Landscaping Performance Research at the ICPE and its Predecessors: A Systematic Literature Review

Alexandru Danciu, Johannes Kroß, Andreas Brunnert, Felix Willnecker, Christian Vögele
fortiss GmbH
Guerickestr. 25, 80805 Munich, Germany
{danciu,kross,brunnert, willnecker,voegele}@fortiss.org

Anand Kapadia, Helmut Krcmar
Chair for Information Systems
Technische Universität München
Boltzmannstr. 3
85748 Garching, Germany
{kapadia,krcmar}@in.tum.de

ABSTRACT

This paper conducts a systematic literature review of papers published in the proceedings of the International Conference on Performance Engineering (ICPE) and its predecessors. It provides an overview of prevailing topics within the community over time. We look at research and contribution facets that have been used to address these topics. Trends are outlined in terms of evaluation methods to validate contributions. The results are complemented with a geographical and organizational dimension. The paper concludes with a look at the top ten contributing countries and organizations for this purpose.

Categories and Subject Descriptors

A.1 [**Introductory and Survey**]; C.4 [**Performance of Systems**]

General Terms

Performance, Theory

Keywords

Systematic Literature Review, Performance Engineering, ICPE, WOSP, SIPEW, Performance Research

1. INTRODUCTION

Many researchers and industry practitioners around the globe have dedicated themselves to performance engineering, due to the complexity of this subject [5]. As a consequence, various workshops and conferences specialized on this field have been established. The principle conference that is focused on the performance of software systems and related questions is the International Conference on Performance Engineering (ICPE). The ICPE was established as a joint meeting of the ACM Workshop on Software and Performance (WOSP) and the SPEC International Performance Evaluation Workshop (SIPEW). In the scope of this conference, domain experts are invited to present and discuss state-of-the-art research results concerning performance measurement, modeling techniques, benchmark design and run-time performance management [1, 2, 3, 4, 11].

Although research in the field of performance engineering is not in its infancy anymore and numerous papers have already been published, a general overview of prevailing topics and methods within the community does not exist. To the best of our knowledge, there has not been any effort to systematically select, synthesize and review existing literature within the ICPE and its predecessors. Therefore, this gap is addressed in this work.

Performance engineering research at the ICPE and its predecessors is analyzed in a systematic literature review. The first WOSP took place in 1998 followed by six WOSP, one SIPEW and five ICPE events at the time of writing this paper. This work analyzes the proceedings of all these events and captures sixteen years of performance engineering research in total.

2. METHODOLOGY

The systematic literature review in this work is conducted following the guidelines provided by Kitchenham and Charters [8]. According to them, a systematic literature review is a "[...] means of identifying, evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest".

2.1 Research Questions

The initial task in a systematic literature review according to Kitchenham and Charters [8] is the definition of research questions (RQ). RQs in general are central drivers of this research methodology and consequently influence the research process heavily. As part of our study, the following three RQs will be answered:

- **RQ 1: Which topics have been addressed in the papers published at the ICPE (respectively at its predecessors) in the time period from 1998 to 2014?**

The goal of this RQ is to get an overview of different subjects that have been published and discussed at the ICPE, WOSP and SIPEW. Therefore, we investigate which specific topics are addressed more frequently by published papers

ICPE'15, Jan. 31–Feb. 4, 2015, Austin, Texas, USA.
Copyright © ACM 978-1-4503-3248-4/15/01 ...$15.00.
http://dx.doi.org/10.1145/2668930.2688039.

and how this focus has shifted over the years. The answer for this RQ can be found in Section 3.1.

- **RQ 2: Which research facets, contribution facets and evaluation methods have been used in papers published at the ICPE and its predecessors?**

To answer this research question an overview of research and contribution facets is given in Section 3.2. This helps to get an overview of how researchers try to tackle topics outlined in Section 3.1. Furthermore, evaluation methods are outlined in Section 3.3 that have been used to validate different contribution types. Over the years, different types of evaluation methods have been established in the performance engineering domain. In this paper, methods are analyzed in terms of frequency, applicability to a certain topic and popularity within the performance engineering community.

