

¹ Andrii Kopp

Ph.D., Associate Professor

² Dmytro Orlovskiy

Ph.D., Associate Professor

^{1,2} *National Technical University “Kharkiv Polytechnic Institute”, Kyrpychova str. 2, Kharkiv, 61002, Ukraine***ENTERPRISE ARCHITECTURE WEB MINING APPORACH**

Abstract. This paper considers the enterprise architecture model extraction from organizational websites in an automatic way to simplify the blueprinting of enterprise architecture landscapes at the conceptual level. Thus, such a technique is proposed to be called “enterprise architecture web mining”.

Keywords: Enterprise Architecture, Business Process, Web Content Mining.

This paper proposes an approach and a software tool for the automatic extraction of Enterprise Architecture (EA) landscapes from websites that nowadays virtually represent organizations on the Internet. It aims at simplifying the procedure of building high-level models in the preliminary stages of EA development. It is well known that today most enterprises offer their products and services on their homepages top-ranked by multiple search engines. Usually, organizational websites contain information not only about offered products or services but also about related activities that allow customers to receive respective products or services (e.g. order, buy, learn, etc.). The study object is the procedure of EA structure extraction from organizational websites that serve as virtual enterprise representations on the Internet. The study subject is the approach and software tool to extract EA landscapes from organizational websites. The study goal is to simplify the process of EA description in the early stages of EA development.

The suggested “EA web mining” is focused on the automatic construction of EA models using corporate websites as sources of data about EA elements and the relationships between them. Hence, the main problem is finding mentions of business processes and other EA elements in HyperText Markup Language (HTML) pages of corporate websites. Whereas the direct search in Google Scholar using the “enterprise architecture web mining” key phrase did not give any results, the “enterprise architecture mining” allowed us to discover several studies in this direction. In [1] the author states that manual maintenance of EA models is costly and time-consuming, so they propose EA mining algorithms and tools based on process mining. The study [2] also considers automatic EA modeling methods that are supposed to reduce the drawbacks of manual EA modeling (error-proneness, time and cost consumption, accuracy, etc.). The systematic review [3] also states that automatic EA modeling could respond challenges of manual EA modeling but this field is still immature and requires further research.

First of all, the web page should be parsed to work with its tags, their attributes, and text content. Then web page tags should be used to extract the data about the

organization's activity described on its web page on the Internet. Using the structured tag data, business activities that help an organization virtually promote its products or services on the Internet should be detected. Finally, using the set of business activities and the previous outcomes, the EA model should be built. The conceptual model of automatic EA model construction using the company's homepage on the Internet is demonstrated in Fig. 1.

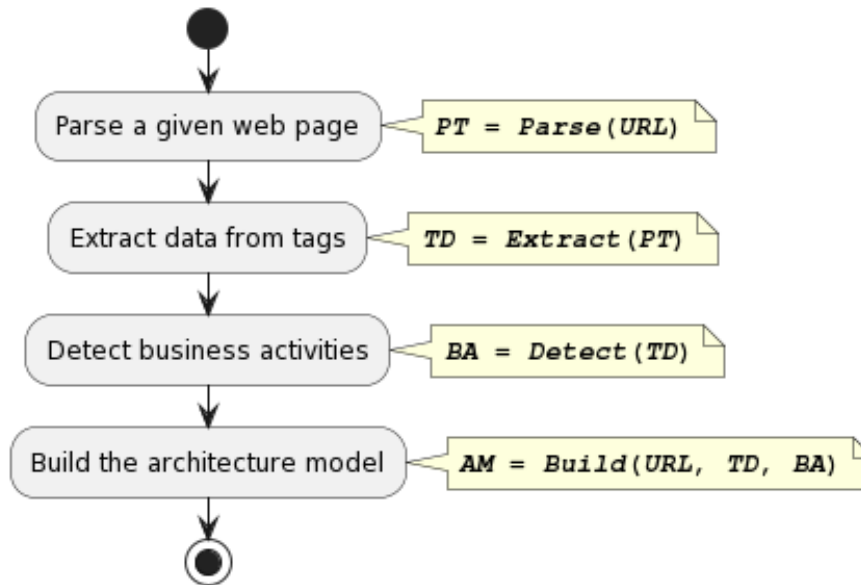


Figure 1 – The conceptual model of “EA web mining” approach

The proposed workflow (Fig. 1) should help automatically build high-level architectural models using only the websites of organizations using the suggested technology we can name “enterprise architecture web mining”. Obtained models may describe landscapes of top-level business processes based on products or services offered to customers on the company's homepage. Moreover, obtained EA models should include application layers to demonstrate website maps, and technology layers to complete the EA cross-layer architecture. However, the most valuable outcome is still a business architecture layer that includes core value-added business processes and the business service offered to the organization's clients. EA models automatically produced using the company's website can help to understand the current state of the enterprise, including its customer relationship strategy, offered products, and services. Then, shortcomings could be detected in such an EA model, and the decisions to improve the enterprise's virtual representation on the Internet could be made.

References:

1. A. Fajri, Enterprise Architecture Mining, MS thesis, University of Twente, 2019.
2. R. Pérez-Castillo, F. Ruiz, M. Piattini, A decision-making support system for Enterprise Architecture Modelling, Decision Support Systems 131 (2020) 113249. doi:10.1016/j.dss.2020.113249.
3. R. Perez-Castillo et al., A systematic mapping study on enterprise architecture mining, Enterprise Information Systems 5(13) (2019) 675–718. doi:10.1080/17517575.2019.1590859.