J Technol Transfer (2008) 33:472–484
DOI 10.1007/s10961-007-9042-2

Structure of the nanoscience and nanotechnology applications literature

Ronald N. Kostoff · Raymond G. Koytcheff · Clifford G. Y. Lau

Published online: 11 July 2007

Abstract The Applications literature associated with nanoscience and nanotechnology research was examined. About 65,000 nanotechnology records for 2005 were retrieved from the Science Citation Index/Social Science Citation Index (SCI/SSCI) (SCI (2006). Certain data included herein are derived from the Science Citation Index/Social Science Citation Index prepared by the Thomson Scientific®, Inc. (Thomson®), Philadelphia, Pennsylvania, USA) Through visual inspection of the Abstract phrases, all the diverse non-medical Applications were identified, and the relationships among the Applications, both direct and indirect, were obtained. The medical applications were identified through a fuzzy clustering process. Metrics associated with research literatures for specific Applications/Applications groups were generated.

Keywords Nanotechnology · Nanoscience · Text mining · Electronics · Applications · Health

JEL Classification O31 · O32 · O33

The views in this paper are solely those of the authors, and do not represent the views of the Department of the Navy or any of its components, or the Institute for Defense Analyses.

R. N. Kostoff (✉) · R. G. Koytcheff
Office of Naval Research, 875 N. Randolph St, Arlington, VA 22217, USA
e-mail: kostofr@onr.navy.mil

C. G. Y. Lau
Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, VA 22311, USA

1 Introduction

In 2003–2005, a comprehensive text mining study was performed to overview the technical structure and infrastructure of the global nanotechnology research literature, as well as the seminal nanotechnology literature (Kostoff et al. 2005a, b, 2006a, b). Based on the global interest generated by these reports, it was decided to update and expand the study using more recent data, a much more comprehensive query, and more sophisticated analytical tools. A detailed report from the updated study is contained in Kostoff et al. (2007).

In the updated study, text mining was used to extract technical intelligence from the open source global nanotechnology and nanoscience research literature (Science Citation Index/Social Science Citation Index (SCI/SSCI) databases) (2006). Identified were: (1) the nanotechnology/nanoscience research literature infrastructure (prolific authors, key journals/institutions/countries, most cited authors/journals/documents); (2) the technical structure (pervasive technical thrusts and their inter-relationships); (3) nanotechnology instruments and their relationships; (4) potential non-medical nanotechnology applications, and (5) potential health applications.

Most importantly, in the updated study, all of the technical structural analyses of the total nanotechnology database show Applications being a key driver in nanoscience and nanotechnology research. The objectives of this paper are to extract the key features of the global nanotechnology Applications literature.

2 Background

The two main components of the present nanotechnology applications study are the text mining analytical procedure and the nanotechnology topical literature. The text mining background will be summarized briefly. The nanotechnology background has been described in detail (Kostoff et al. 2005b, 2006b), and will not be repeated here. The nanotechnology background is updated and expanded in Kostoff et al. (2007). Finally, the relevant technology transfer background issues will be summarized.

2.1 Text mining

A typical text mining study of the published literature develops a query for comprehensive information retrieval, processes the retrieved database using computational linguistics and bibliometrics, and integrates the processed information. In this section, the computational linguistics and bibliometrics are overviewed.

Science and technology (S&T) computational linguistics (Kostoff 2003a; Hearst 1999; Zhu and Porter 2002; Losiewicz et al. 2000) identifies pervasive technical themes in large databases from technical phrases that occur frequently. It also identifies relationships among these themes by grouping (clustering) these phrases (or their parent documents) on the basis of similarity. Computational linguistics can be used for:

- Enhancing information retrieval and increasing awareness of the global technical literature (Kostoff et al. 1997a; Greengrass 1997; TREC 2006)
- Potential discovery and innovation based on merging common linkages among very disparate literatures (Kostoff 2003b, 2006c; Swanson 1986; Swanson and Smalheiser 1997; Gordon and Dumais 1998)

- Uncovering unexpected asymmetries from the technical literature (Kostoff, 2003c; Goldman et al. 1999). For example, Kostoff (2003c) predicted asymmetries in recorded bilateral organ (lungs, kidneys, testes, ovaries) cancer incidence rates from the asymmetric occurrence of lateral word frequencies (left, right) in Medline case study articles.
- Estimating global levels of effort in S&T sub-disciplines (Kostoff et al. 2000, 2004a; Viator and Pastorius 2001)
- Helping authors potentially increase their citation statistics by improving access to their published papers, and thereby potentially helping journals to increase their Impact Factors (Kostoff et al, 2004a, b)
- Tracking myriad research impacts across time and applications areas (Kostoff et al. 2001; Davidse and Van Raan 1997).

