

Please cite this article as:

Wang,Z., Porter,A.L., Kwon, S., Youtie, J., Shapira, P., Carley, S.F., Liu, X., (2019) Updating a Search Strategy to Track Emerging Nanotechnologies, *Journal of Nanoparticle Research* 21: article #199; <https://doi.org/10.1007/s11051-019-4627-x>.

## Updating a Search Strategy to Track Emerging Nanotechnologies

Zhinan Wang, Alan L. Porter, Seokbeom Kwon, Jan Youtie, Philip Shapira, Stephen F. Carley, Xiaoyu Liu

### Abstract

Identifying nanoscience and nanotechnology-related (**nano**) publications in global databases is essential to profile and track research in the ever-changing field. Crafting an effective search query to retrieve as many nano-related publications as feasible, while moderating irrelevant publications, is challenging. This paper reports on a major update to a well-established nano search routine. We offer a 9-module search that significantly augments retrieval from the Web of Science (**WoS**) compared to use of the WoS Nanoscience & Nanotechnology Category (Nano-WC). We compare search results, showing modular overlaps with a “Nano\*” search, and among the modules.

Analyzing the resulting set of 2.2 million nano publication abstract records affirm the tremendous multidisciplinary reach of nano-related research and continued strong growth of the field. By country, China has enhanced its role as the world’s leading producer of nano publications, with slower growth for the US. Relatively high nano publication growth is also evident for India, Iran and several other emerging economies. Analyses of nano publications for 2013-2017 identify more than 20 emerging topics, primarily in the energy and two-dimensional materials domains that are apt to be actively researched in the coming two years or so.

---

Z. Wang

School of Management and Economic, Beijing Institute of Technology, Beijing, China

A. L. Porter (✉)

Search Technology, Inc, Norcross, GA, USA; Georgia Institute of Technology, Atlanta, GA, USA

e-mail: alan.porter@isye.gatech.edu

S. Kwon

Georgia Institute of Technology, Atlanta, GA, USA

J. Youtie

Program in Science, Technology, and Innovation Policy, Georgia Institute of Technology, Atlanta, GA, USA

P. Shapira

Manchester Institute of Innovation Research, Alliance Manchester Business School, University of Manchester, Manchester, UK; Program in Science, Technology, and Innovation Policy, Georgia Institute of Technology, Atlanta, GA, USA

S. F. Carley

Enterprise Innovation Institute, Georgia Tech, Atlanta, GA, USA; Intelligent Information Services Corporation, Atlanta, GA, USA

X. Liu

School of Management and Economic, Beijing Institute of Technology, Beijing, China

**Keywords:** Nanoinformatics, Nanoscience, Nanotechnology, Bibliometrics, Technological emergence, Emergent nano topics, Research trends, Scientometrics, Search strategy

## Introduction

The year 2018 marked 15 years since the authorization of the National Nanotechnology Initiative (NNI). Although funding of most of the multidisciplinary nanotechnology (**nano**) centers has ended, the nanotechnology domain continues to receive both research funding and support for infrastructure. The National Nanotechnology Initiative Supplement to the President's 2019 Budget identifies a \$1.4 billion request, even as it is 5% lower than the 2018 estimated budget. This milestone has triggered a request of the National Academies to conduct a Quadrennial Review of the NNI (National Academies, 2019).

The decline of subsequent publications in a field for which funding is reduced (Glaser and Velarde, 2018) seems a logical consequence. This paper, while not an evaluation of the NNI (see Rogers et al., 2012 for such an evaluation), seeks to understand whether such a decline has occurred with respect to nanotechnology publications and how the portfolio of nanotechnology research has evolved during this period of declining nanotechnology funding in the US, even as other countries such as China continue to maintain, if not increase, funding for nanotechnology R&D (Dong et al., 2016).

This period of declining funding in the US offers a unique window into the evolution of nanotechnology research. To this end, this paper asks three questions. The first is, are the overall parameters of what defines nanotechnology changed? The second concerns whether nanotechnology publications are maintaining growth or are they beginning to decline. The third is, what subtopics within the nanotechnology domain are emerging during this transition period?

A central challenge in nanoinformatics work is to ensure that the underlying search definition for identifying research publications in the dynamic domain of nanotechnology is systematically revalidated and updated to ensure optimal precision and recall. We herein undertake a new round of nanoinformatics work. We do so to serve interests of two ongoing projects – one concerning use of nano facilities and one on indicators of technological emergence (see Funding). For both projects, it is vital to assure that our nano literature coverage reflects current, evolving research emphases. To that end we sought to update a nano search strategy first formulated about a decade ago to retrieve records from the Web of Science (**WoS**) (Porter et al. 2008).

This paper describes our approach to update this nano dataset. We apply a 4-step structural framework to determine how to enhance a query to retrieve nano-related publication abstracts from WoS. A number of tradeoffs are assessed and search results are compared between our 9-module query and alternatives.

The paper then presents results that show the growth trend of nano research from 1991-2017. We explore shifts in disciplinary shares of that nano research over the recent decade. Tabulations also identify shifts “toward the east” in national shares of nano research, as indexed in WoS. We then examine emerging nano topics that suggest promising frontier research for the coming two years. Reflection on these nano research patterns raises interesting challenges.

## Nano Search Background

Various approaches to identify and retrieve nano literature have been used. Huang et al. (2010) review several of them – most of which incorporate “lexical” (term-based) searches. Here we consider some nanosearch variants and a framework for enhancing our search formulation.

An ambitious, pioneering approach by UCLA NanoBank combined multiple methods, including term-based searching with “bootstrapping” from identified core nano publications. Such approaches can “grow” base term sets based on citation patterns to extend linked literature (e.g., papers heavily cited by a core literature kernel and/or papers that cite such core literature). While appealing, such approaches are costly and challenging to validate and maintain, in part because of the changing core literature kernel. NanoBank [<http://www.nanobank.org/>] has been utilized by Zucker et al. (2007) in studying nano funding in relation to regional growth, and by Milojevic' (2012) examining cognitive evolution of the field (NanoBank data through 2004).

We determined here to build upon the term-based, modular approach of our prior nano search (Porter et al. 2008; advanced by Arora et al. 2013; see also Arora et al. 2014). Huang et al. (2010) consider the scope of our term-based procedure as reasonable – i.e., “mainstream” in the realm of nano searches. Stopar et al. (2016) compare three nano searches in WoS, along with a straightforward “nano\*” [nano alone or followed by other characters] by itself; ours is intermediate in adding records beyond a query using “nano\*” alone. Various others have applied close variants of our search (c.f., Levin et al. 2016; Wang et al. 2013). Pero (2013) considers our query along with several others (Glänzel and Meyer 2003; Schmoch and Thielmann 2012), in adopting that of Noyons et al. (2003).

Huang et al. (2015) offer an attractive systematic framework for bibliometric searching (developed for a non-

nano topic) that we find useful. It suggests a 4-step approach that we adapt (Table 1).

**Table 1.** A 4-Step Systematic Search Framework

Step 1.	Consider structural changes in the core lexical search approach
Step 2.	Expand the lexical query (enrich or change the search terms)
Step 3.	Reconsider specialized nano journals, used in a complementary search module
Step 4.	Consider citations to help extend the query.

We consider possible structural modifications to search effectively for nano-related publications in light of scientific and technological advances in the field over the last 10 years in nanomanufacturing, nanoelectronics, nanomedicine, sensors, and other areas<sup>1</sup>. Jo et al. (2016) examine the extent to which other journals link to a set of core nano journals based on citation relationships. Their analyses of nano-related journals suggest five major sub-domains: general, electronics, instruments, materials, and biotechnology. Stopar et al. (2016) grouped nano topics into four clusters: math-phys-chem; geo-eco; bio-med, and econ-soc, but, moreover, focus on how interdisciplinary these are, based on commonality of references cited. Wang et al. (2013) also focus on nano interdisciplinarity, analyzing five nano domains with consideration of terms specific to a domain, as well as terms more general across the domains.

Muñoz-Écija et al. (2017) use a search based on the Web of Science Category for Nanoscience & Nanotechnology, enriched by adding a limited number of most cited papers. They analyze these data to explore research sub-domains, offering a series of visualizations, showing rapid evolution of research interests. They present a set of tables that list most relevant terms for the following nano sub-domains: optics and electronics, materials synthesis and bottom-up processes, biotechnology and biomedicine, microelectronics engineering and top-down processes, physical and mechanical characteristics of materials, and organic electronics. This provides a good resource for one composing a nano search, but one needs to consider how best to adapt the terms (e.g., we inspect candidate terms; try alternative contingency requirements; and inspect 20-record sample abstract sets). For instance, Muñoz-Écija et al.'s organic electronics terms include “fullerene acceptor” (seeming quite nano-specific) and “particulate film” (apt to be used diversely and beyond nano).

This small sampling of nanoinformatics studies that distinguish sub-topics within nano indicates the lack of consensus on parsing nano R&D into topical sub-domains, especially for search query development. Were such sub-domains clearer and well-established, one might pursue sub-domain-specific search strategies. We do not do that, although the modules in Table 2 (next section) reflect some such emphases (e.g., Modules 2,3, and 4), but we organize the other topical modules mainly so as to group terms based on which search contingencies are used (nano-related Modules 5, 6, and 7 differ on contingencies). Module 1 is general nano\* and Module 8 captures nano-related journals.

Relating to search formulation Step 4 (Table 1), we considered backward citation (references) and forward citing as ways to expand our reach. UCLA NanoBank (Darby and Zucker 2003; Zucker and Darby 2007) had considered citation of core nano literature to help determine nano scope. However, we determined not to attempt that here.

Appendix Table A1 shares WoS citation counts for a set of papers pertaining to nano searching (albeit in various ways). Porter et al. (2008) and Arora et al. (2013) are among the heavily cited nano search papers. We selectively examined titles and abstracts of papers that cite these nano search papers, seeking leads on possibilities to extend our Boolean search strategy in various ways (e.g., some of the alternative approaches mentioned herein, additional candidate terms). Some of the papers listed in Table A1 were themselves discovered via this citation tracking approach (a form of bootstrapping).

Another strategic possibility is to alter search lexicon over time for an emerging field such as nano. Zhu et al. (2017) draw upon the keywords devised for nano search used by Chen et al. (2013), enriched for research since 2011 with additional terms to improve capture of “active nano” research (Suominen et al., 2016) and terms oriented to get at advancing subdomains. Likewise, we seek to enrich our previous search strategy, but we go ahead and apply the additional search terms (Table 2, terms in bold) to the whole period (1991-2017). As with Zhu et al., we complement a “nano\*” search module by additionally searching for other nano-relevant terms (see Table 2). Our terms correspond well with their terms (Zhu et al., Tables 6 and 7). Our set is more extensive, with considerable

<sup>1</sup> <https://www.nano.gov/sites/default/files/NNI-FY18-Budget-Supplement.pdf>

variation in exact expressions. For someone considering doing a comprehensive nano search, we note some of their terms that we have not included.<sup>2</sup>

## Data and Methods

### *Toward A Revised Nano Search Formulation*

Mindful of the four-step framework presented in Table 1, we considered search modifications of several sorts. In Step 1, the core lexical search, we experimented with an encompassing set of microscopy terms, in conjunction with our contingency terms (Table 2, Module 4). The rationale was that “any microscopy directed at the molecular level” could pertain to nano. This was not the case; results were excessively noisy. So we retain our approach of combining

- a) nano-oriented terms, limited by requiring:
- b) contingencies that favor molecular environments.

