
Evolutionary overview of the scientific production on Information Literacy indexed in Scopus from 1998 to 2014

Ernesto Galbán Rodríguez¹ & Yohannis Martí Lahera²

¹Departamento de Inteligencia Empresarial, Centro de Ingeniería Genética y Biotecnología, Cuba, ernesto.galban@cigb.edu.cu
²Universidad de La Habana (UH), Facultad de Comunicación, Cuba, ernesto.galban@cigb.edu.cu

Resumen:

INTRODUCCIÓN: La Alfabetización Informacional es un área esencial para la Ciencia de la Información. Su producción científica indizada en Scopus ha sido caracterizada anteriormente, aunque su carácter transdisciplinar requiere de un análisis integrador.


METODOLOGÍA: Se recuperaron los registros que incluyeron la estrategia ‘informationliteracy’ en los campos de resumen, título y palabras clave, y analizaron según indicadores bibliométricos básicos, junto al análisis de agrupamiento con VOS viewer 1.5.7 y mapeos de superposición y estratégicos con SciMAT versión 1.1.03.

RESULTADOS: Se recuperaron 3533 registros, 51 % con la estrategia de búsqueda como palabra clave, grupo que se estableció como núcleo temático. Las doce revistas más productivas correspondieron a las áreas de bibliotecas y sistemas de referencia, medios digitales y comunicación. La producción científica en 2005 duplicó la del año anterior y en
el periodo 2005-2009 fue más del triple del periodo 1998-2004. Ambos periodos se compararon con el de 2010 a 2014, identificándose seis clústeres temáticos, por orden descendente: Tecnología de la información; Bibliotecas académicas; Internet y diseminación de la información; Educación y enfermería; Tecnologías y estándares de la comunicación; y Educación y sistemas bibliotecarios en medicina. Las tecnologías de la información, las encuestas y las bibliotecas emergieron como temas motores recientes. 

CONCLUSIONES: Se muestra un nuevo enfoque evolutivo de la alfabetización informacional como campo temático, que complementa los estudios previos sobre alfabetización informacional y alfabetización en salud.

Palabras clave: alfabetización informacional; bibliometría, producción científica; mapeo de clústeres; mapeo estratégico

Abstract:

INTRODUCTION: Information literacy is a key area of research in Information Science. Previous research has addressed the output of Information Literacy in Scopus. However, this transdisciplinary field requires an integrative view, due to its transversal nature. 

PURPOSE: To analyze the evolution of the scientific production indexed in Scopus on Information Literacy from 1998 to 2014. 

METHODOLOGY: The database was searched for ‘information literacy’ in the title, abstract and keyword fields. Records were analyzed for basic bibliometric indicators, together with field characterization by clustering with VOS viewer 1.5.7 and overlapping and strategic mapping with SciMAT version 1.1.03. 

RESULTS: A total of 3533 were retrieved, 51 % having the search string ‘information literacy’ as keyword and regarded as the main cluster. The 2005 production doubled that of the previous year and that of the 2005-2009 more than tripled that of the previous subperiod (1998-2004). The twelve most productive journals corresponded to library and reference
systems, digital media and communications. Three discrete subperiods were established for evolutionary analyses: 1998-2004, 2005-2009 and 2010-2014. Six keywords clusters were identified in productivity descending order: Information technology; Academic libraries; Internet and information dissemination; Education and nursing; Communication technologies and standards; and Medical education and library systems. Moreover, information technologies, surveys and libraries emerged as recent motor topics.

CONCLUSIONS: A new and evolutionary outlook in the evolution of Information Literacy as a field is provided. Moreover, it complements prior studies on information literacy and health literacy subdomains.

**Keywords:** information literacy; bibliometrics, scientific production; cluster mapping; strategic mapping

---

**Introduction**

Information literacy is a key area of research in Information Science. Stated as the process by which information needs are met by using identified specific skills, knowledge and attitudes (Behrens, 1995; Pinto, 2010), has evolved into a set of ideas involving relationship to metacognition, self-reflection, and successful participation in collaborative work (ACRL, 2015), to synthesise and create information and data to create new knowledge (SCONUL, 2011).

Previous research have addressed the output of Information Literacy in Scopus, such as: Majid et al. (Majid, 2015), in Scopus from 2003 to 2013 (data from November 2013); Aharony et al. (Aharony, 2010), in WoS from 1999 to 2009. Mostly, basic indicators have been analyzed, including: Ndoc, Ncit, top cited journals, top cited papers, top authors, top journals, affiliations, subject areas, output country distribution and publication languages. However, this transdisciplinary discipline requires an integrative view, due to its transversal
nature. In fact, a significant contribution comes from its application in areas other than social sciences, which have been excluded to a certain extent while considering just Library and Information Science journals. Moreover, since core concepts started in the 1970’s and subsequently developed, significant contributions could be excluded while selecting a narrow time window. At the same time, core and peripheral productive areas interacts as in other research areas, a differentiation not always done when analyzing the output data.

