



Universidad de Granada

Ciencias de la Computación e Inteligencia Artificial

**TRATAMIENTO INTELIGENTE
DE DATOS EN PROYECTOS DE
EDIFICACIÓN**

TESIS DOCTORAL

MARÍA MARTÍNEZ ROJAS

Noviembre 2015



Tratamiento Inteligente de Datos en Proyectos de Edificación

Programa de Doctorado
Ciencias de la Computación y Tecnología Informática

Memoria que presenta
María Martínez Rojas
Para optar al título de
Doctora por la Universidad de Granada

Dirigida por
Dr. Nicolás Marín Ruiz
y
Dr. María Amparo Vila Miranda

Ciencias de la Computación e Inteligencia Artificial

Noviembre 2015

Editorial: Universidad de Granada. Tesis Doctorales

Autora: María Martínez Rojas

ISBN: 978-84-9125-368-6

URI: <http://hdl.handle.net/10481/41296>

VISTO BUENO

El Prof. Dr. **Nicolás Marín Ruiz** Titular de Universidad y la Prof. Dra. **María Amparo Vila Miranda** Catedrática de Universidad,

INFORMAN:

Que la memoria titulada:

“Tratamiento Inteligente de Datos en Proyectos de Edificación”

ha sido realizada por **D. María Martínez Rojas** bajo nuestra dirección en el Departamento de Ciencias de la Computación e Inteligencia Artificial de la Universidad de Granada para optar al título de **Doctora por la Universidad de Granada**.

En Granada, a 25 de Noviembre 2015.

Los directores:

Fdo. Nicolás Marín Ruiz

Fdo. María Amparo Vila Miranda

COMPROMISO DERECHOS DE AUTOR

La doctoranda María Martínez Rojas y los directores de la Tesis Nicolás Marín Ruiz y María Amparo Vila Miranda garantizamos, al firmar esta Tesis doctoral, que el trabajo ha sido realizado por el doctorando bajo la dirección de los directores de la Tesis y hasta donde nuestro conocimiento alcanza, en la realización del trabajo, se han respetado los derechos de otros autores a ser citados, cuando se han utilizado sus resultados o publicaciones.

Granada, Noviembre 2015

Los directores:

Fdo. Nicolás Marín Ruiz

Fdo. María Amparo Vila Miranda

El doctorando:

Fdo. María Martínez Rojas

AGRADECIMIENTOS

En primer lugar, me gustaría expresar mi más profundo agradecimiento a mis dos directores de tesis, Amparo y Nicolás. Gracias por vuestro cariño, apoyo, implicación y paciencia durante estos años.

Asimismo, debo agradecer el apoyo económico proporcionado por el programa “Formación de Profesorado Universitario” (AP2009-3089), Ministerio de Educación, Cultura y Deporte (Gobierno de España).

También quiero agradecer a los miembros del departamento de Ciencias de la Computación e Inteligencia Artificial, y en especial a todos los compañeros del grupo de investigación IDBIS, por la cálida acogida y por hacerme sentir como una más.

A mis padres, a los que tengo tanto que agradecer que no habría palabras suficientes para expresarlo. Gracias por inculcarme los valores de la constancia y del esfuerzo, por creer en mí y por ser el espejo en el que mirarme cada día.

A mis hermanos por estar siempre a mi lado. En especial a Javier, por tu paciencia, bondad, generosidad, y por la cantidad de veces que has escuchado “es sólo un ratillo”. Gracias de todo corazón!

A mis compañeros de aparejadores, por el cariño, la confianza y el apoyo que siempre me dais. Especialmente a los que han peleado para darnos la posibilidad de iniciar este camino.

A todos mis amigos, por estar siempre ahí a pesar de no poder vernos todo lo que nos gustaría. A mis compañeros de fatigas Inma y Manu, con los que he compartido tantos ratos durante estos años.

Por último, mi más sentida gratitud a Soto. Gracias por caminar siempre a mi lado, por tu generosidad infinita, por tu paciencia, por tu apoyo, por tus palabras de ánimo y porque a tu lado todo es más sencillo. Gracias por ser como eres! Si me has aguantado este último año, creo que serás capaz de aguantarme toda la vida.

