

Uncovering design topics by visualizing and interpreting keyword data

LEI, Ningrong <<http://orcid.org/0000-0003-0935-9426>>, FAUST, Oliver <<http://orcid.org/0000-0002-0352-6716>>, ROSEN, David William and SHERKAT, Nasser

Available from Sheffield Hallam University Research Archive (SHURA) at:

<https://shura.shu.ac.uk/20860/>

This document is the Accepted Version [AM]

Citation:

LEI, Ningrong, FAUST, Oliver, ROSEN, David William and SHERKAT, Nasser (2018). Uncovering design topics by visualizing and interpreting keyword data. In: MARJANOVIC, D, STORGA, M, SKEC, S, BOJCETIC, N and PAVKOVIC, N, (eds.) DS92: Proceedings of the DESIGN 2018 15th International Design Conference. Design Society, 57-68. [Book Section]

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>



UNCOVERING DESIGN TOPICS BY VISUALIZING AND INTERPRETING KEYWORD DATA

Abstract

This paper presents a bibliometric keyword analysis for the international DESIGN conference. We combined related keywords to form DESIGN topics. After that, we visualized the connections between the topics. Our analysis shows that the web of science database does not contain the DESIGN 2012-14 proceedings. That is relevant for the conference organizers, because content visibility is important. The topic visualization benefits both contributors to and organizers of the international DESIGN conference, because it shows trending topics and it indicates areas with room for improvement.

1. Introduction

The international DESIGN conference is a forum for researchers and practitioners to exchange research ideas across disciplines¹. Since the first issue in 1981, the conference proceedings have grown to a substantial body of literature. The data holds key information about trends and issues in the design world. These trends and issues will shape future contributions and determine coming topics of interest (Small et al., 2014). The main problem is to find appropriate methods to extract relevant information.

Topic Detection and Tracking (TDT) is a promising approach for extracting relevant information from a large body of literature (Allan, 2002). In a practical setting, TDT can be realized by extracting and analysing keywords (Yu et al., 2007). Author supplied keywords are especially interesting, because they reflect best the paper topic (Yoon and Park, 2005). However, author supplied keywords tend to be inconsistent, and there are spelling differences between British and American English. That makes algorithmic keyword analysis error prone and the extracted parameters might not reflect the desired information (Beil et al., 2002). Therefore, extensive keyword cleaning is necessary before the analysis process can start. Once the keywords are cleaned, standard clustering and visualization methods can be used to extract relevant information.

We have conducted a TDT analysis for the international DESIGN conference, based on data available from the web of science database. Our first contribution is that we established the topics from author supplied keywords through a sophisticated refinement process. The second contribution is that we visualized connections between topics with a network diagram. Our analysis revealed data inconsistencies, as well as DESIGN topic-statistics and network structures. To be specific, we found data inconsistencies in the web of science, the database did not include the bibliometric data from the papers published in both 2012 and 2014 conferences. In terms of statistics, the three most frequent topics covered in the conference proceedings are: product, design, and process. The network diagram shows that eco-design and sustainability are peripheral topics in the DESIGN conference. Detecting data inconsistencies helps the conference organisers ensure and indeed establish content visibility in relevant databases. The topic network is important for both management and contributors alike, because it highlights the connections between topics. Contributors can provide a new connection

¹ International DESIGN conference web page (Last accessed 30.10.2017): <http://www.designconference.org/>

between topics or even establish a new topic. The organizing committee might want to steer some topics more to the centre of the map by accepting contributions with these topics.

To demonstrate the fitness of our analysis system, we have organized the remainder of the paper as follows. The next section introduces the materials used for the author supplied keyword analysis. Section 3 presents the analysis results. These findings are discussed in Section 4. We used the topic cluster visualization to synthesise some recommendations for the future direction of the DESIGN conference. The paper concludes with Section 5.

2. Materials

This section describes the materials used for the TDT analysis. The analysis is based on keyword clustering, manual keyword to topic mapping, and algorithms for topic visualization. The flowchart, shown in Figure 1, details the workflow which orders the individual processing steps of the analysis system. The following sections detail the individual processing steps.

