Abstract

Purpose – The objectives of the study were to demonstrate the suitability of methodology based on a text mining toolset for detecting trends in scientific papers and to find trends that were present in the field of project management during the research time span (2000–2019).

Design/methodology/approach – An approach based on text mining tools supported by expert analysis was adopted due to an extensive number of publications in the field of project management. The novelty of the approach lies in the proposed method of trends discovery instead of the commonly used trends predefinition. The use of computer support allowed the full texts of papers, and not only abstracts, to be analysed, which significantly increased the reliability of the achieved results. Overall, 3,544 papers from seven journals were analysed.

Findings – As a result, 43 trends were discovered including seven long-lasting, four declining, 17 emerging and 15 ephemeral trends. Trends were analysed in comparison with the results of previous studies and project management frameworks. New trends and potential gaps were discussed.

Originality/value – The results highlight the topics of research that gain popularity among researchers, and which are related to the current problems that arise in project management. Therefore, the results can help focus studies on the most important areas, as well as find new ones which are not so popular yet. The knowledge of current trends is also important for those project managers who seek to improve the efficiency of their work.

Keywords Systematic review, Research trends, Text mining, Project management

1. Introduction

The importance of project management in management science has been growing in recent years. This can be proved by the number of papers published each year. Moreover, the field is still evolving: old methods undergo modifications, new tools are created and researchers propose new points of view in their studies. Keeping up to date requires the application of methods that will enable practitioners and researchers to monitor current and emerging trends in the field [1].

Systematic or structured literature review methods are widely used in research to uncover new scientific trends and relevant subjects for future research, but they are very time-consuming and require significant effort (O’Mara-Eves et al., 2015; Crossan and Apaydin, 2010; Staples and Niazi, 2007; Tranfield et al., 2003; Kwak and Anbari, 2009; Walker and Lloyd-Walker, 2016; Prater et al., 2017; Miterev et al., 2017). Systematic review is able to provide valuable knowledge about studied phenomena. Therefore, the method is becoming increasingly popular among researchers. However, due to enormous number of papers, classical approach to systematic review without the help of text mining toolset is exposed to selection errors, omissions, simplifications and, consequently, potentially incomplete results. Those problems were discussed in detail in the literature review section. The review of previously published analyses of changes in the field of project management led to the
discovery of significant limitations of methods used. The identified research gap is related to the lack of an unbiased method of trends identification.

We have adopted a working assumption that the early detection of nascent, non-obvious trends in scientific research will allow researchers addressing topics of potentially high value for science and practice. Simultaneously, the identification of declining research topics will enable resources to be moved to less studied subjects. The study should help answer the following questions (Hällgren, 2012):

1. Does full-text analysis of publications using a text mining toolset enable researchers to obtain higher-quality results than in the case of using only abstracts or keywords?

2. How are the interests of researchers changing in the field of project management, and which topics have gained popularity in recent years?

Therefore, we have defined following objectives. The first one is to demonstrate the suitability of the proposed methodology based on a text mining toolset for detecting trends in scientific papers. The second one is to find trends that were present in the field of project management during the time span of the research (2000–2019). This includes discovering the trends that have emerged in recent years and can become an interesting field of study in the near future.

This study contributes to the field of project management research by providing a critical analysis of the existing systematic literature review methodologies and their results, followed by the application of an approach based on text mining and big data toolset. The results present changes in this field and suggest further promising themes of research.

The article fits in with an emerging area of “science of science”, as it reflects a desire to predict new ideas in advance, before they become grounded paradigmatic theories (Clauset et al., 2017). It stems from a post-positivist position (Biedenbach and Müller, 2011).

The paper is organised as follows. In the next section, we present previous systematic literature reviews, and identify and describe their limitations. Then we present methodology that helps solve those limitations using text mining toolset. In the following sections, we present results, discussion and conclusions.

2. Literature review

2.1 Prior literature reviews on project management

Since 1995, we found 17 general literature reviews on project management and 36 thematic studies – dedicated to one subject. The most common subjects were stakeholders (6), project success (3), human capital (3), BIM (3), PPP (2) and megaprojects (2). The researchers adopt different approaches to the acquisition of the papers. The two most common criteria are keywords and a selection of journals. In the case of 11 general and 26 thematic studies, keywords were the main criterion. Titles or abstracts were used as additional criteria in several cases (three general and 14 thematic studies). The main advantage of the approach based on keywords is the possibility of finding papers regardless of the place of their publication. However, this is limited by the improper choice of keywords: a few or even only one. In many cases, keywords do not fully reflect the content of the article (Pollack, 2015). As an example, we found that many papers published in the International Journal of Project Management do not contain the keyword “project management”, even if they are related to this field of study. If the selection of a paper was based on the occurrence of that keyword in the keywords section, title or abstract, as many as 26% of papers related to project management would be omitted (editorials and announcements were excluded). Similar phenomenon occurs in multiple journals.
The approach based on the selection of journal was used in six general and ten thematic studies. Most of them used *International Journal of Project Management, Project Management Journal* and *International Journal of Managing Projects in Business*. Journals dedicated to the field of study usually contain a full cross-section of important themes. However, not in every field can such journals be found. As the publisher’s policy can affect the selection of papers, multiple journals should be analysed in the research.

