Laurea University of Applied Sciences
Ratatie 22, 01300 Vantaa, Finland
Phone +358 (0)9 8868 7150
Fax +358 (0)9 8868 7200
firstname.surname@laurea.fi
www.laurea.fi
Business ID 1046216-1
Domicile Vantaa

This paper was presented at The XXVI ISPIM Conference – Shaping the Frontiers of Innovation Management, Budapest, Hungary on 14-17 June 2015. The publication is available to ISPIM members at www.ispim.org.

Research Topics at ISPIM: Popularity-based Scientometrics keyword analysis

Teemu Santonen*

Laurea University of Applied Sciences
Vanha maantie 9, 02650 Espoo, Finland
E-mail: teemu.santonen@laurea.fi

Steffen Conn

ISPIM
PO Box 18, Worsley, Manchester M28 1XP
E-mail: conn@ispim.org

* Corresponding author

Abstract: This study identifies and analyses emerging innovation management research topics by applying popularity-based scientometrics analysis to ISPIM full academic papers (N=1084) from 2009 to 2014. From a research method point of view this study introduces a novel comprehensive framework for classifying scientometrics studies and proposes "Variant keyword search" – process (VKS) to overcome known synonyms and tackle the separation of multiple terms problem during the keyword harmonization process. Most of the keywords (85 percent) after the basic harmonization process occurred only once and very few are significantly more popular than others. Grounded in the VKS-process the TOP50 research topics are presented. The VKS-process appears to be a powerful tool to compress and reduce useless variation within keywords as the TOP50 list contains 68 percent of all keywords and 60 percent of unique keywords. Unsurprisingly, identified research topics in the TOP50 list can be derived from various mainstream innovation classifications.

Keywords: scientometrics, research topic, popularity-based, ISPIM, keyword analysis, variant keyword search

1 Introduction

The scientific process builds on previous knowledge and itself refines existing knowledge to serve as the foundation for further research. Therefore, an in-depth understanding of scientific knowledge and its evolution in specific research paradigms is important. The concept of research paradigm and paradigm shift was popularized by Kuhn (1962). Later Dosi (1982) proposed that continuous innovation can be regarded as an event within a paradigm, whereas discontinuous innovation could be the starting point for a new paradigm. This fundamental idea has then been applied in various academic fields in order to understand the history, current state and future of the given paradigm (Gupta and Bhattacharya, 2004). The main goal of this study is to evaluate the current

status of the innovation management research paradigm by analysing research topics at ISPIM.

ISPIM as a scientific community (Bourdieu, 2004) connects individual researchers united by a common interest in the topic of innovation management (Trott, 2008; Baregheh et al. 2009) via various ISPIM events and publications including annual conferences, symposiums and forums. So far ISPIM as a network of innovation scholars and practitioners has not been extensively studied, although Santonen and Ritala (2014) recently evaluated the co-authorships relations within the ISPIM community. Their study revealed tight clustering based on geographical and institutional boundaries and found evidence of high-performing authors who span these boundaries using significantly different strategies. As a result we now have a good understanding about the underlying structure of the author network within the ISPIM community, but we do not really know what research topics the ISPIM community covers. Therefore, by way of this study, we are empirically verifying what kind of research themes have been covered by the ISPIM community.

This paper is organised as follows. Following this introduction we discuss the theoretical foundations of our study. Second, we present our data collection and data harmonization process. Third, we present our results and then finally, we conclude with our findings and discussion of theoretical and practical implications of our findings.

2 Theoretical background

2.1 Conference proceedings as a knowledge source

Conference proceedings as a relevant and important knowledge source have been recognized but also criticized by many authors (e.g. Lisée, et al. 2008). It appears that proceedings become obsolete faster and are less cited than scientific literature in general (Goodrum, 2001; Lisée, et al. 2008), but they also have other important objectives. Drott (1995) suggested that there are three main functions for conference publications including improving papers by gathering feedback, stimulating discussion within a research paradigm and providing novel information, which otherwise could be difficult to include other publications. In line with these arguments Montesi and Mackenzie (2008) suggested that conference proceedings can demonstrate an ability to innovate and propose new ideas. Moreover, recent studies have demonstrated that conference proceedings are usually only partially extended into journal publications and the level of utilization varies greatly across disciplines (Miguel-Dasit et al., 2006; Kho and Brouwers, 2009; Aleixandre-Benavent, 2009; Glanzel et al. 2006). As a result conference proceedings may contain insights which are not available elsewhere.