GRACIAS A TODOS

ÍNDICE DE CONTENIDOS

1	INTRODUCCIÓN	1
1.1	Español	1
1.1.1	Objetivos	3
1.1.2	Estructura de la Tesis	5
1.2	English	6
1.2.1	Objectives	7
1.2.2	Thesis structure	9
2	ANTECEDENTES Y PRELIMINARES	11
2.1	Características del sector	12
2.2	Pilares TIC en gestión de proyectos	15
2.2.1	Web, Web Semántica y Computación en la Nube	15
2.2.2	Sistemas de Información	18
2.2.2.1	Modelado de los datos a través de BIM	19
2.2.2.2	Sistemas de Información Geográfica	20
2.2.2.3	Almacenes de datos	21
2.2.2.4	Sistemas de Información Inteligentes	22
2.2.3	Tecnologías para el seguimiento y localización	23
2.3	Manejo de datos en Gestión de Proyectos	24
2.3.1	Información general y gestión del conocimiento	25
2.3.2	Costes	28
2.3.3	Planificación	31
2.3.4	Gestión de riesgos	34
2.3.5	Seguridad	36
2.3.6	Monitorización del proceso	39
2.3.7	Gestión de la calidad	41
2.4	Análisis y motivación de nuestra investigación	43
2.4.1	Análisis del estado actual	43
2.4.2	Necesidad de nuevos enfoques para la gestión de los datos	51
3	MODELO FORMAL PARA LA ESTRUCTURA DEL PRESUPUESTO	57
3.1	Introducción	58
3.2	Descripción del problema	59
3.2.1	El proceso de elaboración del presupuesto	59
3.2.2	Problema de la ausencia de un modelo común	62
3.2.2.1	Ejemplo ilustrativo	63
3.3	Formalización del modelo de datos	65
3.3.1	Descripción de partida	66
3.3.2	Estructura de descomposición del trabajo	67
3.3.3	Relación de Parentesco	68

3.3.4	Elementos de la Estructura de Descomposición del Trabajo	70
3.4	Una estructura de referencia basada en conocimiento experto	72
3.4.1	Nivel L ₂ : Capítulos	73
3.4.2	Nivel L ₃ : Subcapítulos	75
3.4.3	Nivel L ₄ : Descripciones de Partidas	77
3.5	Conclusiones	77
4	MECANISMO DE CLASIFICACIÓN AUTOMÁTICO	79
4.1	Introducción	80
4.2	Extracción de conocimiento para la clasificación	81
4.2.1	Preprocesamiento del texto	82
4.2.2	Cálculo de información relativa a los términos	83
4.2.2.1	Formalización de las medidas	85
4.3	Proceso de clasificación	88
4.3.1	Método de agregación multi-criterio	89
4.3.2	Ejemplo ilustrativo	91
4.4	Experimentación	96
4.4.1	Datos de aprendizaje y prueba	96
4.4.2	Ajuste de los pesos del operador OWA	98
4.4.3	Resultados detallados con la configuración OWA seleccionada	101
4.5	Evaluación	105
4.6	Conclusiones	109
5	APLICACIÓN PARA LA ADQUISICIÓN DE DATOS Y GESTIÓN DE COSTES	113
5.1	Introducción	114
5.2	Sistema de adquisición y edición	115
5.2.1	Módulo de adquisición	116
5.2.1.1	Importación del documento	116
5.2.1.2	Proceso de clasificación	119
5.2.2	Módulo de edición	122
5.2.3	Módulo de recuperación	124
5.2.4	Implementación de la aplicación SIGPE	126
5.2.4.1	Arquitectura del sistema on-line	127
5.2.5	Evaluación de la aplicación	129
5.2.5.1	Resultados	129
5.2.5.2	Cuestionario	132
5.3	Un herramienta flexible para el análisis de costes	135
5.3.1	Estructura multidimensional sobre datos de costes	136
5.3.1.1	Conceptos previos modelo multidimensional difuso	137
5.3.1.2	Propuesta de estructura multidimensional	139
5.3.1.3	Consultas	146

5.3.2	Implementación de la aplicación	150
5.4	Conclusiones	151
6	CONCLUSIONES Y TRABAJOS FUTUROS	153
6.1	Versión en Castellano	153
6.1.1	Conclusiones	153
6.1.2	Trabajos Futuros	158
6.2	English Version	162
6.2.1	Conclusions	162
6.2.2	Future Work	166
6.3	Publicaciones científicas	169
A	MANUAL DE LA APLICACIÓN SIGPE	171
A.1	Inserción de un presupuesto	174
A.2	Edición de un presupuesto	176
A.3	Recuperación/Consulta	181
	 BIBLIOGRAFÍA	 185