The analysis of papers based on abstracts and keywords only was used in seven general and 11 thematic studies. The research performed in the field of biomedicine revealed that “authors report fewer than 7.84% of scientific claims in an abstract” (Blake, 2010). The same result should be expected in the field of management. This claim significantly emphasises the need to analyse the full texts of articles. Therefore, in ten general and 25 thematic studies, the researchers claimed that at least part of papers was scanned. In many papers, that stage was vaguely explained, lacking information about the method (scanning, reading, etc.) and the number of papers. Based on 22 studies where data were available, we found that, on average, 23% of selected papers were analysed (the lowest: 2%). There were no attempts to analyse the full texts of all papers. As Tranfield *et al.* (2003) found, literature reviews related to that field of study were mainly narrative and qualitative. The first stage of analysis is often based on the titles or abstracts of papers. Only after selecting a limited amount of articles can a full text analysis be performed.

The researchers attempted to present the classification of papers based on one of the methods: systematic review, simple statistics, keywords, keywords co-occurrence, network of citations, expert evaluation, own coding, word frequency and term frequency–inverted document frequency (TF-IDF). In recent years, in a growing number of studies, computer applications have been used, including Vosviewer, Framenet, Citespace, Bibexcel, Nvivo, Sci tool, Knime, Leximancer and Leximappe. The criteria of classification were based on papers’ keywords, predefined classes or discovery of classes. The source of predefined classes was PMBOK, APM or IPMA standards, and also own categories. The discovery of classes was performed in most cases using a citation chain or keyword co-occurrence. The authors used different names for classes: topics, clusters, trends, disciplines, categories and areas.

A list of analysed general literature reviews in the field of project management published in 1995–2018 was presented in Table 1. The authors of eight papers performed an analysis of publications over a space of time, which presented trends. Those papers will be discussed in more detail, as their findings are important for our research.

### 2.2 Prior approaches to trends identification

Urli and Urli (2000) used Leximappe to create links between keywords and evaluate their strength. Themes were identified using two measures: centrality and density. As a result, three complementary categories of themes were identified: project tools, activity sectors and particularities of project management. The sample of papers was selected based solely on keywords, which is exposed to the threats discussed earlier.

Crawford *et al.* (2006) used keyword analysis technique to create set of keywords and topics. Subsequently, authors classified papers into topics according to the occurrence of keywords. They evaluated the significance of 18 topics and changes in attention paid by researchers. The authors recognised that abstracts and titles were highly indicative of the content of articles and better than full-text analysis, but no evidence was shown.

Kwak and Anbari (2009) selected 18 mainstream business and academic journals. Authors categorised 537 journal articles into eight predefined so-called allied disciplines.
They concluded that “more publications of PM research in allied disciplines... are being recognised and published in mainstream management journals, and the trends of future research related to PM are strong and healthy”. Classification criteria for mainstream journals were not presented. Therefore, it is unclear why main PM journals were excluded from analysis.
Artto et al. (2009) used co-citation analysis to find relations between papers. High co-citation intensity was used in the cluster indication process. The authors identified four project clusters: product development A, product development B, organisation and product design, knowledge creation and additional 4 programme clusters. The selection of papers was based on the occurrence of “project” and “programme” keywords. In order to limit the complexity of the research, several arbitrary cut-off parameters were used.

Morris (2010) reflected on fundamental research issues facing PM as a discipline. The paper poses questions about where PM research had been heading over the previous five decades and how it can remain relevant in creating value to its clients and key stakeholders. The author concluded that research contribution in this area was often remote from the problems that PM practitioners faced. A shift of interest towards researching organisational phenomena was identified as a possible reason for this. The identification of trends was based on the author’s experience and subjective assessment.

Turner (2010) presented data about citations of papers published in IJPM in three selected years. The number of topics of articles published in this journal grew from 25 in 1997, to 54 in 2007. The author pointed to several emerging topics, such as programme and portfolio management. The study was based on citation analysis.

Polack and Adler (2015) used Sci tool to analyse the keywords co-occurrence network. By identifying keyword bursts, they were able to identify an increasing interest of researchers in certain subjects. The number of presented keywords was arbitrarily limited to 20. The authors indicated the limitations of using keywords and abstracts in the selection process; however, their choice of papers was based on one keyword (project management).

Padalkar and Gopinath (2016) presented an analysis of the evolution of various themes in project management research. The authors adopted a historiographic approach based on prior literature reviews to identify key concepts. Subsequently, they selected 230 papers using an effective annualised citation rate. As a result, 17 themes in three periods were identified. However, the predefinition of key concepts could limit the chance of finding new ones.

