Defining Lines of Research and Curricula with Via Correlation and Factorial Analysis Techniques

Ana Maria Luque Clavijo*[1]; Fernando Palop Marro+,
1 Corresponding author, mail analuque@usantotomas.edu.co
*Posgrados de Ingeniería electrónica y de telecomunicaciones, Universidad Santo Tomás, Colombia
+Departamento de Organización de Empresas, Universitat Politécnica de Valéncia, España

Abstract

Utilities and applications for text mining have led to developments in the methods of analysis for the identification of specific research topics. This paper features a methodological approach that combines correlation analysis and factor analysis.

The proposed methodology provides information for the specific analysis of text-based primary bases (figure 1), through the auto-correlation of keywords, which allows an easy building of relationships and therefore the definition of specific thematic areas of the primary base.

Whilst the contents of said areas are worked out via clustering processes(table 1), which precedes the factorial analysis of secondary databases generated for each specific subject area (figure 2-9). The methodology has been successfully applied in the management of graduate programs in engineering from the Universidad Santo Tomás, for updating curricula and research lines.

Keywords
Text mining, Text clustering, Term clumping, Engineering, Correlation analysis, Factor analysis

Table 1 Size of databases and coverage of thematic specific regions

| Databases	Records	Coverage	Figure
Methodology and standard for project management engineering (Primary DB)	3.862	-	1
Education (Secondary DB)	923	96%	2
Quantitative tools (Secondary DB)	750	91%	3
Impacts estimation (Secondary DB)	603	93%	4
Information sources (Secondary DB)	460	98%	5
Science management (Secondary)	266	100%	6
Quality assurance and quality control (Secondary DB)	1458	96%	7
Industrial management (Secondary DB)	1178	98%	8
Information systems (Secondary DB)	1024	97%	9
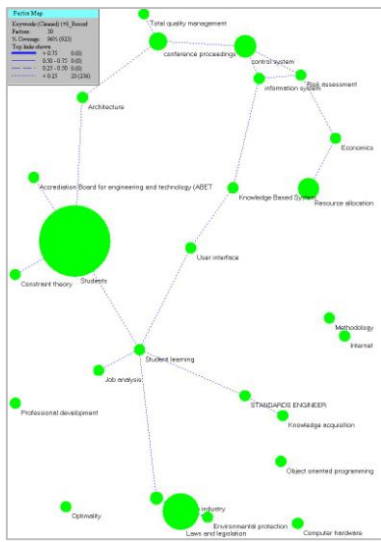
Figure 2 Education factor map
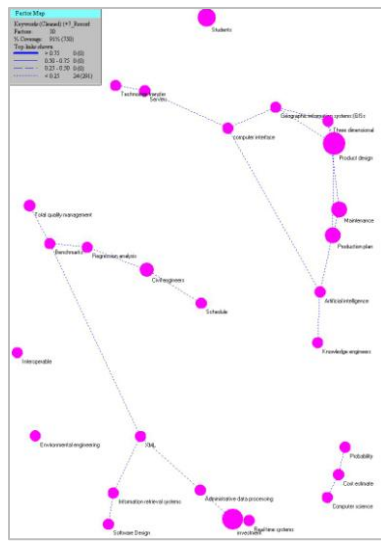
Figure 3 Quantitative tools factor map
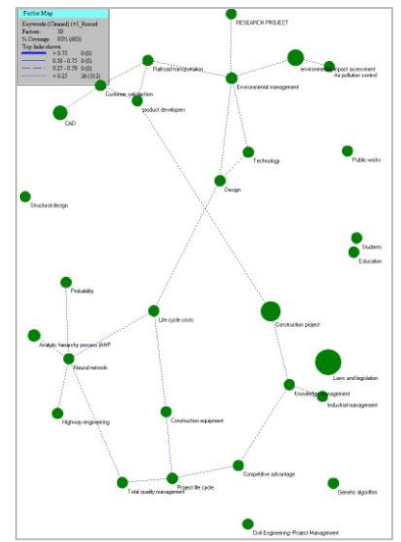
Figure 4 Impacts estimation factor map
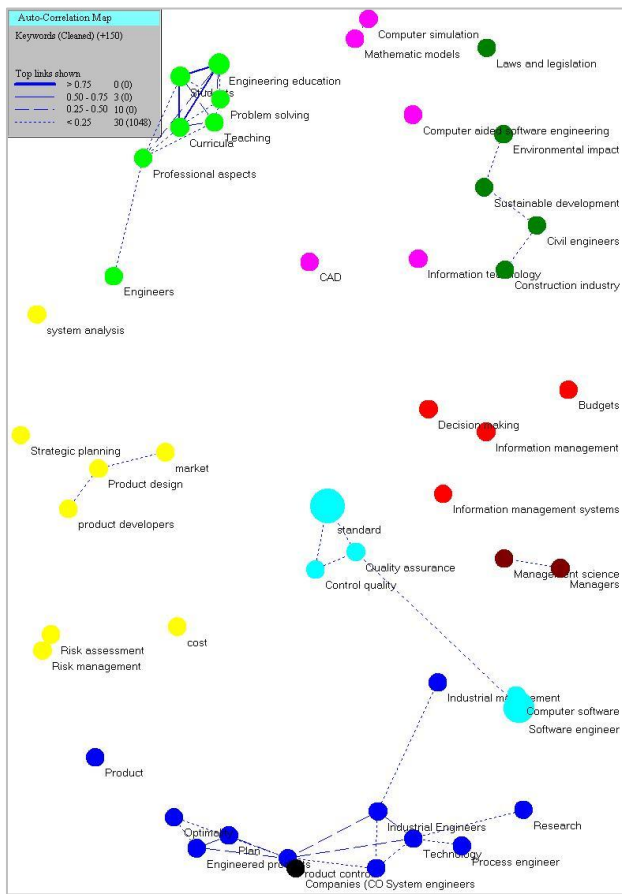
Figure 9 Information systems factor map

