To appear, *R&D Management*

Profiling Leading Scientists in Nano-Biomedical Science: Interdisciplinarity and Potential Leading Indicators of Research Directions

Ruimin Pei

Chinese Academy of Sciences
School of Public Policy
Georgia Institute of Technology
Beijing, China / Atlanta, GA 30332-0345 USA
Tel: (404) 966-2227
Email: peiruimin@gmail.com

Alan L. Porter*
Technology Policy and Assessment Center
Georgia Institute of Technology
Search Technology, Inc., Norcross
Atlanta, USA
Tel: (404) 384-6295
Email: alan.porter@isye.gatech.edu

Ruimin Pei, is a Ph.D. student of Management School, Graduate University of Chinese Academy of Sciences, and a visiting student at School of Public Policy, Georgia Institute of Technology, majoring in management science and engineering.

*Alan Porter is with the Technology Policy and Assessment Center, Georgia Tech, and directs R&D for Search Technology. Current research emphasizes research profiling, innovation indicators, science mapping, and forecasting innovation pathways.

Nano-biomedical science is a promising area in the application of nanotechnology. This paper profiles a group of 21 leading scientists in nanobiomedicine based on high publication rate and high citations. Comparisons with other researchers indicate that the leaders publish more in high impact journals and collaborate more extensively (team science). They reside most heavily in the USA and Western Europe. We compare their research publications using multiple indicators – integration, specialization and a multidisciplinary index. Relations among interdisciplinarity indicators generated support a 3-factor model based on principles of diversity. We locate this research among the disciplines using science overlay mapping. Key term analyses, based on keywords and on natural language processing (NLP) help profile the research emphases of these leading researchers. Such results could serve as leading indicators to help identify directions of future nano-bio development.

1. Introduction

Leading scientists can be taken to be researchers who are particularly active with high influence in a field. We posit a possible association between creativity, prolific publication, and influence, though not necessarily a simple one. Numerous studies on scientific creativity (Buel et al., 1966; Calero et al., 2006; Heinze & Bauer, 2007; Heinze et al., 2007) illustrate that creative scientists in a field move the field in the direction it is already going, or in a new direction (Heinze et al., 2007; Sternberg, 2003). To identify the leading scientists in a field, and profile their research emphases, could contribute to anticipating the development of that field.

Nanotechnology research started in the 1980s, ascended in the 1990s, and escalated dramatically in the 2000s. It receives wide attention not only because of its strategic position in the development of science and technology (e.g., as reflected in major funding initiatives such as the US National Nanotechnology Initiative – NNI), but also because of its high commercial potential. As one of the most important application-oriented fields in nanotechnology, what we term "nano-biomedical science" ("**NBMS**") is an especially promising area for both research and industrial communities.

ESTO (2004) estimates that the market for nanomedicine in 2004 was EUR 5.4 billion, based on 37 nanomedicine-related products already in the market (Meyer et al., 2005). Hullmann (2006) estimates that nanodevices and nanobiotechnology are responsible for the largest shares of around $420 and $415 million, respectively, in the nano-market for the period of 1999-2003, and that the trend of the nano-enabled drug delivery market will increase dramatically. Forecasts indicate that nanobiotechnology will grow sharply, with an overall market impact of nearly $300 billion within the next 12 years (Zhou, 2003). A three-phase model of nanotechnology development suggests the onset of a third phase from 2010 onward in which nanotechnology will become commonplace in manufactured goods, led by healthcare and life science applications entering the pharmaceutical and medical device markets (Hullmann, 2006). NBMS is definitely becoming one of the most promising areas in nanotechnology (ESF, 2005).

Investment from public and private organizations in research in NBMS is also increasing rapidly. National policy making has designated NBMS as an important area for investment in research and innovation through initiatives such as the National Nanotechnology Initiative (NNI) in the US, and similar initiatives in many countries (Sastry et al., 2010). Investment in nanobiosystems research and education in the NNI was around 15% of the US nanotechnology budget in 2002, and has grown to more than 20% in 2010 (Roco, 2003; Sastry et al., 2010). Besides the government investment in nanotechnology, private sector risk capital is also attracted to the nanotechnology market, especially nanobio (with a share of 52% in 2004). Lux Research

also reported in 2007 that venture capital for nanotechnology reached $699 million in 2006 and over half of it went into bionano (Meyer et al., 2005).

NBMS R&D outputs demonstrate exponential growth based on counts of publications and patent applications. For publications, nanobio constituted 25% of the overall nano field in 1991, and grew to 28% in 2003. Nanobiotechnology patents rose over 54-fold from 1990 to 2004. Furthermore, new product introductions, based on press releases, nanobio grew slowly during the 1980s, escalated in the 1990s, and dramatically increased around year 2000 (Grodal & Thoma, 2008).

As aforementioned, NBMS is an important sub-field of nano-science and technology ("**nano**"), which is an emerging domain of scientific endeavor that embraces various research fields, including applied physics, materials science, physical chemistry, condensed matter physics, biochemistry and molecular biology, and polymer science and engineering (Hullmann & Meyer, 2003). Nano is thus notably multidisciplinary, so that one would anticipate important sub-field differences. Yet many bibliometric and other empirical analyses focus on the overall nano field (Braun et al., 1997; Calero et al., 2006; Guan & Wang, 2010; Heinze & Bauer, 2007; Heinze et al., 2007; Hullmann, 2007; Hullmann & Meyer, 2003; Kay & Shapira, 2009; Liu et al., 2009; Meyer, 2006; Meyer & Persson, 1998; Meyer, 2001; Porter & Youtie, 2009).

Drawn by the prominence of NBMS within nano, and its continuing rapid development and strong market prospects, we address this subfield. This paper investigates NBMS using a combination of bibliometric approaches. Elsewhere, we have compared national differences in NBMS research thrusts (Pei et al., to appear)– contrasting China with four other leading countries (USA, Germany, Japan, and the UK). Here we analyze the leading researchers to gain perspective on promising directions for the field.

2. Background

2.1. Defining nano-biomedical science ("NBMS")

Our work is informed by how the existing literature defines the concept of NBMS. NBMS is widely addressed as one of the important application-oriented fields in nanotechnology(ESF, 2005). The Euronano report (Meyer et al., 2005) identifies nanomedicine as one of the four major broad nano fields in exploring the economic potential of nanotechnology. Other major reports and articles related to NBMS include the ESF report (ESF, 2005), Roadmaps report (Meyer & Persson, 1998), European Technology Platform report (Tomellini et al., 2005), Ernest and Shetty (Ganguly et al., 2010), Rafols and Meyer (Rafols & Meyer, 2007) and Wagner *et al.* (Michalakelis et al., 2010).

Although there are notable studies relating to NBMS, there is no agreement as to the definition, terminology, or boundaries for NBMS. As you likely noted in the preceding discussion, terminology varies, which leads to uncertainty on boundaries. Some works use the term nanobiotechnology, while others use the term nanomedicine. For example, the Euronano report (Meyer et al., 2005) defines nanomedicine as the application of nanotechnology to health, specifically, at the nanometric scale, and its impact on the prevention, early and reliable diagnosis, and treatment of diseases. Examples of nanomedicine given in this report are: drug delivery; drugs and therapy; in vivo diagnostics; in vitro diagnostics; biomaterials; tissue engineering / regenerative medicine; active implants. The forward looking report of EMRC (ESF, 2005) defines the field as "*the science and technology of diagnosing, treating and*

preventing disease and traumatic injury, of relieving pain, and of preserving and improving human health, using molecular tools and molecular knowledge of the human body." Ernest and Shetty (Ganguly et al., 2010) consider nano biomedical science as the direct application of the convergence of nanotech and biotech and take the same definition of nanobiomedical science as the EMRC report. Bibliometric analysis might define the boundaries of nanobio, nanobiomedical or nanomedicine from the perspective of search terms, which are then used to address and retrieve publications and patents in the domain. For example, (Grodal & Thoma, 2008) define the boundary of nanobiotechnology using a combination of keywords from the nano and bio fields.

One of the hallmarks of NBMS is its multidisciplinary and interdisciplinary character. Rafols and Meyer (Christodoulos et al., 2010; Rafols & Meyer, 2007) delineate the cross-disciplinary and interdisciplinary properties of NBMS. The authors' 2007 work contributes to the methodologies of measuring interdisciplinarity at the project level. The authors focus on a particular nanotechnology application – molecular motors in bionanotechnology – because bionanotechnology is considered as a cross- or multi- disciplinary area. The 2010 paper uses bionanoscience as a case study to develop a new framework and indicators of interdisciplinarity from the perspectives of diversity and network coherence. Grodal and Thoma (2008) study the knowledge diffusion of nanobiotechnology using keywords for nanobiotechnology coverage. Ernest and Shetty (2005) find that nanobiotechnology is the hybrid of biotechnology and nanotechnology which combines the classical microtechnology with a molecular biological approach. NBMS is an overlap of biotechnology, nanotechnology, and information technology bringing to fruition many important applications in life sciences and medicine.

In our study, we consider NBMS from two perspectives -- the theoretical perspective, taking the relatively comprehensive definition from the EMRC report, and the operational perspective for retrieval of NBMS publication data. From the theoretical perspective, NBMS in this paper is taken as a combination of nanotechnology and biotechnology for medical use, such as to diagnose, treat and prevent disease and traumatic injury, to relieve pain, and to preserve and improve human health.

From the operational perspective, our key NBMS dataset is the intersection between a nano search conducted in the Science Citation Index (of the Web of Science) (Porter et al., 2008b) and a biomedical science categorization. The latter is one of 18 macro-disciplines developed from the 221 subject categories of the Institute for Scientific Information (ISI) by factor analysis according to their journal cross-citation matrix (Rafols et al., 2010b).

2.2.Leading Scientists and Characteristics of Their Research

Attention to the leading scientists and their interactions ("Invisible colleges") and ways to depict these date back to Derek Price and other early bibliometricians (de Bellis, 2009). Extension of the concept of thought leadership at the national and institutional level (Klavans & Boyack, 2008) suggests that the individual publication and citation counts might be used to identify the leading scientists in a field. To do so is to find out the most active, *and/or* the most creative, *and/or* those with the highest influence in the field using appropriate indicators. Publication and citation counts are obvious indicators of prolific scientists in a field (Heinze & Bauer, 2007; Simonton, 1999; Simonton, 2004). Simonton claimed in his studies that the prolific scientists have a higher probability of their work being selected as "creative" by peer evaluation ((Heinze & Bauer, 2007; Simonton, 2004). These selected scientists are believed to have more influence and impact on the field. Peer evaluation is another widely used method to identify the leading scientists in a field, and a study shows how this method works in selecting creative

scientists in a field (Heinze et al., 2007). Bibliometric methods offer advantages over peer review in their replicability (and, possibly, objectivity) and efficiency. With the wide availability of research databases (such as the Web of Science upon which we draw here, Scopus, Chemical Abstracts, EI Compendex, and so forth), and text mining software (e.g., VantagePoint which we use in this work – see www.theVantagePoint.com), data compilation and consolidation (cleaning) are quite accessible. Use of resulting scientometric profiles is especially widespread in Europe, providing research assessment information for researchers, knowledge managers, and policy makers.