We consider applying contingency requirements module by module. As per Table 2, we require the terms in Modules 3 and 6 to co-occur within an abstract record’s topical fields with a term from *Molecular Environment, Inclusive* (MolEnv-I) for a record to be taken. Analogously, terms in Modules 4 and 7 need to co-occur with a *Molecular Environment, Restricted* (MolEnv-R) term.

---

<sup>2</sup> Zhu et al. (2017) search terms include: biomotor, molecular sensor, quantum computing, quantum effect, artificial photosynthes, atom thick layer, biophotonics, cellulose fiber or tube, dendrimers, DNA computing or assembling, molecular system, optoelectronic, optogenetic, proteomic, spintronic, and supramolecul [we include supramolecul\* and quantum\* as contingent terms for several of our search modules].

**Table 2.** Updated nanotechnology definition: modular search query

Set	Emphasis	Contingency with Molecular Environment Terms (I or R)	Terms	# of Records
1	Nano*	No	TS = (nano*)	1505890
2	Quantum	No	TS = (("quantum dot*" OR "quantum well*" OR "quantum wire*") or ("Nuclear Quantum Effect*" OR "quantum capacitance"))	176358
3	Self-assembly	MolEnv-I	TS = (("self assembl*" OR "self organiz*" OR "directed assembl*") AND MolEnv-I)	150199
4	Microscopy and spectroscopy	MolEnv-R	TS = (((("TEM" OR "STM" OR "EDX" OR "AFM" OR "HRTEM" OR "SEM" OR "EELS" OR "SERS" OR "MFM" OR "SAXS" OR "atom* force microscop*" OR "tunnel* microscop*" OR "Probe Microscop*" OR "transmission electron microscop*" OR "scanning electron microscop*" OR "energy dispersive X ray" OR "electron energy loss spectroscop*" OR "single molecule microscop*" OR "focused ion beam" OR "ellipsometry" OR "magnetic force microscop*" OR "Small Angle X-Ray Scattering") OR ("EDS" OR "enhanced raman spectroscop*" or "enhanced raman scattering"))) AND MolEnv-R)	296829
5	Nano-related	No	TS = (("molecul* motor*" OR "molecul* ruler*" OR "molecul* wir*" OR "molecul* devic*" OR "molecular engineering" OR "molecular electronic*" OR "single molecul*" OR "fullerene*" OR "buckyball" OR "buckminsterfullerene" OR "C60" OR "C 60" OR "methanofullerene" OR "metallofullerene" OR "SWCNT" OR "MWCNT" OR "coulomb blockad*" OR "bionano*" OR "langmuir blodgett" OR "coulomb staircase*" OR "PDMS stamp*" OR "graphene" OR "dye sensitized solar cell" OR "DSSC" OR "ferrofluid*" OR "core shell") OR ("layer-by-layer" OR "molecul* junction*" OR NaYF4 OR silicene OR phosphorene OR germanene OR graphyne OR stanene OR borophene OR "transition metal dichalcogenide*" OR CQDs or CQD OR ("nitrogen-doped porous carbon" or "nitrogen-doped carbon" or "N-doped carbon")))	380170
6	Nano-pertinent (highly)	MolEnv-I	TS = (((("NEMS" OR "quasicrystal*" OR "quasi crystal*" OR "quantum size effect" OR "quantum device") OR ("two-dimensional material*" OR "2D material*" OR "2-D material*")) AND MolEnv-I)	12656
7	Nano-pertinent (moderately)	MolEnv-R	TS = (((("biosensor*" OR "sol gel*" OR "solgel*" OR "dendrimer*" OR "CNT" OR "soft lithograph*" OR "electron beam lithography" OR "e beam lithography" OR "molecular simul*" OR "molecular machin*" OR "molecular imprinting" OR "quantum effect*" OR "surface energy" OR "molecular sieve*" OR "mesoporous material*" OR "mesoporous silica" OR "porous silicon" OR "zeta potential" OR "epitax*" OR "atomistic simulat*" or "atomistic model*") OR (Intercalat* OR "molecular* imprint*" OR "hybrid material*" OR C-dot* OR "carbon dot*" OR "molecular simulat*" or "molecular dynamics" or "molecular model*" OR exciton* OR plasmon* OR "spin-orbit coupling"))) AND MolEnv-R)	345900
8	Nano journals	No	SO = (Fullerene* OR "Journal of Nano*" OR Nano*)	117346
9	Nano-WC	No	WC = ("nanoscience & nanotechnology")	478680
		MolEnv-I	(monolayer* OR "mono-layer*" OR film* OR quantum* OR multilayer* OR "multi-layer*" OR array* OR molecul* OR polymer* OR "co-polymer*" OR copolymer* OR mater* OR biolog* OR supramolecul*)	
		MolEnv-R	(monolayer* OR "mono-layer*" OR film* OR quantum* OR multilayer* OR "multi-layer*" OR array*)	
Total			#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9	2313133

Note to table 2: Numbers in the last column give search results on Aug 8, 2018. The total of ~2.3 million is later reduced by removing “nano-noise” and duplicate records to ~2.2 million [2,235,233].

Pertinent to Step 2 in formulating a search (Table 1) -- expanding the lexical query -- Figure 1 in Arora et al. (2013) schematically presents our nano search constituting eight modules – nano\*, six topical modules, and a specialized journal module. That is accompanied by a set of exclusions to remove inappropriate records from the nano\* search portion.

Table 2 herein presents our expanded lexical query – retaining the core from the Arora et al. (2013) nano search (all the modules, using the terms not in boldface<sup>3</sup>; adding a new module #9); plus additional terms added to six of those modules (the **boldface** terms).<sup>4</sup> The contingency conditions, unchanged from Arora et al. (2013) – one set is most constraining [MolEnv-R] and an alternative, more encompassing set of terms [MolEnv-I] – also appear in Table 2. The nano\* exclusions are applied after download of records containing “nano\*” to remove records that are unlikely to relate to nano – e.g., NaNO<sub>2</sub>). These nano exclusion terms remain unchanged from Arora et al. (2013) Table 3 and Table 4 [that they draw from Grienesen and Zhang (2011)], so are not repeated here (but are made available in the Supplemental Materials as Tables S-1 and S-2).

For those interested in particular nano search details, we distinguish various sets of terms by use of parentheses and boldface. [This still allows one to copy & paste these queries directly into a WoS advanced search.] To illustrate, for Module 4 (microscopy and spectroscopy):

- 1) First group has three sub-groups:
  - a. Unchanged terms carried forward from Arora et al. (2013) – e.g., (“TEM”, ... “tunnel\*  
microscop\*”,...)  
[plain font, within first set of parentheses]
  - b. Original terms with modification – e.g., “**Probe Microscop\***”
  - c. Newly added terms that do not show rapid growth for 2008-2017 – e.g., “**Small Angle X-Ray Scattering**”; “**SAXS**”  
[both b. and c. are in **bold font**]
- 2) Second group consists of newly added terms with rapid growth in recent years – e.g., (“**EDS**” OR “**enhanced raman spectroscop\***” or “**enhanced raman scattering**”)  
[these are in the last set of parentheses and are in bold font]

The main intent of these adaptations is to update the search as nano research interests advance and change, and thereby more fully include emerging nanotechnologies. The proportional changes in adding some 35 terms and variants are generally moderate. Modules 1 and 3 add no terms; the new terms in Modules 2, 4, and 5 add less than 10% to the counterpart core module tallies, for years 1991-2017.

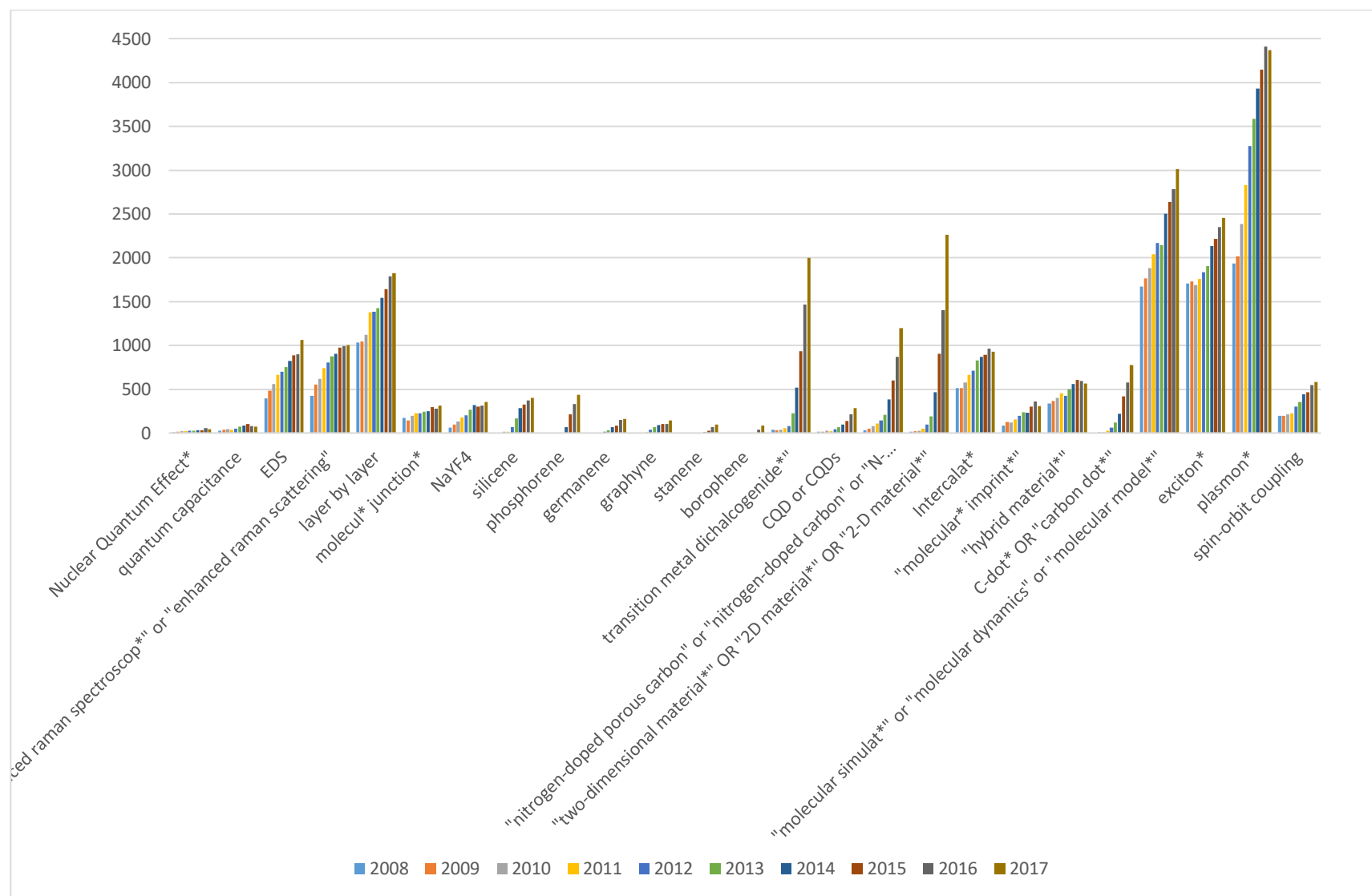
Relating to the growth trend check, this also gives a modest validity test for our earlier work. We would like to see whether these terms have grown strongly in recent years, and were not strongly present (if at all) in earlier years. [Conversely, if new terms have been strongly active for years, we would have liked to have captured them earlier.] We check terms one by one in WoS; the bar graph (Figure 1) shows mostly strong recent growth trends, except for Small Angle X-Ray Scattering” (or “SAXS”) and “atomistic simulat\*” (or “atomistic model\*”) (not shown).

