1-1-1999

Functional Analysis: Deriving Systems Knowledge From Bibliographic Information Resources

R. Watts

A. Porter

Cherie C. Trumbach
University of New Orleans, ctrumbac@uno.edu

Information • Knowledge • Systems *Management* 1 (1999) 45–61
IOS Press

Functional Analysis: Deriving Systems Knowledge from Bibliographic Information Resources

Robert J. Watts[1,*], Alan L. Porter[2], Cherie Courseault[2]

[1]*U.S. Army Tank-automotive and Armaments Command, USA*
[2]*Georgia Tech Technology Policy and Assessment Center, USA*

ABSTRACT: Analyses of particular technology developments provide vital managerial intelligence. Framing of such analyses is key to maximize their managerial value. This paper presents a case analysis of automotive fuel efficiency technologies, posed as a functional analysis, rather than as a technology analysis. This functional perspective provides a systems frame of reference, yielding enhanced managerial knowledge. The case illustrates the use of bibliometric methods to glean systems intelligence from large abstract databases.

INTRODUCTION

Analyses of technological change are useful for compiling pertinent information about particular target technologies. Technology monitoring, technology forecasting, technology assessment, technology roadmapping, and competitive technological intelligence represent forms of technology analyses. The value of any such analyses lies almost solely in their ability to enhance technology management. The premise of this paper is that casting such analyses in terms of target *functional* capabilities provides greater value to management. This perspective requires managers to compare alternative technologies that are able to fulfil particular functional goals.

Such a broader, functional focus poses analytical challenges in finding ways to collect information on multiple technologies and key contextual factors, and then to interpret that information usefully. We present an approach to overcome these challenges that draws upon vast electronic information repositories — namely, bibliographic databases. Searching on target topics in such databases yields huge amounts of information, (i.e. thousands of abstracts). We offer an approach to digest such information through a combination of bibliometric analyses (in essence, listing and counting publication concentrations) and "text mining" (content analysis). This mode of analysis, that we call "Technology Opportunities Analysis" (TOA), helps profile trends in research and development (R&D). It can also identify emerging or unfamiliar research that may intersect one's functional interests.

* Corresponding author: Bob Watts, US Army TARDEC, Warren MI 48397-5000, USA, Tel.: +1 810 574 5012; Fax: +1 810 574 8906; E-mail: wattsb@cc.tacom.army.mil

1389-1995/99/$8.00

METHODOLOGY

Bibliometrics

Many entities are striving to extract higher value information from text sources (cf. Zanasi, 1998, concerning IBM-Europe's text mining). "Information Extraction" is growing explosively, reflecting the increased availability of data sources, and the perceived need to add value to raw data to facilitate management use. TOA, the method utilized in this study, falls within the domain of "Knowledge Discovery in Databases" (KDD), a branch of Data Mining and Information Extraction.

TOA draws upon both statistical and language processing/content analysis approaches. Overall, it fits most closely with "bibliometrics" — the application of statistical analyses to extract information from bibliographic sources. Bibliometrics uses counts of publications, patents, or citations to measure and interpret technological advances. The basic premise of such analyses is that counts of papers or patents provide useful indications of R&D activity and of innovation status, depending on the sources examined. Various forms of bibliometric analysis have emerged. Citation analysis (cf. Garfield et al, 1978) examines referencing patterns among papers and/or patents to detect seminal contributions and interaction patterns, and even to forecast emerging research areas. Patent analysis relates patenting activity to profile company interests and trends. Publication analyses take articles and such as telling indicators of R&D activities.

Bibliometric applications range from strategic (e.g. classifying all British science publications to help formulate national science policy — Cunningham, 1997) through tactical (e.g. providing intelligence on who is doing what on a particular technology — see Kostoff, (1998) for a superb overview). Bibliometrics can help measure and interpret technological advances (cf. Courtial et al., 1997; Ashton et al., 1996). It can support competitive technological intelligence activities (cf. Kostoff, 1994b, 1997), strategic planning, R&D management, intellectual property management, and process improvement (Porter et al., 1998).

A key tenet of bibliometrics is that one can ascertain important links, or relationships, by analyzing: Which topics occur together? Which organizations produce what papers and patents? and Who cites what? (Narin, 1994). Probing such linkages has led to the development of distinct analytical approaches based on entities appearing together — i.e. co-occurrences (Kostoff, 1994a). *Co-citation analysis* identifies pairings of articles jointly cited by later articles (cf. Kostoff, 1998; Small and Griffith, 1988; Franklin and Johnson, 1988). *Co-word analysis*, dating mainly from the 1980s in Europe, looks for words appearing together (Kostoff, 1993). Some co-word analysts focus on keywords (index terms); Kostoff has extended these analyses to whole texts (cf. Kostoff, 1994a). Mapping is particularly useful in facilitating interpretation of the linkages uncovered (Tijssen and van Raan, 1994; Rip, 1988). Specialized software can facilitate such analyses (cf. Dou and Dou, 1995; Kostoff, 1997; Porter and Detampel, 1995; van Raan, 1992).

Technology Opportunities Analysis (TOA)

TOA, the tool suite being applied in the present study, enables bibliometric analyses and mapping. It has considerable in common with:

- *Dataview*, developed by the CRRM at the University of Marseilles (cf. Dou et al., 1991)
- The science mapping tools of the University of Leiden (cf. van Raan, 1992), and
- The Office of Naval Research's co-word analyses (see various Kostoff references, especially Kostoff, 1998).

Georgia Tech's Technology Policy and Assessment Center (TPAC) began developing TOA in 1990 (Porter et al., 1994). Beginning in 1993, TPAC has been developing software to facilitate this process

(Porter and Detampel, 1995). Since 1994, that development has been supported by the Defense Advanced Research Projects Agency (DARPA) with major collaboration by TACOM's National Automotive Center. The software presently includes Georgia Tech's UNIX-based version, TOAK, that provides a research arena in which to experiment with new capabilities, and a more user-friendly Windows version, TOAS, developed at Search Technology, Inc. A project sponsored by NSF beginning in 1998 aims to work with a spectrum of technology managers to learn what analytical results and representations can be most valuable to them (Porter et al., 1998).

TOA applies several statistical techniques, including Principal Components Analysis (PCA), trend forecasting models, and non-linear models (Watts et al., 1997; Cunningham, to appear). These build on text retrieval work on Latent Semantic Indexing led by Bellcore (cf. Deerwater et al. 1990). We further process PCA outputs to represent clusters and links (especially using special forms of multi-dimensional scaling to position clusters and path erasing to depict key links). Taken together, these enable us to generate a family of "technology maps" (Zhu, 1998; Zhu, 1997).

This paper draws particularly upon a proprietary clustering/grouping technique developed under the leadership of the National Automotive Center, and incorporated in TOAS, named principal components decomposition (PCD). PCD classifies and categorizes records (e.g. technological literature or patents) to break out two levels of structure. The structure itself, together with examination of clusters of records, can aid in understanding R&D activity related to the topic of inquiry. The PCD "hierarchy chart" (discussed in the next section) allows the user to investigate relationships between the inductively derived record groups.

The TOA approach is special in that it exploits information for technology management purposes (Newman et al., 1997; van Wyk, 1997). In contrast, most Knowledge Discovery in Databases (KDD) approaches are relatively generic. Most bibliometrics tend toward large-scale tabulations (e.g. national comparisons) with relatively stock measures, applied to research impact evaluation and science policy. We differ in combining the counting and mapping of bibliometrics with the natural language processing of KDD in pursuit of information intended to address specific technology management issues. We differ also in having a "meso" focus — neither trying to retrieve just the few best abstracts on a precise issue, nor profiling entire scientific fields broadly. Instead, we bring information extraction tools to bear to profile specific technologies (e.g. ATM, asynchronous transfer mode).

