

Portland State University

PDXScholar

Engineering and Technology Management
Faculty Publications and Presentations

Engineering and Technology Management

1-1-2005

Mining PICMET: 1997-2005 Papers Help You Track Management of Technology Developments

Alan L. Porter

Robert J. Watts

Timothy R. Anderson

Portland State University, tim.anderson@pdx.edu

Follow this and additional works at: https://pdxscholar.library.pdx.edu/etm_fac



Part of the [Engineering Commons](#)

Let us know how access to this document benefits you.

Citation Details

Porter, A.L., Watts, R.J., and Anderson, T.R., Mining PICMET: 1997-2005 Papers Help You Track Management of Technology Developments, Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, 2005.

This Article is brought to you for free and open access. It has been accepted for inclusion in Engineering and Technology Management Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

Mining PICMET: 1997-2005 Papers Help You Track Management of Technology Developments

Alan L. Porter¹, Robert J. Watts², Timothy R Anderson³

¹Search Technology, Inc., USA

²U.S. Army Tank and Automotive Research, Development & Engineering Center, USA

³Department of Engineering and Technology management, Portland State University, USA

Abstract—PICMET is providing the abstracts from 1997 -- 2005 to enable you to "mine" the contents of these four conferences. This enables you to overview activity in the field of Management of Technology (MOT) and to identify specific research of interest.

I. INTRODUCTION

This effort derives from exploratory analyses of the 1997, 1999, and 2001 PICMET abstract records [18]. That proved sufficiently rich to experiment with providing the PICMET abstracts as a resource to Conference attendees in 2003 [20]. This paper updates that one to include the 369 abstracts accepted for PICMET 2005 as of April 4, 2005. The abstract records of the 1997-2005 conferences were provided by Portland State University. Search Technology, Inc. applied its *VantagePoint* software to "text mine" those records. These are provided to you along with the *VantagePoint Reader* software* that enables you to find materials of interest.

VantagePoint Reader allows you to browse and select subsets of the 2068 abstract records of interest. This works somewhat analogously to *Acrobat Reader* -- you can see trends, investigate "who's doing what?", and dig down to read particular abstracts that you want. But you can't create new files, perform new analyses, or export record sets. To perform such new analyses you would need the *VantagePoint* software [for sale -- see <http://www.theVantagePoint.com>; also available at no cost as *TechOASIS* only to U.S. Government personnel]. Beyond the capabilities you have at hand using the *Reader*, *VantagePoint* is also able to seek relationships based on the co-occurrence of particular terms [2, 5]. Principal components analysis and variants can generate various "maps" -- e.g., of term clusters or authors -- to help perceive activity patterns. At the PICMET tutorial, we illustrate some such analyses.

The next section provides a brief background on text mining in support of R&D management. The following section explains how you can use the PICMET data in support of your research interests.

II. BACKGROUND ON TEXT MINING

* We will provide CDs with the *VantagePoint Reader* software to those attending the tutorial at PICMET. If you miss that, you have the data file on your PICMET CD. **General instructions for obtaining, installing, and using the software are given in this paper's Appendix.**

This section provides a few pointers to those of you who might want to further explore "bibliometrics" and "text mining" for technology management. These studies of science through looking at its outputs (papers) trace back to Price [13]. Bibliometrics counts such publication (or patent) activity to detect trends and changing patterns [10]. Text mining carries this further to process the content of those papers, aided by natural language processing [7, 8].

We are particularly interested in tracking advances in a field (in this case, "MOT") through profiling research activity changes over time [17, 19].[†] Such analyses can "map" changes so as to aid R&D management and other aspects of technology management [9, 14, 15]. They can contribute to research evaluation [6]. The scale of such inquiries can range from very specific technical domains [12] to national R&D profiles [4]. Knowing R&D emphases can also help inform competitive technological intelligence [1] and technology forecasting [16]. A recent book treats technology analyses using such approaches [11].

III. MINING THE PICMET ABSTRACTS

A. Limitations

The data were not all compiled using the same format. Hence, you will see that the records differ. Some information is present only for certain of the conferences. In particular, information on the authors' organization and country is only available since 2001.