Therefore, in this work, we analyzed the scientific production indexed in Scopus on Information Literacy from 1998 to 2014, attending to keywords distribution, core and peripheral output in the field, author number per paper, top authors, scientific output and citation, term co-occurrence analysis followed by clustering and mapping techniques, top journals, top cited documents in the core output.

**Methodology**

The Scopus database was searched for ‘information literacy’ in the title, abstract or keyword fields (database and author’s keywords). Records were downloaded into the EndNote X7 software, normalized and the respective fields exported with the subject bibliography tools for further simple frequency analyses. Microsoft excel was used for histogram representations.

The representativeness of the search strategy was investigated by a Zipf’s frequency of keywords among the records retrieved, to establish its centrality within retrieved results. The productive and cited journals were established by comparison attending to number of articles (Ndoc), cited documents (Ncit), percentage of citation in the domain (%), overall citation, average citation per article, H index and top citation.

The highest citation documents within the core group of documents, established as those including ‘information literacy’ as keyword, were calculated by using the SciMAT software v. 1.1.3, and results were compared to those of the citation of Zipf’s keyword distribution
among records. Top cited authors were further identified, together with the frequency of collaboration, following normalization and authority control.

An overall look at the field’s structuring by keyword analysis was made using the VOSviewer software, version 1.5.7. Keywords were data mined with a minimal term frequency of 3, providing 529 terms.

For evolutionary evaluations, retrieved records were loaded into the SciMAT software, and periods 1998-2004, 2005-2009 and 2010-2014 were set for comparison, based on authors’, indexing and complemented with automatically extracted keywords from titles and abstracts.

The generated overlapping maps, evolution maps, strategic maps (knowledge maps) and cluster networks were analyzed in that order for consistent integration of trends and ideas. The analysis was complemented with review of the represented literature.

**Results and discussion**

The search strategy used, ‘information literacy’ was set due to its specificity, since both words are ubiquitous and inherent to every human activity, regarding its widest application, and their social relevance. The origin of the term and its representativeness was analyzed in the 3533 (Figure 1) records retrieved by Zipf’s distribution, for keywords with a minimal frequency of 50 (Figure 2). The 1815 records including this keyword were regarded as the ‘core’ cluster, with the other regarded as ‘peripheral’. Among them, 241 documents had the listed keywords as the first keyword, for a 13.3 % of all the documents having ‘information literacy’ as secondary keyword in all it’s the identified variants.
Figure 1. Scientific production on Information Literacy, indexed in Scopus, 1998-2014. Data were retrieved in November 17th, 2014. The period was chosen two years after the start of trustable citation statistics at the database.

Figure 2. Zipf's distribution of keywords in the scientific production on Information Literacy, indexed in Scopus, 1998-2014. Data were retrieved in November 17th, 2014. Keywords with a minimal frequency of 50 are shown.
Moreover, the occurrence of secondary terms reinforces the view of information literacy as transversal and supportive for the entanglement of human skills, learning and training in highly specialized areas, particularly in technical and medical sciences.

According to the combined trends of productivity, cited documents and total citation, the period was divided in three discrete subperiods for evolutionary analyses: 1998-2004, 2005-2009 and 2010-2014.

The scientific production steadily increased, with 333 documents indexed from 1998 to 2004. The 2005 production doubled that of the previous year and that of the 2005-2009 more than tripled that of the previous subperiod. The scientific production from 2010 to 2014 subperiod is as much as that of the two previous subperiods (Figure 1).

The twelve most productive journals are related to the field of library and reference systems, digital media and communications (Table 1).

