

TOWARDS A SYSTEMATIC LITERATURE REVIEW OF TEXT SIMILARITY METHODS FOR SEMANTIC QUALITY ASSESSMENT OF BUSINESS PROCESS MODELS

Rudskiyi O.V.¹, Kopp A.M.²

¹ *Master's Student of the SE & MIT Department, NTU «KhPI», Kharkiv, Ukraine*

² *Associate Professor of the SE & MIT Department, Ph.D., NTU «KhPI», Kharkiv, Ukraine*
alex.rudskii@gmail.com<mailto:your@email.com>

Business process modeling emerged to provide a better understanding of business processes in organizations. The result of business process modeling is a process model, which consists of a set of activity models and the execution constraints between them. It is usually illustrated by activities and events that are associated with management flows. Such processes can be modeled using various process modeling languages, also known as techniques or notations [1].

Business Process Model and Notation (BPMN) is a standard for business process modeling that provides a graphical notation for specifying business processes in the form of a Business Process Diagram (BPD) based on traditional flowcharting methods. The goal of BPMN is to support business process modeling for both technical and business users by providing a notation that is intuitive for business users, but at the same time capable of capturing complex process semantics. The BPMN 2.0 specification also provides runtime semantics, as well as mappings between the notation's graphics and other runtime languages, such as the business process runtime language [2].

The main goal of developing BPMN was to create an understandable notation for creating business process models, while providing semantics and underlying mechanisms to handle the complexity inherent in business processes. The approach taken to address these two conflicting requirements was to organize the graphical aspects of the notation into specific categories. This provides a small set of notation categories so that the reader of a BPMN diagram can easily recognize the basic types of elements and understand the diagram [2].

However, creating business process models is a complex and resource-intensive task, sometimes leading to situations where the model itself does not match the textual description of the business process. This can lead to time and financial losses. Therefore, there is a need to analyze the compliance of business process models with their textual descriptions.

Comparing business process models with their textual descriptions has several advantages. First, it can help ensure that the process model accurately reflects the textual description of the process. Second, it can help identify discrepancies between the two representations, which can be used to improve the quality of the process model. Third, it can help identify areas where the textual description of the process is ambiguous or incomplete, which can be used to improve the quality of the textual description [3].

In addition, comparing business process models with their textual descriptions can help ensure that all stakeholders have a common understanding of the process. This is because different stakeholders may have different levels of familiarity with the BPMN notation and therefore may interpret the process model differently [3].

A Systematic Literature Review (SLR) method was used to investigate current methods for comparing texts.

The SLR method identifies, selects, and critically appraises studies to answer a clearly defined question. A systematic review should follow a well-defined protocol or plan that

clearly defines the criteria for the review. It is a comprehensive, transparent search that is conducted in multiple databases and gray literature that can be replicated and reproduced by other researchers. It involves planning a well-designed search strategy that has a specific focus or answers a specific question. A systematic literature review identifies the type of information that has been sought, critiqued, and reported over a known period of time. Search terms, search strategies (including database names, platforms, search dates), and limitations should all be included in the review [1].

To answer the research questions, the following SLR objectives were identified:

- 1) to review articles to identify existing methods for comparing texts;
- 2) to identify weaknesses in the methods in order to eliminate them through additional research;
- 3) to gain new knowledge about text comparison methods that can be used for further research.

For the purposes of the study, the following search string was used: ("allintitle:" + "text" + "similarity" + "site:" + "ieeexplore.ieee.org").

The initial keyword search yielded 107 scientific articles related to text similarity in different languages. After a thorough review of the articles, we excluded articles that were not directly related to the research topic but appeared in the search results due to the coincidence of keywords. In addition, articles were also excluded due to duplication, lack of full text, and if the studies did not address any of the research questions. After all the exclusions, 10 scientific articles remained (Fig. 1).

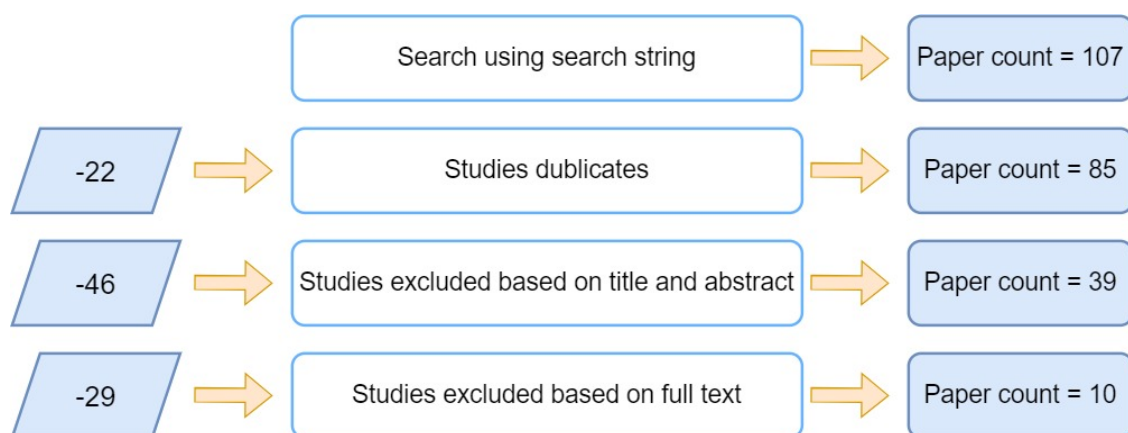


Figure 1 – General systematic literature review scheme

The selected articles (Fig. 1) discuss existing methods for comparing text similarity and the use of different measures of semantic similarity in documents. On the basis of the studied articles, in the following studies it is planned to develop an algorithmic and software tool for improving the semantic quality of business process models by assessing their compliance with textual descriptions of the subject area using natural language processing technologies.

References:

1. Jošt G., Polančič G., Heričko M., Kocbek M. *Business process model and notation: The current state of affairs. Computer Science and Information Systems*. 2015. 12(2). P. 509–539. DOI: 10.2298/CSIS140610006K.
2. Von Rosing M., White S., Cummins F., De Man H. *Business process model and notation-BPMN. The Complete Business Process Handbook: Body of Knowledge from Process Modeling to BPM*. 2014. Vol. 1. P. 429–453. DOI: 10.1016/B978-0-12-799959-3.00021-5.
3. Ottensooser A., Fekete A., Reijers H., Mendling J., Menictas C. *Making sense of business process descriptions: An experimental comparison of graphical and textual notations. Journal of Systems and Software*. 2012. Vol. 85(3). P. 596–606. DOI: 10.1016/j.jss.2011.09.023.