Defence Science Journal, Vol. 58, No. 5, September 2008, pp. 678-685

REVIEW PAPER

Structure of the Anthrax Research Literature

Ronald N. Kostoff[1], Stephen A. Morse[2], and Serkan Oncu[3]

[1]*Office of Naval Research, 875 N. Randolph St., Arlington, VA 22217**
[2]*National Center for Infectious Diseases, Center for Disease Control, Atlanta, GA*
[3]*Faculty of Medicine, Adnan Menderes University, Aydin, Turkey*

ABSTRACT

Text mining was used to extract technical intelligence from the open source global anthrax research literature. An anthrax-focused query was applied to the *Science Citation Index/Social Science Citation Index (SCI/SSCI*) (*SCI*, 2006) databases. The anthrax research literature infrastructure (prolific authors, key journals/institutions/ countries, most cited authors/journals/ documents) was obtained using bibliometrics, and the anthrax research literature technical structure (hierarchical taxonomy) was obtained using computational linguistics/document clustering.

Keywords: Anthrax, *Bacillus anthracis*, anthraxin, bioterrorism, biowarfare, bioweapons, biological weapons, biodefence, biosecurity, biosurety, information technology, text mining, bibliometrics, citation analysis, computational linguistics, clustering, taxonomy

1. INTRODUCTION

Bio-terrorism has become a major concern for the US in the post-9/11 era. The mailings of letters containing spores of *Bacillus anthracis* to the media and members of the US Congress in September and October 2001 resulted in 22 cases of anthrax with 5 deaths, closed part of the US government's operations, and terrorised the American public[1-3]. Importantly, this event affected numerous countries[4] and there are now major world-wide efforts devoted to countering bio-terrorism (and other potential weapons of mass destruction (WMD), such as chemical and nuclear).

One of the less tangible weapons in the arsenal to combat bio-terrorism is intelligence. This has myriad forms, including the direct use of humans to access information, use of sophisticated computer systems to track infrastructure and resource movements, and development of technologies to detect, neutralise, shield and vaccinate against bio-warfare agents. The goal of the present study is to use text mining to extract technical intelligence from the anthrax research literature.

2. BACKGROUND

The study consists of two components: text mining and anthrax. The comprehensive anthrax and text mining backgrounds are described by Kostoff[5], *et al*.

A typical text mining study of the published literature develops a query for comprehensive information retrieval, processes the database using computational linguistics and bibliometrics, and integrates the processed information. For recent studies in which the databases consist of journal article abstracts and associated bibliometric information (authors, journals, addresses, etc.), the final results have included not only identification of the infrastructure (key authors, journals, institutions), but relationships among the technical themes and authors, journals, and institutions as well[6,7].

As a result of renewed interest in anthrax, there have been a number of recent review articles that have provided comprehensive and complementary perspectives on this disease[8-15]. These review articles are structured along traditional lines in that they cover the etiology and pathogenic mechanisms of anthrax, addressing both biological and medical considerations. However, none of these reviews address the infrastructure and technology structure of the anthrax research literature that text mining can provide.

In recent years, especially since the anthrax attack of 2001[1], there has been increased concern over the use of *B. anthracis* and other potentially lethal microorganisms for bio-terrorism. For this reason, related research papers that address various aspects of bio-terrorism have been retained in the database.

3. MATERIALS AND METHODS

The 'anthrax research literature' was defined as published open-literature papers that focused on theoretical, laboratory, biological, clinical, and epidemiological aspects of anthrax, and emphasized the original research literature (Articles) accessed by the *Science Citation Index* (*SCI*) and *Social Science Citation Index(SSCI)*. The anthrax literature was defined operationally by the following query: "anthrax OR anthracis OR anthraxin", where 'anthraxin' refers to a skin test that detects anthrax cell-mediated immunity.

