In this paper, we carry out a study about the main themes treated by the *International Journal of Information Technology & Decision Making* during its first 10 years (2002–2011). The themes are detected, quantified and visualized using an approach that combines performance analysis and science mapping. Bibliometric maps based on co-word analysis will help us to visualize the division of the journal into several subfields and their relationships, providing interesting insight into the main topics being discussed in the journal in these years. In addition, the study will show the most productive themes (according to published papers) and the most impacting ones (according to received citations).

**Keywords:** Science mapping; automatic topic detection.

1. Introduction

On March 2002 the *International Journal of Information Technology & Decision Making* (IJITDM) launched its first issue. More than 400 papers have been published. In general, these published papers are research works that show how...
information technology improves decision techniques as well as how the development of decision-making tools affects the information technology era.

A lot of effort has been done by everyone, including authors and co-authors, anonymous reviewers, editorial board and editorial office, in order to increase the quality of the journal. As a result, nowadays IJITDM is indexed in the most important journal ranking, the Journal Citation Reports (JCR) published by the Thompson Institute for Scientific Information (Thompson ISI), appearing in the JCR-2010 in top positions in different categories (Computer Science-Artificial Intelligence, Computer Science-Information Systems, Computer Science-Interdisciplinary Applications, and Operations Research & Management Science) with a high impact factor (3.139).

In the years 2009 and 2010 (at the beginning of each year), the editor in-chief of the IJITDM has published papers describing briefly the content of the main papers published during the year before. These papers\textsuperscript{58,59} reported the main research trends of the papers published in 2008 and 2009, respectively. These two works, although are very much oriented and detailed, are very much limited due to the size of the period examined, just one year in each case.

In this paper, we try to complete those previous works\textsuperscript{58,59} in order to offer an expeditious perspective of a bigger period, the first decade (2002–2011), of the IJITDM by identifying themes, predicting emerging trends, and determining relationships with other themes/fields. The analysis is done using co-word as a content technical analysis. Co-word has been proved as a powerful technique for describing the research trends in other research fields\textsuperscript{7,13,18,36,40,41,44,69,78}.

The paper is structured as follows. Section 1 describes the methodological framework and the data used. Section 2 visualizes and describes the main themes treated by IJITDM. Section 3 analyzes and determines the influencing themes of the IJITDM, and predicts future research trends. Finally, some conclusions are drawn.

2. Method: Bibliometric Tool and Data

In bibliometrics, there are two main procedures to explore a research field: performance analysis and science mapping.\textsuperscript{68} Performance analysis aims at evaluating groups of scientific actors (countries, universities, departments, researchers) and the impact of their activity on the basis of bibliographic data. Science mapping aims at displaying the structural and dynamic aspects of scientific research, delimiting a research field, and quantifying and visualizing the detected subfields by means of co-word analysis\textsuperscript{14} or documents co-citation analysis.\textsuperscript{61}

2.1. Bibliometric analysis tool

In Ref. 16, we define a bibliometric approach that combines both performance analysis tools and science mapping tools to analyze a research field, to detect and visualize its conceptual subdomains (particular topics/themes or general thematic areas) and its thematic evolution. Co-word analysis is used in a longitudinal
framework which allows us to analyze and track the evolution of a research field. Additionally, we develop a performance analysis of a specific theme or thematic area using different basic bibliometric indicators.

The main steps to develop a bibliometric analysis based on co-word are the following:

1. **To detect the research themes.** To do so, first we have to compute the co-occurrence matrix by assuming that the co-occurrence frequency of two keywords is extracted from the corpus of documents by counting the number of documents in which the two keywords appear together. Second, we have to compute the equivalence index among keywords, called \( e_{ij} = \frac{c_{ij}^2}{c_i c_j} \), where \( c_{ij} \) is the number of documents in which two keywords \( i \) and \( j \) co-occur, and \( c_i \) and \( c_j \) represent the number of documents in which each one appears. At the end of this phase, we cluster keywords to topics/themes by using the simple centers algorithm, which automatically returns labeled clusters, so a post-process to label the clusters is not needed. Through this process of clustering we locate keyword networks that are strongly linked to each other and which correspond to centers of interest or to research problems that are the object of significant interest among researchers. Each cluster network or thematic network is defined as an interconnected network of keywords as it is shown in Fig. 1(a). In it, several keywords are interconnected, where the volume of the spheres is proportional to the number of documents corresponding to each keyword, the thickness of the link between two spheres \( i \) and \( j \) is proportional to the equivalence index \( e_{ij} \).