---

<sup>3</sup> The modules incorporate a few modest refinements:

- a. Module 5 is now Module 4.
- b. replace microscopy by microscop\* in two places - “**single molecule microscop\***” and “**magnetic force microscop\***” (adding 11 records); replacing “scanning probe microscop\*” by “Probe microscop\*”;
- c. remove X-ray photoelectron\* from Module 4 based on examination of sample records of variations of the term (e.g., XPS)
- d. refine variations of enhanced Raman scattering/spectroscopy
- e. add “SAXS” (the acronym for “small angle x-ray scattering”).
- f. remove “NEMS” from Module 7, retaining it in Module 6;
- g. relocate a few microscopy terms to other modules to tune contingencies;
- h. Remove one journal -- “Materials Science Engineering C\*” – as it showed minimal nano hits (1/20 in a test sample) after removal of records captured by the rest of the nano search.

<sup>4</sup> We checked ~135 candidate terms, declining to include most. This was based on search comparisons and inspection of 20-record samples drawn from search results using the given term(s). E.g., some terms judged as not adding sufficient nano coverage after consideration of alternative contingencies and exclusion of nano\* include: STEM, FIB, UPS, SIMS.



**Fig. 1.** Growth Trends of the Added Terms



Relating to search formulation Step 3 (Table 1) – a module that retrieves all records from specialized nano journals -- we applied our core search query for recent years to identify candidate journals to add (anticipating a number of new nano-related sources).

We then shifted gears and compared coverage of our search modules to records identified by searching on the Web of Science Category (“WC”) “Nanoscience & Nanotechnology” (“**Nano-WC**”). Jo et al. (2016) consider the role of this Nano-WC in identifying nano-related journals, noting growth in the number of journals included. Nano-WC rose from 27 journals on inception in 2005 to 92 as of 2017. Using a Nano-WC journal-based search to identify nano-related research can complement our lexical (term-based) search – especially attractive, it is not dependent on appearance of previously identified terms. WoS decides what journals to include in a WC based on editorial knowledge about the domain plus empirical evidence.<sup>5</sup> Hence, publication in a source in the Nano-WC is presumptive evidence of a good likelihood of nano-relevance, though not a guarantee.

One might ask if a simple search just using Nano-WC would do the job? Perhaps, nano as a field has matured such that one can identify relatively well-defined boundaries? Then if Nano-WC reflects those well (and, as per Footnote 4, it seems well-founded), scholars can focus on that content to study nano research activity. This would be attractive in simplifying a lexical (term-based) search subject to the vagaries of selecting search terms. Such a search improves recall of nano-related research based on publication in the Nano-WC sources, whatever terms are used in particular articles.

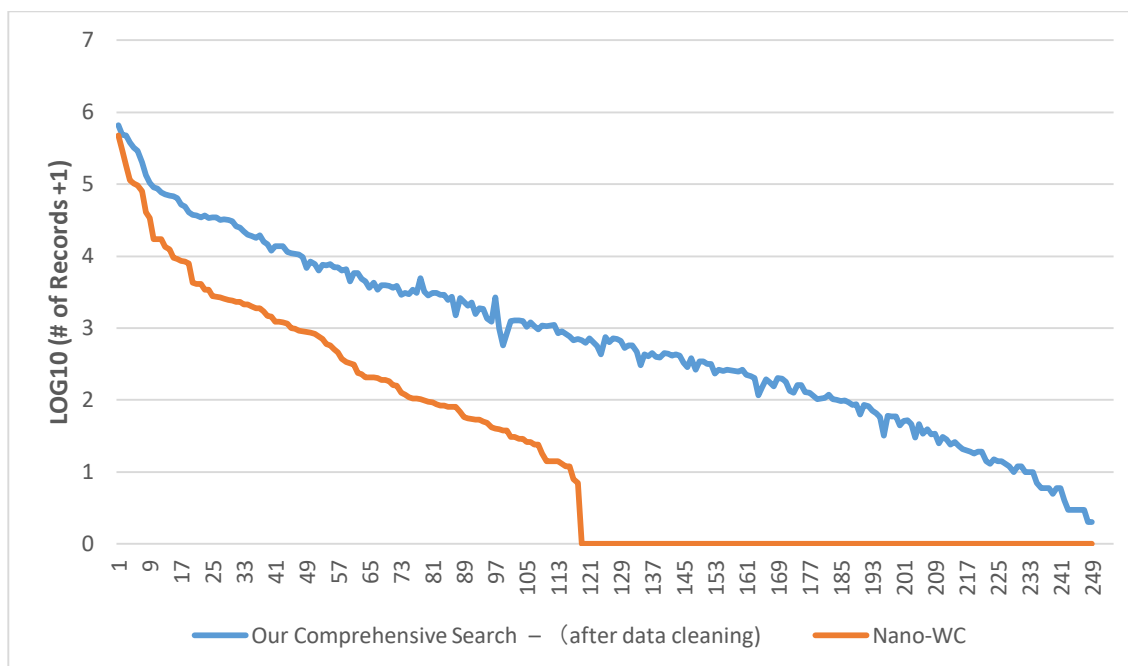
So, does searching on “Nano-WC” suffice to supplant our term-based search? In a word, no. Table 2 indicates tremendous nano content is published beyond the Nano-WC core sources, suggesting that a simple compilation of papers in journals and journal related information is not sufficient for delineating the domain of this dynamic and multidisciplinary field. Nano-WC discovers under 0.5 million records of a total of some 2.3 million. Reliance on Nano-WC alone would miss over three-quarters of what we recognize as nano-related research – i.e., very poor recall. It would also miss diffusion of nano-related research beyond the core sources included in Nano-WC. Such external literature could well be rich in emerging nano research. We are also interested in finding potential new applications beyond nano fields per se. Comparison of topical emphases between the record subsets offers potential insights – e.g., Nano-WC records vs. others.

Figure 2 compares nano publication activity by WC. The horizontal axis orders WCs by nano content in our searches. The vertical axis shows the log of search hits for our comprehensive nano search (top line) vs. log of counts resulting from a Nano-WC search (lower line). Figure 2 indicates that our comprehensive search shows significantly broader reach for nano research than just using Nano-WC.

We show the record count for each WC in the supplemental materials. This allows a reader to identify nano activity in particular WCs – they might do so to investigate content of one or more particular WCs – either to check for spurious inclusion by our search, or to explore nano research in one or more WCs that is of special interest. For instance, were one studying nano research activity pertinent to “Food Science Technology,” running the comprehensive search AND WC = Food Science Technology, would focus on the current records in that intersection (11,474 as of our August, 2018, search).

---

<sup>5</sup> WoS Helpdesk notes factors considered in assigning journals or other sources to WCs: journal subject matter & scope; author & editorial board affiliations; funding acknowledgements; cited & citing relationships to other journals; journal sponsor; journal’s categorizations in other bibliographic databases.



**Fig. 2.** Web of Science Category Distribution Comparison

Others extend a “nano journals set” beyond those included in Nano-WC – e.g., Grieneisen and Zhang (2011) report analyses of 161 journals publishing in 2010-2011. Of course, many nano papers appear in sources not indexed in Nano-WC [e.g., Bartol and Stopar (2015) note that 20% of all 2012 article titles in *Applied Physics Letters* and in *Journal of Applied Physics* contain a nano\* term].

In assessing Nano-WC coverage, we noted that using advanced search in WoS (WC= “nanoscience & nanotechnology”) captures sources well beyond the Nano-WC sources identified in the *Journal Citation Reports (JCR)*. As of 2017, JCR identifies 92 journals in Nano-WC. Some additional journals are included in the WC search; these include some that are no longer published and some new ones that have not been published for three years yet, so as to be included in JCR. Moreover, the Nano-WC search, yielding about 477,000 records for 1991-2017 for the Core Collection, includes hundreds of sources besides those journals. Over 100,000 records from other than the 92 core journals are captured. Small sample inspection of records supported their nano-relevance at well over 95%. We thus incorporate a Nano-WC search module. Table S-3 (Supplemental Materials) shows that a Nano-WC search does capture many papers in diverse WCs, but not as fully as the comprehensive search. E.g., Nano-WC retrieves 603 “Food Science Technology” papers, vs. 11,474 for the comprehensive nano search.

As we moved to use Nano-WC as a search component, we compared it to coverage of search Module 8 – our specialized set of nano-oriented journals. Of some 150,000 articles formerly in Module 8, about 95% were included in Nano-WC. We retain Module 8 in much reduced format [Table 2]. We had also identified 9 new nano-related journals to add, but all are captured by Nano-WC [Module 9] anyway.<sup>6</sup>

We also recognized potential value in using WC searching to improve precision. Somewhat the converse of Nano-WC searching, we considered excluding records that appear in certain other WCs. The logic is that nano terminology has spread so as to be used beyond nano-related research. For one, our nano-pertinent acronyms (e.g., check Module 4 in Table 2 for “TEM,” etc.) are apt to convey diverse meanings in different literatures. The next section details explorations as we sought to improve the search. Therein we discuss possibly excluding noise via use of a list of non-nano WCs, but with concerns to minimize loss of nano-relevant research diffusing into other fields.

<sup>6</sup> The new journals were identified using time span 2015-2017. We exclude those journals captured by the original Module 8 (Sanjay, 2013). The 9 new journals are Applied Nanoscience, E-Journal Of Surface Science And Nanotechnology, European Journal Of Nanomedicine, Ieee Nanotechnology Magazine, International Journal Of Nano Dimension, International Journal Of Nanoscience, International Nano Letters, Proceedings Of The Institution Of Mechanical Engineers Part N-Journal Of Nanomaterials Nanoengineering and Nanosystems, and Small Methods.

## Executing a Revised Search Framework

The previous section has introduced our search framework. We now explain in further detail how we refined the search elements to enrich nano research coverage.

As noted, we implement a systematic reconsideration of our nano search formulation. We assess candidate new terms to augment the nano search query based on their adding significant nano content. But we are also curious as to whether they are adding rapid growth nano research topics (as per Figure 1). In a later section, we go on to investigate “emerging terms and topics” in the augmented nano search dataset,<sup>7</sup> applying a process to generate emergence indicators (Porter et al. 2018).

Table 1 sets out four steps for reviewing the nano search formulation; the prior section explored structural changes (Step 1). Here are two examples pertaining to the decision not to alter the main search strategy regarding contingencies. Review by an Electrical Engineering PhD student pointed out that one MolEnv-R term -- “array” -- has wide non-nano use in telecommunications. Likewise, several MolEnv-I terms are really broad (e.g., “biology\*”). We considered revising our approach to narrow the contingencies, but chose to maintain consistency with our prior search. On the other side of the contingency relationships, we noted that “any” microscopy activity at the molecular level could well be nano-related. Exploring alternative search formulations found that widening Module 4 would have required extreme tightening of the MolEnv-R contingencies. We decided, instead, to retain the current approach of combining nano-oriented terms, co-occurring with moderately restrictive contingency terms.