Interestingly, González-Albo and Bordons (2011) found evidence that the proceedings papers published via ordinary journal issues are similar in terms of structure and impact of research when compared to standard journal articles, whereas proceedings papers published in conference related special issues are less comprehensive and receive fewer citations. However, conference proceedings and the conference related special issues have “time to market” advantage, especially in the fast-evolving disciplines such as software engineering and computer science where it is a critical factor (Montesi and Mackenzie, 2008; Lisée, et al. 2008). In short, most authors argue that conference

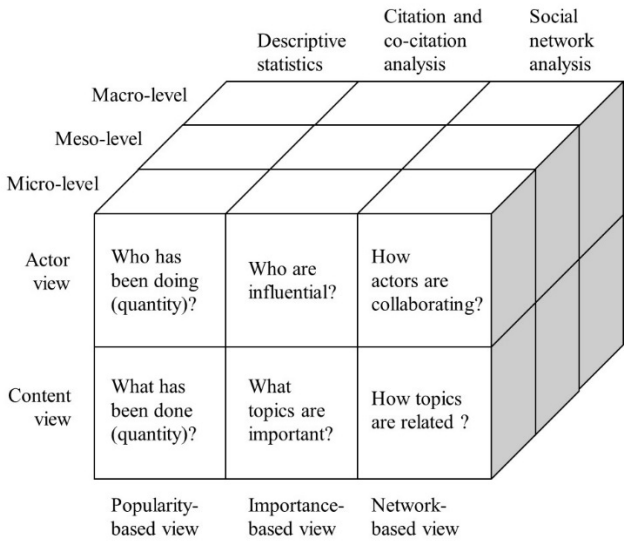
proceedings and journal articles can be considered as complimentary communication channels (González-Albo and Bordons, 2011; Butler and Visser 2006; Godin, 1998). As a result, we argue that ISPIM proceedings could be a store of novel research ideas and a platform to identify and evaluate emerging trends in innovation management research, which may only become visible in formal journal publications several years later.

2.2 Scientometrics as a research method

Typically, methodologies to examine research paradigms as well as our study are grounded in bibliometric analysis (Pritchard, 1969). Bibliometric analysis nowadays is often equated to scientometrics, (Larivière et al. 2012), which as a term we also prefer in our study. Recently Santonen (2015) illustrated a comprehensive framework for classifying various types and combinations of scientometrics studies (Figure 1) in which the cube is divided into 3x2x3 cells in which the three axis are representing following viewpoints:

- **Horizontal axis:** *Popularity* vs. *Importance* vs. *Network* dimensions
- **Vertical axis:** *Actor* vs. *Content* dimensions
- **Depth axis:** *Micro* vs. *Meso* vs. *Macro* dimensions

Figure 1: A comprehensive framework for classifying scientometrics studies (Santonen, 2015)



Broadly speaking, scientometrics studies can be classified as “popularity-based” (Choi et al, 2011) or “social network analysis”-based (later SNA) (Borgatti et al., 1992). Citation and co-citation analysis methods (Pilkington and Meredith, 2009) are also important research methods to conduct scientometrics studies, but depending on the selected methods, they can be considered either as popularity-based or network-based research methods. *Popularity-based studies* are typically analyzing frequency of people,

keywords or other related meta-terms, which have been derived from the context of the research publication.

The recent ISPIM conference proceedings article by Kristansen and Gertsen (2014) on radical innovation definitions is presented as an example of a study utilizing popularity-based method. *Citation and co-citation studies* focuses on the importance and impact of the people or topics by evaluating how much (popularity approach) or by whom (network approach) the particular study is cited (Pilkington and Meredith, 2009). The basic rule is "the greater the amount of citations is, the more influential and important the study is". *SNA-based studies* instead focus on the relationships via publications, which most typically are based on co-authorship (e.g. Su and Lee, 2012) or keywords (Yi and Choi, 2011). Overall, network-based studies have been successfully used to study various types of scientific communities from people's perspective (e.g. Newman 2001, Morlacchi et. al. 2005, Vidgen et. al. 2007) also including innovation communities such as global open innovation research (Su and Lee 2012) and co-authorships relations in ISPIM community (Santonen and Ritala, 2014). Recently, the usefulness of keywords and keyword networks as a fundamental carrier of knowledge has been recognized (Su and Lee, 2010) and related methodologies have been developed (Yi and Choi, 2011). All the above research viewpoints can be applied at the micro-level (e.g. individual authors) or the meso-level (e.g. universities that the authors represent) or at the macro-level (e.g. countries) (Gupta and Bhattacharya, 2004).

Prior scientometrics studies in the context of innovation management have mainly focused on author viewpoints, instead of identifying contextual research topics via keywords. Therefore, we argue that there is a clear research gap, which evaluates the innovation management research paradigm from both a keyword network and popularity point of view.

3 Research methodology

3.1 Objectives of the study

The main goal of this study is to analyze the current status of emerging innovation management research trends by applying popularity-based bibliometric analysis (Choi et al, 2011) to ISPIM proceedings from 2009 to 2014. As a result, we identify the key research topics with the help of keywords and present various descriptive statistics relating to ISPIM publications.

3.2 Data collection

The unit of analysis in this study is an ISPIM conference (and symposium) publication. The proceedings of each ISPIM event are made available to all event delegates and the past proceedings of all events since 2003 are available to ISPIM members via the ispim.org website. The meta-data in the publications post-2008 are in a more robust format and as such in this study we limit our time span to publications from 2009 to 2014.

The ISPIM events during this period produced 1525 publications, which can be further divided into full academic papers (N=1108), short academic papers/Academic Research Development Submission (ARDS) (N=336) or other publications (N=81) such as EU/Funded Project Submission, Practitioner Presentation and Special Interest Group Submission. The detailed descriptions for these submission types are available from the ISPIM website (ISPIM, 2015). As indicated in Table 1, full academic papers represent a great majority (72.7 percent) of the whole dataset and for past two years their share has increased to over 80 percent.