In principle, the leading scientists are those who exert influence in both their own discipline and other disciplines because of their high productivity and creative accomplishments. The latter imply that such scientists link existing knowledge to new concepts, ideas, techniques, questions, phenomena and facts. That suggests that they may be particularly facile in adapting knowledge from related disciplines. Many observers believe that such interdisciplinarity fuels creative problem solving (TheNationalAcademies, 2005). One might anticipate that such work would then be highly cited.

The studies of Levitt and Thelwall (2008; 2009) challenge a simple hypothesis that highly interdisciplinary papers are cited more. They operationalize "multidisciplinary" as publication in journals associated with more than one Subject Category by Web of Science or Scopus, and "mono" as those in single SCs. [Note that this is a different definition than ours.] In social sciences, citation levels for "multi" and "mono" were similar in WOS and Scopus; whereas in life sciences, health sciences, and physical sciences, Scopus citations were considerably higher for "mono" journal articles (Levitt & Thelwall, 2008). In another paper, Levitt and Thelwall (2009) found that 82 highly cited information and library science articles were disproportionately representative of "multidisciplinary" as opposed to "mono" journals. They note confounding factors as well – e.g., if other fields have higher citation rates (say, than library & information sciences), then publication in "multi" journals may give articles an advantage in accruing citations (Levitt & Thelwall, 2009).

Yegros-Yegors et al. (2010) reviewed the issue too. They noted studies showing interdisciplinarity not correlating with citations (Adams et al., 2007; Rinia et al., 2001), and another with mixed effects depending on disciplines (Larivière & Gingras, 2010). Their own findings showed that higher variety in a paper's cited references tending to correlate with higher citation (but not so for better balance or higher disparity – the other two diversity components incorporated in our Integration score – see Section 4.2).

Porter and Rafols (2009) had compiled data on a number of papers published in six SCs in selected years. Intrigued by the challenge of an observant reviewer, we asked Stephen Carley to analyze those data to examine the relationship between interdisciplinarity, as measured by Integration score, and high citation. Correlations between Integration and Times Cited ("TC") average about 0.07 (small). Most telling, the variance in TC explained by Integration score is minimal.[1] Scatterplots of TC on the vertical axis against Integration score horizontally show a smattering of high TC values "boiling" up from a massive, low bell-shaped distribution.

Creative scientists, relative to their peers, belong to professional networks with access to richer and more diverse expertise, and they tend to address a broader disciplinary spectrum in their work (Heinze & Bauer, 2007). According to Burt's theory of structural holes (Burt, 1992; Burt, 2004; Scott, 2000), individuals who occupy a unique position in a network have access to multiple views, information, and perspectives; therefore, they are more likely to develop more novel and better ideas than their peers. In the scientific world, the leading scientists are usually

in a broader collaborative network, from which they can get more information and have more potential to succeed (Abramo et al., 2009; Lee & Bozeman, 2005). Heinze and Bauer think that scientists who interconnect homogeneous groups, such as disciplinary communities or research fields, have a higher probability of exposure to alternative ways of thinking and behaving (Heinze & Bauer, 2007). From the bibliometric standpoint, co-authorship is one of the most tangible and well documented forms of scientific collaboration (Glänzel & Schubert, 2004) and the number of authors is used as a straightforward variable to reflect co-authorship (Sutter & Kocher, 2004). The resulting data provide the basis for social network analyses.

Keyword analysis is a bibliometric approach which is used widely for profiling a field (Chiu & Ho, 2007; McGrath, 1996). It can also be used for predicting the trend of a field (Li et al., 2009; Yi & Xi, 2008). Comparative keyword analysis is an approach to compare the differences between large bodies of text, which is an alternative to the "code and retrieval" approach used in much thematic analysis of qualitative materials (Seale et al., 2006). In order to address the hot topics of the leading scientists in NBMS, we use keyword analysis and comparative keyword analysis. In basic terms, we can compare research topics that are more prominently addressed by leading scientists than by the research community at large; also, we can contrast topics that the leading scientists emphasize more extensively in recent years.

This paper focuses on and profiles the following:

1) The combination of publication activity and citation frequency to identify a group of leading scientists in NBMS;
2) The country, affiliation (organization), and disciplinary distribution of the leading scientists in the field of NBMS;
3) The multidisciplinary and interdisciplinary research attributes of the leading NBMS scientists;
4) The collaboration patterns of the leading NBMS scientists; and
5) The hot topics on which the leading NBMS scientists focus.

3. Methodology and data collection

Bibliometric approaches are used in this paper to identify and profile the leading scientists in NBMS. Specifically, VantagePoint software[2] is used to search, extract, clean, and analyze the data. The number of publications and citations (Heinze & Bauer, 2007), the size of co-authorship networks (Heinze & Bauer, 2007), multidisciplinary indices (Heinze & Bauer, 2007; Heinze et al., 2007), and interdisciplinary indices (Porter et al., 2007; Porter & Rafols, 2009; Porter et al., 2008a) and keyword analysis are used to identify and profile the leading scientists.

Visualization techniques in the scientific world are used to show collaboration networks and disciplinary interaction patterns (Borner et al., 2003; Boyack, 2004; Boyack et al., 2009; Boyack et al., 2005; Chen, 2006; Chen et al., 2002; Klavans & Boyack, 2009; Porter et al., 2007; Porter & Rafols, 2009; Porter et al., 2008a; Porter & Youtie, 2009; Rafols & Meyer, 2010). We use science overlap maps (Rafols & Meyer, 2010) to illustrate cross-disciplinary involvement in NBMS.

The data used in this paper are taken from the Nano datasets developed at Georgia Tech (Porter et al., 2008b). We then separate those records whose journal of publication is assigned by Web of Science ("**WOS**") to certain Subject Categories ("**SCs**"). Our colleagues have taken a full year's WOS journal cross-citation records from the WOS Journal Citation Reports. They

use those to create base science maps (c.f., Figure 3) (Leydesdorff & Rafols, 2009; Porter et al., 2009). The labels shown are generated by factor analysis of the SC cross-citations (derived from journal cross-citations) (Porter & Rafols, 2009). For the 221 SCs in the Science Citation Index and Social Science Citation Index together, we use a set of 18 factors, called "macro-disciplines." One of those is biomedical science, including 24 SCs[3]. The resulting NBMS dataset includes 62,197 records published in the years, 2000-2009, which represents 10.8% of the whole Georgia Tech WOS nano dataset.

According to the definition of leading scientists mentioned above, we identify the leading scientists who are active and creative in the field by exploring the publication counts and citation counts. We later search online and consult knowledgeable persons in NBMS to confirm that these candidates really have significant influence in the field.

4. Results

4.1.Identification and Descriptors of Leading NBMS Scientists

As noted, this paper uses bibliometric approaches to identity a set of leading scientists who may exert significant influence in a field of interest -- NBMS. We begin by distinguishing high publishing researchers (Heinze & Bauer, 2007; Heinze et al., 2007; Simonton, 2004). Then, we use their total citation score to modify our list of leading scientists.

Specifically, we first seek to rank the authors according to their publication counts. We clean up authors who are different persons with the same name (homonyms) as presented in the WOS database (i.e., based on last name and initials). We also seek to consolidate name variations – i.e., the same person with different name variations. This yields 35 candidates with more than 50 publications for the years, 1990-2009, in the NBMS dataset drawn from the Science Citation Index of WOS. However, using the total number of publications as a single predictor of high-impact work is not adequate (Heinze & Bauer, 2007; Simonton, 2004). We found that the leading scientists not only have large number of publications, but also tend to be frequently cited in other papers, which is consistent with Simonton's quality-quantity relation. Of 35 candidates, there are 26 scientists who have at least one publication cited more than 100 times. The largest total citation score among the 35 candidates is 10,746, and the smallest is 259. Finally, we set thresholds to distinguish 20 leading scientists who have both especially prolific publication activity and high citations -- more than 50 publications and more than 1000 total citations; and we add another one who received 21st position with 970 total citations but higher average citations (14.7). Clearly, others could identify a somewhat different set as the "leading" NBMS scientists.

Table 1 offers some descriptive statistics for these leading scientists, ordered on Total Citations (TC) received. Specifically, of the 21 leading scientists, most are from western countries, with two from Japan. Specifically, 9 out of 21 are from the USA, and the others are all from Europe, led by four from Germany.

Table 1. Leading NBMS Scientists' Affiliations and Activity Counts (1990-2009)

Author	Country	Affiliation	TC	# Records	AC
Lieber, C M	USA	Harvard University	10746	60	179.1
Dekker, C	Netherlands	Delft University of Technology	6346	51	124.4
Gaub, H E	USA	MIT	5689	61	93.3
Bustamante, C	USA	University of California, Berkeley	4108	82	50.1
Hansma, P K	USA	Stanford University	3283	51	64.4
Engel, A	Switzerland	University of Basel	3071	97	31.7
Fernandez, J M	USA	Columbia University	3044	70	43.5
Yanagida, T	Japan	Osaka University	2798	101	27.7
Yoshida, M	Switzerland	University of Basel	2596	71	36.6
Spudich, J A	USA	University of California, Santa Barbara	2430	51	47.6
Langer, R	Germany	University of Munich	2307	61	37.8
Aebi, U	Japan	Tokyo Institute of Technology	2062	71	29.0
Muller, D J	Germany	Technical University of Dresden	1908	79	24.2
Couvreur, P	France	University Paris Sud 11	1649	160	10.3
Schulten, K	USA	University of Illinois	1643	65	25.3
Muller, R H	Germany	Free University Berlin	1337	157	8.5
Kreuter, J	Germany	University of Frankfurt	1303	120	10.9
Weissleder, R	USA	Harvard University	1209	70	17.3
Ha, T	USA	University of Illinois	1103	62	17.8
Davies, M C	England	University of Nottingham	1068	56	19.1
Hinterdorfer, P	Austria	University of Linz	970	66	14.7

Note: TC (Total Citation of each author) = $\sum[(\text{Times cited})_i \times (\text{Counts of publication})_i]$; AC (Average Citation of each author) = TC/ Total publication of each author.