Conceptually, TOA can be categorized with innovation process models and technology forecasting, premised on the existence of a degree of orderliness to technological innovation, substitution, transfer, and diffusion processes. Robert Watts (1996) has reviewed a number of models of innovation to focus on three forms of "innovation indicators" to help track innovation progress and prospects:

- ♦ Technology Life Cycle Indicators (e.g. benchmarking where the technology is in development and its likely applications)
- ♦ Innovation Context Indicators (e.g. gauging the socio-economic influences on the technology's development and the competitive environment)
- ♦ Product Value Chain Indicators (e.g. assessing business potential)

The use of tools such as TOA, with other sources of information, enables a context-rich "innovation forecasting" that we seek to develop further (Watts and Porter, 1997).

In brief, here are the main steps in the TOA process:

1. Choose topic — the case example will address "fuel efficiency."
2. Conduct preliminary search — we scan various R&D databases to ascertain which best cover the topic, and then perform a simple search. The resulting search set of abstract records in electronic form are sent to the computer on which the TOA software resides. We process the search set, generating a list of the prominent keywords (subject index terms). We review this list, preferably with the target users, to tune the search.

3. Complete the full search — identify and retrieve the pertinent abstracts from the target databases.
4. Perform basic, "one-dimensional" analysis — make lists for fields of interest, such as date, author(s), affiliation of the author, country, keywords, class codes, etc.
5. Perform "two-dimensional" analyses — construct matrices for items of interest. These are based on co-occurrence — i.e. items that appear together in records more frequently than expected in the sample set imply some relationship.
6. Generate useful outputs — represent the basic research profile; forecast trends; map related technologies.
7. Interpret results — associate the outputs with the issues posed.

THE FUNCTIONAL CHALLENGE: ENHANCING AUTOMOTIVE FUEL EFFICIENCY

The United States Army National Automotive Center (of the Tank-Automotive and Armaments Command — TACOM) faces a challenge in seeking to enhance fuel efficiency for military vehicles. In this case study, fuel efficiency represents a technology-enabled function promoted by several national interests. One aim is the obvious interest in saving fuel to reduce costs and to lower petroleum demand. A less obvious driver is to reduce the amount of fossil fuel burned so as to lighten the carbon dioxide burden on the atmosphere, a major determinant of global warming. These important objectives, however, must be balanced against more paramount military objectives such as "power density".

The Army's interest lies in balancing retrofit and new application opportunities. The existing Army fleet of vehicles and the supporting infrastructure will be maintained, as well as upgraded, for the next 10 to 15 years. Technologies that can extend vehicle life and/or reduce operational costs are highly valued. However, these system considerations reach beyond direct functional objectives. Any change in fuel technology could significantly impact priorities for new vehicle developments for the Army fleet. As an example, changes in automotive engines could reduce or eliminate commercial production of existing fleet engine types, causing component costs to rise and availability to decline. So, in assessing candidate technological changes, we need to consider these system factors.

Innovation in fuel technology is inevitable. Commercial and military interests are pursuing many alternative technologies (ceramic engines, hybrid electrics, etc.). The U.S. Army, specifically, TACOM, like any entity engaging in R&D activities, needs to develop intelligence as to what technologies offer the most promise. Subsequently, they will focus R&D spending on those technologies. Monitoring can identify developers of technology who might be potential partners; it can also identify critical knowledge gaps meriting R&D support. Such information is valuable to any entity, such as TACOM, who receives solicitations to support research. Evaluating the relative merits of competing technology proposals demands knowledge of their relative functional payoffs. One aim of this study is to generate criteria for evaluating alternative researchable technologies against the required functions.

TACOM must understand the technology substitution process in order to forecast what technologies could meet their objectives and what ones have the best prospects. Michael Porter (1985) defines such substitution as "the process by which one product or service supplants another in performing a particular function or functions." These functions can be performed by any technology. For example, writing a paper can be achieved on a typewriter or a computer. The functional objective remains the same, however, the computer has clearly replaced the typewriter as the dominant technology used to perform this function.

We initiate our functional analysis by exploring such potential system changes. A good place to start is to identify the product technology value chain and look for its weaknesses (i.e. cost-sensitivities, energy resource fragilities). Understanding a product technology value chain and its vulnerabilities requires in-depth knowledge and experience in the industry. Monitoring competitor behavior can identify potential changes in these factors. Not surprisingly, most analyses focus *within* the industry — in our case, automotive.

We suggest, however, that it is not enough to search for change within one's industry. Technological advances that could radically impact the targeted functional objectives are not necessarily so limited. We need to seek potential "innovations by invasion" by reaching beyond the immediate industrial context. Recognizing effective external development poses a real challenge. Reaching beyond the familiar industry takes initiative and is often uncomfortable for technologists and managers (e.g. dealing with different jargon, identifying the valuable sources of information). However, broader monitoring, focused on the functional needs, can provide clues that reveal potential functional gains from "outside" resources. We have previously applied TOA in this manner to locate critical thin film ceramic R&D in the microelectronics domain with potential engine applications, leading to major joint projects with TACOM (Watts and Porter, 1997).

A functional perspective directs us beyond the known technological development work within one's industry. It also provides the framework within which to reach "outward" and "downward," so to speak. In our case of fuel efficiency, there is little value in finding out that various others are concerned with the general issue. Rather, the value is added by spotlighting potentially different component and supporting technologies that might better serve the functional aims. Another functional perspective is to be alert for technological systems that could fulfil the functional objective singularly. For instance, emerging recycling approaches might reduce certain waste disposal needs, or life cycle design might eliminate them entirely. Accordingly, a broad perspective of the technology-function interrelationships can help point to next generation technologies that could meet the functional objectives differently — and better.

Identifying and monitoring lead technology developers, within the context of their own competitive markets and economic drivers, may provide early indications of next generation technology capabilities that could be applied to one's own functional ends. For instance, medical and consumer electronics companies are often the technology leaders in various application arenas that offer technology transfer prospects. A "technology delivery system" (TDS) perspective that identifies the enterprise and its key capabilities to deliver a technology-based product, process, or service to market, and a mapping of the vital environmental influences acting upon it has proved helpful in assessing technological development prospects (Wenk and Kuehn, 1977; Porter et al., 1991). Most importantly, TDS consideration points us to consider socio-economic and political factors, as well as technical, factors in our functional analysis of fuel efficiency development. It is clear that socio-economic drivers are key concerns here in assessing technological prospects (e.g. environmental regulation requirements for cleaner engines).

All told, we seek indicators of potential technological changes to enhance fuel efficiency, with a need for consideration of the broader implications of such technologies for military vehicle applications. In essence, this is what we mean by *functional analysis* — analyses of multiple technologies with respect to their potential contributions to accomplish particular system functional objectives. In contrast, *technology analyses* are more often bounded in terms of a particular technology's development and application prospects.

To obtain information to assist in this functional analysis, we choose a bibliometric approach that applies particular software — TOA — described previously. This approach enables us to capture a wide spectrum of R&D information and digest it to yield suitable insights. Through bibliometric analysis, we are able to identify key research areas and the span of industries focusing in these areas.