The iterative search approach of simulated nucleation[16] was used to generate the search query listed above, and

Received 16 July 2007, revised 08 January 2008; *Presently at the MITRE Corporation, McLean, VA 22102

then retrieve the relevant records from the *SCI*/*SSCI* covering the time period from 1991 through 2005 (some bibliometrics updated to 2007). The results of this search were analysed for publication trends, authors, journals, institutions, countries, and citations.

Document clustering[17,18] was used to identify the anthrax literature sub-themes and numbers of records (levels of emphasis) associated with each sub-theme. Sixty-four individual clusters were chosen for the database[19] (1991-2005 articles retrieved from the *SCI/ SSCI*). The clusters are summarized here, and are presented in detail in Kostoff[5], *et al*. CLUTO[17] was used to agglommorate the 64 clusters in a hierarchical tree (taxonomy) structure.

4. RESULTS AND DISCUSSION

4.1 Publication Trends

Figure 1 shows the number of *SCI*/*SSCI* articles retrieved with the above query as a function of time. As the threat of terrorism has increased, the number of research papers per year has increased substantially, and is now an order of magnitude larger than in the early 1990s.

The majority of the anthrax research articles originated in seven countries. The major country producers of anthrax research articles were the USA and France. More significantly, in the seven years since 2000, the anthrax research articles produced in the USA went from ~2.5 times the number of research articles as its nearest competitor to more than 12 times its nearest competitor, reflecting the enhanced investments of the US funding agencies.

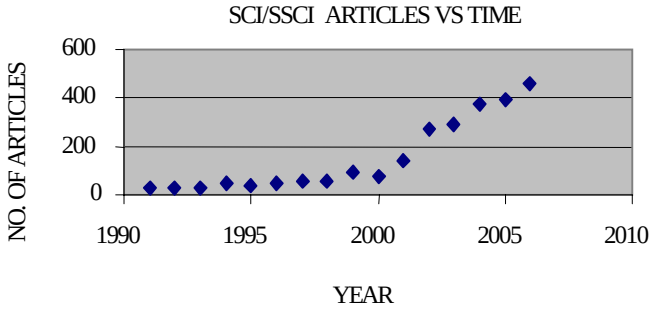

Figure 1. *SCI/SSCI* articles versus time.

4.2 Author Frequency Results

Table 1 lists the 21 authors with the greatest number of anthrax research papers published between 1991 and mid-2007.

Three authors (Leppla, Mock, and Collier) account for 35 per cent of the total number of articles (N=681) published by this group of authors. Thirteen (62 %) of the authors were from the USA, two from France, two from India, and one each from Italy, Norway, UK, and Canada. In previous text mining studies, either the majority or all of the most prolific authors were from universities. However, in the present study, 10 (48 %) of the authors were from research

Table 1. Most prolific anthrax research authors (1991-Mid-2007)

Author	Institution	Country	No. of Papers
Leppla, S.H.	NIH	USA	92
Mock, M.	Institute Pasteur	France	77
Collier, R.J.	Harvard Univ	USA	68
Keim, P.	Northern Arizona Univ	USA	38
Friedlander, A.M.	US Army - Medical Research Inst	USA	36
Fouet, A.	Institute Pasteur	France	33
Bhatnagar, R.	Jawaharlal Nehru Univ	India	30
Little, S.F.	US Army - Medical Research Inst	USA	27
Koehler, T.M.	University of Texas	USA	26
Quinn, C.P.	CDC	USA	26
Singh, Y.	Centre Biochem Tech	India	25
Hanna, P.C.	University of Michigan	USA	23
Turnbull, P.C.B.	Arjemptur Tech, Ltd	UK	23
Ivins, B.E.	US Army - Medical Research Inst	USA	22
Montecucco, C.	University of Padua	Italy	22
Popovic, T.	CDC	USA	22
Tang, W.J.	University of Chicago	USA	21
Moayeri, M.	NIH	USA	19
Baillie, L.W.J.	University of Maryland	USA	17
Kolsto, A.B.	University of Oslo	Norway	17
Mogridge, J.	University of Toronto	CA	17

