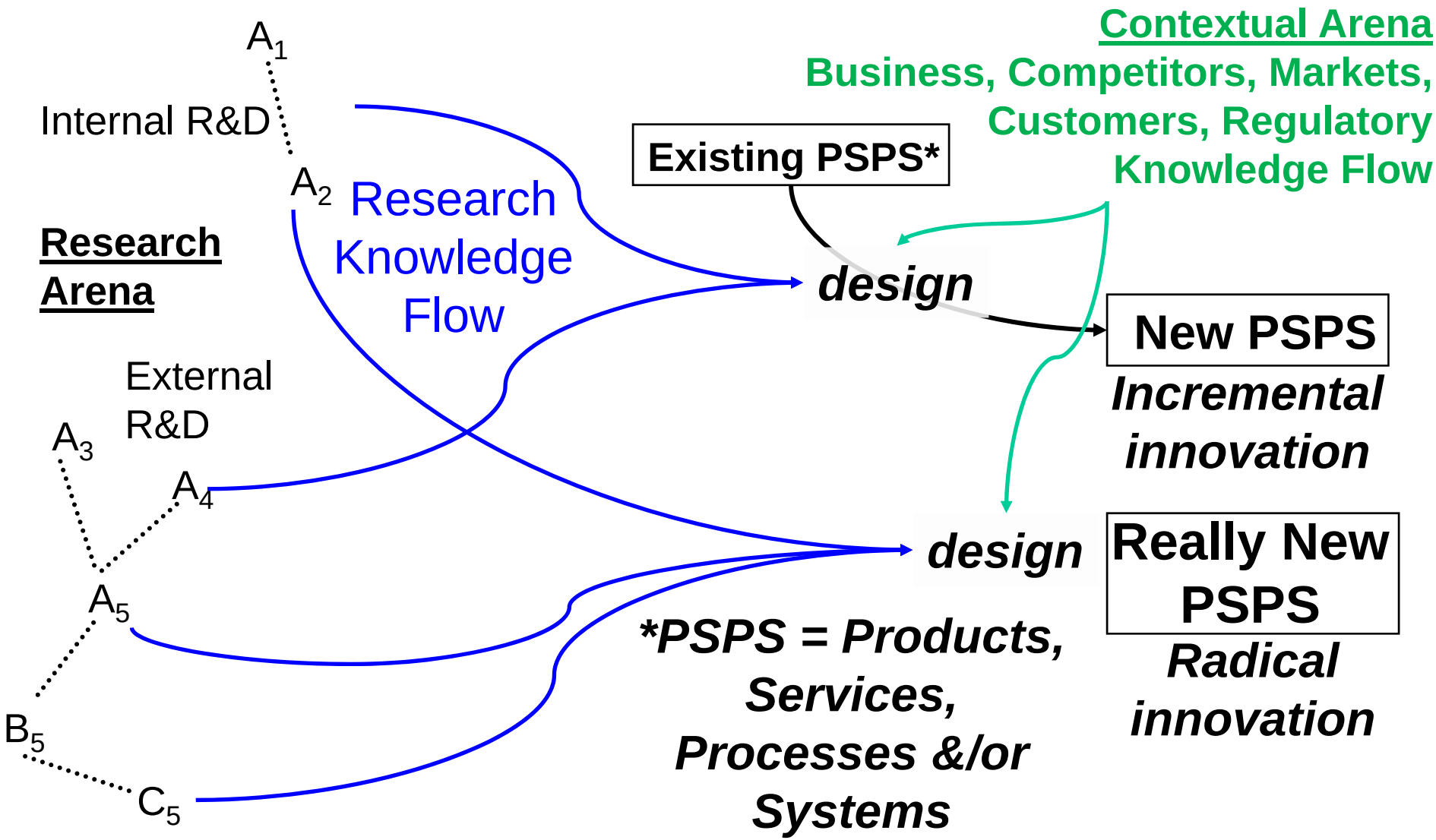
Chesbrough, H. W. (2003). Open Innovation – The New Imperative for Creating and Profiting from Technology, Harvard Business School Press, Boston, MA.

Example: P&G Connect and Develop

- Procter & Gamble's New Model for Innovation
- "35% of company's innovation and billions of dollars in revenue"
- "R&D productivity has increased by nearly 60 percent"
- "R&D investment as a percentage of sales is down from 4.8 percent in 2000 to 3.4 percent today"

"Connect and Develop: Inside Procter & Gamble's New Model for Innovation," Harvard Business Review, Vol. 84, No. 3, March 2006, http://hbswk.hbs.edu/archive/5258.html

Modeling Innovation



Two Innovation Process Dimensions Revisited

- **Extent of change:**
  Continuous vs. Discontinuous (Radical Innovation)
- **Where new knowledge comes from:**
  Internal vs. External (Open Innovation)

Make sense?

Where do your interests fall?

Why Tech Mining?

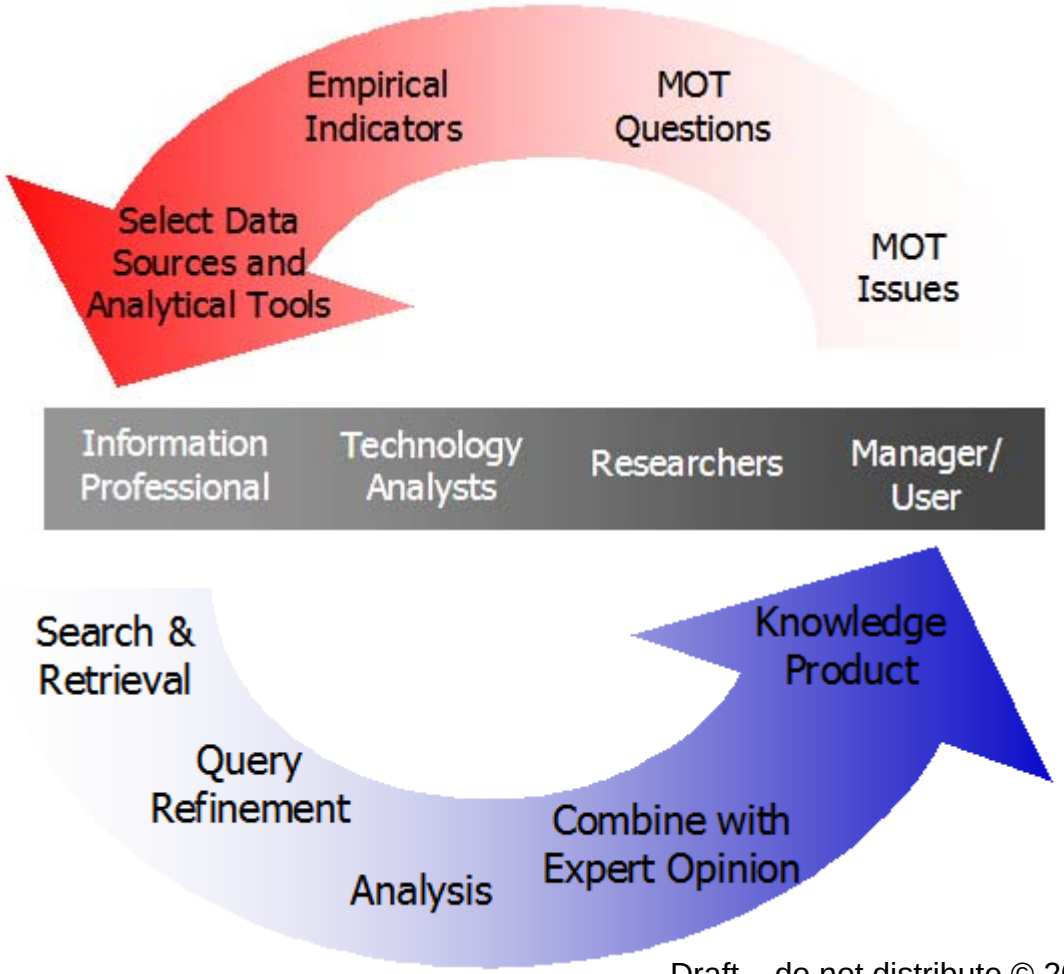
- Welcome to the age of too much information.
- We need to treat: text as data to gain intelligence.
- Mine "ST&I" [Science, Technology & Innovation] information resources to answer technology management questions = Tech Mining.
- Enable Open Innovation

How do you extract effective intelligence from all that Science, Technology & Innovation ("ST&I") information?

Tech Mining

Alan L. Porter and Scott W. Cunningham
John Wiley & Sons Inc., 2005

The Tech Mining Process



Tech Mining Foci

Two main types of analysis

- **Competitive Intelligence**
  - **Focus on one or more target organizations**
- **Technology Intelligence**
  - **Focus on a target technology**

The 13 MOT Issues

- R&D Portfolio Selection
- R&D Project Initiation
- Engineering Project Initiation
- New Product Development
- New Market Development
- Mergers
- Acquisition of Intellectual Property
- Exploiting one's own Intellectual Assets
- Collaboration in Technology Development
- Identifying and Assessing Competing Organizations
- **Tracking and Forecasting Emerging or Breakthrough Technologies**
- Strategic Technology Planning
- Technology Roadmapping

The 39 MOT Questions (Part 1)

- What emerging technologies merit our ongoing attention?
- What facets of this technology development are especially hot?
- What are new frontiers for this technology?
- Are there significant subtypes of the technology?
- How does this technological development fit within the technological landscape?
- What is driving this technological development?
- What are key competing technologies?
- How bright are the development prospects for this technology?
- **What are the likely development pathways for this technology?**
- What are the important component technologies?
- What is the maturity of the component technologies?
- Is there any potential for technology fusion?
- Should we apply for particular patents relating to this technology? (What claims?)
- What does the technology road map look like?

The 39 MOT Questions (Part 2)

- What is the maturation of systems which apply to this technology?
- Which aspects of the technology match our application interests?
- **What are our brightest opportunities in this emerging technology?**
- What societal and market needs do this technology and its applications address?
- What applications offer promise for this technology?
- What are the global opportunities?
- **What is the competitive environment?**
- What is changing in the competitive environment?
- Does this technology offer strong commercialization prospects?
- Who are the available experts?
- **Which universities or research labs lead in this technology?**
- **What are the strengths and gaps within our own organization?**
- Which companies lead in this technology?
- Which companies lead in critical aspects of this technology?
- How strong are the leading companies' R&D teams?
- How do leading companies' development emphases compare to ours?

The 39 MOT Questions (Part 3)

- What other technological strengths does each leading company have?
- Characterize a company's IP relating to this technology.
- What smaller companies or individuals have attractive IP relating to this technology?
- Who is partnering with whom?
- Competitor profiling?
- What companies should we place on watch?
- **Who might be prospects to license our IP?**
- How entrepreneurial is the competitive environment?
- Assessing Competitors

Types of Questions

Text and data mining techniques are good at addressing:

- ***WHO?***
- ***WHAT?***
- ***WHEN?***
- ***WHERE?***

And, especially, combinations of these

Additional questions usually require more human insight:

- **HOW?**
- **WHY?**

Innovation Indicators

- **Technology Life Cycle Indicators**
  - e,g, growth curve location & projection

- **Innovation Context Indicators**
  - e.g., presence or absence of success factors (funding, standards, infrastructure, etc.)

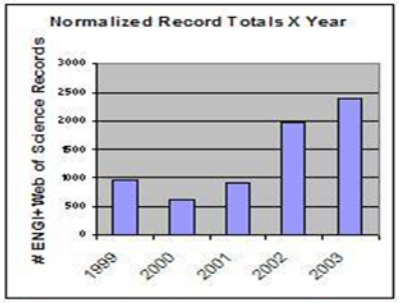
- **Product Value Chain and Market Prospects Indicators**
  - e.g., applications, sectors engaged

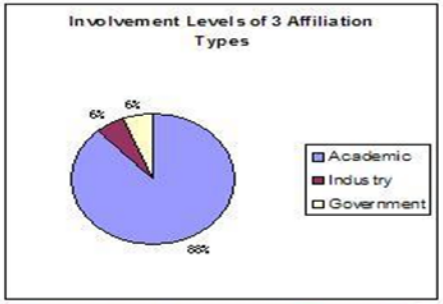
A few of the 200+ MOT Indicators

- Time slice profiles by main topics showing changing research emphases
- Velocity (rate of patenting and rate of change of patenting)
- Research Activity Landscape maps – “blackspaces”?
- Score relative science base (% of patents citing R&D papers, and whose)
- Research networking: Map co-author & co-inventor teaming
- Benchmarking: North Carolina’s relative strengths [academic + industry]