In particular, at this stage, we seek information to guide U.S. Army R&D investment decisions. Downstream, we seek to inform the Army of potential asset type substitutions to initiate appropriate planning. We believe that such techniques can be applied to other functional interests.

FUEL EFFICIENCY TECHNOLOGY OPPORTUNITIES ANALYSIS

The Data

Figure 1 sketches the “information flow” as the case analysis progressed. It is intended to help the reader grasp the overall analytical process and to keep track of what the several figures and tables show.

To initiate this functional analysis on “fuel efficiency,” literature abstract records were obtained from three databases. Two databases, INSPEC and *Engineering Index* (also known as *EI Compendex*; we use “ENGI” for short), consolidate millions of engineering-oriented publication abstracts; the third, *U.S. Patents* (PATS for short), contains all U.S. patent abstracts. These databases, respectively, generally reflect basic research, applied research, and development. The first step in the bibliometric approach is to determine a Boolean search string to find pertinent abstracts in the databases for fuel efficiency. The search strings used in this case (“fuel adjacent to efficiency” or “fuel adjacent to economy” or “fuel adjacent to consumption”) provided:

- 708 INSPEC abstract records for the 1987 to 1997 time period,
- 5978 ENGI abstract records for the period 1985 to 1998, and
- 606 patent abstract records for the period 1986 to 1997.

To the readers not familiar with such databases, we should note their wide accessibility. INSPEC (produced by IEE in London) and ENGI (produced by Engineering Information, Inc.) are proprietary whereas PATS is public domain. Such databases can variously be accessed through world wide web interfaces or through such gateways as “Dialog.” An institution licenses access for proprietary databases, either on a per-abstract basis or unlimited abstracts access. One then, literally, has access to enormous information resources on R&D at one’s “fingertips.” From our personal computers, we can link through the Georgia Tech Electronic Library to these (and some 40 other) databases, perform a search, and retrieve abstracts electronically in approximately one minute. We then initiate analyses using the TOA software.

From the “date indexed” field in the abstracts, the research publication and patent chronologies were derived (Figure 2). Activity, as reflected by the number of documents published each year, has remained relatively constant over the last ten years for basic research (INSPEC search results) and development (PATS). The applied research activity (ENGI) over the same period has increased significantly, albeit with fluctuations. We now turn to explore the nature of this R&D activity.

Emergent R&D Topics

The data just described were sought as a result of a prior PCD analysis. Figure 3 shows a PCD hierarchy chart for diesel engine technology, further elaborated elsewhere (Watts and Porter, 1997). The top structural level, reflecting clusters of related documents, is arrayed across the top row of the chart. For instance, the first cluster, named “Combustion,” implies that a subset of the diesel technology records being analyzed address combustion-related topics. After this level is constructed, similar statistical analyses are performed, using the TOA software, on the documents classified within each top-level group, resulting in the breakout shown under each column heading. For instance,

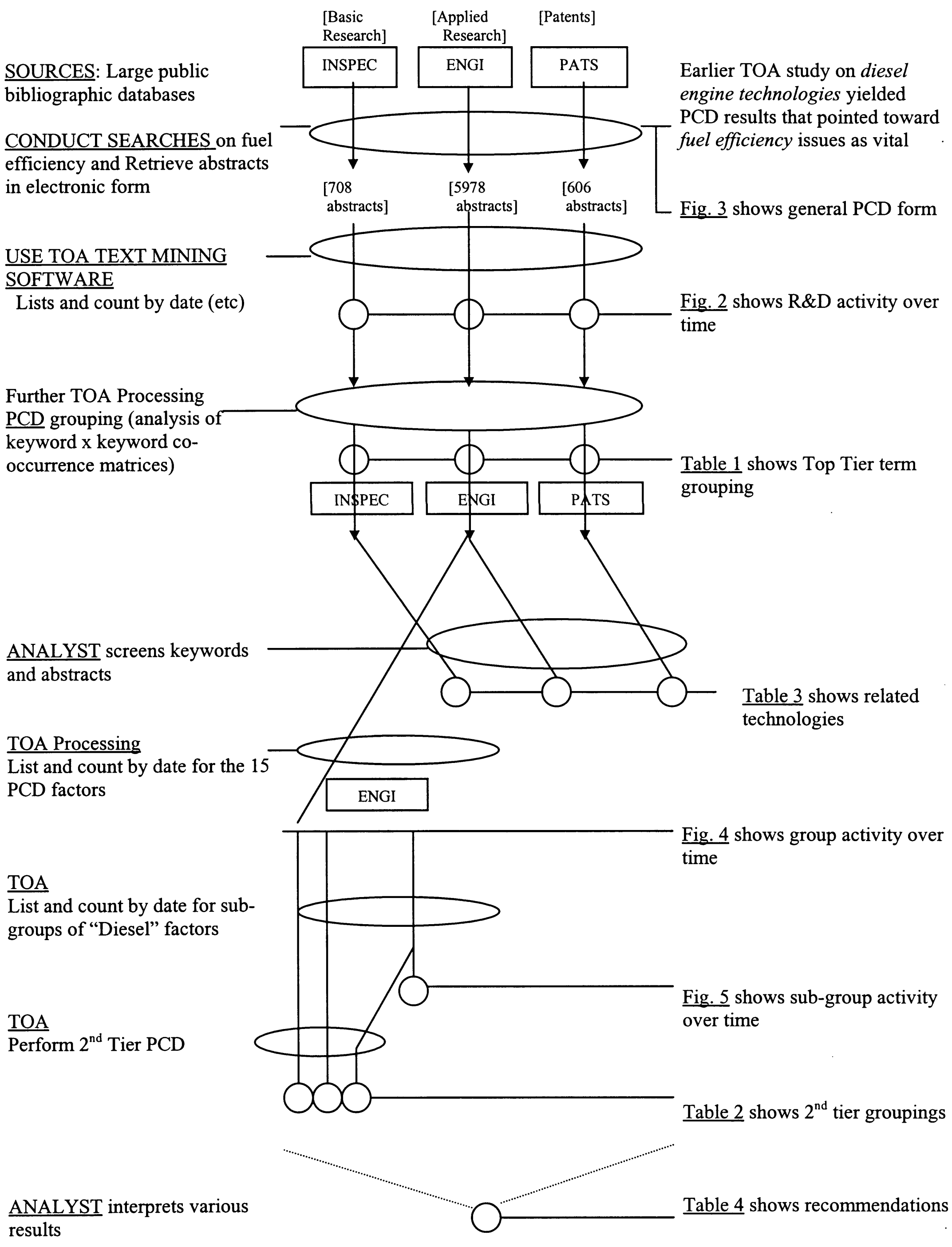


Fig. 1. TOA information flow diagram.

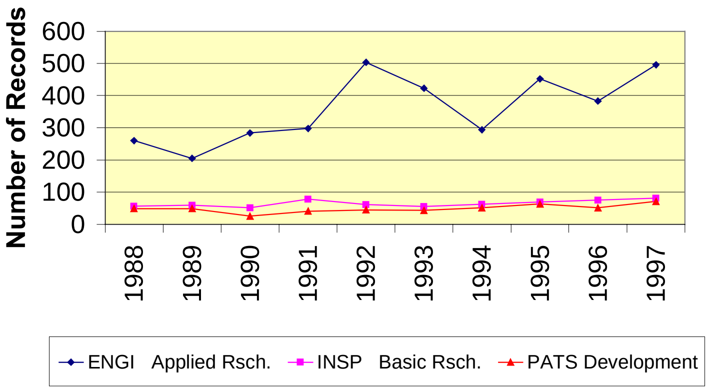


Fig. 2. Research categories publication chronologies.