<table>
<thead>
<tr>
<th>No</th>
<th>Journal</th>
<th>Ndoc</th>
<th>Ncit</th>
<th>% Cit</th>
<th>Cites</th>
<th>Average cites/doc</th>
<th>H Index</th>
<th>Top citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reference Services Review</td>
<td>149</td>
<td>125</td>
<td>83.9</td>
<td>844</td>
<td>5.7</td>
<td>14</td>
<td>97</td>
</tr>
<tr>
<td>2</td>
<td>Journal of Academic Librarianship</td>
<td>119</td>
<td>81</td>
<td>68.1</td>
<td>1296</td>
<td>10.9</td>
<td>20</td>
<td>93</td>
</tr>
<tr>
<td>3</td>
<td>College and Undergraduate Libraries</td>
<td>86</td>
<td>57</td>
<td>66.3</td>
<td>224</td>
<td>2.6</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>Communications in Information Literacy</td>
<td>75</td>
<td>33</td>
<td>44.0</td>
<td>75</td>
<td>1.0</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>Portal</td>
<td>60</td>
<td>54</td>
<td>90.0</td>
<td>607</td>
<td>0.9</td>
<td>14</td>
<td>49</td>
</tr>
<tr>
<td>6</td>
<td>College and Research Libraries</td>
<td>53</td>
<td>47</td>
<td>88.7</td>
<td>862</td>
<td>16.3</td>
<td>18</td>
<td>79</td>
</tr>
<tr>
<td>7</td>
<td>College and Research Libraries News</td>
<td>50</td>
<td>29</td>
<td>58.0</td>
<td>91</td>
<td>1.8</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>Research Strategies</td>
<td>49</td>
<td>48</td>
<td>98.0</td>
<td>411</td>
<td>8.4</td>
<td>11</td>
<td>32</td>
</tr>
<tr>
<td>9</td>
<td>Journal of Library Administration</td>
<td>48</td>
<td>46</td>
<td>95.8</td>
<td>206</td>
<td>4.3</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>Library Review</td>
<td>47</td>
<td>40</td>
<td>85.1</td>
<td>237</td>
<td>5.1</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>11</td>
<td>Public Services Quarterly</td>
<td>47</td>
<td>30</td>
<td>69.8</td>
<td>93</td>
<td>2.0</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>New Library World</td>
<td>46</td>
<td>29</td>
<td>63.0</td>
<td>156</td>
<td>3.4</td>
<td>7</td>
<td>25</td>
</tr>
</tbody>
</table>

* Green, yellow and red cells indicate the first three positions for each indicator in descending order.

Table 1. Most productive journals identified in the scientific production on Information Literacy in Scopus, 1998-2014*
According to the number of articles (Ndoc), cited documents (Ncit), percentage of citation in the domain (%), overall citation, average citation per article, H index and top citation, the most relevant journals were in descending order due to the coincidence of these parameters: the *Journal of Academic Librarianship*, *College and Research Libraries*, *Portal, Research Strategies* and *References Service Review*. In fact, the top journal published two of the identified most cited articles (positions 7th and 8th).

In the case of authorship trends, authors contributing with more than 10 documents were identified. Noteworthy, the overwhelming majority (73.5 %) of the scientific production in the domain is generated by single authors, with 1471 out of the 3489 documents bearing author lists (42.2 %), or by two-authors groups (31.4 %). The other 25.15 % of contributions were produced by 3- to 6-author groups.

Together with terms overlapping maps, the periods were shown to progressively integrate relevant terms, doubling every 5 years, a trend similar but for the first to the second period transition for the case of journals which triplicated, maybe influenced by the increased indexing and the emergence of field descriptors (Figure 3).

Figure 3. Term overlapping map in the scientific production on Information Literacy, indexed in Scopus, 1998-2014. Data were retrieved in November 17th, 2014 and processed with SciMAT v. 1.1.03.
Moreover, the strategic analysis evidenced the diversification of the field in the last five years analyzed (2010-2014), the attention information literacy being fragmented between schools and information technologies (Figure 4).

Figure 4. Strategic map of the scientific production on Information Literacy, indexed in Scopus, 1998-2014. Data were retrieved in November 17th, 2014 and processed with SciMAT v. 1.1.03 for the three subperiods depicted. Progression to the right based on term network centrality and upwards on density. Parameters were: a minimal frequency of two words representing the occurrence of a term in at least two documents, a minimal network size of 2, co-occurrence normalization, clustering by the single centrality algorithm in the range 3-10, core and secondary document mapping, the sum of cites and H index as similarity measures, and the Jaccard’s index and the inclusion index as primary and secondary similarity measures.

Overall, the co-occurrence analysis of terms identified by latent semantics (natural language processing) in the title and abstract fields of the 3533 bibliographic records were grouped into six clusters (terms) in descending frequency (Figure 5): i) Information technology (145);
ii), Academic libraries (121); iii) Internet and information dissemination (108); iv) Education and nursing (83); v) Communication technologies and standards (19); and vi) Medical education and library systems (18). Clusters 1 to 3 were the most interconnected ones, maximal for the first two. Significantly, constructivism emerged as leading theory in cluster vi. This also included a diversification in skills and scenarios, with the emergence of technological and qualitative analytical instruments.

Figure 4. Cluster network of the scientific production on Information Literacy, indexed in Scopus, 1998-2014. Data were retrieved in November 17th, 2014 and processed with VOS viewer version 1.5.7 from the keywords bibliographic fields.

Conclusions

In summary, this study provides a new look onto the evolution of the Information literacy domain as a whole, beyond the classic bibliometric approach based on the output simple frequency of authors, journals and documents. It also complements the view of prior metric
studies differentiating health literacy from information literacy (Pinto, 2012; Pinto, 2013). Notably, the health literacy and information literacy views properly differentiate by clustering and strategic mapping with no need for prior literature segmentation.

**Bibliography**


SCONUL. 2011. SCONUL seven pillars of information literacy. [Online]. Available at: http://www.sconul.ac.uk/groups/information_literacy/papers/Seven_pillars2.pdf