Step 2 – search expansion – drew upon review of nanoinformatic approaches (described in the prior section). However, the main work entailed a bootstrap approach. In April, 2018, we ran the core nano search (applying the terms not in boldface in Table 2) in the WoS Core Collection (eight Citation Indexes)<sup>8</sup> for 2015-2017. We downloaded abstract records and imported into *VantagePoint* software ([www.theVantagePoint.com](http://www.theVantagePoint.com)) for data cleaning, analyses, and visualizations. To identify candidate new search terms, we extracted title phrases using *VantagePoint*’s Natural Language Processing (NLP) routine for recent search results. We calculate a noise ratio of relative frequency (hits/records) in a 50,000 record WoS dataset (approximately random sample for 2016) relative to hits/record in the nano search. Terms in the core search (not in boldface in Table 2) mostly showed <8.5% coverage in the random WoS dataset, so we used this as a screening benchmark. We also examined terms absent in the general WoS dataset but present in the nano dataset. This yielded some 2,260 terms to consider as possible additional search terms. We eliminated most of those – e.g., those appearing in fewer than 3 records, any starting with nano\* (already would be retrieved by the nano\* module), terms covered by our other modules, and general terms (e.g., “templates”).

The remaining ~135 candidate terms and phrases were examined. We gathered and checked brief Wikipedia definitions for many. Then we searched in WoS to identify terms with sizable counts warranting further consideration. We compared 1) hit rates overall; 2) hit rates excluding nano\* occurrences; 3) hits in combination with MolEnv-I; and 4) hits with MolEnv-R. We then examined 20-record abstract record samples (usually the most recent 20 in the WoS search), estimating how many out of a 20-record sample we deemed nano-relevant. As an illustration, “atomic layer deposition” is one of those 135 candidate additional terms. It appeared in 14,572 records (WoS, 1991-2017), of which 8,889 were not captured by a nano\* search, but only 2/20 in the sample examined were deemed on target as nano research, so this term was not added to our query.

---

<sup>7</sup> We have some records downloaded earlier and some currently in the augmented dataset. When we compared WC distribution, we found that for those earlier downloaded records, there are a few of them whose WC has minor changes, for example, WC for "JOURNAL OF HAZARDOUS MATERIALS" was changed from "Engineering, Environmental; Engineering, Civil; Environmental Sciences" to "Engineering, Environmental; Environmental Sciences"; WC for "JOURNAL OF NUCLEAR MATERIALS" was changed from "Materials Science, Multidisciplinary; Nuclear Science & Technology; Mining & Mineral Processing" to "Materials Science, Multidisciplinary; Nuclear Science & Technology";.

<sup>8</sup> In the Georgia Tech interface to WoS, under “More Settings,” we check these 8 databases: Science Citation Index Expanded, 1900-present (SCI-Expanded); Social Sciences Citation Index, 1900-present (SSCI); Arts & Humanities Citation Index, 1975-present (A&HCI); Conference Proceedings Citation Index – Science, 1990-present (CPCI-S); Conference Proceedings Citation Index -- Social Science & Humanities (CPCI-SSH), 1990-present; Book Citation Index – Science, 2005-present (BKCI-S); Book Citation Index – Social Sciences & Humanities, 2005-present (BKCI-SSH); Emerging Sources Citation Index – 1985-present (ESCI). We don’t include the two Chemical Indexes, as all records in testing several of our nano modules are fully included in the eight Core Collection databases.

Another example – “2-D materials” (or two-dimensional materials or 2D materials) appeared in 4,156 records in WoS, 1991-2017. One screening check examined records meeting the MolEnv-I contingency, but not the MolEnv-R one (i.e., “in-between” – records that would not qualify with the stricter contingency). Of 20 records, 14 were deemed nano-related. So, we decided to add this phrasing to the search with a MolEnv-I contingency.

In the screening process, we unearthed various search issues, including:

- How to capture terms containing “nano,” but not at the beginning of the term? We can’t, unless we specify exact term beginnings (WoS advanced search requires at least 3 characters precede a wildcard). We liked “heteronano\*,” but it only occurred in 35 non-nano\* records, so we did not include this term in the search (that yields some 2.2 million records). This also illustrates the tradeoff faced between completeness and simplicity; here we opt for not adding an additional term for so little increase.
- “Atomic scale” offered apparent appeal, but was not judged to associate highly with nano research; we also considered adding it to MolEnv-R, but decided not to do so.

For the 20-record sample judging – this was done mainly by one of us (AP), with a long history of tracking nano R&D through nanoinformatics, but not a nano researcher as such. The general stance was to aim for “nano-relatedness,” as differentiated from work that is predominantly focusing on physics, optics, chemistry, biochemistry, semi-conductor, or other such facets. “Plasmon\*” gives an example in which we try to separate more prominently nano research from such other interests. One could make arguments for much or little plasmon research being “nano.” We opt for a somewhat conservative stance, taking plasmon\* contingent on one or more MolEnv-R terms co-occurring in the abstract record with plasmon. This reduces 95,373 plasmon\* records to 61,773 with a judged sample of 13.5/20 [with “0.5” ratings indicating nano to some degree] as nano-related. We also determined to align treatment of plasmon with exciton, and spin-orbit coupling, and samples of those showed similar nano-relevance. For anyone wanting further illustration of our term assessment processes, Appendix B details a case of determining whether to include particular terms.

Our main search was built through a series of WoS downloads in June-August, 2018, building upon prior search downloads. In conducting the search, we selected the timespan, 1991-2017.<sup>9</sup> We caution that time series analyses could be somewhat distorted as WoS coverage broadens over time, new journals are added, and that 2017 indexing is not fully complete. We selected the eight Citation Indexes under “More settings,” and used “Advanced Search to execute the various modules with contingency conditions. Except for Module 8 concerning journals (search using “SO” for source), and WoSCs (searches using the “WC” field), the topical module searches were all conducted using “TS” – topic search that covers titles, abstracts, Keywords (author), and Keywords Plus.

We chose to extend the search beyond the SCI and SSCI databases of WoS. As per Footnote 7, the eight WoS Core Collection Citation Indexes include conference proceedings papers, books, and emerging sources. We compared topical module search results for the eight Citation Indexes vs. just SCI-SSCI for 1991-2017; all ratios showed modest increase, ranging from 1.07 to 1.17. This seems desirable with the possible gain that some emerging research may appear sooner in conferences in some fields, and possibly in other sources (e.g., ESCI increases inclusion of Chinese sources).

Pertinent to search formulation item #3 (Table 1) – consideration of inclusion of any papers in selected sources – we proceeded with review of various journals for suitability. We checked journals in the core search and removed one -- “Materials Science and Engineering C” – as adding minimal nano-relevant records over and above what the other modules capture.<sup>10</sup> That is, a search in the Source field (SO) for this journal finds 8119 records, of which 4677 are captured anyway by our other search modules. Inspection led us to decide that the remaining 3442 were not sufficiently nano-related to warrant blanket inclusion.

We also scanned our core nano search results for 2015-2017 to identify additional journals to consider adding. This led to nine candidates, but eventually we determined instead to add a WoSC = “nanoscience & nanotechnology” search component instead. That adds some 10% to the overall search and appears well-justified.

Tuning the nano search poses difficult balancing of precision and recall. One seeks to capture a high percentage of nano-related research (high recall), without undue noise (i.e., attain high precision). In focusing the

---

<sup>9</sup> As per Footnote 7, not all the component databases extend back all the way to 1991. We were pleased to learn that WoSC’s have been assigned to nearly all sources, extending well beyond SCI and SSCI.

<sup>10</sup> As we advanced our search formulation, the 20-record sampling adjusted. For initial term screening, as noted, we examined search results for the target term, with and without contingencies -- not screening out records that would also be captured by nano\* or other search modules. By counting records that included such other terms, we got a sense of the target term’s relatedness to nano research. Later, in tuning the search modules in conjunction with each other, we did further sample checks during which we excluded nano\*, particularly. This makes for more strenuous checks. Hence, we don’t assert one given sampling threshold for inclusion of a term.

search formulation on WoS, more so than previously, we can utilize their categorization of sources (journals) – the WCs. We considering doing so in two ways. As discussed, we add “Nano-WC” as search Module #9 to capture records appearing in those nano-related journals, whatever terminology they use. This provides a mechanism to include novel topics pertinent to nano research. The next section describes the second way that we considered (but rejected).

### *Exclusion of Records in Web of Science Categories Seemingly Remote from Nano Interests*

We also considered how to utilize non-nano WCs in best crafting the search. In scanning preliminary results, one of the “red flags” is to find records in fields (WCs) seemingly remote from nano research emphases – e.g., “dance.” On one hand, excluding records in WCs we judged to be non-nano promises good gain in precision. On the other, it could endanger recall of novel diffusion of nano research into seemingly remote fields. We nominated which of some 252 WCs we would consider to be non-nano; then looked at intersections among WCs (a minority of sources are assigned multiple WCs – nano research does so relatively more often). We formed a set of 134 non-nano WCs; then examined the extent that they intersected our nano search modules.

We examined how counts for each search module changed by excluding records in the 134 candidate non-nano WCs. With three exceptions, the exclusion rate is under 2%. Module 1 (nano\*) and Module 3 (self-assembly) lose 3.1%, and Module 5 (nano-related) drops 2.8%. Expert review suggested reducing the 134 “non-nano WC” set as nano research could reasonably occur therein. We further analyzed intersections of records in a non-nano-WC subset with records in a nano\* search. We drew samples from about half of the 134 WCs in question, finding sufficient nano-related research to reduce the WCs. Inspection of a resulting set of 76 refined non-nano WCs found way fewer outtakes resulting from their use in restricting search modules #1-8 -- 6647. Of those, 5308 reflect Module #1 (nano\*) and 1339 come from Modules #2-8. As mentioned, we have a routine (Arora et al. 2013, Table 3; also provided in Supplemental Materials) that scours downloaded records for inclusion of nano noise variants (e.g., nanok in literature as “Nanook of the North”). When applied, this reduces the 5308 nano\* terms by 2226 to 3082. Inspection of 30 of those abstract records found at least 20 terms suggestive of nano-relevant research.