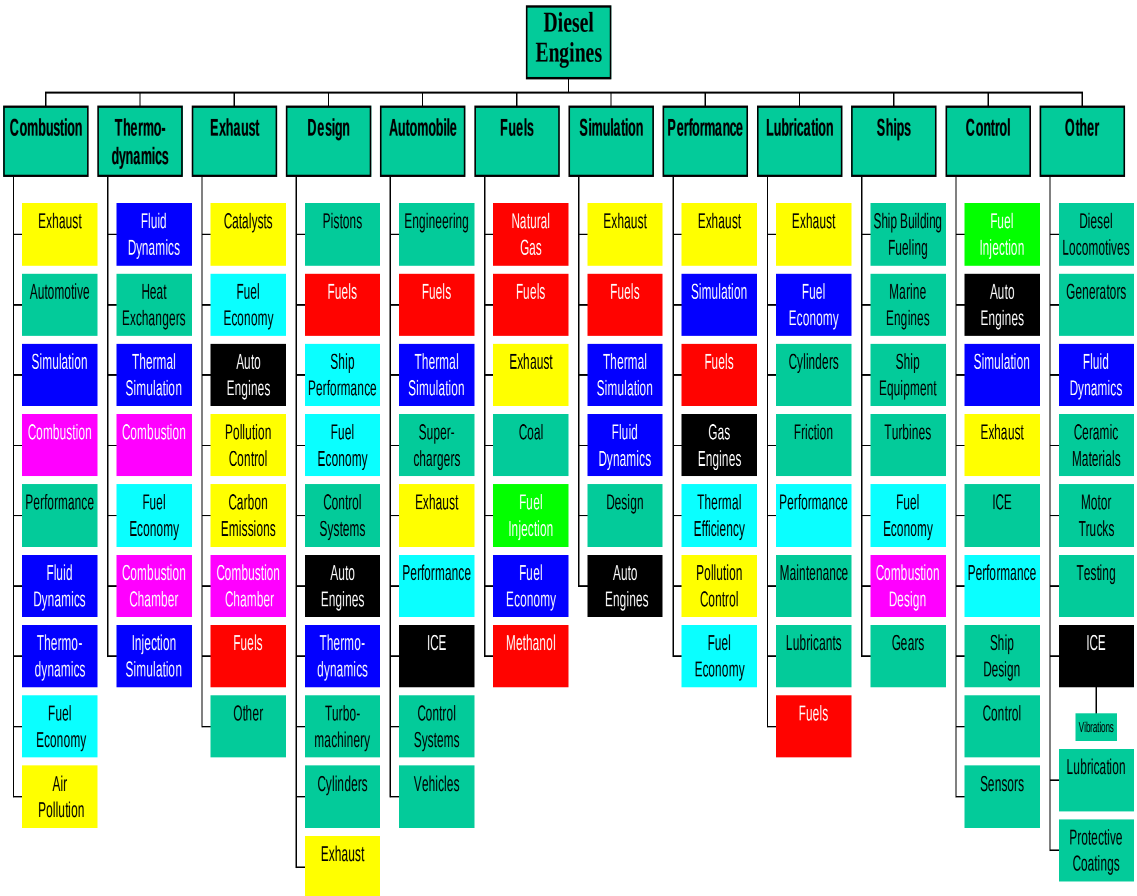


Fig. 3. Principal component decomposition of diesel engine technology.

herein the "Combustion" group statistically subdivides into 9 groups, the first of which emphasizes "Exhaust" issues. In general, this classification can help spotlight potentially relevant research underway, sometimes in quite distinct domains. In this case, Figure 3 helps us frame the search for R&D literature on automotive fuel efficiency, explored in our case analysis in the next section.

To determine and compare the areas of emphasis in each of the R&D categories, the PCD process was next applied separately to the INSPEC, ENGI, and PATS abstract sets. Table 1 presents these results, showing:

1. The emergent research groupings and, for each group, its defining keywords, and
2. Research group size, in terms of both number of records and percentage of that research group.

For instance, in the applied research category (the ENGI search set), factor (FAC 8) is titled "Diesel," the records of which are defined by the frequently occurring keywords: diesel fuels, diesel engines, and air pollution. We create a subset of search records of the 5978 ENGI records that contain one or more of these keywords, thereby constituting the Diesel group. That subset contains 1153 literature abstracts.

Besides the applied research "Diesel" grouping just discussed, this top-level decomposition reveals "combustion" groupings in both the basic research (INSPEC), FAC 4, and development (PATS), FAC 2, categories. These three groups motivate the following analyses.

Table 1
Fuel Efficiency Groups

BASIC RESEARCH GROUPINGS		APPLIED RESEARCH GROUPINGS		PATENT GROUPINGS	
FAC 1 - Road Vehicles road vehicles transportation **(44 Records) (6%)**	**FAC 8 - Aero Controls** optimal control aerospace control **(64 Records) (9%)**	**FAC 1 - Emissions** Particulate emissions Exhaust gases Air pollution control Nitrogen oxides **(595 Records) (10%)**	**FAC 7 - Vehicle Fuels** Fuels Natural gas Vehicles **(677 Records)(11%)**	**FAC 1 - Load** low fuel consumption load **(30 Patents) (5%)**	**FAC 9 - Emissions** improvement emissions reduced fuel consumption **(36 Patents) (6%)**
FAC 2 - Fuel Economies economics thermal power stations fuel energy conservation Air pollution detection and control **(149 Records) (21%)**	**FAC 9 - Diesel-Elec. Vehicles** secondary cells diesel-electric generators electric vehicles **(54 Records) (8%)**	**FAC 2 - Automotive** Automobile engines Automotive engineering Automotive fuels Automobiles **(1159 Records) (19%)**	**FAC 8 - Diesel** Diesel fuels Diesel engines Air pollution **(1153 Records) (19%)**	**FAC 2 - Combustion** combustion chamber combustion water ignition **(87 Patents) (14%)**	**FAC 10 - Operation** engine operation acceleration transmission engine speed vehicle speed operation **(117 Patents) (19%)**
FAC 3 - Turbines gas turbines combined cycle power stations **(48 Records)**	**FAC 10 - Optimize** scheduling optimisation **(40 Records) (6%)**	**FAC 3 - ICE** Internal combustion engines **(766 Records) (13%)**	**FAC 10 - Thermodynamics** Heat transfer **(157 Records) (3%)**	**FAC 3 - Fuels & Lubricants** lubricants fuels **(18 Patents) (3%)**	**FAC 11 - Cylinders** cylinders **(16 Patents) (3%)**
FAC 4 - Combustion combustion **(30 Records) (4%)**	**FAC 11 - Micro-Control** digital control microcomputer applications **(34 Records) (5%)**	**FAC 4 - Aerodynamics** Design Aerodynamics **(371 Records) (6%)**	**FAC 10.2 - Controls** Control systems **(298 Records) (5%)**	**FAC 4 - Speed** sensors speed engines exhaust **(56 Patents) (9%)**	**FAC 14 - Automatic Trans.** control automatic transmission **(48 Patents) (8%)**
FAC 5 - Ship Controls feedback adaptive control control system synthesis ships **(51 Records) (7%)**	**FAC 12 - Digital Simulation** digital simulation **(33 Records) (5%)**	**FAC 5 - Simulation** Computer simulation Mathematical models **(539 Records) (9%)**	**FAC 11 - Turbines** Gas turbines **(249 Records) (4%)**	**FAC 5 - Exhaust Pressure** exhaust gas pressure **(41 Patents) (7%)**	**FAC 15 - SFC** specific fuel consumption **(11 Patents) (2%)**
FAC 6 - Auto Electronics automobiles automotive electronics internal combustion engines **(206 Records) (29%)**	**FAC 13 - Boilers** boilers cogeneration **(49 Records) (7%)**	**FAC 6 - Economies** Performance Costs Economics Gas emissions **(809 Records) (14%)**	**FAC 11.2 - Energy** Energy conservation Energy utilization **(666 Records) (11%)**	**FAC 6 - Fuel Consumption** fuel consumption rate construction **(24 Patents) (4%)**	**FAC 15.2 - Exhaust** gases gasoline exhaust gases **(35 Patents) (6%)**
FAC 7 - Cement Power cement industry power consumption **(24 Records) (3%)**	**FAC 14 - Hydrogen Aircraft** aircraft hydrogen economy **(27 Records) (4%)**		**FAC 12 - Gas** Gasoline **(152 Records) (3%)**	**FAC 7 - ICE** internal combustion engines throttle **(43 Patents) (7%)**	**FAC 16 - Torque** torque **(20 Patents) (3%)**
			FAC 13 - Efficiency Efficiency **(144 Records) (2%)**	**FAC 8 - Performance** housing components performance **(32 Patents) (5%)**	**FAC 17 - Nitrogen** Nitrogen reduction **(31 Patents) (5%)**
	Total "fuel efficiency" file contains 708 abstracts spanning time period of 1987-1997. File groupings capture 543, 77%, of total records.		**Total "fuel efficiency" file contains 5978 abstracts spanning time period of 1985-1997. File groupings capture 4689, 78%, of total records.**		**Total "fuel efficiency" file contains 606 patents spanning time period of 1986-1997. File groupings capture 388, 64%, of total records. ENGI keywords used for PCD record grouping.**