Table 2. Most cited first authors

Author	Current institution	Country	# Cites
Leppla, S.H.	NIH	USA	662
Friedlander, A.M.	US Army - Med Res Inst	USA	429
Turnbull, P.C.B	Arjemptur Tech, Ltd	UK	416
Inglesby, T.V.	Johns Hopkins Univ	USA	346
Ivins, B.E.	US Army - Med Res Inst	USA	321
Welkos, S.L.	US Army - Med Res Inst	USA	254
Singh, Y.	Centre Biochem Tech	India	248
Duesbery, N.S.	Van Andel Res Inst	USA	241
Hanna, P.C.	Univ Michigan	USA	236
Milne, J.C.	Harvard	USA	232
Little, S.F.	US Army - Med Res Inst	USA	224
Klimpel, K.R.	NIH	USA	224
Dixon, T.C.	Duke Univ	USA	217
Brachman, P.S.	Emory Univ	USA	203
Keim, P.	Northern Arizona Univ	USA	191
Helgason, E.	University of Oslo	Norway	179
Pezard, C.	Institute Pasteur	France	175
Vitale, G.	Univ of Udine	Italy	169
Mock, M.	Institute Pasteur	France	168
Uchida, I.	Natl Inst Anim Hlth	Japan	165

institutions. Of the remaining authors, 10 (48%) were from universities, and one (Turnbull) is from industry.

4.3 Author Citation Results

4.3.1 Most Cited First Authors

The 21 authors receiving the most total citations from the articles in the retrieved database are listed in Table 2.

Thirteen of these first authors are from the USA, five are from Western Europe, and two are from Asia. Eight of the authors are from universities, 11 are from research institutions, and one is from industry. Past text-mining studies[20,21] have shown that cited documents tend to be at a more fundamental level than the citing papers. Thus, there is a big proportion of authors from research institutions relative to those from universities [who are engaged in more fundamental (basic) research] is again in stark contrast to these previous text mining studies. One more difference from the past text mining studies is that 10 of the 24 most prolific authors appear with the 20 most-cited first authors. In the past text mining studies *(Kostoff, et al. 1999, 2002)*, perhaps one or two authors would be in common between the two lists. The reasons for this high degree of overlap are not clear, but may reflect a highly in-bred community. It may also reflect the relatively high fraction of research institutes, where typically (not always) researchers have a higher first author fraction than from universities.

A caveat is that the *SCI* downloads citations by first author only. Thus, authors who may publish many papers, but who tend to get listed behind the first author, will be under-represented in this tabulation.

To identify the authors most associated with the highly cited papers, the 100 anthrax-related documents cited most highly (as listed in the *SCI*) were retrieved (in January 2007), and the author frequency was extracted. This method of author extraction included all of the authors, and was not limited to the first author. The central authors are clearly evident from these results (Table 3). Many more bibliometric quantities were extracted for the 100 most cited anthrax papers. Kostoff [5] *et al.* have done detailed analyses of these bibliometrics.

Table 3. Authors appearing on the 100 most cited anthrax papers

Author	No. of Papers
Collier, R.J.	16
Leppla, S.H.	15
Mock, M.	10
Klimpel, K.R.	8
Friedlander, A.M.	7
Hanna, P.C.	6
Fraser, C.M.	5
Jiang, L.X.	5
Koehler, T.M.	5
Arora, N.	4
Ivins, B.E.	4
Keim, P.	4
Kolsto, A.B.	4
Little, S.F.	4
Okstad, O.A.	4
Read, T.D.	4
Andersen, G.L.	3
Fouet, A.	3
Fouts, D.E.	3
Helgason, E.	3
Jackson, P.J.	3
Milne, J.C.	3
Mogridge, J.	3
Montecucco, C.	3
Mourez, M.	3
Pitt, M.L.M.	3
Popovic, T.	3
Quinn, C.P.	3
Ravel, J.	3
Vitale, G.	3

*totals are > 100 due to multiple authorship

The citation data for authors in Table 2 represent citations generated only by the specific records extracted from the *SCI*/*SSCI* database for this study. These data do not represent all of the citations received by these papers as they could have been cited additionally by papers in other technical disciplines. The next metric provides examples of the magnitude of these differences.