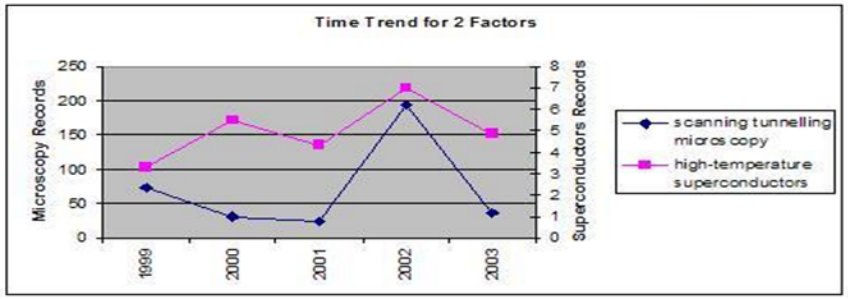
One Pagers

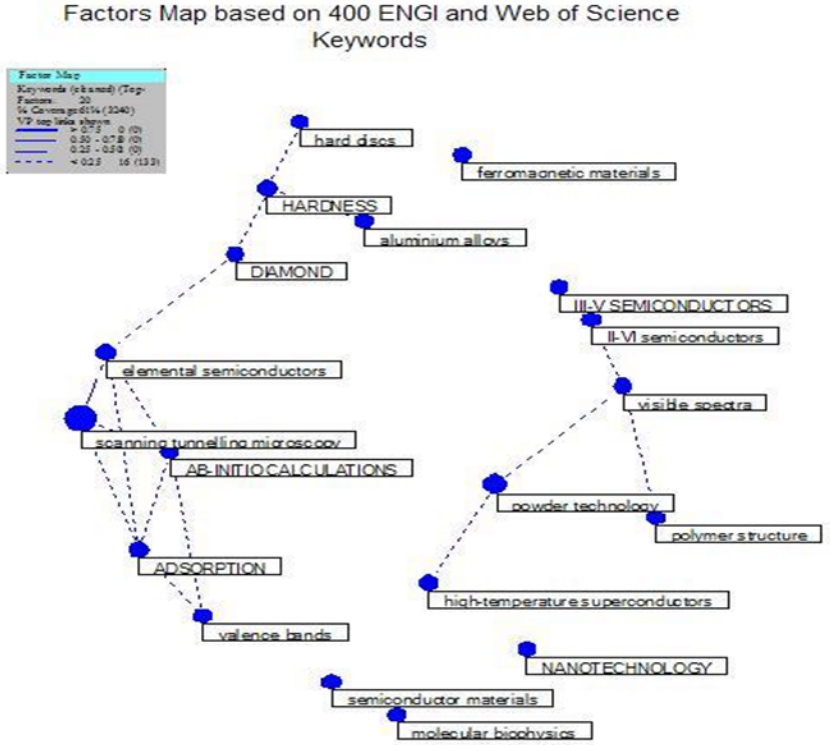
Nano Surface Technologies at a Glance





Country	# Records	Country	# Records
Japan	1140	Germany	445
USA	781	UK	261
China	644	South Korea	255





Factor Name	USA Affiliation	Author
scanning tunnelling microscopy[31]	Sandia Nat. Labs., Livermore, CA, USA [3]; Dept. of Phys. & Astron., Ohio Univ., Athens, OH, USA [3]; Dept. of Phys. & Astron., Univ. of Toledo, OH, USA [3]	Bartelt, N C [3]; de la Figuera, J [3]; Clarke, R [3]
powder technology[11]	Dept. of Phys. & Astron., Rutgers Univ., Piscataway, NJ, USA [3];	Kear, B H [4]; Skandan, G [3]; Sadangi, R K [2]

How to do Tech Mining: 8-steps

1. Spell out the questions and how to answer them
2. Get suitable data
3. Search (iterate)
4. Import into text mining software (VantagePoint)
5. Clean the data
6. Analyze & interpret
7. Represent the “innovation indicators” well – communicate!
8. Standardize and semi-automate where possible

Mining for what?
Patterns (+ “gold nuggets”)

Use Co-word Bibliometrics/Co-occurrence statistics to find relationships

- Count the number of times words appear together in a set of documents
- The higher the co-occurrence, the stronger the potential relationship

Word 1  **Word 2**

What to mine? Fielded Text Data

What does the raw fielded text data data look like?

- Fielded
- Delimited
- Structured

Author(s) AU:	Aragane, J.; Urushibata, H.; Murahashi, T.
Affiliation AF:	Central Res. Lab., Mitsubishi Electr. Corp., Hyogo, Japan
Title TI:	Proton deficiency in a phosphoric acid fuel cell
Journal JN:	Journal of the Electrochemical Society
Date DA:	Jan. 1995
Record Type RT:	Journal paper
Subject(s) SU:	electrochemical electrodes. fuel cells. voltammetry (chemical analysis)
Abstract AB:	In the in situ cyclic voltammetry that we have developed, the anode potential was shifted, and hysteresis of the anode potential appeared during the scanning of the cathode potential when the hydrogen partial pressure was decreased to ca. 5% at the anode. It was concluded that proton deficiency was responsible for the hysteresis at the anode. Further, at this hydrogen pressure in real single-cell operation, the cathode potential (iR-free) deviated from Nernst's law, which showed the interaction between the anode potential and the cathode potential. This means that anode polarization influenced cathode polarization under a low hydrogen partial pressure where the proton deficiency occurred
Class. Codes CC:	A8630G. A8280F. A8245. B8410G
Date Indexed DI:	9502

Tech Mining – 6 information types

Technical Information

A. ST&I (Science, Technology & Innovation) Databases (e.g., Web of Science, INSPEC, Micropatents)

B. Internet Sources (e.g., Googling)

C. Technical Expertise

Contextual Information

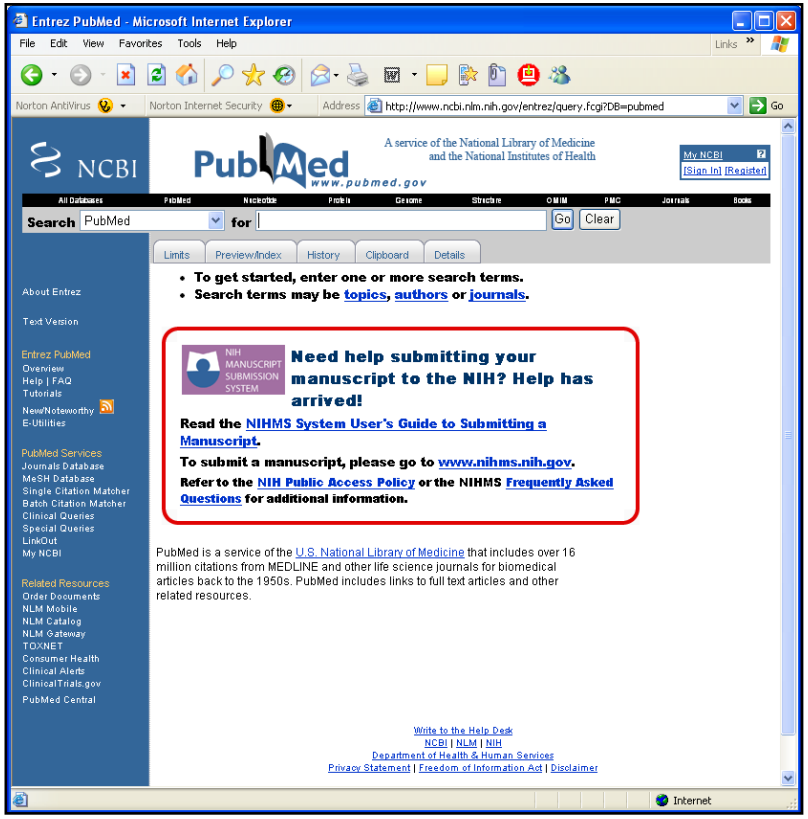
D. Business, competition, customer, popular, policy content Databases (e.g., Lexus-Nexus, Factiva)

E. Internet Sources (e.g., blogs, website profiling)

F. Business Expertise

The Search

- Where to Search
  - Any Structured Text Databases
    - Dialog, STN, Micropatent, Excel,etc...
- How to Search
  - Standard Search Interfaces
- Search Strategy
  - Inclusive/Cast a broad net
- Downloading Search Results
  - The more fields the better
    - Dialog - format 9, Tags On, Highlighting Off
    - STN – IALL or ALL



How to do Tech Mining: 8-steps

1. Spell out the questions and how to answer them
2. Get suitable data
3. Search (iterate)
4. Import into text mining software (VantagePoint)
5. Clean the data
6. Analyze & interpret
7. Represent the “innovation indicators” well – communicate!
8. Standardize and semi-automate where possible

Questions/Comments?

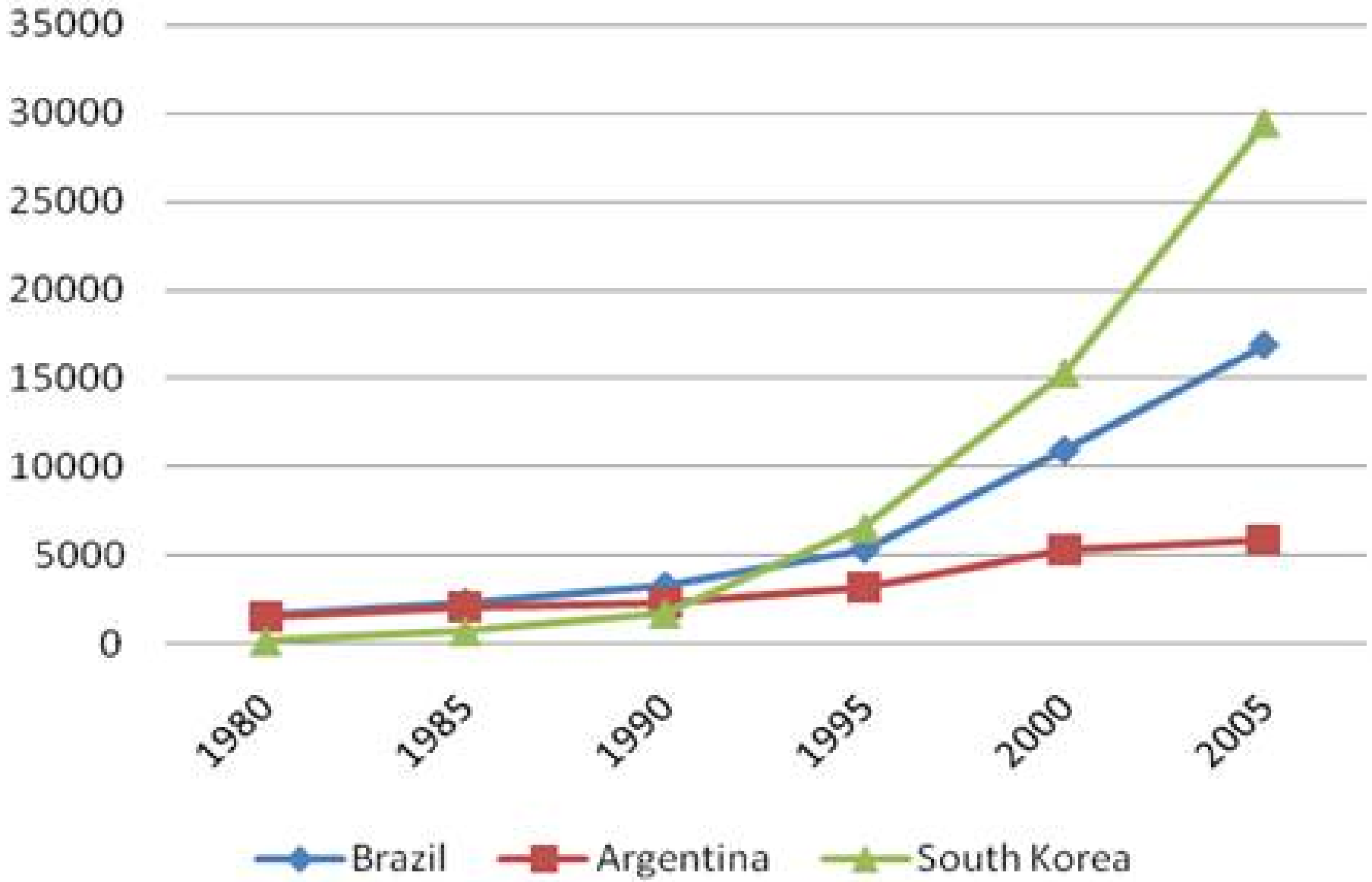
2. Tech Mining Tales

a. “Research Profiling” Country study: Brazil
b. Biomaterials Patenting: Technology Opportunities Analyses
c. NanoBioSensors: Innovation Path Mapping

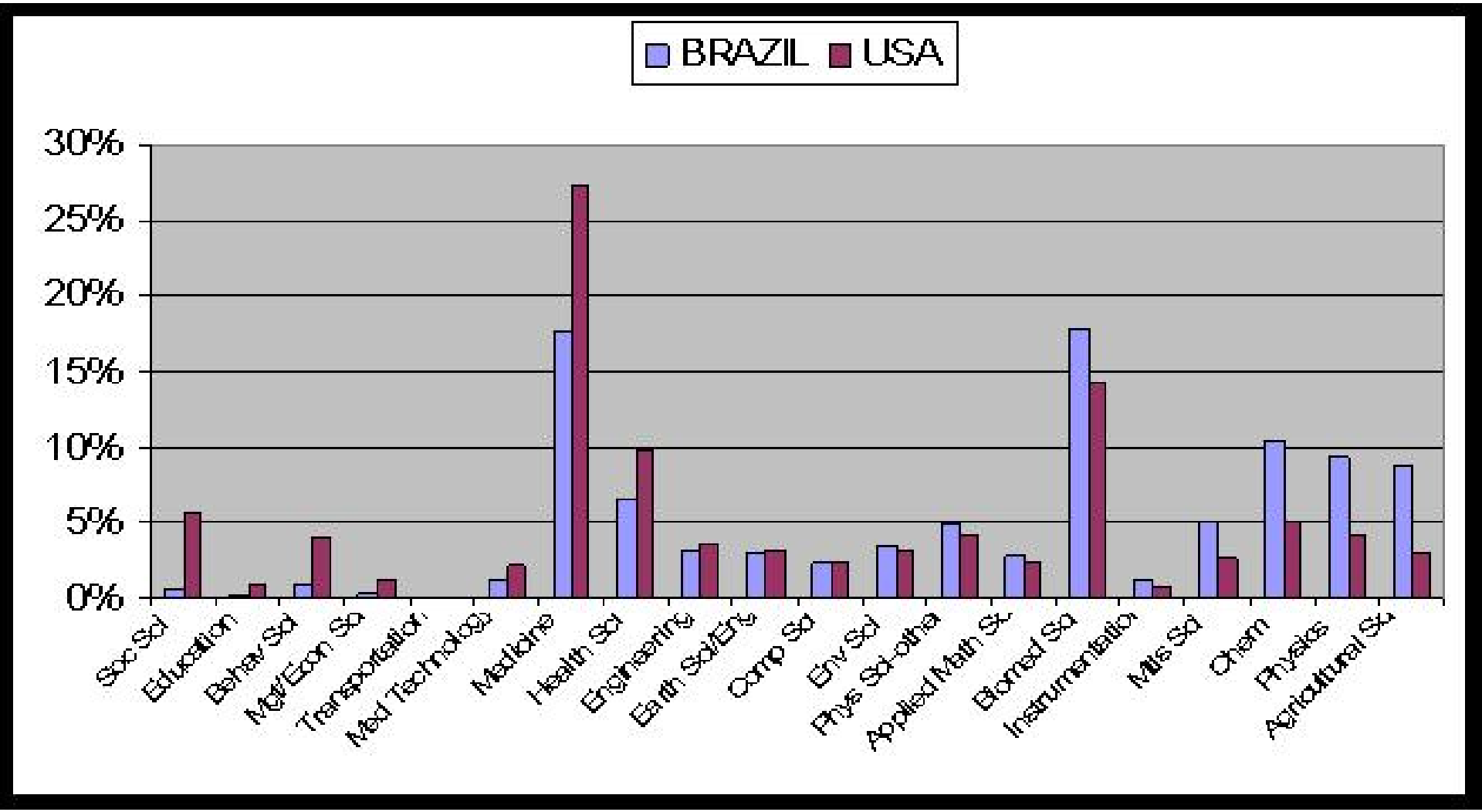
Brazil Country Study, 2007
Search Technology, Inc., for US Navy