2.3 Limitations of prior literature reviews

The analysis of prior literature reviews in the field of project management, which have been mentioned above, allows detecting limitations, including:

(1) significant reduction in the number of full texts analysed,

(2) omission of papers related to the field of study due to relying on keywords only,

(3) inclusion of papers faintly bound to the field of study because of too wide search criteria,

(4) analysis of abstracts which can be poor representation of the full texts,

(5) predefined themes or clusters can prevent the discovery of new ones,

(6) classification conducted by researchers without the help of quantitative methods can induce significant bias,

(7) lack of quantitative data to describe behaviour of trends in time.

The methodology adopted in this study was designed to overcome such limitations. Increased access to full-text scientific articles and literature published online facilitates the use of text mining methods. An analysis of a large quantity of textual data and the identification of...
potentially interesting research subjects and trends can be done with the help of computer tools. The automation of these tasks significantly reduces the workload required to find and read every article, select articles for research, identify publications on similar topics or discard irrelevant articles from analysis. Text mining methods limit the impact of researchers’ bias in the process of keyword definition, and make it possible to go beyond the framework of a given paradigm for a specific field of science.

3. Methodology

The most popular methods of systematic literature review, including those using text mining tools, require the predefinition of keywords that will be used in the categorisation process. Thus, a researcher is able to formulate a hypothesis about such an issue as the popularity of certain keywords. The discovery of new, not predefined, keywords is difficult due to researcher bias, limited knowledge and the adopted approach. In order to discover trends rather than assume them, the approach has to be changed. A researcher has to find keywords in data, and analyse and understand them. The hypotheses can be formulated only after completing the whole process. That is similar to the approach applied in grounded theory. In the proposed methodology, we use well-known text mining tools and complement them with our own method of trends discovery. A diagram of the proposed approach compared with systematic literature review steps is presented in Figure 1. The approach is consistent with general rules for systematic literature reviews (Ananiadou et al., 2009).

<table>
<thead>
<tr>
<th>Typical SLR steps</th>
<th>This research</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Planning</strong></td>
<td></td>
</tr>
<tr>
<td>Define research question</td>
<td>Y</td>
</tr>
<tr>
<td>Select method used for review (text-mining, full text analysis)</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Selection</strong></td>
<td></td>
</tr>
<tr>
<td>Define inclusion criteria (journals listed in Scopus)</td>
<td>Y</td>
</tr>
<tr>
<td><strong>Extraction</strong></td>
<td></td>
</tr>
<tr>
<td>Collection of articles (3544 scientific papers)</td>
<td>Y</td>
</tr>
<tr>
<td>Preprocessing (conversion of papers to bags of words)</td>
<td>Y</td>
</tr>
<tr>
<td>Clustering (TF-IDF and HDBSCAN)</td>
<td>X</td>
</tr>
<tr>
<td><strong>Execution</strong></td>
<td></td>
</tr>
<tr>
<td>Trend discovery (43 trends)</td>
<td>Y</td>
</tr>
<tr>
<td>Analysis and interpretation</td>
<td>Y</td>
</tr>
</tbody>
</table>

Source(s): own study

Figure 1.
Diagram of the method and steps of systematic literature review (SLR)
3.1 Collection of papers

In order to meet research objectives using the described methodology, it was necessary to collect an extensive set of papers on project management. The criteria for choosing papers were as follows:

1. the main subject of a journal related to project management,
2. listed in Scopus database,
3. availability of the full text of an article,
4. published between 2000 and 2019,
5. only scientific papers were included.

To ensure thorough coverage of the field, we chose all published articles that met the criteria from seven journals. Similar approach was used by i.a. (inter alia (among others)): Walker and Lloyd-Walker (2016, 2018), Uchité et al. (2016), Padalkar and Gopinath (2016), Turner (2010), Crawford et al. (2006). To decrease the risk of bias related to journal policy, multiple journals were analysed, as well as multiple volumes:

1. *International Journal of Project Management* (IJPM) – years 2000–2019, 1707 papers,
2. *Project Management Journal* (PMJ) – years 2000–2019, 693 papers,
5. *International Journal of Information Technology Project Management* (IJITPM) – years 2010–2019, 198 papers,

As mentioned earlier, many studies on extensive data repositories are limited to the keywords, titles or abstracts of scientific papers. The use of exploratory data analysis (EDA) tools on short pieces of text is less time-consuming and requires less computing power. It can, however, lead to the misunderstanding of data. As a title usually contains 5–9 words, the relative importance of the single occurrence of a word is higher than in the case of the full text. A specific structure of abstracts leads to the use of words that highlight the terms that are important for authors, but not necessarily the most important terms in the paper. They are sometimes used to increase readership. Therefore, this research study analyses only the texts of papers (without titles, abstracts, keywords and references).

The full texts of papers published in aforementioned journals were downloaded from academic databases. No duplicates were found. Editorials, calls for papers, errata and book reviews were excluded from download. In all, 3,541 scientific papers were subject to further analysis. Titles, abstracts, keywords and references were removed from each paper.