Table 2
Fuel efficiency (ENGI); 2nd tier sub-groups

Diesel Sub-Group		Automotive Sub-Group		Other Sub-Group	
FAC 1 - Simulation Mathematical models Computer simulation (77 Records) (7%)	**FAC 8.2 - Particulates** Air pollution control Particulate emissions Gas emissions Environmental impact (193 Records) (17%)	**FAC 1 - Manufacture** AUTOMOBILE MATERIALS Automobile manufacture (72 Records) (6%)	**FAC 9 - Exhaust** AIR POLLUTION Exhaust gases (139 Records) (12%)	**FAC 1 - Furnaces** COMBUSTION Furnaces FURNACES, HEATING GLASS FURNACES (159 Records) (11%)	**FAC 10 - Iron** IRON ORE TREATMENT IRON ORE SINTER (46 Records) (3%)
FAC 2 - Automobile Automobile engines AUTOMOBILES Automotive fuels (201 Records) (17%)	**FAC 9 - Exhaust** Nitrogen oxides Carbon monoxide Hydrocarbons (132 Records) (11%)	**FAC 2 - Energy** ENERGY CONSERVATION (45 Records) (4%)	**FAC 10 - Transmissions** Automobile transmissions (62 Records) (5%)	**FAC 2 - Electric Power** ELECTRIC POWER PLANTS Fossil fuel power plants Electric power generation (86 Records) (6%)	**FAC 11 - Metallurgy** FURNACES, METALLURGICAL Steelmaking (59 Records) (4%)
FAC 3 - Combustion Fuel injection Combustion (270 Records) (23%)	**FAC 10 - Pistons** PISTONS Ceramic materials HEAT TRANSFER (88 Records) (8%)	**FAC 3 - Methanol** Gasoline Methanol (100 Records) (9%)	**FAC 11 - Control Systems** Control systems Sensors (104 Records) (9%)	**FAC 3 - OIL** Oil economy (27 Records) (2%)	**FAC 12 - Spacecraft** SPACECRAFT Optimization (92 Records) (6%)
FAC 4 - Supercharger SUPERCHARGERS AND SUPERCHARGING Supercharging (57 Records) (5%)	**FAC 11 - Performance** Performance (107 Records) (9%)	**FAC 4 - Simulation** Mathematical models Computer simulation (81 Records) (7%)	**FAC 12 - Fuel Utilization** TRANSPORTATION Fuels Energy utilization (99 Records) (9%)	**FAC 5 - Kilns** Cement plants KILNS (62 Records) (4%)	**FAC 13 - Electric Heat** HEATING Electric power systems Industrial economics (93 Records) (6%)
FAC 5 - Ships SPECIFIC FUEL CONSUMPTION Ship propulsion Marine engines (129 Records) (11%)	**FAC 12 - Control Systems** Design Control systems (90 Records) (8%)	**FAC 5 - Exhaust Control** Natural gas Air pollution control Exhaust systems (engine) Particulate emissions (112 Records) (10%)		**FAC 6 - Aircraft** AIRCRAFT ENGINES, JET AND TURBINE Aircraft engines Aircraft (113 Records) (8%)	**FAC 14 - Ships** MAINTENANCE Ship propulsion Ships (120 Records) (8%)
FAC 6 - Fuels Fuels Natural gas (91 Records) (8%)		**FAC 6 - Design** VEHICLES Design AERODYNAMICS (155 Records) (13%)		**FAC 7 - Boilers** BOILER FIRING Boilers POWER PLANTS (93 Records) (6%)	**FAC 15 - Tires** TIRES MOTOR TRUCKS (45 Records) (3%)
FAC 7 - Lubrication EiRev LUBRICATING OILS (75 Records) (7%)		**FAC 7 - Diesel** DIESEL FUELS Performance Diesel engines Fuel injection (258 Records) (22%)		**FAC 8 - Blast Furnaces** BLAST FURNACES BLAST FURNACE PRACTICE (79 Records) (5%)	**FAC 15.2 - Trans/Material** Automobile transmissions Automobile materials (56 Records) (4%)
FAC 8 - ICE INTERNAL COMBUSTION ENGINES Hydrocarbons (211 Records) (18%)		**FAC 8 - Combustion** Fuels Internal combustion engines Lubricating oils EIREV Combustion (324 Records) (28%)		**FAC 9 - Waste Heat** WASTE HEAT UTILIZATION HEAT EXCHANGERS (57 Records) (4%)	
	Total "Diesel" file contains 1153 records. File groupings capture 937, 81%, of file records.		**Total "Auto" file contains 1159 records. File groupings capture 926, 80%, of file records.**	**FAC 9.2 - Manufacture** Manufacture (26 Records) (2%)	**Total "Other" file contains 1477 records. File groupings capture 964, 65%, of file records.**

"Combustion" appears as a distinct group in both the basic research (INSPEC) and development (PATS) categories (Table 1). However, the top-level applied research (ENGI) document groupings do not include a combustion group per se. This anomaly justified more in-depth analyses. Figure 3, together with automotive engine expert opinion, suggests potential cross-links. Therefore, a second tier PCD decomposition was performed on three of the applied research (ENGI) record groupings (Table 1) — "Diesel" (FAC 8), "Automotive" (FAC 2), and "Other" (the set of the ungrouped literature abstract records) (Table 2).