Together with the whole network of interconnected themes and keywords a second network is built, based on the documents linked to each thematic network. In this second network, documents with keywords associated to any detected thematic network are linked to it. Given a thematic network, a document \( d \) is associated to it if

![Fig. 1. A thematic network (a), and the strategic diagram (b).](image-url)
d has at least two keywords presented in the thematic network. These second networks of documents are used in the third phase.

(2) *To build strategic diagrams.* In the clustering process we obtain a set of interconnected networks or themes. Then, in this context each keyword network or theme can be characterized by two parameters (Centrality and Density)\(^{13}\):

- **Centrality:** It measures the degree of interaction of a network with other networks and it can be defined as: 
  \[ c = 10 \times \sum e_{kh}, \]
  with \( k \) a keyword belonging to the theme and \( h \) a keyword belonging to other themes. Centrality measures the strength of external ties to other themes. We can understand this value as a measure of the importance of a theme in the development of the entire research field analyzed.

- **Density:** It measures the internal strength of the network and it can be defined as:
  \[ d = 100 \times \frac{c_{ij}}{w}, \]
  with \( i \) and \( j \) keywords belonging to the theme and \( w \) the number of keywords in the theme. Density measures the strength of internal ties among all keywords describing the research theme. This value can be understood as a measure of the theme’s development.

A Strategic Diagram is a two-dimensional space built by plotting themes according to their centrality and density rank values. As an example, in Fig. 1(b) a strategic diagram is presented. Thus, with both parameters a research field can be understood to be a set of research themes, mapped in a two-dimensional space and classified into four groups\(^{13}\):

- Themes in the upper-right quadrant are both well developed and important for the structuring of a research field. They are known as the *motor-themes* of the specialty given that they present strong centrality and high density.
- Themes in the upper-left quadrant have well-developed internal ties but unimportant external ties and so are of only marginal importance for the field. These themes are very specialized and peripheral in character.
- Themes in the lower-left quadrant are both weakly developed and marginal. The themes of this quadrant have low density and low centrality, mainly representing either emerging or disappearing themes.
- Themes in the lower-right quadrant are important for a research field but are not developed. So, these quadrant groups are transversal and general, basic themes.

Furthermore, the strategic diagrams can be enriched by adding a third dimension in order to show more information. So, for example, the themes can be represented as a sphere, its volume being proportional to different quantitative (or qualitative) data, such as, for example, the number of documents associated with the theme or the number of citations received by the documents associated with the theme.

(3) *To carry out a performance analysis.* In this phase, we can measure (quantitatively and qualitatively) the relative contribution of themes and thematic
areas to the whole research field, detecting the most prominent, productive, and highest-impact subfields. To do so, we use the following bibliometric indicators applied to the different detected themes: the number of published documents, the number of received citations, and the h-index.\textsuperscript{3,10,27}

We should point out that the co-word analysis was carried out with the software CoPalRed.\textsuperscript{17,19,41} It is based on the simple center algorithm to detect the themes through different subperiods of years. The plotting of the themes in the strategic diagram, the drawing of the thematic networks were done with specific ad-hoc software.

\textbf{2.2. Data set}

We downloaded them from ISI Web of Science\textsuperscript{a} (ISIWoS) for the years 2004 to 2011 (inclusive). For the two first years (2002 and 2003) data was manually downloaded from the IJITDM’s website.\textsuperscript{b} The data set includes 415 papers including articles, letters, proceeding papers and reviews. We have not considered editorial material. In Figs. 2 and 3 the distribution of published documents and citations per year are shown.