Evaluative bibliometrics (Narin 1976; Garfield 1985; Schubert et al. 1987) uses counts of publications, patents, citations and other potentially informative items to develop science and technology performance indicators. Its validity is based on the premises that (1) counts of patents and papers provide valid indicators of R&D activity in the subject areas of those patents or papers, (2) the number of times those patents or papers are cited in subsequent patents or papers provides valid indicators of the impact or importance of the cited patents and papers, and (3) the citations from papers to papers, from patents to patents and from patents to papers provide indicators of intellectual linkages between the organizations that are producing the patents and papers, and knowledge linkage between their subject areas (Narin et al. 1994). Evaluative bibliometrics can be used to:

- Identify the infrastructure (authors, journals, institutions) of a technical domain,
- Identify experts for innovation-enhancing technical workshops and review panels,
- Develop site visitation strategies for assessment of prolific organizations globally,
- Identify impacts (literature citations) of individuals, research units, organizations, and countries

2.2 Technology transfer

In its modern form, nanotechnology has been around for about 15 years. It has the status of an emerging technology, and many papers/books have been written promoting its Applications potential in many areas. One goal of this study was to document the Applications potential. A second goal was to identify some of the science and infrastructure markers of potential nanotechnology Applications, so that the science of nanotechnology/nanoscience could be accelerated to advanced levels of development. This article is intended to facilitate the nanotechnology transition process by identifying the significant application areas.

In 1997, a Special Issue of the Journal of Technology Transfer (edited by the first author) addressed accelerated conversion of science to technology (Kostoff 1997b). Its articles emphasized the importance of potential downstream users of science to become involved with the science development as early and broadly as possible, in order to direct the science toward potential user needs, and smooth the eventual transition to actual Applications. The important first step in this conversion process is to identify the science relevant to specific desired Applications, and identify the associated infrastructure. Once contact is made between the on-going science and the potential user, then the full technology transfer process can be initiated.

This article will provide such information of importance to the nanotechnology technology transfer community. It will identify the main nanotechnology Applications, in both technology and in health, from today's vantage point, as well as the related science and infrastructure.

3 Approach

The first part of the following approach describes how the main nanotechnology non-medical Applications were identified, as well as their direct and indirect relationships. The second part of the approach addresses identification of the medical Applications.

3.1 Non-medical Applications

The total nanotechnology records for 2005 ($\sim$65,000) were retrieved using a novel 300+ term query (Kostoff et al. 2007). A phrase frequency analysis was performed of this total database, and hundreds of thousands of phrases were generated. All single word, adjacent double word, and adjacent triple word phrases were extracted and corrected to eliminate phrases containing trivial words at the beginning or end, and their occurrence frequencies were recorded. The phrases were then inspected visually, starting from the highest frequency. Approximately 60,000 phrases were examined visually. Every non-medical Applications-related phrase was extracted. Then, the root phrase for each Application (e.g., cataly*, tribolog*, etc.) was inserted into the phrase search engine, and all variants of the Applications terminology were retrieved, including the lowest frequency variants. Approximately 860 phrases resulted. Additionally, phrases related to materials, properties, phenomena, and nanostructures were extracted during the visual inspection process. The non-medical Applications phrases were related to each other and to common materials, properties, phenomena, and nanostructures with the use of correlation maps and co-occurrence matrices.

3.2 A different procedure was used for the medical Applications

A document fuzzy clustering analysis (Karypis 2006), where documents are divided into groups based on their text similarities and where documents can be assigned to more than one group, was performed on the $\sim$65,000 total nanotechnology records retrieved for the overall nanotechnology study. The resulting hierarchical taxonomy was inspected visually, and the largest sub-network that included all medical Applications (hereafter called the Health sub-network) was identified. A meta-level taxonomy of the Health sub-network (the highest two hierarchical levels) was generated, then a taxonomy of the elemental (lowest level) clusters was generated. These clusters were analyzed for infrastructure and technical content.

For analytical purposes, non-medical applications were segregated from medical applications by design, although there was of necessity some small inclusion of medical applications in the non-medical component. In the remainder of this paper, the non-medical applications will be referred to as Applications, and the medical applications will be referred to as Health.

The 401 most cited nanotechnology papers published since 1991 were also retrieved. A short Applications-related analysis was done of this retrieved database as well.