- Searches in Web of Science and EI Compendex
  - 17,965 articles with a Brazilian author in 2006 [SCI/SSCI]
  - 21,753 articles, 2003-2007 (partial) [Compendex]
  - Additional 5-year snapshots
- Overview of Brazilian S&T Institutions
- Resource for the Navy to locate potential research partners in Brazil

SCI/SSCI Publication Trend



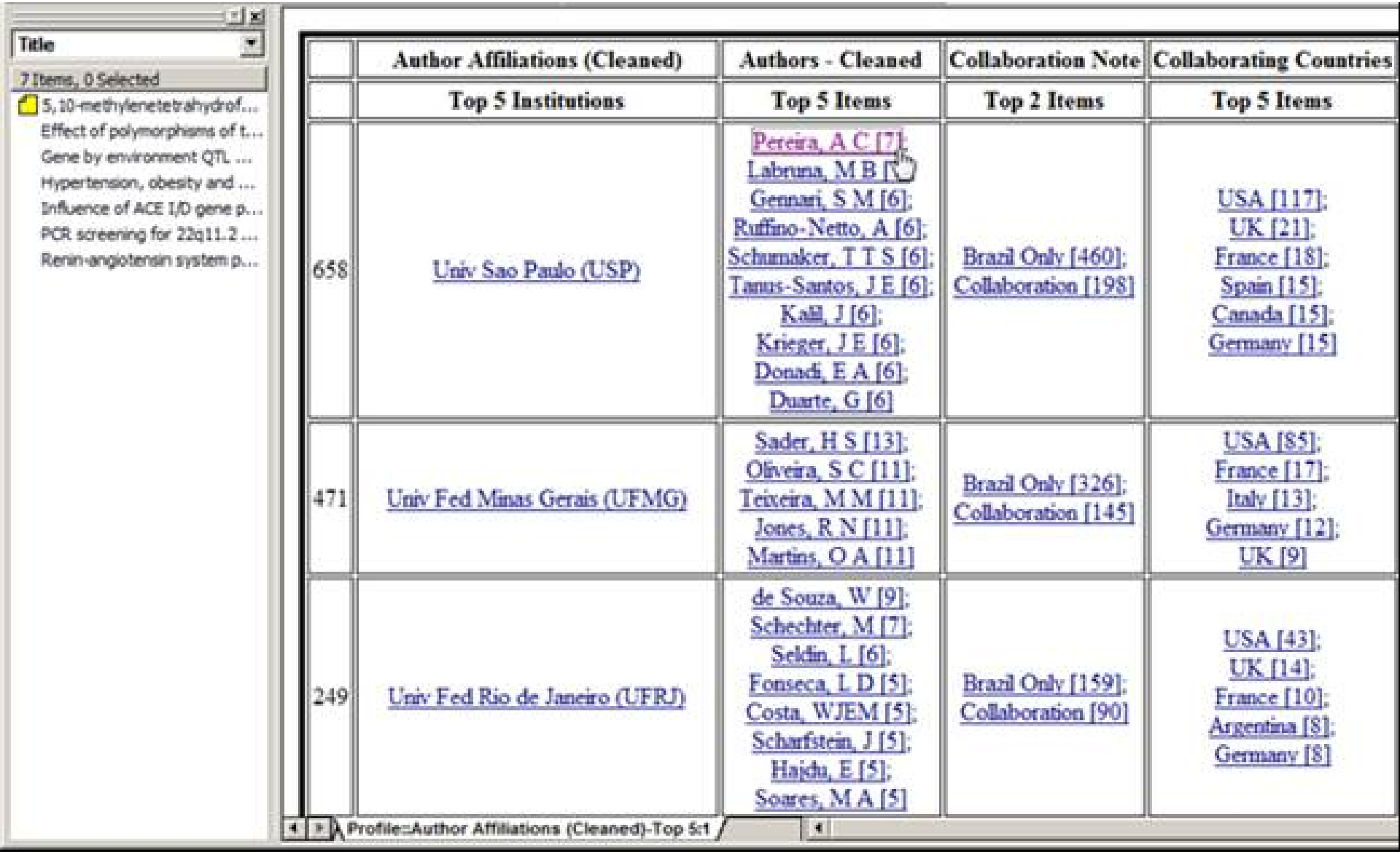
Relative Research Emphases [SCI/SSCI]



Top 10 Collaborating Countries	# SCI/SSCI, 2005
Brazil	16936
USA	1965
France	694
UK	671
Germany	645
Canada	372
Argentina	343
Italy	324
Spain	296
Netherlands	211
Japan	207

Collaborating Country	Taxonomy Themes	Journals
# with Brazil		SCI/SSCI 2006
USA [2049]	Microbiology, Infectious Diseases & Treatments [449] Diagnosis and Treatment of Chronic Diseases [395] Computer Hardware, Networks, and Algorithms [234] Molecular & Cell Biology; Biophysics [220] Particle Physics & Field Effects [206]	Physical Review B [35] Physical Review Letters [34] Physical Review D [26] Physics Letters B [18] Astrophysical Journal [17]

Sample Research Profile for Leading Organizations



	Author Affiliations (Cleaned)	Authors - Cleaned	Collaboration Note	Collaborating Countries
	Top 5 Institutions	Top 5 Items	Top 2 Items	Top 5 Items
658	Univ Sao Paulo (USP)	Pereira, A C [7]; Labruna, M B [6]; Gennari, S M [6]; Ruffino-Netto, A [6]; Schumaker, T T S [6]; Tanus-Santos, J E [6]; Kalil, J [6]; Krieger, J E [6]; Donadi, E A [6]; Duarte, G [6]	Brazil Only [460]; Collaboration [198]	USA [117]; UK [21]; France [18]; Spain [15]; Canada [15]; Germany [15]
471	Univ Fed Minas Gerais (UFMG)	Sader, H S [13]; Oliveira, S C [11]; Teixeira, M M [11]; Jones, R N [11]; Martins, O A [11]	Brazil Only [326]; Collaboration [145]	USA [85]; France [17]; Italy [13]; Germany [12]; UK [9]
249	Univ Fed Rio de Janeiro (UFRJ)	de Souza, W [9]; Schechter, M [7]; Seldin, L [6]; Fonseca, L D [5]; Costa, WJEM [5]; Scharfstein, J [5]; Hajdu, E [5]; Soares, M A [5]	Brazil Only [159]; Collaboration [90]	USA [43]; UK [14]; France [10]; Argentina [8]; Germany [8]

Brazil Country Study

- Profile the research activity of an entity of interest [a country, a company, a research group]
- Match against your own interests & strengths
- Identify promising areas & partners for collaboration
- Comments or questions?

- Innovation Mapping - White space Analysis for Biomaterials in Complex Patent Landscapes

Alan L. Porter, Georgia Tech
alan.porter@isye.gatech.edu
Michael Kayat, UTEK Corporation
mkayat@utekcorp.com

ICIC – 2007

Premises

- The Challenge: Expedite Innovation
- The Foundation: Innovation Process Modeling
- The Tools: Tech Mining
- The Result: Innovation Mapping for intelligence & foresight [illustrated for Collagen Gel Opportunities]

How to Enhance Innovation?

- Manage based on strong competitive technical intelligence - CTI
  - Those who do so will win
  - Those who do not will lose
- Innovation Mapping can show the way
  - Understand the system & its key leverage points
  - Identify technology opportunities

Technological Innovation: The Conceptual Bases

- Recognize Technological Capabilities
- Focus on changes in function – of products, processes, or services
- Draw upon models of technological change
  - Innovation (life cycle) processes
  - Technology substitution, transfer & diffusion
- Promote “OI” – Open Innovation

Innovation Mapping Elements I

Technological Landscape

- Technological Advance
  - Capabilities
  - Applications
- Competitive/Collaborative Milieu
  - Key players
  - Profile their strengths & orientation
- Contextual Influences?
  - Stakeholders & Concerns
  - Regulations, standards, funding
- Future prospects

Innovation Mapping II: Market Prospects
(not the emphasis in today's presentation)

- Market Opportunities
  - Sectors & Locations
  - Forecast
- Customer Needs
  - Currently identified & extrapolated
  - Lead users
- Innovation Implementation
  - External obstacles
  - Internal obstacles

“White Space” Analysis – Misnomer?

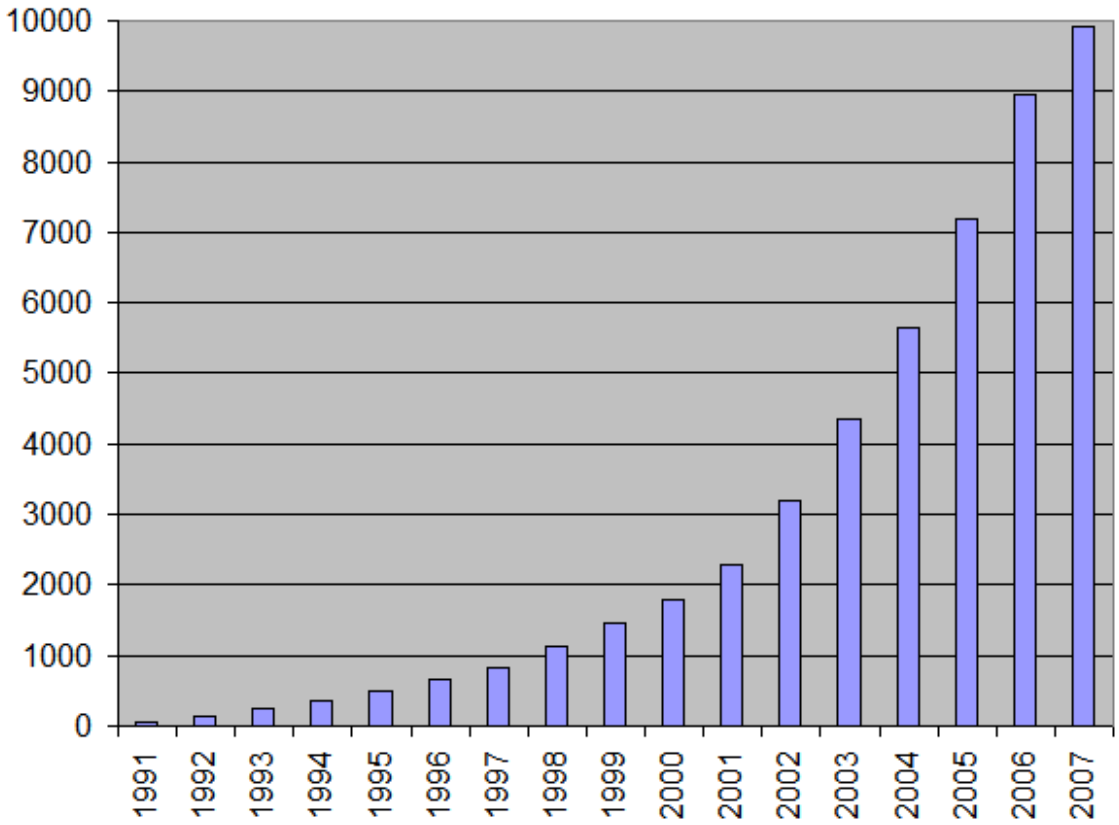
- Complex, multidimensional milieu
- Reduction to 2-D or 3-D is precarious
- Finding what’s missing (“not there”) is dicey
- Better to focus on “what is” along selected dimensions
- Provide derived empirical knowledge to a diverse expert body with requisite domain knowledge to stimulate discourse

Case: Polymer Biomaterials

- “Are there any new market spaces for [*your idea here*] which look relatively free of existing IP?"
- Market Prospects: A Quick Glance
  - Implants: global spending nearly $120 billion/year
  - Biocompatible materials market projected to $12 billion in 2008
  - Biomaterial polymers reached $7 billion in 2003

Polymer Biomaterials

- Micropatents search yielded ~10,000 patents (not comprehensive)
- This constitutes the broad picture
- Could extend by examining research funding, research publications, business activity, etc. searches & analyses (not today!)