So, we determined *not* to blanket remove any records by applying the exclusion WCs. Upon reflection, we determined that the existing screening process to remove nano noise was a better approach than excluding based on publication in sources indexed in WCs seemingly unrelated to nano. We also note that as we came down to a reduced level of exclusion WCs, the number of records removed is almost trivial (e.g., 6647 – 2226 = 4421 out of some 2.2 million, or ~0.2%).<sup>11</sup>

## **Results**

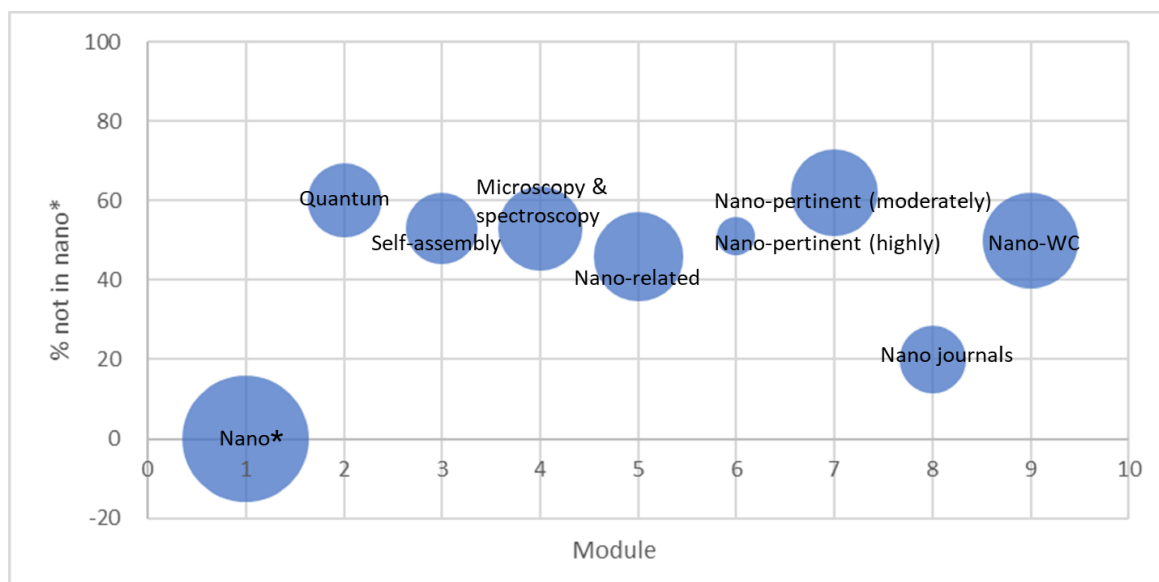
### *Nano Search Comparisons*

In advancing a revised nano search, we check how it differs from simple nano\* or Nano-WC searches alone and from our prior searches. We compare the new search (S3) to nano\* alone and to the prior Arora et al. (2013) version (S2). [We also compared application of the original Porter et al. (2008) (S1), but that is not much different from S2 in terms of records retrieved.] These comparisons are based on current searches for 1991-2017 in the 8 Core WoS Databases, as discussed.

Figure 3 depicts the degree of overlap of the search modules with nano\*. The y-axis shows the proportion of each module that doesn’t overlap with nano\*. Most modules show roughly half overlap; this indicates strong relevance to nano-research and useful extension of the nano search by inclusion of the modules. We have also inspected the extent to which Modules 2-9 intersect with each other (Figure 4), but do not attempt to show that in Figure 3.

---

<sup>11</sup> We note that conduct of the search queries yields ~2.3 million records prior to removal of nano noise terms and duplicates, which reduces the total to ~2.2 million. To check, we ran our nano noise removal routine on the 1339 records from Modules #2-8 that application of the 76 “non-nano WC” set would remove. Only 13 intersect with our nano noise removal routine. We determine not to remove these 1339 records (0.06% of 2.2 million nano records) in favor of search parsimony.



**Fig. 3.** Degree of Overlap of the Nano Search Modules with a Nano\* Search

Note to Fig.3.: Number of publications for a module is denoted as the area, the diameter for each module is calculated by dividing the area by Pi first, then taking the square root of the result as the radius, finally doubling the radius.

Table 3 describes how many records are gathered by alternative search formulations. The top row gives total publication records, as retrieved by Nano\* itself, and by our three successive search versions. Of interest:

- About 65% of the records in Search 3 would be captured by nano\* alone
- Search 3 augments Search 2 by about 10% (236,805)
- Search 3 also leaves out 16,452 Search 2 records

**Table 3.** # of Records in Common among Alternative Nano Search Formulations (Update on Aug 8, 2018)

	Nano*	S1	S2	S3
<b>Total by Search:</b>	1505890	1949902	2092780	2313133
<b>Exclude Nano*</b>		444012	586890	807243
<b>Exclude S1</b>			144856	366307
<b>Exclude S2</b>		1978		236805
<b>Exclude S3</b>		3076	16452	

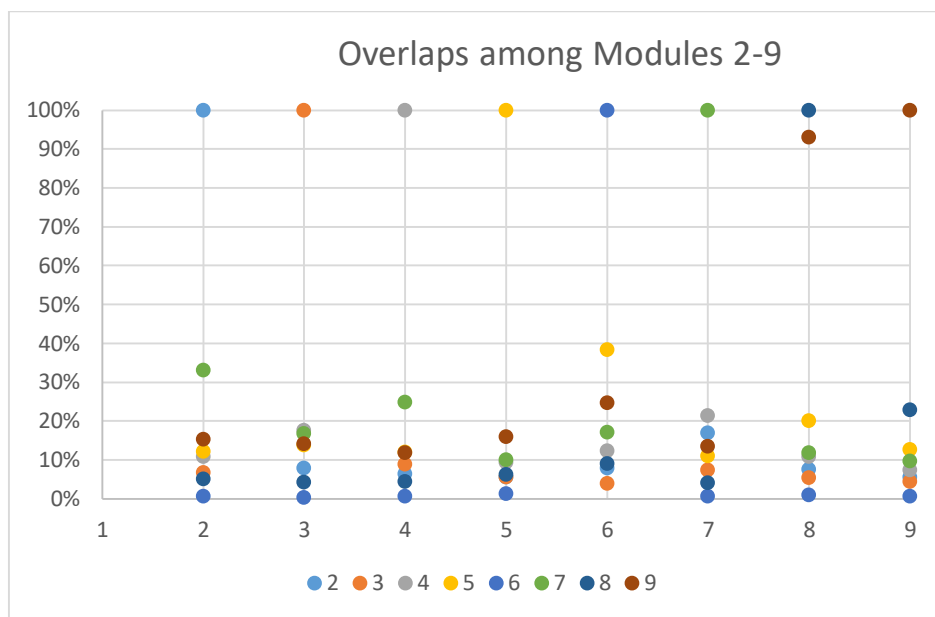
**Note to Table 3:** Search counts change daily, so are not permanent. Nano\* is a search in WoS for 1991-2017, as described. S1 is our current application of the Porter et al. (2008) version nano search; S2 is that slightly modified and reported by Arora et al. (2013); S3 is an in-progress version of the current search (Table 2).

Comparing S2 to S3, we considered both a) records from S2 not picked up by S3, and b) records newly added by S3. Intermediate results showed striking numbers for (a) when we used the 134 exclusion WCs; however, those dropped drastically to just 16,452 records left out when we determined not to apply those exclusions. Of the 16,452 that are not captured by other modules, 3,442 are in *Materials Science Engineering C*. Another 13,029 reflect removing X-ray photoelectron\* as a search term. In both cases our sampling checks judged that these were suitably left out. In this regard the new search appears superior.

Re: b) records added, S3 adds some 236,805 records anew of a total of ~2.3 million; S2 would give ~2.09 million records as run in August 2018. Of the 236,805 added records, 182,109 (77%) come from the addition of Module #9, Nano-WC. As discussed, WoS development of this source category provides a good rationale for inclusion. The remaining 54,288 added records reflect the new search terms incorporated (bold in Table 2). Each of those terms was vetted by sample judgment to support inclusion as nano-relevant. Thus, the newly added terms appear well justified. We believe S3 is superior to our prior versions, and to simpler alternatives of nano\* by itself, or Nano-WC alone.

Focusing again on the current search (S3), Figure 3 illustrates how component Modules 2-9 overlap with a nano\* search. One might additionally inquire as to how much Modules 2-9 overlap with each other. Figure 4 depicts

this. The vertical axis shows the percent of records in common between particular modules. The horizontal axis shows this separately for each module. Overlaps are not symmetrical. Figure 4 shows the extent to which records retrieved by the Module named along the horizontal axis are contained in the search results of the other Modules spread out along the vertical dimension.



**Fig. 4.** Overlaps among Modules 2-9

To illustrate, consider Module 2. It overlaps 33% with Module 7; 15% or less with the others (of Modules 3-9). Figure 3 separately presents overlaps with Module 1 (nano\*). In general, most Module 2-9 overlaps are below 20%, supporting the logic of including all of them in a comprehensive nano search. Module 6 is sizably included (38%) in Module 5 (both called “nano-pertinent”), where substantial overlap seems reasonable. [Conversely, only 1% of Module 5 is captured also by the much smaller Module 6 – see Table 2.] The extreme exception is that Module 8 (journals whose name contains nano\* or fullerene\*) is 93% contained by search Module 9, Nano-WC. We retain Module 8 mainly for potential future searches, to cover potential new nano journals and should one adapt this search query to databases other than WoS.

#### *Distribution of Nano Research across Disciplines*

Nano research is incredibly widely distributed. Table S-3 (Supplemental Materials) lists the frequencies by Web of Science Categories (WCs), our means of operationalizing “disciplines” (albeit, one could consider them to be sub-disciplines, such as “physics, applied” – incidentally, just ahead of “Nano-WC” in frequency of nano-related hits). We find nano-related publications in almost all of the WCs (249 of 252). The top WCs in our comprehensive search for 1991-2017 are shown in Table 4.

**Table 4.** Nano Research Publications, by Leading Web of Science Categories (WCs), 1991-2017

WC	Number
Materials Science, Multidisciplinary	655,477
Physics, Applied	485,619
Nanoscience & Nanotechnology	471,835
Chemistry, Physical	376,562
Chemistry, Multidisciplinary	325,497
Physics, Condensed Matter	292,150
Engineering, Electrical & Electronic	202,289
Optics	134,129

Polymer Science	104,906
-----------------	---------

Note: Search conducted on August 8, 2018.

Table 4 includes the WCs with over 100,000 records. Each of the top six represents over 10% of the search total (2,235,233). However, recognize that many WoS sources are indexed in multiple WCs. The extent of that varies by research area, with some 40% of journals assigned to more than a single WC. [So, a tally of WCs would be expected to approximate 1.4 times the number of source hit counts.] However, nano research categorization indicates more extensive multidisciplinary – the tally of WC assignments for the 2.2 million records is 4,622,496, so these sources average ~2.1 WCs each.

The list of the top six WCs lends additional support to the sense of nano research being multidisciplinary, with two WCs including “Multidisciplinary” in their names and Nano-WC being especially multidisciplinary. Moreover, nano research appears in many other fields -- 47 WCs have over 10,000 nano records for 1991-2017. Extending further, 108 WCs show over 1,000 records; 184 WCs, over 100. Nano research is extremely diversified!

*Focusing on the recent decade, we sample every 5 years: 2007; 2012; 2017.* The purpose of this five-year sampling analysis is to further validate this search approach by examining categorizations of nano research and any notable shifts. Table 4 showed leading WCs for the full 1991-2017 period (Supplemental Table S-3 lists record counts for all WCs containing nano research, as based on our search). Note that there are two chemistry sub-fields included in Table 4. Here in Table 5, we consolidate sub-fields, based simply on names (inclusively; some WCs are included in multiple categories here). For instance, we search on “bio” in WC names, and tally the 17 WCs containing it, including “Materials Science, Biomaterials,” “Psychology, Biological” (present in 2007 & 2012), etc. The intent is to step back to consider major disciplinary groupings into which nano is dispersing, as shown in Table 5.

