Pre-publication Version
Porter, A.L., Garner, J., Newman, N.C., Carley, S.F., Youtie, J., Kwon, S., and Li, Y. (to appear), National nanotechnology research prominence, *Technology Analysis and Strategic Management;* available at: https://doi.org/10.1080/09537325.2018.1480013.

National Nanotechnology Research Prominence

Alan L. Porter[1 2], Jon Garner[2], Nils C. Newman[2], Stephen F. Carley[2], Jan Youtie[1], Seokbeom Kwon[1], Yin Li[1]

Contact Alan Porter: alan.porter@isye.gatech.edu; Search Technology, Inc., 6025 The Corners Pkwy., Norcross, GA, USA 30092.

Notes on contributors

Alan Porter is Professor Emeritus of Industrial & Systems Engineering, and of Public Policy, at Georgia Tech, where he is Co-director of the Program in Science, Technology & Innovation Policy (STIP). He is also Director of R&D for Search Technology, Inc., Norcross, GA . He is author or co-author of some 270 articles and books, including *Tech Mining* (Wiley, 2005) and *Forecasting and Management of Technology* (Wiley, 2011). Publications are available at: http://www.researchgate.net/profile/Alan_Porter4,

Jon Garner is a Senior Research Analyst at Search Technology. He earned a B.S. degree in Computational Media at Georgia Institute of Technology in 2009. His publications can be found on Google Scholar at *https://scholar.google.com/citations?user=f8mD5-gAAAAJ&hl=en&oi=ao*.

Nils C. Newman is the President of Search Technology, Inc., Norcross, GA. He has a Bachelor of Mechanical Engineering (1989) and an MS in Technology and Science Policy (1993) from Georgia Tech. Mr. Newman is currently pursuing a PhD in Economics from MERIT at the University of Maastricht in the Netherlands.

Stephen Carley is a Research Scientist at Search Technology. He recently finished a Ph.D. at the Georgia Institute of Technology. He's been published over a dozen times in peer-reviewed journals and has presented at more than a dozen conferences.

Jan Youtie is a Director of the Program in Science, Technology and Innovation Policy and Social and Ethical Implications Coordinator of the Southeastern Nanotechnology Infrastructure Coordinator at Georgia Institute of Technology. Her research interests are in emerging technology measurement and analysis, manufacturing, and innovation-based economic development. Her publications can be found at http://works.bepress.com/jan_youtie/ and on Google Scholar a thttps://scholar.google.com/citations?user=MYOwwTUAAAAJ&hl=en.

Seokbeom Kwon is a doctoral student of the School of Public Policy at Georgia Institute of Technology. He earned a master's degree in Technology Management for Innovation from University of Tokyo Japan in 2013, and B.S. degrees in Electrical Engineering and Management of Technology from the Seoul National University, Korea in 2011.

Yin Li is a PhD Candidate at School of Public Policy, Georgia Institute of Technology.

ORCiDs:
Alan Porter: https://orcid.org/0000-0002-4520-6518
Jon Garner: https://orcid.org/0000-0002-1403-3132

---

[1] Program in Science, Technology & Innovation Policy (STIP), Georgia Institute of Technology, Atlanta GA, USA 30332.
[2] Search Technology, Inc., Norcross GA, USA 30092.

Nils Newman: https://orcid.org/0000-0003-2717-1267
Stephen Carley: https://orcid.org/0000-0003-3859-1857
Jan Youtie: https://orcid.org/0000-0002-9759-5252
Seokbeom Kwon: http://orcid.org/0000-0003-4074-9342
Yin Li: https://orcid.org/0000-0002-2699-9938

National Nanotechnology Research Prominence

ABSTRACT:
A new bibliometric technique enables one to distinguish high emergence topical content. This technique can be applied to sets of research publication abstracts reflecting a given technical domain (here, nanotechnology) to score cutting edge research terms. The resulting high emergence terms warrant special consideration in setting R&D priorities. The researchers (individuals, organizations, or countries) whose publications address those emergent terms heavily deserve consideration as possible leaders in that technical domain.
This paper studies nanotechnology research publications using the new emergence scoring in conjunction with established bibliometric publication and citation measures. Findings challenge U.S. superiority in cutting edge nanotechnology research. China shows strongest at addressing emergent nanotechnology topics, followed by the U.S., South Korea, India, and, surprisingly, Iran.

Keywords
emerging technology, technical emergence indicators, bibliometrics, nanotechnology

1. Introduction

Understanding how scientific advances translate to technological innovation continues to be a major concern for both Management Of Technology and for Science, Technology & Innovation (**ST&I)** policy (Wagner and Popper 2003).

For technology management, knowledge of the status and prospects of developmental progress in a target domain is vital to inform R&D priorities. It is also essential in anticipating the timing and best prospects for commercial innovation deriving from R&D advances. Indicators of technological emergence could provide vital information to generate technology roadmaps (c.f., Phaal et al. 2004) and contribute to technology forecasting (c.f., Roper et al. 2011).

Concerning ST&I policy, ramifications of technological development abound for national economic competitiveness and national security. Tracking the progression of target science and technology topics can help avoid technological surprises (Reardon 2014). The nanotechnology case illustrated in this paper focuses on national level indicators of degree of engagement of research frontiers.

Our 'Co-Lab' Program,[3] together with various colleagues, continues to pursue 'tech mining' -- i.e., text analyses of ST&I information (Porter and Cunningham 2005). Such empirical approaches have been applied to track emerging technologies in various studies (c.f., Guo et al. 2012; Zhang et al. 2013; Zhang et al. 2014; Huang et al. 2014; Geum et al., 2015). We also recognize the value of combining expert knowledge with such data-based approaches (Huang et al., 2012).

Over the past seven years, several of us have been pursuing development of technical emergence indicators. These reflect an extension of the tech mining approach to key on changes in topical emphases of research publication or patent abstract compilations for a target domain under study. The following section elaborates on this development.

In this paper we pursue an enriched bibliometric approach to help track emerging technologies for the case of nanotechnology ('nano') -- a general purpose technology of prominent policy interest to multiple economies (Youtie, Iacopetta, and Graham 2008; Graham and Iacopetta 2014). Nano has drawn attention as an emerging science/technology domain with earth-changing prospects since the 1980's (Drexler 1986). Many efforts have contributed to tracking nanotechnology development (c.f., Walsh 2004). Nano is proclaimed as a U.S. priority dating from the National Nanotechnology Initiative (**NNI**) of 2000 through the present (see www.nano.gov for the 2016 NNI Strategic Plan). We (Porter and Cunningham 1995; Porter et al. 2008; Arora et al. 2012; Arora et al. 2014) and others (Grieneisen and Zhang 2011; Zitt and Bassecoulard 2006; Mogoutov and Kahane 2007; Huang, Notten and Rasters 2010) have used bibliometrics to track evolution of the 'nano' research domain. Further efforts have sought to track nano subfields (Subramanian et al. 2010; Ma et al. 2015; Li, Porter and Wang 2017); others have pursued application/domain advances (Cunningham and Porter 2011; Shapira and Youtie 2008; Shapira and Wang 2010). Youtie et al. (2008) generated national nano comparisons that provide somewhat of a baseline for the present analyses.