Application Domains

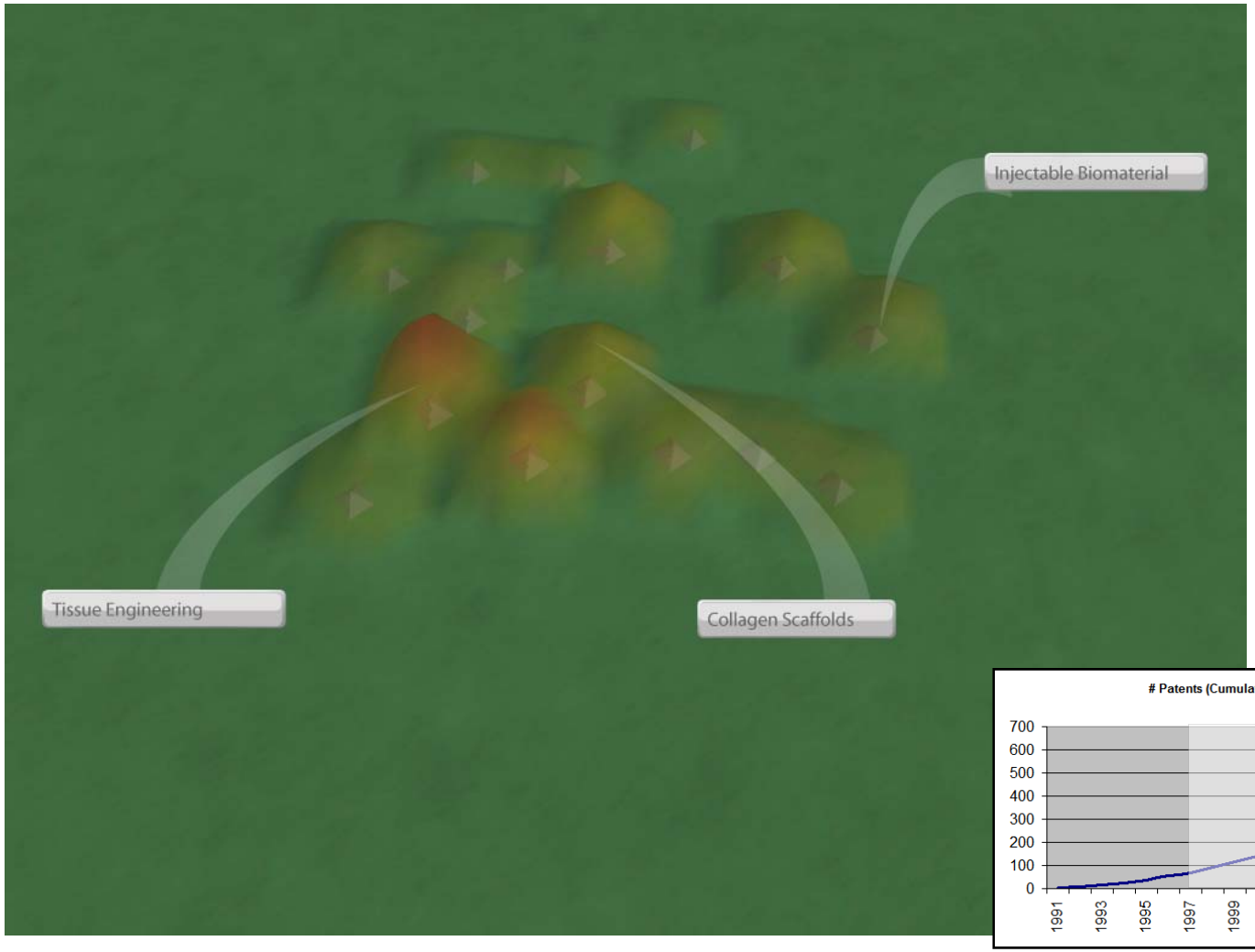
Leading International Patent Classes (IPC codes)

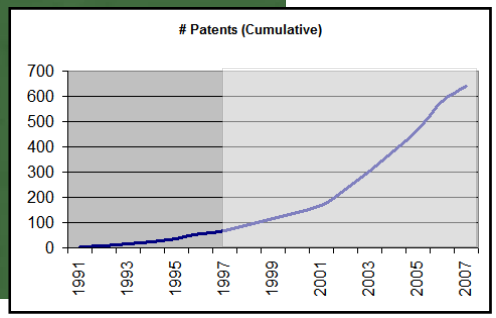
IPC Classes	**#**
A61K-Preparations For Medical, Dental, Or Toilet Purposes	4148
A61L-Methods Or Apparatus For Sterilising Materials Or Objects In General; Disinfection, Sterilisation, etc.	4043
A61F-Filters Implantable Into Blood Vessels; Prostheses; Orthopaedic, Nursing Or Contraceptive Devices; etc.	2782
C12N-Micro-Organisms Or Enzymes; Compositions Thereof; etc.	1477
A61B-Diagnosis; Surgery; Identification	1214

Focusing: For this illustration

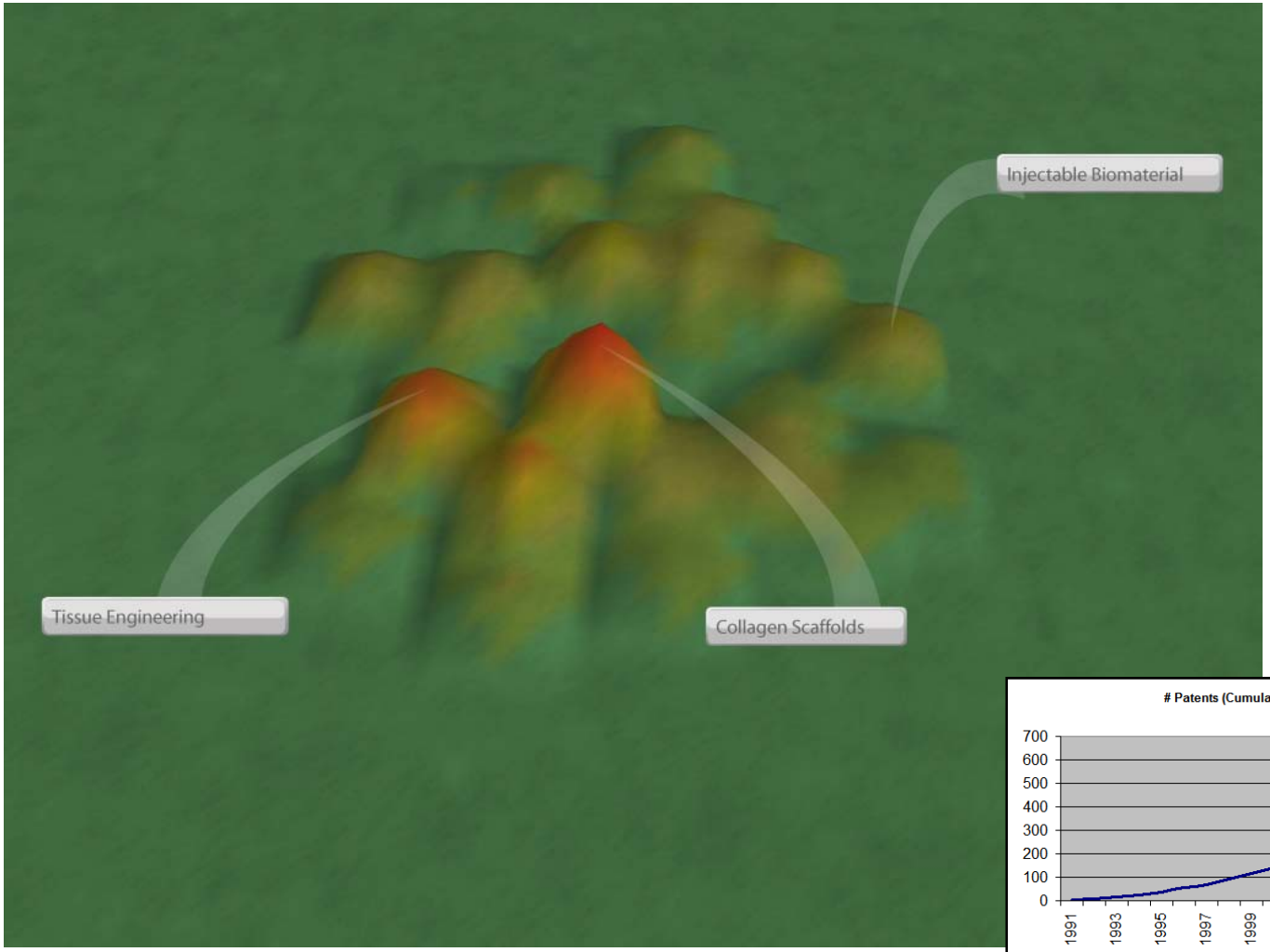
- Multidimensional – various ways to cut 10,000 Biomaterials Patent set
- We selected on 2 dimensions:
  - Technology Type: Fibrous structural proteins [searched these patent records for collagen, fibrillin, laminin, proteoglycan, elastin, ECM, …]
    - ~2200 patents
  - Target Application Biosystem: skin [or derm] in claims
    - ~640 patents

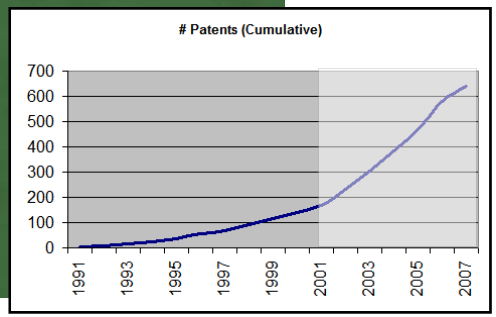
Polymer Biomaterials : fibrous structural proteins : skin 1991-1997 (68 records)



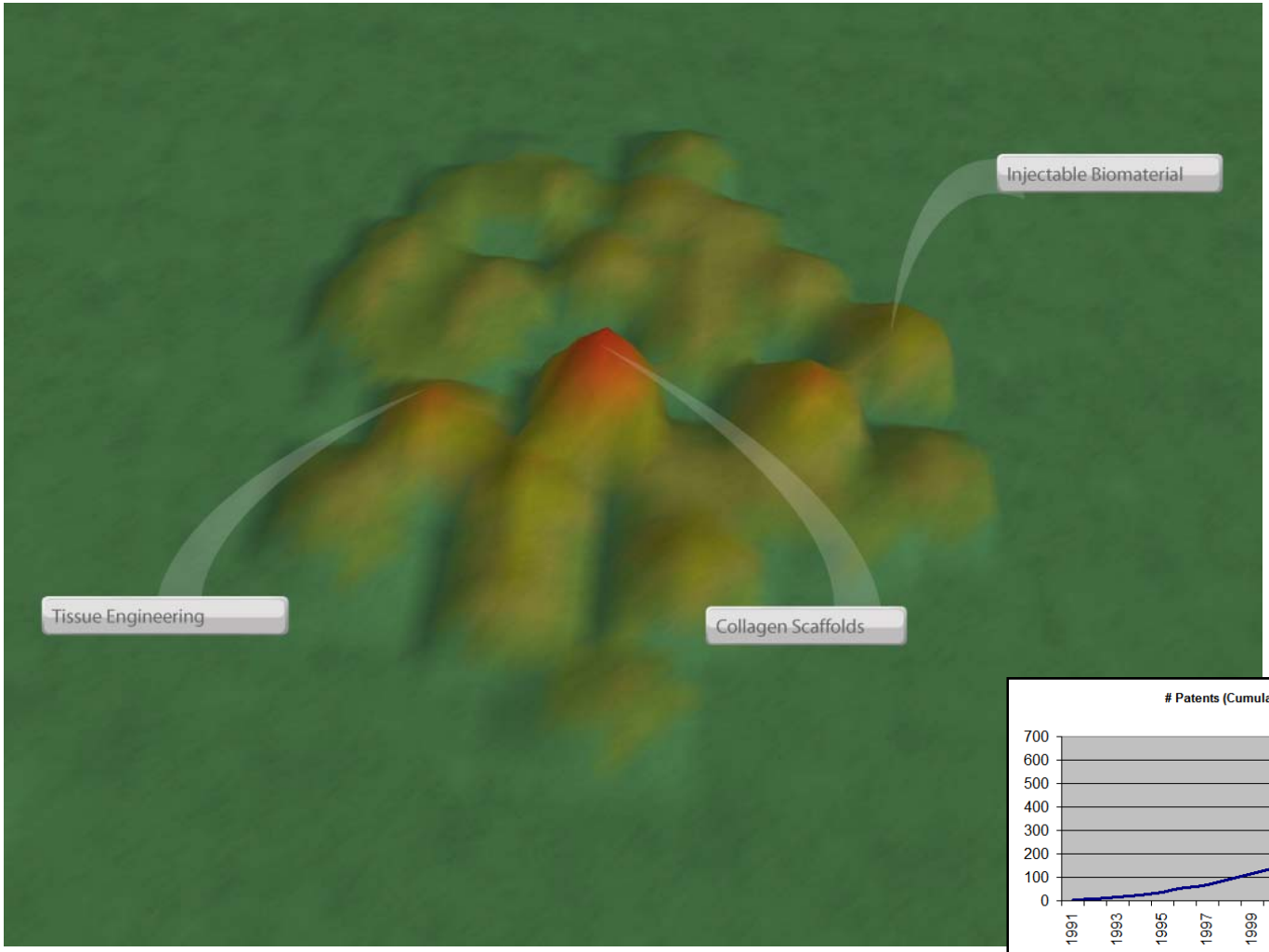


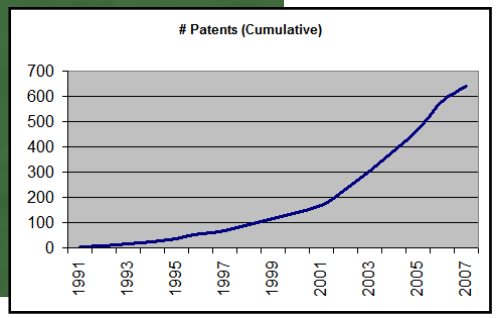
Polymer Biomaterials : fibrous structural proteins : skin 1991-2001 (168 records)





Polymer Biomaterials : fibrous structural proteins : skin 1991-2007 (640 records)