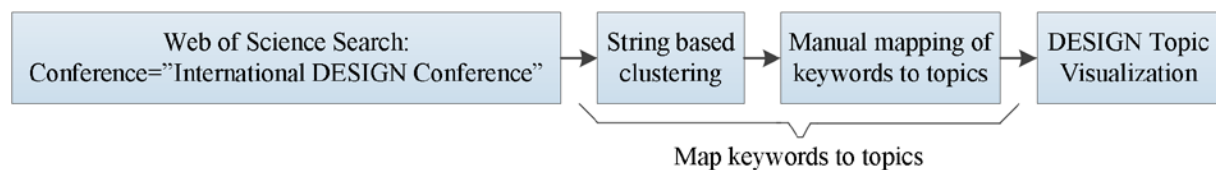
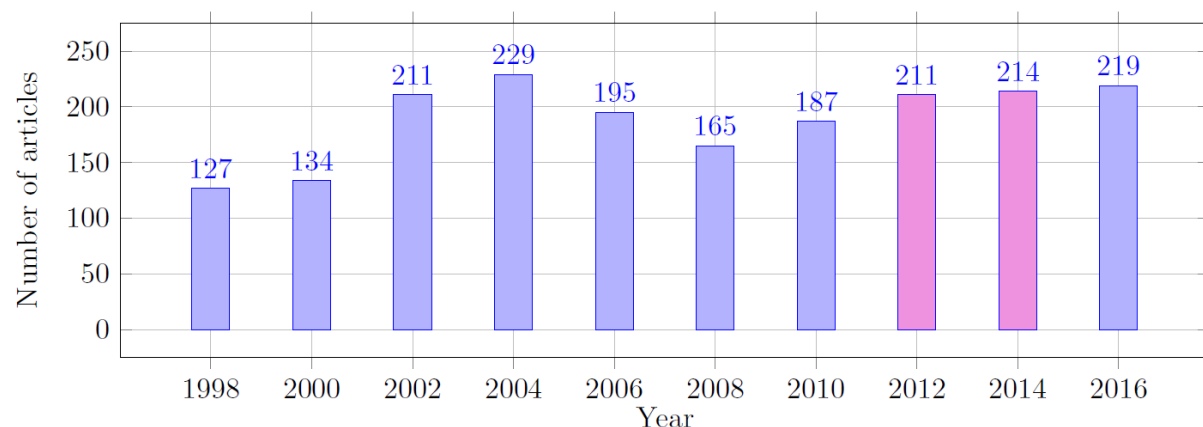


Figure 1. Flowchart of the keyword analysis

2.1. Data

The data was captured with a straight forward search in the web of science database². The search for 'International DESIGN Conference' resulted in a set with 1467 database entries. Closer inspection of this set revealed that all entries from the conferences in 2012 and 2014 were missing. To correct the problem, we extracted the missing bibliographic data from the scopus³ database. Figure 1 provides an overview of the database entries, which represent the papers published in the DESIGN conference from 1998 to 2016. We found 1892 database entries, both overall number and yearly distribution matches the numbers published on the DESIGN conference web page⁴. Each entry contains, amongst other bibliometric information, the keywords of a conference contribution. All keywords were extracted, and duplicates were removed. The result of these actions was a set of 4479 distinct keywords. We have selected the 500 distinct keywords which appeared most often in the 1892 papers, i.e. these papers have the highest weight. These keywords constitute the raw material for the keyword refinement process.



² Web page (Last accessed 30.10.2017): <http://ipscience.thomsonreuters.com/product/web-of-science/>

³ Web page (Last accessed 30.10.2017): <https://www.elsevier.com/en-gb/solutions/scopus>

⁴ Web page (Last accessed 30.10.2017): <http://www.designconference.org/tradition>

Figure 2. Yearly distribution of database entries that represent the number of articles published in the international DESIGN conference from 1998 until 2016. The database entries from web of science are in blue and the database entries from scopus are in pink.

2.2. Map keywords to topics

Author supplied keywords are free text, i.e. any character string can be used as a keyword. There is no standardisation. Therefore, the author supplied keywords are quite noisy, i.e. there are many minor spelling differences between individual keywords. To address this problem, we aim to find groups of similar keywords and map them to one topic (Suominen and Toivanen, 2016). As such, a topic is defined as an abstract semantic concept in a collection of keywords. To facilitate the aim, we established a two-step process, as shown in Figure 1. In the first process step, we order the keywords based on their character string similarity. This step provides a coarse keyword ordering which gives an overview of potential topics. In the second process step, we manually mapped the pre-ordered keywords to DESIGN topics.