4.4 Most Cited Documents

Table 4 lists the 20 most cited documents. The column headed '#Cites' reflects the citations from the retrieved documents only, whereas the column headed 'Total SCI Cites' reflects citations from all documents contained in

Table 4. Twenty most cited anthrax document references.

Document Reference	#Cites	Total *SCI* cites	Max *JRNL* cites
DUESBERY, N.S.,1998, *SCIENC*E, V280, P734	196	297	2659
(Proteolytic inactivation of mapkk by anthrax lethal factor)			
LEPPLA, S.H., 1982, *P NATL ACAD SCI USA*, V79, P3162	190	357	3943
(Anthrax toxin edema factor)			
FRIEDLANDER, A.M., 1986, *J BIOL CHEM*, V261, P7123	178	255	1745
Mmacrophage sensitivity to anthrax lethal toxin)			
DIXON, T.C., 1999, *NEW ENGL J MED*, V341, P815	173	270	5067
(Anthrax)			
INGLESBY, T.V., 1999, *JAMA-J AM MED ASSOC*, V281, P1735	167	293	871
(Anthrax as bio-weapon-medical management)			
PETOSA, C., 1997, *NATURE*, V385, P833	146	272	2749
Anthrax toxin protective antigen crystal structure)			
GREEN, B.D., 1985, *INFECT IMMUN*, V49, P291	123	189	266
(Capsule plasmid in bacillus-anthracis)			
KLIMPEL, K.R., 1992, *P NATL ACAD SCI USA*, V89, P10277	123	211	2338
(Anthrax toxin protective antigen protease activation)			
BRADLEY, K.A., 2001, NATURE, V414, P225	121	181	4733
(Identification of the cellular receptor for anthrax toxin)			
HELGASON, E., 2000, *APPL ENVIRON MICROB*, V66, P2627	121	194	194
(Bacillus anthracis, bacillus cereus, and bacillus thuringiensis - one species)			
MILNE, J.C., 1994, *J BIOL CHEM*, V269, P20607	121	173	1458
(Oligomer formation by anthrax protective antigen)			
JERNIGAN, J.A., 2001, *EMERG INFECT DIS*, V7, P933	120	196	196
(Bioterrorism-related inhalational anthrax)			
INGLESBY, T.V., 2002, *JAMA-J AM MED ASSOC*, V287, P2236	113	175	2656
(Anthrax as a biological weapon)			
MESELSON, M., 1994, *SCIENCE*, V266, P1202	112	245	2620
(The sverdlovsk anthrax outbreak of 1979)			
MOCK, M., 2001, *ANNU REV MICROBIOL*, V55, P647	111	149	316
(Anthrax)			
PEZARD, C., 1991, *INFECT IMMUN*, V59, P3472	109	139	492
(Contribution of individual toxin components to virulence of bacillus-anthracis)			
SAMBROOK, J., 1989, *MOL CLONING LAB MANU*	105	>75000	NA
(Molecular cloning lab manual)			
HANNA, P.C., 1993, *P NATL ACAD SCI USA*, V90, P10198	103	156	2086
(Macrophage role in anthrax)			
MIKESELL, P., 1983, *INFECT IMMUN*, V39, P371	97	172	477
(Plasmid-mediated toxin production in bacillus-anthracis)			
READ, T.D., 2003, *NATURE*, V423, P81	96	181	554
(The genome sequence of bacillus anthracis ames)			

the *SCI/SSCI* database. Finally, the right-most column labeled 'Max JRNL Cites' is the maximum number of citations received by any paper published in that journal for that year. Thus, the first paper listed (published in *Science* in 1998) was cited 196 times by other papers in the retrieved anthrax-specific database, and was cited 297 times by all the papers in the *SCI/SCSI*. The highest cited paper published in *Science* in 1998 received 2659 cites (citation data circa early 2006).