The size of the sample should be considered sufficient. The authors carried out an experiment in which the number of papers examined was gradually increased. Changes in results between corpuses consisting of approximately 2,200 (three journals), 2,800 (four journals) and 3,541 papers were insignificant. Therefore, further increase of the sample was not justified.
Additional sources were used to obtain supplementary data related to papers, for example, Crossref. Each paper was converted into a text file and prepared for automatic analysis using computer algorithms. In this research, algorithms have been created using Python libraries, including grobid, nltk, scikit-learn, hdbscan and scipy (Lopez, 2009; Pedregosa et al., 2011; McInnes et al., 2017; Jones et al., 2001).

3.2 Search for the most important terms
The simplest method of finding which ideas, subjects or terms are gaining popularity among scientists is to count how frequently they appear in scientific papers. However, that approach has some disadvantages. First, in return, a researcher gets not only keywords, but also other less-important words, as the most popular English words are “the” and “of”. Moreover, many keywords are insignificant for that type of analysis – in the case of project management, the word “project” occurs in every paper. Therefore, the terms that appear in all papers contribute little to the analysis, those which appear in a moderate number of texts can provide valuable information and those which appear only in few documents are important only for those documents (Cong et al., 2016; Salton and Yang, 1973). A solution to that problem is offered by the term TF-IDF method. It takes into account the term frequency, and also the number of documents in which the term occurs. In order to assess the importance of each term, it uses the following formula:

\[ w_{ij} = tf_{ij} \cdot \log \left( \frac{N}{df_i} \right) \]

where:
- \( w_{ij} \) – result for term \( i \) in document \( j \),
- \( tf_{ij} \) – number of occurrences of term \( i \) in document \( j \),
- \( df_i \) – number of documents containing term \( i \),
- \( N \) – number of documents in corpus (set of documents).

The result of the TF-IDF method is a set of terms which have been assigned an important value. The most important terms describe the content of the analysed text. They can be used as keywords instead of those proposed by the authors of the paper. Each analysed document can be presented as a vector consisting of all the terms and corresponding values. A comparison of the similarity of all documents in the corpus is possible using those vectors, for example, cosine similarity method (Mihalcea et al., 2006).

The quality of the results can be improved owing to proper text preparation. As TF-IDF analyses separate terms, the papers should be converted into “bags of words” – all words should be converted into singular, present tense, and all punctuation marks should be removed, as well as frequent, but not important, words (stop words, e.g. “the”, “of”, “a”, “for”). That, however, can lead to the misinterpretation of some terms, as, for example, “quality management” is not the same as “quality of management”. To avoid this, apart from single words (unigrams), multiword terms should also be put into the bag of words: bigrams, trigrams, tetragrams. In the case of the last two, internal stop words should not be removed. The TF-IDF method is widely used in EDA, but it has some limitations (Zhang et al., 2011):

1. it is not derived from a mathematical model, but usually explained using Shannon’s information theory,
2. in large data sets, it requires extensive computation,
3. it is not able to discover synonyms and polysemous words,
4. individual words can have different meanings depending on the context.

Most of those limitations can be minimised during the preparation stage of the analysis. Synonymy and polysemy in scientific papers have a limited impact due to more precise language and terms used by authors.
Some other methods can be used instead of TF-IDF, such as latent semantic indexing (LSI) or latent Dirichlet allocation (LDA). They can explain the meaning of text, but the format of results makes their use in the further steps of our cluster analysis difficult.

3.3 Discovery of thematic groups (clusters)
The vectors created using TF-IDF can be used to discover the thematic groups of papers. There are two main approaches to the discovery process – partitioning and hierarchical clustering. The result of the former approach is partitions where each element of the corpus belongs to one partition. The result of the latter is clusters containing some elements and free elements which do not belong to any cluster. The chosen approach should be based on the set objectives. If the objective is to put all elements into some category, partitioning is the best option. The side effect is that papers on completely different topics have to be classified into one from particular categories. Usually, a number of categories have to be predefined, which additionally limits the effectiveness of this approach. If the objective is to find the most important groups, clustering should be used. It facilitates focusing only on those papers that concern very similar topics. Therefore, clustering is better when searching for trends. In such an analysis, only real, strong thematic groups should be discovered. The examples of partitioning methods include k-means, affinity propagation, spectral clustering and agglomerative clustering, while the examples of clustering are mean shift (based on k-means), DBSCAN, Optics and HDBSCAN (Jain, 2010; McInnes et al., 2017).

The HDBSCAN (hierarchical density-based spatial clustering of applications with noise) is a relatively new method developed on the basis of DBSCAN (McInnes et al., 2017). It discovers clusters based on density (groups of elements that are close to each other, e.g. papers on similar topics), and limits the number of required predefined variables to only one – the minimum size of the cluster. The HDBSCAN algorithm takes each element and checks at what distance it can find the required number of similar elements. Then, the results for all elements are compared, and the densest areas are detected as clusters. Density and the number of elements in the cluster can differ. This helps to decrease the probability of a situation occurring where two clusters merge because of the element which is at a similar distance to both of them. This situation could happen, for example, when one scientific paper equally emphasises two different topics. It was one of the limitations of the older DBSCAN method. The HDBSCAN algorithm can be modified by several optional parameters, of which the most important is the cluster selection method, allowing larger but more diverse clusters, or smaller but more homogenous ones. In the research, both approaches were tested, and the latter was chosen, as it delivered better results for trends analysis.