There are two main sections to this article: Applications (non-medical) and Health. The Applications include:

- Lists of the key nanotechnology applications, with emphasis on those referenced most frequently.
- Key findings of co-occurrence matrices, showing the relation of the major nanotechnology applications to materials, properties, phenomena, and nanostructures.
- Results from a factor matrix analysis, showing a quantitative description of the relationships among technical thrust areas.

The Health section uses fuzzy clustering to generate a medical Applications taxonomy, and then generates the infrastructure and technical thrusts of each medical cluster.

4 Results

4.1 Nanotechnology Applications

4.1.1 Nanotechnology Applications types

If the full 860 Applications phrases identified from the total retrieval phrase frequency analysis are used as the contents of an Applications taxonomy, then a flat Applications taxonomy (one level only) can be constructed (Table 1). The contents of this table serve as starting points for many of the analyses in this section.

Six generic Applications dominate the first tier, focusing on catalysts, lasers, devices, sensors, electrodes, and copolymers. In general, many of the applications are typical of the most advanced and exotic chemicals and materials. It is not surprising that the applications are exploiting the unique properties of chemicals and materials at the nanoscale, in areas such as catalysis and sensing, information, electronics, anti-corrosive coatings, tribology, and lubricants.

Additionally, in order to assess the importance of applications relative to impact, the 401 most cited nanotechnology papers were examined for Applications phrases, down to a phrase record frequency of two. Few were found; highly cited papers tend to be at a very fundamental level, and focus heavily on the science relative to the Applications. In other words, science-oriented papers tend to be cited more than applications-oriented papers. Applications mentioned in these highly cited papers include: Catalysis/Photocatalysis; Device(s); Transistors (Field-Effect); Optoelectronics; Light-Emitting Diodes. Many of these are in the electronics and photonics application areas.

4.1.2 Applications-measured quantity co-occurrence matrices

An interesting nanotechnology research question revolves around which Applications are associated with specific input quantities (e.g., materials, properties, phenomena, nanostructures, etc). This section shows those Applications commonly associated with different materials, properties, phenomena, and nanostructures.

Six of the most widely referenced nanotechnology Applications were matrixed with: (a) the materials-related terms strongly associated with each Application; (b) the materials properties-related terms strongly associated with each Application; (c) the nanoscale

Table 1 One level taxonomy of Applications

Catalysts (Photo, Electro, Platinum, Bimetallic, Oxide)

Lasers (Deposition, Ablation, Sapphire, Excimer, Semiconductor, Laser Tweezers, Desorption Ionization, Quantum Dot, Vertical-Cavity Surface-Emitting, Pump, Distributed Feedback, Solid-State, Quantum Cascade, Quantum Well, Edge-Emitting, Waveguide, Matrix Assisted)

Sensors (Glucose/Amperometric/SPR/DNA Biosensors, Immunosensors, Gas, Chemical, Optical, Pressure, Electrochemical, Temperature, pH, Humidity, Oxygen, Force)

Electrodes (Gold, Glassy Carbon, Gate, Composite, Graphite, Platinum, ITO, TiO2, Enzyme, Ferromagnetic, Carbon Paste, Diamond, Calomel, Photo, CNT, SnO2, BDD, Silver, Copper)

Copolymers (Block, Graft, Amphiphilic)

Electrolytes (Poly, Polymer, Composite, Gel, YSZ)

Lithography (Electron Beam, Photo, Nanoimprint, Soft, Optical, Nanosphere, Dip-Pen Nano, Deep Ultraviolet, Interference, Scanning Probe, X-Ray, EUV, AFM, Immersion, Projection, Stereo, Interferometric)

Diodes (Light-Emitting, Laser, Photo, Schottky, Barrier, Tunneling, Junction, P-I-N, Wave)

Corrosion (Resistance/Protection/Inhibition)

Storage (Hydrogen, Charge, Data, Energy, Information, Oxygen, Ion)

Tribology (Wear Resistance/Rate/Mechanisms, Friction Coefficient, Lubrication, Lubricant Films, Solid Lubricants, Scratch Resistance)

Solar Cells (Dye-Sensitized, Photovoltaics, Organic, Silicon, Thin Film, Polymer, Photoelectrochemical, Hybrid, Heterojunction)

Transistors (Field-Effect, MOSFETs, Single-Electron, Thin Film, Heterojunction Bipolar, Electron Mobility)

Detectors (Photo, Infrared, QWIPs, UV)