**Table 5.** Major Disciplinary Groupings of Nano Research for the Recent Decade

Disciplinary Group	Sub-fields Combined			Rank			Percentage			Nano Publications		Research
	2007	2012	2017	2007	2012	2017	2007	2012	2017	2007	2012	2017
<b>Chemistry</b>	13 <sup>12</sup>	13	13	2	1	1	36%	42%	44%	35,055	63,588	94,688
<b>Materials Science</b>	8	8	8	3	2	2	34%	37%	37%	33,471	55,865	80,345
<b>Physics</b>	11 <sup>13</sup>	11	11	1	3	3	39%	33%	28%	38,335	49,124	59,875
<b>Engineering</b>	18	18	18	5	5	5	25%	20%	19%	24,922	30,500	41,342
<b>Bio</b>	17	17	16	7	7	7	8%	8%	8%	7,386	11,642	17,012
<b>Nano-WC</b>	1	1	1	4	4	4	28%	24%	21%	28,346	36,070	45,906
<b>Other</b>	155	165	171	6	6	6	9%	11%	14%	9,012	16,384	30,490
<b>WOSC total#</b>	215	225	230							--		
<b>Nano Search Total #</b>						--				98,427	149,950	216,459

Table 5 shows disciplinary engagement over this recent decade, as the number of nano publications indexed by WoS more than doubled from 98,427 in 2007 to 216,459 in 2017. Materials Science and Biological nano publication shares hold quite steady over the decade. However, there are notable shifts – note that the share of publications associated to a degree with physics dropped from 39% to 28%, while that associated with chemistry rose from 36% to 44%. Engineering nano publication share declines as well. Also notable, the share of “Other” nano publications that fall outside these six groupings (the five disciplinary groups and Nano-WC) increased from 9% to

<sup>12</sup> To give a sense of the inclusion -- of those for 2017, 7 subfields are “Chemistry\*” for a total of 80,697; others are Electrochemistry (9958 records); Chemical Engineering (9436); Physics, Atomic, Molecular & Chemical (5665); Biochemistry & Molecular Biology (4440); Biochemical Research Methods (2558); and Geochemistry & Geophysics (236).

<sup>13</sup> Of those for 2017, 8 are “Physics, XX” for a total of 56,509; the additional three are Biophysics (3132), Geochemistry & Geophysics (236), and Astronomy & Astrophysics (217).



14%. Given the increase in total publications, this reflects a marked increase in publications – for the “Other” group from 9012 papers (of 98,427 in 2007) in 2007 to 30,490 (out of 216,459) in 2017. It seems that nano interests and capabilities are increasing and diffusing.

As stated, many individual WoS records are categorized in more than one WC (based on their publication source), so the tally of the 2017 records’ WCs considerably exceeds 216,459 (likewise, the percentages sum to more than 100%). Table 5 is meant to explore the extent to which nano-related research connects with various macro-scale disciplinary groupings. Nano research (based on our nano search in WoS for 2017—last column of Table 5) is most prominent in chemistry, materials science, physics, and nano-WC, with substantial engineering activity as well.

As noted, the “Other” for 2017 is also substantial – 30,490 publications. The Top 10 WCs with the most nano research in this category are listed here:

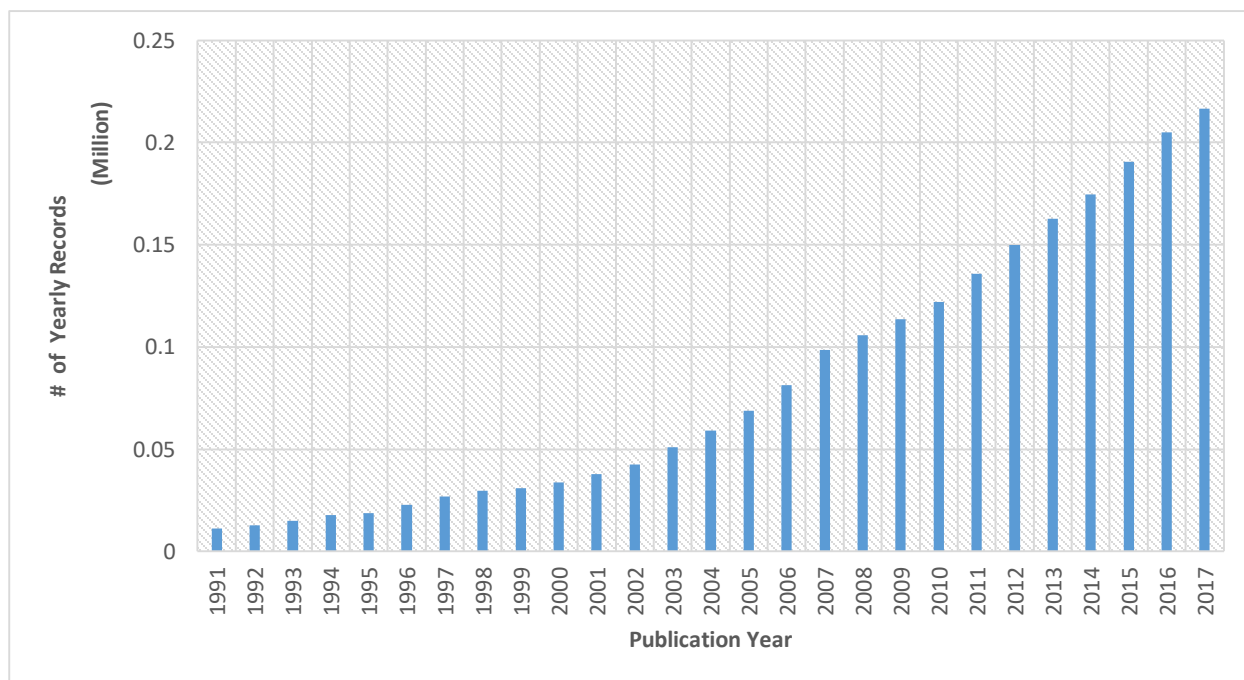
- Optics	10,095
- Polymer Science	9,778
- Energy & Fuels	7,798
- Multidisciplinary Sciences (e.g., papers in <i>Nature</i> )	6,790
- Pharmacology & Pharmacy	5,042
- Instruments & Instrumentation	4,552
- Environmental Sciences	3,584
- Mechanics	2,421
- Thermodynamics	2,078
- Crystallography	1,921
-	

This breakout of 2017 “Other” suggests that WCs could be further consolidated – e.g., optics could be grouped with other physical science or technology research.

In sum, the nano research presence is increasing in the life sciences relative to the physical sciences/technology and engineering.

### *Nano Publication Trend*

Figure 5 displays the growth rate for nano research. It shows continuing strong growth over 25 years or thereabouts. Nano publication has more than doubled in the recent time period (2007-2017).

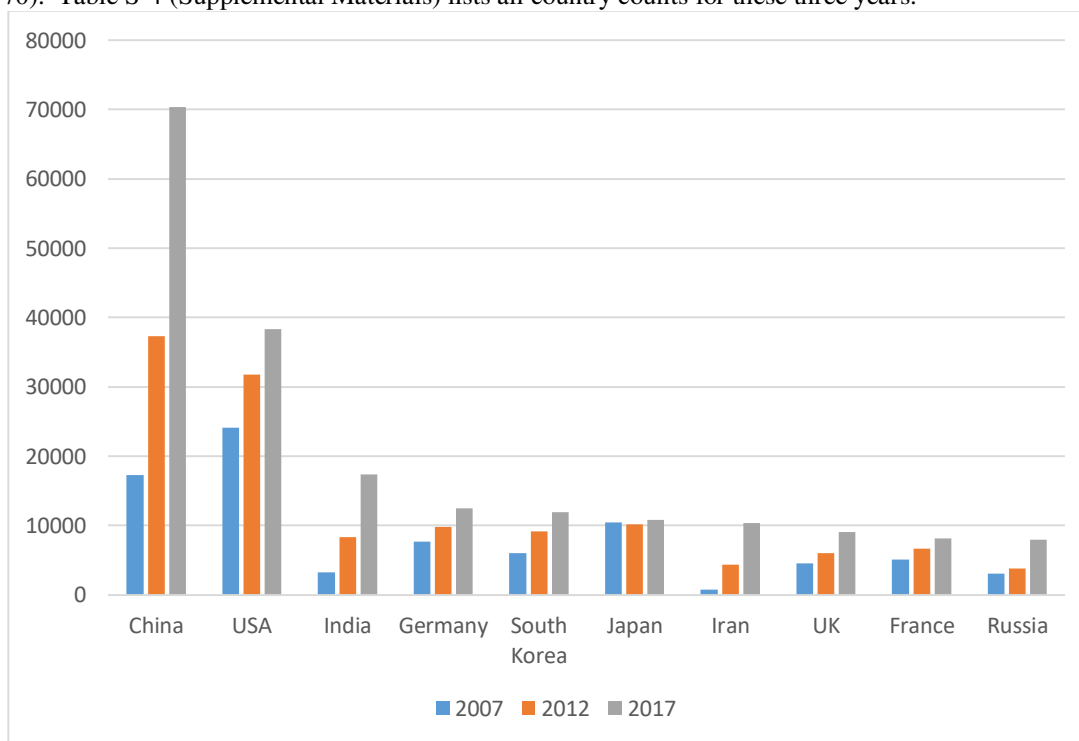


**Fig. 5. Nano Publication Trend**

[Based on August 8, 2018 comprehensive nano search (see Table 2) in Web of Science]

### *Nano Research: By Country*

Figure 6 shows the “Top 10” country publication counts for 2007, 2012, and 2017. China has the most publications, followed by the US and India. Although the number of publications associated with US authors grew by 59% over this time period, there are other intriguing national gainers. Especially notable are China, India, and Iran. Outside the Top 10, notable gainers from 2007 to 2017 include: Saudi Arabia (from 58 to 3496 publications indexed by WoS), Malaysia (from 176 to 3020), Turkey (from 588 to 2969), Pakistan (from 94 to 2100), and Indonesia (from 20 to 970). Table S-4 (Supplemental Materials) lists all country counts for these three years.



**Fig. 6.** Nano Research Publications Authored by the Top 10 Countries for 2007, 2012 & 2017

WoS provides a field of author affiliations. We clean these data to an extent, so publication counts per organization are only an estimate.<sup>14</sup> Tables S-5, S-6, and S-7 (Supplemental Materials) list the Top 200 nano-publishing organizations (with ties) for each of these three years. In all cases, counts reflect the number of publications that include at least one author or co-author giving that affiliation. Those are provided in case one wanted to search to see how active a particular research organization has been in nano publication. [They are not directly usable for comparison as they are not fully cleaned and mix individual (e.g., a university) with collections of institutions (e.g., Chinese Academy of Sciences).

### *Nano Emergence*

Previous efforts of this research group to develop search strategies for defining the nanotechnology research domain stopped short of using the search strategies to identify emerging research sub-domains. A contribution of this development of a more comprehensive nano search is to rectify this omission by applying a method for indicating

<sup>14</sup> Using *VantagePoint*'s list cleanup (fuzzy matching) routine, Organization Names.fuz (ignore department), results were checked for organizations showing over 400 nano publications for 2017. This was saved as a thesaurus and applied, as well, to the 2007 and 2012 data. So, cleaning is not fully vetted, but, for example, this consolidates 44,043 down to 34,074 organizations for 2017. Results should be treated as estimates.

trending topics in nanotechnology. We are not the first group to tackle the subject of identifying emerging topics in this domain. Pero (2013) offers a 2-dimensional framework to gauge emerging nanotechnologies. At a much more aggregate level than our emerging terms, they consider research publication and patenting, to focus on six sub-domains showing high growth: carbon nanotubes, polymer nanotubes, polymer particles, core shell, silica, and silver. Parsing into four quadrants, they note gold as more active in research; graphene and titania, more in patenting.

