J Nanopart Res (2014) 16:2194
DOI 10.1007/s11051-013-2194-0

PERSPECTIVES

Measuring the development of a common scientific lexicon in nanotechnology

Sanjay K. Arora · Jan Youtie · Stephen Carley · Alan L. Porter · Philip Shapira

Received: 26 June 2013 / Accepted: 5 December 2013 / Published online: 24 December 2013

Abstract Over the last two decades, nanotechnology has not only grown considerably but also evolved in its use of scientific terminology. This paper examines the growth in nano-prefixed terms in a corpus of nanotechnology scholarly publications over a 21-year time period. The percentage of publications using a nano-prefixed term has increased from <10 % in the early 1990s to nearly 80 % by 2010. A co-word analysis of nano-prefixed terms indicates that the network of these terms has moved from being densely organized around a few common nano-prefixed terms such as "nanostructure" in 2000 to becoming less dense and more differentiated in using additional nano-prefixed terms while continuing to coalesce around the common nano-prefixed terms by 2010. We further observe that the share of nanotechnology papers oriented toward biomedical and clinical medicine applications has risen from just over 5 % to more than 11 %. While these results cannot fully distinguish between the use of nano-prefixed terms in response to broader policy or societal influences, they do suggest that there are intellectual and scientific underpinnings to the growth of a collectively shared vocabulary. We consider whether our findings signify the maturation of a scientific field and the extent to which this denotes the emergence of a shared scientific understanding regarding nanotechnology.

Keywords Scientific lexicon · Bibliometrics · Science and technology policy

S. K. Arora (✉) · S. Carley · A. L. Porter · P. Shapira
School of Public Policy, Georgia Institute of Technology, Atlanta, GA 30332-0345, USA
e-mail: sanjayk.arora@gatech.edu

J. Youtie
Enterprise Innovation Institute, Georgia Institute of Technology, Atlanta, GA 30308, USA

P. Shapira
Manchester Institute of Innovation Research, Manchester Business School, University of Manchester, Manchester M13 9PL, UK

Introduction

Growth in nanotechnology publication and patent production continues apace (Roco 2010; Arora et al. 2012). Yet debate persists about the extent to which terminology in nanotechnology is representative of a distinct set of scientific phenomena or whether it is the result of more sociological processes. On one hand, Kuhn (1996) posits that, in "normal science," research generates a collectively shared and esoteric vocabulary that helps scientists describe phenomena and diffuse new insights. As fields develop, the trend toward increasing specialization produces an evermore specialized and fragmented vocabulary. On the other hand, some science and technology scholars maintain that nano-prefixed terminology represents a "catchphrase" for a set of political, social, and economic dynamics which inform yet are still distinct from the intellectual pursuit of knowledge. The debate is not only about the maturation of the nanotechnology domain but, as we argue, also about the maturation of the terminology used to distinguish the domain.

This paper seeks to contribute to this debate by providing bibliometric evidence of the growing use of nano-prefixed terminology in scientific articles. Although we cannot ascertain nanotechnology researchers' motivations for using nanotechnology and related terms from this bibliometric approach, our results do suggest that terms beginning with a nano prefix are used in an increasing share of research articles' title and abstract content. Moreover, these nano-prefix terms are becoming more differentiated and application-oriented.

Literature review

The debate about the nomenclature of nanotechnology begins from questions about the definition of nanotechnology. Although nanotechnology has a specific definition related to size, novelty, and engineering and manipulation (NSET 2007), intricacies in the definition occur. The classic work by Wood et al. (2003) highlights these intricacies by acknowledging that field definitions based on scientific size boundaries have blurred lines in part because of scale variability in the materials themselves. The authors go on to underscore the importance of taking sociological as well as scientific considerations into account in defining nanotechnology. This perspective is echoed by Grieneisen and Zhang (2011) who observe that definitions of nanotechnology are evolving from size-dependent definitions to characterizations that enable journal editor flexibility in accepting articles that do not specifically mention size dimensions (see also Braun et al. (2007) on the gatekeeping function of editors of nanotechnology journals).

This debate is furthered through a set of articles concerning the extent to which nanotechnology relates to something distinctive, or whether it is a renaming of existing scholarly work. The article by Macnaghten et al. (2005) about social values embedded in nanotechnology raises issues as to whether the science behind nanotechnology is really distinctly novel compared to previous scientific work. Likewise, Khushf (2004), whose research provides evidence for convergence of technologies at the nanoscale, also acknowledges that some scientists may be using nano-prefixed terms to obtain funding for research that in the past has or could have received support from established grant programs. Restated, scientists may be relabeling their research to make their work more attractive to the contemporary funding environment. More recently, Hodge and colleagues point toward the potential for rebranding research as nanotechnology and the implications of this rebranding for discussions about risk and regulatory approaches (Hodges et al. 2013). Other authors wonder if there is a "nano-hype" in overselling the distinctiveness of nanotechnology, including its terminology. The use of the "nano-hype" phraseology has occurred in both a discourse-theoretical context, such as the Wullweber (2008) article that describes nanotechnology as an "empty signifier," as well as in a bibliometric context, such as the Meyer (2007) article which shows little integration between various nanotechnology subfields (see Porter and Youtie 2009a, b for an alternative finding showing increased integration of nanotechnology articles). In sum, this set of works questions the validity of terminology associated with nanotechnology as being representative of a new wave of scientific inquiry but instead sees it as a manifestation of societal factors.