Table 1 Distribution of ISPIM full academic papers 2009-2014

		2009	2010	2011	2012	2013	2014	Total
Full Academic Paper	N	168	164	183	165	251	177	1108
Full Academic Paper	%	62.9	69.8	73.2	63.5	83.4	83.5	72.7
Annual total	N	267	235	250	260	301	212	1525
Annual total	%	17.5	15.4	16.4	17.0	19.7	13.9	100

In this study we focus only on full academic papers, which by definition are "accomplished, substantial and complete academic research results of an empirical or theoretical nature" (ISPIM, 2015, p. 3). In all, our dataset included 1108 full papers, but 24 of these submissions did not include any keywords and therefore were omitted resulting in a total of 1084 full academic papers. Strictly speaking this collection of ISPIM proceedings might not qualify as a big dataset (McAfee, 2012) when compared to some Library and Information Science studies covering more than century of publication data (Larivière et al. 2012). However, we argue that this dataset provides a rather comprehensive perspective and a good starting point to evaluate the most current research topics in the field of innovation management.

Each paper's keywords were manually coded to a database by a research assistant. Then 20 percent of the papers were double-checked by the two authors of this study in order to verify that the data was coded correctly. During this verification process no coding errors were detected.

3.3 Keyword cleaning and harmonization process

Since the format and definition of keyword is not unambiguous, there are multiple variations on how to list keywords with similar meanings. In order to overcome this problem a multi-staged keyword harmonization process was conducted as proposed in similar studies (e.g. Choi et al. 2011, Yi and Choi (2011). Practically speaking the harmonization process was conducted as a manual, iterative process with the help of various Excel formula rules and NVIVO word frequency and finding matches analysis, since this process cannot be automated reliably. The refinement of keywords included the following harmonization rules:

PHASE 1: Removing hyphens, **spaces and other special characters**: Software applications, which were used for analysis, have trouble to correctly interpret hyphens, spaces and all other special characters. Therefore all the particular characters were removed from the keywords and afterwards the list of unique keywords were constructed.

PHASE 2: Standardization into a singular form: Keywords having singular and plural forms were standardized into a singular form including popular terms such as innovation(s), service(s) and enterprise(s).

PHASE 3: Removing redundant keywords and avoidance of abbreviations: Popular examples for PHASE3 include: small and medium sized enterprise(s) (SMEs or SME) → small and medium sized enterprises and R&D → Research and development. Also in many cases abbreviations had been added as a separate keyword, which were removed.

PHASE4: Unification of synonyms and separation of multiple terms. When compared with previous phases, the execution of PHASE4 is not a straightforward process and not well described in prior studies at a practical level (e.g. Choi et al. 2011, Yi and Choi, 2011). Since our data includes scientific terms and definitions, finding an unambiguous process for PHASE4 is extremely difficult. It is a well-known fact that academics are keen to define intentionally and unintentionally parallel terms which are closely related. For example Kristiansen and Gertsen (2014) recent study relating definition of radical innovation identified various synonyms such as exploratory, discontinuous, really new, strategic and breakthrough innovation. Therefore, the incautious unification of synonyms might reduce the weak signals as to the particular terms and their variations that are dominant among innovation researchers.

The separation of multiple terms has a somewhat similar problem as the unification of synonyms, since it is not always obvious which terms one can separate. For example many multiword terms such "business model innovation" and "small and medium sized enterprises" are well established terms and cloud be kept as a single term or could be separated depending on what viewpoints needs to be highlighted. To overcome this problem, we introduce following novel *Variant keyword search* – process for PHASE4, which partially also eliminates the need for PHASES 1 to 3.

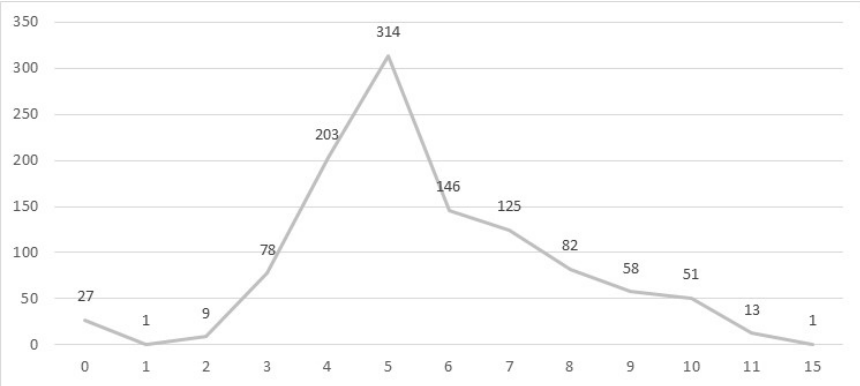
- PHASE 4.1: All keywords were separated into individual words. For example "business model innovation" is treated as three separate words "business", "model" and "innovation". This phase detects which individual words are dominant keywords.
- PHASE 4.2: Using individual words (from PHASE 4.1) as search terms when searching original keyword list (from PHASE 3). This phase detects and links all keyword variants for search terms including possible spelling errors, spelling differences and combined terms such as "new product development" and "product development" are linked to "product" research topic.
- PHASE 4.3: As an iterative process 1) removing irrelevant and illogical keyword variant links such as "pr**open**sity to innovate" link from "open" search term and 2) identifying possible wildcard search terms in order eliminate additional spelling differences such as "technology" and "technological" combined as *techno** wildcard search term.
- PHASE 4.4: Creating research topic ranking list by calculating 1) the sum of all linked research terms, 2) the number of variant terms and 3) identifying the most frequent individual keywords within a research topic.