- **RQ 3: Who are the top ten countries and organizations in terms of the quantity of articles published at the ICPE and its predecessors?**

RQ 3 aims to identify how papers published in the proceedings of the ICPE are distributed among countries and organizations. The resulting analysis includes research and publication activity, from these countries and organizations, at the ICPE and its predecessors from 1998 to 2014. First, the geographical perspective is outlined in Section 3.4. Second, the organizational perspective is adressed in Section 3.5.

2.2 Data Sources and Paper Selection

Only papers that have been published at the ICPE, WOSP and SIPEW were considered in a first step. The initial set of papers contained 471 publications in total. It was predominantly available online in the ACM Digital Library[1] and SpringerLink[2]. Papers with illegible writing due to formatting issues such as overlapping characters were replaced. The replacements were taken from other digital libraries and were then checked on correspondence in order to avoid distorting the outcomes.

After having set a solid base for the research by establishing an initial set of papers, the set needed to be filtered to acquire meaningful results. The filter process was divided into two steps which had been specified with different goals and exclusion criteria. The first exclusion criterion (EC1) to be applied was the removal of all invited talks, keynote addresses and editorial articles, as they did not provide any further benefit for this research. By doing so the initial set could be reduced to 448 papers.

For this research, great focus was set on prevailing topics and evaluation methods. Therefore, the published papers were analyzed in terms of giving information on the formulated RQs which could mainly be found in research, industrial and work-in-progress/vision tracks. As a consequence, all demonstration papers, posters and tutorials were removed from the set (EC2). Papers published in workshops that were held alongside with the ICPE are also not considered. This reduced the set to 388 papers.

After the initial selection of papers, the next step in the systematic literature review is to assess the quality of the primary studies [8]. However, as our goal is to provide a complete overview of the conference, we did not exclude any papers based on quality assessment scores. That is why all 388 papers have ultimately been moved to the data extraction process.

[1]http://dl.acm.org/

[2]http://link.springer.com/

2.3 Data Extraction and Synthesis

Data was extracted from a total of 388 included papers. Due to space limitations, we provide an online accessible list of the papers on our website[3] instead of in this paper. We created a data extraction scheme in an Excel spreadsheet with respect to the previously stated RQs. This proved to be a necessary step as we could easily compute frequencies, filter for relevant information and analyze relationships between the different RQ findings. The scheme is divided into different sections. It contains generic (paper ID, title, authors, type of report, year, conference) as well as specific information (organization, evaluation method, contribution facet, research type facet, lifecycle phase, domain, system under study).

3. RESULTS

In this section results of the systematic literature review are presented. The research questions outlined in Section 2 are answered in a chronological order.

3.1 Topics at the ICPE

Obtaining a deep understanding of topics discussed at a conference and the evolution of these topics over several years is a difficult task. New technologies and trends have a constant influence on topics addressed by researchers. Thus, the focus of the conference is shifting from year to year. The N-Gram analysis is employed in this section to provide a solution to this problem and reveal trends within the conference from 1998 to 2014 [12, 6].

An N-Gram analysis is a technique used in the field of natural language processing for identifying the frequency of the occurrence of words or combinations of words [9]. An N-Gram represents a sequence of n words which is extracted from a body of text. For example, the phrase "software performance management" can be divided into three 1-Grams ("software", "performance", "management"), two 2-Grams ("software performance", "performance management"), and one 3-Gram ("software performance management").

In order to perform the analysis we follow the approach of Soper and Turel [12] and first establish a corpus of text. The corpus consists of the collection of 388 selected articles. All articles were available as PDF documents. We converted each document to a parsable text file. In order to prevent distortion of results we removed in several post-processing steps any unnecessary data such as author information, keyword lists, the bibliography, the appendix, page numbers and citation references. The resulting text files were then grouped by the year of the publication to enable an analysis run for each conference edition.