Perspectives that maintain and describe the distinctive scientific phenomenon behind various nano-prefixed terms are in contrast. For example, even as they suggest that it is difficult to distinguish terms such as nanobiotechnology and bionanotechnology, Baneyx and Park (2013) go on to provide the scientific underpinning for differences between these two terms. They indicate that nanobiotechnology refers to tools and devices for biotechnology research, while bionanotechnology reflects self-assembly of biological matter for nano-enabled products. Grieneisen (2010) shows that the number of journals with nanotechnology in their title has grown exponentially, especially after 1997; more than 90 % of these journals are not open access, meaning that they may use peer review to ensure that articles are scientifically relevant to the nanotechnology domain and use appropriate methodologies. The article further observes that more recent journals with nanotechnology in their title use more qualifying and modifying terms alongside the nano-prefixed field term and demonstrate greater levels of specialization. An update to this work demonstrates that the ever-increasing ranks of nanotechnology journals also have higher citation rates (Grieneisen and Zhang 2012). Broadly concluding from the initiatives of Nano2's workshops and synthesis of nanotechnology scientific developments and applications (Roco et al. 2011), Hersm (2011) describes nanotechnology as moving into a "posthype era" in which its applications are widely pervasive throughout scientific and industrial sectors.

While the evidence is mixed as to whether the growth of nano-prefixed terms arises out of sociological drivers or the emergence of distinctive scientific phenomena (or both), many studies continue to use nano-prefixed and related content in bibliometric analyses of patents and publications to understand the dynamics of nanotechnology research and innovation systems. The set of works that use bibliometric content analysis to draw implications about the nanotechnology scientific enterprise is specific to the approach taken in this paper. Huang et al. (2003) were among the first to use nano-prefixed terms to delineate trends in the field. Shapira et al. (2010) indicate that scientometric works on nanotechnology constitute an increasing share of citations in social scientists' articles about the domain: social science studies about nanotechnology terminology attract increasing attention by other social scientists, suggesting that the terms themselves are relevant for the field. Wang et al. (2013) use a keyword mining technique to demonstrate that nanotechnology-related keywords in five different fields are becoming more diverse and interacting across the five fields. Cacciatorie et al. (2012) find growth in the use of the terms nanotechnology and nanotech (excluding terms such as iPod, Tata, and MP3) in social media, though not in print media.

This second set of studies suggests that there is validity in analyzing patterns of development of nano-prefixed content for understanding trends in scientific research. Our work begins from this point, drawing on a tradition of co-word analysis based on the notion that authors choose words from a limited vocabulary, that they use different terms to describe different phenomena and concepts, and that this type of understanding is shared and implemented in similar ways across a research community (DeBellis 2009). In this vein, Watts and Porter (1997) contend that shifts in the specialization of vocabulary used by scientists and practitioners present one way to measure the maturation of the underlying science and related technologies.

Data and methods

To examine how nano-prefixed terminology has changed over time, we use article metadata extracted from the Science Citation Index through a multi-term and multi-stage Boolean search strategy. The search terms in this strategy include both nano-prefixed terms and a large array of non-nano-prefixed terms indicated to be of relevance to nanotechnology research based on expert interviews and surveys, a review of search terms used in prior bibliometric research, an application of precision and recall measures, and a comparison of "hits" from the various search strategies (Porter et al. 2008; Arora et al. 2012). This approach results in a nanotechnology domain composed of more than 800,000 publications and was tested by Huang et al. (2010), who found that the search strategy was comparable in results to several other bibliometric approaches to defining nanotechnology.

This paper furthers research about nanotechnology terminology by examining two research propositions. The first is that nano-prefixed terms capture an increasing share of research publications in the domain. The second is that nano-prefixed terms demonstrate greater specialization and applied research orientation in recent years. For this work, we compare publications in the nanotechnology domain which include nano-prefixed terms in their metadata (of which there are approximately 20,000 such terms) and those which similarly include terms relevant to the domain that lack a nano prefix (of which there are approximately 80 such terms). The nano-prefixed terms exclude records that just reference nanometer, nanosecond, nanogram, and the like or compounds involving NaNO. Terms relevant to the domain without a nano prefix include nanoscale materials (such as graphene and fullerenes), instruments, processes (such as self-assembly and microscopy), and a set of other terms that refer to the molecular environment.

To isolate the population of records containing "nano" in the corpus, we search for this term in the abstract, title, and keyword metadata fields (collectively referred to herein as "keywords") of scholarly publications within VantagePoint[1], a software application for the analysis of structured and unstructured text. We find that, after 1991, greater than 90 % of articles have

[1] http://www.thevantagepoint.com/.

abstracts and 100 % have titles. The percent of articles with keywords (provided by the Web of Science and/or by authors) grows from 82 to 94 % between 1991 and 2010. We then remove various punctuation characters and proceed to count the number of terms containing "nano" across all 21 publication years. The unit of observation is keyword occurrence, while the unit of analysis is the keyword itself.

We use three approaches to explore the growth of nano-prefixed terminology. The first is to look at the share of articles captured by nano-prefixed terms by year. Our second approach is related to the first and makes use of "keyword groups" of nano-prefixed terms. Keyword groups aggregate similar words by employing a series of thesauri to map keyword instances to a meta-level category. For example, we combine instances of "nanotubes-doping," "nanotubes-doped," "nanotube-doped," and "doped-nanotube" under a common term, "doped-nanotube," given that these collectively embody the same concept. This method reduces redundancy and noise that comes with working with large data sets such as ours. We also collapse on chemical compounds (both in short-form notation and spelled out) and on term inflections to further reduce variability. These steps are implemented as part of a multi-phased cleaning process called term clumping (O'Brien et al. 2013). This process results in 13,499 term groups.