As mentioned above, the keywords of the documents are used to develop our analysis. Due to the fact that the majority of the data have been downloaded from the ISIWoS, the author-provided keywords and the Keywords Plus of the documents are jointly used. A normalization process is carried out prior to this over the keywords, where the plural and singular forms of the keywords are joined. The acronyms are also joined with the respective keywords. Finally, we have obtained 330 valid documents published in IJITDM from 2002 to 2011.

\textsuperscript{a}http://scientific.thomson.com/products/wos/.
\textsuperscript{b}http://www.worldscinet.com/ijitdm/.
keywords for our study. We consider valid keywords to be those that present both frequency and co-occurrence greater than 2.

In this study the citations of the documents are also used. We have considered for each paper the citations received until November 15th 2011 (the data were downloaded on that date). The citations that we take into account proceed from the ISIWoS (for the papers published from 2004) and Scholar Googlec (for the papers manually downloaded from the journal’s website during the years 2002 and 2003).

3. Visualizing the Main Themes Treated by IJITDM

Using the approach described above, 17 themes were automatically detected by CoPalRed, these are: AGENTS, ANALYTICAL-HIERARCHY-PROCESS, CONSENSUS, DATA-ENVELOPMENT-ANALYSIS, DATA-MINING, DECISION-SUPPORT-SYSTEM, GROUP-DECISION-MAKING, INTERNET-QUALITY-OF-SERVICE, INTERPRETIVE-STRUCTURAL-MODELING, MANAGEMENT, OPTIMIZATION, PERFORMANCE-EVALUATION, SHAPLEY-VALUE, SOFTWARE-DEFECT-PREDICTION, TECHNOLOGY-ACCEPTANCE, TIME-SERIES, UNCERTAINTY.

In order to analyze the most highlighted themes treated by IJITDM for the studied period (2002–2011), two kinds of strategic diagrams were built using CoPalRed. In the first one, the volume of the spheres is proportional to the number of documents associated with each theme; and in the second one, it is proportional to the number of citations received by the documents of each theme. These strategic diagrams are shown in Figs. 4 and 5, respectively, and whose main aspects are:

- According to the number of documents, the themes ANALYTICAL-HIERARCHY-PROCESS (19 papers), DATA-ENVELOPMENT-ANALYSIS (17), MANAGEMENT

\[\text{http://scholar.google.com.}\]
(14), DATA-MINING (14) and PERFORMANCE-EVALUATION (11) are the five more productive one.

- Taking into account the citations, the themes DATA-MINING (with 156 citations), ANALYTICAL-HIERARCHY-PROCESS (106), CONSENSUS (93), DATA-ENVELOPMENT-ANALYSIS (75), MANAGEMENT (48) and TIME-SERIES (48) are the six most cited themes. On the other hand, the documents associated to the theme INTERPRETIVE-STRUCTURAL-MODELING has not cited.

- On the other hand, analyzing the ratio documents/citations, the themes CONSENSUS (with 18.60 citations per paper), GROUP-DECISION-MAKING (13.33), DATA-MINING (11.14) and TIME-SERIES (9.60) have been the most impacting themes.

- With respect to the h-index, the themes ANALYTICAL-HIERARCHY-PROCESS (with 7), DATA-ENVELOPMENT-ANALYSIS (6), CONSENSUS (5) are the three one with higher value. We have to remark that the theme CONSENSUS has not cited.
has obtained the highest h-index value that it is possible with its five associated papers.

- Analyzing their strategic location in the strategic diagrams, the themes ANALYTICAL-HIERARCHY-PROCESS, PERFORMANCE-EVALUATION, MANAGEMENT, DECISION-SUPPORT-SYSTEM AND UNCERTAINTY are supposed transversal and basic-themes; whereas CONSENSUS, GROUP-DECISION-MAKING and SOFTWARE-DEFECT-PREDICTION are supposed motor-themes, given that they present strong centrality and high density, i.e., they are understood as very much developed themes with also high importance/weight in the development of other themes and of the whole journal in the analyzed period.

- Moreover, themes like: INTERNET-QUALITY-OF-SERVICE, INTERPRETIVE-STRUCTURAL-MODELING, SHAPLEY-VALUE, TECHNOLOGY-ACCEPTANCE and TIME-SERIES are considered as very specialized and low interconnected one, being INTERNET-QUALITY-OF-SERVICE and INTERPRETIVE-STRUCTURAL-MODELING the most isolated ones.
Due to their strategic location, the themes DATA-ENVELOPMENT-ANALYSIS, AGENTS and DATA-MINING have to be considered as new and emerging themes.

