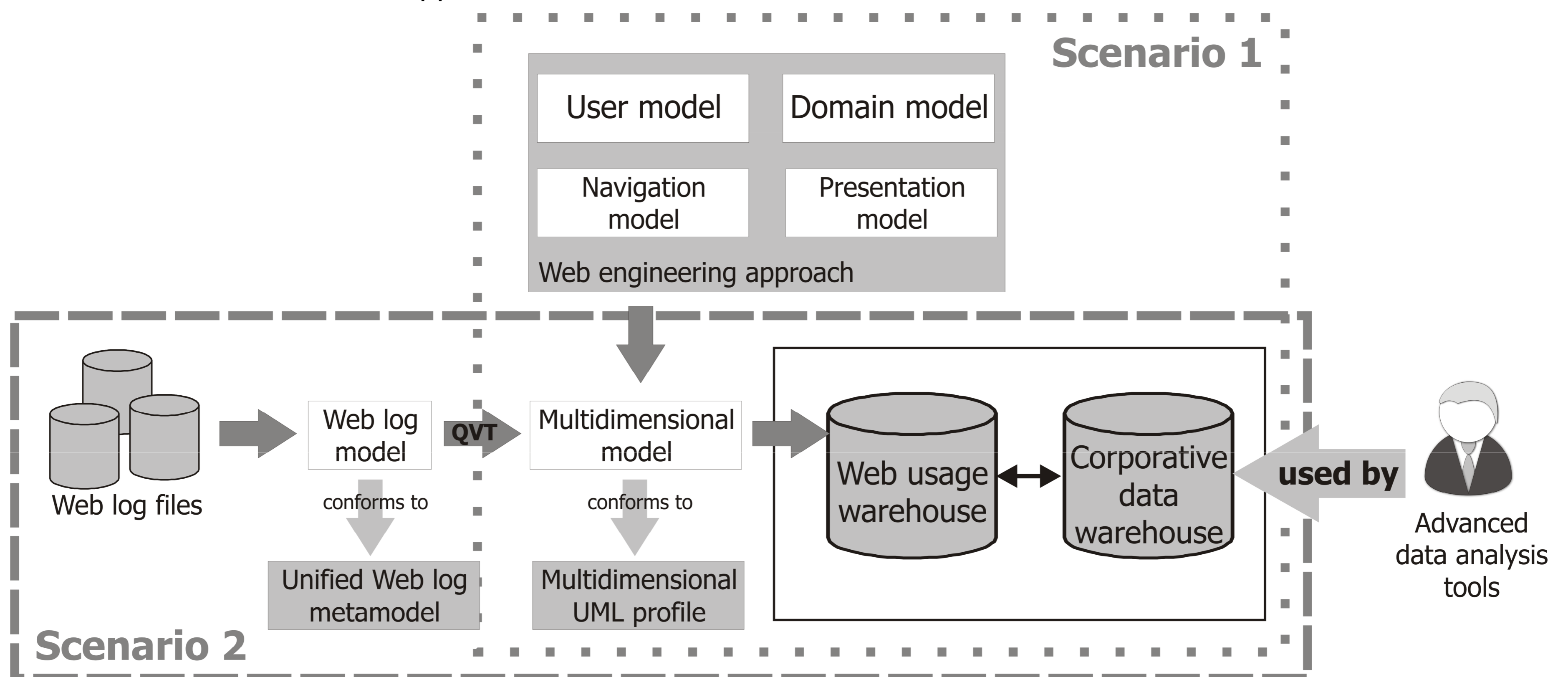


# Towards a model-driven framework for Web usage warehouse development

Paul Hernández, Octavio Glorio, Irene Garrigós and Jose-Norberto Mazón  
DLSI/Lucentia, University of Alicante, Spain  
{phernandez, oglorio, igarrigos, jnmazon}@dlsi.ua.es

## Motivation

- **Web usage analysis** is the process of finding out what users are looking for on the Internet.
- Analyzing the usage of a website is a **key issue for a company to improve decision making** → evolution of a website
- **Advanced data analysis tools** requires the development of a data warehouse or data mart to structure data in a multidimensional model
- **Commercial tools (e.g. Google Analytics)** some have drawbacks
  - Significant **limitations performing advanced analytical tasks**.
  - **Uselessness** when trying to understand **users navigational patterns**.
  - Inability to integrate and correlate information from **different sources**.
  - Unawareness of the **conceptual schema** of the application.
- Several approaches propose **multidimensional schema** to analyze Web usage by using Web log data, **however**:
  - Multidimensional elements are informally chosen according to a specific format
  - The resulting multidimensional model may be incomplete
- **Web engineering** methods should provide techniques for designing a multidimensional model of the Web usage at the same time that the rest of the Web application



## Our approach

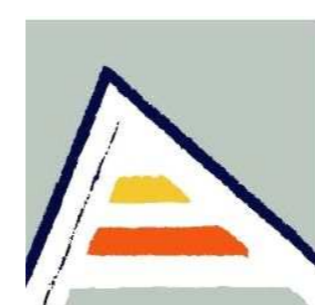
Our **model-driven framework for developing a Web usage warehouse** considers two different scenarios

### 1. Web usage warehouse within model-driven Web engineering

- Several conceptual models when designing a Web application (navigation model, user model, etc.)
- A **multidimensional model should be derived from Web models** to represent and understand Web usage
  - Multidimensional concepts (facts, dimensions, hierarchies, etc.) should be identified within the conceptual models of a Web application to build a Web usage warehouse in an integrated and structured manner

### 2. Web usage warehouse from Web log data

- A Web usage warehouse is developed directly Web log files by using a **Web log metamodel**
  - Elements and semantics that allow building a conceptual model from Web log files
  - Representing the interaction between raw data elements (i.e. the client remote address) and usage concepts (i.e. session, user)



Universitat d'Alacant  
Universidad de Alicante