Next, we map each article (through its unique identifier) to its set of keyword groups and produce a co-occurrence list of keyword groups found within the same article. For example, a given article may contain one or more nano-prefixed terms such that each of the corresponding keyword groups are linked in a co-occurrence relationship. We present three co-occurrence networks from 2000, 2005, and 2010 to show trends in the evolution of keyword connections over time. The maps are created in VOSviewer, a freely accessible software package for visualizing bibliometric data (van Eck and Waltman 2010). VOSviewer implements a mapping algorithm, VOS, which seeks to minimize the Euclidean distance between similar elements, in this case keyword groups, in a graph.[2]

In addition, we provide year-by-year measures of two common network topography measures: density and average (global) clustering coefficient. These measures, along with the maps, help convey how individual nano-prefixed terms attach to the larger network year-over-year. Network density, D_t, is defined as:

$$D_t = \frac{\alpha_t}{n_t(n_t - 1)/2}$$

where t is a year, α is the number of observed edges between keyword groups, and n is the number of keyword groups (Borgatti et al. 2013). Network density is a measure of the extent to which a set of terms (in this case, nano-prefixed terms) co-occur. If the network is denser, each term co-occurs with every other term but if the network is less dense, that can be an indicator of the introduction of new and, we posit, more differentiated terms. The network's average clustering coefficient C_t is the average of the sum of each keyword group v's local clustering coefficient $C_{v,t}$ (c.f. Watts and Strogatz 1998) where t is a publication year. The local clustering coefficient of a keyword group v is higher when the term groups $\{j,k, v \neq j,k\}$ with which v co-occurs also co-occur with one another in that same year. In other words, the local clustering coefficient provides a measure of whether a local network is more dense (e.g., shared or common) or less dense (e.g., specialized or differentiated). These measures are implemented in Gephi (Bastian et al. 2009), an open-source graph visualization and manipulation software tool.

The third approach measures whether articles are becoming more application-oriented. In this analysis we focus on a case example of biomedical and medicine applications. We perform this analysis by reporting on results for three select meta-disciplines (Rafols et al. 2010), biomedical sciences and clinical medicine (to represent biomedical and medicine applications) and materials science (to represent general materials investigators). Within the caveat that some materials science may be application-oriented while some biomedical and medicine research may be knowledge- or subject-driven matter rather than application-driven, the aim of this analysis is to juxtapose basic materials science publications with more applied biomedical- and clinical-related research over time. We present these findings from 2000 to 2010 inclusive to illustrate recent trends since

[2] VOS is subject to the constraint that the average distance between elements $\{i, j, i < j\}$ is equal to one (van Eck et al., 2010). Similarity is defined as the "association strength" of the co-occurrence relationship between $\{i, j\}$ weighted by how often $\{i, j\}$ co-occur with all keyword groups in the publication year.

the establishment of national nanotechnology initiatives in various countries.

Results

The first set of findings is based on year-to-year comparisons of publications that match "nano*" vis-à-vis records matched exclusively by seven other components of the search strategy. This approach allows us to capture at an aggregate level the influence of an expanding scientific vocabulary in defining and bounding a field of inquiry over two decades. The second set of findings reveals time-varying structural features of keyword group co-occurrence networks. The final set of findings includes a series of analyses which suggest the emergence of terms with increasing levels of specificity and applied research foci drawing on the case of biomedical and clinical medicine applications.

The persistent upward rise over time of nanotechnology records has been noted in many bibliometric studies on nanotechnology (e.g., see Arora et al. 2012). Figure 1 shows a steep growth curve in total publications, especially since 2001–2002, but a smooth, exponential-like increase in the absolute number of publications matching "nano*" in abstracts, titles, and keywords is also apparent. When inspecting the percent of articles matching "nano*" in the entire corpus, we see significant increases from 1993 to 2005, with a slightly higher rate of growth around 1995. After 2005, the share of publications matching "nano*" increases, but at a decreasing rate. From 2009 to 2010 the absolute increase in the percent of articles matching "nano*" is 1.5 %, the smallest difference between any two years since 1991–1992. These results suggest that most recent publication records identified by the search strategy outlined in Arora et al. (2012) contain a prominently featured nano-prefixed word in the title, abstract, or keyword fields. It may be the case that the same nano-prefixed terms account for the majority of the increase in articles captured by the nano* search query year-over-year. If this were true, we might expect to see (a) low growth in the number of term groups appearing year-over-year and (b) a disproportionate number of records identified by only the top echelon of term groups (vis-à-vis less frequent nano*-prefixed terms). Table 1 shows that this is not the case. The growth rate of new term groups increases by double digits for most years between 1991 and 2008, and three continuous time periods exhibit growth rates exceeding 20 % per year: 1991–1995, 1997–1998, and 2001–2005. In addition, there exists a sizeable increase in the number of records not identified exclusively by top echelon terms (i.e., those term groups whose frequency is in the top 20 on a yearly basis). As the number of publications grows throughout the sample time frame, records matching top 20 and non-top 20 keyword groups increase substantially. For example, in 2000, 8,516 articles matched non-top 20 (keyword group) terms; by 2010, this increased to 57,908, an increase of over 580 %. Unsurprisingly, the number of articles matching the top 20 terms also increases substantially in the same time frame, from 6,891 to 46,112 (569 %). However, if the overall growth in articles matching nano*-prefixed terms were attributable to only commonly occurring keyword groups, the vast majority of articles would match exclusively in the domain of the top echelon keyword groups. This pattern suggests that as agreement on the top terms (proxied by the percentage of articles' metadata incorporating these terms) rises, so too does the number of keyword groups, suggesting ever-greater differentiation of terms in the nanotechnology research domain.

