

# *A Semantic Middleware to enhance current Multimedia Retrieval Systems with Content-based functionalities*

**a dissertation submitted to the Department of Computer Languages and Systems of the Computer Science Faculty of the University of the Basque Country in partial fulfillment of the requirements for the degree of Doctor of Philosophy  
by Gorka Marcos Ortego**

## **ABSTRACT**

This work reviews the information retrieval theory and focuses on the revolution experimented in that field promoted by the digitalization and the widespread use of the multimedia information. After analyzing the trends and promising results in the main disciplines surrounding the content-based information retrieval field, this thesis proposes a reference model for Multimedia Information Retrieval that aims to contextualize the thesis contributions. According to this reference model, this work proposes an architecture for a component named “Semantic Middleware” that aims to centralize the main semantic services to be provided during the indexing, storage, search, retrieval and consume of the multimedia elements. This architecture has been designed from a pragmatic point of view, aiming to facilitate the enhancement of the current systems with content-based functionalities. The architecture proposal includes a set of key design criteria for a right deployment. In order to validate this thesis, two real complementary deployments have been performed and reported in this work.

## RESUMEN

En esta tesis, el autor revisa la teoría de recuperación de información y el impacto que ha tenido en dicho ámbito la digitalización de los contenidos y el incremento exponencial de la presencia de la información multimedia. Tras analizar las tendencias y los prometedores resultados de las principales disciplinas del ámbito científico denominado “content-based retrieval”, esta tesis propone un modelo de referencia para la recuperación de contenido multimedia. Este modelo sirve de marco de referencia sobre el que contextualizar las contribuciones de esta tesis. Precisamente sobre ese modelo, esta tesis propone una arquitectura para un componente denominado “Semantic Middleware” y que tiene como objetivo centralizar la provisión de los servicios semánticos que se puedan llegar a necesitar durante la catalogación, el almacenamiento, la recuperación y el consumo de los contenidos multimedia. La definición de esta arquitectura se ha realizado desde una perspectiva pragmática, teniendo como principal objetivo la mejora de los sistemas de información actuales con funcionalidades “content-based”. La propuesta de arquitectura de esta tesis va acompañada de un conjunto de recomendaciones de diseño para facilitar una correcta materialización de dicho módulo. La validación de la arquitectura y del conjunto de recomendaciones se ha realizado sobre dos desarrollos de este componente en dos entornos industriales complementarios.

## LABURPENA

Tesi honetan egileak informazioaren berreskuratzearen teoria aztertzen du, baita edukien digitalizazioak eta multimedia informazioaren presentziaren areagotze esponenzialak aipaturiko eremuan izandako eragina ere. Joerak eta “*content-based retrieval*” izeneko zientzia-eremuko diziplina nagusien emaitza oparoak aztertu ondoren, tesi honek multimedia edukien berreskurapenerako erreferentzia eredu bat proposatzen du. Eredu hau erreferentzi markotzat hartu daiteke tesi honen ekarpenak testuinguruan jartzeko. Hain zuzen ere eredu horren gainean tesi honek “*Semantic Middleware*” izeneko osagai batentzako arkitektura proposatzen du eta multimedia edukien katalogazio, biltegiatze, berreskuratze eta kontsumoan zehar behar izan daitezkeen zerbitzu semantikoaren hornidura zentralizatzea du helburu. Arkitektura honen definizioa ikuspuntu pragmatiko batean oinarritu da, helburu nagusitzat “*content-based*” funtzionalitateekin egungo informazio sistemen hobekuntza duelarik. Tesiaren arkitektura proposamenari aipaturiko moduluaren gauzatze egokia errazteko diseinu gomendio bilduma bat gehitzen zaio. Arkitekturaren eta gomendio bildumaren baliozkotzea osagai honen bi garapenen inguruan garatu da bi industri giro osagarrietan.