We generate emergence indicators for nano to spotlight 'hot' research *topics* within the domain. We then tally which *players* (individuals, organizations, or countries) most actively address those emergent topics. Publication activity indicates extent of participation in the research domain; citations to those publications indicate research influence (albeit imperfectly). We then combine these results – tech *emergence scores* in conjunction with *publication counts* and *citations accrued* -- to provide enriched perspectives on the cutting edge players in nanotechnology. Here, we key on national level comparisons that pose intriguing results with potential policy implications. We also anticipate applications of the emergence scores in R&D management and innovation management.

2. Technical Emergence Indicators

2.1 The Emergence Indicators Model

---

[3] A collaboration of the Georgia Tech Program in Science, Technology & Innovation Policy (STIP), the Knowledge Management and Data Analysis Lab of Beijing Institute of Technology, and the Manchester Institute of Innovation Research (MIoIR) of the University of Manchester, along with partner organizations such as Search Technology, Inc.

The Georgia Tech Technology Policy and Assessment Center,[4] generated national 'high tech' indicators for twenty years, with U.S. National Science Foundation support. Those entailed gathering of secondary analyses in the form of various time series pertaining to technological development. [They did not focus on specific technological domains as we do here.] A particularly telling result was to note China's overtaking of the U.S. in several guises about a decade ago (Porter et al. 2009). Also, as mentioned, Youtie et al. (2008) have used publication and citation data to compare countries' nano R&D endeavours. Technical emergence indicators provide a new dimension to help gauge frontier R&D activity.

We note various contributions to tracking technological emergence. In the broader sense, this relates to technology forecasting (c.f., Martino 2003; Roper et al. 2011), future-oriented technology analysis (Coates et al. 2001; Technology Futures Analysis Methods Working Group 2004; http://foresight.jrc.ec.europa.eu/), and foresight (Irvine and Martin 1984; Martin 2010) scholarship. Further work in this broad domain addresses "emerging technologies" as such (Daim et al. 2006; Bengisu and Nekhili 2006; Cozzens et al 2010; Choi 2016; Carbonell et al., 2017).

The roots of our tech emergence indicators trace back to the U.S. Intelligence Advanced Research Projects Activity (IARPA) Foresight and Understanding from Scientific Exposition (**FUSE**) Program that funded development of emergence indicators.[5] Newman and Porter participated in conceptualizing bases for emergence with the SRI group working in the FUSE program (Alexander et al. 2012). Suominen and Newman (2017) explore conceptual foundations and alternative ways to consider tech emergence, drawing on FUSE and on Rotolo et al. (2015).

To model and measure technical emergence, we follow Rotolo et al. (*31*), operationalizing four criteria – term *novelty, persistence*, and *growth*, plus research *community* formation. The next sub-section addresses how we operationalize those four criteria as a software function generates emergence scores for each of a list of candidate terms extracted from the set of records under study.[6]

2.2 Calculating Emergence Indicators and Scores

Several of us have been working to devise practical technological emergence indicators. Search Technology (2016) set out an algorithmic process to generate emergence indicators using R&D publication or patent abstract record sets. Carley et al. (to appear – 'a') present computational explorations leading to a basic algorithm with available variations. The data model uses a 3-year base period (e.g., that paper reported on 2001-2003 in the case of Dye-Sensitized Solar Cells -- DSSCs) contrasted to a 7-year active period (e.g., 2004-2010). They validated the emergent terms in the sense that such

---

[4] The active participants in TPAC coincide heavily with those in the Program for Science, Technology & Innovation Policy (STIP).

[5] www.iarpa.gov/index.php/research-programs/fuse.

[6] Analyses performed with *VantagePoint* desktop software (www.theVantagePoint.com); the routine is incorporated as an analytical option in the software (essentially, a script to facilitate the calculations).

terms tend to be more actively pursued in a subsequent time period (e.g., 2011-2013) than are other terms used in the R&D abstracts in the active period. Carley et al. (to appear – 'b') then investigated the 'staying power' – i.e., to what extent high emergence terms, and, moreover, the authors and research organizations using them, evidence stable behavior over time. They compiled abstract records for an extended period of DSSC research (1991-2014) and examined fifteen sequential 10-year periods (1991-2000; 1992-2001; 1993-2002; etc.). The emergence indicators showed reasonable persistence over successive periods

The emergence indicators algorithm, enabled in a scripted computing routine, distinguishes high emergence terms using a 10-year dataset to address the four criteria noted. Again, the first 3 years provide a base period; the last 7 years constitute an active period in which to gauge growth in term usage. *Novelty* reflects in a requirement that an emergent term not appear frequently in the base period records. *Persistence* translates into thresholds of a minimum number of years (3) and records (7) in which the term must appear. *Growth* translates into a combination of three trend components that distinguish terms that show accelerating research activity recently (detailed in Carley et al. to appear – a; Garner et al. 2017). *Community* thresholds are set to assure that the term is not just used by a very limited research group.

Terms included in the base emergence analyses are single word or multi-word noun phrases drawn from the Web of Science (**WoS**) records' abstract fields, using *VantagePoint's* Natural Language Processing (**NLP**) that is tailored for scientific discourse. Those terms are pruned by removing those appearing only once, and screening out a few very high frequency and noise terms. One could opt for alternative terms (and the computational script readily provides for this) – e.g., keywords are provided by various databases as well and these can be combined with NLP phrases if desired.

The computational process to generate the emergence indicators entails two stages: 1) distinguishing *emergent terms* (topical words or phrases showing accelerated usage in recent years), and 2) *players* (researchers, research organizations, countries) whose publication or patent records are particularly rich in the inclusion of those emergent terms.

Through a series of case sensitivity analyses, we have settled upon base model values for the emergence calculations. We have evolved a threshold value to distinguish topical terms showing 'high emergence' (i.e., terms increasingly prevalent in recent publication abstract records in the dataset, not 'brand new' terms, as we require persistence too). The threshold is 1.77 (the square root – **SQRT** -- of Pi) – inspection of the behavior of various thresholds for some six different technical domains suggested that a threshold of somewhere in the 1.5 to 2 range performed well.[7]

We have explored alternative ways to tally players' emergent term usage, settling on a SQRT transformation to reduce over-emphasis on a few exceptional terms (Garner et al. 2017). [To illustrate such over-emphasis, in analyzing 'big data' records from WoS for 2003-2013, the term 'big data' scored 46.7, while no other term scored above 8.6.] We then tally the usage of such high emergence terms to score the degree to which authors, research organizations, or countries' research publications evidence cutting-edge content. For example, in considering country, the script requires at least 45% of records and 10 total records to have an emergent term.

[7] The "SQRT of Pi" choice was whimsical, not mathematical; it split the range of well-performing cutoffs.

We are using two alternative ways to tally players' usage of emergent terms. A 'total' measure – the sum of SQRT (Emergence Scores) above threshold (1.77) -- counts each time a high emergence term appears in a distinct record authored by a researcher from that player (but not tallying multiple occurrences within records). Again, we consider three levels of players – individuals (authors or inventors), R&D organizations, and countries (based on affiliation location); in this paper our main interest is in country level comparisons of nano R&D. We give full weight to a country appearing as an author affiliation. So, if a particular article has one Chinese and five American co-authors, it credits equally to each country's emergence score. Our second way to tally emergent activity is a 'normalized' score that divides the total score by the SQRT (number of records) (Garner et al. 2017). The normalized score seems preferable for the national comparisons here.