The first process step starts by establishing the Levenshtein distance between the keyword strings (Levenstein, 1965; Levenshtein, 1966). Based on that information, we generated an agglomerative hierarchical cluster tree (Anderberg, 2014). The cluster tree was cut into individual clusters. To be specific, we were cutting the cluster tree in 500 different ways yielding 1 to 500 clusters. That resulted in a list of 500 differently pruned cluster trees. For each cluster, within that list, we calculated the within-clusters sum of squares (Kodinariya and Makwana, 2013). For each list entry, i.e. for each pruned cluster tree, the individual calculation results were summed up to form the total within-clusters sum of squares. The blue graph, in Figure 3, shows the total within-clusters sum of squares for each list entry, ordered by the number of clusters in the pruned cluster tree. The mapping results in a strictly decreasing knee graph. To find the right number of clusters, we evaluated the knee point of that graph. As such, the knee point is furthest away from the line through the points (0,2.1446e+04) and (500,0). The dashed orange graph indicates the distance to that line for each point of the knee graph (blue solid line). The maximum appears for 126 keyword clusters. Hence, the knee point analysis suggests that the optimal number of keyword clusters is 126. These clusters were used to order the author supplied keywords.

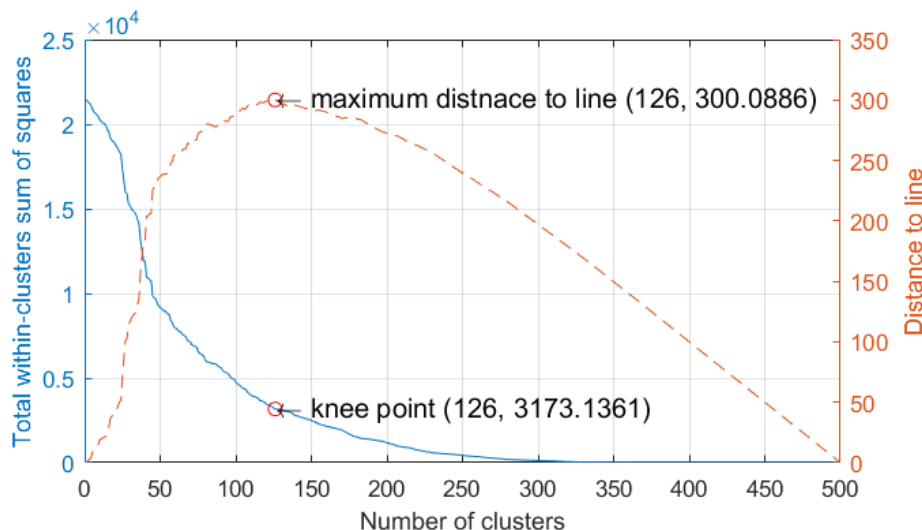


Figure 3. Knee graph analysis plots. The left scale is for the solid blue line plot. The right scale is for the orange dashed line plot.

After the initial ordering, we manually created a thesaurus file which maps the author supplied keywords to broader topics. These topics were ranked in order of their occurrence. In this study, we consider only topics that appear more than 12 times in the DESIGN conference publications. Introducing the threshold led to 65 topics which represent the 412 most important (weight) keywords. Tables 2 to 15, in the appendix, show the mapping of the author supplied keywords to topics. For

example, the first row in Table 2 details design, which is broad topic. The thesaurus file mapped the following 18 author supplied keywords to the design topic: creative design, design, design activities, etc. These 18 keywords appear 195 times in the DESIGN conference publications, i.e. their weight is 195. Hence, topic design crossed the threshold of 12 and therefore it is included in our analysis.

2.3. DESIGN topic visualization

Mapping and clustering are often used in bibliometric research to visualize some structural aspects from a large body of literature (van Eck and Waltman, 2014). Distance- and graph-based maps are two commonly used map types (Garfield, 2009). Distance-based maps establish the distance between two items. To be specific, the distance reflects the strength of the relation between the items (Van Eck and Waltman, 2010). We used the VOSviewer software to carry out the distance-based topic clustering (van Eck and Waltman, 2014; Van Eck and Waltman, 2011; Waltman et al., 2010).