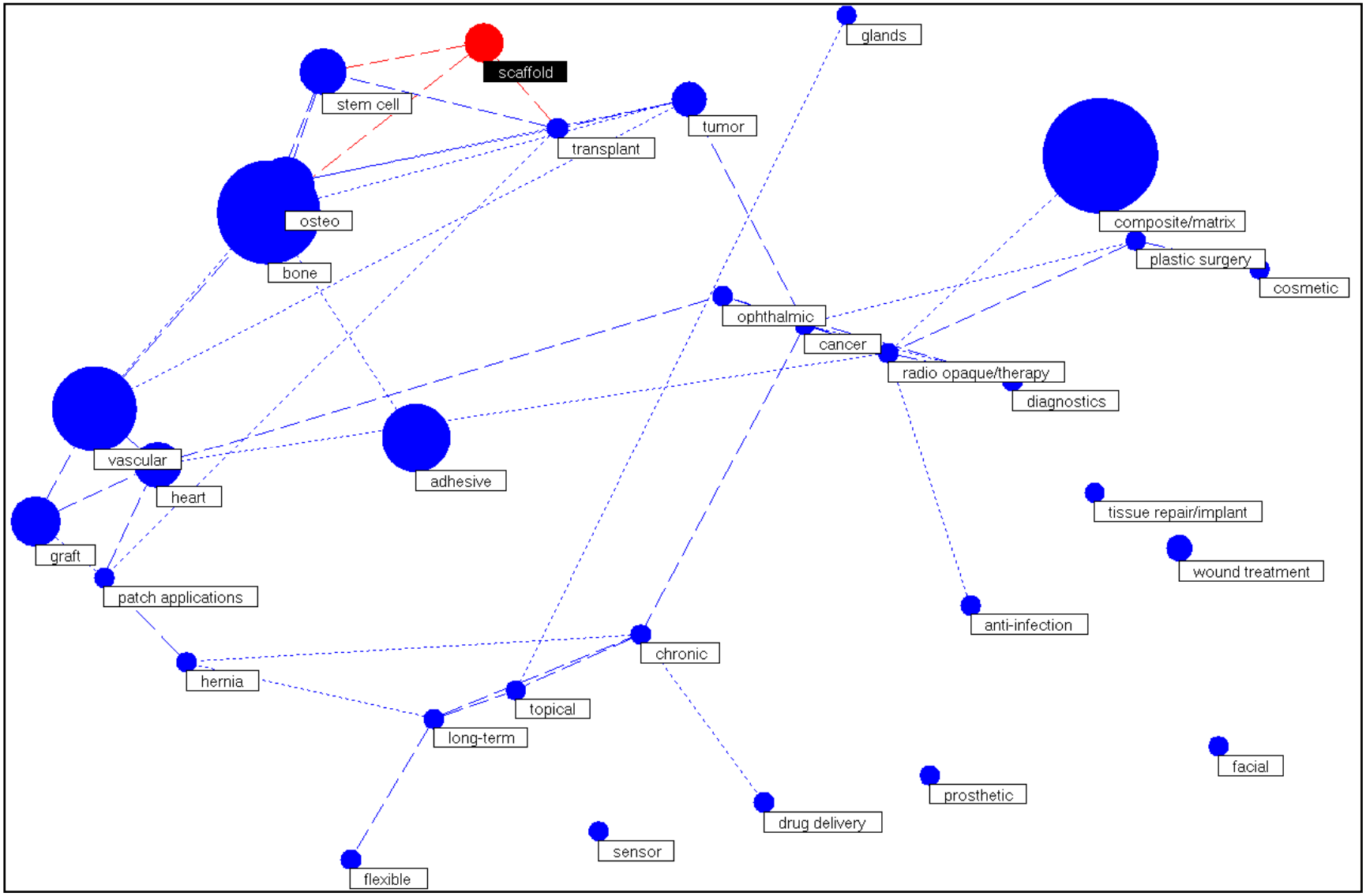




Topic Detection

- Patent records lack keywords
- Class codes are very helpful, but not highly specific
- One approach: "entity extraction" – apply a dictionary or rule-set to get at key phrases
  - e.g., Extracellular matrix (ECM) classes of biomolecules
- Another: apply a general-purpose natural language processor to extract terms (noisy); browse and classify large collections interactively.
  - e.g., Application/property terms in Claims

Application/Property Term Associations



Multi-dimensional Views

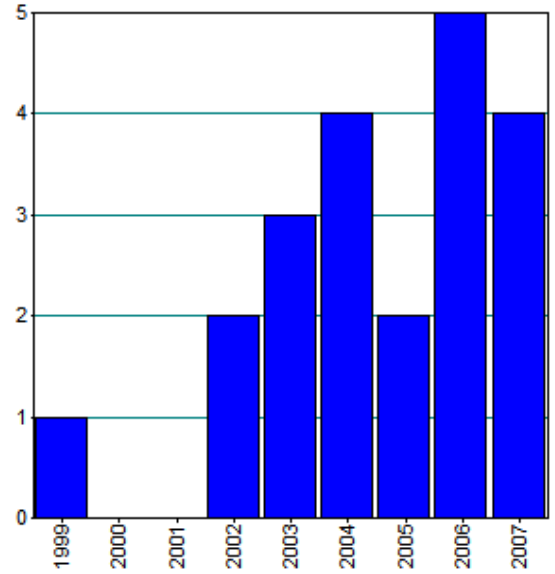
		laminin [57]; elastin [50]; Chondroitin [46]; fibronectin [41]; glycosaminoglycans [37]; keratan [31]; proteoglycan [31]; heparan [17]	
156	graft		8% of 156
123	scaffold	collagen [103]; Hyaluronic [72]; elastin [54]; fibronectin [50]; glycosaminoglycans [45]; laminin [41]; Chondroitin [31]; proteoglycan [25]; heparan [21]; keratan [18]	11% of 123
110	tumor	collagen [85]; fibronectin [55]; Hyaluronic [50]; elastin [36]; glycosaminoglycans [30]; Chondroitin [21]; laminin [17]; keratan [12]; proteoglycan [10]; heparan [8]	5% of 110
		collagen [62];	

Profile::Ap-Property Terms from Claims-Top 19:2

Assignee / Applicant (Cleaned) (1)

3	↑↑↑	CorMatrix Cardiovascular, Inc
2	↑↑↑	Celxcel Pty LTD
2	↑↑↑	Encelle, Inc
2	↑↑↑	NVR LABS, LTD
2	↑↑	The Board of Regents, The University of Tex
1	↑↑	ASTACHOV, Liliana
1	↑↑	CHAIKOF, Elliot, L
1	↑↑	Emory University
1	↑↑	FAUCHER, Keith, M
1	↑↑	FENG, June

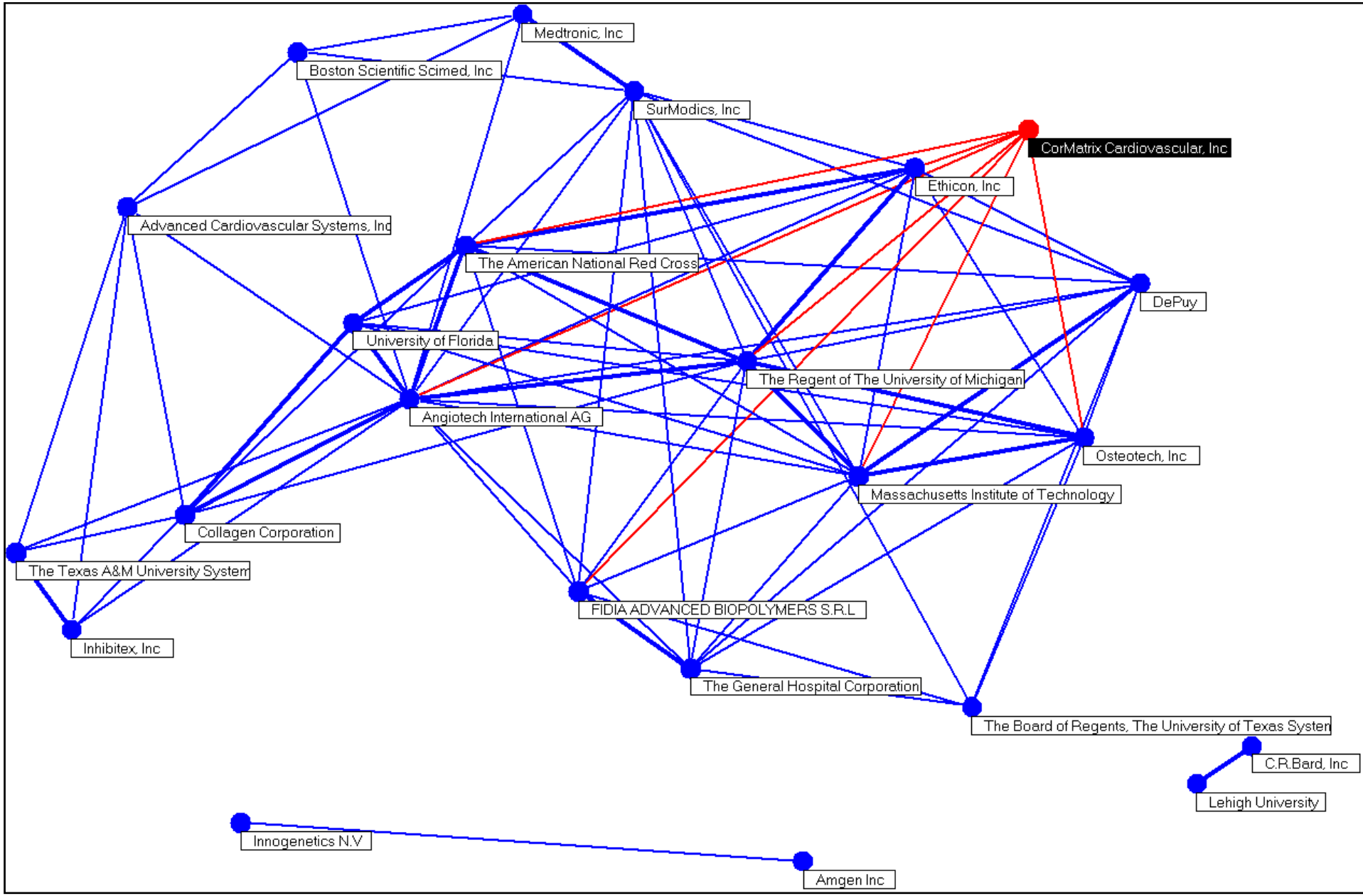
Publication Year



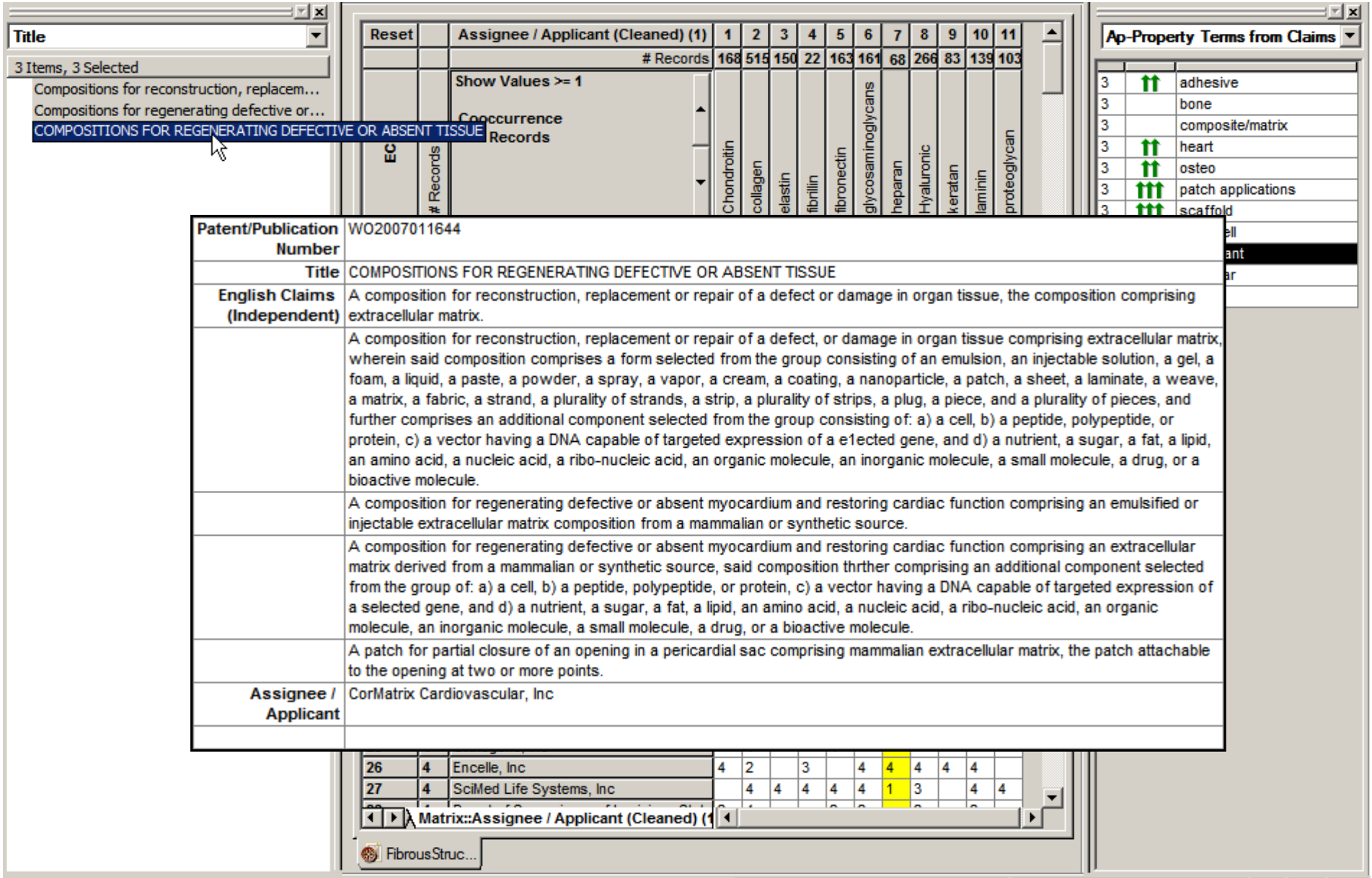
What – Who - When

	Property Terms in Claims	Assignee / Applicant	Publication Year (2005-7)
212	adhesive	Angiotech International AG [11]; FIDIA ADVANCED BIOPOLYMERS S.R.L [11]; Massachusetts Institute of Technology [9]; DePuy [8]; The Regent of The University of Michigan [8]	26% of 212
156	graft	Medtronic, Inc [17]; Angiotech International AG [6]; SurModics, Inc [5]; Boston Scientific Scimed, Inc [5]; Baxter International Inc [4]; Orthogene, Inc [4]; Advanced Cardiovascular Systems, Inc [4]	25% of 156
123	scaffold	CorMatrix Cardiovascular, Inc [7]; Osteotech, Inc [6]; FIDIA ADVANCED BIOPOLYMERS S.R.L [6]; The Regent of The University of Michigan [6]; Massachusetts Institute of Technology [5]; Ethicon, Inc [5]	29% of 123
110	tumor	FIDIA ADVANCED BIOPOLYMERS S.R.L [8]; The Regent of The University of Michigan [7]; Osteotech, Inc [6]; Ethicon, Inc [6]; Massachusetts Institute of Technology [4]; Orthogene, Inc [4]	21% of 110
83	wound treatment	FIDIA ADVANCED BIOPOLYMERS S.R.L [8]; Amgen Inc [5]; Innogenetics N.V [5]; JSF Consultants Ltd [4]; The Regent of The University of Michigan [3]; The Board of Regents, The University of Texas System [3]	13% of 83