The journal impact factor reflects the average number of citations to articles published in science and social science journals. Impact factor is frequently used as a proxy for the relative importance of a journal within a field. Generally, the higher the impact factor of a journal, the higher its papers' presumed "quality" is, and the larger their impact on research. Table 2 compares the 21 leading NBMS scientists to others in the field in terms of journal impact factor [The high impact factor journals concentrate in Multidisciplinary Sciences SC journals, such as Nature and Science.] The results show that the average percentage of publications in high impact factor journals by the leading scientists (9.6%) is more than twice as large as that of the other NBMS scientists (4.4%); so also is the weighted impact factor by publications (2.79 vs. 1.28). This result supports the argument by Simonton that highly prolific scientists are more successful in producing high-impact work compared with their less productive peers (Simonton, 2004).

Table 2. Comparison of distribution of the high impact factor journals between leading and other scientists

Impact factor (2008)	Journal	Leading scientists			Other scientists		
		# Records	% of total publications	Mean impact factors weighted by publications	# Records	% of total publications	Mean impact factors weighted by publications
74.575	CA-Cancer J. Clin	0	0.000%	0.000	3	0.004%	0.003
35.423	Nat. Rev. Mol. Cell Biol	3	0.188%	0.067	30	0.040%	0.014
35	Physiol. Rev	0	0.000%	0.000	9	0.012%	0.004
31.434	Nature	54	3.379%	1.062	1147	1.527%	0.480
31.253	Cell	6	0.375%	0.117	146	0.194%	0.061
30.762	Nat. Rev. Cancer	0	0.000%	0.000	8	0.011%	0.003
30.259	Nature Genet	0	0.000%	0.000	17	0.023%	0.007
30.016	Annu. Rev. Biochem	5	0.313%	0.094	29	0.039%	0.012
28.69	Nat. Rev. Drug Discov	1	0.063%	0.018	21	0.028%	0.008
28.103	Science	70	4.380%	1.231	1675	2.231%	0.627
27.553	Nat. Med	2	0.125%	0.034	28	0.037%	0.010
24.962	Cancer Cell	0	0.000%	0.000	7	0.009%	0.002
24.185	Nat. Rev. Genet	0	0.000%	0.000	8	0.011%	0.003
22.731	Annu. Rev. Cell Dev. Biol	1	0.063%	0.014	26	0.035%	0.008
22.297	Nat. Biotechnol	10	0.626%	0.140	104	0.138%	0.031
22.192	Annu. Rev. Plant Physiol. Plant Molec. Biol	0	0.000%	0.000	1	0.001%	0.000
21.936	Pharmacol. Rev	1	0.063%	0.014	5	0.007%	0.001
21.561	Annu. Rev. Pharmacol. Toxicol	0	0.000%	0.000	3	0.004%	0.001
	Total	153	9.574%	2.791	3267	4.351%	1.275

Note: the number of total publications of the leading scientists is 1,598; the number of total publications of the other scientists in NBMS is 75,091. The mean impact factors weighted by publications in different groups are calculated as the percentage of total publications of each paper multiplied by its journal impact factor. Publications are gathered for 1990-2009. Journal ISSN numbers are associated to their Journal Impact Factor for 2008.

4.2.Analyses of Interdisciplinarity

We next turn to the researchers' cross-disciplinary and collaboration patterns. "Multidisciplinary" can be defined as "a basic situation in which elements from different disciplines are present," while "interdisciplinary" is "a more advanced stage of the relationship between disciplines in which Integration between them is attained" (Porter et al., 2007). Integration of research-based knowledge is the key concept (National Academies, 2005). In the present paper, we address interdisciplinarity in terms of the spectrum of scholarly work. We operationalize this in terms of the WOS Subject Categories.

ISI (Thomson Reuters) assigns journals to SCs based on journal-to-journal citation patterns and editorial judgment. There are arguments concerning the accuracy of these classifications (Bensman & Leydesdorff, 2009; Pudovkin & Garfield, 2002). However the SCs remain the most widely used and most easily accessible, with assignment errors generally among neighboring SCs. Global science maps have been shown to be relatively robust to such classification issues

(Rafols & Leydesdorff, 2009; Rafols et al., 2010a). We use these in our indices and science overlay maps here.

As shown in Table 3, the dominant disciplines[4] of the 21 leading NBMS scientists are pharmacology & pharmacy [5 authors out of 21], biochemistry & molecular biology [8 authors], biophysics [7 authors], and multidisciplinary Sciences [3 authors]. Based on their individual dominant discipline, we calculate the Multidisciplinary Index using the following formula based on SCs (Heinze & Bauer, 2007):

$$\text{Multidisciplinary Index} = (1 - GINI) \sum_{i=1}^{n} \mathrm{SC_i}, \qquad (1)$$

where SC_{i} denotes the number of Subject Categories, $GINI = 1 - \frac{1}{n}(2\sum_{i=1}^{n-1} wSC_i + 1)$, where w_i is the cumulative percentage of the publication counts in the ith SC to the total publication counts.

The index combines the number of subject categories in which scientists publish and the concentration of their publications across these subject categories using a GINI coefficient. The index increases when scientists publish in different subject categories; it decreases when they publish most of their work in few subject categories. Consequently, if scientists publish across various subject categories but show a high concentration in a few, they receive lower values than those scientists with a more equal distribution in their publishing activity (Heinze & Bauer, 2007).

Interdisciplinarity is measured using interdisciplinary research ("IDR") indicators (Porter et al., 2007; Porter & Rafols, 2009; Porter et al., 2008a; Rafols & Meyer, 2010). One of the IDR indicators – Integration ("I") "measures the diversity of knowledge sources, as shown by the breadth of references cited by a paper" (Porter & Rafols, 2009). However, interdisciplinarity is not only reflected by the diversity of source knowledge, it is also reflected by Specialization ("S"), which "reflects the breadth of Subject Categories in which a researcher (or group of researchers) publishes" (Porter et al., 2008a). Different from Integration, Specialization draws upon the publications' Subject Categories rather than their references' SCs, so Integration and Specialization can potentially behave independently (Porter et al., 2008a). The Specialization scores and Integration scores presented in Table 3 are calculated by combining the whole body of work of each author for the given period. Entries are ordered by Integration scores, ranging from most to least interdisciplinary.

Table 3. Discipline profiling, multidisciplinary and interdisciplinary indexes by authors

Author	I	S	MI	Ave. # of Authors	Subject Category Top 5 Items	Keywords Top 5 Items
Weissleder, R [70]	0.61	0.46	2.63	5.6	Biochemical Research Methods [30]; **Biochemistry & Molecular Biology** [29]; Chemistry, Organic [20]; Chemistry, Multidisciplinary [19]; Oncology [13]	cancer [15]; nanoparticle [15]; peptide [15]; in-vivo [15]; superparamagnetic iron-oxide [10]; cells [10]
Davies, M C [56]	0.61	0.53	2.52	6.6	**Pharmacology & Pharmacy** [22]; **Biochemistry & Molecular Biology** [18]; Chemistry, Multidisciplinary [16]; Microscopy [9]; Chemistry, Organic [5]	atomic force microscopy [19]; AFM [12]; DNA [10]; surface [9]; adhesion [8]; protein [8]
Langer, R [61]	0.6	0.46	2.69	7.2	**Pharmacology & Pharmacy** [18]; Chemistry, Multidisciplinary [16]; Biotechnology & Applied Microbiology [14]; Multidisciplinary Sciences [9]; Oncology [9]	nanoparticle [19]; in-vivo [10]; gene delivery [9]; liposome [8]; nanosphere [8]; drug-delivery [8]; protein [8]
Couvreur, P [160]	0.59	0.56	2.06	6	**Pharmacology & Pharmacy** [128]; Chemistry, Multidisciplinary [51]; Biochemistry & Molecular Biology [18]; Biochemical Research Methods [10]; Chemistry, Medicinal [8]	nanoparticle [88]; liposome [30]; nanocapsule [24]; polyalkylcyanoacrylate nanoparticle [23]; nanosphere [21]
Lieber, C M [60]	0.59	0.86	1.67	4.1	**Multidisciplinary Sciences** [44]; Biochemistry & Molecular Biology [8]; Biophysics [3]; Biotechnology & Applied Microbiology [3]; Microscopy [2]	carbon nanotube [16]; devices [8]; quantum wires [7]; growth [7]; atomic force microscopy [4]; silicon nanowires [4]; building-blocks [4]; transistors [4]; transport [4]; silicon [4]; heterostructures [4]
Hansma, P K [51]	0.59	0.76	2.3	6.5	**Biophysics** [23]; Microscopy [11]; Multidisciplinary Sciences [11]; Biochemistry & Molecular Biology [10]; Cell Biology [4]	atomic force microscopy [11]; resolution [9]; protein [8]; DNA [7]; molecules [6]; scanning tunneling microscopy [6]
Kreuter, J [120]	0.57	0.55	1.98	5.1	**Pharmacology & Pharmacy** [102]; Chemistry, Multidisciplinary [29]; Chemistry, Applied [15]; Engineering, Chemical [14]; Biochemistry & Molecular Biology [7]	nanoparticle [81]; delivery [24]; rat [19]; doxorubicin [18]; microsphere [18]
Muller, R H [157]	0.55	0.64	1.8	3.8	**Pharmacology & Pharmacy** [145]; Chemistry, Multidisciplinary [33]; Chemistry, Medicinal [16]; Engineering, Chemical [10]; Chemistry, Applied [9]	Solid lipid nanoparticle[54]; SLN [33]; controlled drug-delivery[30]; nanoparticle [27]; nanosuspension [20]
Gaub, H E [61]	0.52	0.81	2.27	5	**Biophysics** [30]; Multidisciplinary Sciences [14]; Biochemistry & Molecular Biology [14]; Cell Biology [8]; Microscopy [4]	microscopy [12]; cells [12]; spectroscopy [11]; AFM [11]; atomic force microscopy [10]
Hinterdorfer, P [67]	0.49	0.66	2.55	6.9	**Biophysics** [24]; **Biochemistry & Molecular Biology** [24]; Biochemical Research Methods [10]; Microscopy [9]; Cell Biology [8]; Chemistry, Multidisciplinary [8]	atomic force microscopy [21]; AFM [20]; spectroscopy [16]; DNA [12]; molecular recognition [12]
Dekker, C [51]	0.46	0.92	2.23	5.2	**Multidisciplinary Sciences** [23]; **Biophysics** [15]; **Biochemistry & Molecular Biology** [15]; Cell Biology [7]	protein [7]; mechanism [7]; binding [6]; transport [6]; escherichia-coli [6]
Schulten, K [65]	0.4	0.91	1.9	4.5	**Biophysics** [48]; Biochemistry & Molecular Biology [17]; Cell Biology [9]; Multidisciplinary Sciences [7]; Biochemical Research Methods [3]	molecular-dynamics [16]; atomic force microscopy [12]; protein [11]; molecular-dynamics simulations [9]; mechanism [9]