In general, the most cited anthrax documents receive relatively low numbers of citations (<10 of the highest cited papers) when published in the broad multi-disciplinary journals (e.g, *Science*, *Nature*), but seem to perform much better when published in specialty journals (e.g., *Infection and Immunity, Annual Review of Microbiology*).

The twenty most cited publications appear to be highly applied. Additionally, the ratio of citations by other papers in the retrieved anthrax database to total citations as listed in the *Science Citation Index* (the ratio of the left to middle numerical columns) is, on average, higher than in previous text mining studies conducted by the first author. This reflects the highly applied literature, where most of the

citing papers are within the focused topical area, and the highly cited documents are not sufficiently fundamental to be cited outside the specific anthrax literature of interest (with exceptions, e.g., Sambrook's Manual).

The twenty most cited documents can be divided into four topical groups. The largest group contains nine documents (45) focused on the three proteins (PA, EF,and LF) that comprise the two binary anthrax toxins (LT [PA + LF] and ET [PA + EF]). The next largest groups consist of five documents focused on bio-terrorism, post-exposure prophylaxis and medical management of anthrax and five documents focused on broader biological aspects of *B. anthracis* and its pathogenesis. These documents have a medical, epidemiological, and highly experimental focus. Fundamental theory, computer modelling, and access to other literatures are not evident from the citations.

4.5 Journals with Most Anthrax Research Papers

Table 5 presents the 20 scientific journals publishing the most anthrax research papers.

Infection and immunity stands out in terms of the number of papers published, with 50 more papers than its nearest competitor. Many of the journals are highly specialised, and appear quite applied. The technical emphases of these journals are medicine (primarily infectious diseases), microbiology, and chemistry.

Two time bands of publications were analysed: pre 2001, and post 2000. Journals that published substantially more papers after 2000 included *Emerging Infectious Diseases, Journal of Clinical Microbiology, Biochemical and Biophysical Research Communications, Analytical Chemistry, Antimicrobial Agents and Chemotherapy, Journal of Infectious Diseases, and Biosecurity and Bioterrorism.* Journals that published substantially fewer papers after 2000 included *Journal of Applied Microbiology, FEMS Microbiology Letters, and Molecular Microbiology*. The latter are all microbiology journals, while the former cover a broader variety of topics.

Table 5. Journals containing most anthrax research papers (1991-mid-2007)

	# Papers		
Journal	**<2001**	**>2000**	**Total**
Infection and Immunity	39	118	157
Journal of Bacteriology	14	87	101
Applied and Environmental Microbiology	13	79	92
Journal of Biological Chemistry	15	69	84
Vaccine	11	73	84
Proceedings of the National Academy of Sciences of the United States of America	14	61	75
Journal of Applied Microbiology	39	29	68
Emerging Infectious Diseases	5	61	66
Journal of Clinical Microbiology	7	50	57
Biochemical and Biophysical Research Communications	4	44	48
Fems Microbiology Letters	16	23	39
Analytical Chemistry	1	37	38
Molecular Microbiology	15	19	34
Biochemistry	9	24	33
Antimicrobial Agents and Chemotherapy	0	29	29
Journal of Infectious Diseases	1	22	23
Clinical Infectious Diseases	5	17	22
Biosecurity and Bioterrorism-Biodefense Strategy Practice and Science	0	21	21
Cellular Microbiology	2	19	21

Table 6. Most cited journals in the anthrax literature (1991-2005)

Journal	**# Cites**
Infection and Immunity	2829
Journal of Bacteriology	2585
Journal of Biological Chemistry	2347
Proc of the National Academy of Sciences- USA	2049
Nature	1365
Science	1286
Molecular Microbiology	1286
Applied Environmental Microbiology	1154
JAMA-Journal of the American Medical Association	1120
Journal of Clinical Microbiology	799
Biochemistry-US	738
Vaccine	723
Emerging Infectious Diseases	721
MMWR-Morbid Mortal W	592
FEMS Microbiology Letters	564
Nucleic Acids Research	561
New England Journal of Medicine	519
Journal of Applied Microbiology	518
Gene	511
Journal of Infectious Diseases	488

4.6 Most Cited Journals

Table 6 contains the 20 journals cited most frequently in the specific records extracted for this study.