Attending to the kind of themes, all the detected themes, except SOFTWARE-DEFECT-PREDICTION, INTERNET-QUALITY-OF-SERVICE, PERFORMANCE-EVALUATION and MANAGEMENT are considered as “methodological” or “theoretic” themes rather than “practical” or “applied” one.

Performance data, including quantity of associated documents, sum and average of citations and h-index, for these detected themes are summarized in Table 1.

### 3.1. Visualizing the most highlighting themes

In this section, we draw the most highlighting themes according to the number of associated documents, citations and locations at the strategic diagrams.

- The theme CONSENSUS, whose thematic-network is shown in Fig. 6, is related with the keywords (or topics): aggregation, assessments, labels, owa-operators, consistency and operators.

The four most cited paper associated to this theme are:


<table>
<thead>
<tr>
<th>Theme</th>
<th>Documents</th>
<th>h-Index</th>
<th>Average Citations</th>
<th>Total Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGENTS</td>
<td>6</td>
<td>2</td>
<td>2.17</td>
<td>13</td>
</tr>
<tr>
<td>ANALYTICAL-HIERARCHY-PROCESS</td>
<td>19</td>
<td>7</td>
<td>5.58</td>
<td>106</td>
</tr>
<tr>
<td>CONSENSUS</td>
<td>5</td>
<td>5</td>
<td>18.60</td>
<td>93</td>
</tr>
<tr>
<td>DATA-ENVELOPMENT-ANALYSIS</td>
<td>17</td>
<td>6</td>
<td>4.41</td>
<td>75</td>
</tr>
<tr>
<td>DATA-MINING</td>
<td>14</td>
<td>4</td>
<td>11.14</td>
<td>156</td>
</tr>
<tr>
<td>DECISION-SUPPORT-SYSTEM</td>
<td>8</td>
<td>3</td>
<td>3.50</td>
<td>28</td>
</tr>
<tr>
<td>GROUP-DECISION-MAKING</td>
<td>3</td>
<td>2</td>
<td>13.33</td>
<td>40</td>
</tr>
<tr>
<td>INTERNET-QUALITY-OF-SERVICE</td>
<td>2</td>
<td>2</td>
<td>4.50</td>
<td>9</td>
</tr>
<tr>
<td>INTERPRETIVE-STRUCTURAL-MODELING</td>
<td>2</td>
<td>0</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>MANAGEMENT</td>
<td>14</td>
<td>4</td>
<td>3.43</td>
<td>48</td>
</tr>
<tr>
<td>OPTIMIZATION</td>
<td>8</td>
<td>3</td>
<td>5.88</td>
<td>47</td>
</tr>
<tr>
<td>PERFORMANCE-EVALUATION</td>
<td>11</td>
<td>4</td>
<td>3.00</td>
<td>33</td>
</tr>
<tr>
<td>SHAPLEY-VALUE</td>
<td>3</td>
<td>2</td>
<td>4.33</td>
<td>13</td>
</tr>
<tr>
<td>SOFTWARE-DEFECT-PREDICTION</td>
<td>6</td>
<td>3</td>
<td>3.50</td>
<td>21</td>
</tr>
<tr>
<td>TECHNOLOGY-ACCEPTANCE</td>
<td>3</td>
<td>1</td>
<td>1.67</td>
<td>5</td>
</tr>
<tr>
<td>TIME-SERIES</td>
<td>5</td>
<td>4</td>
<td>9.60</td>
<td>48</td>
</tr>
<tr>
<td>UNCERTAINTY</td>
<td>7</td>
<td>3</td>
<td>3.43</td>
<td>24</td>
</tr>
</tbody>
</table>

Xu, ZS, AN APPROACH TO PURE LINGUISTIC MULTIPLE ATTRIBUTE DECISION MAKING UNDER UNCERTAINTY. IJITDM 4:2 (2005) 197–206. Times cited:15

The thematic-network of the theme ANALYTICAL-HIERARCHY-PROCESS is shown in Fig. 7. This theme was mainly related with the topics: preference, decision and multicriteria-decision-making.