The sample was divided into groups that contained papers published in 5-year overlapping periods, starting with 2000–2004 and ending with 2015–2019. Each paper was assigned to all the groups into which it fitted. Cluster analysis was performed in every group separately. Figure 2 presents the number of articles in each period (corpus).

Due to limited length of the paper, it is not possible to describe HDBSCAN algorithm in detail. Full documentation with examples and comparison to other methods can be found on dedicated website (hdbscan.readthedocs.io).

3.4 Discovery of trends
Clusters are static and require further analyses to discover trends. Cluster analysis can be performed for sets of papers published each year, and then compared. Such an approach has some limitations, as in each year, entirely different sets of clusters are discovered. Better results can be achieved if each corpus contains papers published in several subsequent years, and each next cluster analysis is shifted one year further. Each step of analysis leads to the discovery of slightly different clusters in which paper tracking is possible. Thanks to
tracking, an evolution of clusters can be observed, and trends can be identified. As a result of the analysis, several types of trends can be discovered:

1. long-lasting trends that exist and evolve during the studied period,
2. declining trends which end during the studied period,
3. emerging trends which begin during the studied period,
4. ephemeral trends which begin and end during the studied period.

In this research, each corpus contained papers published in subsequent five-year periods. The minimum size of the cluster was set to five on the basis of the results of experiments (one paper per year on average, written by different authors). The adoption of a larger cluster size would lead to lower resolution and the hiding of small emerging trends which are visible when a smaller cluster size is used. An excessively low cluster size, however, would lead to the discovery of a large number of small trends that would be difficult to analyse. An optimum cluster size can depend on the corpus size. Therefore, it should be determined experimentally.

Proposed approach enables algorithm to track and discover trends. Moreover, it may deliver better results than analysis performed by human, as it is free of bias. Previous knowledge and beliefs of experts may be a limitation in such study.

It needs to be highlighted that papers not included in clusters are not less important. For example, an average number of citations of papers in clusters and out of clusters is similar (analysed yearly on the basis of data obtained from CrossRef.org). The reason for not being included in thematic groups is the selected minimum size of cluster. Papers which are out of clusters touch upon important topics studied by very few researchers. Among them, future trends can be found that have not been discovered in the current research.

3.5 Trends verification, description and interpretation
This step has to be performed by researchers without the help of automatic algorithms. First, they must decide whether trends have been identified correctly. Two types of errors can be
found at this stage: unjustifiable merger of two clusters and the omission of papers on the subject. Both errors may indicate that the parameters of HDBSCAN should be corrected. The objective of the description phase is to find the best suited names of trends. Trends can be described on the basis of the most important terms found during the TF-IDF analysis of papers that constitute the trend. The interpretation phase should help to highlight changes within the trends and try to predict their future evolution.

4. Results
4.1 Discovery of clusters
The number of clusters found using HDBSCAN has been presented in Figure 3. It was not possible to include every cluster in the trends due to several reasons. In order to discover the trend, the flow of papers between at least two clusters in subsequent periods was required. Therefore, clusters that had no successors were omitted. Moreover, the initial assumption was that trends have to involve multiple independent authors, which leads to the omission of clusters that consisted of papers written by only one author. Thus, the number of clusters cannot be directly compared with the number of trends.

It should be noted that the growing number of papers presented in Figure 2 had a small impact on the number of clusters, which indicates a limited influence of the corpus size on the results.

As mentioned earlier, the fact that the paper was not included in one of the clusters was related only to the predefined minimum cluster size. Overall, approximately 50% of papers were included in clusters, which allowed us to identify the main trends of research in the field of project management. The subject of the other papers was so diverse that the algorithm did not include them in the identified clusters (too large distance between vectors representing papers), nor did it create new ones (too few similar papers).

4.2 Presentation of trends
In the next step, the flow of articles in clusters between periods was analysed, leading to the discovery of trends. In several cases, the algorithm identified that the theme of the trend
evolved over time, leading to division into subthemes. For the sake of clarity, they were presented as one trend. Due to the limited size of the paper, the evolution of those trends will be a subject of further research.

The discovered trends were named on the basis of the most important keywords (results from TF-IDF), as well as titles and abstracts of included papers. The list of trends on a timescale is presented in Table 2. The list was ordered by the year of the first occurrence of the trend and its length.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.
Trends in research on project management in years 2000–2019
Trends discovery led to the identification of 43 trends: 7 long-lasting trends, 4 declining trends, 17 emerging trends and 15 ephemeral trends. The year of the trend discovery indicates an increase in popularity among researchers, not the creation of the idea behind it. For example, the first papers on megaprojects (trend 40) were published several years before 2011.