Bustamante, C [82]	0.36	0.87	2.47	4.9	**Multidisciplinary Sciences** [28]; Biophysics [27]; Biochemistry & Molecular Biology [26]; Cell Biology [12]; Microscopy [4]	single molecule [20]; DNA [16]; force [13]; optical tweezers [11]; protein [11]
Fernandez, J M [70]	0.35	0.89	2.22	4.8	**Biophysics** [34]; Multidisciplinary Sciences [25]; Biochemistry & Molecular Biology [16]; Cell Biology [7]; Biotechnology & Applied Microbiology [2]	elasticity [18]; titin [14]; stability [13]; single molecule [11]; protein [10]
Engel, A [97]	0.33	0.88	2.15	5.9	**Biochemistry & Molecular Biology** [66]; Cell Biology [33]; Biophysics [33]; Microscopy [10]; Microbiology [4]; Multidisciplinary Sciences [4]	atomic force microscopy [26]; transmission electron-microscopy [18]; protein [18]; resolution [17]; escherichia-coli [14]
Muller, D J[80]	0.32	0.86	2.23	5	**Biochemistry & Molecular Biology** [50]; Biophysics [33]; Cell Biology [26]; Microscopy [4]; Biochemical Research Methods [4]; Biotechnology & Applied Microbiology [4]	atomic force microscopy [39]; bacteriorhodopsin [19]; angstrom resolution [16]; escherichia-coli [16]; purple membrane [16]
Yoshida, M [71]	0.3	0.8	2.48	6	**Biochemistry & Molecular Biology** [31]; Biophysics [24]; Cell Biology [22]; Multidisciplinary Sciences [14]; Oncology [5]	ATP synthase [19]; F1-ATPase [17]; rotation [17]; F-1-ATPase [15]; resolution [13]
Aebi, U [71]	0.25	0.89	2.13	5.8	**Biochemistry & Molecular Biology** [47]; Cell Biology [32]; Biophysics [16]; Microscopy [3]; Multidisciplinary Sciences [3]	protein [25]; transmission electron-microscopy [18]; atomic force microscopy [16]; in-vitro [15]; electron microscopy [13]
Ha, T [62]	0.23	0.92	2.27	4.6	**Biophysics** [27]; Biochemistry & Molecular Biology [20]; Multidisciplinary Sciences [17]; Cell Biology [8]; Biochemical Research Methods [2]	resonance energy-transfer [15]; protein [11]; mechanism [11]; single-stranded-DNA [7]; crystal-structure [7]; escherichia-coli [7]; single molecule [7]; translocation [7]
Spudich, J A [51]	0.23	0.9	2.46	4	**Biochemistry & Molecular Biology** [21]; Cell Biology [19]; Multidisciplinary Sciences [15]; Biophysics [14]; Biology [6]	actin [16]; mechanism [12]; movement [11]; Motor [11]; molecular motor [9]
Yanagida, T [101]	0.18	0.83	2.26	4.9	**Biophysics** [48]; Biochemistry & Molecular Biology [28]; Biology [18]; Multidisciplinary Sciences [12]; Cell Biology [10]	movement [25]; force [18]; steps [15]; actin [15]; myosin [14]; binding [14]
Mean	**0.44**	**0.76**	**2.25**	**5.3**	**----**	**----**
Standard Deviation	**0.15**	**0.16**	**0.27**	**0.2**	**----**	**----**
Average for all NBMS papers	**0.27**	**---**	**2.72**	**4.8**	**Biochemistry & Molecular Biology [29.17%]; BIophysics [14.82%]; Pharmacology & Pharmacy [13.29%]; Multidisciplinary Sciences [10.84%]; Biotechnology & Applied Microbiology [10.48%]**	protein [7.69%]; nanoparticle [6.02%]; expression [4.44%]; in-vitro [4.32%]; cells [4.05%]; atomic force microscopy [3.36%]; binding [3.32%]; in vivo [3.09%]

Note: I→ Integration; S→ Specialization; MI → Multidisciplinary Index; the **bold** subject categories are the dominant disciplines. We also examined the average Integration of each paper, and the results show that there is a highly positive relationship between the average Integration and the combined Integration (0.988).

We see a striking range in degree of interdisciplinarity among these leading NBMS researchers. Our main indicator of interdisciplinarity is the Integration score (c.f., Porter et al., 2007). To provide a benchmark at the individual paper level, 775 papers published in 2005, sampled from the Web of Science Medicine, Research & Experimental Subject Category, show an average Integration score of 0.44 (standard deviation of 0.11) (Porter and Rafols, 2009). So

our author average is coincidentally also 0.44, but the individual author averages reach well over a standard deviation above and below this benchmark.[5]

Integration reflects diversity in references; Specialization reflects (1- diversity) in publications; and the Multidisciplinary Index reflects the diversity in publication. Specifically, Integration score is a particular way of looking at interdisciplinarity in terms of disciplinary diversity, as discussed in Rafols and Meyer (2009). The formula for computing Integration is:

$$I = 1 - \left[\frac{\left[\sum \left(f_i \times f_j \times COS\left(SC_i - SC_j \right) \right) \right]}{\sum \left(f_i \times f_j \right)} \right]$$

$$where \;\; i = row; \; j = column; \; f = \; frequency$$

Where the Subject Categories (SCs) are those cited in a given paper. The summation is taken over the cells of this SC x SC matrix.

"COS (SCi – SCj)" is the cosine measure of association between any two SCs.

The cosine measure is a statistical relative of correlation measures. It can be calculated from a series of paired observations (x, y) as:

$$\frac{\sum_i \left(x_i y_i \right)}{\sqrt{\sum_i x_i^2 \sum_i y_i^2}}$$

Conceptually, it can be thought of as the angle between two vectors – the closer they align, the closer the cosine value approaches 1. The cosine values are based on a year's worth of cross-journal citation in the Web of Science (2007), as calculated by Loet Leydesdorff.

We anticipate that Integration will be positively related to the Multidisciplinary Index and negatively related to Specialization; the Multidisciplinary Index should be more strongly and negatively related to Specialization, as both deal with diversity of publication outlets.

As shown in Table 4, among our 21 leading scientists, Integration is negatively associated with Specialization ($r = -0.728$, $p<0.01$), which is consistent with our expectation and previous research (Porter et al., 2008a). But it is more interesting that we can identify three groups of scientists with different patterns – Group I is in the upper right with high Integration and high Specialization; Group II is in the upper left of Figure 1 with high Integration and low Specialization; and Group III is in the lower right with low Integration and high Specialization. Examining within each sub-group, we do not find significant relations between Integration and Specialization. We also calculated the correlations between counts, Integration and Specialization, and there are not significant relationships, neither of subject categories and topics which we examined.

Table 4: Correlations among interdisciplinary indicators and number of authors

	Specialization	Integration	Multidisciplinary Index	Average Number of Authors
Specialization	1	-.728**	-0.323 (p=0.154)	-0.521*
Integration	-0.728**	1	-0.049 (p=0.833)	0.347 (p=0.123)
Multidisciplinary Index	-0.323 (p=0.154)	-0.049 (p=0.833)	1	0.604**
Average Number of Authors	-0.521*	0.347 (p=0.123)	0.604**	1

* p<0.05; ** p<0.01

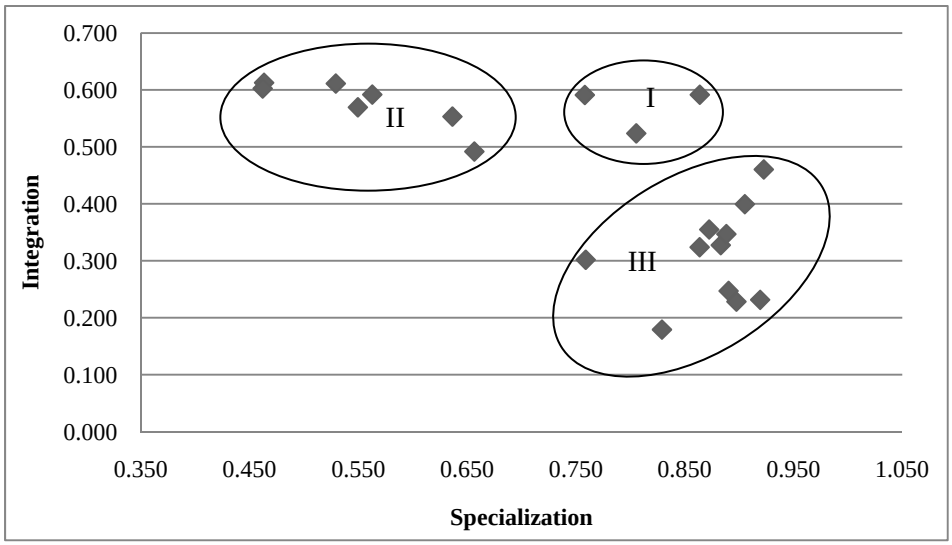

Figure 1. Scatter plot of Specialization vs. Integration for the 21 leading NBMS scientists

Figure 2 presents the relationship of the Multidisciplinary Index, Specialization and Integration. There are no significant overall correlations between Integration and Specialization with the Multidisciplinary Index, which differs from our expectations.