Etching (Chemical, Reactive Ion, Electrochemical, Dry, Plasma, Wet, Isotropic/Anisotropic, Sputter, ICP, Photo, Silicon, HF, Anodic, Oxide)

Waveguides (Optical, Ridge, Planar, Photonic Crystal)

Batteries (Lithium-Ion)

Capacitors (Super, MOS, Electrochemical, MIM, Ferroelectric, Platinum, Film, PZT, Silicon, Double Layer, Embedded)

Motors (Molecular, Brownian)

Gate (Dielectrics, Insulators, Stacks)

Scaffolds (Tissue Engineering, Composite, PLGA)

Chips (Sensor, Bio, Microfluidic)

Hard Disk (Drives)

Fuel Cells (Oxide, Methanol, Polymer Electrolyte)

Circuits (Integrated)

Electromechanical Systems (Micro, Nano)

Adhesives (Self-Etch, Resins, Conductive, Polyurethane)

Piezoelectric (Ceramics, Quartz Crystal)

Actuators (Piezoelectric)

Resonators (Nanomechanical, Dielectric, Ring, Quartz)

Recording (Magnetic Media, Optical, Data, Holographic)

Cements (Resin, Bone)

Molecular Sieves (Mesoporous, Carbon)

Memory (Random Access, Nonvolatile Devices, Ferroelectric, Optical, Flash)

Transducers (Signal, Ultrasonic)

Reactors (Nano, Micro)

Field Emitters (Arrays, CNT, Field Emission Gun)

Table 1 continued

Superconducting Quantum Interference Device (SQUID)
Filtration (Gel, Ultra)
Displays (Flat Panel, Liquid Crystal)
Coatings (Antireflection)
Superconductors (Thin Films, Wires)
Microlenses (Arrays)
Dechlorination/Generators/Inductors/Explosives/Micromirror/Quantum/Computer/Remote Sensing/Robotics/Defluorination/Optoelectronics/Switching/Imprinting/Screen Printing/Oxidation Resistance/Spintronics/Injection Molding/Photosensitizers/Bearings/Plastics/Computers/Resistors/Micromanipulator

phenomena-related terms strongly associated with each Application; (d) the nanostructure-related terms strongly associated with each Application. While many terms were listed in each matrix, a few were dominant, and will be summarized.

The pervasive materials, materials properties, phenomena, and nanostructures related to the most frequently mentioned non-medical nanotechnology Applications were identified using the Techoasis software (SEARCH 2006), as follows:

- TiO_2, Pt, Si, gold, and polymers tend to stand out as the most pervasive material types
- Morphology, thickness/diameter/particle size, optical properties, catalytic performance, and electrochemical properties tend to stand out as the most pervasive material properties
- Deposition, absorption, oxidation, immobilization, catalysis, degradation, and self-assembly tend to stand out as the most pervasive nanoscale phenomena
- Thin films, nanowires, nanotubes (especially carbon), and self-assembled monolayers tend to stand out as the most pervasive nanostructures

4.1.3 Nanotechnology Applications taxonomies

Maps were constructed to show groupings of related Applications into broader thematic areas. An autocorrelation map of the most widely referenced non-medical Applications showed five weakly connected sub-networks:

- Electronic Devices and Components
- Optical Switching
- Tribology and Corrosion
- Optoelectronic Sensors
- Electrochemical Conversion and Catalysis

Factor analyses were performed to show non-medical thematic areas from a slightly different perspective. A six-factor analysis showed the following themes:

- Factor 1: Optoelectronics
- Factor 2: Tribology
- Factor 3: Lithography
- Factor 4: Control Systems
- Factor 5: Devices
- Factor 6: Microsystems

The six main non-medical Applications thrusts identified above were augmented by three important, but non-networked thrusts, and the nine resulting themes were related to science and infrastructure by co-occurrence matrices. The three additional themes were catalysts, sensors/detectors, and electrochemistry. "Catalysts" was selected due to its high frequency. "Sensors" was the combination of the phrases "sensors" and "detectors," which are both high frequency and very similar. "Electrochemistry" showed up as a separate network branch on the auto-correlation map, and the combination of "batteries"/ "capacitors"/"fuel cells" was selected to represent "electrochemistry." For non-medical Applications:

- The USA leads in total Applications publications and in six out of nine themes in the high-tech research areas such as devices, sensors, and lithography. China leads in publications in three traditional area themes such as catalysis, tribology, and electrochemistry.
- In total Applications, two of the top three institutions are Chinese. However, the USA is well represented by the large University of California and University of Illinois state university systems.
- The journal Applied Physics Letters appears in the top layer in seven of the nine themes and is by far the leader in total Applications publications. Journal of Physical Chemistry B appears in four of the nine themes, as does Journal of Applied Physics.
- For total Applications, the key underlying science areas include XRD, TEM, films, SEM, XPS, electron microscopy, AFM, fabrication, thickness, growth, hydrogen, substrate, carbon nanotubes, microstructures, nanoparticles, particles, diameter, TiO_2, deposits, coatings, electrodes, silicon, CO, infrared spectroscopy FTIR, electrons, biosensors, catalytic activity, oxidation, silica, thin films, nanotubes, silicon substrates, PL, photocatalytic activity, crystals, Raman spectroscopy, mechanical properties, particle sizes, proteins, catalysis, sol-gel, gold, storage, metals, optical properties, annealing, adsorption, platinum, polymer, corrosion, quantum dots.
- Instrumentation and the associated growth of nanostructures dominate the science efforts at present.

The science areas associated with the specific applications areas are as follows:

Catalysis: Catalysis, XRD, TEM, XPS, SEM, TiO_2, hydrogen, infrared spectroscopy, FTIR, CO, water, carbon nanotubes, electron microscopy,
Devices: Fabrication, films, electrons, XRD, AFM, thickness, transistors, structures, silicon, growth, substrate, electrodes, TEM, channels, nanotubes,
Optoelectronics: Films, electron transport, XRD, TEM, fabrication, growth, solar cells, electrons, Thickness, PL, AFM, optical properties, SEM, carbon nanotubes,
Sensors: Biosensors, detection limit, films, materials, fabrication, proteins, electrodes, XRD, AFM, SEM, hydrogen peroxide, enzymes
Lithography: Fabrication, photolithography, electron beam lithography, substrate, AFM, films, silicon substrates, silicon, thickness, deposits, structures,
Tribology: wear resistance, tribological properties, friction coefficient, coatings. Friction, hardness, SEM, XRD, films, microstructures, electron microscopy,
Electrochemistry: XRD, electrodes, SEM, cyclic voltammetry CV, TEM, electrochemical properties, XPS, films, carbon nanotubes, electrolytes, hydrogen
Microsystems: Fabrication, substrate, films, silicon, nanotechnology, deposits, channels, proteins, SEM, AFM, particles, structures, microchannels,

Control Systems: SEM, thickness, XRD, fabrication, films, AFM, microstructures, mechanical properties, structures, TEM, hardness, friction, silicon substrates,

4.2 Nanotechnology health

4.2.1 Nanotechnology health types

The document clustering approach used to identify the Health types was a recent algorithmic upgrade of our CLUTO software package (Karypis 2006) called fuzzy clustering, where a record could be assigned to multiple clusters. Fuzzy clustering, compared to non-fuzzy clustering, is important for articles that have multiple thrusts, such as Health applications articles in a research database.

There were 256 elemental clusters specified for the algorithm. Of these, 22 were in the Health sub-network. Of these 22 elemental clusters, 19 related directly to Health. The resultant 19 clusters are of different types. Some address specific Health problems (e.g., Tumor Treatment, Sentinal Lymph Node Cancer), some address Health Treatment mechanisms (e.g., Drug Release, Drug Delivery), some address biomaterial types (e.g., Cells, DNA, Biofilms, Virus Proteins, Amyloid Fibrils), but most are Health-related phenomena and processes (e.g., Peptide Sequences, Binding and Affinity, Detection, Sensing). A medical Applications categorization constructed from visual inspection of the fuzzy clustering categories showed five thematic categories: Cancer Treatment; Sensing and Detection; Cells; Proteins; DNA. The higher-level taxonomy categories will now be discussed, followed by a discussion of the elemental clusters.

4.2.2 Higher level taxonomy categories

In the highest-level category in the Health sub-network, the USA appears to have a commanding lead ($\sim$3/1) over its nearest competitor (China). However, these results must be considered in context. First, in total SCI articles, the USA had about four times as many records as China when these data were obtained. Second, for overall nanotechnology, the USA had about 25% more records than China for 2005. Third, for nanotechnology instrumentation, China actually had 25% more records than the USA. Fourth, relative to China, the USA had a commanding lead in overall biomedical articles, as our recent text mining study on China showed (Kostoff et al. 2006d). When all these facts are integrated, it appears that China is placing substantial emphasis on its nanotechnology medical research relative to its overall medical research.