Our emergence analyses are much finer-grain. We ran *VantagePoint's* emergence indicator script on the new 2.2 million record nano dataset. The process begins by extracting cleaned and consolidated abstract and title NLP (Natural Language Processing) phrases. Carley et al. (2018) and Porter et al. (2018) describe the next steps in the indicator methodology that separates terms or phrases evidencing four attributes associated with emerging R&D terms – 1) term novelty, 2) persistence, and 3) strong recent growth, with 4) existence of a research community. Candidate terms had to meet established threshold filters for novelty, persistence, and community; three trend components were combined to emphasize recent acceleration in term usage in the abstract records (see Porter et al., 2018).

The emergence indicator calculation uses a three-year base period again which to compare a 7-year active period to derive an indicator value. Our analysis sought to emphasize the most recent terms over the 2013-2017 timeframe, so we used six-month increments, rather than whole years, to represent the timing of the base and active periods. The resulting emergence indicator is based on a pattern of increasing topic usage in the active period over that of the base period, not on review by nano researchers. Nevertheless, we consulted with two senior researchers at Georgia Tech in the management team of the Southeastern Nanotechnology Infrastructure Corridor (SENIC) to understand and interpret the trends underlying these emergent terms.

Table 6 indicates the high loading terms that consolidate to form each topic. 22 Emergent topics are named based on their high loading terms. The terms suggest that two main trends are prominent in emerging research areas: new designs for energy applications and two-dimensional materials.

**Table 6.** Highly Emerging Nano Terms with their EScore for 2013-2017

<b>Emergence Terms</b>	<b>Emergence Scores</b>
hydrogen evolution reaction (HER)	16.428
oxygen evolution reaction (OER)	14.362
sodium ion battery	10.409
lithium ion battery (LIB)	9.444
graphitic carbon nitride, g-C <sub>3</sub> N <sub>4</sub>	8.664
two-dimensional (2D) material	7.904
wearable electronics	7.614
carbon dot	4.269
transition metal dichalcogenide	3.872
lithium sulfur (Li-S) battery	3.391
shuttle effect	2.844
metal-air battery	3.377
additive manufacture	3.02
photothermal therapy (PTT)	2.99
PTB7-Th	2.973
tungsten(IV) selenide	2.503
non-fullerene acceptor	2.198
soft robot	2.158
high performance asymmetric supercapacitor	2.012
triboelectric nanogenerator (TENG)	1.826
black phosphorus (BP)	1.816
cellulose nanofiber (CNF)	1.809

The energy applications include: sodium ion batteries, lithium ion batteries, shuttle effect (for Li-S), metal-air battery, perovskite solar cells, PTB7-Th for solar cells, non-fullerene acceptors, triboelectric nanogenerators (Suominen et al., forthcoming), and cellulose nanofibers. Some of these applications, such as lithium ion batteries and

cellulose nanofibers (George and Sabapathi 2015), are old technologies which are benefitting from recent improvements. Others such as perovskite solar cells are applications that are more cost effective.

The emergent two dimensional materials include graphitic carbon nitride or g-C<sub>3</sub>N<sub>4</sub> (Ong et al. 2016), black phosphorus, transition metal dichalcogenide, and tungsten selenide. Efforts to identify, characterize, and develop these two-dimensional materials have been particularly prominent over the 2013-2017 timeframe. A subsequent exploration of the 2017-2018 period identified a further class of two-dimensional materials, MXene, as being on the rise in scholarly publications, although this term was not as widespread in the 2013-2017 period.

It is interesting that these research trends encompass relatively fewer devices outside the energy domain. Wearable electronics, soft robots, and additive manufacturing occupy the list as devices. But for the most part, rising interest continues to be in the materials area rather than macro integration in devices. Likewise, we do not see a profusion of nanobiotechnology applications on the list. Wearable electronics for monitoring metabolic processes and photothermal therapy are two that appear, and the battery designs could have medical applications. Nevertheless, the list has a notable lack of biological applications.

It is interesting that hydrogen evolution reaction and oxygen evolution reaction are so conspicuous on the list. It is possible that these reactions have become more prominent as a result of using nanocatalysts in the energy area and for generating two-dimensional materials, even though these reactions are not necessarily linked to the energy or materials domains.

## Discussion

### *Searching for Nano Publications for Nanoinformatics Purposes*

Returning to our original research questions, we do observe relative to the first research question sufficient dynamism in the nanotechnology domain to warrant a review of the search terminology. We present an updated and enhanced nano search formulation that retrieves 2.2 million records from Web of Science (WoS) for the period, 1991-2017. The text discusses and exemplifies issues in conforming the search query so as to retrieve many nano-related records (strong recall) without undue noise (reasonable precision). However, the dataset surely includes records that many observers would consider noise. Furthermore, there is no “right” formulation of exactly what constitutes nano research. While certainly imperfect, the results presented indicate that this record set provides a good representation of nano research publication suitable for further analyses of the field.

Table 2 presents our comprehensive nano search query (suitable for pasting into WoS advanced search), with record counts for the nine modules. Figures 3 and 4 show how many records are in common between particular search modules. Results support interpretation that the modules each add value without drifting far afield.

As discussed, a number of alternative nano search approaches are reasonable. We considered the argument that, perhaps, nano research has matured and its lexicon stabilized so that just searching for the Nano-WC would suffice. However, our results show that this only captures about 22% of the 2.2 million records of our comprehensive search, with limited coverage of many fields showing substantial nano activity (Figure 2). Such nano research not published in sources indexed in Nano-WC could well include some emerging areas, such as new materials, tools (microscopy), and applications. We see strong value in the 9-module nano search for nanoinformatics research.

We also contrast search results for this comprehensive query (2.2 million records) to a “nano\*” standalone query (1.5 million, prior to cleaning). The new comprehensive search augments nano\* by over 50%. Again, this supports use of the more encompassing search for most nanoinformatic purposes.

We hope that others pursuing nanoinformatics will consider use or adaptation of this comprehensive nano search. Unlike our prior search (Arora et al. 2013), this formulation takes advantage of database particulars to enhance search performance. As discussed, we experimented rather extensively with a module to exclude papers appearing in apparently non-nano WCs, but determined that unduly constrained the search. Our current screening out of “nano noise” through excluding records that include certain nano\* variants accomplishes this quite well.

Module 9 retrieves records based on WoS categorization of the source in which the record is published – “Nano-WC.” As discussed, this substantially augments our otherwise term-based searching to capture likely nano research unanchored by use of particular terminology. Including Module 9 contributes some 8% of the 2.2 million nano publications. We present a case that combining a module to retrieve Nano-WC, together with the other modules offers a strong base for nanoinformatic analyses.

A cost to inclusion of the Nano-WC search module is that application of the search to other databases e.g., *Scopus*, *Inspec*, *PatStat* (possibly combining lexical with International Patent Class information) or *MEDLINE* (tailoring with MeSH categories) requires revision of the query. Such application would require tuning, anyway, as search interfaces

differ. We retain Module 8, source (journal) based search, in reduced form (compared to Arora et al. 2013), to facilitate such query adaptation for other data resources or analytical aims. Namely, one could extend the search's source listing to supplant Module 9 in searching elsewhere than in WoS. In, 2019, we are adapting the search to retrieve nanopatents in *Derwent Innovation Index (DII)*. We apply Modules 1-7 in the DII topic search (TS) field and add a module to capture patent families using the nanotechnology patent class code B82Y.

For those planning a nano search, downloading from WoS poses challenges. WoS, under its licenses, imposes limits on record downloading. But for many purposes, sub-datasets should be more suitable. We envision ourselves or others adapting this nano search strategy to such ends. For example, within WoS, one might compose the nano search (Table 2) using Advanced Search – call it search “S3.” Then one could generate particular search intersections -- e.g., S3 with “bio” WCs; then download just those records. Or, one could focus on particular nano elements, drawing terms, along with contingencies or such, from Table 2 selectively.

Furthermore, manipulating (cleaning, consolidating, subdividing, analyzing, visualizing) 2.2 million records under the Windows operating system poses additional challenges. To generate the results presented herein, where viable, we operated on reduced datasets – e.g., three individual years of data to consider change over the recent decade. For analyses that demanded the full span of 1991-2017, we took steps to expedite processing.<sup>15</sup>

### *Substantive Findings: Nano Research Patterns*

As to our second research question, we observe that nano research does not show signs of “topping out” (an S-shaped growth curve maturing). Nano publication continues to increase approximately linearly, and strongly. Nano-related publications indexed by WoS more than double from 2007 to 2017. The overall trend may be exaggerated somewhat as WoS coverage by the eight Core Collection databases that we searched has expanded over time and overall annual publication counts in WoS are increasing, although not so strongly.

Tables 5 shows some important shifts among fields in which nano research appears over the past decade. Again, expanding WoS coverage could contribute to this, but we don't think that is a major factor. As per Table 5, within the physical sciences, chemistry's share is up, while physics and engineering are down (but still very strong). As discussed, this does not reflect a decline in physical science numbers of papers, just a relative shift in share of all nano-related publications.

Who are the leading “players” in nano research? Profiling nano research publication generally shows marked growth by a number of Asian and Middle-Eastern countries – but not Japan (Figure 6, with Table S-4). We point particularly to China, India, and Iran among the leaders showing greatest growth, plus Saudi Arabia, Malaysia, Turkey, Pakistan, and Indonesia. US-authored publications continue to rise, but perhaps reflecting a decline in nanotechnology R&D funding in the US, the slope is flattening compared to the overall publication trends.

The third research question is addressed in the presentation in Table 6 of emerging nano research terms, deriving from calculations of emergence indicators based on strong publication growth rate over the prior 10 half-year periods, along with a process of threshold filtering, as discussed. Energy applications and two-dimensional materials dominate the list. These emerging nano topics are offered for consideration as researchers and research managers plan their nano research. We postulate these will tend to be strongly represented in research publications of the coming two years.

**Acknowledgements** Dr. David Gottfried, Associate Director of the Southeastern Nanotechnology Infrastructure Corridor (SENIC) at Georgia Tech, provided vital nuanced guidance reviewing our tentative decisions regarding nano search modifications and interpretation of the emerging terms with high EScores. Professor Oliver Brand, Executive Director of SENIC, also provided guidance with interpreting these emerging terms. Other Georgia Tech researchers, including Prof. Z.L. Wang and an ECE PhD student Lei Yan, aided us in understanding nano sub-fields.

### **Compliance with ethical standards**

**Funding** We acknowledge support from the US National Science Foundation under Award # 1542174 (NNCI: Southeastern Nanotechnology Infrastructure Corridor), and under Award #1759960 (Indicators of Technological Emergence). 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.

---

<sup>15</sup> We opened just the necessary fields in the *VantagePoint* data file; minimized memory requirements; and opted for computers with relatively high performance. Some analyses still required hours.

**Conflict of interest** The authors declare that they have no conflict of interest.

**Research involving Human Participants and/or Animals** This article does not contain any studies with human or animal subjects.