The selected topics are visualized in a two-dimensional co-occurrence map. That map is constructed, from the 1892 conference publications, in the following way: the larger the number of papers which contain the same two topics together, the closer those terms will appear in the map (Kim et al., 2016; Khan and Wood, 2015). That distance is also used to establish an agglomerative hierarchical cluster tree (Anderberg, 2014). Pruning that cluster tree establishes the topic clusters.

3. Results

Figure 4 shows the co-occurrence map, which visualizes the relationship between DESIGN topics. Each node represents one topic and the node size indicates how often the topic appears in the 1892 analysed conference contributions. The distance between two nodes indicates how often the topics co-occur in the papers. That co-occurrence is also indicated by the edges between the nodes. To be specific, the edge line width indicates the link strength between two nodes. The node colour indicates the cluster to which the topic belongs. Overall there are 14 different topic clusters.

3	6	life cycle design, dfx, eco-design, sustainability, design tool, change management	268	54
4	6	product development, management, decision making, interdisciplinary, strategy, risk	273	31
5	9	cad, engineering practice, visualization, analysis, data analysis, design problem, fea, product, property	370	88
6	2	process, methodology	144	16
7	3	human factor, systems engineering, team	83	19
8	6	design theory, modelling, system, synthesis, requirement, prototyping	153	36
9	4	innovation, ideation, cognitive, evaluation	164	20
10	5	collaborative design, user centered design, industrial design, virtual reality, automotive	130	18
11	6	modularization, product family, product architecture, structure, cost, customization	148	25
12	5	design engineering, communication, information, integration, product data management	100	14
13	4	knowledge, research, empirical study, design practice	190	27
14	1	design education	70	3
Total 65			2357	397

4. Discussion

Figure 4 shows ten research hot-spots within the DESIGN conference, namely: design, process, product development, design method, knowledge, innovation, management, Computer Aided Design (CAD), engineering practice, and design education. From 1998 to 2016 the conference organizing committee called consistently for papers in the following seven areas: design process, product development, design method, design knowledge, management, engineering practice, and design education. Apart from the generic topic of design, innovation and CAD, there is a good correspondence between the research hot-spots and the conference areas. The following paragraphs discuss that relationship in more detail by establishing the connection between the DESIGN topic clusters and the conference areas.

Cluster 1, shown in Figure 4 and described in Table 1, contains the topic of design, which is the first research hot-spot. It is not surprising that design, with a weight of 195, is the main topic of the DESIGN conference.

Cluster 2 has seven topics including design method, conceptual design, mechatronics, optimisation, simulation, product service system, and service. It contains the research hot-spot of design method. This topic has the highest weight (109) within the cluster. That research hot-spot corresponds to the conference core area with the same name. This topic is well connected with conceptual design, product service engineering, engineering design in mechatronics, simulation, and optimization.

Cluster 3 does not contain a research hot-spot. In general, the topics grouped in this cluster, focus on tools for Design for X (DfX) and sustainability. Topics, such as eco-design, sustainability, and life cycle design, are all related to design methods which take into consideration social awareness of environmental impact (Lei and Moon, 2015). These topics link to the conference core area of sociotechnical issues. The change management topic focuses on change and how to manage it. This topic relates to a wide range of conference core areas, because change and how to cope with it comes up naturally in papers with an applied nature.

Cluster 4 contains two research hot-spots, namely product development and management with the weights of 110 and 82 respectively. These two hot-spots are related to the conference core area of organization & management.

Cluster 5 contains the research hot-spots of CAD and engineering practice. As such, the topics in this cluster are quite diverse. A possible connection between the topics comes from the fact that CAD, visualization, Finite Element Analysis (FEA) and data analysis are used in engineering practice. The topic of engineering practice is linked to the conference core area of engineering design practice.

Cluster 6 contains the research hot-spot process and topic methodology with the weight of 121 and 23 respectively. The link between the topics of process and methodology is strong. That can be interpreted as: analysis is a prerequisite for discussing process.

Cluster 7 has three topics including systems engineering, team, and human factor. These topics reside in the intersection which is formed the conference core areas of systems engineering and human behaviour.

The topics in Cluster 8 are rather diverse. The cluster links the topics of modelling, system, prototyping, synthesis, design theory, and requirements. Consequently, the topics in this cluster touch a range of conference core areas including design processes, theory, and research methods.