The USA has substantial institutional representation in the top ten (California, Texas, Harvard, Illinois). These university publication numbers include all the state campuses. Thus, University of California system includes UCB, UCSB, UCSF, etc.

While the leading journals have a strong chemistry component, a number of them cross disciplines among physics, chemistry, biology, and materials.

The science associated with the total Health-type applications in the highest-level category can be divided into four major categories: instrumentation, materials, structures, phenomena. The key elements of each of these categories are as follows:

Instrumentation: surface plasmon resonance, atomic force microscopy, scanning electron microscopy, transmission electron microscopy, differential scanning calorim-

etry, X-ray photoelectron spectroscopy, fourier transform infrared spectroscopy, quartz crystal microbalance, magnetic resonance imaging, confocal laser scanning, enzyme linked immunosorbent assay, laser scanning microscopy, X-ray diffraction, mass spectrometry.
Materials: protein, DNA, peptides, drugs, bovine serum albumin, poly ethylene glycol, single stranded DNA, double stranded DNA, green fluorescent protein, lipids, human serum albumin, *Escherichia coli*, antibodies, tissues, enzymes, genes, oligonucleotides, gold, nucleic acid.
Structures: cells, membranes, surfaces, nanoparticles, self-assembled monolayers, cell surfaces, endothelial cells, receptors.
Phenomena: fluorescence, interaction, polymerase chain reaction, dynamic light scattering, resonance energy transfer, particle size, drug release, cell adhesion, binding, affinity, gene expression, transfection efficiency.

The highest-level category is divided by the fuzzy clustering algorithm into two categories, with one category being about seven times the size (number of records) as the other category. The larger category is centered around cells, proteins, and membranes, while the smaller category is strongly focused on DNA. The larger category's main journals (Langmuir 185, Biomaterials 120, Journal of Physical Chemistry 112, Analytical Chemistry 108, Journal of Biological Chemistry 97, Biophysical Journal 95) focus on chemistry, physics, biology, and materials, while the smaller category's main journals (Langmuir 30, Nano Letters 29, Nucleic Acids Research 27, Analytical Chemistry 27, Journal of the American Chemical Society 21, Journal of Nanoscience and Nanotechnology 21) focus on chemistry and nanotechnology. The only journal in common at the top is Langmuir.

The larger category's main country performers (USA 1867, P.R. China 620, Japan 608, Germany 561, England 323, France 301, South Korea 299) are remarkably similar to the smaller category's main country performers (USA 273, P.R. China 123, Japan 106, Germany 78, England 46, France 43, South Korea 33).

4.2.3 Lower level taxonomy categories

For a detailed description of the lower level taxonomy categories, see Kostoff et al. (2007).

For medical Applications, analysis of 19 thematic categories obtained from fuzzy clustering of the total 2005 nanotechnology database revealed the following:

- The USA is the publication leader in total Health types, and in all the thematic areas as well, most by a wide margin. China was the second most prolific in seven thematic areas, Japan in six, Germany in four, and England in two.
- The University of California system led in five clusters, the Chinese Academy of Science led in four, and the National University of Singapore led in three. The University of California and the Chinese Academy of Science were the most prolific in the non-medical Applications as well, but their orders were reversed. The National University of Singapore is a prolific contributor, especially in pharmaceuticals and biomaterials.
- The journal Langmuir contains the most articles in total Health, and is in the top layer of ten of 19 themes. The only journals in common in the top layers of Applications and Health are Langmuir and Journal of Physical Chemistry B.
- For total Health, the key underlying science areas include cells, proteins, DNA, membranes, binding, drugs, fluorescence, peptides, nanoparticles, detection, lipids, antibodies, immobilization, tissues, receptors, enzymes, genes, drug delivery, self

assembly, cell surface, detection limit, *Escherichia coli*, amino acid, molecular weight, particle size, real time, serum albumin, drug release, cell line, cell adhesion, DNA molecules, endothelial cells, surface plasmon resonance, atomic force microscopy, scanning electron microscopy, transmission electron microscopy, differential scanning calorimetry, X-ray photoelectron spectroscopy, bovine serum albumin, poly ethylene glycol, single stranded DNA, double stranded DNA, green fluorescent protein, fourier transform infrared spectroscopy, quartz crystal microbalance, polymerase chain reaction, self assembled monolayer, magnetic resonance imaging, confocal laser scanning, dynamic light scattering, enzyme linked immunosorbent assay, resonance energy transfer, extracellular matrix, laser scanning microscopy, human serum albumin, and poly lactic acid.