ÍNDICE DE FIGURAS

Figura 1	Distribución de tecnologías por tareas del ámbito de la gestión de proyectos	45
Figura 2	Esquema del proceso de elaboración del presupuesto	60
Figura 3	Estructura de descomposición del trabajo del ejemplo 2	68
Figura 4	La estructura de descomposición de trabajo (WBS)	70
Figura 5	Elemento de la estructura de descomposición de trabajo (WBS)	71
Figura 6	Propuesta de estructura jerárquica para los presupuestos en proyectos de edificación en España	73
Figura 7	Proceso de clasificación	78
Figura 8	Proceso de clasificación	88
Figura 9	Resultados por capítulos en el nivel L ₂ (capítulos)	101
Figura 20	Resultados por subcapítulos en el nivel L ₂ (capítulos)	102
Figura 11	Resultados para el conjunto de proyectos en el nivel L ₂ (capítulos)	103
Figura 12	Resultados para el conjunto de proyectos en el nivel L ₃ (subcapítulos)	104
Figura 13	Ejemplo de presupuesto elaborado con la herramienta Cype y exportado a formato Excel	118
Figura 14	Ejemplo de presupuesto elaborado con la herramienta Presto y exportado a formato Excel	118
Figura 15	Ejemplo del proceso del módulo de adquisición .	119
Figura 16	Pantalla para insertar un nuevo proyecto en la aplicación	121
Figura 17	Pantalla que muestra la información del proyecto y permite iniciar el proceso de clasificación	121
Figura 18	Pantalla que muestra la información del proyecto y la distribución de las descripciones de partida .	122
Figura 19	Pantalla con los capítulos desplegados	124
Figura 20	Pantalla con menú desplegable para modificar el resultado	125
Figura 21	Pantalla con la lista de proyectos almacenados . .	126
Figura 22	Pantalla de consulta	127
Figura 23	La arquitectura cliente-servidor de la aplicación <i>SIGPE</i>	128
Figura 24	Cubo para el análisis de costes	140

Figura 25	Valores y agrupación del nivel <i>tamaño</i> en la dimensión <i>proyecto</i>	141
Figura 26	Valores y agrupación del nivel <i>tamaño</i> en la dimensión <i>empresa</i>	144
Figura 27	Valores y agrupación del nivel <i>fases</i> en la dimensión <i>actividad</i>	144
Figura 28	Beneficio medio en los proyectos de viviendas unifamiliares en función de la tipología del promotor - Crisp	147
Figura 29	Beneficio medio en la fase inicial en función del tamaño de la empresa	148
Figura 30	Desviación de coste máximo en cada fase según las tipologías de los proyectos	149
Figura 31	Desviación máxima de coste en proyectos de gran tamaño según las fases de la obra	149
Figura 32	Desviación de coste media en proyectos de gran tamaño según los capítulos del presupuesto	150
Figura 33	Inicio de sesión en <i>SIGPE</i>	173
Figura 34	Ventana principal en <i>SIGPE</i>	174
Figura 35	Insercción de un nuevo proyecto en <i>SIGPE</i>	175
Figura 36	Información del proyecto una vez importado documento del presupuesto.	176
Figura 37	Información del proyecto y la distribución de las descripciones de partida en la estructura de referencia.	177
Figura 38	Lista de proyectos almacenados para un usuario .	178
Figura 39	Pantalla con la lista de capítulos para el proyecto de ejemplo	179
Figura 40	Capítulo 8 desplegado	180
Figura 41	Pantalla con menú desplegable para modificar el resultado	181
Figura 42	Procedimiento para la modificación de una descripción de partida	182
Figura 43	Apariencia del módulo de "Recuperación"	183
Figura 44	Ilustración de un ejemplo de búsqueda en <i>SIGPE</i>	184