Processing of the records is not 100% error-free. You may see mistakes in identifying "authors" as being the same or different persons; likewise, for organizations. The "MOT-relevant" phrases are solely the judgment of Porter and Watts. They reviewed frequently occurring phrases from the titles and abstracts (i.e., those occurring 5 or more times) and selected 406 that they thought might hold interest to some MOT colleagues.

B. Using the PICMET Abstract Data

First, let us examine a summary of the data. Fig. 1 shows the content on your CD. The fields that you can access are listed. We have parsed the titles and abstracts to list the noun phrases contained in the 2068 papers. We have names of authors, their organizations, and countries. And we have the year in which the paper was presented.

[†] See also Georgia Tech's Technology Policy and Assessment Center website: [//tpac.gatech.edu](http://tpac.gatech.edu), particularly pages on "Technology Opportunities Analysis" and "HotTech."

Number of records:	2068	Source database:	PICMET + PICMET 1997-1999 + PICMET 2003 + PICMET 2005	
Source date:		Source file:	C:\Documents and Settings\aporter.SEARCH\My Documents\Alan\Papers-in-pr...	
Field	Number of items	Derived	Data Type	Meta Tags
~Raw Record	2068			
Abstract	2060			
Abstract (NLP) (Phrases) + Title (NLP) (Phrases)	30067	*		
Author	2842			
Country	89			
Organization of Author	784			
Organization-AcIndGov	265			
Title	2065			
Title Phrases	4787	*		
Year	5		Year	

Figure 1. Summary of imported PICMET records.

Fig. 2 shows the top of Author list as a matrix by year. Clicking on an author's name shows each of the paper titles authored. Since this table is ordered by decreasing # of Records, we can see that Hannu Jaakkola had the most presentations at PICMET between 1997 and 2005 and is closely followed by David Probert. The table also makes it

visible how the number of presentations changed from year to year. When you click on Hannu Jaakkola, he is highlighted and all of his 18 paper titles are listed on the left. If you wanted to view just his presentations from one year, clicking on the corresponding cell to the right of his name would show just those papers for that year.

Reset		Author	1	2	3	4	5
		# Records	481	449	401	368	369
		Show Values >= 1					
		Cooccurrence					
		# of Records	1997	1999	2001	2003	2005
1	20	Jaakkola, Hannu	3	3	10	2	2
2	18	Probert, David		3	4	5	6
3	16	Wilemon, David L.		4	6	3	3
4	16	Xu, Qingrui		2	6	5	3
5	16	Basoglu, Nuri			2	5	6

Figure 2. Listing of Authors by Year, sorted by record frequency.

The titles of the presentations for the highlighted author are listed on the left. Clicking on any of the paper titles opens a window (adjustable size) containing the information on the submission. This is illustrated by Fig. 3 which shows

the complete information contained in the record as well as any additional notes that might be available or that you might want to add yourself.

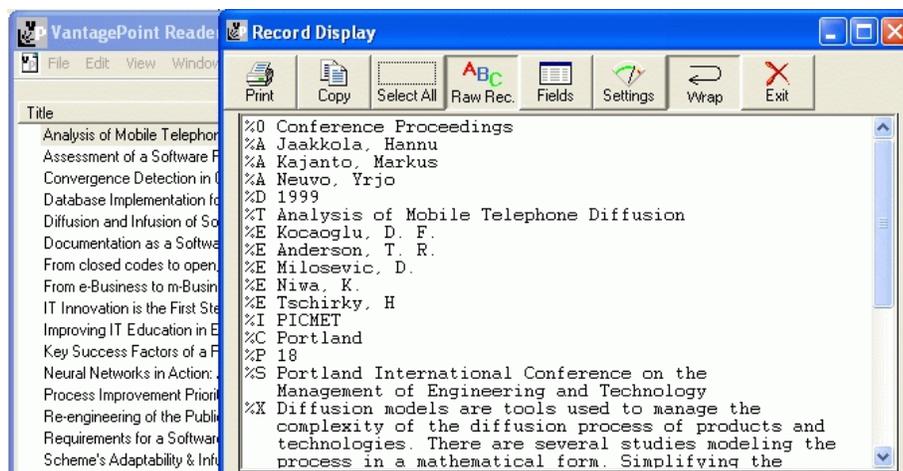


Figure 3. Detailed record/article information.