To test this proposition more rigorously, we turn to three network maps based on co-word analysis, i.e., co-occurring keyword groups in 2000, 2005, and 2010, to assess how new keyword groups attach to the established keyword groups (Figs. 2, 3, and 4). As noted above, VOSviewer uses the VOS mapping algorithm, which has been shown to evenly distribute co-occurring elements across a graph without imposing an artificial structure (van Eck et al. 2010). The map in 2000 shows 262 keyword groups located in a dense core of some well-known "top 20" terms, including nanocrystalline, nanostructure, and nanotube. In 2005, this core group of well-established terms moves toward the periphery to accommodate a burgeoning set of new nano-prefixed keywords (a total of 890 keyword groups); by 2010 this trend is even more apparent (with a total of 1,447 keyword groups). The visualizations indicate that, while keyword groups appear in the dense web within the established lexicon, new terms appearing at the periphery increasingly attach to a few established keywords. For many non-top 20 keyword groups, differentiation occurs within an integrated web of keyword co-occurrence

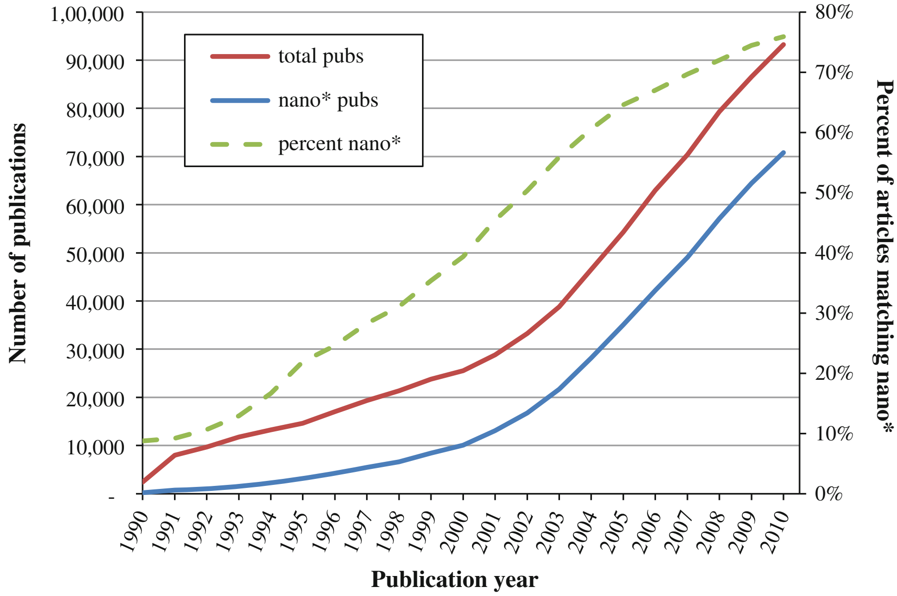

Fig. 1 Absolute number and percentage of publication records matching "nano*" by publication year. Total publications (*red dotted line*) identified through a multi-modular search approach (see Arora et al. 2012). *Source* web of science. (Color figure online)

Table 1 Descriptive measures of keyword groups and percent of articles matching nano*-prefixed terms, by top-20 and non-top-20 keyword groups, by year

Year	Keyword groups	←Percent growth (%)	Total articles in corpus[a]	Matched by non-top 20 terms	←Percent matched (%)	Matched by top-20 terms	←Percent matched (%)
1990	21		2,383	171	7	130	5
1991	49	133.3	8,002	659	8	498	6
1992	70	42.9	9,666	918	9	707	7
1993	89	27.1	11,760	1,363	12	1,099	9
1994	112	25.8	13,257	1,988	15	1,573	12
1995	150	33.9	14,636	2,845	19	2,297	16
1996	159	6.0	17,021	3,633	21	2,986	18
1997	205	28.9	19,328	4,766	25	3,874	20
1998	253	23.4	21,345	5,678	27	4,549	21
1999	287	13.4	23,797	7,244	30	5,893	25
2000	319	11.1	25,512	8,516	33	6,891	27
2001	383	20.1	28,842	11,072	38	8,842	31
2002	500	30.5	33,237	14,281	43	11,516	35
2003	618	23.6	38,801	18,449	48	14,858	38
2004	795	28.6	46,533	24,098	52	19,354	42
2005	961	20.9	54,329	29,799	55	23,647	44
2006	1,072	11.6	62,976	35,617	57	28,351	45
2007	1,175	9.6	70,380	41,313	59	32,685	46
2008	1,353	15.1	79,287	47,653	60	38,068	48
2009	1,471	8.7	85,681	53,294	62	42,441	50
2010	1,517	3.1	91,301	57,908	63	46,112	51

Source web of science

[a] Indicates all records matching the entire nanotechnology research domain, as defined by Arora et al. (2012), and as depicted in Fig. 1