The three most cited papers:


DATA-ENVELOPMENT-ANALYSIS was strongly related with the topics: benchmarks, efficiency and banks. The network of interconnected keywords associated to this theme is shown in Fig. 8.

The three most cited papers:


On the other hand, DATA-MINING was associated with clustering, databases, knowledge-management, multi-aspect-analysis, discovery and fuzzy-linear-programming. The associated thematic network is drawn in Fig. 9.

The three most cited papers:


The theme GROUP-DECISION-MAKING was connected (see Fig. 10) with words, fuzzy-linguistic-modeling and preference-relation.
The two most cited papers:


- MANAGEMENT was a practical theme which was related with the topics: business-strategies, project-management, simulation, systems and sites and multicriteria-decision-support-systems. Its associated thematic network is drawn in Fig. 11.

The three most cited papers:

The theme PERFORMANCE-EVALUATION was mainly related with the topics: trust, acceptance, ranking, impact and market. Its associated thematic network is drawn in Fig. 12.

The four most cited documents associated to PERFORMANCE-EVALUATION are:

• TIME-SERIES was related with fuzzy-regression, forecasting and rate-prediction as it can be observed in Fig. 13.

The three most cited documents are:


In this section, we introduce a simple analysis that could be done in order to detect the most beneficial themes for the journal. For this analysis, we only use the last two years of the journal, i.e., 2010 and 2011.
In Table 2 a list of the themes associated to the documents published during 2010 and 2011 is shown. This list includes quantitative and qualitative measures as the number of documents (\#Doc) associated to each theme, references of each document, the citations and citation average (Average). The information included in this table allows us to distinguish to type of themes, hot-themes and cold-themes. For us, a hot-theme is that one with a citation average higher (or equal) to 1.00, i.e., themes with at least one cite per each associated document. These hot-themes are those themes giving positive citation to the journal, helping to increase the impact factor of the journal, whereas, the cold-themes are themes penalizing the journal, decreasing its impact factor. In this sense, themes as SOFTWARE-DEFECT-PREDICTION (with 4.00 cites per document), GROUP-DECISION-MAKING (2.25 cites per document), MANAGEMENT (1.75), ANALYTICAL-HIERARCHY-PROCESS (1.56) and PERFORMANCE-EVALUATION (1.43) are supposed, in general, very beneficial themes for the journal, clearly contributing to a high impact factor.

This \textit{a posteriori} classification could be used by the reader as a predictive estimation about what themes will be beneficial for the journal in next years (2012 or
This predictive analysis could be even done and used by the journal editorial board to help them to select the priority thematics for the journal, given preference to documents associated to presumably hot-themes. Anyway, this predictive analysis (and the associated editorial decisions) has to be carefully done otherwise it is also possible to find cold-themes with some particular highly cited...
Table 3. Most cited papers and their associated themes.