Assignees based on Shared Topical Claims



And Eventually… locate & study the “nuggets”



Biomaterials in Complex Patent Landscapes

- Focus on the “MOT” questions to be answered
- Patent landscaping can show “blackspaces” of high activity
- Tech Mining can provide “breakout” details on players and topics of particular concern
- Comments/questions?

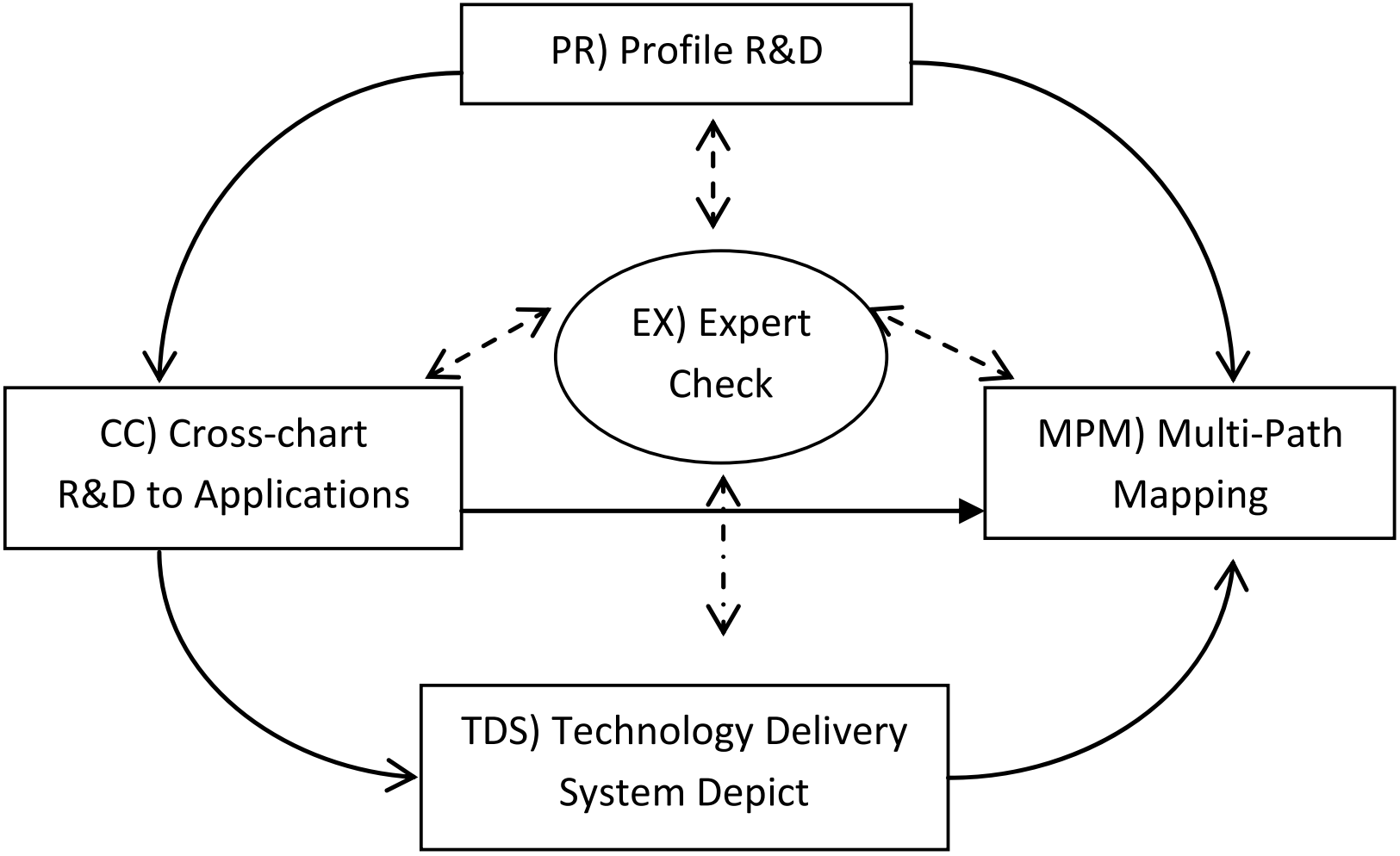
Nano-enhanced Biosensors Case

1. **Pursuing 2 "zoom in" analyses within Nano this year**
   - **Nano interests [to get to local knowledge]**
   - **2 Beijing Institute of Technology PhD students ~research methods + dissertation**
   - **Lu Huang leading this study**
   - **Ying Guo leading on nano-enhanced, thin-film solar cells**
2. **In stages, via papers**

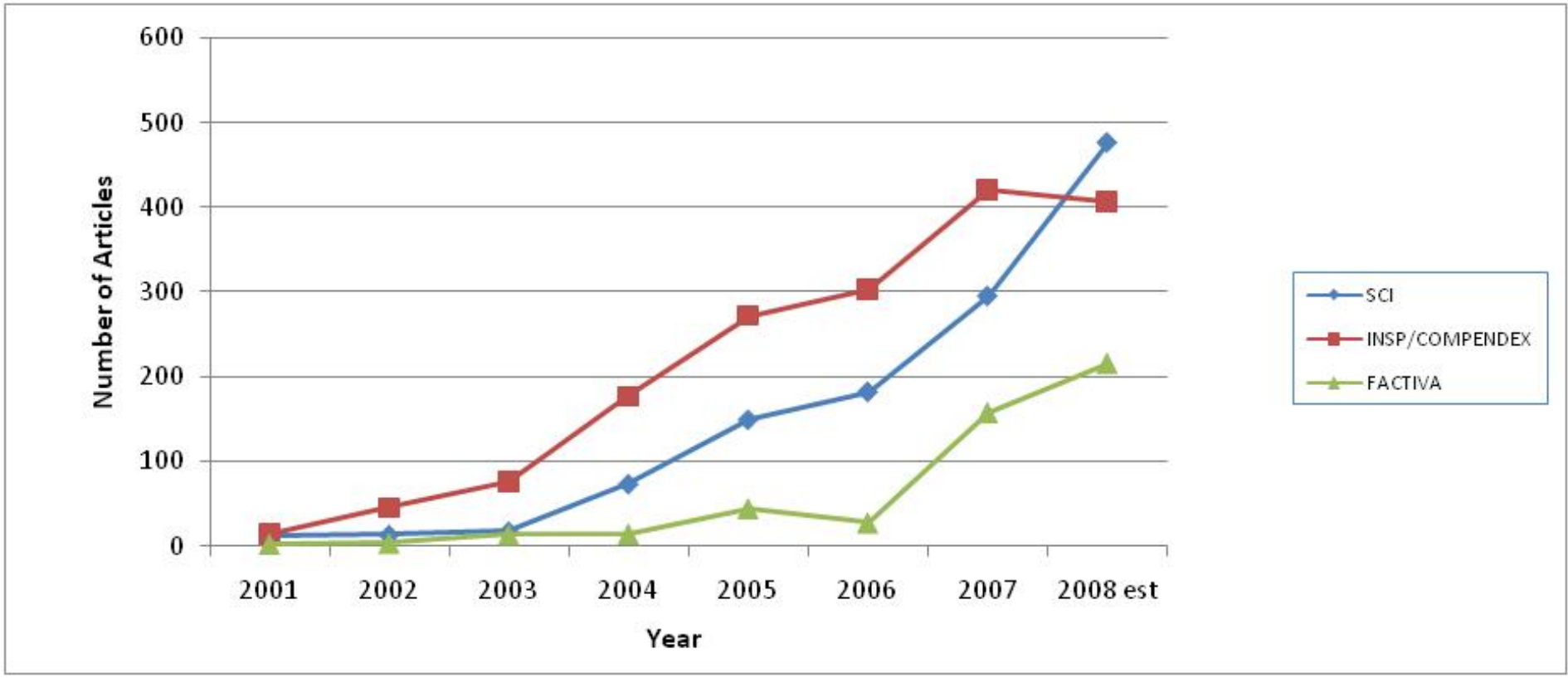
Search & Retrieval

1. **Databases**
   - **Began with the GT nano search sets**
   - **Science Citation Index (SCI – a component of the Web of Science),**
   - **INSPEC & EI Compendex**
   - **Factiva [no GT nano search set here]**
2. **Most of these examples draw from SCI**