The journals can be divided into four major groups. The first group comprises four most cited journals (*Infection and Immunity, Journal of Bacteriology, Journal of Biological Chemistry, and Proceedings of the National Academy of Sciences*); the second group contains the next five journals (*Nature, Science, Molecular Microbiology, Applied and Environmental Microbiology,* and *JAMA*); the third group contains the next four journals (*Journal of Clinical Microbiology*, *Biochemistry*-US, *Vaccine*, and *Emerging Infectious Diseases*); and the fourth group contains the remainder. The first group and most of the second group consist of basic science journals, whereas about half of the third and fourth groups address clinical/epidemiological/public health-related issues.

4.7 Institutions Producing Most Anthrax Reserch Papers

The 21 institutions producing the most anthrax research papers are shown in Table 7.

Table 7. Institutions producing most anthrax research papers (1991-2005)

		# Papers		
Institution	**Country**	**<2001**	**>2000**	**Total**
US Army	USA	46	99	145
NIH	USA	32	101	133
Institute Pasteur	FRANCE	65	42	107
Harvard University	USA	27	62	89
CDC	USA	4	85	89
University of Texas	USA	11	32	43
US Navy	USA	4	37	41
Johns Hopkins University	USA	3	29	32
Northern Arizona University	USA	13	17	30
US FDA	USA	1	26	29
University of Maryland	USA	5	20	25
Jawaharlal Nehru University	INDIA	4	20	24
University of Chicago	USA	1	22	23
University of Michigan	USA	4	18	22
Louisiana State University	USA	11	10	21
University of Padua	ITALY	11	8	19
Israel Institute of Biological Research	ISRAEL	1	17	18
Public Health Lab Service	UK	17	1	18
Stanford University	USA	1	17	18
Emory University	USA	1	15	16
University of Oklahoma	USA	0	16	16
Lawrence Livermore National Laboratory	USA	0	16	16

Five of the top 10 are government research institutions, and seven of the top 22 are government laboratories. The total number of papers was divided into those published pre-2001 and post-2000. Institute Pasteur was the most prescient of the leading producers, having generated about 2/3 of its articles before 2001. Other forward-looking organisations in this field include Northern Arizona University, Louisiana State University, and University of Padua. The bulk of the organisations (those with 5 or less articles pre-2001) increased their rate of publication of anthrax articles reactively after the anthrax attacks in the USA in 2001.

Five institutions stand out in terms of productivity: US Army (including all variants); NIH (including all institutes); Institute Pasteur; Harvard University; and CDC (a public health agency-includes all variants).

Thirteen are universities, and the others are research institutions (with the exception of CDC). The fraction of research institutions is much higher than in previous text mining studies[6,20,21]. Seventeen of these prolific institutions are located in the USA. Of the remaining four, two are in Western Europe, and one is in India and Israel apiece.

To identify preferred institutional publishing venues, an institution-journal co-occurrence matrix was generated. Based on the leading journals in which the anthrax research is published it can be deducted. The US Army emphasizes vaccines and microbiology; NIH emphasises biochemistry and infection; Institute Pasteur emphasizes bacteriology, microbiology, and infection; Harvard emphasises biochemistry and infection; and CDC emphasises the public health aspects of infectious diseases and epidemiology.

Table 8. Countries producing most anthrax research papers (2005-mid-2007).

Country	**# Papers**
USA	1821
France	189
England	180
Germany	125
Canada	92
India	80
Italy	68
Israel	54
Japan	54
South Korea	52
P. R. China	40
Australia	36
Turkey	35
Netherlands	29
Norway	29
Russia	28
Belgium	27
Switzerland	27
Brazil	20
Denmark	18

4.8 Countries Producing Most Anthrax Research Papers

The 20 countries producing the most anthrax research papers are shown in Table 8.

