

Software-Assisted Knowledge Generation in the Cultural Heritage Domain: A Conceptual Framework

Patricia Martín Rodilla

Computer Science Ph.D., Departamento de Sistemas Informáticos y Computación DSIC. Centro de Investigación en Métodos de Producción de Software PROS Universitat Politècnica de València. Camino de Vera s/n Edificio 1F – 46022, Valencia, España

Researcher, Instituto de Ciencias del Patrimonio (Incipit). Consejo Superior de Investigaciones Científicas (CSIC). Avda. de Vigo s/n. Campus Sur. C.P. 15705 Santiago de Compostela (A Coruña), España.

Advisors

Prof. Dr. Óscar Pastor López

Dr. César González Pérez

Abstract

Software Engineering provides a repository of techniques, methods and tools to manage, process, use and exploit information. In recent decades, this corpus has not only been applied to domains that traditionally act as a receivers of software solutions, but also it has been expanded and enriched by contributions from other disciplines and domains with needs related to the information produced. One of the most common needs in these disciplines is the software assistance to experts or domain professionals in performing processes ranging from the analysis of raw gathered data to the generation of new knowledge based on these, thus allowing the continuous advance of the discipline. Good examples are the software assistance in contexts of genomics research and in business decision-making processes. In order to assist knowledge generation processes through software, it is necessary a deep understanding of the Software Engineering corpus as well as the particularities of the domain assisted and how knowledge is generated inside it.

This situation appears too in the Cultural Heritage domain, whose professionals produce and manage large amounts of data about evidences of our past and present, from which they create new knowledge that constitutes the knowledge about heritage of a particular community. This knowledge defines the community at present, and is transmitted to the present and future generations. Despite their relevance and the regular application of Software Engineering solutions to the Cultural Heritage domain, the knowledge generation process in Cultural Heritage poses a challenge for Software Engineering, mainly due to the low presence of formal studies of the process, making it difficult to assist it through software. The lack of formal studies implies that we do not know which particular processes in Cultural Heritage we must assist and what should be the appropriate assistance in each case. Furthermore, the Cultural Heritage domain and, in general, the humanities, possesses some particular characteristics that are especially difficult to deal with by software, such as the presence of high subjectivity, the fact that much information is uncertain or vague, and the importance of the temporal aspect in the information.

In order to address these two challenges from a transdisciplinary perspective, this thesis presents a conceptual framework based on software models for the construction of software solutions to assist to the knowledge generation process in Cultural Heritage. Firstly, the thesis conducts a deep exploration of the knowledge generation processes in Cultural Heritage, whose inputs are mainly textual sources. As a result, the thesis proposes a methodology and a modelling language to use discourse analysis in Software Engineering. By using this approach, it is possible to relate elements of a text with the domain entities that are referenced and the argumentative mechanisms used during the knowledge generation process and captured in the text. Subsequently, the thesis proposes a conceptual framework whose implementation allows to manage the domain particularities mentioned above, providing a software assistance to the Cultural Heritage professionals through information visualization techniques.

The proposed conceptual framework has been validated in two complementary ways. On the one hand, we have developed a full case study in the Cultural Heritage domain, for which we have instantiated all the software models proposed as part of the framework to represent a real-world scenario. This case study application has revealed the potential of the framework in terms of conceptual representation, technical support and software-assistance definition mechanisms. On the other hand, the proposed software models have been implemented as a functional iOS application prototype. The prototype has been validated empirically against professionals in Cultural Heritage, comparing the performance of knowledge generation processes using the proposed framework to the conventional ways without software assistance. The empirical validation has revealed how the proposed framework provides a robust solution—based on the software models defined as part of the conceptual framework—for the construction of software systems to assist to the knowledge generation process in the Cultural Heritage domain. In addition, the thorough validation performed in these two complementary forms has detected areas for improvement and future research lines that will allow us to conceive other software-assistance possibilities, not only looking for software-assisted knowledge generation in Cultural Heritage, but also in related but unexplored domains.