ÍNDICE DE TABLAS

Tabla 1	Hallazgos de las TIC en relación a las tareas de gestión en construcción	51
Tabla 2	Bases de precios	62
Tabla 3	Estructuración en el primer nivel de organización del documento de presupuesto correspondiente a tres proyectos de edificación de viviendas unifamiliares	64
Tabla 4	Ejemplo de descripción de partida de trabajos del terreno en los tres proyectos de ejemplo	65
Tabla 5	Ejemplo de descripción de partida de geotextil en los tres proyectos de ejemplo	65
Tabla 6	División de capítulos y subcapítulos propuesta	76
Tabla 7	Número de apariciones del término <i>tabique</i> en los subcapítulos del capítulo de <i>albañilería</i>	86
Tabla 8	Número de apariciones del término <i>tabique</i> en cada uno de los capítulos de la estructura de referencia.	86
Tabla 9	Cálculos de $F^{T^{odo}}(\delta_1)$ (Importancias)	93
Tabla 10	Cálculos de $F^{T^{odo}}(\delta_1)$ (Importancia general e idoneidad)	94
Tabla 11	Cálculos de $F^{GT^2}(\delta_1)$ (Importancias)	95
Tabla 12	Cálculos de $F^{GT^2}(\delta_1)$ (Importancia general e idoneidad)	96
Tabla 13	Combinaciones para el vector de pesos del OWA ₁ que agrega los valores de las tres importancias	99
Tabla 14	Los cinco mejores resultados para el nivel de capítulo y subcapítulo para ambos conjuntos de test	100
Tabla 15	Influencia de las <i>Importancias</i> en el proceso de clasificación	105
Tabla 16	<i>Precision, Recall</i> y <i>F-measure</i> para cada Capítulo	106
Tabla 17	<i>Precision, Recall</i> y <i>F-measure</i> para cada subcapítulo	109
Tabla 18	Número de términos y número de descripciones de partidas para cada capítulo (L2) y subcapítulo (L3) de la estructura de referencia	111
Tabla 19	Resultados obtenidos para el nivel de capítulo y subcapítulo mediante la interacción de los usuarios con la aplicación SIGPE	131
Tabla 20	Cuestionario sobre la aplicación SIGPE	133

Tabla 21	Media aritmética (μ) y desviación típica (σ) de las puntuaciones obtenidas por los usuarios expertos sobre el cuestionario de la Tabla 20.	134
Tabla 22	Relación de parentesco de las fases en la dimensión Actividad	145
Tabla 23	Medidas sobre el cubo de costes	146
Tabla 24	Estructura de descomposición de trabajo (WBS) del proyecto de ejemplo correspondiente a una vivienda unifamiliar aislada.	172

INTRODUCCIÓN

1.1 ESPAÑOL

El sector de la construcción ha evolucionado en las últimas décadas de manera muy marcada, así como la cantidad de datos que se generan e intercambian en cualquier proyecto de edificación. Esos datos son necesarios para gestionar las diferentes tareas necesarias para cumplir con los objetivos del proyecto. El origen de estos datos es muy diverso debido principalmente a las características propias del sector, tales como, la cantidad de agentes que intervienen en el proceso, el carácter temporal de los proyectos y la diversidad de herramientas que se utilizan para generarlos. Por ejemplo, para almacenar los datos de los proyectos se suelen utilizar distintas bases de datos operacionales, documentos en formato digital, e incluso, documentos en soporte papel. Como consecuencia, los proyectos de edificación están asociados a conjuntos de datos muy grandes, irregulares y dispersos.

Sin embargo, el acceso a los datos de los proyectos y la gestión de los mismos son tareas fundamentales para el funcionamiento diario de las empresas del sector. En el caso de las empresas constructoras, el acceso adecuado a esos datos es esencial para alcanzar con éxito los objetivos principales del proyecto; cumplir con la finalización de un proyecto en el plazo establecido, cumplir con las especificaciones de diseño atendiendo a un cierto nivel de calidad y ajustarse al presupuesto establecido. Es por ello que el almacenamiento de los datos que se generan a lo largo del ciclo de vida del proyecto es una tarea de vital importancia en el sector de la construcción. Por un lado, para facilitar la operación diaria de la empresa en la ejecución de los proyectos y, por otro lado, para dar soporte a la toma de decisiones futuras en base a experiencias previas.

Para almacenar los datos de los proyectos, inicialmente se disponía de sistemas que los almacenaban en diferentes ubicaciones o almacenes de datos dispersos, o incluso en formato papel. Para solventar esto, se implantaron sistemas para la centralización de documentos. Estos sistemas han contribuido a mejorar el acceso e intercambio de dichos documentos pero, sin embargo, no son más que simples repositorios

que centralizan grandes cantidades de documentos. Además, la información almacenada en estos sistemas, a pesar de estar centralizada, es heterogénea, ambigua y se encuentra desestructurada debido a la propia naturaleza de los datos de los proyectos así como a la falta de estándares de referencia que regularicen dicha información.