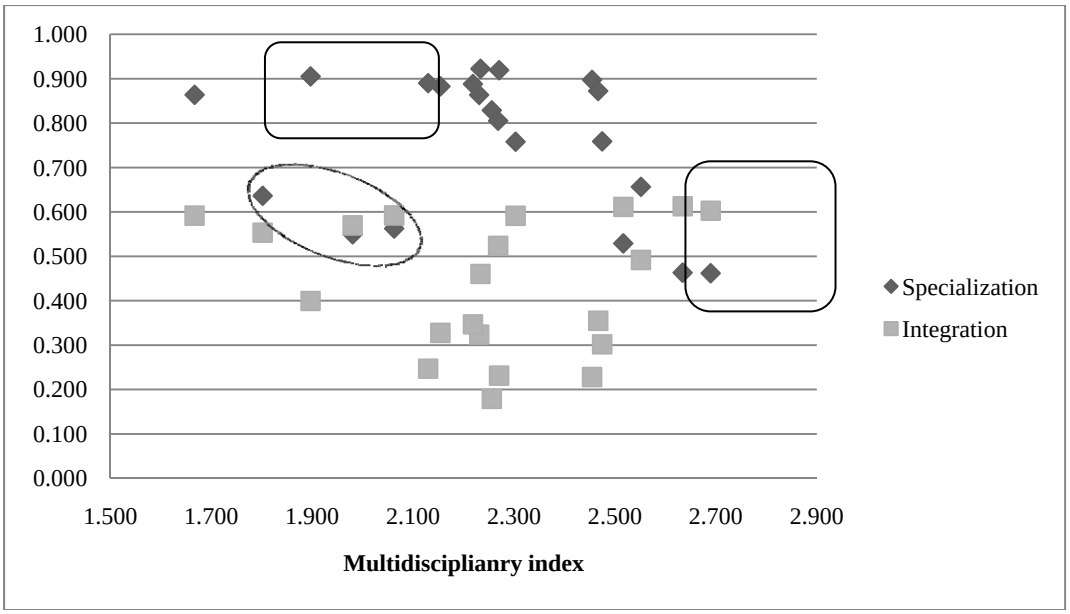

Figure 2. Scatter plot of Multidisciplinary Index and the interdisciplinarity indexes of the 21 leading scientists

Publication behaviors (reflected in Specialization) and citation behaviors (reflected in Integration) could be quite independent in some situations (Porter et al., 2007; Porter et al., 2008a). We wondered why Specialization has no significant relationship with the Multidisciplinary Index here? Consider their formulas – Specialization is

$$S = [\frac{[\sum(f_i \times f_j \times Cos(SC_i - SC_j))]}{\sum(f_i \times f_j)}] \qquad (2)$$

where i=row; j=column; f=frequency; SCs are the Subjective Categories in which the set of papers was published; COS(SC_i-SC_j) is the cosine measure of association between the two SCs, which can be calculated from series of paired observations (Porter et al., 2008a).

Comparing Equations (1) and (2), a key difference from Specialization is that the Multidisciplinary Index is based on a Gini coefficient. The Gini coefficient was developed to measure the degree of concentration (inequality) of a variable in a distribution of its elements (in this paper, the authors' publication distribution over SCs), while Specialization is based on the degree of association among different SCs based on the extent of their cross-citation in a full year's WOS publications.

As presented in Table 3, the range of Specialization ("S") is from 0.46- 0.92 and the range of the Multidisciplinary Index ("MDI") is from 1.69 to 2.69. Based on Figure 2, we selected individuals to examine more closely.

The anticipated negative correlation between S and MDI suggests two "well-behaved" types -- Lo S--Hi MDI and Hi S—Lo MDI (rectangles in Figure 2). We selected Langer and Weissleder from the former, and their publication patterns fit the expectation – relatively broad distribution (24 or 25 SCs) without an overly prominent peak. We chose Shulten to represent the second group, and found an expected distribution – only 7 SCs, with a sharp peak (48 Biophysics papers, followed by 17 Biochemistry papers). Furthermore, the disparity between Biophysics and Biochemistry is exceptionally low – they show a correlation of 0.97 based on the SC cross-citation pattern.

We identify three outliers representing the "Lo S—Lo MDI" region – Muller, R H (1.802, 0.636), Kreuter, J (1.980, 0.550), and Couvreur, P (2.062, 0.563) (ellipse in Figure 2). They are the top 3 publishers in our sample; their dominant subject categories are all Pharmacology & Pharmacy, with over 100 papers in that SC. Their MDI should be relatively low, reflecting this concentration. Their concentration pattern for all 3 is similar: 145 (Pharm) - 33(Chem,Multi) - 16(Chem, Med); 102(Pharm) – 29(Chem,Multi) -15 (Chem,Applied); 128 (Pharm) – 51(Chem,Multi) -18(Biochem);. Yet their S value is not so high because they do publish considerably in other SCs, and the correlations among these are moderate.

MDI reflects disciplinary distribution, whereas S taps into disciplinary diversity (e.g., how distinct the SCs in which an author publishes are from each other). Stirling (2007) offers a framework for diversity that serves well to understand this distinction. He posits three components of diversity – *variety* (here reflecting in the number of SCs represented), *balance* (the distribution among those SCs), and *disparity* (how similar those SCs are to each other). The Integration and Specialization score formulas address all three of these; the Multidisciplinary Index addresses only two (variety and balance). So we believe the differing behavior between the Multidisciplinary Index and Specialization is due to the latter incorporating *disparity* among the SCs in which an author publishes. We take Bustamante and Spudich to represent "Hi S—Hi MDI" outliers. Both show four prominent SCs, and these have high intercorrelations. This fits our interpretation well – were disparity among SCs taken into account in the MDI metric, their score would be considerably reduced.

Figure 3 shows the science overlay map for the publications by the 21 NBMS researchers and science overlay maps of the top two authors separately (Figure 3a and Figure 3b). The overall map shows, generally, that the leading scientists focus on the biomedical sciences and nearby macro-disciplines, like infectious diseases and clinical medicine, which is consistent with the results of the previous study of NBMS (Pei et al., to appear). And again, it testifies that nanotechnology engages a diverse set of disciplines, including physics, chemistry and biology (Ernest & Shetty, 2005) to generate new research fields.

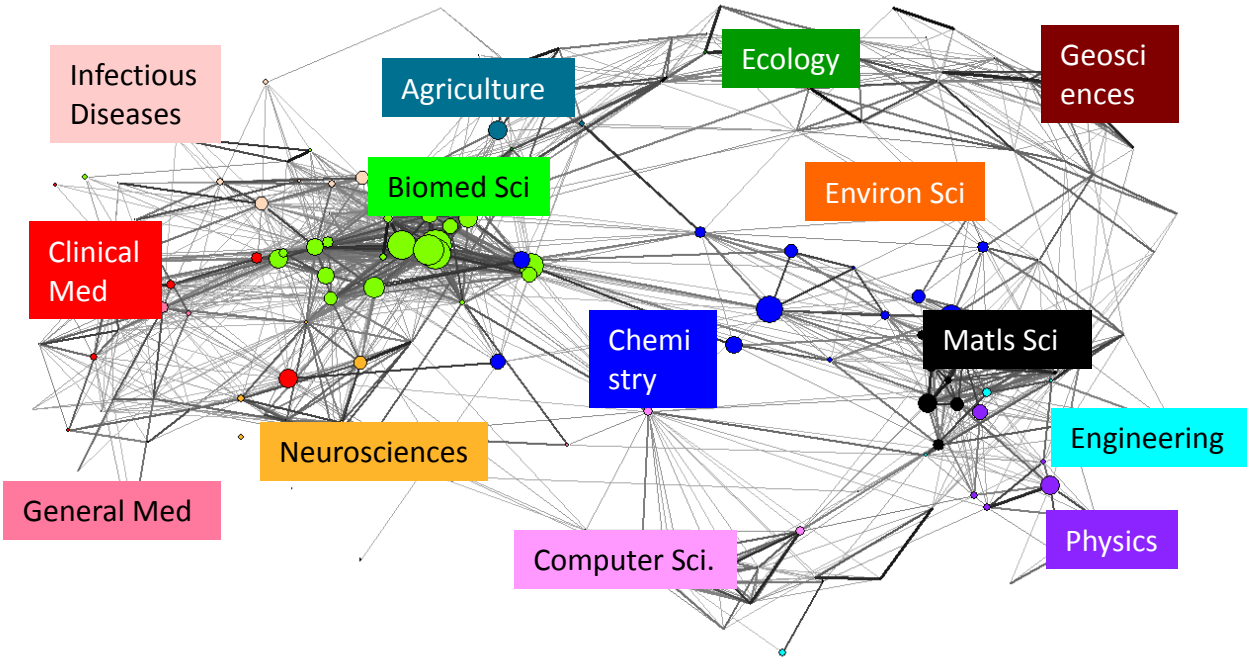

Figure 3. Publication profiles of the leading scientists overlaid on the map of science

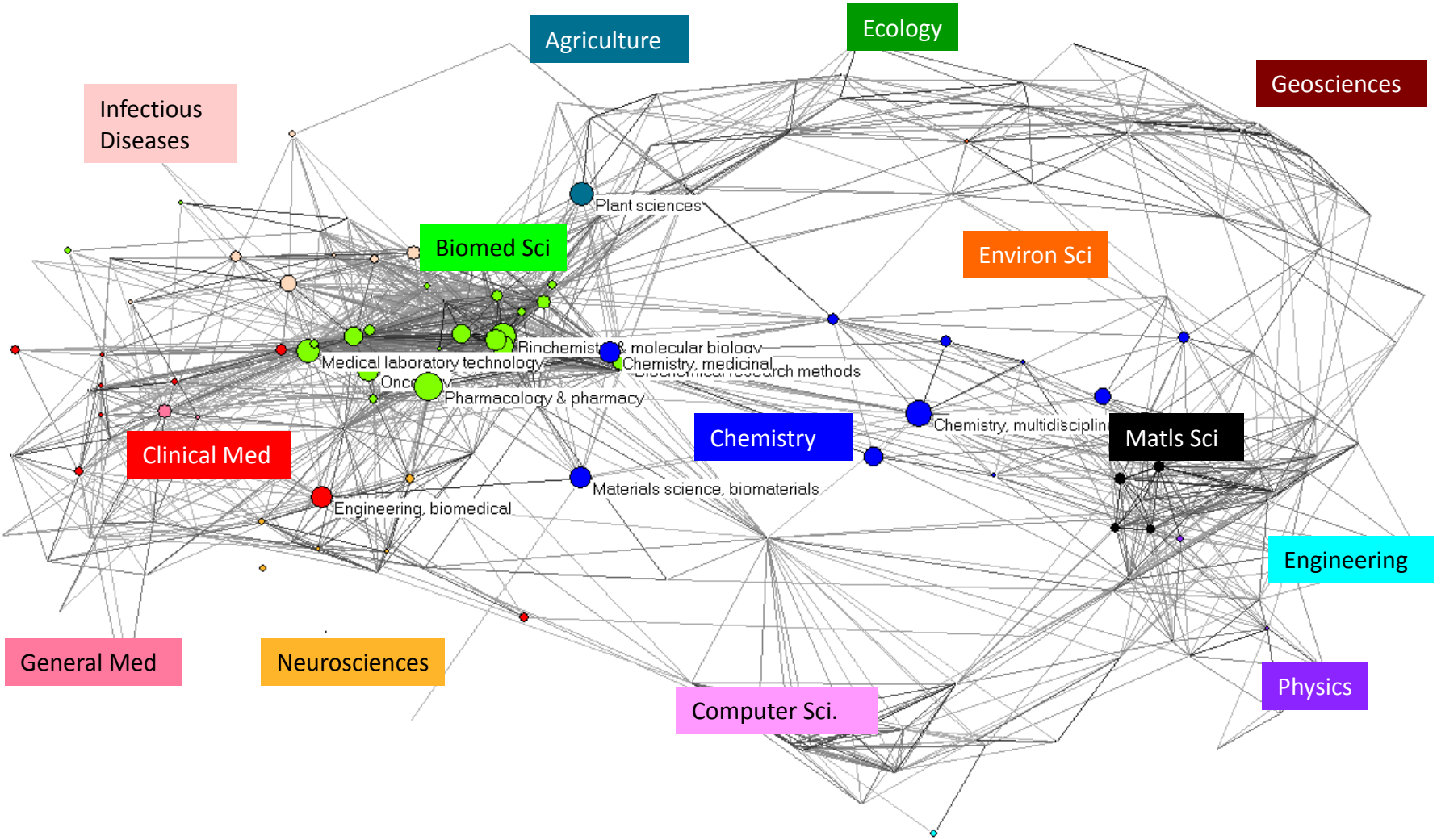

Figure 3a. Publication profiles of Couvreur's publications overlaid on the map of science