The N-Gram analysis is supported by a variety of tools. We used the freeware tool AntConc[4] because it is easy to use and well documented. For each analysis run, the user can specify the minimum and a maximum length of N-Grams to be considered. When a sequence of words occurs more frequent than a single word, the sequence receives a higher rank within the analysis results. The results provided by the tool consist of the absolute frequency of each N-Gram. Since

[3]http://pmw.fortiss.org/research/icpe/

[4]http://www.antlab.sci.waseda.ac.jp/software.html

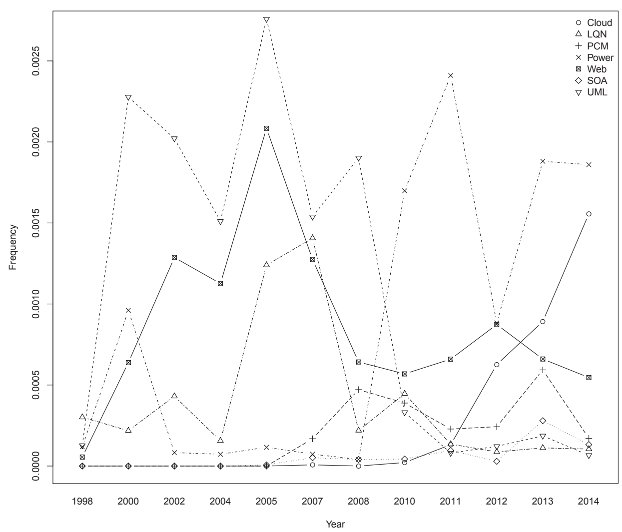

Figure 1: Frequency of terms

the relevance of each N-Gram depends on the size of the text corpus, we calculate relative frequencies and, thus, make the results for each conference edition comparable. Multiple occurrences of the same N-Gram within one article are counted separately. Since word classes such as articles are among the most frequent ones, a filtering needs to be performed to select only content-relevant keywords.

We first performed an analysis for each conference edition to include N-Grams with a minimum length of n=1 to include the most frequent keywords and maximum length of n=4 to limit the expense for the calculation. The most frequent N-Grams identified during the analysis have a length of n=1. However, some of the highest ranked results are of limited value for describing topics addressed by the research. Therefore, words such as *performance*, *system*, *software*, *server*, *model* or *data* were not considered. The frequency of the occurrence of these keywords remains constant over all conference editions[5]. Some of the most relevant topics and their evolution over time are shown in Figure 1.

The terms *Power* and *Cloud* are among the most frequent N-Grams in 2014. The frequency of the term *Power* first peaked in 2000 and then increased significantly after 2008. The term *Cloud* was first used in 2007 in a context different than cloud computing. Only in 2010 the term was first used in this context and its frequency continued to increase every year. During the transition from WOSP/SIPEW to ICPE between 2008 and 2010 only the terms *Power* and *UML* display a significant change.

To gain more insight on the conference topics we performed a second N-Gram analysis to include N-Grams with a minimum length of n=2 and maximum length of n=4. The top ten most frequent word combinations identified for each year at the WOSP/SIPEW and the ICPE are are shown in Tables 1 and 2. The values displayed in the tables represent relative frequencies. Among the most frequent N-Grams in 2014 are *Energy Consumption* and *Power Consumption* having an absolute frequency of 155 and 87 respectively. While used during every edition since 2002, the term *Garbage Collection* is included in 2014 for the first time in the top ten list having 58 occurrences.

[5]In 2008, the WOSP and SIPEW were held separately. Due to different profiles of the two workshops, subsuming the papers of both events to one text corpus would not provide consistent results. To avoid a disruption in the course of trends over time the two conferences cannot be both considered separately for the year 2008. Therefore, the SPIEW 2008 was not considered here.