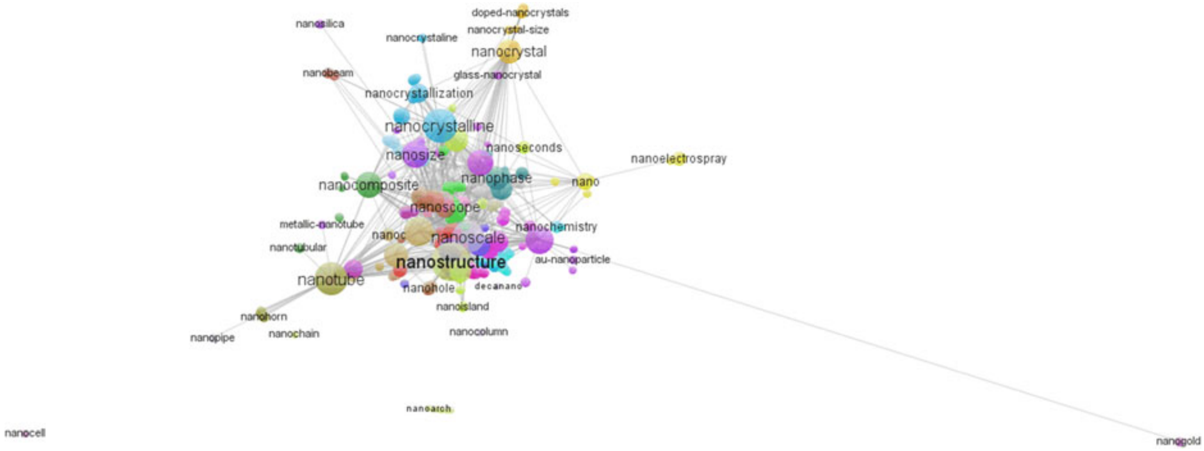

Fig. 2 Nano keyword group co-occurrence map, 2000. The total number of groups equals 319 but only 262 groups and 700 edges are shown because unconnected groups are not mapped here. *Colors* represent clusters, which are not interpreted. (Color figure online)

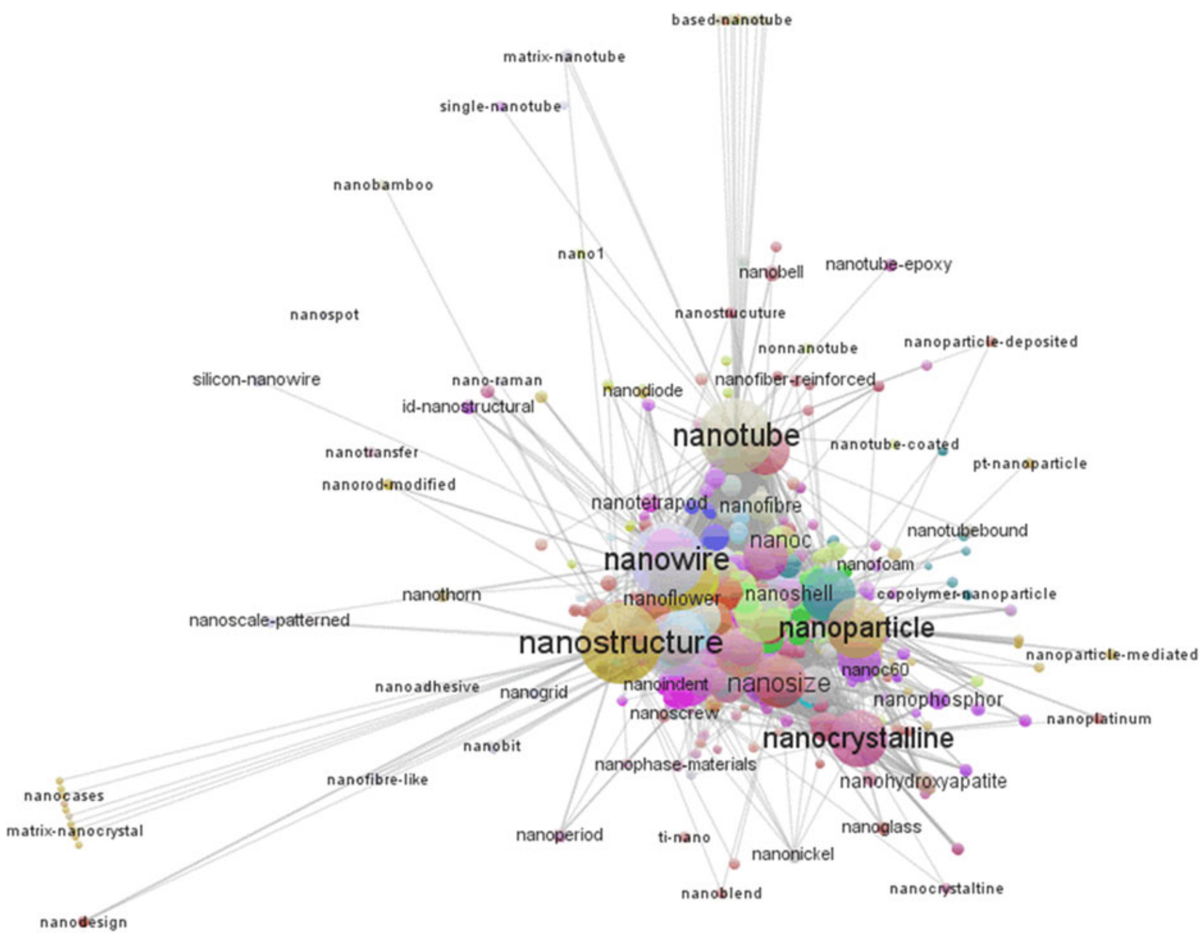

Fig. 3 Nano keyword group co-occurrence map, 2005. The total number of groups equals 961 but only 890 groups and 5,500 edges are shown because unconnected groups are not mapped here. *Colors* represent clusters, which are not interpreted. (Color figure online)