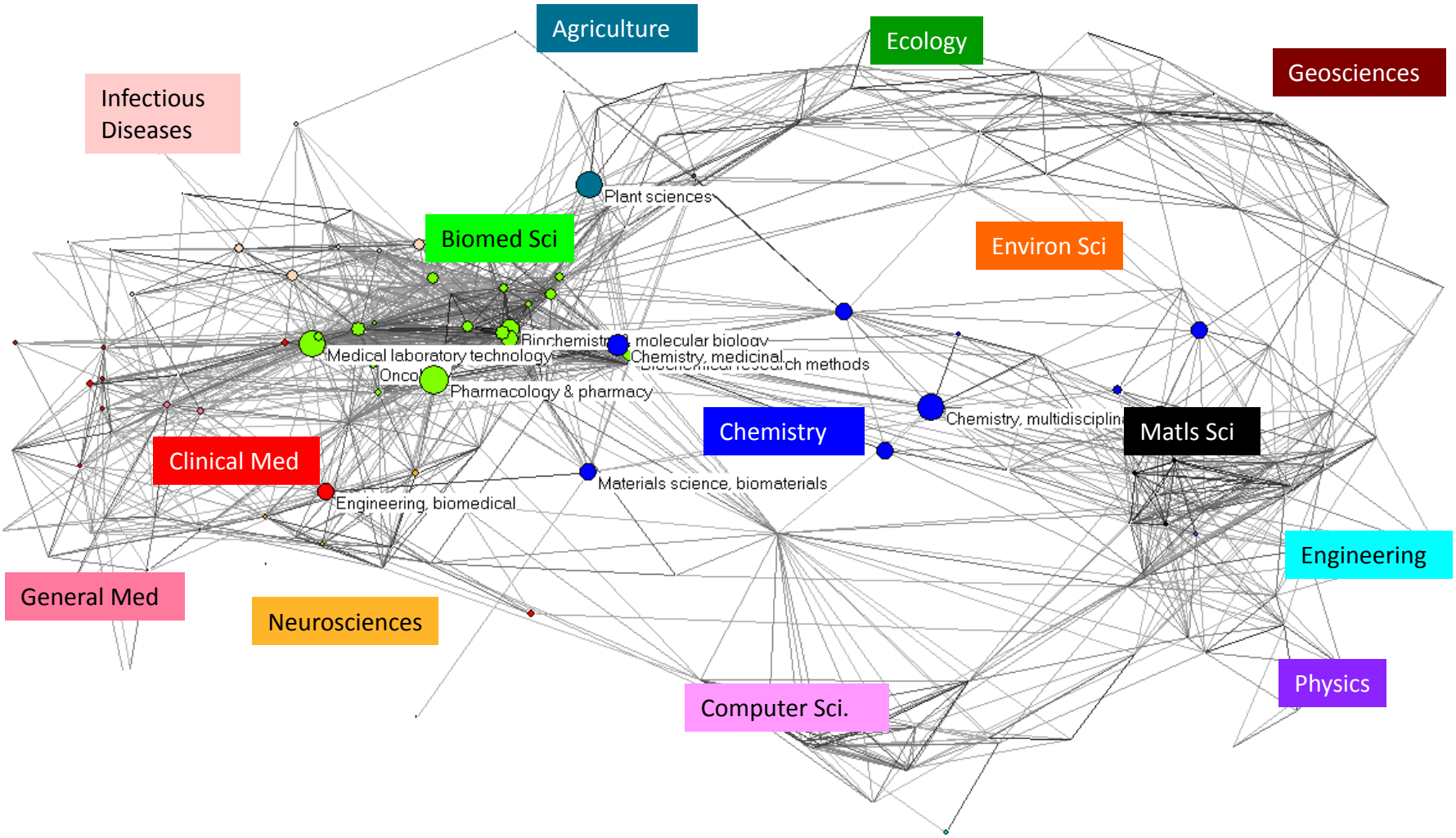

Figure 3b. Publication profiles of Muller's publications overlaid on the map of science

4.3.Collaboration Patterns

Co-authorship analysis considers the social aspect, focusing on research practices instead of information content (Schummer, 2004). This is similar to social network analysis, but answers specific questions about collaboration groups (Boyack, 2004). One could assume that, in general, each co-author has made a substantial contribution to the paper, so we expect that the larger the number of authors of a paper, the higher the possibility that it cites from different fields. For each author, the average number of authors of his or her papers generally reflects his or her collaborative patterns. The more he or she collaborates with other researchers, the higher the possibility that his or her research would be located across a relatively broad disciplinary spectrum. Therefore, we expect that the average number of authors is related to the multidisciplinary and interdisciplinary indices.

As shown in Table 4, the average number of authors, which is an index reflecting an author's research networking, is negatively related to Specialization and positively related to the Multidisciplinary Index. Although the coefficient of Integration and average number of authors is also relatively high (0.347), this is not significant at the 0.05 level (2-tailed). This is partly consistent with our expectation.

Meanwhile, the average number of authors per paper for the leading scientists group is 5.26, which is larger than the average number for the NBMS field without the leading scientists — 4.75 ($t=7.742$, $p<0.001$). This provides preliminary evidence that the leading scientists in NBMS collaborate somewhat more extensively than the other scientists in NBMS. In Figure 4, the distribution of number of authors of each publication shows that for leading scientists, the skewness (0.210) is larger than for the other scientists (0.049). And the mode of the leading scientists group is 5, higher than for the other researchers (mode =3). This result supports the expectation that the leading scientists have larger collaboration networks than their peer scientists.

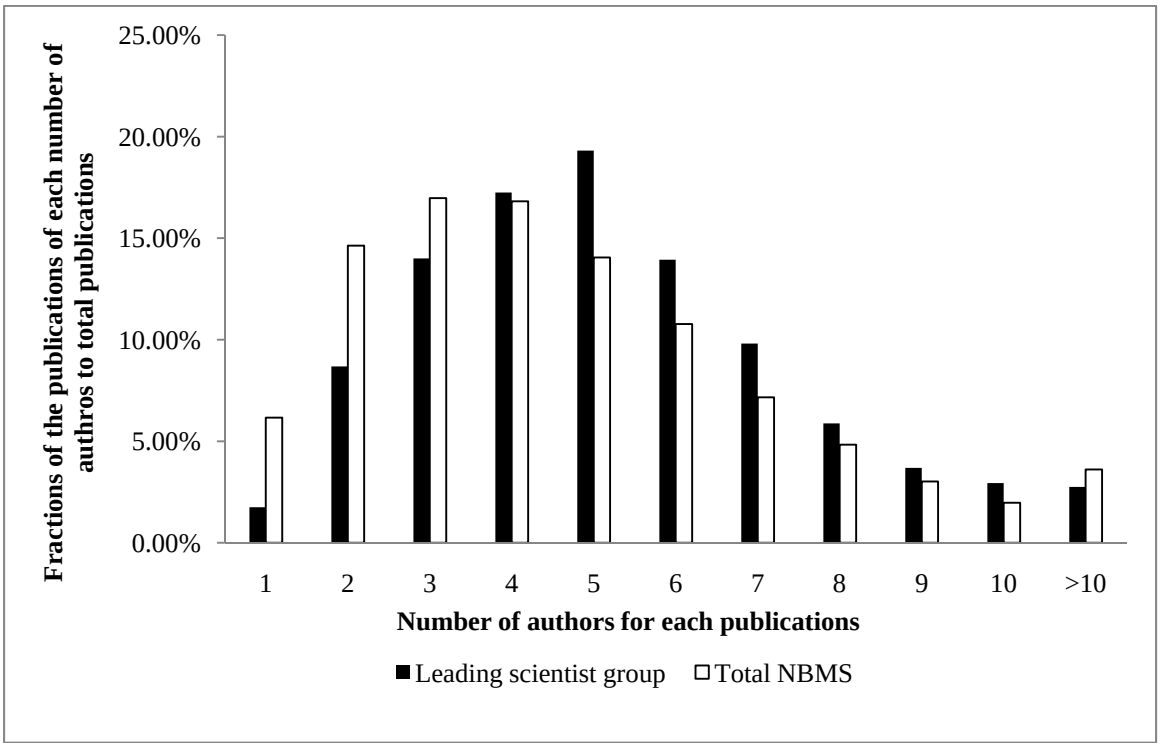

Figure 4. Collaboration patterns of leading scientists and the other authors in NBMS

4.4.Keyword Analyses

The next section considers the "hot topics" on which the leading scientists focus. "Hot topics" studied in this paper are somewhat similar to the concept of research front (Price, 1965). A research front represents "the state-of-art thinking of a research field" (Chen, 2006), which builds on recent work. In this study, we identify the hot topics in two different ways. First, we compare the list of the merged keywords (the keywords from authors and WOS "keywords-plus") for 2008-2009 publications with those in the previous years to see if there are any keywords that first show up, or increase rapidly, in recent years. [As an astute reviewer noted, "keywords-plus" derive from titles of cited articles, so inherently lag. Nonetheless, we find that tracking recent shifts in this keyword combination can help identify possibly noteworthy emerging topics.] The ratio of each keyword in 2008-2009 to those in the previous years and the frequency of each keyword are calculated and ranked. Subsequently, we omit the common words and categorize the remaining keywords based on their meaning. Our other way to seek hot topics is to compare the keywords of the leading scientists to those of the other scientists for the period of 2008-2009. We identify the keywords that only appear in the leading scientists group, omitting the common words. Again, we classify the remaining terms into different categories based on their meaning.