4 Results

4.1 Descriptive statistics

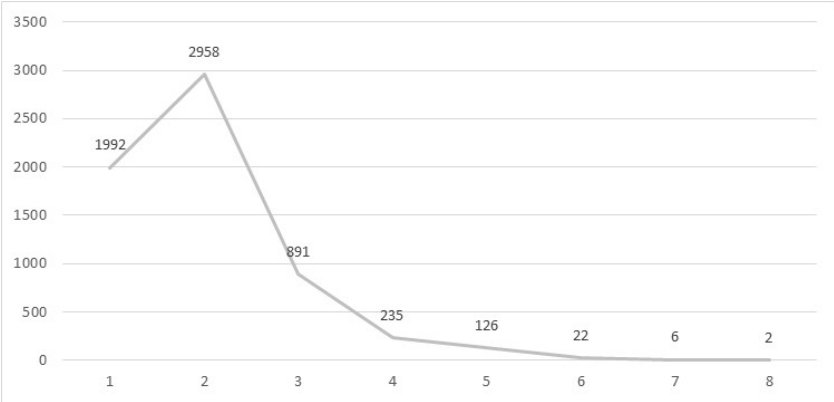
Number of keywords in a paper. The ISPIM author guidelines instruct authors to add about 10 keywords to a paper. In Figure 2 we have presented the distribution of the number of keywords per paper. According to the most commonly used normality test – Kolmogorov-Smirnov (Kolmogorov 1933, Smirnov, 1939) and Shapiro-Wilk tests (Shapiro and Wilk, 1965) – the data clearly violated normal distribution assumption ($p < 0.001$). However, skewness value 0.290 passed the conservative threshold +/- 0.5 threshold criteria and kurtosis value 0.631 liberal interpretation +/- 1.0 threshold criteria for normal distribution (e.g. Cramer, 1997 and SPSS manual). The papers included 6232 keywords (in total) and 3147 different keyword variants after the data harmonization process PHASE3. On average, papers had 5.62 keywords while the standard deviation was 2.13. The average is a bit higher than in some other academic fields (Emrouznejad et al. 2008; Su and Lee, 2009) and high ranking journals (Yi and Choi, 2012), but interestingly about the same as in MIS research (Choi et al. 2011) and when users are providing keywords for interest in online social networks (Bhattacharyya et al. 2009). Most typically papers included four (18.3 percent) or five (28.3 percent) keywords. To conclude, we argue that in terms of number of keywords our dataset is comparable to other scientific domains.

Figure 2 The distribution of the number of keywords per paper



Number of words in keywords. According to Lee et al. (2010, p. 691) "a keyword is the most basic fundamental carrier of knowledge". However, as discussed in the keyword cleaning and harmonization process section individual keywords can be formed from single or multiple words such as "open innovation", which has two words, but forms a single term in the innovation management domain. In Figure 3 we have presented the distribution of the number of words within each keywords. About half (N=2958, 47.5 percent) of the all keywords were formed of two words and about one third (N=1992, 32.0 percent) had only one word. Together these represent nearly 80 percent of all keywords and fit the 80/20 rule well.

Figure 3 The distribution of the number of words in keywords



Keyword frequency and popularity. About half of the keywords (N=3147, 50.5 percent of all keywords) had different writing format, but most of them appeared only once (N=2400, 38.5 percent of all keywords, 76.2 percent of unique keywords), twice (N=718, 11.5 percent / 22.8 percent) or three times (N=152, 7.3 percent / 14.5 percent) as a keyword. On the contrary few keywords occurred as clearly more popular than others: highest frequency N=224 (innovation), second highest N=191 (open innovation) and third highest N=99 (innovation management) already less than half than in the case of the most popular keyword. The evaluation of the annual distribution of keyword occurrence in Table 1 reveals that about 85 percent of annual keywords occur only once. As a result, keyword distribution in terms of "unique keywords" and "keyword popularity" appears to follow a kind of long tail profile in which a large number of occurrences are far from the "head" of the distribution (Anderson, 2006).

