

Documenting and predicting topic changes in Computers in Biology and Medicine: A bibliometric keyword analysis from 1990 to 2017



Oliver Faust

Department of Engineering and Mathematics, Sheffield Hallam University, UK

ARTICLE INFO

Keywords:

Research trends
Topic analysis
Topic detection and tracking
Text mining
Computers in biology and medicine

ABSTRACT

The Computers in Biology and Medicine (CBM) journal promotes the use of computing machinery in the fields of bioscience and medicine. Since the first volume in 1970, the importance of computers in these fields has grown dramatically, this is evident in the diversification of topics and an increase in the publication rate. In this study, we quantify both change and diversification of topics covered in. This is done by analysing the author supplied keywords, since they were electronically captured in 1990. The analysis starts by selecting 40 keywords, related to Medical (M) (7), Data (D) (10), Feature (F) (17) and (AI) (6) methods. Automated keyword clustering shows the statistical connection between the selected keywords. We found that the three most popular topics in CBM are: Support Vector Machine (SVM), Electroencephalography (EEG) and IMAGE PROCESSING. In a separate analysis step, we bagged the selected keywords into sequential one year time slices and calculated the normalized appearance. The results were visualised with graphs that indicate the CBM topic changes. These graphs show that there was a transition from Artificial Neural Network (ANN) to SVM. In 2006 SVM replaced ANN as the most important AI algorithm. Our investigation helps the editorial board to manage and embrace topic change. Furthermore, our analysis is interesting for the general reader, as the results can help them to adjust their research directions.

1. Introduction

Documenting the use of computers in bioscience and medicine is a very dynamic endeavour. Therefore, Computers in Biology and Medicine (CBM) is a journal which was set-up as a forum to publish scientific articles and reviews.¹ The content areas include medical disease diagnosis [1–4], medical data [5], information processing [6–8] and dissemination [9]. Medical disease creates the need to build physical problem solutions and computer methods realize the required functionality [10,11]. The problem solutions can take the form of biochemical [12], biocontrol [13], neural simulation [14] and automatic computer analysis systems [15–17]. Keeping track of topic changes in that scientific area is important for steering the use of computing machinery in medicine and biology towards novel and forward thinking applications. However, the diverse and dynamic nature of the forum makes it difficult to track and analyse topics over time.

Bibliographic research aims to provide an overview of trends and issues encountered in dynamic literature [18–20]. As such, it is a meta-analysis method which is applied to a substantial body of research literature [21,22]. Not only the topics and the writing style evolve, also the bibliographic features of the documents change over time. For

example, the author supplied keywords in CBM are only captured since 1990. Before that, Topic Detection and Tracking (TDT) [23] was difficult and error prone, because a third party had to extract the topic from the paper text. The ability to conduct TDT studies is significant for meta research on science, technology and policy [24,25]. TDT tools can be used to profile research [26,27], document trends and topics [28–30] as well as analyse research impact [31,32]. Linear models can be used to predict incremental change, but they underperform when it comes to predicting disruptive and revolutionary events [33]. Unfortunately, traditional methods use linear models on static data [34]. For example, keyword cluster analysis is static, because the clusters do not reflect change over time. Therefore, these methods cannot be used to document and predict topic changes [33].

To address the issues raised above, we have analysed the topics covered in CBM with static and dynamic keyword analysis methods. We have applied static frequency and cluster analysis to author supplied keywords from all papers published in CBM since 1990. The frequency analysis shows that Support Vector Machine (SVM), Electroencephalography (EEG), and IMAGE PROCESSING are the most widely used keywords in CBM publications. The cluster analysis reveals the structure within the keyword co-occurrence matrix. With a second analysis step,

¹ URL (01.09.2017): <https://www.journals.elsevier.com/computers-in-biology-and-medicine/>.