Fig. 4 lists the organizations with the greatest number of records (papers with at least one author affiliated with that organization). Since the 1997 and 1999 data did not include author affiliation, these results are only for the 2001-03-05

conferences. One interesting thing to note is that the top seven most active institutions come from six different countries. This highlights the international nature of PICMET and the field of Technology Management.

Reset		Organization of Author	1	2	3	4	5
		# Records	481	449	401	368	369
Year	# Records	Show Values >= 1					
		Cooccurrence # of Records	1997	1999	2001	2003	2005
1	64	Portland State University			34	16	14
2	48	ETRI			10	16	22
3	30	ZHEJIANG UNIVERSITY			9	6	15
4	24	Stevens Institute of Technology			3	7	14
5	22	Swiss Federal Institute of Technology Zuri			4	13	5
6	21	University of Cambridge			4	10	7
7	20	National Chiao-Tung University			3	8	9

Figure 4. Listing of presentations by institution.

Fig. 5 shows how we can examine areas to see which technology management terms are the most frequently used and also examine the changing patterns of usage over time. Be wary – our term parsing has evolved somewhat and the records were analyzed at different points in time. We have aggressively combined term variations here. Still, there are over 28,000 terms and phrases listed. The Appendix offers

pointers on searching. This figure highlights that the term “china” is being used more frequently. This approach could be used to identify subfields of technology management that are quickly developing as well as those that may be on the decline. Such information could help you situate your own research in "ripening" areas. The Title View here shows titles of China-affiliated papers from 2005.

Reset		Abstract + Title Phrases (cleaned)	1	2	3	4	5
		# Records	481	449	401	368	36
Year	# Records	Show Values >= 1					
		Cooccurrence # of Records	1997	1999	2001	2003	2005
62	69	Informational Technology	9	19	14	15	12
63	68	china	12	8	9	15	24
64	66	university	12	16	10	15	13
65	66	technology news	15	6	5	23	17
66	66	future	8	12	17	16	13
67	65	country	10	11	8	19	17
68	64	practices	7	10	18	16	13
69	64	surveys country	18	12	13	9	12

Figure 5. Phrases used in the abstracts and titles of presentations.

Fig. 5 showed how term usage changes by year. Additional *VantagePoint Reader* file tabs (that open other worksheets, much like *MS Excel* worksheets) would allow the researcher to examine the usage of these or other terms by author, organization, or country. Both industry analysts and academicians could find this to be an important tool to highlight opportunities for rich new areas of research, point out potential collaborators, and warn of competitors.

The issue of collaboration is an interesting one. To investigate this, a matrix showing the co-authorship relationship among countries is given in Fig. 6. [We could

have shown association among authors or organizations.] Highlighted are British joint papers. On the right, we've opened two Detail Windows. These present auxiliary information (fields of your choosing) for those highlighted cells. Here we see authors and dates of those collaborations. [The Up Arrows in the Authors Detail indicate that the frequency of these authors, compared to their frequency in the whole dataset is especially high. That is not meaningful for this breakout, but could be, were we exploring what MOT topics a country or organization emphasizes.]