3. Data and Analyses

3.1 The Data

The present analyses address *nanotechnology* research publication abstracts retrieved from the Web of Science ('WoS') Core Collection. The search was conducted in May, 2016, for publications dating from 1991. Comparison of various nano search strategies finds our approach to be central and robust (Huang, Notten and Rasters 2010). Our strategy is based closely on that described in previous publications (Porter, Youtie and Shapira 2008; Arora et al. 2012; Arora et al. 2014). The gist of it is to separately search and retrieve abstract records based on terms appearing in the WoS 'topic' field (title, abstract, keywords) in 8 modules ('nano*' and 7 others – see Appendix). Occurrence of some terms suffices to designate a record as nano-related. Other terms have additional contingencies (e.g., requiring co-occurrence with other terms indicative of molecular scale focus) to include a record as nano. After download, additional steps are taken to exclude certain records from the nano* set (see Appendix).

We divided the nano WoS abstract records into three 10-year periods (with some overlap):

1) **1991-2000 – 184, 077 records**, reflecting early nano research growth (WoS data are somewhat discontinuous to extend back further, and this period covers up to the instigation of the U.S. NNI in 2000)
2) **1998-2007 – 465,454 records**; we chose this period as it gives an active period (2001-2007) that starts just after the NNI initiation
3) **2006-2015 – 1,178,694 records**, the most recent data available with reasonably complete WoS data coverage as of our retrieval of the ~1.6 million overall records in May, 2016.

3.2 Emergent Term Analyses

Our analyses entail a few distinct elements of note:

1) Data cleaning – after compiling the nano WoS dataset and separating it into the three time periods noted, we did selective cleaning operations – i.e., consolidation of term variations and of organizational name variations. Most importantly we tried various manipulations to determine a suitable set of topical terms for the emergent term indicator development.
2) Bibliometric calculations – we tallied the number of publications and citations (using the WoS 'TC' field – Times Cited). These operations are straightforward, so are not detailed here. We gave full credit among co-authors (i.e., no attempt to apportion publication or citation credits).
3) Emergence scoring, at two stages:

a) Terms – we generated an emergence score for each topical term included (treated in this subsection)
b) Players – we tally the degree to which countries and R&D organizations' abstract records contain high emergence terms (treated in the next sub-section).

We compared nano results described here against those when using alternative term handling. Recall that our basic approach analyses Title and Abstract NLP phrases with minimal cleaning. The alternative here entailed combining title and abstract NLP phrases, and application of five 'stopword' thesauri to consolidate terms. Sensitivities vary according to granularity. Individual papers flagged as rich in emergent terms vary most – by roughly half – but those are not of prime interest here. The 'Top 100' emergent terms shift somewhat – e.g., changing the input term set replaced 13 of those terms. The Top 100 research organizations change by only 3 for the recent period (20 for the earliest period). And, of main concern here, the Top 100 countries change by 1 for the recent period (2 for the middle period and 0 for the earliest period). So for the present country comparisons, results are not very sensitive to term cleaning and consolidation, but for term analyses, they are more sensitive, but not unduly so.

The 'Top 10' *high emergence terms* for each period are shown in Table 1. The first two terms are obviously variants of 'atomic force microscopy' – reflecting sensitivity to data cleaning (recall previous paragraph). The 'hot topics' (concentrations of emergent terms) vary by period, with versions of 'atomic force microscopy' prominent in the early period; 'nanoparticles' and other 'nanostructures,' lead in the middle period; and 'graphene' variations most notable in the recent period. Inclusion of a cancer term at #10 in the latest period hints at the advent of nano-biomedical developments becoming prominent in nano research publication. In general, the progression of high ESc terms across the periods seems sensible. These are not the "hottest, new" terms one could anticipate; this reflects the nature of the criteria for emergence in our model – namely requirements for term novelty, but also growth and persistence, as well as research community in place. Those criteria require time to build up.

[The Supplemental Materials provide a table of all the emergent terms with their emergence scores, containing 178 terms exceeding the 1.77 threshold for 1991-2000; 438 terms for 1998-2007; and 763 for 2006-2015.]

Table 1. Top 10 High Emergence Nanotechnology Terms by Time Period

1991-2000 High Emergence Terms	**Emergence Scores**	**1998-2007** Emergence Scores	**Emergence Scores**	**2006-2015** High Emergence Terms	**Emergence Scores**
atomic force microscopy	23.68	nanoparticles	63.10	graphene	55.41
atomic force microscopy (AFM)	20.53	nanowires	20.04	g(-1)	48.15
nanoparticles	19.81	gold nanoparticles	19.97	great potential	42.13
carbon nanotubes	17.85	field emission	18.92	graphene oxide (GO)	32.28
x-ray diffraction (XRD)	15.48	CNTs	17.89	synergistic effect	31.97
mechanical properties	13.10	carbon nanotubes (CNTs)	17.85	electrochemical performance	31.59
particle size	12.37	detection limit	17.52	visible light irradiation	24.93
nanotubes	11.90	nanocomposites	16.66	cancer cells	24.85
XRD	10.25	Catalyst	15.64	Specific capacitance	22.59
Quantum dots	9.87	Nanostructures	14.28	Graphene oxide	21.99

3.3 Emergent Player Analyses

Our prevailing interest here is not in the emergent terms per se, but rather to use them to identify *research players* whose publications evidence greatest attention to cutting edge topics (i.e., high emergence score terms). Experimentation led us to adopt two measures to compare players (i.e., individuals, organizations, and/or countries):

- **Total** = Summation of SQRT (Emergence Score = **ESc**) above the chosen threshold [SQRT (Pi) = 1.77], counting each time a term was used in a distinct record; but not crediting multiple occurrences within a record
  ∑ (SQRT(ESc) x # records for that ESc term) – summation over all high ESc terms

- **Normalized** = Summation (as in '1') divided by the SQRT of the number of records..
  [∑(SQRT(ESc) x # records for that ESc term) – summation over all high ESc terms] / SQRT (# of records of that player)

One, or the other, measure (or both, to offer two-dimensional comparisons) could be most suitable for any given analysis. For this study, comparing countries, we favor the normalized ESc tally. To illustrate the calculation, glance ahead to Table 2 to consider Iran's 49 nano papers as indexed by WoS for 1991-2000. The computer script looks for any of the 178 high emergence terms for that period (for all countries; the Top 10 of those 178 appear in Table 1) in Iran's 49 abstract records. It tallies up each time a term is used in a distinct record. What we find is one term, 'XRD' [ESc of 10.245], appearing in 5 Iranian authored (or co-authored) papers, plus 4 other high emergence terms in 1 record each, with EScs of 12.366, 5.5.96, 4.836, and 1.966. So, taking the SQRT of each ESc, multiplying by the number of records containing that term (5 for XRD; 1 for each of the others), and summing, we get a Total ESc for Iran for the early period of 25.49. Dividing that by the SQRT (49) = 7, gives a Normalized ESc of 3.64 (rounded to 4 in Table 2).

Table 2. National Nanotechnology R&D Profiles: Publications, Citations & Emergence Intensity

Country	**# Records**			**Times Cited -- Mean**			**Emergence Scores**		
	1991-2000	**1998-2007**	**2006-2015**	**1991-2000**	**1998-2007**	**2006-2015**	**1991-2000**	**1998-2007**	**2006-2015**
China	9760	64046	287807	19	25	13	121	598	1201
USA	43706	123205	248979	54	48	22	177	442	485
So. Korea	3709	23044	70985	21	25	13	76	268	416
India	2376	12307	66054	25	27	10	63	227	396
Iran	49	1389	30963	11	22	8	4	88	331

4. Results

4.1 Emergent R&D Organizations

Table 3 shows the 'Top 20' R&D organizations for the early and late periods, based on emergence scores. Note that high emergence correlates, but imperfectly so, with high publication activity. Comparing these time periods that are 15 years apart, both the reappearances and the changes are interesting. Of the 20 highest nano emergence organizations in the early period, 6 appear in the late period as well. Only one non-Chinese organization (National University of Singapore) shows in both lists. Indeed, the late period list shows extreme Chinese dominance. Besides the National University of Singapore, only Islamic Azad University in Iran (a private university system) makes the list.