The USA has dominated the output, contributing over 64 of the open research article literature on anthrax (N=2858) between 1991 and 2007. Following are France (N=189), England (N=180), and Germany (N=125), with other countries far behind. The low numbers of articles from China are interesting. In recent text mining studies of different technical disciplines (e.g., Kostoff[6], *et al*), China has been near the top in country, institution, and author listings. Here, China appears 11 in the top 20 countries.

5. COMPUTATIONAL LINGUISTICS RESULTS

5.1 Anthrax Taxonomy

The first three levels of the hierarchical taxonomy of the retrieved anthrax literature are presented in Table 9. The following describe the first two levels of the hierarchical taxonomy of the anthrax literature. Subsequently, the third level is treated as a flat taxonomy (eight separate non-hierarchical categories). The themes of the elemental clusters in each category are shown in greatly expanded form, including bibliometrics for each cluster, in Kostoff[5],*et al*.

The first taxonomy level (N=1779 records with abstracts) can be sub-divided into two categories: Anthrax clinical medicine/animal epidemiology; bio-terrorism (N=461 records)

Table 9. Anthrax literature taxonomy (records with abstracts - 1991-2005)

Level I	Level II	Level III
Anthrax clinical medicine/ animal epidemiology; bio-terrorism (461)	Anthrax bioterrorism (219)	Biological agent threat/ attack/ detection (97)
		Planning/ surveillance/ communication/ preparedness/ response for bioterrorist attacks (122)
	Anthrax clinical medicine/ animal epidemiology (242)	Evolution, transmission, and impact of infectious disease on animal populations (108)
		Inhalation and cutaneous anthrax, and anthrax meningitis and meningoencephalitis (134)
Anthrax biology (1318)	Anthrax spore detection/ prevention (956)	Vaccination/ immunisation and spore detection (498)
		Bacillus cereus/ anthracis strain identification (458)
	Toxin lethality pathways (362)	Binding of anthrax lethal toxin to host cell receptors (228)
		Lethal toxin inactivation of macrophages and protein kinase (134)

and anthrax biology (N=1318 records). The anthrax clinical medicine/animal Epidemiology; bio-terrorism category focuses on anthrax patient modalities of treatment, as well as more general public health preparedness and emergency care issues resulting from potential bio-terrorist attacks. The anthrax biology category focuses on the mechanisms and pathways, from detection to final intoxification. The boundaries between the two categories are relatively sharp, reflecting the quality of the clustering approach used, and especially the tripling of trivial words compared to past text mining studies.

For the second level taxonomy, each first level category is divided into two sub-categories. Anthrax Clinical Medicine/ Animal Epidemiology; Bioterrorism is divided into Anthrax Bioterrorism (N=219 records) and Anthrax Clinical Medicine/ Animal Epidemiology (N=242 records), while Anthrax Biology is divided into Anthrax Detection/ Prevention (N=956 records) and Toxin Lethality Pathways (N=362 records). The Anthrax Bioterrorism category focuses on potential public health responses to bio-terrorist attacks such as preparedness, emergency care, and the required underlying logistics, while the Anthrax clinical medicine/ animal epidemiology category focuses on non-bioterrorism clinical medicine and bioterrorism case studies in treating anthrax. The Anthrax detection/ prevention category focuses on vaccine development for protection and spore/strain identification for detection, while the toxin lethality pathways category focuses on the binding, activation, and delivery of the toxins to the cell at both the aggregate toxin level and the component factor level.

The second level categories are further sub-divided to form eight third level categories. The third level is discussed in detail, including themes and metrics, by Kostoff[5], *et al*.

In summary, the open anthrax scientific literature has expanded enormously since the terrorist attacks in the USA in 2001, with much of the expansion coming from the USA itself. In contrast to some of the emerging physical sciences areas like nanotechnology, the leading nations in anthrax paper production (and most cited papers as well) are the Western democracies, not the Far Eastern nations.