3.2 Research and Contribution Facets

This section outlines different kinds of research and contribution facets of papers published in the WOSP/SIPEW and ICPE proceedings. It broadens the understanding of how researchers tried to address topics outlined in Section 3.1. Petersen et al. [10] propose a systematic map to classify and structure studies and their fields in the area of software engineering. For their map and its visualization they categorize studies in the following three different facets:

- Variability context facet - categorization for different topics among studies
- Research facet - classification for the type of research such as evaluations, proposals or experience papers
- Contribution facet - attribution of papers' outcomes such as tools or models

Since we have already analyzed major topics in Section 3.1, the focus here lies on the research and contribution facets. For the research facet, Petersen et al. [10] differentiate between the following research types:

- Validation research - assessment of new techniques with example experiments
- Solution proposal - suggestion of a solution for an existing issue
- Philosophical paper - taxonomy or framework for existing subjects
- Experience paper - personal experience and guide for techniques in practice
- Evaluation research - assessment of already implemented techniques
- Opinion paper - personal opinion about methods and techniques

These categories are used to classify all papers that are included according to our selection in Section 2.2. The results are illustrated in Figure 2. If multiple research types were covered by a given paper, only the focused aspect was considered for the classification. Similarly to Petersen et al. [10], they are presented in a bubble chart showing the number of papers for each category with a corresponding bubble size. The research facets are aggregated per year to give an indication about the progress. Evaluation and validation research are the most common types with a total amount of 109 and 113 each and constantly appear over the years. Solution proposals are also very frequent since the second WOSP in 2000.

Petersen et al. [10] also considered the contribution facet of papers. Such contributions facets are methods, metrics, models, processes or tools [10]. All papers are classified according to their contribution facet. Since multiple papers contain two contribution facets and, thus, are counted twice, the total number of contribution types does not represent the total amount of papers. The results are presented in Figure 3. As before, a bubble chart is used to illustrate the amount of occurrences of each contribution in relation to the year.

Table 1: Most frequent keyword combinations at the WOSP/SIPEW (values need to be multiplied by 10^{-3})

WOSP 1998		**WOSP 2000**		**WOSP 2002**		**WOSP 2004**	
Queueing Network	0.619	Performance Model	0.792	Response Time	1.459	Response Time	0.644
Software Architecture	0.619	Software System	0.647	Performance Model	0.674	Web Service	0.452
Response Time	0.603	Execution Time	0.485	Autonomous Service	0.514	Performance Analysis	0.431
Task Graph	0.595	Software Architecture	0.477	Performance Analysis	0.409	Performance Model	0.369
Service Time	0.532	Performance Engineering	0.429	Real Time	0.393	Software Performance	0.333
Performance Requirements	0.453	Performance Analysis	0.38	Sequence Diagram	0.382	Operational Profile	0.322
Server Subsystem	0.373	Object Oriented	0.356	Use Case	0.382	Component based	0.312
Mean Service	0.357	Optimal Shutdown	0.356	Web Server	0.371	Class Diagram	0.307
Component Model	0.342	Response Time	0.356	Data Structure	0.321	Content Location	0.27
Use Case	0.334	Service Time	0.348	Data Flow	0.293	Software Component	0.265
WOSP 2005		**WOSP 2007**		**WOSP 2008**		**SIPEW 2008**	
Performance Model	1.308	Response Time	1.143	Performance Model	1	Response Time	1.444
Response Time	0.709	Web Service(s)	0.835	Performance Analysis	0.643	User Behavior	0.565
Software Performance	0.675	Performance Model	0.557	Execution Time	0.634	SPEC CPU	0.532
UML Model	0.579	Queueing Network	0.543	Case Study	0.569	Timing Behavior	0.532
Software System	0.552	LQN Model	0.543	Use Case	0.561	Calling Context	0.5
Redundant Computation	0.539	Performance Engineering	0.44	Software Performance	0.513	Resource Demands	0.424
Software Archtecture	0.396	Service Time	0.418	Performance Modeling	0.48	Context Analysis	0.413
Web Service	0.396	Business Process	0.403	Response Time	0.472	Composite Service	0.402
Acitivity Diagram	0.327	Performance Analysis	0.359	Meta Model	0.391	Trace Context	0.402
Covering Arrays	0.327	Execution Time	0.352	Model Transformation	0.358	Behavior Model	0.391