Cluster 9 contains the research hot-spot innovation, which has a high cumulative weight of 94. This is surprising, because in the years before 2018, the DESIGN conference management did not specifically call for papers that focus on design innovation. The organizing committee recognized that discrepancy as well. In 2018 design innovation was included in the call for paper as a new core area. Our clustering results support and indeed predicted that extension of the conference areas. We have mapped the following keywords to the topic of innovation: creativity, design creativity, eco-innovation, innovation, open innovation, innovation management, innovation methods, and innovation process, as stated in Row 1 of Table 10.

Cluster 10 addresses multidisciplinary aspects of design, by systematically discussing methods and tools which are relevant for both industrial and engineering designers. There is a strong link between industrial design and user centered design topics. That's because industrial design focuses principally on aesthetic and user-interface aspects of products (Laursen and Møller, 2016). Apart from automotive, other topics in this cluster correspond well with the conference core area of industrial design. The automotive topic shows little relationship with the other topics in the cluster. It might be the case that automotive was mapped to Cluster 10 because many authors used automotive related design examples.

Cluster 11 is focused on mass customization and product family design. Mass customization has received much attention and popularity from both industry and academia alike (Pine and Davis, 1999, Lei and Moon, 2014). Product family design has been recognized as an effective way of achieving mass customization. It provides an architecture that is based on platforms or modules which benefit from the economies of scale and, at the same time, leverage the costs of delivering variety.

Cluster 12 relates topics of information, integration, product data management, communication, and design engineering. The papers, which contain these keywords, fit into the conference core area of information and knowledge.

Cluster 13 includes the research hot-spot of knowledge. Knowledge and research are wide topics that correspond to the conference core areas of information and knowledge as well theory and research methods. The organizing committee focuses, through the call for papers, on documents which establish credible and convincing theories as well as research approaches by experiments. That explains why empirical studies fall into the same cluster as knowledge and research.

Cluster 14 contains only one topic which corresponds directly the conference area of design education. The co-occurrence map, shown in Figure 4, indicates that design education has weak links with many topics. Hence, papers with a focus on design education discuss other DESIGN topics from an educational perspective.

We have selected the web of science as our data source, because only this source had the required level of detail and the correct data format. The dependence on one source introduces a single point of failure. Indeed, the missing conference proceedings from 2012 and 2014 are a problem of the web of science database. We were not able to correct the problem while conducting the study. We had to work around that problem by downloading the missing data from the scopus database.

Another limitation is the arbitrary selection of 500 important keywords out of 4479 author supplied keywords that were found in the dataset. An incomplete literature survey shows that other published studies use considerably fewer keywords. For example, Zhang et al. (2017) used only the 100 most important keywords for their TDT study. In our case, the 100 most important author supplied keywords cover only 51.96% of the papers, i.e. 51.96% of all conference papers include at least one keyword from the list. This coverage was not acceptable for us. The default setting in the VOSviewer suggests using 500 author supplied keywords for the TDT analysis. In our case these 500 keywords cover 76.05% of the papers. Mapping 397 of these 500 keywords into 65 topics decreased the coverage to 73.97%. We considered this as an acceptable coverage.

5. Conclusion

In this paper we uncover DESIGN topics by visualizing and interpreting keyword data. The data comes from the web of science and scopus database. It contains all database entries for papers published in the international DESIGN conference from 1998 to 2016. These database entries are the raw material for a bibliometric analysis. Through string clustering and manual interventions, we mapped keywords into DESIGN topics. A co-occurrence map visualizes the connections between these topics.

Our work established the DESIGN topics from author supplied keywords. The selected 65 topics cover 73.97% of the papers published in the international DESIGN conference proceedings and available on the web of science database. Furthermore, the clusters correspond well with the DESIGN conference hotspots. The combination of these facts makes our results significant for all conference areas. Our analysis shows that the ten main topics are: design, process, product development, design method, knowledge, innovation, management, CAD, engineering practice, and design education. These topics correspond well with the conference areas outlined by the organizing committee in their call for papers. Visualizing the connections between the topics is beneficial for both conference organizers and participants. The DESIGN topic visualization can provide the conference organizers with key information for the selection of the conference topics. Conference participants can select a topic for or indeed create a new topic through their future research.