Innovation Pathway Modeling

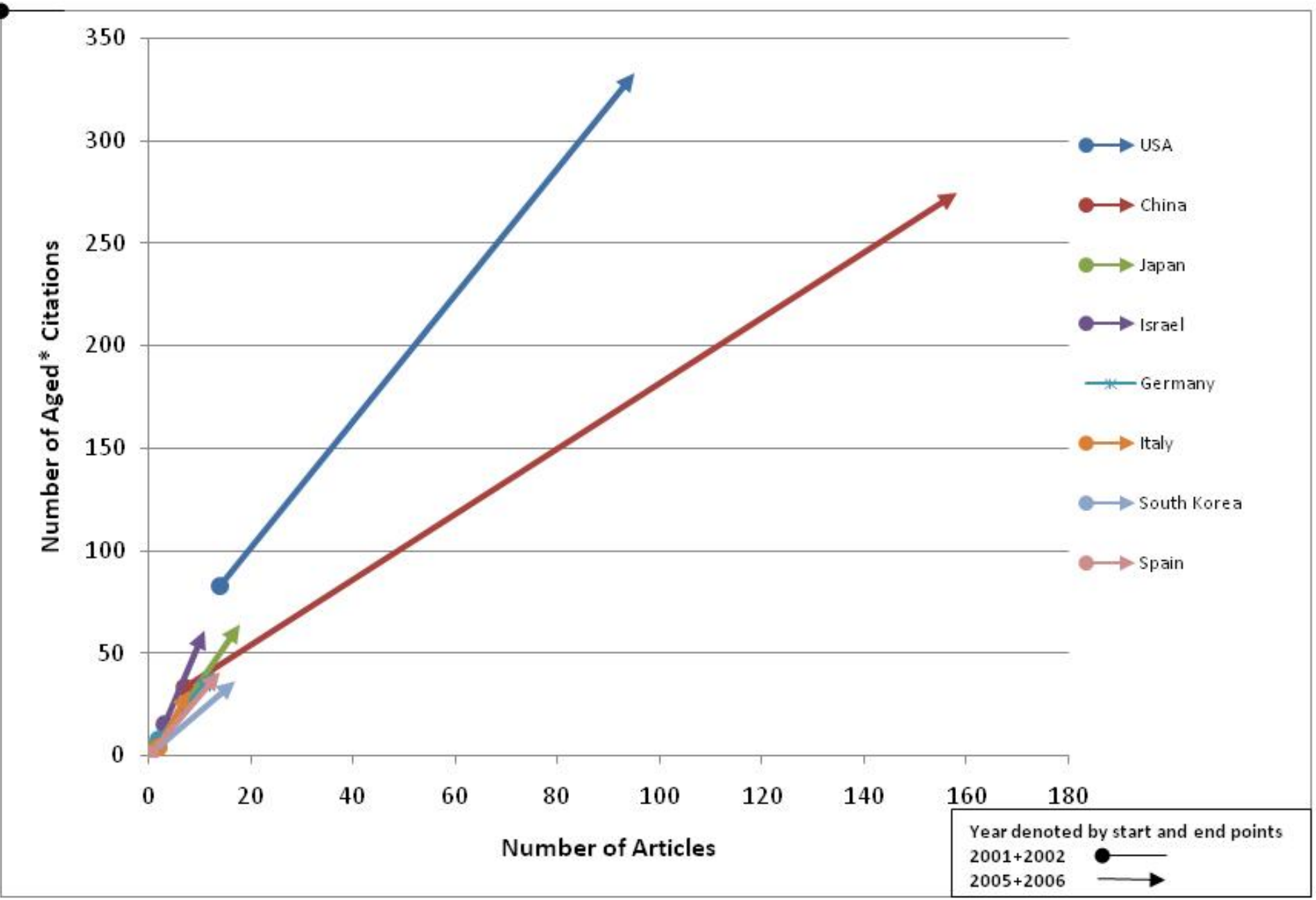


NanoBioSensor Activity Trends

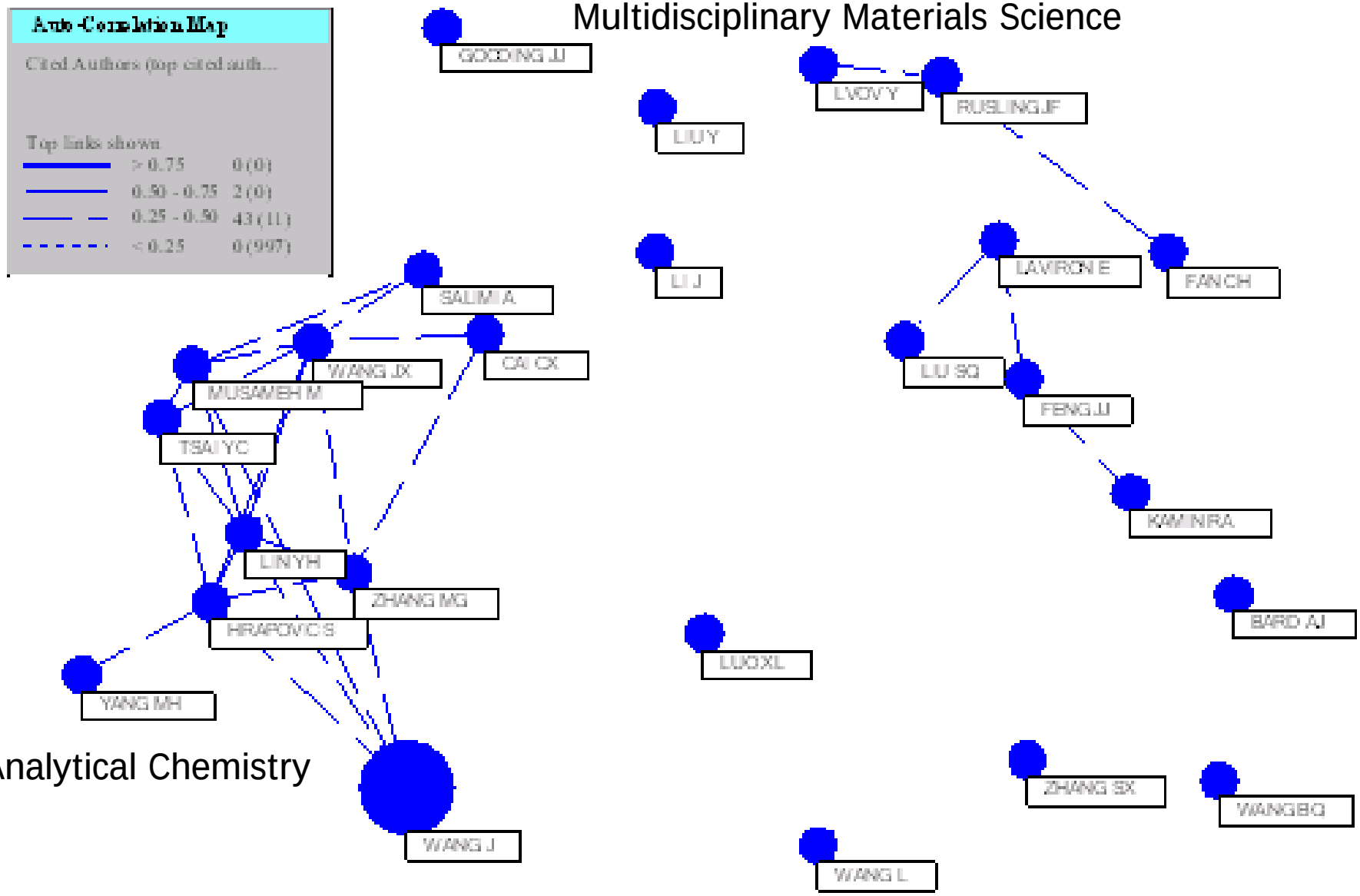


Special interest: Take-offs and Lags
(~3 years to business/popular attention)

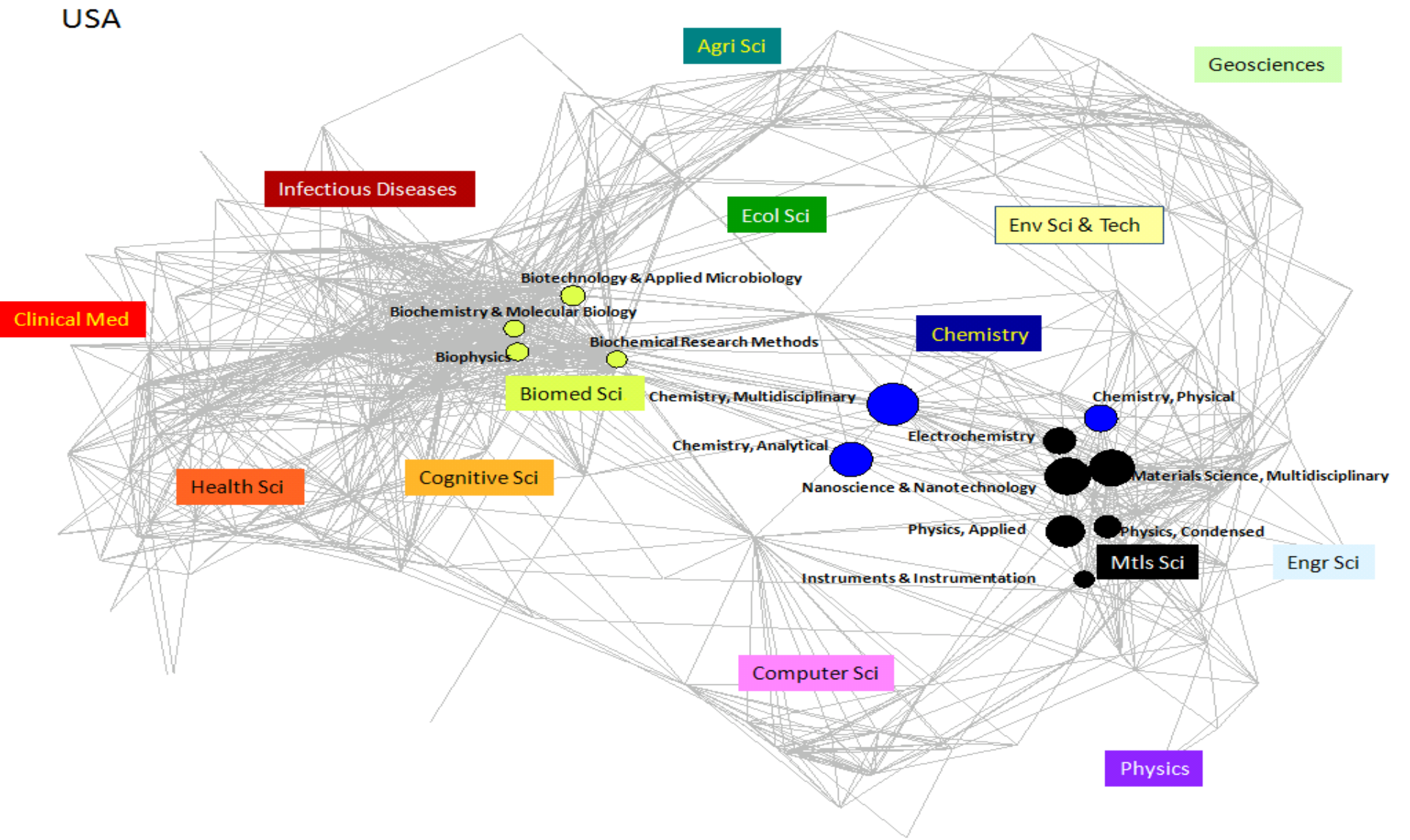
NanoBioSensor -- National Quantity + Quality Indicators



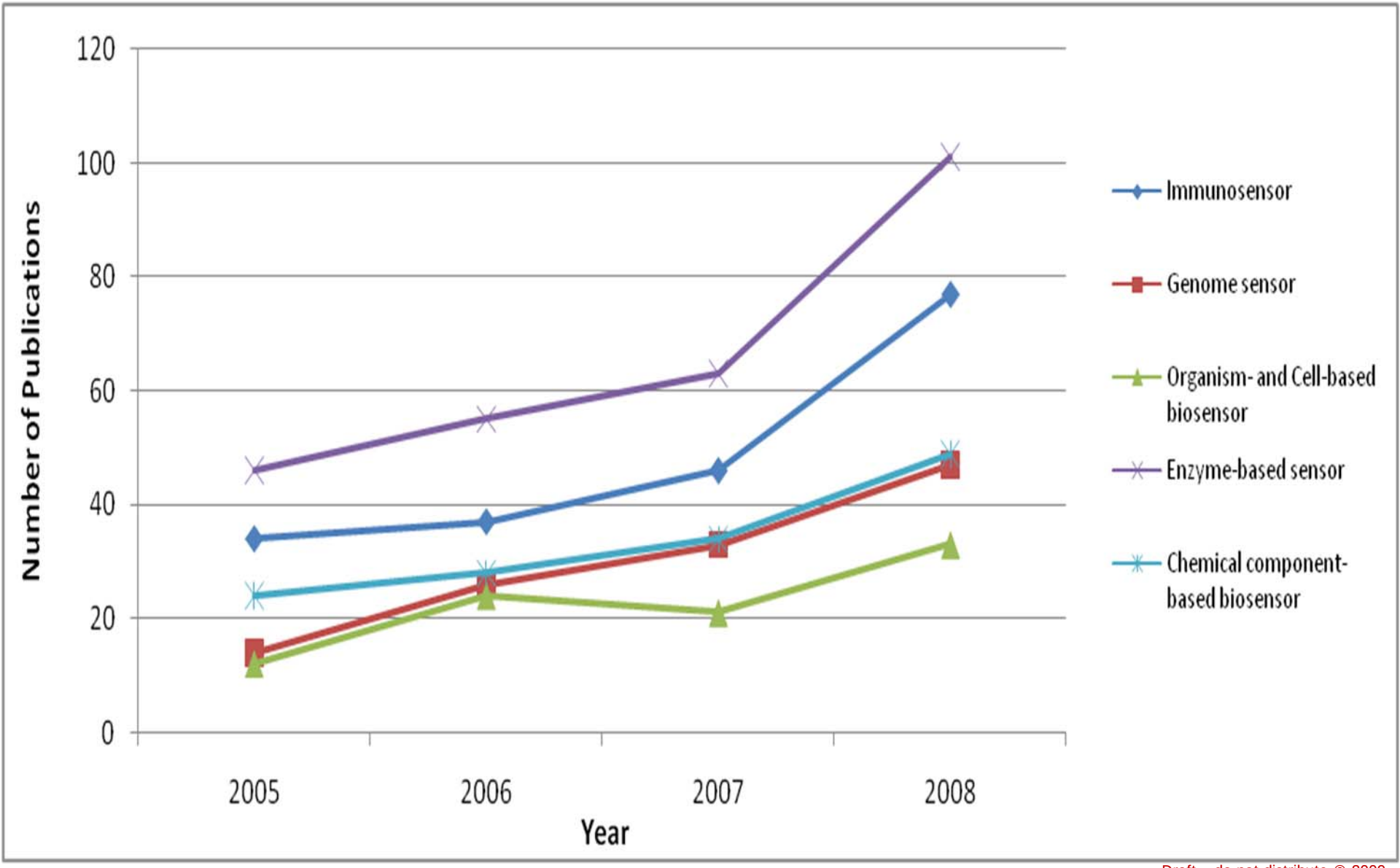
Research Network Mapping: Portion of Co-citation Map



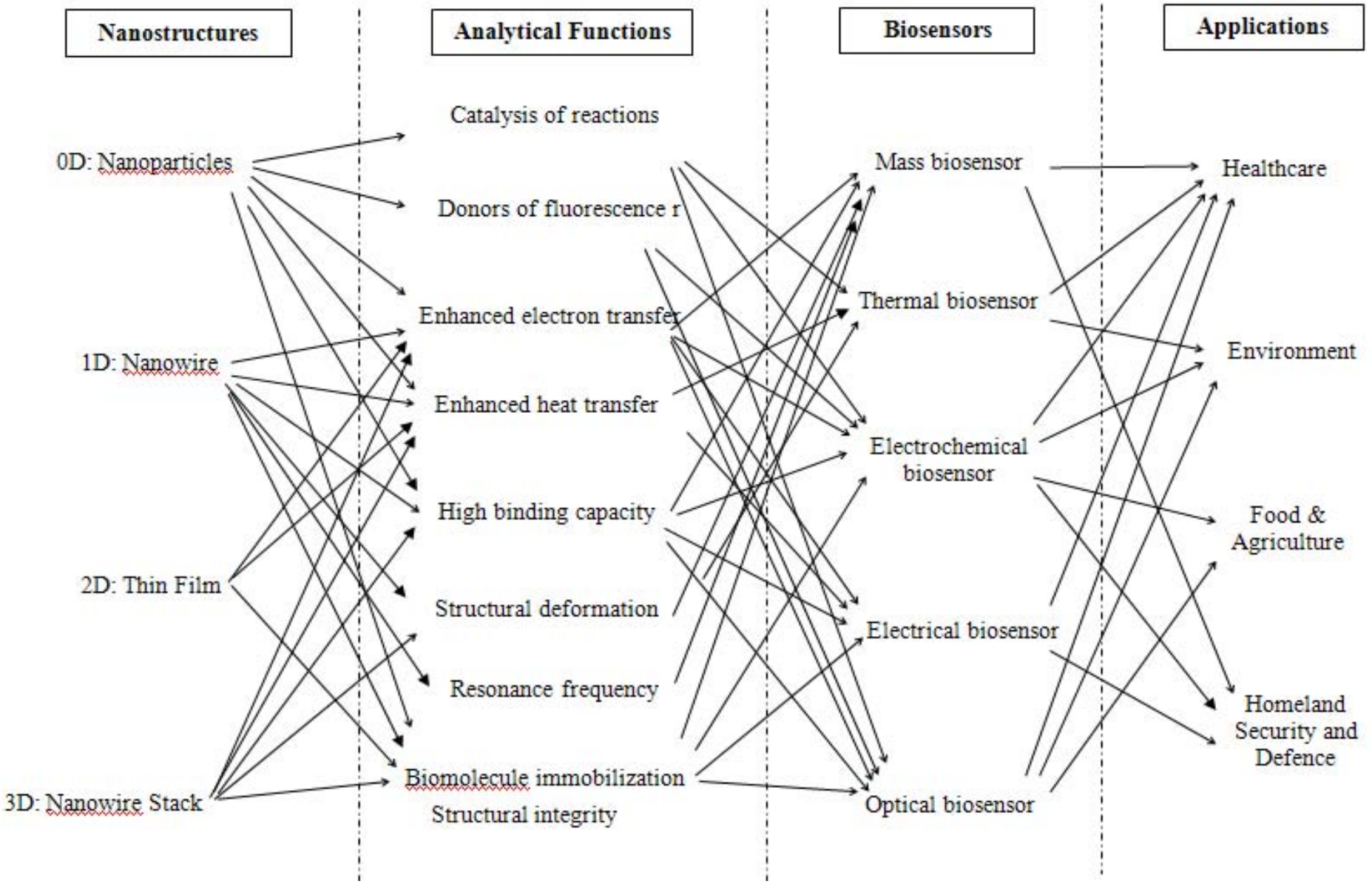
Locating US Nanobiosensor Research in the Map of Science