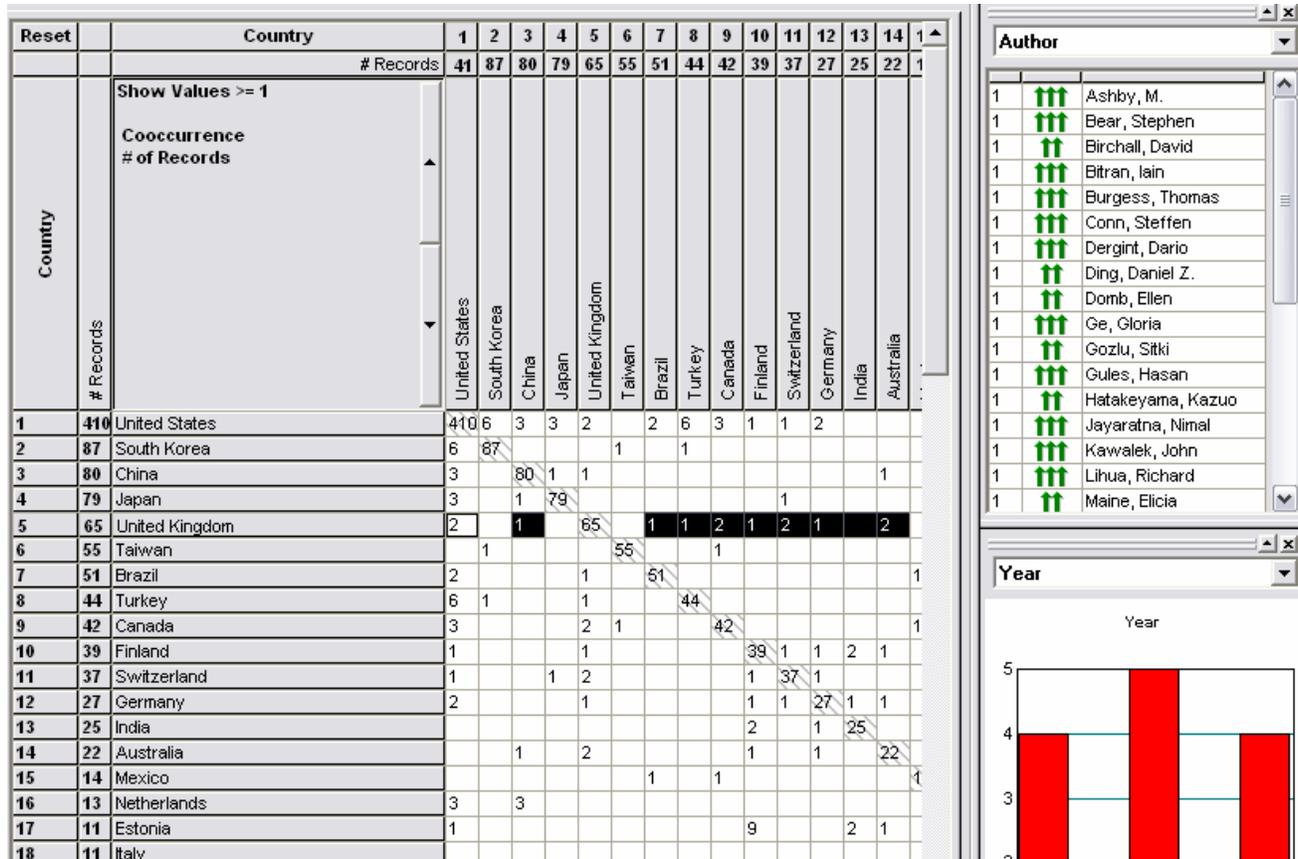


Figure 6. Matrix showing co-authorship of presentations by Country.

We are keenly interested in developing “innovation indicators” [11]. These seek to track empirical changes relating to an emerging technology’s maturation, contextual influences, and market prospects to help forecast its development. We generate one such indicator in the *VantagePoint Reader* file to illustrate. This is the extent of publication by sector: academic, industry, and government. We have not exhaustively classified each organization, but could easily get about half. The matrix by Year suggests some decline in industrial participation at PICMET. That by Country shows Japan with the strongest industrial presence at PICMET. The matrix of top MOT terms by organizational type indicates the extent of industrial interest in various

PICMET topics (markets and services lead with 9 each) (Fig. 7). “Competitiveness” is the leading term in governmental papers at PICMET. Not surprisingly, academic contributions constitute the largest share.

Lastly, Fig. 8 illustrates a factor map (interactive in the software). Using a Principal Components Analysis, presentations were clustered by their usage of 211 interesting MOT terms (in our judgment) in the title and abstracts. Nodes represent principal components, or “factors,” named after the most central term (abstract & title MOT phrases). The pull-down boxes give the terms most associated together to form that factor. Note the variations on New Product Development clustered together.