Table 3. Top R&D Organizations based on Nanotechnology Publication Emergence Scores

	1991-2000			2006-2015		
	Organization	# Records	ESc	Organization	# Records	ESc
1	Chinese Acad Sci	2671	71.2	Chinese Acad Sci	46712	436.9
2	Natl Univ Singapore	652	42.5	Nanyang Technol Univ	8813	218.0
3	Nanyang Technol Univ	305	40.4	Univ Chinese Acad Sci	3587	211.2
4	Osaka Univ	1374	39.8	Zhejiang Univ	9145	211.0
5	Russian Acad Sci	3232	39.1	Tsinghua Univ	9158	202.6
6	Univ Tokyo	1847	37.9	Fudan Univ	6705	201.0
7	Tohoku Univ	1488	36.5	Hunan Univ	3477	195.6
8	CNRS	1846	34.8	Cent S Univ	3406	194.8
9	Univ Sci & Technol China	622	34.3	Univ Sci & Technol China	7339	193.9
10	Ecole Polytech Fed Lausanne	660	34.2	Natl Univ Singapore	8484	186.8
11	Nanjing Univ	754	33.2	Shanghai Jiao Tong Univ	6896	186.7
12	Penn State Univ	828	32.8	Soochow Univ	3963	182.3
13	CSIC	777	32.7	Jilin Univ	7998	176.3
14	Univ Paris 06	764	31.6	Nanjing Univ	7615	175.9
15	Univ Calif Berkeley	1285	31.5	Shandong Univ	5066	174.3
16	CNR	1020	30.9	Nankai Univ	3955	172.0
17	Fac Univ Notre Dame Paix	210	30.7	Wuhan Univ	4172	171.8
18	Kyoto Univ	1060	30.1	Tianjin Univ	4763	171.5
19	Harvard Univ	706	29.6	Jiangsu Univ	2190	167.7
20	Jilin Univ	557	28.3	Islamic Azad Univ	6415	165.4

Organizational emergence scores can also provide pointers to interesting research concentrations. To illustrate, Figure 1 profiles leading Iranian universities based on nano emergence scores for the latest period (but note that only one makes the 'Top 20' in Table 3). The University of Tehran Medical School, with the highest ESc, is spotlighted. We see 1105 nano-related WoS publications with an affiliated author. Truncated on the left is a Title field; if a topic looked to be of special interest, one could 'double-click' it to open that abstract record. Just to the right of the organization name is the Emergence score; note that the Medical School is second only to Islamic Asad University in the Iranian data.

Four detail windows appear on the right side. 'Emergence Scores' shows the prevalence of high emergence terms used in the University of Tehran Medical Science 1106 paper abstract records. 'NPs' (nanoparticles) and nanoparticles are both top 5, suggesting an emphasis in their research. More notably, we see seven biomedical phrases in these top 15 terms (cytotoxicity, toxicity, antibacterial activity, cancer cells, cellular uptake, vivo studies, stem cells), pointing to a likely emphasis in their 'Medical Science' nano-research. The window to the right shows leading authors on those papers; one could go on to generate a network map to help identify research groups and collaborations among them. Another window shows the 'Countries' of co-authors; interestingly, the U.S. leads in collaborating with Iran. 'Times Cited' finds 11 of the 1106 papers cited more than 100 times – potentially worth exploring in depth to understand 'who's doing what?' [We note, just to illustrate, that the top cited paper (358 times) is a multinational collaboration, whereas the second most cited (342 times) is Iranian authored only.]

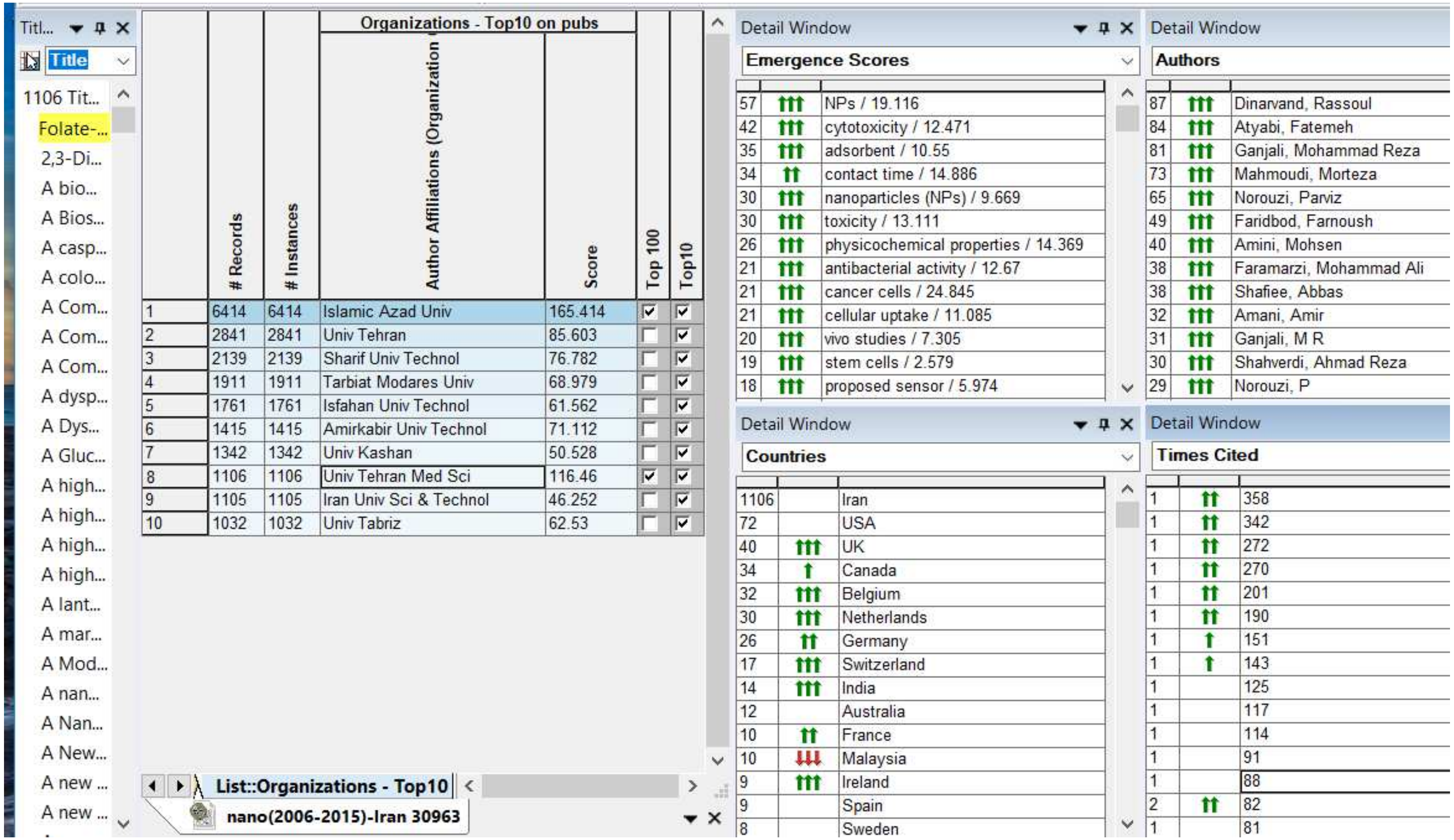

Figure 1. Breaking Out Information on Leading Iranian Nano R&D Organizations

4.2 Country Comparisons

A growing disparity between China and the U.S caught our attention in exploring nano emergence results for leading research organizations. We noted:

- For 1991-2000, of the Top 100 research organizations, the U.S. led with 27; China showed 14.
- For 1998-2007, of the Top 100, 42 are Chinese; 22 are U.S.
- For 2006-2015, of the Top 100 research organizations, China leads with 73, the U.S. shows 3.[8]

Organization-level emergence scores thus suggest that China is outpacing the US on the nano research frontiers. Publication-based differences are less pronounced than these emergence score based comparisons. For 2006-2015, the Top 100 organizations by articles indexed in WoS include 34 Chinese and 22 American. And, on average, American nano papers (2006-2015) remain the most heavily cited (Table 2).

Table 2 tallies publications, times cited, and emergence scores for the top five countries in the latest period. Their order based on emergence scores is: 1) China, 2) USA, 3) South Korea, 4) India, 5) Iran [plus 6) Singapore -- not shown]. Ordering is based on emergence scores for 2006-2015 (last column). In contrast, the top 'nano emergence' countries in 1991-2000 were heavily weighted toward the 'usual suspects' – see Supplemental materials – Table on More Countries (all 3 periods). Namely, the top six countries based on nano emergence scoring in the earliest period were: 1) USA, 2) Japan, 3) China, 4) Germany, 5) France, and 6) UK.

The 'Top 10' order based on publication activity in the latest period is largely similar to that tracking back to the earliest period: 1) China, 2) USA, 3) Japan, 4) Germany, 5) South Korea, 6) India, 7) France, 8) UK, 9) Italy, and 10) Taiwan (see Supplemental Materials for additional countries). Were we to go by research publications in the same earliest period, the top six hold steady (with China at fifth) for Period 1. So, national nano publication levels have changed much less over this span than have the countries dominating high nano emergence topical research.

Another surprise to us is the prominent presence of Iran. As per Table 2, Iran was 'nowhere' in the 1991-2000 nano domain based on publications (49) or emergence scores (4). By 2006-2015, Iran shows fifth globally on emergence scores, 13th on total nano publications indexed by WoS, and quite respectably on citations accrued by their publications.

Research publications provide a basic indicator of topical engagement. Obviously, different images show if one were to normalize by population (e.g., Singapore would stand forth here). Times Cited provide a complementary indicator reflecting research influence (with various issues, including differing field citation norms and increasing citation rates over time). Furthermore, citations are notoriously skewed [e.g., 15,306 cites to one 2007 graphene paper (Geim and Novoselov 2007) leads our 3d period], time-sensitive (accruing cites over time, so avoid direct comparison across our three time periods), and malleable (e.g., one could remove self-citations). That said, our mean [and median (not shown)] Times Cited values for the five countries are quite consistent. The U.S. continues to lead on this measure, but the gap is narrowing.

The Emergence indicator adds perspective to publications and citations by

---

[8] Also in the Top 100 was 'Northeastern University,' but this confounds publications of the U.S. and Chinese universities sharing this name in WoS.

focusing on recently accelerating topics research (Table 2 and Figure 2). Figure 2 shows striking gains by other countries, led by China, vis-à-vis the U.S. in frontier research on nano, based on this indicator.

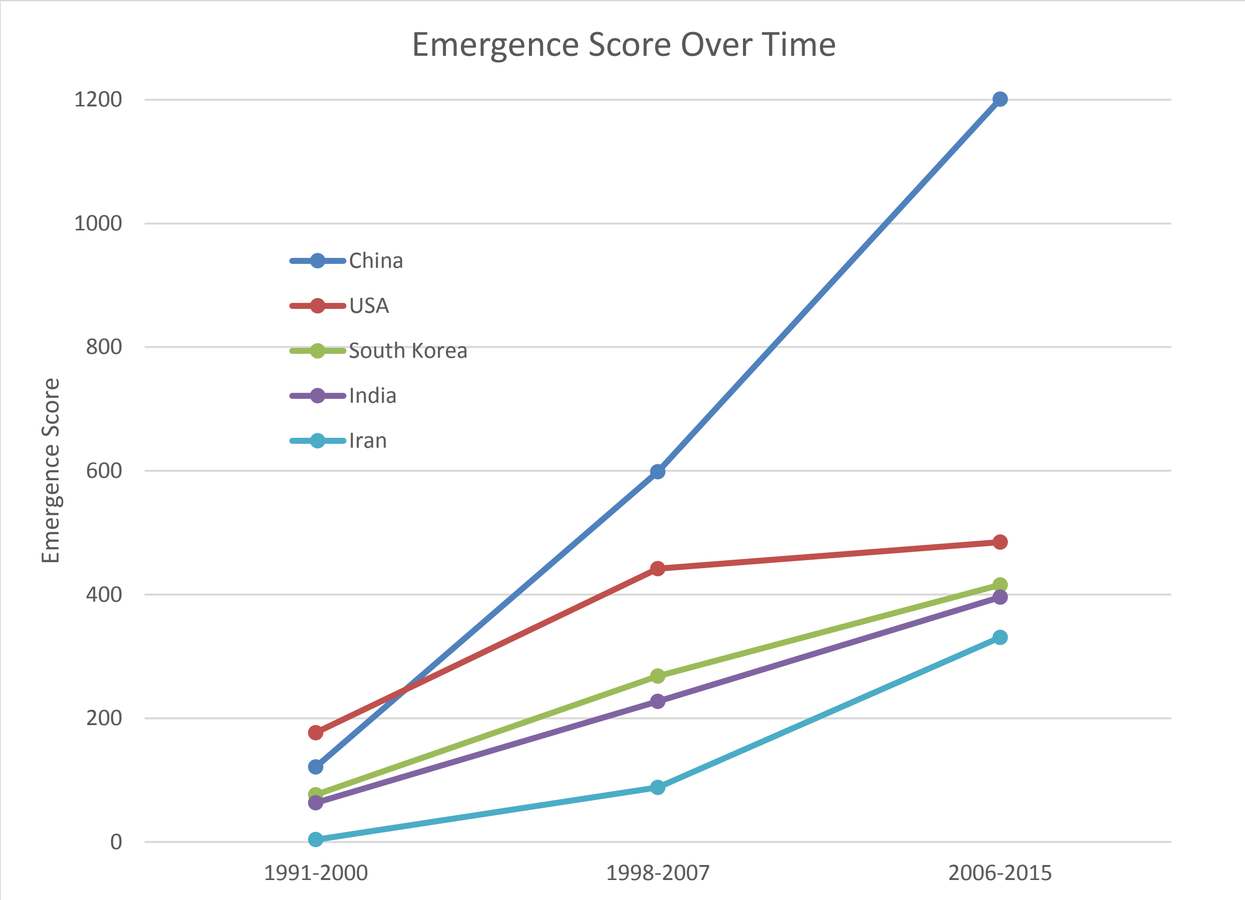

Figure 2. National Nano Emergence Scores by Time Period

5. Conclusions and Discussion

We focus on two aspects of this research: 1) a new bibliometric or 'tech mining' emergence scoring capability (tool) to help distinguish cutting edge research content and the researchers most heavily engaged with those topics, and 2) exploration of national nanotechnology research prominence, based on application of the emergence scoring. Here, discussion of these two aspects is necessarily intertwined in the case analyses. We hope this discussion identifies limitations to be overcome, but moreover illuminates opportunities to develop the emergence indicators further and apply them in support of diverse technology management and ST&I policy purposes. We also hope that the international nanotechnology comparisons serve to enrich policy perspectives.