Table 1 The annual distribution of the keyword occurrence (KW N = Number of keywords)

KW	2009		2010		2011		2012		2013		2014		MEAN	
N	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1	566	84.10	497	82.97	617	85.69	575	85.69	838	84.82	651	86.11	635.6	84.9
2	67	9.96	60	10.02	67	9.31	53	7.90	80	8.10	60	7.94	64.0	8.9
3	17	2.53	16	2.67	15	2.08	18	2.68	26	2.63	19	2.51	18.8	2.5
4	9	1.34	10	1.67	7	0.97	10	1.49	14	1.42	8	1.06	9.8	1.3
5	4	0.59	4	0.67	4	0.56	5	0.75	8	0.81	4	0.53	5.0	0.7
6	5	0.74	4	0.67	1	0.14	2	0.30	6	0.61	2	0.26	3.0	0.5
>6	5	0.74	8	1.34	9	1.25	8	1.19	16	1.62	12	1.59	10.6	1.3
All	673	15.27	599	13.59	720	16.34	671	15.23	988	22.42	756	17.15	746.8	16.7

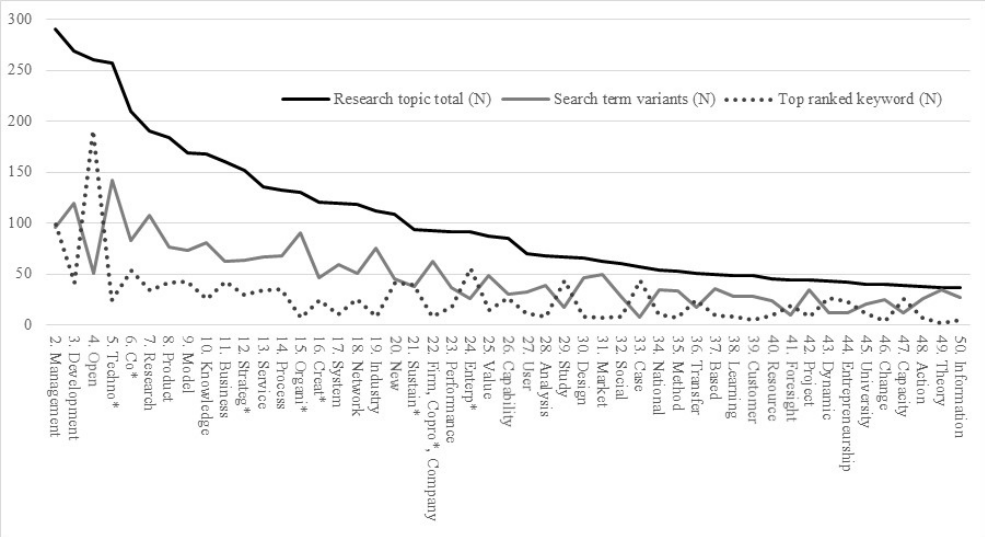
The annual distribution of the TOP20 keywords (table omitted due to space limitations) was also evaluated. The analysis revealed that the most popular keywords mainly increased linearly and the most popular keyword "innovation" had on the average 37.3

hits per year, second most popular keyword “open innovation” 31.8 words and third popular “innovation management” 16.5 words. This indicates that most popular keywords at ISPIM are rather stable. Based on the above descriptive profiles, we argue that there is an obvious need for our *Variant keyword search* – process in order to verify whether research topics are genuinely scattered or these findings are caused by a high variation in keyword writings formats.

4.2 Most popular keywords (TOP50)

To identify the most popular research topics, *Variant keyword search* – process was conducted as described in *3.3 Keyword cleaning and harmonization process* – section. In all our 6232 keywords (in total) and 3147 different keywords included 11836 words in total which were combined from 1938 different words. In Appendix Table 2, we have presented the TOP50 ISPIM research topics. In Figure 4, we have visualised the TOP50 distribution of research topic in total (N), number of search term variants (N) and top ranked keyword (N). INNO* research topic, which frequencies are significantly higher than others (research topic in total (N) = 1370, search term variants (N) = 402 and top ranked keyword (N) = 224, innovation), was excluded to keep the Figure 4 interpretable.

Figure 4. TOP50 research topic distribution

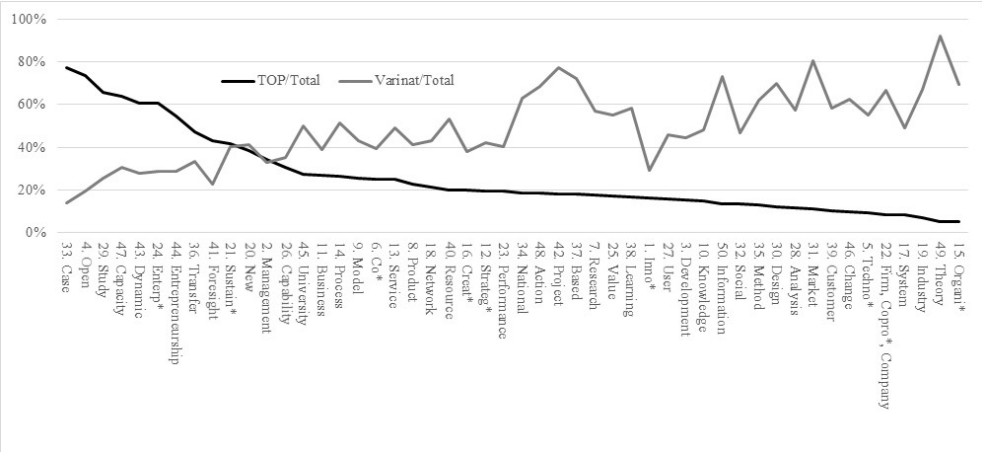


As a result the TOP10 ranking research topics (and their most common keyword) were inno* [N=1370] (innovation, N=224), management [N=290] (innovation management, N=99), development [N=269] (new product development, N=42), open [N=260] (open innovation, N=191), techno* [N=257] (technology transfer, N=24), co* [N=210] (collaboration, N=53), research [N=190] (research and development, N=34), product [N=184] (new product development, N=42), model [N=169] (business model, N=43) and knowledge [N=168] (knowledge management, N=25).