Esta heterogeneidad se refleja notablemente en uno de los documentos más importantes del proyecto, el presupuesto. La información contenida en este documento aporta información de gran interés en el ámbito de la gestión del proyecto ya que, además de suministrar la información económica de la obra, establece la división y definición de las distintas actividades que se tienen que realizar para completar el objetivo del proyecto. Sin embargo, el actual proceso que se sigue para la elaboración de este documento, junto con la falta de estándares que definan una estructura común, hace que cada presupuesto sea diferente tanto en su estructura como en la redacción y definición de las actividades que contiene. De nuevo, la información contenida en este documento plantea ambigüedad y heterogeneidad provocando que la gestión de la información de los proyectos en base a los presupuestos sea ineficiente, además de complicar la toma de decisiones, aspecto fundamental durante el proceso de construcción.

Las Tecnologías de la Información y la Comunicación (TIC) y, en concreto, los Sistemas de Información juegan un papel fundamental para afrontar las necesidades de gestión de información en las organizaciones. En el sector de la construcción, la adopción de estas tecnologías se ha producido de manera más paulatina que en otros sectores, ya que un requisito fundamental para lograr el tratamiento automatizado de los datos con garantías, es disponer de datos estructurados y unificados en un almacén de datos centralizado. Es más, el éxito de un Sistema de Información integrado depende en gran medida del desarrollo de métodos de adquisición de datos, de modelos adecuados para representar la información, así como de almacenes de datos que permitan el almacenamiento y recuperación de dicha información. Sin embargo, como hemos mencionado, la información asociada a los proyectos de edificación y, en particular, al presupuesto, es ambigua, imprecisa y heterogénea. Por consiguiente impone grandes limitaciones en el desarrollo de estos modelos.

En este contexto, esta Tesis aborda un reto importante en cuanto al desarrollo de herramientas que permitan integrar los datos asociados al documento del presupuesto en un almacén de datos de manera estructurada. Por un lado, la extracción de información a partir de conjuntos de datos heterogéneos, eliminando particularidades y ambigüedades del lenguaje para su posterior integración en un almacén de datos centralizado es un aspecto determinante en este reto. Por otro lado,

el desarrollo de herramientas que permitan el tratamiento y acceso a la información de forma flexible a los usuarios que intervienen en la gestión del proyecto con el fin de dar soporte a la toma de decisiones.

1.1.1 *Objetivos*

El contexto anterior, nos ha impulsado a afrontar el reto de la construcción de un Sistema de Información que permita el tratamiento automatizado de los datos contenidos en uno de los documentos más importantes dentro de la gestión de los proyectos de edificación. En el marco de esta propuesta, se plantean los siguientes objetivos principales:

1. Proponer una estructura de referencia que pueda ser utilizada para la construcción de almacenes integrados de datos de documentos de presupuesto en proyectos de edificación.
2. Aportar una metodología que permita automatizar la adquisición de los datos contenidos en el documento del presupuesto y la clasificación de éstos en la citada estructura de referencia.
3. Desarrollar herramientas para manejar la información ya estructurada y unificada, tanto desde el punto de vista operacional como desde el punto de vista analítico, para facilitar la toma de decisiones en el ámbito de los proyectos de edificación.

Para el cumplimiento de los objetivos generales se plantean los siguientes sub-objetivos de forma más específica:

- a) Analizar las características del sector de la construcción, pues condicionan el funcionamiento del proceso de edificación y la gestión de la información que se genera durante el ciclo de vida de éste. Este estudio debe ser afrontado partiendo desde una perspectiva general, de modo que se puedan identificar las principales necesidades en la gestión de los proyectos.
- b) Realizar una revisión bibliográfica de las propuestas presentes en la literatura que incorporan el uso de las TIC para afrontar la gestión de los datos en las distintas tareas de la gestión de proyectos. Para ello se analizarán las propuestas tanto desde el punto de vista de las principales tecnologías que se han utilizados para permitir un manejo de los datos adecuado, como desde el punto de vista de las principales tareas dentro del ámbito de la gestión de proyectos de construcción. Se prestará especial atención a la identificación de retos abiertos en relación

con la gestión eficiente y automática de los datos asociados a los proyectos de edificación.