Observation of the "other" group's record sub-groups (far right section of Table 2) shows a particular feature of this analysis process — these sub-groups reflect different topics than those of the top-level applied research categories (mid-section of Table 1). If the top groupings (Table 1) reflect the R&D mainstream, then emerging knowledge may be located in the other second-tier groupings (Table 2). Methods for extracting this new knowledge would include performing sub-tier grouping, as depicted in Table 2. Another method to seek emerging knowledge is to compare keyword lists for the most recent time period to earlier lists — deliberately seeking terms recently introduced to the domain of inquiry for further investigation (this is illustrated below).

Let's return to the issue of who's researching "Combustion". A motivation to generate the applied research sub-groupings (Table 2) was the lack of an applied research "Combustion" group at the top level (Table 1, mid-section). We do find sub-groups of both the applied research Diesel and

Automotive groups that contain a combustion focus. Examination of the abstract records of these two combustion sub-groups in Table 2 (Diesel FAC 3 and Automotive FAC 8) finds only 29 of the hundreds of abstracts common to both groups. Put differently, these two sub-groups are quite distinct. Such a large difference between the sub-groups, indicates a possible source for new knowledge/innovations for parties engaged in the other technology sub-group. Without going into details, this is a vital interim product of TOA. It is through examination of such focused groupings of publications or patents that one identifies research pockets to be assessed for potential value in meeting one's functional objectives.

EXPLORING R&D TRENDS

Both degree of interest and maturity of a subject area (a technology) can be gauged by a chronological analysis of publication or patent activity. Patterns can be interpreted in terms of empirical evidence of life cycle progression. Figure 3 provides such a representation for the 15 top-level applied research (ENGI) record groups (from the mid-section of Table 1). The figure provides four time slices — 1985–87, 1988–91, 1992–94, and 1995–98. These are used because they lessen the noisiness of annual fluctuations and because the overall time series shows, first, a decline from the 1985–87 period to 1988–91, and, second, an increase since then (recall Figure 2's ENGI trend). Observation reveals that the applied research publication activity related to three groups B- economies, emissions, and simulation — has significantly increased from 1988 to 1998. These categories call for further analysis to identify the reasons for growth. Conversely, publication activity has declined notably for two groups — internal combustion engines (FAC 3 - ICE) and vehicle fuels (FAC 7).

Similar publication activity trends can be observed (Figure 4) for the sub-groups under "Diesel", Figure 5. Again, further analyses would be necessary to attempt to explain the relevance of the decline in publication activity level (e.g. perhaps the subject technologies have matured and moved to the next life cycle stage, or possibly the research trail has proven unproductive and been largely

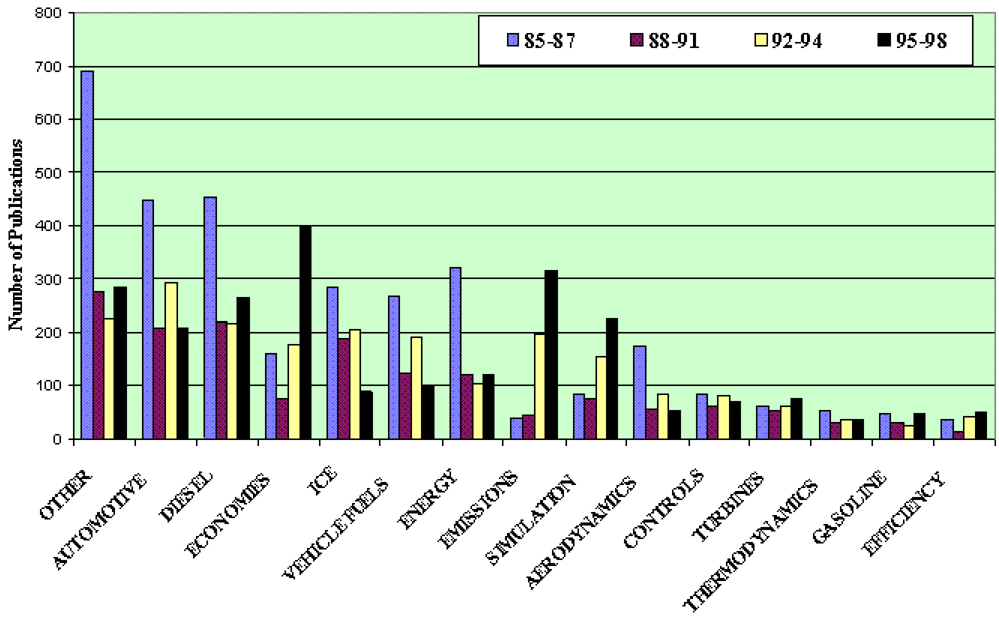


Fig. 4. Fuel Efficiency (ENGI); top level groupings' publication chronologies.

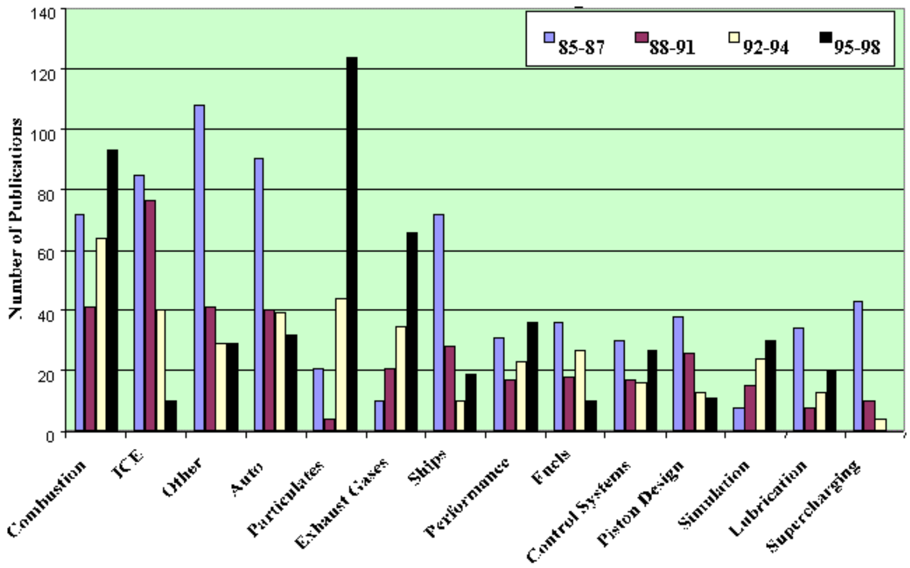


Fig. 5. Fuel efficiency (ENGI); diesel sub-groups' publication chronologies.

abandoned). Most germane to the goals of this functional analysis, Figure 4 exposes the emerging emphases of diesel research since 1995 — especially particulates and exhaust gas research thrusts, Figure 5. Note that these sub-groups are dominant both in terms of activity (number of 1995–98 articles including their keywords) and in rate of growth (increase from 1988–91, particularly). These spotlighted R&D areas warrant in-depth examination of the research underway and its potential ramifications for our fuel efficiency interests.

In our analyses, we performed a variety of such breakouts to examine clusters of basic research and patent activity as well. For instance, we noted a spurt of basic combustion-related research in 1995–98 that merited further examination. We also found that combustion patents comprise one of the three main categories of intellectual property being protected within the overall fuel efficiency search set. From this record grouping, five combustion patent sub-groups were generated (not shown). Further investigation helped isolate technologies that might offer complementary insertion potential for current fleet diesel engines. The results will be discussed in the conclusion section.