Table 2: Most frequent keyword combinations at the ICPE (values need to be multiplied by 10^{-3})

ICPE 2010		**ICPE 2011**		**ICPE 2012**		**ICPE 2013**		**ICPE 2014**	
Response Time	0.842	Response Time	0.66	Response Time	0.694	Response Time	0.963	Energy Consumption	0.856
Non Determinism	0.67	Power Consumption	0.512	File System	0.544	Energy Consumption	0.739	Performance Model	0.685
Page Coloring	0.54	Performance Model	0.403	Software Performance	0.326	Live Migration	0.605	Load Test	0.619
Calling Context	0.439	Power Savings	0.367	Control Charts	0.306	Performance Model	0.534	Power Consumption	0.481
Execution Time	0.425	Product Form	0.358	Web Server	0.292	Performance Regression	0.437	Response Time	0.481
Power Consumption	0.353	Execution Time	0.351	Software System	0.263	Power Consumption	0.355	Execution Time	0.436
Workload Intensity	0.317	Data Center	0.348	Stack Distance	0.263	Time Series	0.348	System Performance	0.381
Data Item	0.288	Delay Tolerant	0.261	Access Control	0.258	Web Server	0.262	Performance Degradation	0.376
Performance Model	0.26	Delay Sensitive	0.232	Monitoring Mechanism	0.238	System Architect	0.239	Performance Metrics	0.359
Performance Signature	0.238	System Performance	0.229	Data Access	0.219	System Performance	0.239	Garbage Collection	0.321

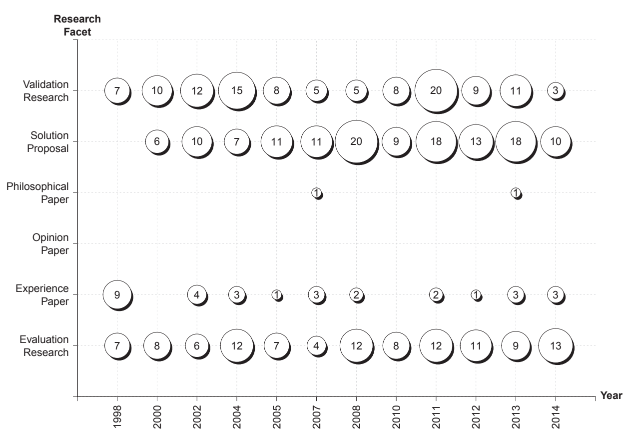

Figure 2: Research facets

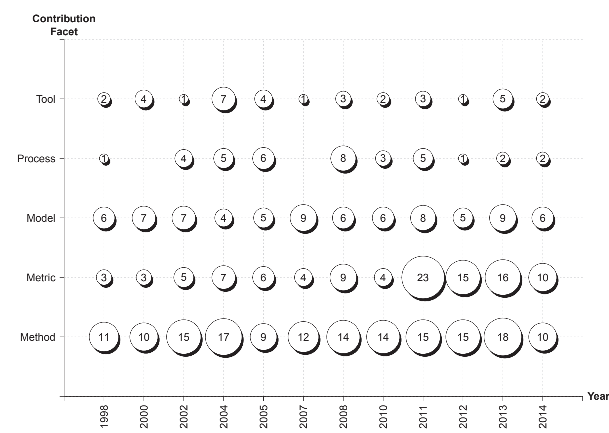

Figure 3: Contribution facets

The results show that each contribution facet is appearing at almost every edition. However, tools as well as processes are the minority group and occur several times not at all or only once per year. Methods represent the majority group and are contributed in 160 papers.