6. References

- Allan J. (2002) Introduction to Topic Detection and Tracking. In: Allan J. (eds) *Topic Detection and Tracking. The Information Retrieval Series*, vol 12. Springer, Boston, MA. doi: 10.1007/978-1-4615-0933-2_1
- Anderberg, M. R. (2014), *Cluster analysis for applications: probability and mathematical statistics: a series of monographs and textbooks*, Vol. 19, Academic press.
- Beil, F., Ester, M. and Xu, X. (2002), Frequent term-based text clustering, in “Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining”, ACM, pp. 436–442. doi: 10.1145/775047.775110
- Garfield, E. (2009), “From the science of science to scientometrics visualizing the history of science with histcite software”, *Journal of Informetrics* 3(3), 173–179. doi: 10.1016/j.joi.2009.03.009
- Khan, G. F. and Wood, J. (2015), “Information technology management domain: Emerging themes and keyword analysis”, *Scientometrics* 105(2), 959–972. doi: 10.1007/s11192-015-1712-5
- Kim, W., Khan, G. F., Wood, J. and Mahmood, M. T. (2016), “Employee engagement for sustainable organizations: Keyword analysis using social network analysis and burst detection approach”, *Sustainability* 8(7), 631. doi: 10.3390/su8070631
- Kodinariya, T. M. and Makwana, P. R. (2013), “Review on determining number of cluster in k-means clustering”, *International Journal* 1(6), 90–95.
- Laursen, E. S. and Müller, L. (2016), “Sharing the design intent between industrial designers and engineering designers”, in *DS 84: Proceedings of the DESIGN 2016 14th International Design Conference*.
- Lei, N., & Moon, S. K. (2014). “Decision support systems design for data-driven management”. In *ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (pp. V02AT03A016-V02AT03A016). American Society of Mechanical Engineers.
- Lei, N., Moon, S. K., & Rosen, D. W. (2015). “Redefining product family design for additive manufacturing”. In *DS 80-4 Proceedings of the 20th International Conference on Engineering Design (ICED 15) Vol 4: Design for X, Design to X*, Milan, Italy, 27-30.07. 15.
- Levenshtein, V. I. (1966), “Binary codes capable of correcting deletions, insertions, and reversals”, in *Soviet physics doklady*, Vol. 10, pp. 707–710.
- Levenstein, V. (1965), “Binary codes capable of correcting spurious insertions and deletions of ones”, *Problems of Information Transmission* 1(1), 8–17.
- Pine, B. J. and Davis, S. (1999), *Mass customization: the new frontier in business competition*, Harvard Business School Press.
- Small, H., Boyack, K. W. and Klavans, R. (2014), “Identifying emerging topics in science and technology”, *Research Policy* 43(8), 1450–1467. doi: 10.1016/j.respol.2014.02.005
- Suominen, A. and Toivanen, H. (2016), “Map of science with topic modeling: Comparison of unsupervised learning and human-assigned subject classification”, *Journal of the Association for Information Science and Technology* 67(10), 2464–2476. doi: 10.1002/asi.23596

- Van Eck, N. J. and Waltman, L. (2010), "Software survey: Vosviewer, a computer program for bibliometric mapping", *Scientometrics* 84(2), 523–538. doi: 10.1007/s11192-009-0146-3
- Van Eck, N. J. and Waltman, L. (2011), "Text mining and visualization using vosviewer", arXiv preprint arXiv:1109.2058.
- van Eck, N. J. and Waltman, L. (2014), "Visualizing bibliometric networks", in *Measuring scholarly impact*, Springer, pp. 285–320. doi: 10.1007/978-3-319-10377-8_13
- Waltman, L., van Eck, N. J. and Noyons, E. C. (2010), "A unified approach to mapping and clustering of bibliometric networks", *Journal of Informetrics* 4(4), 629–635. doi: 10.1016/j.joi.2010.07.002
- Yoon, B. and Park, Y. (2005), "A systematic approach for identifying technology opportunities: Keyword-based morphology analysis", *Technological Forecasting and Social Change* 72(2), 145–160. doi: 10.1016/j.techfore.2004.08.011
- Yu, H., Zhang, Y., Ting, L. and Sheng, L. (2007), "Topic detection and tracking review", *Journal of Chinese information processing* 6(21), 77–79.
- Zhang, Y., Chen, H., Lu, J. and Zhang, G. (2017), "Detecting and predicting the topic change of knowledge-based systems: A topic-based bibliometric analysis from 1991 to 2016", *Knowledge- Based Systems* 133, 255–268. doi: 10.1016/j.knosys.2017.07.011

[Corresponding author data - this will be inserted automatically]

[---]

[---]

[---]

[---]

[---]

[Do not delete or modify the author's data area]

Appendix: Mapping author supplied keywords to cluster topics

Table 2. Mapping author supplied keywords into the topics of cluster 1. Where W is the keyword weight and N is the number of keywords mapped to a topic.