We continue the TOA analysis by considering further the applied research "Diesel" group of literature abstracts. Because of relevance to the Army — diesel is the primary vehicular power source — we continue to mine for innovations here. For present case illustration, we focus on the combustion research sub-group under "Diesel" (Table 2). The records from organizations that published in both the 1992–1994 and 1995–1998 time frames were subjected to detailed review of the abstract phrases describing their work. We feel these organizations demonstrate current knowledge and longevity. Our investigation focused on recent terms — i.e. terms used during the 1995–1998 period but not during 1992–94. We then mined the segmented abstracts' texts. The results are discussed in the Conclusion section. (Automated processes to replicate the procedures followed are being developed for future versions of the TOA software.)

We also applied "list comparison" of the more recent terms to create a sub-group of the applied research (ENGI) Diesel group. This process focuses the analysis on the 191 records in the 1995–98 period containing new keywords. These abstracts appear highly relevant (based on expert/peer review opinions) toward defining state-of-the-art research.

Table 3
Combustion technologies relating to fuel efficiency

Combustion Focus (INSP, ENGI & PATS)		
Basic Research - Top Level: Combustion Category & Several "Control" Groupings	**Applied Research - Top Level: Diesel & Control Categories, Combustion 2nd Tier Group**	**Dev./Patents - Top Level: Combustion Category, Subject Index Terms as Group Pointers**
Variable Displacement	**Electronic Control - Injection Timing & Duration**	**Combustion Chamber Turbulance Device**
Adaptive Fuel Control	**Stratified Charge Engine**	**Swirl Control for Intake Air**
Plasma Jet Ignition	**Variable Resonance Intake Manifold**	**Throttle Sensed Actuation for Fuel Injectors**
Fuel Oxidation Catalysts	**Injection of Dilute Oxidizers**	**Metered/Controlled Fuel Injector**
Stratified Diesel	**Fuel Heating System**	**Inlet Valves with Igniter**
Exhaust Gas Recirculation	**Common-Rail Fuel Injection**	**Stratified Combustion**
Combustion Simulation & Measurement Techniques	**Exhaust Gas Recirculation w/Intercooler**	**Hybrid Engine/Air Compressor (Multiple Contol Req.)**
Regeneration/Low Heat Rejection	**Variable Geometry Turbocharger**	**Fuel Catalyzers/Atomizers/Mods**
INSP "NOT" ENGI Keywords: **Chemically reactive flow, flow simulation flow control, control system analysis, digital control, chemical variables control, flow visualization, computerized control, multivariable control systems, flow through porous media, swirling flow & gas mixtures**	**Diesel(92-94) vs (95-98): "AND" Affiliations Using "NOT" Keywords. NLP Phrases used for Technologies/Approach Selection.**	**Diesel Keywords "AND" Comparison to Patent Abstract NLP Phrases used to Perform Principle Component Decomposition.**

Table 4
Fuel efficiency innovation forecast focused combustion recommendations

- **Electronic Control** is Enabling Technology
  1. New Engines
     Removes "Need" for Mechanical Linkage between Fuel/Air Delivery and Power Generation Systems (e.g., Camless Engine)
     Requires **Sensor Technologies**
  2. Existing Engines
     Enables **Smarter Fuel/Air Delivery** (Electronic Unit / Common Rail Injection Systems)
     Requires leverage of Technologies for **Improved Gas Mixture**, Igniters for **Stratified Combustion**, and **Impact/Simulation Models** for Preliminary Assessments
  3. Both: **Fuel Heating Systems & Catalyst** Study
- Expand TOA for More Fuel Efficiency Technologies

CONCLUSION

Table 3 summarizes the combustion-related technological approaches to fuel efficiency being pursued. These reflect search and analysis of three major R&D databases, cumulatively reflecting some 10,000,000 publications and patents over the past dozen years or so. The emphasis on combustion research results from both the interests driving this study and the empirical emphases seen in the literature. The first row in Table 3 provides the top-level observations related to the general emphases of basic research (estimated as publication activity in INSPEC), applied research (via publication activity indexed in ENGI), and development (gauged by patent activity in U.S. Patents).

The technological themes, documented in the middle eight rows of Table 3, identify specific approaches being developed, determined by manually reviewing abstracts from the 1995–97 time period. The number of publications and patents to be sought for detailed review, and possible networking with the authors, has been reduced to only the most relevant through the PCD and time segmentation processes. The last row in Table 3 synopsizes the key techniques applied to segment the relevant literature abstracts. Again for our basic research category (first column of Table 3), we list the INSPEC combustion group keywords that can be used as pointers to literature abstracts that document research on technologies yet to have transitioned into the applied research arena. These documents are relevant for TACOM, to help target its R&D investments at research areas both highly relevant to functional objectives and showing distinct developmental promise.

Table 4 presents selected strategic planning recommendations made to TACOM based upon this analysis and the objective of investing in technologies that both complement national fuel efficiency and Army diesel engine power density goals. The potential of electronic controls to achieve combustion fuel efficiency represents the most notable trend observed. Controls and/or control capabilities enable almost every technological approach (Table 3) or its system integration. Sensor technologies represent a critical complementary technology to controls. The remaining technology investment opportunities listed in Table 4 follow directly from observations listed in Table 3.

This TOA has provided TACOM management with function-oriented technology recommendations (i.e. R&D emphasis should be placed on electronic controls, sensor technology, etc.). This brings us back to the functional perspective with which we undertook this research. The next analysis step would be to perform more specific TOA-based analyses on each of these recommendations from a narrower, technology perspective. That process would begin by searching for each technology afresh. The keywords, or alternatives such as title words or abstract phrases, from the records would be reviewed to identify those terms that describe the technology space (e.g. for electronic controls pertinent to fuel efficiency). A literature search using these terms would be performed in suitable databases, and the TOA process would be repeated on each specific technology (e.g. as depicted in Figure 2 for diesel engine technologies). Recall that the literature abstracts used for this paper were obtained using the functional search terms related to fuel efficiency, not the specific combustion technologies identified through the analyses. For instance, rather than focus just on literature on "electronic controls" appearing within search sets conducted on "fuel efficiency," TACOM should reach out to profile electronic controls R&D broadly, then ascertain the technology with most potential for automotive engine fuel efficiency needs. This process will increase the potential for gaining insights from outside industries, as proved successful in our previous work (Watts and Porter, 1997).

TOA capabilities have been applied to identify promising technologies for fuel efficiency in this study. These are not the only TOA tools. TOA and other text mining software enables convenient concept mapping and generation of various innovation indicators (TPAC website, 1998). The selection of the tools and integrated processes that will be automated in TOAS over the next three years will be determined in conjunction with Beta site, mission-related applications. That is, we will work to enhance the TOA tools and their ease of application (automating where possible) based on feedback from various technology managers as to what information they want. A long-range interest is in knowledge warehousing — i.e. retaining the results of specific analyses such as this, for later updating and cross-linking with other analyses.

This study was not restricted to just using TOA. We have alluded to, but not emphasized, the importance of obtaining substantive expertise to balance "text mining." The purpose of the types of analyses in this paper are to help meet the information needs of strategic R&D program managers (assessing alternative technologies for possible developmental funding) and competitive technological intelligence users (to ascertain the strategies of in target technological domains). In this case, the fuel efficiency analysis is informing current research procurements, but the developmental outcomes are not yet known.

This case analysis illustrates emerging analytical capabilities to exploit increasingly available electronic information resources concerning technological development. The TOA tools applied herein are most effective in profiling and identifying relationships from extensive searches of bibliographic databases. In so doing, they enable the generation of effective intelligence based on patterns discerned from thousands of publications and patents. Such TOA results are well complemented by Internet probes. For instance, TOA can be used to identify prolific researchers on a topic. One may want to follow up by seeking the worldwide web sites of those researchers to identify more recent work and to make direct contact.