In Figure 5 we have presented the relative share of the most common keywords and search term variants compared to the research topic total. On average the most common

keyword within each research topic represents about a one-quarter share (26 percent, Std.dev = 18.7 percent) and the number of search term variants covered nearly half (48.8 percent, Std.dev =17.1 percent) of the research topic total.

Figure 5. TOP50: Relative share of the most common keyword and search terms variants from the research topic total.



The two extreme examples were 1) *"case" - research topic* in which the most common keyword "case study" covered 77.2 percent of the whole research topic and had only 8 variant search terms which represented 14.0 percent share and 2) *theory - research topic* in which the most common keyword "Actor network theory" covered only 5.4 percent of the whole research topic and had only 34 variant search terms which represented 91.9 present share. Basically these findings indicate that within TOP50 list *research topic N* - variable correlates with the *number of search terms* – variable (0.960**, **. sig. 0.01 level) and *top ranked keyword (N)* – variable, (0.806**). Also *the number of search terms* - variable and *top ranked keyword (N)* – variable are correlating (0.670**).

4.3 Effectiveness of Variant keyword search – process (VKS)

To verify whether the distribution of keywords was caused by a high variation in keyword writing formats or genuinely scattered research topics 1) the total number of keywords and 2) the number unique keywords covered by TOP50 research topics list was evaluated. It appears that the total number of keywords in the TOP50 list was 4239 (vs. starting point 6232 keywords) and the number of unique keywords was 1880 (vs. starting point 3147). About one third of the keywords (N=850) were included in two or more research topics such as *new product development* - keyword was included in "new", "development" and "product" research topics. It appears that *Variant keyword search* – process (VKS) was a fairly powerful tool to compress and reduce useless variation within keywords, since with only 50 research topics we were able to cover 68 percent of all keywords and 60 percent of unique keywords without losing significant information.

5 Conclusions

This study empirically verified what kind of research topics have been covered by ISPIM proceedings during the period 2009 to 2014. Beyond obvious innovation and innovation management related keywords, the TOP50 list was populated by research topics, which keywords can be derived from various innovation classifications such as product, process, market, organizational keywords based on Schumpeter (1934) typology or Chesbrough's open innovation (Chesbrough and Bogers, 2014) as an example. As a result most of the papers follow traditional innovation management research themes and portray a kind of "incremental innovation" based research approach. Focusing on well-known prior topics is not a big surprise since by definition science is grounded in prior knowledge.

However the remaining set of keywords, which represents about a third of keywords, still contain plenty of possible emerging research ideas, which have gained less interest. The evaluation of these themes is beyond the scope of this paper, which focused on the key research topics. In the literature review section we argued that "*ISPIM proceedings could be a store of novel research ideas and a platform to identify and evaluate emerging trends in innovation management research, which may only become visible in formal journal publications several years later.*" This study provides a starting point to verify this claim. As a future study it is proposed that ISPIM research topic profile is compared to top tier innovation management related journals in order to verify, whether the research topic profiles differ or whether certain topics appear in ISPIM before the journals. Moreover, additional analysis on the remaining keyword set could reveal more fledgling research topics, which may later gain popularity. ISPIM organizers are encouraged to use our result as a tool to develop forthcoming events and encourage novel research topics postings, which are beyond traditional innovation themes. Finally, this study can help practitioners to understand the innovation management research paradigm as a whole.

References and Notes

Aleixandre-Benavent, R., Gonzalez-Alcaide, G., Miguel-Dasit, A., Navarro-Molina, C., Valderrama-Zurián, J. C. (2009). Full-text publications in peer-reviewed journals derived from presentations at three ISSI Conferences. Scientometrics, 80(2), 407–418

Anderson, Chris (2006). The Long Tail: Why the Future of Business is Selling Less of More. New York, NY: Hyperion

Baregheh, A., Rowley, J., Sambrook, S. (2009). Towards a multidisciplinary definition of innovation. Management decision, 47(8), 1323-1339.

Bhattacharyya, P., Garg, A., Wu, S. F. (2009, July). Social network model based on keyword categorization. In Social Network Analysis and Mining, 2009. ASONAM'09. International Conference on Advances in (pp. 170-175). IEEE.

Borgatti, S.P., Everett, M.G. and Freeman, L.C. (1992). Ucinet – Guide – Ucinet for Windows: Software and Social Network Analysis. Harvard, MA: Analytic Technologies.

Bourdieu, P. (2004). Science of Science and Reflexivity. Cambridge: Polity.