<table>
<thead>
<tr>
<th>Document (Reference)</th>
<th>Year</th>
<th>Citations</th>
<th>Associated Theme/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>2008</td>
<td>57</td>
<td>DATA-MINING</td>
</tr>
<tr>
<td>75</td>
<td>2006</td>
<td>53</td>
<td>DATA-MINING</td>
</tr>
<tr>
<td>11</td>
<td>2009</td>
<td>35</td>
<td>CONSENSUS // GROUP-DECISION-MAKING</td>
</tr>
<tr>
<td>26</td>
<td>2004</td>
<td>31</td>
<td>DATA-MINING</td>
</tr>
<tr>
<td>60</td>
<td>2005</td>
<td>29</td>
<td>OPTIMIZATION</td>
</tr>
<tr>
<td>73</td>
<td>2005</td>
<td>21</td>
<td>CONSENSUS</td>
</tr>
<tr>
<td>54</td>
<td>2002</td>
<td>16</td>
<td>TIME-SERIES</td>
</tr>
<tr>
<td>4</td>
<td>2009</td>
<td>15</td>
<td>CONSENSUS // UNCERTAINTY</td>
</tr>
<tr>
<td>74</td>
<td>2005</td>
<td>15</td>
<td>CONSENSUS</td>
</tr>
<tr>
<td>55</td>
<td>2005</td>
<td>15</td>
<td>ANALYTICAL-HIERARCHY-PROCESS</td>
</tr>
<tr>
<td>1</td>
<td>2006</td>
<td>14</td>
<td>ANALYTICAL-HIERARCHY-PROCESS //</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DATA-ENVELOPMENT-ANALYSIS</td>
</tr>
<tr>
<td>34</td>
<td>2005</td>
<td>14</td>
<td>TIME-SERIES</td>
</tr>
<tr>
<td>30</td>
<td>2007</td>
<td>13</td>
<td>DECISION-SUPPORT-SYSTEM //</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PERFORMANCE-EVALUATION //</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TIME-SERIES</td>
</tr>
<tr>
<td>32</td>
<td>2009</td>
<td>12</td>
<td>DATA-ENVELOPMENT-ANALYSIS</td>
</tr>
<tr>
<td>2</td>
<td>2008</td>
<td>11</td>
<td>ANALYTICAL-HIERARCHY-PROCESS</td>
</tr>
<tr>
<td>63</td>
<td>2009</td>
<td>11</td>
<td>MANAGEMENT</td>
</tr>
<tr>
<td>23</td>
<td>2007</td>
<td>10</td>
<td>ANALYTICAL-HIERARCHY-PROCESS</td>
</tr>
</tbody>
</table>

Fig. 14. The influence zone of IJITDM in 2010 and 2011.
documents, as it is shown in Table 3. In Table 3, the historical (considering papers published from 2002 to 2011) most cited papers are shown. We can see as relatively recent papers (2009) as Refs. 4 and 11 (associated to CONSENSUS) and Ref. 63 (associated to MANAGEMENT) are highly cited papers.

We think all this information should be jointly taken into account for determining the “best” themes for the IJITDM. So, we define the Influence-Zone as the part of the strategic diagram that groups motor and basic-themes and simultaneously hot-themes with themes with highly cited papers in the last three or four years before 2009, 2010 and 2011 in our case. This influence zone is drawn in Fig. 14.

So, we think that future (mainly 2012 or 2013) papers associated to the influence themes: SOFTWARE-DEFECT-PREDICTION, GROUP-DECISION-MAKING, CONSENSUS, MANAGEMENT, ANALYTICAL-HIERARCHY-PROCESS, PERFORMANCE-EVALUATION and DECISION-SUPPORT-SYSTEM will be very beneficial papers, providing quantity and specially quality and cites to the journal.

5. Conclusion

The purpose of this paper has been to offer an expeditious perspective of the first decade (2002–2011) of the IJITDM by identifying themes, predicting emerging trends, and determining relationships with other themes/fields.

Our analysis has to allow to detect the influence themes of the journal, i.e., those themes with a much relevant role in future years for the journal.

All this has been done using co-word as content analysis technique. However, this task is not free of difficulties due to the biases involved in an analysis of this kind. The main one is that the analysis focuses on priority themes (just those automatically detected by CoPalRed) and will inevitably exclude those that are of only marginal importance. Nonetheless, the analysis serves to legitimize discussions about general trends of the journal.

Since the analysis is constrained by factors such as sample size and the period examined, among others, its “applicability-generalizability” must be further reviewed and tested in the future, preferably at regular intervals.

Finally, experts and novices could use these data, results and maps to understand the current state of the art with regard to IJITDM and predict where future research will lead.

Acknowledgments

This work was partially supported by the Projects of Ministry of Education and Science TIN2010-17876, Spain, Excellence Andalusian Projects TIC-5299 and 5991, Spain, National Natural Science Foundation of China (Grant Nos. 70901015, 70621001, 70921061, 71110107026), China, the CAS/SAFEA International Partnership Program for Creative Research Teams, China and National Science
Foundation, EPSCoR Program, USA, the Fundamental Research Funds for the Central Universities and Program for New Century Excellent Talents in University (NCET-10-0293).

References