Figure 1 Auto-Correlation Map primary database
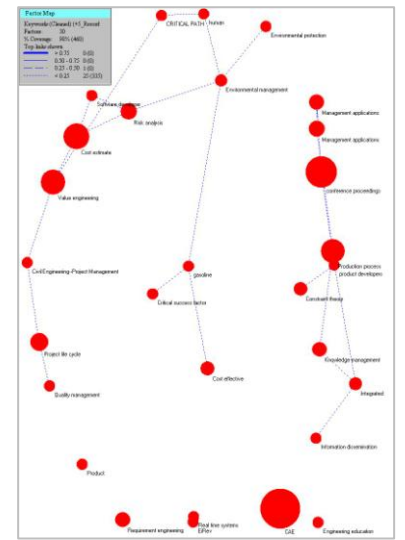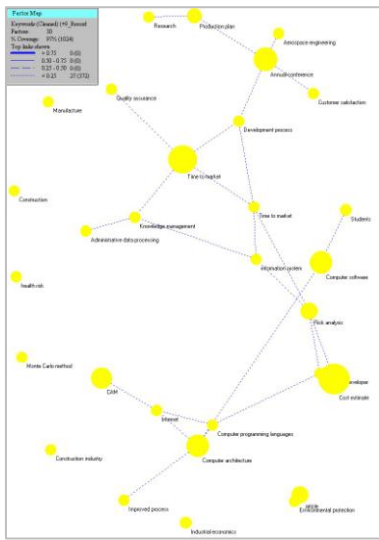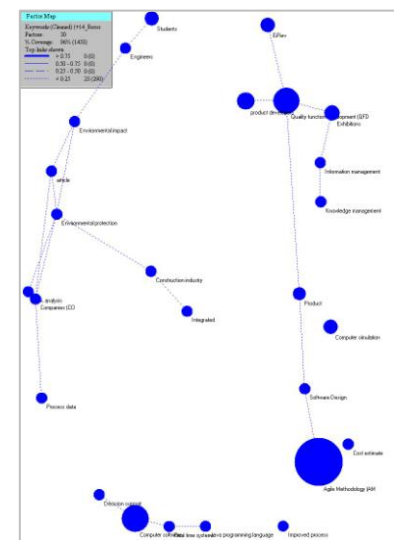
Figure 5 Information sources factor map
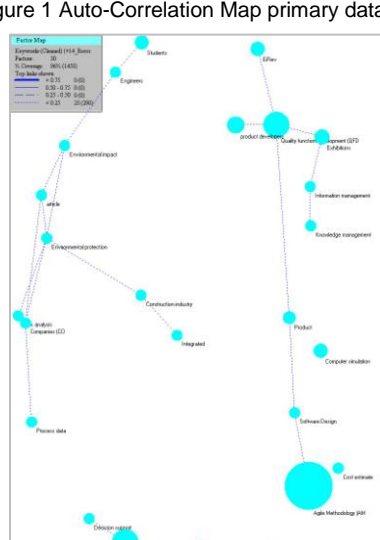
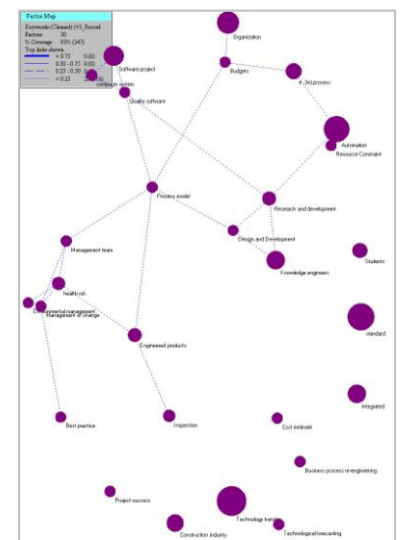Figure 8 Industrial management factor map

Figure 7 Quality assurance and quality factor map

Figure 6 Science management factor map