The long-lasting trends are not homogenous and change over time. The evolution of trends can be tracked using keywords of high importance in subsequent years. The importance of the keywords was evaluated by TF-IDF algorithm and averaged for each cluster. Comparison of cluster keywords reveals new topics within the trends (Table 3). The evolution of the trends may lead to the disappearance of earlier topics or to their parallel development. Examples of the former are programme management and knowledge management, while the example for the latter is public-private partnership.

The declining and ephemeral trends are associated with issues that are of less interest to researchers, were solved or there has been a change in the approach to their study. Disappearing of certain trends is a normal phenomenon in science. Disappearing of the trend can be predicted to a certain degree, thanks to observation of decreasing average number of papers (Table 4).

Finally, 17 trends were found that emerged during the analyzed period and are still actively studied. These trends present a wide scope of research topics in the field of project management.

Table 2 presents list of the trends; however, it is unclear what is the dynamics of development of the currently active trends (long-lasting or emerging ones). Observation of changes in average number of papers per year (ANPY) in subsequent periods allows answering this question. In Table 4, all active trends were divided into three groups, in which ANPY in recent years was decreasing, but the increase was stable. Additionally, to depict strength of those trends, the average number of papers in the last studied period (2015–2019) was presented.

The highest and still increasing activity is assigned to five trends related to EVM, portfolios, stakeholders, governance and megaprojects. On the other end, low and decreasing activity can be observed in two trends: programme management and leadership, and teamwork and team performance. Analysis of that table can hint researchers which trends are worth undertaking further research. As an example, below we present an analysis of a
group of trends with increasing activity in recent years for which the average number of articles published annually is four or more.

Articles identified in trend “Earned Value based project monitoring” relate to effective schedule and costs monitoring. In recent years, researchers put more emphasis on using statistical methods and forecasting. Lipke et al. (2009) proposed extension of EVM with statistical prediction to enable more accurate forecasts that will help managers in decision-making. Kwak and Anbari (2009) reviewed history and current practices of using EVM in government, as well as provided recommendations to improve and broaden applications of the method in future projects. Vanhoucke (2012) analysed experiments performed in simulation study on the efficiency of project control techniques in order to find causes of successes and failures of using EVM in different projects. Wide overview of literature on earned value management was presented by Willems and Vanhoucke (2015). Our review indicates that this issue is still valid as the subject of research and publication.

Next trend identified in the group “Project Portfolio Management” includes early-year papers related to formalisation of portfolio management and its relation to strategy. In the following years, we identified more papers related to risk and agile approach in portfolio management. Meskendahl (2010) found that many companies encounter problems with implementation of strategy and propose conceptual model on the relationship between strategic orientation, project portfolio management and success. Teller et al. (2012) complement that result with more elaborate study on formalisation and relations between individual projects and whole portfolio. Topic of risk in portfolio management is presented, among others, by Teller et al. (2014) and Martinsuo et al. (2014). Authors of the first paper propose adopting a wide perspective to risk management that allows covering all projects in portfolio. Formal approach to risk management on both project and portfolio level was found to be positively associated with success of whole portfolio. The second paper explores uncertainties in project portfolio management and how managers frame them and deal with
them. The growing complexity of portfolios as well as the need for increased agility (Stettina, Hörz, 2015) demands finding new approaches to project portfolio management, which makes this topic an interesting area of further study.

Stakeholder management (identified as trend since 2009) at first included methods of stakeholder analysis and then expanded to problems of communication with specific types of stakeholders. Aaltonen (2010) discusses interpretation processes leading to recognition of external stakeholders environment. Based on four interpretation modes, the author presents the importance of understanding factors that help understand stakeholders. Butt et al. (2016) focus on stakeholder communication with special respect to process of project change. They use case study method to find how communication routines affect stakeholder involvement. In recent years, growing number of authors mention local communities as stakeholders, which suggests that in future more research studies on local communities will be performed and published.

For the trend “Governance and governmentality”, quite a large number of scientific articles was identified. Ahola et al. (2014) examined general governance research on governance published outside the area of project management journals, and related its concept to definitions presented within the PM literature. Too and Weaver (2014) presented comprehensive literature review on project/programme management, governance, portfolio management and project management office. Authors analysed what scope of the concept of “governance” is present in other thematic areas undertaken in the PM literature. They also studied the relationships between these concepts. Müller et al. (2016) proposed practical framework for implementing proper governance processes within organisation. Main conclusions presented in the article refer to identified sub-dimensions of governance, that is, sovereignty (external autonomy and internal control), governance mechanisms (balance of trust and control) and governance institutions (external control). The issue of trust and ethical considerations for project governance is also taken in another article identified in this trend by Müller et al. (2014).