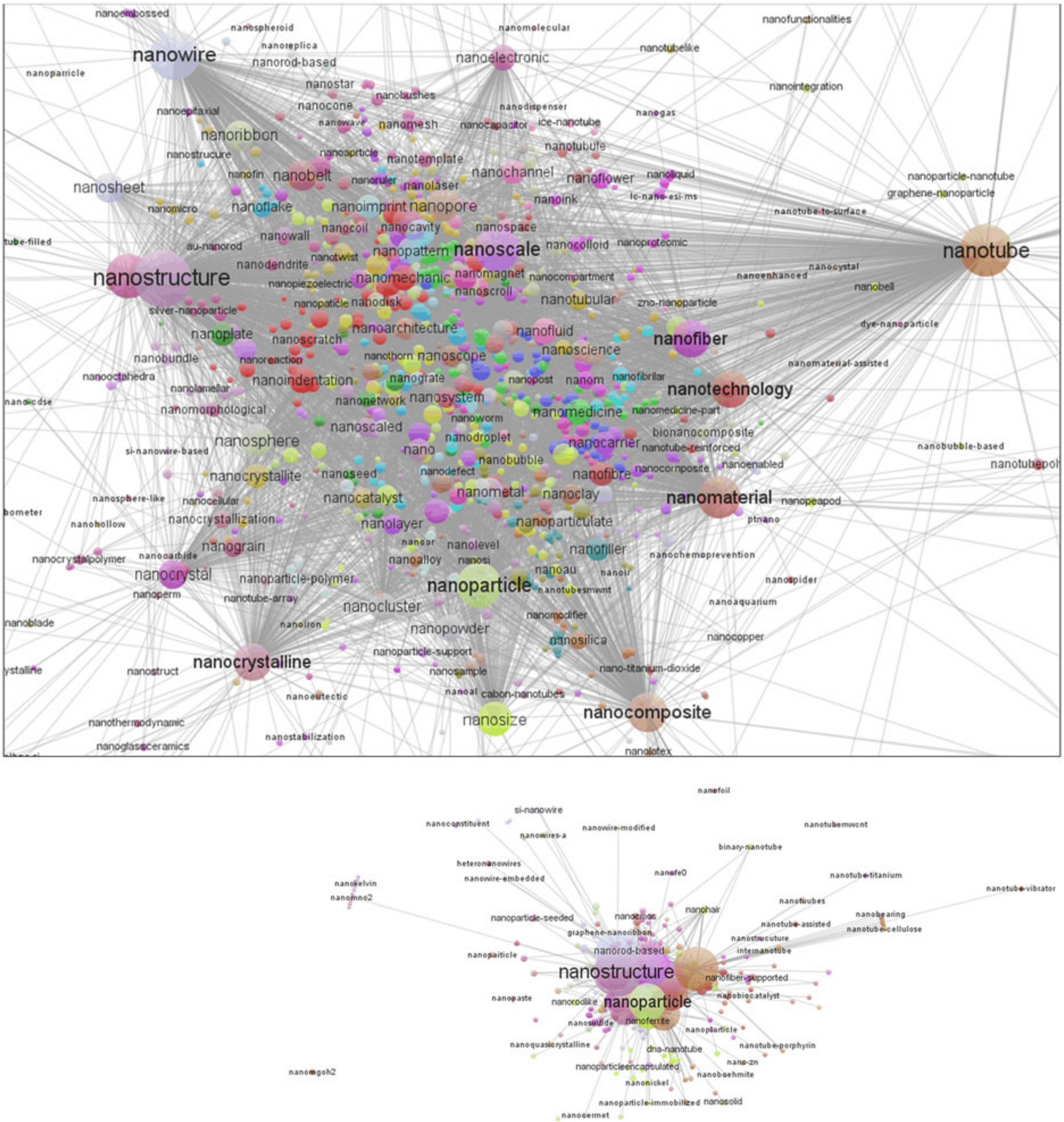

Fig. 4 Nano keyword group co-occurrence map, 2010, zooming-in on the dense center of the graph. 1,447 groups and 7,200 edges are shown in the *inset*. Missing groups are not connected to the main component (*inset*). *Colors* represent clusters, which are not interpreted. (Color figure online)

relationships, suggesting that term development is based to some extent on a common intellectual platform of related concepts. One may question whether these findings are an artifact of the illustrations, but we reiterate that the mapping algorithm does not impose an artificial structure.

An interpretation of the findings of the mapping content relative to our research questions suggests that new terms attach to old ones, but not to such an extent that the network becomes unreachable without the presence of the most well-known, well-established keyword groups. In fact, as the number of keyword groups increases, the density of the network decreases—at least since 1996 (Table 2). In addition, keyword groups attach to small clusters of similar concepts; this is evident through the upward trending clustering coefficients, which by 2010, reaches 79 %. This increase in clustering is not merely a function of an increasing number of keyword groups because the network density measure declines over time.