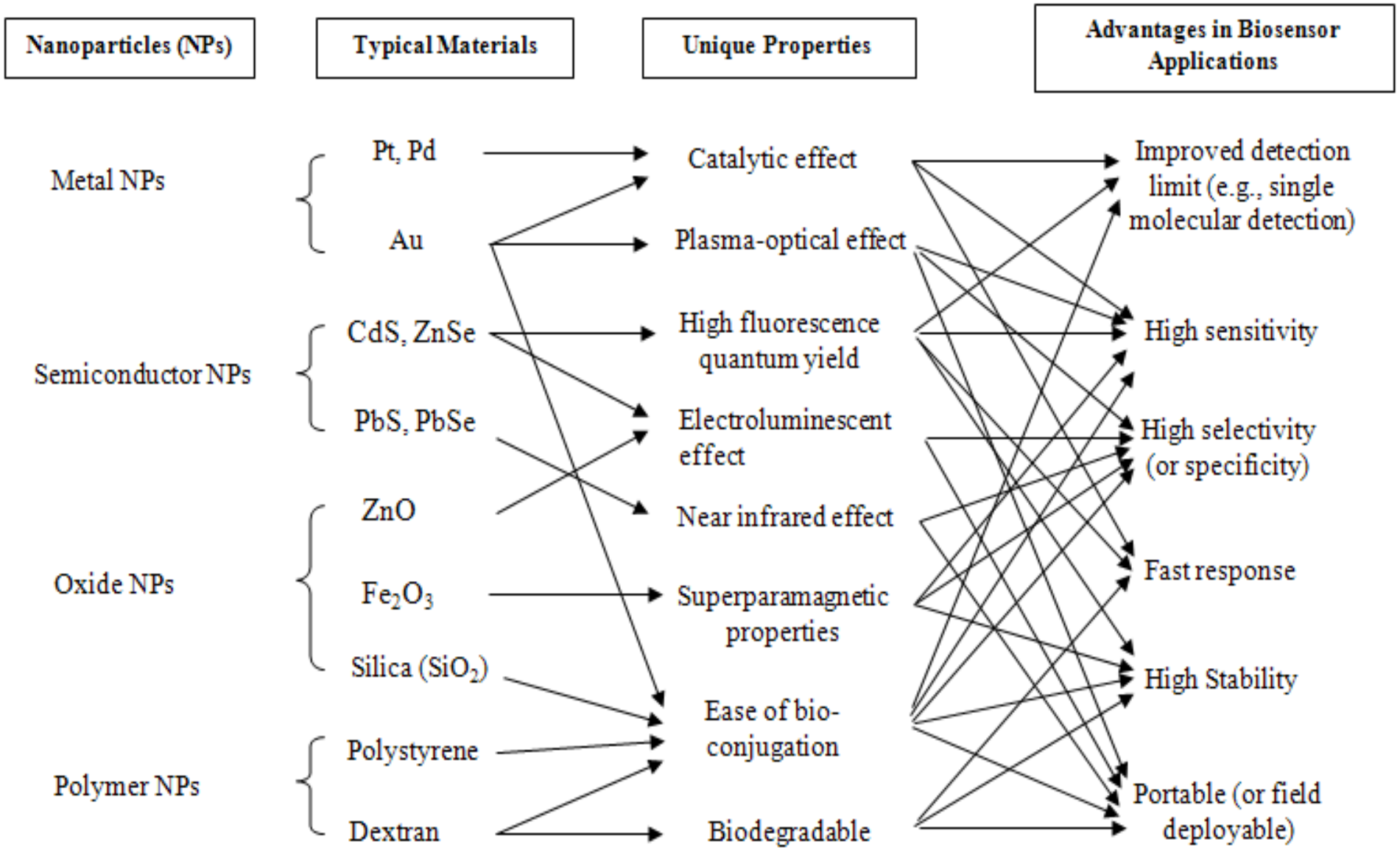
Biosensor Application Emphases



Technology-Function-Application ["T-F-A"] Cross-chart



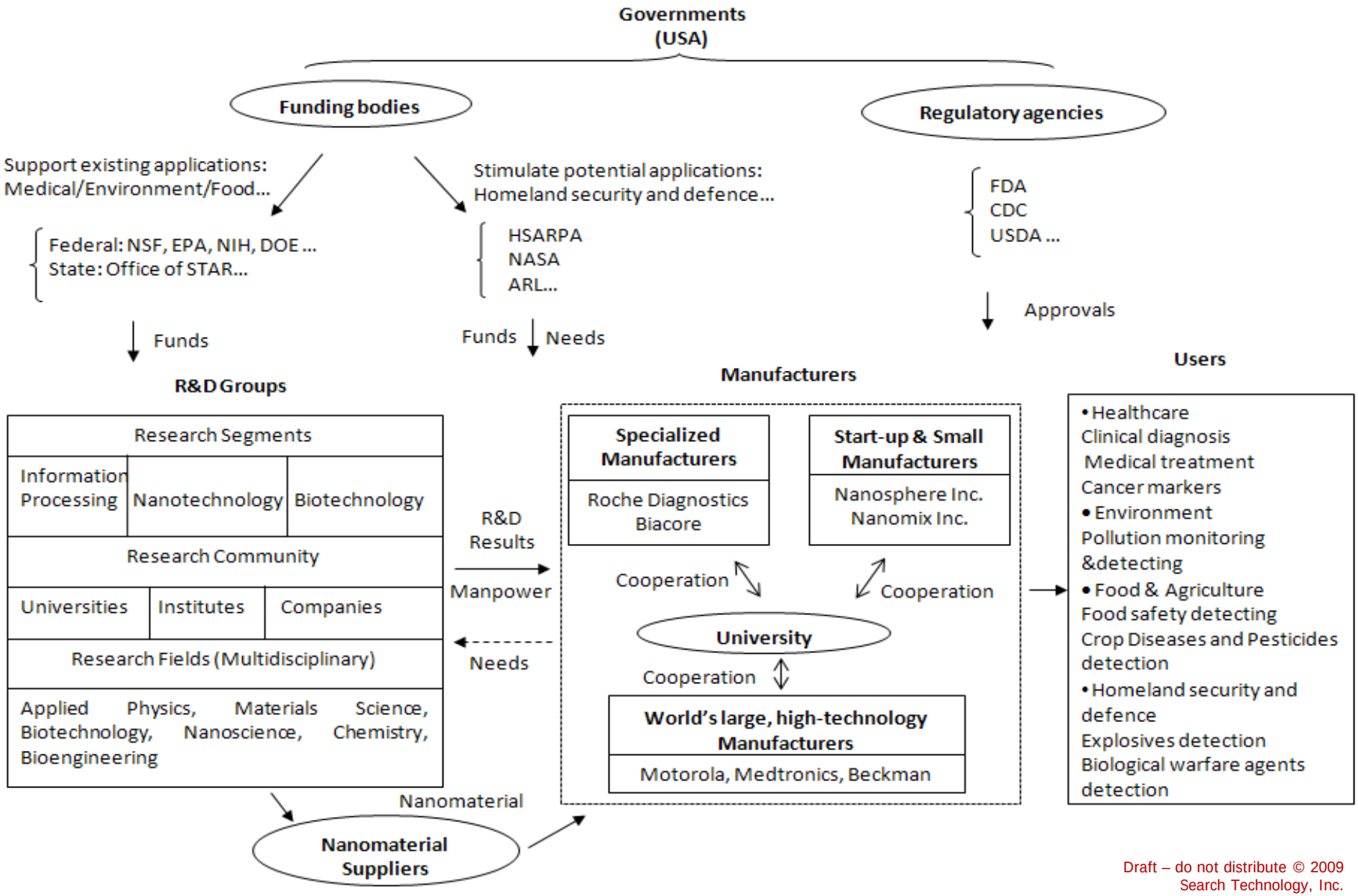
More Specific Nanoparticles T-F-A Cross-chart



Zoom in further: One Type of Nanoparticle -- Gold

Functions	Biosensor	Sensor advantages	Typical examples
Biomolecule immobilization	Enzyme-based sensor	Improved stability and sensitivity	Bienzyme amperometric biosensor using gold nanoparticle-modified electrodes for the determination of inulin in foods (c.f., Manso J , 2008).
Catalysis of reactions	Enzyme-based sensor	Improved sensitivity and selectivity	Glucose biosensor based on Au nanoparticles (c.f., Xian YZ, 2005).
Labeling biomolecules	Immunosensor	Improved sensitivity Indirect detection	Horseradish peroxidase-functionalized gold nanoparticle label for amplified immunoanalysis based on gold nanoparticles/carbon nanotubes hybrids modified biosensor (c.f., Rongjing Cui, 2008).
Enhancement of electron transfer	Genome sensors	Improved sensitivity Direct electrochemistry of proteins	Colloid Au-enhanced DNA immobilization for the electrochemical detection of sequence-specific DNA(c.f., Cai H, 2001).

Technology Delivery System (general)



Actor Analysis

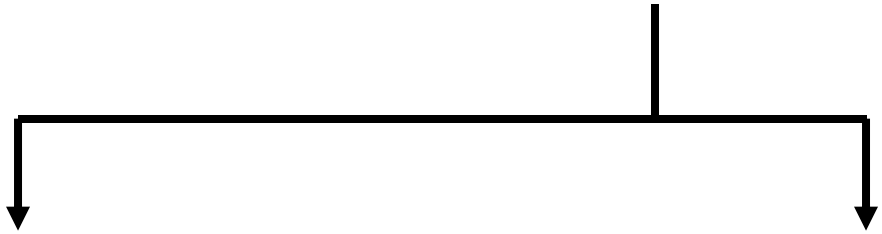
	Supports	Barriers
Governments	• Strong financial support	• High regulatory barriers
R&D groups	• Steep increase in literatures • Multidisciplinary cooperation • Strong cooperation with manufacturers	• Far away from commercialization • Lack of good integration of biosensor into easy-to-use systems
Manufacturers	• Promising market prospects • Strong cooperation with universities • Ever-growing number of companies offering nanomaterials	• Separate market segments • High standards of door-step to markets • High cost with needed performance • Scaling up manufacture of nanomaterials
Users	• Plenty of needs • Plenty of potential users	• Needs beyond present ability • Safety • Use friendly

Nano-enhanced Biosensors Case

1. Trends & Quality Indicators
2. Mapping
   a. Research networks
   b. Science overlay maps
3. Innovation Pathway Mapping
   1. T-F-A Cross-charting
   2. Enhanced TDS
4. Comments/questions?

The Tech Mining Approach:

- Define the Management of Technology (MOT) Issues
- Break out particular MOT Questions
- Identify candidate empirical Indicators



- Identify appropriate Data Source(s)
- Identify appropriate Analytical Tool(s)

- Design Effective composite Representations

Competitive Technical Intelligence (CTI):
Who, What, When, Where?

- Profile R&D Domain(s) of concern:
  - Who?
  - What?
  - When?
  - Where?
- Map Relationships: Network Analyses & Science Overlay Maps
- Analyze Trends: What's Hot & What's Coming
- Develop Innovation Indicators
- Then, locate "nuggets" to be read

Resources

- *Open Innovation: The New Imperative for Creating and Profiting from Technology*, by Henry Chesbrough, Harvard Business School, Cambridge, MA (paperback edition), 2006.
- *Tech Mining* by Alan Porter and Scott Cunningham, Wiley, 2005.
- www.theVantagePoint.com – offers multiple papers and some case analyses, including:
  - Lu Huang et al., Identifying Emerging Nanoparticle Roles in Biosensors, *IAMOT Proceedings, 2009* [case & methods illustration from this + an article building upon this]
  - Various "Tech Mining" papers

Open Discussion

1. Introduction
   a. Open Innovation
   b. Tech Mining

2. Tech Mining Tales
   a. “Research Profiling” Country study: Brazil
   b. Biomaterials Patenting: Technology Opportunities Analyses
   c. NanoBioSensors: Innovation Path Mapping

3. How could these tools work for you?