The last trend taken into consideration due to the growing interest among scientific community includes subject of “Megaprojects” identified since 2011. In the first years, researchers focused mainly on high level of difficulty of managing large, extensive projects, which manifests itself, among others, mainly by cost overruns. However, recently, there are growing number of papers that study social responsibility and environmental issues in megaprojects. Eweje et al. (2012) surveyed 69 managers of megaprojects and demonstrated that information feed to project managers has significant impact on the strategic value created by those projects. Authors also observed important risk areas that influence value creation. Flyvbjerg (2014) presented an overview of rules that are important in managing as well as studying megaprojects. The author defined what type of projects should be included in megaprojects category, what are the reasons to undertake their implementation and what are typical phenomena related to them. Zhou and Mi (2017) studied publications related to social responsibility in context of megaprojects and applied systematic review method based on abstracts and keywords. The importance and the number of megaprojects grow each year, and we believe that that it will lead also to growing of number of studies related to discussed trend in future

5. Discussion
5.1 Evaluation of the results adequacy
The results of this research are generally convergent with the trends and themes identified in previous studies in the field of project management. However, such a comparison is not easy due to different research assumptions. An analysis of the used methodologies has shown that only three studies presented in Table 2 delivered results in a form that allowed direct comparison. Even then, several limitations apply:
only adjacent or overlapping periods could be compared,
(2) different naming convention was applied by authors,
(3) lack of information of which papers were included in a certain trend or theme,
(4) different trends and themes identification criteria.

The study by Morris (2010) covered the years 1950–2009. In seven periods, 15 trends based on own reflection and experience were identified. Eight trends were active in the last period that began in 2000: philosophy, assurance/governance, benefits/value, phasing/scheduling, performance management, organisation, knowledge management/organisational learning, people. On the basis of trends description, it was found that seven of them correspond to the trends discovered in this study. The study by Turner (2010) covered three years: 1987, 1997 and 2007, and identified over 60 topics in nine classes. The topics were identified if at least one paper published in IJPM was included. It was found that out of 33 trends active in 2007 in our study, 23 were also presented by Turner. These examples prove that combination of text mining methods proposed in this paper is able to deliver expert knowledge about changes that take place in the analysed field of study. Moreover, the results can be easily repeated in subsequent years and serve as monitoring tool.

The study by Padalkar and Gopinath (2016) covered the years 2000–2015, which is very similar to the time frame of this study. The authors found 17 themes present in project management literature. In this study, more than twice as many trends were found, including all mentioned in a study by Padalkar and Gopinath. It proves, that proposed approach returns much more detailed results than systematic literature review performed without help of advanced text mining toolset. Larger sample size and using quantitative methods lead to better distinction of similar but separate trends.

The above comparison with previous studies presents a relative convergence with prior results, which allows the formulation of the claim that the methodology used in this research returns results consistent with other approaches. At the same time, it allows the discovery of more detailed trends and the indication of their time frame.

5.2 Novelty of proposed approach
In the literature review section, several limitations of previous systematic reviews were identified. This study was designed to overcome all of them:

(1) The labour intensity was reduced by using text mining algorithms and computer-aided analysis. That allowed authors to significantly increase the number of analysed papers.

(2) The problem of too wide or too narrow selection of papers was solved by the identification of leading journals dedicated to the field of study, instead of search based on keywords.

(3) The risk of errors during classification based on abstracts was mitigated by analysis of full texts.

(4) The risk of researchers’ bias was mitigated by using algorithms that allowed resignation from predefinition of the trends. That also allowed discovery of new trends that were not identified in previous studies.

(5) The data delivered by algorithms allowed not only to identify the trends, but also to describe their behaviour in time. This can be an important clue for researchers.

Analysis of large samples of papers is possible on regular desktop computers, thanks to very efficient algorithms. Today, the main role of the researcher lies less in analysis and more in
the design of the study and interpretation of the results. It requires new skills, going beyond the field of study. It should be noted that computer-aided systematic reviews will not eliminate classical ones for a long time due to limitations of algorithms. They work best with big samples that would require enormous workload of large teams of researchers.

5.3 Coverage of PM frameworks – potential gaps
Several authors used project management frameworks, mainly PMBOK, to predefine themes (Themistocleus and Wearne, 2000; Zobel and Wearne, 2000; Koopenborg and Opfer, 2002; Padalkar and Gopinath, 2016). The frameworks present knowledge areas, processes or competences that are important in project management practice. Their comparison with trends can show which of them were studied in recent years and reveal gaps. The subject of the comparison included trends, not the individual papers, which impacted on the results. Two frameworks were compared: PMBOK 6 (2017) and IPMA Individual Competence Baseline 4 – ICB4 (2018).

The PMBOK 6 contains 49 processes in 10 knowledge areas. The trends were matched to six knowledge areas: project schedule management, project resource management, project communications management, project risk management, project procurement management and project stakeholder management. Overall, 14 trends were related to PMBOK processes and knowledge areas. Additional trends were related to PMI standards on programmes and portfolios. No trends corresponding with project integration, scope and quality management were found.

The ICB4 contains three competence areas with 29 competence elements: perspective – five elements, people – ten elements, practice – 14 elements. The comparison shows that 23 of the competence elements were related to the trends discovered in this study. Competences related to compliance, conflicts, personal integrity, resourcefulness, scope, selection and balance were not assigned to any trends. Overall, 31 trends were related to ICB4 competence elements.