Formulating analyses to cast a broader net, as illustrated by the case's focus on a functional target, rather than a specific technology, can empower a more systemic management perspective. Such a system perspective is well-served by capabilities to digest large amounts of information. Recognition of actionable knowledge from exploiting such information resources will distinguish effective technology managers as the "Information Economy" unfolds.

REFERENCES

Ashton, W.B., Johnson, A.H. & Stacey, G.S. (1996). Monitoring science and technology for competitive advantage. *Competitive Intelligence Review. 7*, S115–S126.

Courtial, J.P., Sigogneau, A. & Callon, M. (1997). Identifying strategic sciences and technologies through scientometrics. In W.B. Ashton & R.A. Klavanas (Eds.), *Keeping Abreast of Science & Technology: Technical Intelligence for Business.* (pp. 337–371).

Cunningham, S.W. (1997). *The content evaluation of british scientific research.* PhD. Thesis. Science Policy Research Unit, University Sussex, Brighton, UK.

Cunningham, S.W. (in press). A maximum likelihood approach to the mapping of text. *Journal of the American Society for Information Science.*

Dou, H. & Dou, J.-M. (1995). Virtual research universities and research management. *Technology Analysis and Strategic Management. 7*(1), 109–122.

Dou, H., Quoniam, L. & Hassanaly, P. (1991). The scientific dynamics of a city: a study of chemistry in Marseille from 1981 to the present. *Scientometrics. 22*(1), 83–93.

Franklin, J.J. & Johnson, R. (1988). Co-citation bibliometric modeling as a tool for S&T policy and R&D management: issues, applications and developments. In A.F.J. van Raan (Ed.), *Handbook of Quantitative Studies of Science and Technology.* New York: North-Holland.

Garfield, E., Malin, M.V. & Small, H. (1978). Citation data as science indicators. In Y. Elkana, J. Lederberg, R.K. Merton, A. Thackray & H. Zuckerman (Eds.), *Toward A Metric of Science: The Advent of Science Indicators.* New York: Wiley.

Kostoff, R.N. (1993). Co-word analysis. In B. Bozeman & J. Melkers (Eds.), *Evaluating R&D Impacts: Methods and Practice*. (pp. 63–78) Boston: Kluwer.

Kostoff, R.N. (1994a). Database tomography: origins and applications. *Competitive Intelligence Review. 5*(1), 48–55.

Kostoff, R.N. (1994b). Research impact assessment: problems, progress, promise. In *Fourth International Conference on Management of Technology.* Miami, FL 423–433.

Kostoff, R.N. (1997). Database tomography for technical intelligence: analysis of the research impact assessment literature. *Competitive Intelligence Review. 8*(2), 63–79.

Kostoff, R.N. (1998). *Science and Technology Metrics*. This and other insightful reviews by Kostoff are available on his website: http://www.dtic.mil/dtic/kostoff/index.html.

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

Newman, N.C., Porter, A.L. & Cunningham, S.W. (1997). Technology opportunity analysis for malaysia. *Portland International Conference on Management of Engineering and Technology (PICMET).* Portland, OR.

Porter, A.L., Roper, A.T., Mason, T.W., Rossini, F.A. & Banks, J. (1991). *Forecasting and Management of Technology*. New York: Wiley.

Porter, A.L., Jin, X.-Y., Gilmour, J.E., Cunningham, S.W., Xu, H., Stanard, C. & Wang, L. (1994). Technology opportunities analysis: integrating technology monitoring, forecasting & assessment with strategic planning. *SRA Journal (Society of Research Administrators). 21*(2), 21–31.

Porter, A.L. & Detampel, M.J. (1995). Technology opportunities analysis. *Technological Forecasting and Social Change. 49*, 237–255.

Porter, M.E. (1985). *Competitive Advantage.* (pp. 273–314). New York: The Free Press.

Rip, A. (1988). Mapping of science: possibilities and limitations. In A.F.J. van Raan (Ed.), *Handbook of Quantitative Studies of Science and Technology.* New York: North-Holland.

Small, H. & Griffith, B. (1988). The structure of scientific literatures. *Science Studies. 4*, 7–40.

Tijssen, R.J.W. & van Raan, A.F.J. (1994). Mapping changes in science and technology. *Evaluation Review. 18*(1), 98–113.

TPAC Website — Technology Policy and Assessment Center: http://tpac.gatech.edu [note “TOA” section and examples on “Fuel Cells” and “KDD/Data Mining”].

Van Raan, A.F.J. (1992). Advanced bibliometric methods to assess research performance and scientific development: basic principles and recent practical applications. *Research Evaluation. 3*(3), 151–166.

Van Wyk, R.J. (1997). Innovation forecasting. *Technological Forecasting and Social Change. 56*, 25–47.

Watts, R.J. & Porter, A.L. (1997). Innovation forecasting. *Technology Forecasting and Social Change. 56*, 25–47.

Watts, R.J. (1996). *Defining next generation technologies.* Management of Technology Thesis, National Technological University, Fort Collins, CO.

Watts, R.J., Porter, A.L., Cunningham, S. & Zhu, D. (1997). TOAS intelligence mining; analysis of natural language processing and computational linguistics. In J. Komorowski & J. Zytkow (Eds.), Principles of Data Mining and Knowledge Discovery. *First European Symposium PKDD ’97.* 323–334.

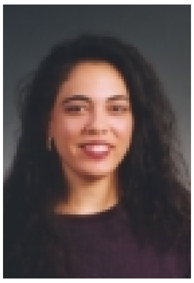
Wenk, E., Jr. & Kuehn, T.J. (1977). Interinstitutional networks in technological delivery systems. In J. Haberer (Ed.), *Science and Technology Policy* (pp.153–175). Lexington: Lexington Books.

Zanasi, A. (1998). Competitive intelligence through data mining public sources. *Competitive Intelligence Review. 9*(1), 44–54.

Zhu, Donghua, (1998). Summary of mapping improvements and development. TPAC Internal paper.

Zhu, Donghua, (1997). Comparison in grouping techniques (No. I–No. V). TPAC Internal paper.

Alan L. Porter received the B.S. degree in Chemical Engineering from Caltech and the Ph.D. from UCLA in Engineering Psychology. He has served on the faculty of the University of Washington and Georgia Tech, where he is currently professor of Industrial & Systems Engineering, and Public Policy. He directs the Technology Policy and Assessment Center at Georgia Tech. Current research interests focus on development of text mining tools in support of technology management.



Cherie R. Courseault received a B.S. degree in Industrial Engineering from Stanford University. Afterwards, she worked for Ernst & Young LLP in Management Consulting: Process Improvement. She is currently a PhD student at Georgia Tech in Industrial Engineering, with a focus on utilizing bibliometric analysis in technology management and a research assistant for the Technology Policy and Assessment Center.

Robert J. Watts currently works for the Tank-automotive and Armaments Command (TACOM), National Automotive Center (NAC), managing the program planning for the next generation engine technologies and the development of a competitive intelligence software system. Prior to his NAC assignment, he served as Science Adviser to the Commanding Generals of both VII Corps and 7^{th} Army Training Command (7ATC). He has a bachelor of science in aeronautical engineering from Ohio State University and a masters in industrial engineering from Texas A&M. He received a second masters in management of technology from the National Technological University (NTU). The synopsis of his NTU master's thesis received "Best Student Paper" recognition at the Portland International Conference on Management of Engineering and Technology (PICMET '97).