Table 2 Whole network measures across the nano keyword group co-occurrence map, 1990–2010

Year	Keyword groups	Co-occurrences	Network density (%)	Average clustering coefficient (%)
1990	21	4	1.90	n/a
1991	48	41	3.63	49
1992	70	58	2.40	43
1993	89	143	3.65	63
1994	112	172	2.77	62
1995	150	314	2.81	64
1996	159	383	3.05	62
1997	205	523	2.50	67
1998	253	657	2.06	69
1999	287	879	2.14	73
2000	319	1,084	2.14	72
2001	383	1,443	1.97	68
2002	500	2,260	1.81	73
2003	618	3,212	1.68	78
2004	795	4,631	1.47	75
2005	961	6,337	1.37	77
2006	1,072	7,703	1.34	77
2007	1,175	8,916	1.29	77
2008	1,353	10,402	1.14	77
2009	1,471	11,886	1.10	79
2010	1,517	12,896	1.12	79

Measures computed in Gephi. *Source* web of science

We conclude our findings with an analysis of application orientation: we compare the number of records matching nano* in the biomedical science and clinical medicine categories with the number of records matching nano* in materials science from 2000 to 2010 (Table 3). Results indicate that the percentage of records in materials science, while over 70 % of the entire corpus, begins to decline in 2006. Conversely, the percentages in biomedical science and clinical medicine rise over time beginning in 2001–2002. In 2001, the percentage of articles in biomedical sciences was 4.0 % of the corpus in that year; by 2010, the share increased to 9.3 %. In absolute terms, this represents an increase in 489 publications in 2000 to 6,411 publications in 2010. Similarly for clinical medicine, the share of publications for clinical medicine increased from 0.9 % of the entire corpus in 2001 to 2.5 % in 2010. The increase in terms in the absolute numbers is 89 in 2000 versus 1,703 in 2010. These findings suggest that nano-prefix terms increasingly appear in more application-oriented journals, at least in case of the nano-bio subdomain.

Discussion and conclusion

This paper has presented an assessment of the use of nano-prefixed terms over 21 years of publication data in the nanotechnology domain. It has proposed that nano-prefixed words have followed a pathway toward ever greater usage, and that increasing specialization has occurred in this terminology. We examine this proposition by parsing our more than 800,000 nanotechnology publication data set (which was developed from an eight-module, two-stage Boolean search strategy) into two parts: those containing nano* terms and those that do not (but contain other nano-related

Table 3 Percentage of records matching nano* keywords and belonging to select meta-discipline categories, 2000–2010

Meta-discipline	2000 (%)	2001 (%)	2002 (%)	2003 (%)	2004 (%)	2005 (%)	2006 (%)	2007 (%)	2008 (%)	2009 (%)	2010 (%)
Biomedical science	4.3	4.0	4.5	4.2	5.4	5.7	6.1	7.4	7.2	8.3	9.0
Clinical medicine	0.9	0.9	1.1	1.2	1.2	1.3	1.6	1.7	1.9	2.3	2.5
Materials science	76.5	78.7	77.8	78.2	78.2	76.6	80.8	79.9	78.9	77.6	76.7
# Nano* records	10,0062	13,122	16,800	21,712	28,211	35,142	42,266	49,113	57,182	63,820	69,275

Source web of science. Total records indicates the number of all records matching nano* keywords in the corpus terms without the nano prefix). We find that the percentage of articles with nano* terms in their titles, abstracts, and keywords has increased from fewer than 10 % in the early 1990s to nearly 80 % by 2010. In addition, we show through network analysis that new key terms represent increasing specialization even as there is coalescence around the top 20 most common terms. We also show that nano-prefixed terms are becoming increasingly application-oriented, based on a case study of biomedical and medicine subdomains.

We recognize that other factors may limit the extent to which these results reflect scientific trajectories. To the degree that nanotechnology is a "catchphrase" that represents political, social, and economic roots rather than intellectual ones, our results have limitations. In addition, the use of nanotechnology may well be a natural response to the availability of a new pool of funding, rather than a distinctive scientific trend. It is difficult to parse out the effect of these factors from scientific trends, particularly with the methods used in this paper. While there are studies that emphasize the "hype" factor in nanotechnology, there is a large segment of studies that use similar methods to the content analysis used in this work. This latter set of studies continues to distinguish trends in the use of nano-prefixed terms in scientific articles. We judge that the scientific terminology used in a data set of more than 800,000 journal publications, the majority of which are peer reviewed, is unlikely to be solely a reflection of a popularized catchphrase. That does not mean there are no societal influences at work, just that such influences are unlikely by themselves to explain our results.

To further probe the factors behind these trends, future work could deconstruct the observed phenomena by, for example, isolating and tracking effects across subject categories, journal titles, institutions, and author collaborations. This work also draws attention to the need to further refine the methodological and theoretical motivations for pairing keyword occurrences, which are more closely tied to a particular article and less amenable to drawing interferences at a term level. Interviews could also be conducted to distinguish scientific and other motivational factors in the selection of terms.

What do the findings of growth and settling of new nanotechnology terms suggest about the field? The scientific community appears to have adopted a shared perspective on a scientific terminology. At the same time, this terminology in the nanotechnology domain is becoming more specialized around particular areas of application and further scientific discovery.

Our findings offer evidence about the development and maturation of the nanotechnology research domain over an extended period of time. This assertion does not intimate that future developments in nanotechnology will be incremental. It is reasonable to expect the continued development of significant innovations that leverage existing research pathways but possibly do not require a wide variety of new terms to describe and explain.

As argued by (Guston and Sarewitz 2002), by continuing to monitor scientific progress in publication output, a real-time technology assessment in nanotechnology may alert policymakers to specific science and technology policies that can further intended benefits of public programs. We hope that future iterations of lexical analysis will continue to enable retrospective evaluation studies, as well as forward-looking technology roadmaps. In this way, lessons learned in nanotechnology may be generalized to other domains of emerging science and technology in order to stimulate an interactive and constructive relationship between producers and funders of the original scientific work.