Abstract + Title Phrases (cleaned)		1	2	3
# Records		893	107	68
# Records	Show Values >= 1	Academic	Industry	Govt
	Cooccurrence # of Records			
119	markets	63	9	3
92	services	47	9	3
93	competitiveness	45	8	9
52	product development	24	8	1
61	Japanization	34	7	2
130	knowledge	66	6	3
47	operations	26	6	2
89	quality-	35	6	3
48	managing knowledge	33	5	2
69	competitive advantages	43	5	1
31	Cultures	16	5	

Figure 7. Extent of academic, industrial & governmental topical activity

The top-most blue circle in Fig. 8 corresponds to the "Roadmaps" group. The terms listed in the adjacent pull-down box are those that show the strongest statistical association with this factor in the collection of 2068 presentations.

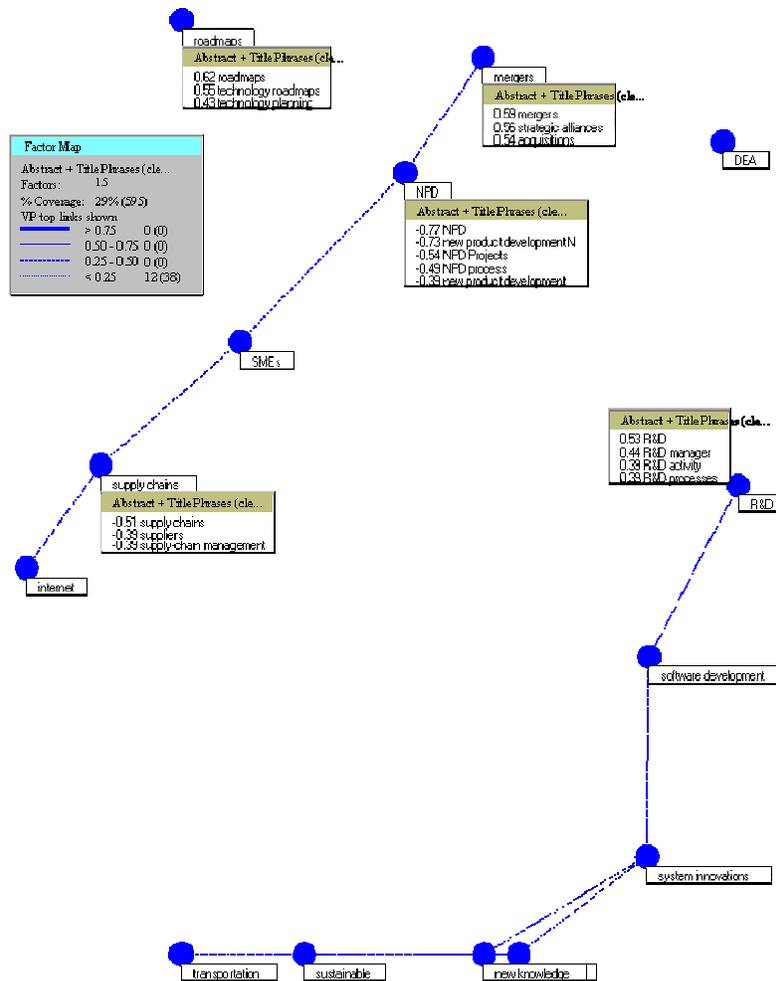


Figure 8. Factor map of MOT topical clusters

The map used Multi-Dimensional Scaling (MDS) to locate nodes. In this example there are 14 nodes, so accurate scaling in 14 dimensions cannot be retained in a 2-D representation. To better depict distance among the nodes, links are shown. These reflect a path-erasing algorithm to indicate greater or lesser similarity. To the left in the *VantagePoint Reader*, you can see the list of papers with their relevance scores to a given topic on which you click.

The List of Abstract (NLP)(Phrases) + Title (NLP) (Phrases)(Cleaned) shows the group on which the map is based (top MOT). It also gives the terms that cluster together on each of the 14 factors. So, you could pursue who is researching any one of these thrusts by examining that group.

There are many possible mapping variations using *VantagePoint*. Our paper at PICMET 2005 explores ways to enhance the representations possible by using externally provided keywords [21; see also 3]. We have done such an analysis on the PICMET abstracts through 2003 using "MOT" terms. This includes more of the records than does the map in Fig. 8. Such maps help one perceive the overall research field; they can also be used to focus into subsets of most relevance to your work.