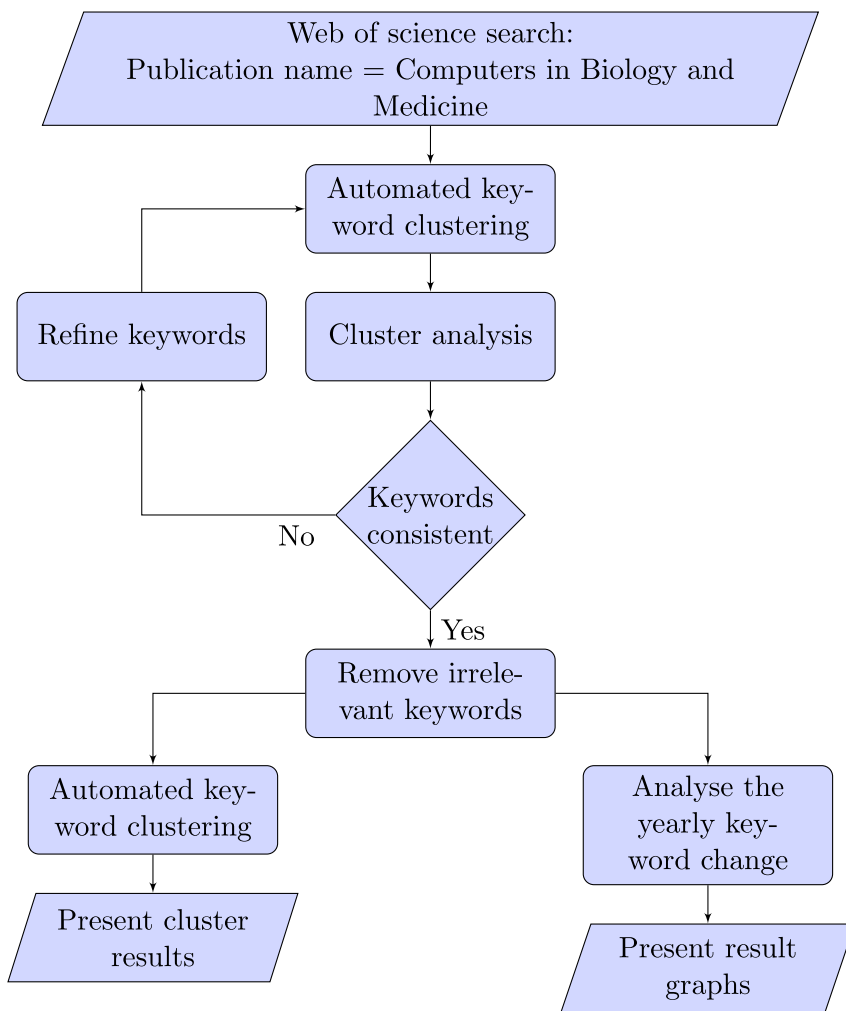


Fig. 1. Flowchart of the keyword analysis.

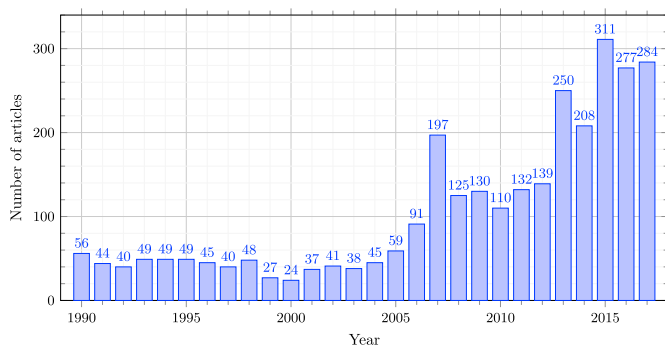


Fig. 2. Number of articles published in from 1990 (Volume 20) to 2017 (Volume 87).

we found the normalized appearance of the most important keywords in yearly intervals from 1990 to 2017. Analysing the normalized keyword appearance reveals the topic change dynamics. In order to interpret these dynamics, we put forward that future topics in CBM will be influenced by Medical (M) needs and advances in Artificial Intelligence (AI), Feature (F) extraction and medical Data (D) acquisition. Based on these four categories, we found that the most striking topic change happened in AI, namely the transition from Artificial Neural Network (ANN) to SVM. The static and dynamic bibliographic analysis results can serve as a basis for the editorial board to keep CBM relevant for the advancement of science

Table 1
Corrections done to the original author keywords.

Keyword	Correction
ALGORITHMS	→ ALGORITHM
SUPPORT VECTOR MACHINE	→ SVM
SUPPORT VECTOR MACHINE (SVM)	→ SVM
SVMS	→ SVM
NUERAL NETWORKS	→ NN
NUERAL NETWORK	→ NN
FINITE ELEMENT METHOD	→ FEM
FINITE ELEMENT ANALYSIS	→ FEM
ELECTROCARDIOGRAM	→ ECG
ELECTROCARDIOGRAM (ECG)	→ ECG
ELECTROENCEPHALOGRAM	→ EEG
ELECTROENCEPHALOGRAM (EEG)	→ EEG
MAGNETIC RESONANCE IMAGING	→ MRI
COMPUTED TOMOGRAPHY	→ CT
ARTIFICIAL NEURAL NETWORK	→ ANN
ARTIFICIAL INTELLIGENCE	→ AI
VIRTUAL REALITY	→ VR
MODELING	→ MODELLING
HEART RATE VARIABILITY	→ HRV
AUTONOMOUS NERVOUS SYSTEM	→ ANS
GENETIC ALGORITHM	→ GA

and technology. The bibliometric research results are interesting for the general reader as well, because they reveal both statistical connections between keywords and trending topics. That information is useful when it comes to deciding on what technology to focus.

Download English Version:

<https://daneshyari.com/en/article/6898906>

Download Persian Version:

<https://daneshyari.com/article/6898906>

[Daneshyari.com](https://daneshyari.com)