3.3 Evaluation Methods

In this section we investigate evaluation methods used for research and contribution facets outlined in the previous section. The methods are categorized according to the design science theory of Hevner et al. [7]. This theory categorizes evaluation methods based on the validation of IT artifacts. Each of these categories are described in Table 3. According to Hevner et al. [7], these “IT artifacts can be evaluated in terms of functionality, completeness, consistency, accuracy, performance, reliability, usability, fit with the organization, and other relevant quality attributes“. All papers are categorized according to these methods and the results are presented in Figure 4. The categorization is based on the evaluation type mentioned by authors of a paper. If multiple evaluation methods were applied in a given paper, only the focused method was considered for the classification.

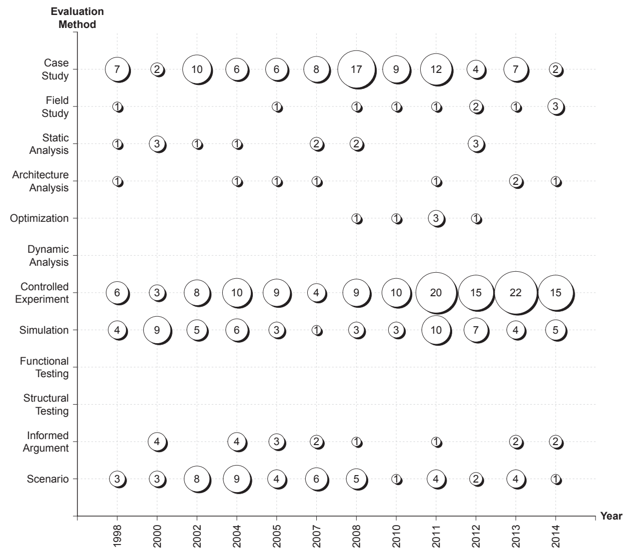

Figure 4: Evaluation methods

Table 3: Design evaluation methods [7]

Category	Method	Paper
1. Observational	**Case Study:** Study artifact in depth in business environment	90
	Field Study: Monitor use of artifact in multiple projects	11
2. Analytical	**Static Analysis:** Examine structure of artifact for static qualities (e.g., complexity)	13
	Architecture Analysis: Study fit of artifact into technical information system architecture	8
	Optimization: Demonstrate inherent optimal properties of artifact or provide optimality bounds on artifact behavior	6
	Dynamic Analysis: Study artifact in use for dynamic qualities (e.g., performance)	0
3. Experimental	**Controlled Experiment:** Study artifact in controlled environment for qualities (e.g., usability)	131
	Simulation: Execute artifact with artificial data	60
4. Testing	**Functional (Black Box) Testing:** Execute artifact interfaces to discover failures and identify defects	0
	Structural (White Box) Testing: Perform coverage testing of some metric (e.g., execution paths) in the artifact implementation	0
5. Descriptive	**Informed Argument:** Use information from the knowledge base (e.g., relevant research) to build a convincing argument for the artifact's utility	19
	Scenarios: Construct detailed scenarios around the artifact to demonstrate its utility	50

The ICPE submissions can be categorized into nine out of twelve distinct evaluation methods: case studies, field studies, static analysis, architecture analysis, optimizations, controlled experiments, simulations, informed arguments and scenarios. No paper was submitted using one of the three remaining categories: functional testing, structural testing and dynamic analysis. It is important to note that Hevner et al. [7] refer to the performance of the artifact itself when he talks about dynamic analysis (e.g., performance). Thus, he does not refer to the performance of a system that is analyzed using the artifact.

The controlled experiment is the most popular evaluation method. 131 of 388 (34%) categorized papers used this evaluation method. The case study is equally popular with a maximum of 17 publications in 2008 and 90 publications (23%) in total. The interest in contributing research with this evaluation method decreases slowly while controlled experiments became more popular recently. A similar development can be observed for scenario based evaluations; both seem to get unpopular while controlled experiments are rising. Today, the most common methods (controlled experiments and case studies) count for over 55% of all papers.