Topic	W	N	Author supplied keyword
design.	195	18	creative design, design, design activities, design plans, design thinking, distributed design, early design phases, early stages of design, embodiment design, evolutionary design, inclusive design, lightweight design, limit state design, methodical design, participatory design, product design, robust design, designer

Table 3. Mapping author supplied keywords into the topics of cluster 2.

Topic	W	N	Author supplied keyword
design method.	109	7	boundary element method, design method, design methodology, design methods, method, methods, numerical methods
conceptual design.	47	2	concept selection, conceptual design
optimisation.	24	6	design optimisation, optimal design, optimisation, optimization, self-optimization, topology
simulation.	31	2	simulation, simulation-based design
mechatronics.	31	3	mechatronic, mechatronic design, mechatronics
service.	16	3	service, service design, service engineering
product service system.	19	4	product service system, product service systems, product-service system, product-service systems

Table 4. Mapping author supplied keywords into the topics of cluster 3.

Topic	W	N	Author supplied keyword
-------	---	---	-------------------------

eco-design.	44	9	ecodesign, eco-design, ecologies, ecology, environment, environmental impact, environmental impacts, bio-inspired design, global warming
life cycle design.	70	15	lean product development, life cycle, life cycle analysis, life cycle assessment, life cycle assessment (lca), life cycle design, life cycle thinking, life-cycle design, plm, product life cycle, product life thinking, product lifecycle, product lifecycle management, product upgrade, useful life
sustainability.	29	5	design for sustainability, sustainability, sustainable design, sustainable development, sustainable product development
dfx.	51	11	assembly-oriented design, design for disassembly, design for environment, design for quality, design for safety, design for x, design for x (dfx), dfx, maintainability, reliability, stability
design tool.	47	9	design structure matrix, design structure matrix (dsm), design tool, design tools, dsm, fmea, matrix, qfd, quality function deployment
change management.	27	6	change prediction, change propagation, change management, engineering change, engineering change management, product improvement

Table 5. Mapping author supplied keywords into the topics of cluster 4.

Topic	W	N	Author supplied keyword
product development.	11	03	integrated product development, new product development, product development
management.	82	11	agile, complexity, complexity management, design management, lean, management, portfolio management, process management, project management, requirements management, structural complexity management
strategy.	16	4	design roles, design strategy, strategic design, strategy
risk.	16	5	risk, risk management, risk assessment, organizational risk, software risk
decision making.	18	3	decision making, decision-making, decision-making theory
interdisciplinary.	18	5	disciplinary boundaries, interdisciplinary, interdisciplinary collaboration, interdisciplinary design, multidisciplinary

Table 6. Mapping author supplied keywords into the topics of cluster 5.

Topic	W	N	Author supplied keyword
cad.	82	16	cad, cad modelling, cad systems, cad/cam, cad/cam/cae, cae, caid, cax, computer, computer aided design, computer-aided design, computer aided design (cad), computer aided tolerancing, computer graphics, computer support, transfinite surface patches
engineering practice.	74	13	assembly, engineering, engineering design, engineering workbench, packaging, plant engineering, re-engineering, reverse engineering, rolling contact, rotational connections, simultaneous engineering, software engineering, technical drawing
visualization.	17	3	visual presentation, visualisation, visualization
analysis.	70	16	analysis, analyze, change impact analysis, competitiveness, mission analysis, nonlinear analysis, numerical analysis, process analysis, protocol analysis, similarity, spreadsheets, structural analysis, task analysis, tolerance analysis, traceability, uncertainty
data analysis.	14	8	data mining, digitisation, electronic images, genetic algorithm, neural networks, point-colud data, product data, run-time version of nn
design problem.	16	3	contact problems, design problem, problem solving
expert system.	11	3	expert system, expert systems, rule-based es
fea.	34	5	fea programs, fem, finite element analysis, finite element method, finite elements method
product.	25	11	advanced ceramic, clutch, ec5, gears, mobile robots, part, russian software, science-based products, screw compressor, shipbuilding, spc
property.	38	13	colour, constraints, feature, friction, haptics, impact damage, motion, pressure distribution,

properties, stiffness, tolerances, waste, wear

Table 7. Mapping author supplied keywords into the topics of cluster 6.