The leading scientists are focusing on several health issues like cancer (cancer, breast cancer, prostate cancer, cell lung cancer, prostate cancer, tumor, tumor cells, tumor bearing, glioblastoma), brain and neuro-diseases (blood brain barrier, neurodegenerative), heart diseases (carcinoma, myocardial infarct). These diseases are complex and hard-to-cure, and they need early diagnoses and in-vivo therapy. The results of keyword analyses show that the research of the leading scientists is increasingly focused on these health issues.[6]

Second, thrusts identified from the keywords analysis include: variations on force spectroscopy (atomic force microscope/y, molecule force spectroscopy, single molecule force spectroscopy, force-clamp spectroscopy, parelectric spectroscopy, ATPase, ATP synthase, F-1-ATPase, Na+-ATPase), optical tweezers (optical tweezers, laser tweezers), drug delivery (drug delivery, in-vivo delivery, local delivery), molecular motors (molecular motor, F-1 motor, rotary motor , flagellar motor, , ATPase, ATP synthase, F-1-ATPase, Na+-ATPase), heat therapy (hyperthermia, thermal coagulation therapy, thermotherapy). These topics show up increasingly in recent years in the 21 leading scientists' work.

Third, nanotechnology enhances techniques for diagnosis and therapy for some complex and hard-to-cure diseases. The nanoparticles identified in the leading scientists' recent research include nanocrystals, nanosuspension, magnetic nanoparticles, solid lipid nanoparticles, cyanoacrylate nanoparticles, and quantum dots. We nominate these NBMS sub-topics as strong candidates for further exploration.

To get after "new" technologies, we made a group of Natural Language Processing (NLP) title or abstract phrases that appeared from 2-9 times in the leading researchers' abstract records. We did this after confirming that analyses of the high frequency NLP terms did not identify new topics (i.e., those first appearing in the most recent years). We cleaned this list in VantagePoint (fuzzy matching). Then we selected 59 candidate "relatively rare" terms (appearing 7 or more times). We used "Find" to locate other versions and to combine variations of terms (e.g., 31 variants on "single-molecule"). We composed groups for each new term or phrase; then made a field from those group names (i.e., consolidating term variations within each).

The resulting 49 terms or phrases include variations of the terms so that the frequency now ranges from 149 to 7 occurrences. Incorporating variations also turns up that most of these terms have been in use for years. Here are some terms that these 21 leading researchers introduced into their research more recently:

- 2007: gemcitabine
- 2006: siRNA (small interfering RNA)
- 2005: ssDNA (single strand)
- 2004: NhaA (a single transmembrane protein)

[All other terms date to 1999 or earlier, with the largest number, 11, from 1991.]

This process is presented as a simple pilot. It should be adjusted to one's driving interests (e.g., focusing on a particular sub-domain of interest). It does "rediscover" more general terms that were screened out by focus on these relatively low frequency terms (e.g., "colloidal drug carriers" generalizes back to "drug delivery"). Performing these key term scans and screens in consultation with field experts should prove better at identifying salient "hot" topics.

5. Conclusions

This paper explores the possibility of identifying leading researchers in a field and profiling their research endeavors. The leading scientists may influence the development of the field, so that tracking their research emphases might provide valuable intelligence concerning emerging directions.

We think the following are notable findings of this study:

(1) Using bibliometric methods, rather than peer review (Heinze & Bauer, 2007), offers a more practical way to address the leading scientists in a field. Presenting the country and affiliation distribution of the leading scientists offers interesting information on the focal locations of NBMS research. In contrast to a profile of NBMS research not keying on the leading researchers, the US here shows as more prominent and China as less so.

(2) This paper examines the relations of multiple interdisciplinarity indicators and collaboration patterns of the leading scientists in NBMS. The results are somewhat unexpected -- Integration and Specialization have no statistically significant relations with the Multidisciplinary Index (for this small sample of 21 authors). We explain the observation that Specialization is not significantly related with the Multidisciplinary Index based on the former incorporating a disparity dimension. We plan to investigate this further in other data sets.

(3) The average number of authors is positively related with the Multidisciplinary Index and negatively related with Specialization. The more collaborative the authors, the more they may widen the disciplinary spectrum in which their papers are published. The non-significant relationship between average number of authors and Integration is somewhat strange to us, but it shows a relatively high coefficient ($r= 0.347$, $p=0.123$).

(4) The leading scientists have larger collaboration networks than the other NBMS researchers, which is consistent with our expectation.

(5) Keyword analyses identify a number of candidate hot topics based on relative emphases of the leading NBMS scientists in recent years. PhD students in the area reviewed these and found them credible as possible leading indicators. [Review by leading researchers would strengthen this case.] Such results may help predict future NBMS R&D emphases. Such analyses could be highly valuable to R&D managers.

Our results show characteristics of the leading scientists in NBMS from the perspective of distribution of their countries and regions, of their collaboration patterns, and of their disciplinary spectrum. Visualization techniques help show the diffusion of NBMS findings across the sciences. We suggest that analysis of a field in terms of its leading researchers offers an interesting alternative to overall research profiling (Porter et al., 2002). In particular we see promise in ascertaining their recent research emphases as leading indicators of prospective R&D opportunities for the field in question.

Acknowledgements

This research was undertaken at Georgia Tech drawing on support from the National Science Foundation (NSF) through the Center for Nanotechnology in Society (Arizona State University; Award No. 0531194); and the Science of Science Policy Program -- "Measuring and Tracking Research Knowledge Integration" (Georgia Tech; Award No. 0830207). The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Our sincere thanks to three PhD students -- Hui Zhang whose major is biological sciences in the National Institute of Biological Sciences (NIBS) in Beijing; Xiaobo Hu, an exchange student at Georgia Tech, whose major is Chemical Engineering; and Yanni Lin, a Ph.D. student at Georgia Tech, whose interest is "Quantum dot labeling of proteins in nucleus." They reviewed the findings, particularly the results of keywords analyses. We also thank Stephen Carley, Georgia Tech, and Jon Garner, Search Technology, Inc., for calculating associations between Integration scores and Times Cited.

Endnotes:

1. Precisely, correlations are 0.067 for 6806 papers published in 1995 and 0.081 for 6817 papers published in 2005. They show mixed effects -- ranging from slightly negative (-0.09 and -0.20 for Medical Research & Experimentation in 1995 and 2005) to slightly positive (+0.15 for Mathematics in 1995 and also in 2005). R^2 is very low: 0.0045 for all 6806 papers in 1995 and 0.0066 for the 6817 papers together in 2005. Within the two disciplines with the extreme relationships (still small), R^2 is 0.0074 and 0.041 for Medical R&E, and 0.23 and 0.023 for 1995 and 2005, respectively. We also looked at samples of 1602 and 182 biosciences articles published 1999-2002; correlation between TC and Integration score was, respectively, 0.014 and 0.023. Again, highly cited articles appear to bubble up more frequently from middle level Integration score papers (moderate interdisciplinarity) than from very high or low ones.
2. VantagePoint is text mining software; variations are available as Thomson Data Analyzer and TechOASIS [www.theVantagePoint.com].
3. These are cell biology, biochemistry & molecular biology, developmental biology, biophysics, multidisciplinary sciences, genetics & heredity, medicine, research & experimental, biology, anatomy & morphology, andrology, microscopy, biotechnology & applied microbiology, biochemical research methods, endocrinology & metabolism, mathematical & computational biology, reproductive biology, physiology, pathology, oncology, medical laboratory technology, toxicology, pharmacology & pharmacy, obstetrics & gynecology, and medicine, legal.
4. The dominant disciplines are identified as the SC in which the authors published the most papers.
5. We calculate the Integration score based on the diversity of each author's collective references (all of his/her papers combined); if we calculate, instead, based on each paper individually, the mean for the 21 authors is 0.39 – again, not too different from the benchmark dataset.
6. Three PhD students, two from Biosciences and one from Chemical Engineering, reviewed this section and were comfortable with the thrusts represented as "hot topics," and emerging NBMS tools to which nano could contribute.

References

Abramo, G., D'Angelo, C. A. and Di Costa, F. (2009) Research collaboration and productivity: is there correlation? *Higher Education*, **57**, 155-171.

Adams, J., Jackson, L. and S., M. (2007) Bibliometric analysis of interdisciplinary research: Report to the Higher Education Funding Council for England.

Bensman, S. J. and Leydesdorff, L. (2009) Definition and Identification of Journals as Bibliographic and Subject Entities: Librarianship Versus ISI Journal Citation Reports Methods and Their Effect on Citation Measures. *Journal of the American Society for Information Science and Technology*, **60**, 1097-1117.

Borner, K., Chen, C. M. and Boyack, K. W. (2003) Visualizing knowledge domains. *Annual Review of Information Science and Technology*, **37**, 179-255.

Boyack, K. W. (2004) Mapping knowledge domains: Characterizing PNAS. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 5192-5199.

Boyack, K. W., Borner, K. and Klavans, R. (2009) Mapping the structure and evolution of chemistry research. *Scientometrics*, **79**, 45-60.

Boyack, K. W., Klavans, R. and Borner, K. (2005) Mapping the backbone of science. *Scientometrics*, **64**, 351-374.

Braun, T., Schubert, A. and Zsindely, S. (1997) Nanoscience and nanotechnology on the balance. *Scientometrics*, **38**, 321-325.

Buel, W. D., Albright, L. E. and Glennon, J. R. (1966) A Note on Generality and Cross-Validity of Personal History for Identifying Creative Research Scientists. *Journal of Applied Psychology*, **50**, 217-&.

Burt, R. S. (1992) *Structural Holes: The Social Structure of Competition*. London: Harvard University Press.

Burt, R. S. (2004) Structural Holes and Good Ideas. *The American Journal of Sociology*, **110**, 349-399.

Calero, C., Buter, R., Valdes, C. C. and Noyons, E. (2006) How to identify research groups using publication analysis: an example in the field of nanotechnology. *Scientometrics*, **66**, 365-376.

Chen, C. M. (2006) CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, **57**, 359-377.