Simulations were popular in the beginning and got a small renaissance in 2011. In the years from 2004 to 2010, this evaluation method was not very popular with only one to three papers per year. In the last years, the interest in simulation based evaluations is again slowly decreasing.

Informed arguments and static analysis are very exotic for ICPE papers. Field studies are rare, even if this method is used more frequently in 2014. We only found eight papers using architecture analysis as evaluation method and six papers in total using optimizations. We see that evaluation methods are more popular that rely on only one artifact, like case studies and controlled experiments. This evaluation popularity is related to the contribution facet of papers submitted to the ICPE. As seen in Section 3.2, model and method contributions are very popular. Such contributions require a case study or experiment as an evaluation. Process or tool contributions are easier evaluated in a field study for example but these contribution types are rare for the ICPE and its predecessors.

3.4 Geographical Perspective

A large number of different countries have contributed to publications over the years. In summary, 33 countries have been involved. Table 4 shows the top ten countries ranked by their total amount of publications. The metric publications includes exclusive as well as joint publications. If, for instance, one paper was published by three authors from USA and one from Germany, the number of publications will be increased by one for both countries since authors from both countries contributed to the publication. Therefore, the total amount of publications in Table 4 is not equal to the total amount of publications of all editions. Furthermore, the share of each country to the total amount of papers is listed.

Table 4: Top 10 contributing countries

Rank	Country	Publications	Share	Cooperation
1	USA	130	33.51%	40
2	Germany	67	17.27%	23
3	Canada	61	15.72%	12
4	Italy	52	13.40%	23
5	UK	41	10.57%	14
6	Spain	20	5.15%	13
7	Australia	9	2.32%	2
	Netherlands	9	2.32%	5
	India	9	2.32%	2
	Switzerland	9	2.32%	5

The first rank is represented by USA with 130 publications followed with a large distance by Germany and Canada with 67 and 61 publications respectively. An analysis of the number of papers published by countries hosting the ICPE indicates that hosting countries publish more papers that on average. Except for three events, the host countries have published twice as many papers than usual.

As publications include joint publications between countries, the number of papers in cooperation is listed in Table 4 as well. A remarkable value is presented by Canada, which counts 12 joint publications and, thus, presents the lowest proportion of papers in cooperation in relation to their publications with 20%. In contrast, Spain contributed 13 papers in cooperation and, therefore, has the biggest proportion with 65%.

3.5 Organizational Perspective

The evaluation of the research activity from an organizational perspective is performed in a similar way as in Section 3.4. Table 5 lists the top 10 contributing organizations ranked by the amount of publications. Its listed metrics publications, share and cooperation are defined in the same way as in Table 4, only applied to organizations instead of countries.

Carleton University constitutes the first rank with 38 publications. Rank two and three are placed by Karlsruhe Institute of Technology (KIT) and University of L'Aquila with 24 and 20 publications followed by Imperial College London and University of Rome Tor Vergata with 16 and 12 publications on rank four and five. Beginning from rank six, the remaining organizations count less than ten publications. Although the USA is ranked first by the number of publications, none of the top six organizations belongs to this country.

Table 5: Top 10 contributing organizations

Rank	Organization	Country	Publications	Share	Cooperation
1	Carleton University	Canada	38	9.79%	12
2	Karlsruhe Institute of Technology	Germany	24	6.19%	18
3	University of L'Aquila	Italy	20	5.15%	14
4	Imperial College London	UK	16	4.12%	3
5	University of Rome Tor Vergata	Italy	12	3.09%	5
6	University of Zaragoza	Spain	9	2.32%	4
7	AT&T Labs	USA	8	2.06%	3
	Hewlett-Packard Laboratories USA	USA	8	2.06%	5
	University of the Balearic Islands	Spain	8	2.06%	8
10	George Mason University	USA	7	1.80%	3
	Oracle Corporation USA	USA	7	1.80%	6
	Performance Engineering Services	USA	7	1.80%	7
	SAP Research Karlsruhe	Germany	7	1.80%	5
	University of Oldenburg	Germany	7	1.80%	6