Topic	W	N	Author supplied keyword
process.	121	12	design process, design processes, development process, documentation, process, process change, process design, process improvement, process modeling, process modelling, product development process, standardization
methodology.	23	4	design research methodology, integrating design methodology, methodology, research methodologies and methods

Table 8. Mapping author supplied keywords into the topics of cluster 7.

Topic	W	N	Author supplied keyword
systems engineering.	26	4	complex systems, systemic approach, systems engineering, systems engineering (se)
team.	20	5	design teams, teams, teamwork, virtual team building, virtual teams
human factor.	37	10	anthropology, anthropometry, culture, emotional design, ergonomics, experience, healthcare, human behaviour in design, social media, human engineering

Table 9. Mapping author supplied keywords into the topics of cluster 8.

Topic	W	N	Author supplied keyword
modelling.	44	15	3d modelling, applied signposting model (asm), concept models, design models, functional modelling, geometric modelling, metamodel, model, model-based engineering, modelling, numerical modelling, physical models, product modelling, reference models, virtual models
prototyping.	24	5	prototypes, prototyping, rapid prototyping, rp, virtual prototyping
system.	29	9	dynamics, particle system, system, system dynamics, system architecture, systematic design, systems, technical system, technical systems
synthesis.	13	2	computational design synthesis (cds), synthesis
design theory.	31	2	c-k design theory, design theory
requirements.	12	3	customer requirement, requirements, software requirements

Table 10. Mapping author supplied keywords into the topics of cluster 9.

Topic	W	N	Author supplied keyword
innovation.	94	8	creativity, design creativity, eco-innovation, innovation, open innovation, innovation management, innovation methods, innovation process
ideation.	29	4	brainstorming, idea generation, triz, intuition
cognitive.	13	4	design cognition, cognitive capabilities, cognitive products, cognitive technical systems
evaluation.	28	4	design evaluation, evaluation, evaluation criteria, idea evaluation

Table 11. Mapping author supplied keywords into the topics of cluster 10.

Topic	W	N	Author supplied keyword
collaborative design.	43	4	collaboration, collaborative design, collaborative design process, co-design
industrial design.	26	2	industrial design, industrial design engineering
user centered design.	30	5	usability, user centered design, user experience, user-centered design, user-centred design

design.			
support.	10	2	decision support, design support
virtual reality.	16	3	augmented reality, virtual reality, vrml
automotive.	15	4	automotive, automotive industry, powertrain, racing car

Table 12. Mapping author supplied keywords into the topics of cluster 11.

Topic	W	N	Author supplied keyword
cost.	13	4	cost, cost estimation, cost/benefit analysis, target costing
customization.	13	3	customization, mass customization, product configuration
modularization.	36	5	modular architecture, modular function deployment, modularisation, modularity, modularization
product architecture.	25	4	architectural design, architecture, naval architecture, product architecture
structure.	16	4	product structure, product structuring, structure, timber structures
product family.	32	5	platform, product family, product platform, variant design, variant management

Table 13. Mapping author supplied keywords into the topics of cluster 12.

Topic	W	N	Author supplied keyword
design engineering.	28	3	collaborative engineering, concurrent engineering, design engineering
information.	20	3	information, information flow, information retrieval
integration.	18	3	integral design, integrated design, integration
product data management.	16	3	edm, pdm, product data management
communication.	18	2	communication, design communication

Table 14. Mapping author supplied keywords into the topics of cluster 13.

Topic	W	N	Author supplied keyword
empirical study.	33	7	case studies, case study, empirical study, experiment, industrial case study, observational studies, scenarios
design practice.	12	2	design practice, reflective practice
research.	33	5	action research, design research, solution search, design science, empirical research
knowledge.	103	13	design knowledge, knowledge, knowledge base, knowledge based engineering, knowledge engineering, knowledge management, knowledge processes, knowledge representation, knowledge sharing, knowledge structuring, knowledge support, knowledge transfer, method transfer

Table 15. Mapping author supplied keywords into the topics of cluster 14.

Topic	W	N	Author supplied keyword
design education.	70	3	design education, education, engineering education