IV. CONCLUSION

This paper demonstrates how text mining software can be used to examine trends and patterns of research in the fields of technology management using software developed specifically for these types of knowledge mining applications. Specific instructions on installing and using the *VantagePoint Reader* are provided in the Appendix.

LET US KNOW WHAT YOU THINK OF THIS!

Tell us what was and what was not useful. Was a particular function difficult for you to use? Should PICMET do this again in 2007? If so, is there additional information you would like from future PICMET abstract compilations? Please e-mail your suggestions to Alan Porter at: aporter@searchtech.com

REFERENCES

- [1] Ashton, W.B., A.H. Johnson, and G.S. Stacey, "Monitoring Science and Technology for Competitive Advantage," *Competitive Intelligence Review* 7: S115-S126, 1996.
- [2] Callon, J., J.P. Courtial, and F. Laville, "Co-word Analysis as a Tool for Describing the Network of Interactions Between Basic and Technological Research -- The Case of Polymer Chemistry," *Scientometrics* 2(1), 155-205, 1991.
- [3] Chen, H., K.J. Lynch, B. Koushik, and T.D. Ng, "Generating, Integrating, and Activating Thesauri for Concept-based Document Retrieval," *IEEE Expert*, 25-34, 1993 (April).
- [4] Cunningham, S.W., *The Content Evaluation of British Scientific Research*. PhD. Thesis. Science Policy Research Unit. University of Sussex, Brighton, UK, 1997.
- [5] Kostoff, R.N., "Co-word Analysis," in B. Bozeman, and J. Melkers, (Eds.), *Evaluating R&D Impacts: Methods and Practice*, Kluwer, Boston, 63-78, 1993.
- [6] Kostoff, R.N., "Research Impact Assessment: Problems," Progress, Promise, in *Fourth International Conference on Management of Technology*. Miami, FL, 423-433, 1994.
- [7] Kostoff, R.N., and E. Geisler, "Strategic Management and Implementation of Textual Data Mining in Government Organizations," *Technology Analysis & Strategic Management* 11(4), 493-525, 1999.
- [8] Kostoff, R.N. "Technology Watch and Evaluation," http://www.onr.navy.mil/sci_tech/special/technowatch/, March 13, 2003.
- [9] Leydesdorff, L., "Various Methods for the Mapping of Science," *Scientometrics* 11, 295-324, 1987.
- [10] Narin, F., D. Loivastro, and K.A. Stevens, "Bibliometrics -- Theory, Practice and Problems," *Evaluation Research* 18(1), 65-76, 1994.
- [11] Porter, A.L., and S.W. Cunningham, *Tech Mining*, Wiley, New York, 2005.
- [12] Porter, A.L., A. Kongthon, J.C. Lu., "Research Profiling: Improving the Literature Review," *Scientometrics* 53, 351-370, 2002.
- [13] Price, D. de S., *Big Science, Little Science*, Columbia University Press: New York, 1963.
- [14] Rip, A., "Mapping of Science: Possibilities and Limitations," in A.F.J. van Raan, (Ed.), *Handbook of Quantitative Studies of Science and Technology*, North-Holland, New York, 1988.
- [15] Tijssen, R.J.W. , and A.F.J. van Raan, "Mapping Changes in Science and Technology," *Evaluation Review* 18(1), 98-113, 1994.
- [16] Watts, R.J., and A.L. Porter, "Innovation Forecasting," *Technological Forecasting and Social Change* 56, 25-47, 1997.
- [17] Watts, R.J., and A.L. Porter, "Requirements-based Knowledge Discovery for Technology Management," *Portland International Conference on Management of Engineering and Technology (PICMET)*, Portland, OR, 2001.
- [18] Watts, R.W., and A.L. Porter, "Tracking the Evolution of Management of Technology (MOT)," *International Association for the Management of Technology*, Miami, 2002.
- [19] Watts, R.W., and A.L. Porter, "R&D Cluster Quality Measures and Technology Maturity," *Technological Forecasting and Social Change*, 70 (8), 735-758, 2003.
- [20] Watts, R.W., A.L. Porter, and T.R Anderson, "Mining PICMET: 1997-2003 Papers Help You Track Management of Technology Developments," *Portland International Conference on Management of Engineering and Technology (PICMET)*, Portland, OR, 2003.
- [21] Watts, R.W., and A.L. Porter, "Technical Text Content Clustering Enhanced by Protected Natural Language Processing," *Portland International Conference on Management of Engineering and Technology (PICMET)*, Portland, OR, 2005.

