

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

**Patricia Martín Rodilla**

Estudiante de Doctorado en Informática, 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

Investigadora, 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.

Directores

Prof. Dr. Óscar Pastor López

Dr. César González Pérez

## Resumen

La ingeniería del software ofrece un repositorio de técnicas, métodos y herramientas como soluciones para el manejo, tratamiento, uso y explotación de información. En las últimas décadas, este corpus no sólo ha sido aplicado a dominios tradicionalmente receptores de soluciones software, sino que se ha expandido y enriquecido con aportaciones de diversas disciplinas y dominios con necesidades relacionadas con la información que producen. Una de las necesidades más habituales es la asistencia a los profesionales de dichas disciplinas durante el proceso evolutivo que realizan desde el análisis de los datos más primarios hasta la generación de conocimiento nuevo que permita avanzar en la disciplina involucrada. Buenos ejemplos son la asistencia software en contextos de investigación genética y en asistencia a la toma de decisiones de negocio. Para poder asistir mediante software este proceso, es necesario un conocimiento profundo del corpus de ingeniería del software, pero también de las especificidades del dominio a asistir y cómo se genera conocimiento en él.

Este es el caso del Patrimonio Cultural, cuyos profesionales producen y manejan ingentes cantidades de datos acerca de evidencias sobre nuestro pasado y presente, y desde los cuáles descubren y generan conocimiento nuevo, que supone la herencia cultural propia de una comunidad. Este conocimiento define la comunidad en el presente y es transmitido a las generaciones presentes y futuras. Pese a su relevancia y a la habitual aplicación de determinadas soluciones de ingeniería software en el dominio, el proceso de generación de conocimiento en Patrimonio Cultural representa en sí mismo un reto para la ingeniería del software, debido fundamentalmente a la poca presencia de estudios formales acerca del mismo, lo que dificulta su asistencia mediante software. Esto implica que no sabemos qué tipo de subprocesos debemos asistir mediante software ni cuál es la asistencia más adecuada. Además, el corpus actual en ingeniería del software debe soportar especificidades del dominio patrimonial y, en general, de las humanidades, como son la presencia de una alta subjetividad, el hecho de que mucha información es incierta o vaga, y la importancia del aspecto temporal en los datos.

Con el objetivo de abordar estos dos retos desde una perspectiva co-investigadora y transdisciplinar, la presente tesis doctoral presenta un marco conceptual basado en modelos software para la construcción de soluciones software que asistan a la generación de conocimiento en Patrimonio Cultural. En primer lugar, la tesis explora a fondo el proceso de generación de conocimiento en Patrimonio Cultural, cuyas fuentes eminentemente textuales han dado lugar a la propuesta de una metodología completa y un lenguaje de modelado para utilizar análisis del discurso en ingeniería del software. Esta propuesta permite que se puedan relacionar elementos de un texto con las entidades del dominio que se referencian, así como los mecanismos argumentativos que se emplean. Posteriormente, la tesis propone un marco conceptual completo cuya implementación permite gestionar las especificidades del dominio antes señaladas, ofreciendo una asistencia mediante técnicas de visualización de información software a los especialistas en Patrimonio Cultural.

El marco conceptual propuesto ha sido validado de dos maneras complementarias. Por un lado, se ha desarrollado un caso de estudio patrimonial completo, para el cual se han implementado todos los modelos software del marco conceptual propuesto, representando un escenario de aplicación completo del mundo real. Este caso de estudio ha permitido comprobar la potencia del marco conceptual propuesto en cuanto a representación, soporte y definición de mecanismos de asistencia software. Por otro lado, los modelos software que conforman el marco conceptual propuesto han sido implementados en un prototipo funcional en forma de aplicación iOS. Esto ha permitido contar con una implementación real de asistencia software en Patrimonio Cultural. Dicha solución se ha validado empíricamente con profesionales del dominio, comparándola con los modos de generación de conocimiento habituales sin dicha asistencia. La validación empírica ha permitido comprobar cómo el marco conceptual propuesto constituye una solución sólida para la construcción, a partir de los modelos software especificados en el mismo, de sistemas software para asistir a la generación de conocimiento en Patrimonio Cultural. Además, el profundo trabajo de validación en ambos sentidos ha permitido detectar áreas de mejora y líneas futuras de trabajo que permitan implementar otros tipos de asistencia software, buscando que la asistencia software a la generación de conocimiento sea una realidad no sólo en Patrimonio Cultural, sino en dominios similares inexplorados.