Butler, L., Visser, M.S. (2006). Extending citation analysis to non-source items. Scientometrics, 66(2), 327–343.

Chesbrough, H., Bogers, M., (2014) Explicating open innovation: clarifying an emerging paradigm for understanding innovation. In: Chesbrough, H., Vanhaverbeke,W., West, J. (Eds.), New Frontiers in Open Innovation. Oxford University Press, Oxford.

Choi, J., Yi, S., Lee, K. C. (2011). Analysis of keyword networks in MIS research and implications for predicting knowledge evolution. Information & Management, 48(8), 371-381.

Cramer, D. (1997). Basic statistics for social research. Step-by-step calculations and computer techniques using Minicab. London: Routledge.

Drott, M. C. (1995). Reexamining the role of conference papers in scholarly communication. JASIS, 46(4), 299-305.

Emrouznejad, A., Parker, B. R., & Tavares, G. (2008). Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. Socio-economic planning sciences, 42(3), 151-157.

Glaenzel, W., Schlemmer, B., Schubert, A., & Thijs, B. (2006). Proceedings literature as additional data source for bibliometric analysis. Scientometrics, 68(3),457–473.

Godin, B. (1998). Measuring knowledge flows between countries: The use of scientific meeting data. Scientometrics, 42(3), 313–323.

González-Albo, B., Bordons, M. (2011). Articles vs. proceedings papers: Do they differ in research relevance and impact? A case study in the Library and Information Science field. Journal of Informetrics, 5(3), 369-381.

Goodrum, A. A., MacCain, K. W., Lawrence, S., Giles, C. L. (2001). Scholarly publishing in the Internet age: A citation analysis of computer science literature. Information Processing and Management, 37(5), 661–675.

Gupta, B. M., Bhattacharya, S. (2003). Bibliometric approach towards mapping the dynamics of science and technology. DESIDOC Journal of Library & Information Technology, 24(1).

ISPIM, (2015), Guidelines for authors, presenter, moderators, & facilitators: Instructions and guidelines – 13 October 2014 – Budapest and future, http://ispim.org/wp-content/uploads/ISPIM_Guidelines.pdf

Kho, M. E., Brouwers, M. C. (2009). Conference abstracts of a new oncology drug do not always lead to full publication: Proceed with caution. Journal of Clinical Epidemiology, 62, 752–758.

Kolmogorov, A.N. (1933). Sulla Determinazione Empirica di una Legge di Distribuzione, Giornale dell' Istituto Italiano Delgi, Attuari 4:83–91.

Kristiansen, J. N., Gertsen, F. (2014), Reviewing Radical Innovation Typology: Remedying Innovation Type Confusion. in K.R.E. Huizingh; S. Conn; M. Torkkeli & I. Bitran (ed.), The Proceedings of The XXV ISPIM Conference. Dublin, Ireland.

Kuhn, T. S. The structure of scientific revolutions. (Chicago, IL: University of Chicago Press, 1962)

Larivière, V., Sugimoto, C. R., Cronin, B. (2012). A bibliometric chronicling of library and information science's first hundred years. Journal of the American Society for Information Science and Technology, 63(5), 997-1016.

Lee, P. C., Su, H. N., Chan, T. Y. (2010). Assessment of ontology-based knowledge network formation by Vector-Space Model. Scientometrics, 85(3), 689-703.

Lisée, C., Larivière, V., Archambault, É. (2008). Conference proceedings as a source of scientific information: A bibliometric analysis. Journal of the American Society for Information Science and Technology, 59(11), 1776-1784.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., Barton, D. (2012). Big Data. The management revolution. Harvard Bus Rev, 90(10), 61-67.

Miguel-Dasit, A., Marti-Bonmati, L., Sanfeliu, P., Aleixandre, R. (2006). Scientific papers presented at the European Congress of Radiology 2000: Publication rates and characteristics during the period 2000–2004. European Radiology, 16(2), 445–450.

Montesi, M., Mackenzie Owen, J. (2008). Research journal articles as document genres: exploring their role in knowledge organization. Journal of Documentation, 64(1), 143-167.

Morlacchi, P., Wilkinson, I. F. and Young, L. C. (2005). Social networks of researchers in B2B Marketing: A case study of the IMP Group 1984–1999. Journal of Business-to-Business Marketing, Vol. 12, pp. 3–34.

Newman, M. E. (2001). The structure of scientific collaboration networks. PNAS, Vol. 98, pp. 404–409.

Pilkington, A., Meredith, J. (2009). The evolution of the intellectual structure of operations management—1980–2006: A citation/co-citation analysis. Journal of Operations Management, 27(3), 185-202.

Pritchard, A. (1969). Statistical bibliography or bibliometrics?. Journal of documentation, (25), 348-349.

Santonen, T., (2015) Unpublished funding application.

Santonen, T., Ritala, P. (2014). Social Network Analysis Of The Ispim Innovation Management Community In 2009–2011. International Journal of Innovation Management, 18(01).

Schumpeter, J., 1934, The theory of economic development, Harvard Economic Studies, Cambridge, Mass (1934)

Shapiro, S. and Wilk, M.B. (1965). An analysis of variance test for normality ~complete samples. Biometrika 52:591–611.