These results require further context. The three major institutions discussed are of different size, have different funding levels, and have different manpower and other resources. For example, in 2005, there were 3,399 Articles and Reviews in the SCI/SSCI that contained at least one author with a National University of Singapore address, and there were a total of 6,622 authors listed on these records. The corresponding numbers for the other major institutions are: Chinese Academy of Science, 14,347 records, 19,089 authors; University of California system, 27,954 records, 84,667 authors.

Thus, for the National University of Singapore to be the publication leader in three thrust areas requires a considerable concentration of its relative modest resources relative to the other major institutions.

For the technology transfer community, these results contain some important messages. First, while there are some pervasive infrastructure results throughout the elemental clusters (e.g., USA is always most productive, China, Japan, Germany, England typically rank high), there are many individual differences. To understand the specific research infrastructure related to specific Health applications, disaggregated evaluations are necessary. While the present analysis had a reasonable level of disaggregation, users interested in very specific medical applications will want to conduct much more disaggregated analyses. There are substantial differences between the overall nanotechnology Health results and very specific Health applications results.

Additionally, while there are some instruments that pervade the different elemental clusters, there are substantial instrumentation, material, nanostructure, and phenomenological differences among the clusters. Again, the individual cluster research can differ substantially from the overall nanotechnology Health applications average. Users who are interested in tracking the nanotechnology Health-related research for technology transfer purposes are well-advised to conduct specific analyses of the above type for each application. For investors, identifying which research areas pervade multiple applications would be extremely valuable, and the same recommendations are made as for technology transfer application.

While the instrumentation and nanostructures are similar in the Applications and Health phrases, the phrases uniquely contained in the Health themes include proteins, drugs, antibody, bacteria, DNA, Ph, peptides, tissues, drug release, drug delivery, genes, cytotoxicity, lever, brain, bone, etc. Thus, while there are fundamental research issues that span both types of applications, mainly in the techniques, there are many unique issues when specific materials and phenomena are encountered.

5 Summary and conclusions

The study has identified the main nanotechnology Applications, both medical and non-medical, as well as the related science and infrastructure. These relationships will allow the potential user communities to become involved with the Applications-related science and performers at the earliest stages, to help guide the science conversion towards specific user needs most efficiently.

References

Davidse, R. J., & Van Raan, A. F. J. (1997). Out of particles: Impact of CERN, DESY, and SLAC research to fields other than physics. *Scientometrics, 40*(2), 171–193.
Garfield, E. (1985). History of citation indexes for chemistry—a brief review. *JCICS, 25*(3), 170–174.
Goldman, J. A., Chu, W. W., Parker, D. S., & Goldman, R. M. (1999). Term domain distribution analysis: A data mining tool for text databases. *Methods of Information in Medicine, 38*, 96–101.
Gordon, M. D., & Dumais, S. (1998). Using latent semantic indexing for literature based discovery. *Journal of the American Society for Information Science, 49*(8), 674–685.
Greengrass, E. (1997). Information retrieval: An overview. National Security Agency. TR-R52-02-96.
Hearst, M. A. (1999). Untangling text data mining. In: *Proceedings of ACL 99, the 37th Annual Meeting of the Association for computational linguistics*, University of Maryland, June 20–26.
Karypis, G. (2006). CLUTO—A clustering toolkit. http://www.cs.umn.edu/cluto.
Kostoff, R. N., Eberhart, H. J., & Toothman, D. R. (1997a). Database Tomography for information retrieval. *Journal of Information Science, 23*(4), 301–311.
Kostoff, R. N. (1997b). Accelerating the conversion of science to technology: Introduction and overview. In R. N. Kostoff (Ed.), *Journal of Technology Transfer*, Special Issue on Accelerating the Conversion of Science to Technology, 22:3, Fall 1997.
Kostoff, R. N., Green, K. A., Toothman, D. R., & Humenik, J. A. (2000). Database Tomography applied to an aircraft science and technology investment strategy. *Journal of Aircraft, 37*(4), 727–730.
Kostoff, R. N., Del Rio, J. A., García, E. O., Ramírez, A. M., & Humenik, J. A. (2001). Citation mining: integrating text mining and bibliometrics for research user profiling. *Journal of the American Society for Information Science and Technology, 52*(13), 1148–1156.
Kostoff R. N. (2003a). Text mining for global technology watch. In M. Drake (Ed.), *Encyclopedia of Library and Information Science* (2nd ed., Vol. 4, pp. 2789–2799). New York, NY: Marcel Dekker, Inc.
Kostoff, R. N. (2003b). Stimulating innovation. In L. V. Shavinina (Ed.), *International handbook of innovation* (pp. 388–400). Oxford, UK: Elsevier Social and Behavioral Sciences.
Kostoff, R. N. (2003c). Bilateral asymmetry prediction. *Medical Hypotheses, 61*(2), 265–266.
Kostoff, R. N., Shlesinger, M., & Tshiteya, R. (2004a). Nonlinear dynamics roadmaps using bibliometrics and Database Tomography. *International Journal of Bifurcation and Chaos, 14*(1), 61–92.
Kostoff, R. N., Shlesinger, M., & Malpohl, G. (2004b). Fractals roadmaps using bibliometrics and database tomography. *Fractals, 12*(1), 1–16.
Kostoff, R. N., Stump, J. A., Johnson, D., Murday, J., Lau, C., & Tolles, W. (2005a). The structure and infrastructure of the global nanotechnology literature. DTIC Technical Report Number ADA435984 (http://www.dtic.mil/). Defense Technical Information Center, Fort Belvoir, VA, 2005.
Kostoff, R. N., Murday, J., Lau, C., & Tolles, W. (2005b). The seminal literature of global nanotechnology research. DTIC Technical Report Number ADA435986 (http://www.dtic.mil/). Defense Technical Information Center, Fort Belvoir, VA. 2005.
Kostoff, R. N., Stump, J. A., Johnson, D., Murday, J., Lau, C., & Tolles, W. (2006a). The structure and infrastructure of the global nanotechnology literature. *Journal of Nanoparticle Research,* 8(3–4), 301–321.
Kostoff, R. N., Murday, J., Lau, C., & Tolles, W. (2006b). The seminal literature of global nanotechnology research. *Journal of Nanoparticle Research, 8*(2), 193–213.
Kostoff, R. N. (2006c). Systematic acceleration of radical discovery and innovation in science and technology. *Technological Forecasting and Social Change, 73*(8), 923–936.
Kostoff, R. N., Briggs, M., Rushenberg, R., Bowles, C., & Pecht, M. (2006d). The structure and infrastructure of Chinese Science and Technology. DTIC Technical Report Number ADA443315. (http://www.dtic.mil/). Defense Technical Information Center, Fort Belvoir, VA.