APPENDIX: HOW TO USE THIS RESOURCE:

First, obtain the VantagePoint Reader from Search Technology, Inc. The reader is available to PICMET 2005 participants free of charge by e-mailing picmet2005@searchtech.com providing a) your name, b) your current corporate or educational affiliation, and c) your current location (city/state/country). Second, open the "PICMET-97-99-01-03-05" file from the 2005 PICMET Conference CD-ROM.

You now have access to a set of sheets -- rather analogous to worksheets in *MS Excel*.

Look at the "Summary" sheet -- it indicates the available fields derived from the records and the number of entities in each. Want to know how many papers were presented at PICMET? Here's an answer. After the summary, the following "worksheets" provide a series of lists and matrices.

You can use this information to get a "bird's eye" perspective of what is happening in our research arena. For instance, from the **lists**, you can generate:

- "who" information -- which authors and organizations have been most active at PICMET?
- "what" information -- which terms & phrases have been most popular, and
- "when" information -- looking at changes over time (conference **year**)
- "where" information -- which countries contribute most actively.

[But note the caveat on some information not being available equally for all 5 conferences.]

The **Record View Window** enables you to "dig down" into the information to locate details on topics of interest. For interest, you can select one or more entries from a list, or one or more cells in a matrix. To the left, you will see the titles of the abstracts that pertain (**Title Window**). If you double-click on a title, you will open up a record window to see that abstract.

Matrices combine any two of the lists:

1. authors
2. organizations
3. countries
4. abstract & title phrases,
5. abstract & title phrases deemed most interesting for MOT (406 phrases)
6. year

For instance, you might want to see who from South Korea has addressed "supply chain management" at

PICMET? In the "*Matrix: Abstract+Title PhrasesXCountry*" sheet, you can sort the entries in the "South Korea" column (by double clicking in this column's header) and locate "supply chain management."

Or, to check how "hot" supply chains are at PICMET, you could go to the "*Abstract+Title Phrases X Year*" matrix. Double-clicking on the Abstract+Title Phrases column alphabetizes. Double-clicking again reverses the order, making it easy to find the supply chain items and see that interest remains strong.

"Detail Views," on the right side of the screen, show aspects concerning entities you select ("click on"). Under the "View" menu, be sure "Details Window" is checked. Then check "Create Details Window." You can have several Detail Views open and you can reposition them as you want (e.g., two columns of two views each).

Suppose you're planning to visit Japan and want to interact with researchers analyzing "Japanization." Go to the "*Matrix: Abstract+Title Phrases X Country*" sheet. Sort the entries in the Japan column. Go to the "Japanization" row cell -- 28 abstracts are there. In the Detail View to the Right, pull down to see whatever field you want -- for instance, "Organization" shows Japan Sci & Tech Corp. and Tokyo Inst. Of Technology each with 3 related articles. If you click on these in the Detail View, the titles of their articles are highlighted in the Title View (to the left -- you can open or shut such Views from the "View" menu). You can double-click on titles to read the abstracts in a Pop-up Window.

You can add more Detail Views. For instance, a second Detail View might show "Year" while the first Detail View showed "Author." You can resize windows as you like.

Note *VantagePoint Reader "Help"* on the top menu. Menu pull-downs also enable you to perform various usual functions. Under the **"File"** menu, you can

- Open and Save the PICMET abstracts file, incorporating changes you make
- Print

Under **"Edit"** you can:

- Copy -- to paste sheets or portions -- e.g., say you wanted to copy the "Top 10" MOT terms X Year into *MS Excel* to generate trends in MOT emphases
- Find

Under **"Window"** you can:

- Have 2 versions of the file open to facilitate comparisons (either as 2 windows or as a cascade or tile)
- See the list of available worksheets (and reorder them as desired).