References

Arora SK, Porter AL, 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. doi:10.1007/s11192-012-0903-6

Baneyx F, Park J-K (2013) Nanobio versus Bionano: what's in a name? Biotechnol J 8(2):158–159

Borgatti SP, Everett MG, Johnson JC (2013) Analyzing social networks. SAGE, London, p 304

Braun T, Zsindely S, Dióspatonyi I, Zádor E (2007) Gatekeeping patterns in nano-titled journals. Scientometrics 70(3):651–667

Cacciatore M, Anderson A, Choi D-H, Brossard D, Scheufele D, Liang X, Ladwig P, Xenos M, Dudo A (2012) Coverage of emerging technologies: a comparison between print and online media. New Media Soc 14(6):1039–1059

DeBellis N (2009) Bibliometrics and citation analysis: from the science citation index to cybermetrics. Scarecrow Press, Lanham, p 450

Grieneisen M, Zhang M (2011) Nanoscience and nanotechnology: evolving definitions and growing footprint on the scientific landscape. Small 7(20):2836–2839

Grieneisen M, Zhang M (2012) The ongoing proliferation of nano journals. Nat Nanotechnol 7(5):273–274

Grienseisen M (2010) The proliferation of nano journals. Nat Nanotechnol 5:825

Guston D, Sarewitz D (2002) Real-time technology assessment. Technol Soc 24(1–2):93–109. doi:10.1016/S0160-791X(01)00047-1

Hersm M (2011) Nanoscience and nanotechnology in the posthype era. ACS Nano 5(1):1–2

Hodges G, Maynard A, Bowman D (2013) Nanotechnology: rhetoric, risk and regulation. Sci Pub Policy. doi:10.1093/scipol/sct029

Huang Z, Chen H, Yip A, Ng G, Guo F, Chen Z-K, Roco MC (2003) Longitudinal patent analysis for nanoscale science and engineering: country, institution and technology field. J Nanopart Res 5(3–4):333–363

Huang C, Notten A, Rasters N (2010) Nanoscience and technology publications and patents: a review of social science studies and search strategies. J Technol Transf 36(2):145–172

Khushf G (2004) A hierarchical architecture for nano-scale science and technology: taking stock of the claims about science made by advocates of NBIC convergence. In: Baird D, Nordmann A, Schummer J (eds) Discovering the nanoscale. IOS Press, Amsterdam, pp 21–33

Kuhn TS (1996) The structure of scientific revolutions (3rd edn). University of Chicago Press, Chicago, IL

Macnaghten P, Kearnes M, Wynne B (2005) Nanotechnology, governance, and public deliberation: what role for the social sciences? Sci Commun 27(2):268–291

Meyer M (2007) What do we know about innovation in nanotechnology? Some propositions about an emerging field between hype and path-dependency. Scientometrics 70(3):779–810

NSET (2007) The national nanotechnology initiative: research and development leading to a revolution in technology and industry. Subcommittee on Nanoscale Science, Engineering and Technology, Committee on Technology, National Science and Technology Council, Executive Office of the President, Washington, DC

O'Brien JJ, Carley S, Porter AL (2013). Keyword field cleaning through ClusterSuite: a term-clumping tool for VantagePoint software. Poster presented at Global Tech Mining Conference, Atlanta, GA

Porter AL, Youtie J (2009a) Where does nanotechnology belong in the map of science? Nat Nanotechnol 4:534–536

Porter AL, Youtie J (2009b) How interdisciplinary is nanotechnology? J Nanopart Res 11(5):1023–1041

Porter AL, Youtie J, Shapira P, Schoeneck DJ (2008) Refining search terms for nanotechnology. J Nanopart Res 10(5): 715–728

Rafols I, Porter AL, Leydesdorff L (2010) Science overlay maps: a new tool for research policy and library management. J Am Soc Inf Sci Technol 61(9):1871–1887. doi:10.1002/asi.21368Roco

Roco M (2010) Long view of nanotechnology development at NSF. 2010 NSF nanoscale science and engineering grantees conference, Arlington, Virginia, 6–8 Dec 2010

Roco M, Mirkin C, Hersam M (2011) Nanotechnology research directions for societal needs in 2020: retrospective and outlook. Springer, New York

Shapira P, Youtie J, Porter AL (2010) The emergence of social science research in nanotechnology. Scientometrics 85(2):595–611

Van Eck NJ, Waltman L (2010) Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 84(2):523–538. doi:10.1007/s11192-009-0146-3

van Eck NJ, Waltman L, Dekker R, van den Berg J (2010) A comparison of two techniques for bibliometric mapping: multidimensional scaling and VOS. J Am Soc Inf Sci Technol 61(12):2405–2416. doi:10.1002/asi.21421

Wang L, Notten A, Surpatean A (2013) Interdisciplinarity of nano research fields: a keyword mining approach. Scientometrics 94:877–892

Watts RJ, Porter AL (1997) Innovation forecasting. Technol Forecast Soc Change 56:25–47

Watts DJ, Strogatz SH (1998) Collective dynamics of "small-world" networks. Nature 393(6684):440–442. doi:10.1038/30918

Wood S, Jones R, Geldart A (2003) Nanotechnology: from the science to the social. The Economic and Social Research Council, London

Wullweber J (2008) Nanotechnology an empty signifier à venir? Sci Technol Innov Stud 4(1):27–45