Kostoff, R. N., Koytcheff, R. G., & Lau, C. G. Y. (2007). Structure of the global nanoscience and nanotechnology research literature. DTIC Technical Report Number ADA461930 (http://www.dtic.mil/) Defense Technical Information Center, Fort Belvoir, VA.

Losiewicz, P., Oard, D., & Kostoff, R. N. (2000). Textual data mining to support science and technology management. *Journal of Intelligent Information Systems, 15*, 99–119.

Narin F. (1976). *Evaluative bibliometrics: the use of publication and citation analysis in the evaluation of scientific activity* (monograph). NSF C-637. National Science Foundation. Contract NSF C-627. NTIS Accession No. PB252339/AS.

Narin, F., Olivastro, D., & Stevens, K. A. (1994). Bibliometrics theory, practice and problems. *Evaluation Review, 18*(1), 65-76.

Schubert, A., Glanzel, W., & Braun, T. (1987). Subject field characteristic citation scores and scales for assessing research performance. *Scientometrics, 12*(5–6), 267–291.

SCI. (2006). Certain data included herein are derived from the Science Citation Index/Social Science Citation Index prepared by the Thomson Scientific®, Inc. (Thomson®), Philadelphia, Pennsylvania, USA: ©Copyright Thomson Scientific® 2006. All rights reserved.

SEARCH (2006). *TechOasis*. Norcross, GA: CopyRight, Search Technology Inc.

Swanson, D. R. (1986). Fish Oil, Raynauds Syndrome, and undiscovered public knowledge. *Perspectives in Biology and Medicine, 30*(1), 7–18.

Swanson, D. R., & Smalheiser, N. R. (1997). An interactive system for finding complementary literatures: a stimulus to scientific discovery. *Artificial Intelligence, 91*(2), 183–203.

TREC (Text Retrieval Conference). (2006). Home Page, http://trec.nist.gov/.

Viator, J. A., & Pestorius, F. M. (2001). Investigating trends in acoustics research from 1970–1999. *Journal of the Acoustical Society of America, 109*(5), 1779–1783, Part 1.

Zhu, D. H., & Porter, A. L. (2002). Automated extraction and visualization of information for technological intelligence and forecasting. *Technological Forecasting and Social Change, 69*(5), 495–506.