Chen, C. M., McCain, K., Boyack, K. W., Lin, X. and Morris, S. A. (2002) Mapping the knowledge. In: Toms, E. G., ed. *Asist 2002: Proceedings of the 65th Asist Annual Meeting, Vol 39, 2002*. Medford: Information Today Inc, 511-512.

Chiu, W. T. and Ho, Y. S. (2007) Bibliometric analysis of tsunami research. *Scientometrics*, **73**, 3-17.

Christodoulos, C., Michalakelis, C. and Varoutas, D. (2010) Forecasting with limited data: Combining ARIMA and diffusion models. *Technological Forecasting and Social Change*, **77**, 558-565.

de Bellis, N. (2009) *Bibliometrics and Citation Analysis*. Lanham, MD: The Scarecrow Press.

Ernest, H. and Shetty, R. (2005) Impact of Nanotechnology on Biomedical Sciences: Review of Current Concepts on Convergence of Nanotechnology With Biology *Journal of Nanotechnology Online*.

ESF. (2005) ESF Forward Look on Nanomedicine. Le Bischenberg: European Medical Research Councils (EMRC)

Ganguly, I., Koebel, C. T. and Cantrell, R. A. (2010) A categorical modeling approach to analyzing new product adoption and usage in the context of the building-materials industry. *Technological Forecasting and Social Change*, **77**, 662-677.

Glänzel, W. and Schubert, A. (2004) Analysing Scientific Networks Through Co-Authorship. In: Moed, H. F., Glänzel, W. and Schmoch, U., eds. *Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in Studies of S&T Systems* Dordrecht ; Boston Kluwer Academic Publishers.

Grodal, S. and Thoma, G. (2008) Cross-pollination in science and technology: Concept mobility in the nanobiotechnology field *NBER Conference on Emerging Industries: Nanotechnolgy and NanoIndicators*. Cambridge, MA, USA.

Guan, J. and Wang, G. (2010) A comparative study of research performance in nanotechnology for China's inventor–authors and their non-inventing peers. *Scientometrics*.
Heinze, T. and Bauer, G. (2007) Characterizing creative scientists in nano-S&T: Productivity, multidisciplinarity, and network brokerage in a longitudinal perspective. *Scientometrics*, **70**, 811-830.
Heinze, T., Shapira, P., Senker, J. and Kuhlmann, S. (2007) Identifying creative research accomplishments: Methodology and results for nanotechnology and human genetics. *Scientometrics*, **70**, 125-152.
Hullmann, A. (2006) The economic development of nanotechnology -- An indicators based analysis *Nano S&T - Convergent Science and Technologies*.
Hullmann, A. (2007) Measuring and assessing the development of nanotechnology. *Scientometrics*, **70**, 739-758.
Hullmann, A. and Meyer, M. (2003) Publications and patents in nanotechnology - An overview of previous studies and the state of the art. *Scientometrics*, **58**, 507-527.
Kay, L. and Shapira, P. (2009) Developing nanotechnology in Latin America. *Journal of Nanoparticle Research*, **11**, 259-278.
Klavans, R. and Boyack, K. W. (2008) Thought leadership: A new indicator for national and institutional comparison. *Scientometrics*, **75**, 239-250.
Klavans, R. and Boyack, K. W. (2009) Toward a Consensus Map of Science. *Journal of the American Society for Information Science and Technology*, **60**, 455-476.
Larivière, V. and Gingras, Y. (2010) On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology*, **61**, 126-131.
Lee, S. and Bozeman, B. (2005) The impact of research collaboration on scientific productivity. *Social Studies of Science*, **35**, 673-702.
Levitt, J. M. and Thelwall, M. (2008) Is multidisciplinary research more highly cited? A macrolevel study. *Journal of the American Society for Information Science and Technology*, **59**, 1973-1984.
Levitt, J. M. and Thelwall, M. (2009) The most highly cited Library and Information Science articles: Interdisciplinarity, first authors and citation patterns. *Scientometrics*, **78**, 45-67.
Leydesdorff, L. and Rafols, I. (2009) A Global Map of Science Based on the ISI Subject Categories. *Journal of the American Society for Information Science and Technology*, **60**, 348-362.
Li, L. L., Ding, G. H., Feng, N., Wang, M. H. and Ho, Y. S. (2009) Global stem cell research trend: Bibliometric analysis as a tool for mapping of trends from 1991 to 2006. *Scientometrics*, **80**, 39-58.
Liu, X., Zhang, P. Z., Li, X., Chen, H. C., Dang, Y., Larson, C., Roco, M. C. and Wang, X. W. (2009) Trends for nanotechnology development in China, Russia, and India. *Journal of Nanoparticle Research*, **11**, 1845-1866.
McGrath, W. E. (1996) The unit of analysis (objects of study) in bibliometrics and scientometrics. *Scientometrics*, **35**, 257-264.
Meyer, M. (2006) Knowledge integrators or weak links? An exploratory comparison of patenting researchers with their non-inventing peers in nano-science and technology. *Scientometrics*, **68**, 545-560.
Meyer, M., Makar, I., Rafols, I., Olsen, D., Wagner, V., Zweck, A., Porter, A. L. and Youtie, J. (2005) Euronano: nanotechnology in Europe: Assessment of the current state, opportunities, challenges and socio-economic impact.
Meyer, M. and Persson, O. (1998) Nanotechnology - Interdisciplinarity, patterns of collaboration and differences in application. *Scientometrics*, **42**, 195-205.
Meyer, M., Persson, O., Power, Y. (2001) Nanotechnology Expert Group and Eurotech Data, Mapping Excellence in Nanotechnologies.
Michalakelis, C., Varoutas, D. and Sphicopoulos, T. (2010) Innovation diffusion with generation substitution effects. *Technological Forecasting and Social Change*, **77**, 541-557.
Pei, R. M., Youtie, J. and Porter, A. L. (to appear) Nano-biomedical Science in China: Growth and Challenges. *Technology Analysis & Strategic Management*.

Porter, A. L., Cohen, A. S., Roessner, J. D. and Perreault, M. (2007) Measuring researcher interdisciplinarity. *Scientometrics*, **72**, 117-147.
Porter, A. L., Kay, L., Carley, S. and Rafols, I. (2009) "How to" Overview.
Porter, A. L., Kongthon, A. and Lui, J. C. (2002) Research profiling: Improving the literature review. *Scientometrics*, **53**, 351-370.
Porter, A. L. and Rafols, I. (2009) Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics*, **81**, 719-745.
Porter, A. L., Roessner, J. D. and Heberger, A. E. (2008a) How interdisciplinary is a given body of research? *Research Evaluation*, **17**, 273-282.
Porter, A. L. and Youtie, J. (2009) Where does nanotechnology belong in the map of science? *Nature Nanotechnology*, **4**, 534-536.
Porter, A. L., Youtie, J., Shapira, P. and Schoeneck, D. J. (2008b) Refining search terms for nanotechnology. *Journal of Nanoparticle Research*, **10**, 715-728.
Price, D. J. D. (1965) Networks of scientific papers. *Science*, **149**, 510-&.
Pudovkin, A. I. and Garfield, E. (2002) Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology*, **53**, 1113-1119.
Rafols, I. and Leydesdorff, L. (2009) Content-Based and Algorithmic Classifications of Journals: Perspectives on the Dynamics of Scientific Communication and Indexer Effects. *Journal of the American Society for Information Science and Technology*, **60**, 1823-1835.
Rafols, I. and Meyer, M. (2007) How cross-disciplinary is bionanotechnology? Explorations in the specialty of molecular motors. *Scientometrics*, **70**, 633-650.
Rafols, I. and Meyer, M. (2010) Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience. *Scientometrics*, **82**, 263-287.
Rafols, I., Porter, A. L. and Leydesdorff, L. (2010a) Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*.
Rafols, I., Porter, A. L. and Leydesdorff, L. (2010b) Science overlay maps: A new tool for research policy and library managment. *Journal of the American Society for Information Science & Technology*.
Rinia, E. J., van Leeuwen, T. N., van Vuren, H. G. and van Raan, A. F. J. (2001) Influence of interdisciplinarity on peer review and bibliometric evaluations in physics research. *Research Policy*, **30**, 357-361.
Roco, M. C. (2003) Converging science and technology at the nanoscale: opportunities for education and training. *Nature Biotechnology*, **21**, 1247-1249.
Sastry, R. K., Rashmi, H. B., Rao, N. H. and Ilyas, S. M. (2010) Integrating nanotechnology into agri-food systems research in India: A conceptual framework. *Technological Forecasting and Social Change*, **77**, 639-648.
Schummer, J. (2004) Multidisciplinarity, interdisciplinarity, and patterns of research collaboration in nanoscience and nanotechnology. *Scientometrics*, **59**, 425-465.
Scott, J. (2000) *Social network analysis : a handbook*. London; Thousand Oaks; New Delhi: Sage.
Seale, C., Ziebland, S. and Charteris-Black, J. (2006) Gender, cancer experience and internet use: A comparative keyword analysis of interviews and online cancer support groups. *Social Science & Medicine*, **62**, 2577-2590.
Simonton, D. K. (1999) *Origins of Genius: Darwinian Perspectives on Creativity*. New York: Oxford University Press.
Simonton, D. K. (2004) *Creativity in Science: Chance, Logic, Genius, and Zeitgeist*. Cambrige: Cambridge University Press.
Sternberg, R. J. (2003) *Wisdom, Intelligence, and Creativity Synthesized*. Cambridge: Cambridge University Press.
Stirling, A. (2007) A general ramework for analysing diversity in science, technology and society. *Journal of the royal society interface*, **4**, 707-719.

Sutter, M. and Kocher, M. (2004) Patterns of co-authorship among economics departments in the USA. *Applied Economics*, **36**, 327-333.
TheNationalAcademies. (2005) *Facilitating Interdisciplinary Research*. Washington, DC: National Academies Press.
Tomellini, R., Faure, U. and Panzer, O. (2005) European Technology Platform on NanoMedicine: Nanotechnology for Health *Office for Official Publications of the European Communities*: European Commission.
Yegros-Yegors, A., Amat, C. B., d'Este, P., Porter, A. L. and Rafols, I. (2010) Does interdisciplinary research lead to higher scientific impact? *Science and Technology Indicators (STI) Conference*. Leiden, The Netherlands.
Yi, H. and Xi, Z. (2008) Trends of DDT research during the period of 1991 to 2005. *Scientometrics*, **75**, 111-122.
Zhou, W. (2003) Ethics of nanobiotechnology at the frontline. *Santa Clara Comput High Technol Law J*, **19**, 481-489.