Smirnov, N. V., (1939). Estimate of Deviation Between Empirical Distribution Functions in two Independent Samples (in Russian), Bulletin Moscow University 2 (2), 3-16.

SPSS Inc., PASW® Statistics 18 Brief Guide, SPSS Inc., Chicago, USA.

Su, H. N., & Lee, P. C. (2009, August). Knowledge map of publications in research policy. In Management of Engineering & Technology, 2009. PICMET 2009. Portland International Conference on (pp. 2507-2516). IEEE.

Trott, P. (2008). Innovation management and new product development. Pearson education.

Vidgen, R., Henneberg, S. and Naudé, P. (2007). What sort of community is the European Conference on Information Systems? A social network analysis 1993–2005, European Journal of Information Systems, Vol. 16, pp. 5-19.

Wilcox, R.R. (1990). Comparing the means of two independent groups. Biometrical Journal 32:771–780.

Wu, B., Xiao, H., Dong, X., Wang, M., Xue, L. (2012). Tourism knowledge domains: A keyword analysis. Asia Pacific Journal of Tourism Research, 17(4), 355-380.

Yi, S., & Choi, J. (2012). The organization of scientific knowledge: the structural characteristics of keyword networks. Scientometrics, 90(3), 1015-1026.

Zhu, W., & Guan, J. (2013). A bibliometric study of service innovation research: based on complex network analysis. Scientometrics, 94(3), 1195-1216.

Appendix: Table 2: TOP50 ISPIM research topics

	Research topic total (N)	Search term variants (N)	Top ranked keyword (N)	Top ranked keyword name	Top / Total (%)	Variants / Total (%)
1. Inno*	1370	402	224	Innovation	16.4	29.3
2. Management	290	96	99	Innovation management	34.1	33.1
3. Development	269	120	42	New product development	15.6	44.6
4. Open	260	51	191	Open innovation	73.5	19.6
5. Techno*	257	142	24	Technology transfer	9.3	55.3
6. Co*	210	83	53	Collaboration	25.2	39.5
7. Research	190	108	34	Research and development	17.9	56.8
8. Product	184	76	42	New product development	22.8	41.3
9. Model	169	73	43	Business model	25.4	43.2
10. Knowledge	168	81	25	Knowledge management	14.9	48.2
11. Business	160	62	43	Business model	26.9	38.8
12. Strateg*	152	64	30	Strategy	19.7	42.1
13. Service	136	67	34	Service innovation	25.0	49.3
14. Process	132	68	35	Innovation process	26.5	51.5
15. Organi*	130	90	7	Organizational innovation	5.4	69.2
16. Creat*	121	46	24	Co-creation	19.8	38.0
17. System	120	59	10	Innovation ecosystem	8.3	49.2
18. Network	118	51	25	Network	21.2	43.2
19. Industry	112	75	8	University-industry collaboration	7.1	67.0
20. New	109	45	42	New product development	38.5	41.3
21. Sustain*	94	38	39	Sustainability	41.5	40.4
22. Firm, Copro*, Company	93	62	8	Corporate foresight	8.6	66.7
23. Performance	92	37	18	Innovation performance	19.6	40.2
24. Enterp*	91	26	55	Small and medium sized enterprises	60.4	28.6
25. Value	87	48	15	Value creation	17.2	55.2

Table continues in the next page.

	Research topic total (N)	Search term variant (N)	Top ranked keyword (N)	Top ranked keyword name	TOP/ Total (%)	Variants/ Total (%)
26. Capability	85	30	26	Dynamic capability	30.6	35.3
27. User	70	32	11	User involvement	15.7	45.7
28. Analysis	68	39	8	Patent analysis	11.8	57.4
29. Study	67	17	44	Case study	65.7	25.4
30. Design	66	46	8	Design	12.1	69.7
31. Market	62	50	7	Emerging market	11.3	80.6
32. Social	60	28	8	Social media	13.3	46.7
33. Case	57	8	44	Case study	77.2	14.0
34. National	54	34	10	National innovation system	18.5	63.0
35. Method	53	33	7	method	13.2	62.3
36. Transfer	51	17	24	Technology transfer	47.1	33.3
37. Based	50	36	9	Resource based view	18.0	72.0
38. Learning	48	28	8	learning	16.7	58.3
39. Customer	48	28	5	Customer	10.4	58.3
40. Resource	45	24	9	Resource based view	20.0	53.3
41. Foresight	44	10	19	Foresight	43.2	22.7
42. Project	44	34	8	Project management	18.2	77.3
43. Dynamic	43	12	26	Dynamic capability	60.5	27.9
44. Entrepreneurship	42	12	23	Entrepreneurship	54.8	28.6
45. University	40	20	11	University	27.5	50.0
46. Change	40	25	4	Change management	10.0	62.5
47. Capacity	39	12	25	Absorptive capacity	64.1	30.8
48. Action	38	26	7	Action research	18.4	68.4
49. Theory	37	34	2	Actor network theory	5.4	91.9
50. Information	37	27	5	Information technology	13.5	73.0