- c) Proponer un modelo formal que permita organizar la información del presupuesto para su posterior procesamiento y/o toma de decisiones. Para poder afrontar el objetivo del tratamiento automatizado de los datos contenidos en el dato del presupuesto con garantías, es un requisito fundamental disponer de datos estructurados y unificados. En la práctica actual, debido a la falta de estándares de referencia y a la forma de elaborar este documento, los datos contenidos en este documento están muy desestructurados.
- d) Instanciar el modelo anterior para proponer una estructura de referencia para la organización de documentos de presupuesto en España. Como se ha comentado, la falta de estándares es una de las principales razones que conducen a la diversidad de este documento, tanto en la forma de organizar las actividades como en el marco lingüístico utilizado para definirlos. La estructura de referencia considerará aspectos generales que tienen en común la mayoría de los presupuestos en España y permitirá la organización de la información de manera estructurada y homogénea.
- e) Proponer una metodología para la clasificación de la información extraída del documento del presupuesto en la citada estructura de referencia. Para facilitar a los usuarios la tarea de almacenamiento de presupuestos en un repositorio común para su posterior procesamiento y/o toma de decisiones, nos planteamos la propuesta de una metodología que permita realizar este proceso de manera casi automática.
- f) Desarrollar una herramienta de adquisición y edición de presupuestos amigable y fácil de utilizar, basada en las propuestas anteriores que pueda utilizarse como mecanismo de adquisición de datos en un repositorio integrado de proyectos de edificación.
- g) Hacer una propuesta que permita el uso de un repositorio de datos integrado para el soporte para la toma de decisiones en gestión de costes en proyectos de edificación. Para llevar a cabo dicha tarea con éxito se desarrollará un modelo de datos multidimensional para la gestión de datos relativos a costes, que permita analizar dichos datos desde diferentes perspectivas.

1.1.2 Estructura de la Tesis

Según los objetivos planteados, esta Tesis se organiza como sigue:

- Al cumplimiento de los objetivos a) y b) se dedica el **Capítulo 2** de la memoria. En primer lugar se detallan las principales características del sector de la construcción para poner en antecedentes al lector, aportándole una idea clara de los problemas actuales y reconocidos que existen en relación a la gestión de los datos en los proyectos de construcción. Posteriormente, se procede a analizar las aportaciones existentes en la literatura que incorporan el uso de las TIC para afrontar las principales tareas dentro de la gestión de los proyectos.
- En el **Capítulo 3** se abordan los objetivos c) y d). En primer lugar se propone un modelo formal para el almacenamiento integral y estructurado de los datos asociados al presupuesto de proyectos de edificación y, en segundo lugar, en base al modelo propuesto e integrando conocimiento experto se presenta una estructura de referencia para proyectos españoles, la cual permite organizar la información del presupuesto de forma estructurada.
- El **Capítulo 4** se dedica al objetivo e), en el que se propone un mecanismo de clasificación automática de las descripciones de partida contenidas en el documento del presupuesto dentro de una estructura común de referencia. Con este mecanismo se aborda la incorporación de los datos en un almacén común de manera integrada, independientemente de la herramienta con la que ha sido elaborado.
- Al cumplimiento de los objetivos f) y g) se dedica el **Capítulo 5**, en el que se presentan dos herramientas que permiten a los usuarios la gestión de los datos contenidos en el presupuesto de manera integrada y estructurada. Por un lado, se desarrolla una herramienta que permite extraer y almacenar datos del documento del presupuesto en un repositorio común de manera automática. Y por otro lado, se propone una estructura de datos multidimensional que permite el acceso de manera flexible e intuitivo a datos de costes para dar soporte a su procesamiento analítico en el proceso de toma de decisiones.
- Finalmente, en el **Capítulo 6** se recogen las principales conclusiones de la investigación realizada y las futuras líneas de investigación.

Además, se incluye el **Apéndice A**, en el que se presenta un manual de la aplicación *SIGPE* mediante la inserción de un proyecto real.

1.2 ENGLISH

The construction sector has significantly evolved in recent decades, in parallel with a dramatic increase in the amount of data generated and exchanged in any building project. These data are necessary to manage the different tasks needed to accomplish the objectives of the project. However, the origin of these data is very diverse, mainly due to the sector's characteristics, such as the number of participants involved in the process, the temporary nature of projects, and the diversity of tools used to generate them. For example, the storage of project data has often carried out by means of different operational databases, the use of documents in digital format, and even paper documents. As a result, construction projects are associated with large, irregular, and scattered datasets.