Present results offer an enriched set of reproducible 'tech emergence indicators,' focusing on national leadership in nanotechnology. The emergence indicator set reflects *the degree to which a player's abstract records discuss topics that show accelerating attention by the research community in the last few years.* Tests on three other emerging technologies (Garner et al. 2017) serve to help validate the emergence indicators in that they show that high emergence terms receive disproportionate research attention in a follow-on 3-year period. Current results' national and organizational emergence findings add intelligence value to standard publication and citation metrics. Such tools can help identify challenges to nations' economic trajectories and security interests. Further development and assessment of emergence scoring is in order to understand their meaning and implications.

By generating high emergence term sets, one has an appealing filter to pursue topical comparisons. We inspected terms particular to the countries of special interest, the U.S. and China. Eyeballing terms unique to the U.S. (vs. China) for Period 3 did suggest greater U.S. biomedical emphases as evidenced in terms such as mouse model, clinical translation, and antimicrobial activity making a list of emergent terms based on U.S.-only authored research abstracts; those terms were not prominent based on a set of Chinese-only authored abstracts. [That is not to say there is not strong Chinese attention to such topics, only that they did not show the emergence patterns.]

Conversely, it is interesting to compare the number of emergent terms unique to one or another entity. Comparing the U.S. and China:

	# Unique to the U.S.	# Unique to China
➢ Period 1	101	423
➢ Period 2	80	1004
➢ Period 3:	26	1097

Another hypothesis that we explored is whether the U.S. might be so far ahead that its topical terms don't show forth as emergent within the global nano research dataset. On first inspection, that does not seem to be the case. Comparing U.S.-only records vs. non-US, or comparing U.S. vs. full sets, finds them to be relatively similar. [We ran various 'list comparisons' to explore term differences.]

The high emergence scoring terms offer a rich resource for further probing. For instance, given the national level explorations advanced here, one might want to compare topical emphases among select countries. Given interest in China, Iran, and the U.S., such investigation could point to researchers sharing interests in certain cutting edge technological topics. That, in turn, might be pursued via invitations to explore shared interests at workshops, etc. [Or, should national security issues be entangled, perhaps not.]

Emergent terms are those with accelerating attention in recent years. One should be somewhat cautious with respect to the concerns noted about term cleaning – i.e., how term variations are treated affects what topics are included as emergent. A topic for future research is to experiment further with term consolidation variations to provide more informative term sets to analyze for emergence. That said, the emergent terms provide a nice shopping list for consideration of potential dissertation topics, R&D prioritization, etc. A ready extension is to sift out the abstract records containing concentrations of the emergent terms of special interest for review, likely leading to further information retrieval of the most promising full papers for study.

These results question U.S. leadership in nano research. Certainly, the emergence scores don't tell the complete story, but they challenge the assumption that U.S. leadership is a given. Publication data indicate that Chinese nano-research activity exceeds American as of the 2006-2015 period. Moreover, research organization prominence on the basis of emergence score is shocking (73 Chinese vs. 3 American). Might a would-be nano graduate student seeking to do research at the frontier find keener prospects at a Chinese university? We're not prepared to assert that, but potential collaboration could well be advantageous.

Why the U.S. is no longer the leader based on emergence is a challenging question. Perhaps, the U.S., Japan, Germany, France, and the UK find richer prospects in pursuing nano application more than fundamental nano-research? Consistent with such an interpretation, recent period U.S. nano emergent terms show strong biomedical emphases.

The fact that Iran has risen so dramatically based on publication rate and that those publications assert a strong role in emerging nano topics is startling. We don't have background in studying Iran, but a visiting PhD student from Iran found the results reasonable. We also note interesting research on technological development in Iran (Ghazinoory and Soofi 2017; Soofi 2017). Monitoring the R&D position of particular nations may have a place in national security interests. Emergence scoring, together with publication and citation measures, could aid in exploring within-country concentrations of emergent topics within a domain such as nano, directing attention to the organizations and researchers performing that frontier R&D. Certainly, one would want to complement such empirical results with expert opinions by those knowledgeable about the technology and attendant delivery system aspects.

We wish to facilitate others' exploration of the emergence indicators introduced here and elsewhere (Carley et al. in press 'a' and 'b'; Garner et al. 2017). The emergence algorithm is open. The script incorporated in VantagePoint software eases its use, but is not required. Our explorations show promise in use with research publication or patent abstract record sets.

We recognize that "emergence" is a term in wide use with varied meanings. Even "tech emergence" as used in scholarly literatures has multiple nuances (Li et al. under review). Most critically, we point out that the emergence indicators and emergence scoring presented here pertain to R&D. We are distinguishing topical emphases in that context, so these tools do not pertain to technological innovation in the sense of pointing to commercialization on the brink.

We also recognize that an easy association with "tech emergence" is the notion of rapid identification of "the newest" topics just as they first attract attention. Our emergence indicators do not do that. Rather, our model of tech emergence rests on the four criteria noted -- *novelty, persistence*, *growth*, and *community.* To demonstrate growth requires some time; persistence and community, even longer. So, as noted, our basic algorithm addresses a 3-year base and 7-year active period. The algorithm could be adapted to shorter intervals, say "months." But it does not appear relevant to instant monitoring of breakthrough happenings.

At the *Portland International Conference on Management of Engineering and Technology (PICMET)* Conference in 2017, a keynote presentation and four sessions comprised papers on 'Indicators of Technical Emergence,' followed by a Panel session exploring next steps in developing indicators of emergence.[9] We invite readers attracted to emergence indicator potential to track and join in further development and assessment to make these into effective technology management and ST&I policy tools.

Acknowledgements

This material is based upon work supported by the National Science Foundation under EAGER Award #: 1645237 for a project, "Using the ORCID ID and Emergence Scoring to Study Frontier Researchers." Any opinions, findings, and conclusions or

[9] A special issue on tech emergence indicators deriving from these and additional papers is planned. It is to be edited by Yi Zhang and colleagues, in *Technological Forecasting & Social Change.*

recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Alexander, J.J., Chase, J., Newman, N., Porter, A.L., & Roessner, J.D. (2012). Emergence as a conceptual framework for understanding scientific and technological progress. *PICMET (Portland International Conference on Management of Engineering and Technology)*, Vancouver (2012). https://www.researchgate.net/profile/Jeffrey_Alexander/publication/248391856_Emergence_as_a_Conceptual_Framework_for_Understanding_Scientific_and_Technological_Progress/links/550847050cf26ff55f80c7b9.pdf

Arora, S.K., Porter, A.L., Youtie, J. & Shapira, P. (2012). Capturing new developments in an emerging technology: an updated search strategy for identifying nanotechnology research outputs. *Scientometrics,* **95** (1), 351–370 (2012). DOI:10.1007/s11192-012-0903-6.

Arora, S.K., Youtie, J., Carley, S., Porter, A.L., & Shapira, P. (2014). Measuring the development of a common scientific lexicon in nanotechnology. *Journal of Nanoparticle Research,* **16**, 2194 (2014). DOI: 10.1007/s11051-013-2194-0.