The comparison shows that only part of the research topics is related to subjects covered by project management frameworks. Many studies are related to theoretical and practical issues which were not yet included in the frameworks. Some of them take a different perspective and evaluate projects as a whole, such as megaprojects and disaster recovery projects. There are, however, some elements in the frameworks that have been less frequently studied by researchers in recent years. They can become interesting topics for future studies.

5.4 Insights for future research
The information about trends dynamics presented in Table 4 offers insights for future research. We identified five trends that are more popular than others, and in the same time their popularity grows. They are related to EVM, portfolios, stakeholders, governance and megaprojects. In the last studied period, each of these topics was discussed, on average, in more than four papers a year in the studied journals. They are also often theme of presentations at international conferences. We expect that in coming years, the popularity of these trends will at least continue.

Three other trends, namely, project complexity, role of project owner and building information modelling, also grow, however, at slower pace. The first one is associated with a wider problem of complexity in management, the second with agile approach and the last with specific information systems. The reason for slower growth of those trends can result from less close connection with the field of project management.

The last three growing trends – disaster recovery, sustainable project management and agile portfolio management – were discovered in recent years. They were not found in presented previous systematic reviews. The disaster recovery trend is associated with a
growing perception of the need to respond appropriately to the effects of disasters. The sustainable project management trend indicates the growing role of sustainability in infrastructure and development projects. The agile portfolio management trend is associated with barriers that agile approach faces at the strategic level of the corporations. The studies included in those trends are an researchers’ attempt to answer contemporary problems related to project management.

## 6. Conclusions

Two objectives of the study were defined: the demonstration of the suitability of proposed methodology and the discovery of research trends in the field of project management. The analysis of previous studies on the subject revealed their limitations. They were discussed in the literature review section of this paper. To overcome those limitations, this study was based on the analysis of full texts of articles. It was found that the proposed methodology was more resistant to bias; delivered more detailed results; trends were discovered, not predefined; and it allowed the indication of the time frame of the trends. It is also possible to scale the level of detail using only one parameter – minimum size of the cluster. Therefore, the quality of the results is higher than in the case of previous studies. This has been achieved by using techniques well described in literature (TF-IDF), modern (HDBSCAN), as well as a proposed new approach to the identification of trends.

One of the more significant findings to emerge from this study is that commonly used methods of systematic literature review based on the analysis of citations, keywords and abstracts discard a significant number of scientific publications. These findings have significant implications for the understanding of how authors should look for interesting research directions and publications that correspond to their research area.

This study establishes a quantitative framework for detecting trends in scientific publications, allowing a significant reduction in the time and effort necessary for performing a manual systematic review of the literature. The insights gained from this study may also be of assistance to scientists and practitioners interested in developing their understanding of which directions in PM is the research growing in popularity. The findings of this study have a number of practical implications, especially for researchers struggling to read all newly published articles in their field of interest. They can adopt approach presented in this paper to develop their research toolkit used for information retrieval, scientometric analysis, structured reviews or research field scanning.

There are some limitations in the study: the number of papers and the minimum cluster size. The scope of this study was limited to seven journals and articles since 2000. It should be emphasised that the subject of project management is also discussed in other management journals, books, manuals, training materials and so forth. This approach, however, let the authors bypass the problems related to searching by keywords discussed in the introduction. Despite this limitation, the sample size in this research is significantly larger than in most of other studies. Based on extensive search in scientific databases, it was assumed that the sample was large enough to allow conclusions to be drawn on its basis. The minimum cluster size in HDBSCAN had to be defined by the authors. Its reduction could lead to the discovery of less popular trends, but, at the same time, it would make their analysis and presentation impossible due to their large number.

Another potential limitation of this study is the use of unsupervised machine learning methods for detecting patterns of articles presented in literature. This causes that often topics that seem important from the point of view of researchers active in the field are omitted (e.g. innovation management, narratives and storytelling). The lack of these topics in the research is not due to the fact that authors diminish their significance, but it is a direct result of adopting such study method.
This paper contributes to a growing pool of systematic reviews published in PM literature offering a new understanding that should help to improve predictions of the impact of particular identified trends in publications on the further development of this discipline. It also develops a methodology used in the preparation of systematic literature reviews.

**Highlights**

1. Limitations of systematic literature reviews were identified.
2. Forty-three trends in project management journals were discovered in 2000–2019.
3. Dynamics of currently active trends were presented.
4. Discussion includes predictions for discovered trends.
5. A method of automatic trends discovery, instead of their predefinition, was proposed.
6. The study shows the usefulness of text mining in systematic reviews.

**Note**

1. Trend in this paper is understood as subject studied by multiple researchers for several consecutive years. For each trend, it is possible to identify its time span, strength (average number of papers per year) and dynamics. In some cases, it is possible to identify more detailed topics within the trend.

**References**


**Corresponding author**

Sławomir Wawak can be contacted at: wawaks@uek.krakow.pl