However, the access and management of project data are essential tasks to support daily operation in organizations. In the case of construction companies, proper data access plays a key role to ensure the success in project objectives: completing the project on time and within budget limitations while meeting the established quality requirements and other specifications. This is why the storage of data generated throughout the project life cycle is a task of a vital importance in the construction sector. On the one hand, to facilitate the daily operation of the company in the project execution and, additionally, to support the decision making process based on previous experiences.

Originally, project data were stored in different locations, in scattered repositories, and even in a paper format. To overcome this problem, systems that allowed a centralized management of documents were introduced. These systems have contributed to improve the access and exchange of those documents; however, they are merely repositories that centralize large amounts of documents. In addition, the information stored in these systems, despite being centralized, is heterogeneous and ambiguous due to the unstructured nature of project data and to the lack of reference standards to regularize such information.

This heterogeneity is notably present in one of the most important documents, the bill of quantities, which contains information of great interest in the construction domain. This document, in addition to providing economic information, establishes the division and definition of the different activities which need to be implemented in order to accomplish the project objectives. However, due to the current process to elaborate it along with the lack of standards to establish a common structure, each bill of quantities is different both as regards the structure and the definition of the activities. Once again, the information

contained in this document shows ambiguity and heterogeneity, which causes that information management is not efficient. Moreover, it complicates the decision making process, an essential aspect during the construction life cycle.

Information and Communication Technologies (ICT) and concretely, Information Systems, play a fundamental role to address the needs of information management in organizations. In the construction domain, the adoption of these technologies has been exploited more slowly than in other industries. The main reason is that an essential requirement to achieve an automatic data management is that data have to be structured and unified in a centralized repository. Moreover, the success of an integrated information system largely depends on the development of methods for data acquisition, suitable models to represent the information, and data warehouses that allow the storage and retrieval of such information. However, as we have mentioned, the information associated with building projects and, in particular, with the bill of quantities document, is ambiguous, imprecise, and heterogeneous. Therefore, it imposes major limitations in the development of these models.

In the previously presented context, this Thesis poses an important challenge regarding the development of tools to integrate data related to the bill of quantities document in a structured way in a common repository: on the one hand, the extraction of information from heterogeneous datasets by eliminating language ambiguities for further integration into a centralized repository is a key issue of this challenge, and on the other hand, the development of tools that allow the handling and access to information in a flexible way for users involved in the project management in order to support decision making.

1.2.1 *Objectives*

The previous context has encouraged us to face the challenge of building an information system that allows the automated treatment of the data contained in one of the most important documents in the management of building projects: the bill of quantities. In the framework of this proposal, the following objectives are posed:

1. Proposing a reference structure that can be used for the construction of integrated repositories of bill of quantities documents of building projects.

2. Providing a methodology for automating the acquisition of the data contained in the bill of quantities document and classifying them in the cited reference structure.
3. Developing tools for the management of the information when it is structured and unified, both from the operational and from the analytical point of view, in order to facilitate the decision making process in the building projects domain.

For the fulfillment of these objectives, we pose the following sub-objectives:

- a) Analysing the characteristics of the construction sector, since they determine the functioning of the building process and the management of the information generated during the project life cycle. This study should be addressed starting from an overall perspective, so as to identify the main needs in the project management.
- b) Reviewing the literature regarding the current proposals that incorporates the use of the ICT to address data management in the different tasks of the project management. To do this, we put the focus on the main ICT tools that have been used in project management research and, secondly, we analyze ICT-based proposals in relation to the project management tasks that they try to improve. We will pay special attention to the identification of open challenges in relation to the efficient and automatic data management associated with building projects.
- c) Proposing a general (formal) model for the organization of bill of quantities information for further processing and/or decision-making. To meet the goal of automated processing with guarantees, an essential requirement is to have structured and unified data. In current practice, due to the lack of reference standards and to the way in which this document is developed, the data are highly unstructured.
- d) Presenting a reference structure for organizing the Spanish bill of quantities document based on the previous general model. As mentioned before, the lack of standards is one of the main reasons that lead to the diversity as regards this document, both in the structure and in the linguistic framework used to define them. The reference structure will consider general aspects shared by most bill of quantities in Spain. Moreover, this structure will allow us to organize information in a structured and homogeneous way.