Bengisu, M., & Nekhili, R. (2006). Forecasting emerging technologies with the aid of science and technology databases. *Technological Forecasting and Social Change*, 73 (7), 835-844.

Carbonell, J., Sanchez-Esguevillas, A. and Carro, B. (2017). Easing the assessment of emerging technologies in technology observatories. Findings about patterns of dissemination of emerging technologies on the internet, *Technology Analysis and Strategic Management*, doi.org/10.1080/09537325.2017.1337886.

Carley, S.F., Newman, N.C., Porter, A.L. & Garner, J. (to appear - a). An indicator of technical emergence, *Scientometrics*.

Carley, S.F., Newman, N.C., Porter, A.L., & Garner, J. (to appear – b). A measure of staying power: Is the persistence of emergent concepts more significantly influenced by technical domain or scale? *Scientometrics*.

Choi, H. (2017). Ready-steady-go for emerging technologies in post catch-up countries: a longitudinal network analysis of nanotech in Korea, *Technology Analysis and Strategic Management*, 29 (8), 946-959.

Coates, V., Faroque, M., Klavins, R., Lapid, K., Linstone, H.A., Pistorius, C., and Porter, A.L. (2001). On the future of technological forecasting, *Technological Forecasting and Social Change*, 67 (1), 1-17.

Cozzens, S., Gatchair, S., Kim, K-S., Ordonez, G., Porter, A.L., Lee, H.J., Kang, J. (2010). Emerging technologies: Quantitative identification and measurement, *Technology Analysis and Strategic Management*, 22 (3), 361-376.

Cunningham, S., & Porter, A.L. (2011). Bibliometric discovery of innovation and commercialization pathways in nanotechnology. *Proceedings of the Portland International Conference on Management of Engineering and Technology (PICMET)*, Portland OR. http://ieeexplore.ieee.org/xpls/abs_all. jsp?arnumber=6017795.

Daim, T. U., Rueda, G., Martin, H., & Gerdsri, P. (2006). Forecasting emerging technologies: Use of bibliometrics and patent analysis. *Technological Forecasting and Social Change*, 73 (8), 981-1012.

Drexler, K.E. 1986. *Engines of Creation: The Coming Era of Nanotechnology*. Garden City, NY: Anchor Press/Doubleday.

Garner, J., Carley, S.F., Porter, A.L., & Newman, N.C. (2017). Technological Emergence Indicators using Emergence Scoring. *PICMET (Portland International*

Conference on Management of Engineering and Technology), Portland, OR (July, 2017).
Geim, A.K., & Novoselov, S. (2007). The rise of graphene. *Nature Materials,* **6** (3), 183-191.
Geum, Y., Lee, H., Lee, Y., & Park, Y. (2015). Development of data-driven technology roadmap considering dependency: An ARM-based technology roadmapping. *Technological Forecasting and Social Change*, **91**, 264-279.
Ghazinoory, S., & Soofi, A.S. (2017). An introduction to the special issue on science and technology in Iran, *Technological Forecasting and Social Change* 122, 104-106.
Guo, Y., Ma, T., Porter, A. L., & Huang, L. (2012). Text mining of information resources to inform Forecasting Innovation Pathways. *Technology Analysis & Strategic Management,* **24** (8), 843-861.
Graham, S.J., & Iacopetta, M. (2014). Nanotechnology and the emergence of a general purpose technology. *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, (115/116), 25-55 (2014).
Grieneisen, M.L., & Zhang, M. (2011). Nanoscience and nanotechnology: evolving definitions and growing footprint on the scientific landscape. *Small,* **7** (20), 2836–2839.
Huang, C., Notten, A., & Rasters, N. (2010). Nanoscience and technology publications and patents: a review of social science studies and search strategies. *Journal of Technology Transfer,* **36** (2), 145–172.
Huang, L., Guo, Y., Porter, A. L., Youtie, J., & Robinson, D. K. (2012). Visualising potential innovation pathways in a workshop setting: the case of nano-enabled biosensors. *Technology Analysis & Strategic Management*, **24** (5), 527-542.
Huang, L., Zhang, Y., Guo, Y., Zhu, D., & Porter, A. L. (2014). Four dimensional Science and Technology planning: A new approach based on bibliometrics and technology roadmapping. *Technological Forecasting and Social Change,* **81**, 39-48.
Irvine, J. and Martin, B.R. (1984). *Foresight in Science: Picking the Winners*, London. Pinter Publishers.
Li, M., Porter, A.L., and Suominen. (under review), Insights into relationships between disruptive technology/innovation and emerging technology: A bibliometric perspective, *Technological Forecasting and Social Change,*
Li, M., Porter, A.L., & Wang, Z.L. (2017). Evolutionary trend analysis of nanogenerator research based on a novel perspective of phased bibliographic coupling, *Nano Energy,* **34**, 93-102.
Ma, J., Porter, A.L., Aminabhavi, T. & Zhu, D. (2015). Nano-enabled drug delivery systems for brain cancer and Alzheimer's Disease: Research patterns and opportunities, *Nanomedicine: Nanotechnology, Biology and Medicine,* **11** (7), 1763-1771.
Martin, B.R. (2010). Foresight in science and technology, *Technology Analysis & Strategic Management* 7 (2), 139-168.
Martino, J (2003). A review of selected recent advances in technological forecasting. *Technological Forecasting and Social Change*, 70 (8), 719-733.
Mogoutov, A., & Kahane, B. (2007). Data search strategy for science and technology emergence: a scalable and evolutionary query for nanotechnology tracking. *Research Policy,* **36** (6), 893–903.
PCAST, (2012). *Report to the President and Congress on the Fourth Assessment of the National Nano- technology Initiative*, President's Council of Advisors on Science and Technology. Washington: Executive Office of the President (2012). http://www.whitehouse.gov/administration/eop/ostp/pcast/ docsreports.