**Informed consent** None.

## References

- Arora SK, Porter AL, Youtie J, Shapira P (2013) Capturing new developments in an emerging technology: an updated search strategy for identifying nanotechnology research outputs. *Scientometrics* 95(1):351–370. doi:[10.1007/s11192-012-0903-6](https://doi.org/10.1007/s11192-012-0903-6)
- Arora S, Youtie J, Carley S, Porter A, Shapira P (2014) Measuring the development of a common scientific lexicon in nanotechnology. *J Nanopart Res* 16(1), 2194; DOI:10.1007/s11051-013-2194-0
- Bartol T, Stopar K (2015) Nano language and distribution of article title terms according to power laws. *Scientometrics* (2015) 103:435–451. doi: 10.1007/s11192-015-1546-1
- Carley SF, Newman NC, Porter AL, Garner J (2018) An indicator of technical emergence. *Scientometrics* 115 (1): 35-49; <http://link.springer.com/article/10.1007/s11192-018-2654-5>.
- Chen H, Roco MC, Son J, Jiang S, Larson CA, Gao Q (2013) Global nanotechnology development from 1991 to 2012: patents, scientific publications, and effect of NSF funding. *J Nanopart Res* 15:1–21
- Darby MR, Zucker LG (2003) Grilichesian breakthroughs: Inventions of methods of inventing and firm entry into nanotechnology. NBER Working Paper, No. 9825
- Dong, H., Gao, Y., Sinko, P. J., Wu, Z., Xu, J., & Jia, L. (2016). The nanotechnology race between China and the United States. *Nano Today*, 11(1), 7-12.
- George, J., & Sabapathi, S. N. (2015). Cellulose nanocrystals: synthesis, functional properties, and applications. *Nanotechnology, science and applications*, 8, 45.
- Glänzel W, Meyer M (2003) Patents cited in the scientific literature: an exploratory study of “reverse” citation relations. *Scientometrics* 58 (2): 415-428
- Gläser, J. & Velarde, K.S. (2018) Changing Funding Arrangements and the Production of Scientific Knowledge, *Minerva* (56), 1. <https://doi.org/10.1007/s11024-018-9344-6A>
- Grieneisen, ML, 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. *J Technology Transfer* 36(2): 145–172
- Huang, Y., Schuehle, J., Porter, A. L., & Youtie, J. (2015). A systematic method to create search strategies for emerging technologies based on the Web of Science: illustrated for ‘Big Data’. *Scientometrics*, 105(3), 2005-2022.
- Jo, H, Park Y, Kim SE, Lee H (2016) Exploring the intellectual structure of nanoscience and nanotechnology: journal citation network analysis. *J Nanopart Res* 18: 167. doi: 10.1007/s11051-016-3473-3
- Levin L, Jensen P, Kreimer P (2016) Does size matter? The multipolar international landscape of nanoscience. *PLoS ONE* 11(12): e0166914. doi:10.1371/journal.pone.0166914
- Milojević S (2012) Multidisciplinary cognitive content of nanoscience and nanotechnology. *J Nanopart Res* 14(1):685. doi:10.1007/s11051-011-0685-4
- Muñoz-Écija T, Vargas-Quesada B, Chinchilla-Rodríguez Z (2017) Identification and visualization of the intellectual structure and the main research lines in nanoscience and nanotechnology at the worldwide level. *J Nanopart Res* 19:62. doi: 10.1007/s11051-016-3732-3 24
- National Academies (2019). Quadrennial Review of the National Nanotechnology Initiative, <https://www8.nationalacademies.org/pa/projectview.aspx?key=51396>
- Noyons ECM, Buter RK, van Raan AFJ, Schmoch U, Heinze T, Hinze S, Rangnow R (2003) Mapping excellence in science and technology across Europe. Final Report of Project EC-PPLS CT-2002-

- 0001 to the European Commission  
<http://mydigitalart.webs.com/documents/Mapping%20Excellence%20in%20Science%20and%20Technology%20across%20Europe%20-%20Life%20Sciences.pdf>
- Ong, W. J., Tan, L. L., Ng, Y. H., Yong, S. T., & Chai, S. P. (2016). Graphitic carbon nitride (g-C<sub>3</sub>N<sub>4</sub>)-based photocatalysts for artificial photosynthesis and environmental remediation: are we a step closer to achieving sustainability?. *Chemical reviews*, 116(12), 7159-7329.
- Pero MH (2013) Identifying emerging technologies: An application to nanotechnology. 24<sup>th</sup> International Society for Professional Innovation Management (ISPIM) Conference, Helsinki, Finland
- Porter AL, Garner J, Carley SF, Newman NC (2018) Emergence scoring to identify frontier R&D topics and key players, *Technological Forecasting and Social Change*,  
<https://doi.org/10.1016/j.techfore.2018.04.016>
- Porter AL, Youtie J, Shapira P, Schoeneck DJ (2008) Refining Search Terms for Nanotechnology. *J Nanopart Res* 10(5):715-728
- Rogers, J. D., Youtie, J., & Kay, L. (2012). Program-level assessment of research centers: Contribution of Nanoscale Science and Engineering Centers to US Nanotechnology National Initiative goals. *Research Evaluation*, 21(5), 368-380.
- Schmoch U, Thielmann A (2012) Cyclical long-term development of complex technologies - premature expectations in nanotechnology?". *Research Evaluation* 21 (2): 126-135
- Stopar K, Drobne D, Eler K, Bartol T (2016) Citation analysis and mapping of nanoscience and nanotechnology: identifying the scope and interdisciplinarity of research. *Scientometrics* 106:563–581. Doi: 10.1007/s11192-015-1797-x
- Suominen A, Li Y, Youtie, J, Shapira P (2016) A bibliometric analysis of the development of next generation active nanotechnologies. *J Nanopart Res* 18(9); doi: 10.1007/s11051-016-3578-8
- Suominen, A., Peng, H., and Ranaei, S. (forthcoming), Examining the dynamics of an emerging research network using the case of triboelectric nanogenerators, *Technological Forecasting and Social Change*.
- Wang L, Notten A, Surpatean A (2013) Interdisciplinarity of nano research fields: a keyword mining approach. *Scientometrics* 94:877–892. doi: 10.1007/s11192-012-0856-9
- Zhu H, Jiang S, Chen H, Roco MC (2017) International perspective on nanotechnology papers, patents, and NSF awards (2000-2016). *J Nanopart Res* 19: 370-; doi.org/10.1007/s11051-017-4056-7
- Zucker LG, Darby MR (2007) Nanobank Data Description release 2.0 (beta-test). UCLA Center for International Science, Technology and Cultural Policy and Nanobank, Los Angeles, 17 Jan 2007–2 Feb 2009
- Zucker LG, Darby MR, Furner J, Liu RC, Ma H (2007) Minerva unbound: knowledge stocks, knowledge flows and new knowledge production. *Res Policy* 36(6):850–863

## Appendices

### Appendix A.

Table A1 lists a number of papers pertaining to nano searching and citations that they have accrued.

**Table A1.** Citations to Nano-search Related Papers

Authors	Published	Cites	Cites/Yr
Arora et al.	2013	56	11.2
Bartol & Stopar	2015	3	1.0
Chen et al.	2013	11	2.2
Darby & Zucker	2003	--	--
Glänzel & Meyer	2003	42	2.8
Grieneisen & Zhang	2011	20	3.3
Huang et al.	2010	65	8.1
Jo et al.	2016	4	2.0
Levin et al.	2016	0	0.0

Muñoz-Écija et al.	2017	1	1.0
Noyons et al.	2003	--	--
Pero	2013	0	0.0
Porter et al.	2008	197	19.7
Schmoch & Thielmann	2012	0	0.0
Stopar et al.	2016	1	0.5
Wang et al.	2013	13	2.6
Zhu et al.	2017	0	0.0
Zucker & Darby	2007	--	--
Zucker et al.	2007	96	8.7

Note: Search in the 8 WoS Core Collection databases on July 13, 2018. Since that reflects ~1/2 of 2018, we subtract publication year from 2018 in estimating cites/year. “—” indicates not found for these non-journal publications.

## Appendix B. An In-depth Example: Assessing whether to add certain emerging nano research terms

This section details one case segment. The main purpose is to give a feel to the nature of our testing processes.

Given our interest in emerging nano research topics, we did another “bootstrap” in which we generated a set of 142 emergent terms using *VantagePoint*’s emergence indicator script on the recent core nano download but to apply to the full 1991-2017 time period. Most were either already covered by the core nano search phrasing or were too general for a nano search (e.g., “energy storage”). Six terms were then screened, as detailed here. Our intent is really to illustrate our screening process that was used to revise the nano search by assessing ~135 candidate terms as noted.

The six candidate terms, based on accelerating research attention on them, as meeting our criteria to be categorized as “emerging terms” (Carley et al. 2018; Porter et al. 2018) are:

- Self-healing
- Molecularly imprinted polymers
- Metal-organic framework
- Hexagonal boron nitride
- Molecular docking
- Transition metal dichalcogenide

“Self-healing” has uses outside the realm of nano-oriented materials – e.g., in psychology. In our checking procedure, we now searched WoS – 8 core databases, not including the additional Chemical Indexes, for 1991-2017 [searches conducted as of 7/9/2018 or thereabouts]. Of 9,123 records, 3,522 are in the MolEnv-I set too, after excluding any in “nano\*.” Given concerns about other uses of the term, we focused down on the 826 in the MolEnv-R set, after excluding any in nano\*. We sampled the first (most recent) 20 abstracts and would only classify ~3 as nano-related. The criterion that the records be nano-related, while not appearing in the “nano\*” search set, is stiff (see Footnote 7); nonetheless, 3 of 20 is low. So we do not add “self-healing” to the search algorithm.

“Molecularly imprinted polymers” was similar to a term already being added to the search. This review led us to tune the search string to: “molecular\* imprint\*”.

“Metal–organic frameworks” – quoting Wikipedia: “(MOFs) are compounds consisting of metal ions or clusters coordinated to organic ligands to form one-, two-, or three-dimensional structures.” We separately searched on “metal-organic framework\*”, and on MOF, and on MOFs. We found other uses of the acronyms, but most records containing them also contained the full phrase, suggesting they are suitable. For instance, of 8,583 records with WOFs, only 10% were not also in the “metal-organic framework\*” set. We considered the full captured set (33,936 records for metal-organic frameworks; reduced to 24,333 by excluding any in “nano\*”). We also examined reduced sets contingent on MolEnv-I, not nano\* (20,516); and another on MolEnv-R, not nano\* (1,903). Examination of a 20-record sample of MolEnv-R, not nano\* judged 11/20 as nano-related. We determined to use this, rather than risk considerably remote possibilities were we to use the much larger MolEnv-I or unrestricted sets. Including the acronyms in the search string yielded a moderate 1,992 records. Some of those would already be captured by other modular terms in our search, even though they are not in the nano\* set.

Similar screening led to addition of “hexagonal boron nitride” and its acronyms, h-BN and hBN to the search, contingent on co-occurring in the MolEnv-I set. We thereby get 2,277 records not already captured in the nano\* set.



The acronyms do reflect some other uses. And the records reach into tribology, various thin film aspects, semiconductor development, etc., but are generally highly nano-related -- given the exclusion of records already captured by our nano\* search. Different searchers could reach other decisions on what to include here.

“Molecular docking” retrieved 15,331 records, of which 14,711 fall outside the nano\* records set. So little overlap with nano\* results posed doubts about nano-relevancy. Examination of a subset of 513 in MolEnv-I, not nano\*, found a lot of quantum chemistry, but not much that seemed directly nano. We chose not to include this term in the nano search algorithm.

And lastly of the six candidates, “transition metal dichalcogenide” monolayers were incorporated into the search, but the TMD acronym was not, based on too many other uses. The MolEnv-I contingency made little difference – reducing 5,967 to 5,077 records. We examined 20 abstracts from the set not also in the MolEnv-R contingency group, deeming ~14/20 as nano. Based on that, we include the term without contingencies.