4. CONCLUSION

In this work, we have conducted a systematic review of literature on performance research at the ICPE, WOSP and SIPEW. Most articles published at the WOSP and SIPEW are focused on the system development phase. Since the inception of the ICPE in 2010, an increasing number of papers address the system operation phase which results in a well-balanced conference profile. Our N-Gram analysis revealed a constant shift of the conference focus towards the latest technologies such as cloud computing. Due to this shift, the conference is increasingly addressing the most relevant performance goals for these technologies, e.g., optimizing resource consumption.

The proportion of research and contribution types published has remained constant over the years and only slight shifts can be observed. While metrics are being contributed more frequently since 2011, philosophical papers continue to be underrepresented. The ICPE community would, however, greatly benefit from more research which provides taxonomies for the generated knowledge and summarizes existing findings within the performance engineering field.

Trends found in Section 3.3 show that evaluations based on one IT artifact are very popular nowadays, while evaluations of multiple artifacts or architectures are rare. As contribution type methods and models are very popular, it would be of great benefit to use the results of such research to create tools and processes as well as to evaluate results in broader environments.

The data and conducted analysis in Section 3.4 indicate a positive influence on the number of publications of a country if it is the host of a conference. An explanation for this fact could be an increased amount of submissions due to lower travel costs. In our opinion, the conference organizers should consider this in order to increase the involvement of certain countries.

5. REFERENCES

[1] *WOSP/SIPEW '10: Proceedings of the First Joint WOSP/SIPEW International Conference on Performance Engineering*, New York, NY, USA, 2010. ACM. 488103.

[2] *ICPE '11: Proceedings of the 2nd ACM/SPEC International Conference on Performance Engineering*, New York, NY, USA, 2011. ACM. 594112.

[3] *ICPE '12: Proceedings of the 3rd ACM/SPEC International Conference on Performance Engineering*, New York, NY, USA, 2012. ACM. 594122.

[4] *ICPE '14: Proceedings of the 5th ACM/SPEC International Conference on Performance Engineering*, New York, NY, USA, 2014. ACM. 594142.

[5] A. Brunnert, C. Vögele, A. Danciu, M. Pfaff, M. Mayer, and H. Krcmar. Performance management work. *Business & Information Systems Engineering*, 6(3):177–179, 2014.

[6] S. Demeyer, A. Murgia, K. Wyckmans, and A. Lamkanfi. Happy birthday! a trend analysis on past msr papers. In *Proceedings of the 10th Working Conference on Mining Software Repositories*, MSR '13, pages 353–362, Piscataway, NJ, USA, 2013. IEEE Press.

[7] A. R. Hevner, S. Ram, S. T. March, and J. Park. Design science in information systems research. *MIS Quarterly*, 28(1):75–105, 2004.

[8] B. Kitchenham and S. Charters. Guidelines for performing systematic literature reviews in software engineering. Technical report, Technical report, EBSE Technical Report EBSE-2007-01, 2007.

[9] C. D. Manning and H. Schütze. *Foundations of Statistical Natural Language Processing*. Cambridgeand MA : The MIT Press, 1999.

[10] K. Petersen, R. Feldt, S. Mujtaba, and M. Mattsson. Systematic mapping studies in software engineering. In *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering*, EASE'08, pages 68–77, Swinton, UK, 2008. British Computer Society.

[11] S. Seelam, editor. *ICPE '13: Proceedings of the 4th ACM/SPEC International Conference on Performance Engineering*, New York, NY, USA, 2013. ACM. 594132.

[12] D. S. Soper and O. Turel. An n-gram analysis of communications 2000–2010. *Commun. ACM*, 55(5):81–87, May 2012.