Phaal, R., Farrukh, C. J. P., & Probert, D. R., (2004). Technology roadmapping—A planning framework for evolution and revolution, *Technological Forecasting and Social Change,* **71,** 5–26.
Porter, A.L., & Cunningham, S. (1995). Whither Nanotechnology? A Bibliometric Study. *Foresight Update*, **21**, 12-15.
Porter A.L., & Cunningham S.W. (2005). *Tech mining: exploiting new technologies for competitive advantage.* New York: Wiley.
Porter, A.L., Newman, N.C., Roessner, J.D., Johnson, D.M., & Jin, X-Y. (2009). International high tech competitiveness: does China rank #1?, *Technology Analysis and Strategic Management*, **21** (2), 173-193.
Porter, A.L., & Youtie, J. (2009). Where does nanotechnology belong in the map of science?, *Nature-Nanotechnology,* **4**, 534-536.
Porter, A.L., Youtie, J., Shapira, P. & Schoeneck, D.J. (2008). Refining search terms for nanotechnology. *Journal of Nanoparticle Research,* **10** (5), 715–728.
Reardon, S. (2014). Text-mining offers clues to success: US intelligence programme analyses language in patents and papers to identify next big technologies. *Nature,* **509** (7501), 410-411.
Roper, A.T., Cunningham, S.W., Porter, A.L., Mason, T.W., Rossini, F.A., & Banks, J. (2011). *Forecasting and Management of Technology, 2d edition,* New York: John Wiley.
Rotolo, D., Hicks, D., & Martin,B.R. (2015). What is an emerging technology? *Research Policy,* **44,** 1827-1843.
Search Technology, Inc., Patent pending – Technological Emergence Scoring and Analysis Platform, US Patent and Trademark Organization (#62420295, Nov. 10, 2016).
Shapira, P., & Wang, J. (2010). Follow the money. *Nature* **468** (7324), 627–628.
Shapira, P., & Youtie, J. (2008). Emergence of nanodistricts in the United States, *Economic Development Quarterly,* **22** (3), 187–199.
Soofi, A.S. (2017). A comparative study of Chinese and Iranian Science & Technology, and techno-industrial development policies, *Technological Forecasting & Social Change*, 122, 107-118.
Suominen, A., & Newman, N. (2017). A critical evaluation of the technological emergence concept. *PICMET (Portland International Conference on Management of Engineering and Technology)*, Portland, OR (July).
Subramanian, V., Youtie, J., Porter, A.L., & Shapira, P. (2010). Is there a shift to 'active nanostructures'? *Journal of Nanoparticle Research,* **12** (1), 1–10.
Technology Futures Analysis Methods Working Group [Alan L. Porter, Brad Ashton, Guenter Clar, Joseph F. Coates, Kerstin Cuhls, Scott W. Cunningham, Ken Ducatel, Patrick van der Duin, Luke Georghiou, Ted Gordon, Hal Linstone, Vincent Marchau, Gilda Massari, Ian Miles, Mary Mogee, Ahti Salo, Fabiana Scapolo, Ruud Smits, and Wil Thissen] (2004). Technology futures analysis: Toward integration of the field and new methods, *Technological Forecasting and Social Change,* 71, 287-303.
Wagner, C.S. & S. W. Popper, S.W. (2003). Identifying critical technologies in the United States: A review of the federal effort. *Journal of Forecasting,* **22** (2-3), 113-128.
Walsh, S. T. (2004). Roadmapping a Disruptive Technology: A Case Study - The Emerging Microsystems and Top-down Nanosystems Industry. *Technological Forecasting & Social Change*, 71, 161-185.
Youtie, J., Iacopetta, M., & Graham, S.J. (2008). Assessing the nature of nanotechnology: can we uncover an emerging general purpose technology? *The Journal of Technology Transfer,* **33** (3), 315-329.

Youtie, J., Shapira, P. & Porter, A.L. (2008). Nanotechnology publications and citations by leading countries and blocs. J*ournal of Nanoparticle Research,* 1**0**, 981–986.
Zhang, Y., Guo, Y., Wang, X., Zhu, D., & Porter, A. L. (2013). A hybrid visualisation model for technology roadmapping: bibliometrics, qualitative methodology and empirical study. *Technology Analysis & Strategic Management*, **25** (6), 707-724.
Zhang, Y., Zhou, X., Porter, A. L., & Gomila, J. M. V. (2014). How to combine term clumping and technology roadmapping for newly emerging science & technology competitive intelligence:'problem & solution' pattern based semantic TRIZ tool and case study. *Scientometrics*, **101** (2), 1375-1389; DOI: 10.1007/s11192-014-1262-2.
Zitt, M., & Bassecoulard, E. (2006). Delineating complex scientific fields by an hybrid lexical-citation method: an application to nanosciences. *Information Processing and Management,* **42** (6), 1513–1531.

Appendix

The search strategy is fully detailed in Arora et al. (2012). Key features are the use of 8 discrete search modules:

1. Nano*
2. Quantum (quantum dot or quantum well or quantum wire)
3. Self-assembly (self assembl* or self organiz* or directed assembl*)
4. Nano-related (some 25 discrete terms such as molecule* motor or C60 or graphene)
5. Microscopy and spectroscopy (some 25 discrete terms such as TEM or x-ray photoelectron)
6. Nano-pertinent (NEMS or quasicrystal or quasi-crystal or quantum size effect or quantum device)
7. Nano-pertinent (some 20 additional terms such as solgel or molecular sieve or epitax*)
8. Nano journals (some 14 particular journals such as *Nature Nanotechnology*).

Critically important, modules 1, 2, 4, and 8 stand as is. Modules 3, 5, 6, and 7 are contingent searches at either of two levels: a) requiring co-occurrence of a restricted set of terms suggestive of a molecular environment focus (e.g., monolayer) or a more inclusive set of terms (adding terms such as co-polymer).

The next step, after download, is to exclude records containing certain terms from the nano* record set. Records containing terms such as nanometer and no other nano terms were excluded. Records containing some other terms were excluded no matter what other terms were present (e.g., plankton).

Supplemental Materials

These are prepared as an MS Excel file with multiple worksheets:

1. Top nanotechnology research organizations -- # of records and ESc's
2. Top countries for each of the 3 periods -- # records and ESc's
3. All emergent terms for each of the 3 periods -- # of records containing each term and ESc's
4. All countries for each of the 3 periods – ESc's
5. Top 5 countries for each of the 3 periods – providing additional statistical data such as medians to complement means

Not in the paper:

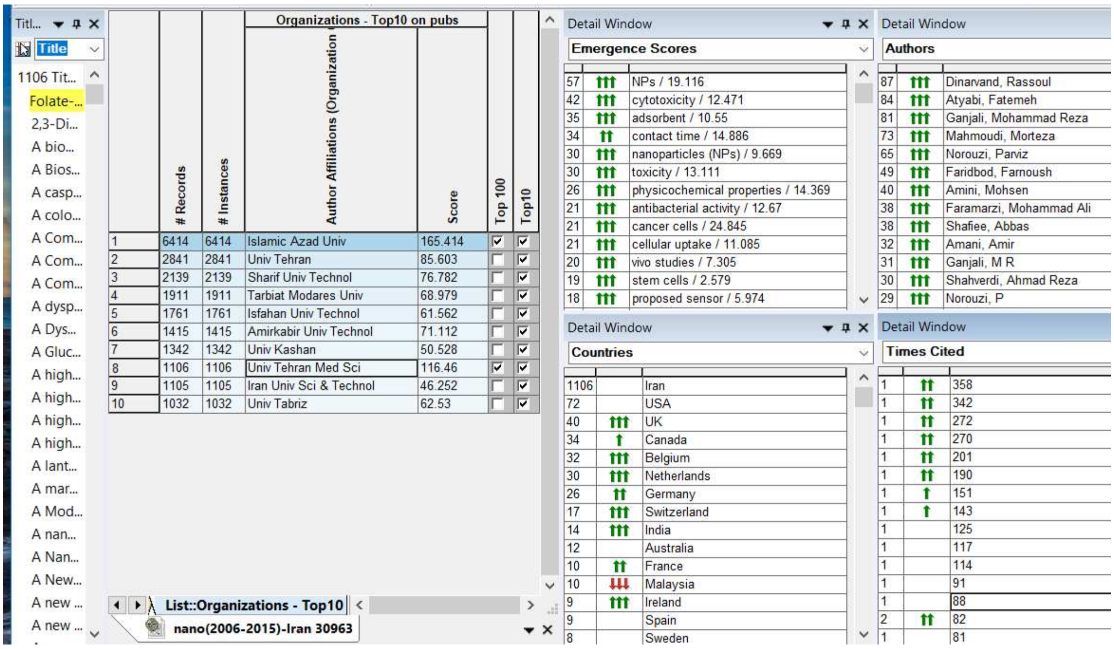

Figure 2. Leading Iranian Nanotechnology R&D Organizations based on EScores for 2006-2015, with Additional Information Breakouts