REFERENCES

1. Jernigan, D.B.; Raghunathan, P.L. & Bell, B.P. Investigation of bioterrorism-related anthrax, United States, 2001: Epidemiologic findings. *Emerg. Infect. Dis.*, 2002, **8**, 1019-028.
2. Hsu, V.P.; Lukacs, S.L. & Handzel, T. Opening a *Bacillus anthracis*-containing envelope : The public health response. Capitol Hill, Washington, D.C. *Emerg. Infect. Dis.*, 2002, **8**, 1039-043.
3. Morse, S.A.; Kellogg, R.B. & Perry, S. Detecting biothreat agents: Laboratory response network. *ASM News,* 2003, 69, 433-37.
4. Polyak, C.S.; JT. Macy, M. Irizarry-De La Cruz, et al. 2002. Bioterrorism-related anthrax: International response by the Centers for Disease Control and Prevention. *Emerg. Infect. Dis.*, **8**, 1056-059.
5. Kostoff, R.N.; Morse, S. & Oncu, S. Text mining of the anthrax literature. DTIC Technical Report No. ADA451571 Defense Technical Information Center, Fort Belvoir, VA. 2006c. (http://www.dtic.mil/).
6. Kostoff, R.N.; Stump, J.A.; Johnson, D.; Murday, J.; Lau, C.G.Y. & Tolles, W. The structure and infrastructure of the global nanotechnology literature. *J. Nanoparticle Res.*, 2006a, **8**(3-4), 301-21.
7. Kostoff, R.N.; Tshiteya, R.; Bowles, C.A. & Tuunanen, T. The structure and infrastructure of the Finnish research literature. *Technol. Anal. Strategic Manage.*, 2006b, **18**(2), 187-220.
8. Dixon, T.C.; Meselson, M.; Guillemin, J. & Hanna, P.C. Anthrax. *New England J. Medicine,* September 1999, **341**(11), 815-26.
9. Mock. M. & Fouet, A. Anthrax. *Ann. Rev. Microbiology,* 2001, **55**, 647-71.
10. Gardner, R.A. Anthrax (*Bacillus anthracis*). *Cambridge Sci. Abs.*, 2001.
11. Khanna, H. & Singh, Y. War against anthrax. *Molecular Medicine,* 2001, **7**(12), 795-96.
12. Koehler, T.M. (Ed). Anthrax. Current Topics in Microbiology and Immunology Series, Vol. **271**, Springer, 2002.
13. Oncu, S.; Oncu, S. & Sakarya, S. Anthrax-an overview.

Med. Sci. Monit. (Poland), November 2003, **9**(11), RA 276-83.
14. Lindler, L.E.; Lebeda, F.J. & Korch, G.W. (Eds) Biological weapons defense: Infectious diseases and counter-bioterrorism. Humana Press, Totowa, N.J., 2005.
15. Anderson, B.; Friedman, H. & Bendinelli, M. (Eds). Microorganisms and bioterrorism. Infectious Agents and Pathogenesis Series. Springer, 2006.
16. Kostoff, R.N.; Eberhart, H.J. & Toothman, D.R. Database tomography for information retrieval. *J. Inf. Sci.*, 1997a, **23**(4), 301-11.
17. Karypis, G. CLUTO–A clustering toolkit. 2006. http://www.cs.umn.edu/~cluto.
18. Zhao, Y. & Karypis, G. Empirical and theoretical comparisons of selected criterion functions for document clustering. *Machine Learning*, 2004, **55**(3), 311-31.
19. Science Citation Index, 2006. Thomson Scientific®, Inc. Philadelphia, Pennsylvania, USA.
20. Kostoff, R.N.; Shlesinger, M.F. & Tshiteya, R. Nonlinear dynamics roadmaps using bibliometrics and database tomography. *Int. J. Bifurcation and Chaos*, January 2004a, **14**(1), 61-92.
21. Kostoff, R.N.; Shlesinger, M.F. & Malpohl, G. Fractals roadmaps using bibliometrics and database tomography. *Fractals*, March 2004b, **12**(1), 1-16.