- e) Proposing a methodology for the classification of the information extracted from the bill of quantities document in the proposed reference structure. To facilitate users the storage of bill of quantities documents in a common repository, we will develop a methodology which allows the quasi-automatic execution of this process.
- f) Developing a user friendly tool for the acquisition and edition of bill of quantities documents based on the previous proposals, which can be used as a mechanism for data acquisition in an integrated repository of building projects.
- g) Developing a proposal to allow the use of an integrated data repository to support decision making regarding cost management in building projects. To successfully accomplish this task, we will present a multidimensional data model to manage cost data from several perspectives.

1.2.2 *Thesis structure*

According to the previous objectives, this Thesis is structured as follows:

- **Chapter 2** will be devoted to accomplish objectives a) and b) in the memory. Firstly, the characteristics of the construction sector are presented in order to identify the current and known problems in this sector. Subsequently, we will proceed to analyze existing literature contributions that incorporate the use of ICT to address the main tasks within the project management.
- **Chapter 3** proposes a general model for integrated and structured data storage associated with the bill of quantities document. In addition, a reference structure based on the proposed model and incorporating expert knowledge is proposed. This structure allows us to organize the information in Spanish bill of quantities documents in an integrated way.
- **Chapter 4** is devoted to introduce a mechanism for the quasi-automatic classification of work descriptions contained in the bill of quantities document in a reference structure. This mechanism addresses the incorporation of data in a common integrated repository, regardless of the tool that has been used to elaborate the document.

- **Chapter 5** is devoted to present two tools that allow users the management of data contained in the bill of quantities document in an integrated and unified way. Firstly, a tool which allows to extract and store bill of quantities data in a common repository is presented. Secondly, we propose a multidimensional data model that enables flexible and intuitive access to cost data so as to support analytical processing in the decision-making process.
- Finally, in **Chapter 6**, the main conclusions of the research and future lines of work are presented.

In addition, the **Appendix A** is included which presents the manual of the tool *SIGPE* by means of the insertion of a real project.

ANTECEDENTES Y PRELIMINARES

RESUMEN: Este capítulo pretende poner en antecedentes al lector de las características generales del sector de la construcción y del estado actual de la investigación en cuanto al uso de las Tecnologías de la Información y de la Comunicación (TIC) para proporcionar acceso a los datos y aprovechar dicho acceso para dar soporte a la toma de decisiones en las principales tareas de la gestión de los proyectos.

Para ello, en primer lugar se detallan las principales características del sector de la construcción que motivan la dificultad inherente a la gestión de los proyectos en dicho ámbito. A continuación, se realiza una revisión bibliográfica completa de las distintas propuestas existentes en la literatura en relación con la gestión de datos en los proyectos. Por último, a la luz de dicha revisión, se realiza un análisis crítico y se motiva la investigación realizada en esta memoria.

ÍNDICE

2.1	Características del sector	12
2.2	Pilares TIC en gestión de proyectos	15
2.2.1	Web, Web Semántica y Computación en la Nube	15
2.2.2	Sistemas de Información	18
2.2.2.1	Modelado de los datos a través de BIM	19
2.2.2.2	Sistemas de Información Geográfica	20
2.2.2.3	Almacenes de datos	21
2.2.2.4	Sistemas de Información Inteligentes	22
2.2.3	Tecnologías para el seguimiento y localización	23
2.3	Manejo de datos en Gestión de Proyectos	24
2.3.1	Información general y gestión del conocimiento	25
2.3.2	Costes	28
2.3.3	Planificación	31
2.3.4	Gestión de riesgos	34
2.3.5	Seguridad	36
2.3.6	Monitorización del proceso	39
2.3.7	Gestión de la calidad	41
2.4	Análisis y motivación de nuestra investigación	43
2.4.1	Análisis del estado actual	43
2.4.2	Necesidad de nuevos enfoques para la gestión de los datos	51

2.1 CARACTERÍSTICAS GENERALES DEL SECTOR DE LA CONSTRUCCIÓN.

La industria de la construcción tiene peculiaridades específicas que es necesario resaltar, pues condicionan su estructura y funcionamiento. Entre ellas cabe destacar las siguientes:

- Existe una separación física entre el lugar de ejecución del proyecto y la sede central de la empresa, o incluso entre dos proyectos de la misma empresa. Esta circunstancia dificulta la comunicación y transmisión de información entre los distintos puntos de trabajo [6, 164, 203].
- El proceso de ejecución es amplio [6], oscilando desde los seis o siete meses para proyectos de poca envergadura, hasta 2 o 3 años para proyectos más complejos, superando esta duración en algunos casos. Al tratarse de un proceso tan largo en el tiempo, es muy posible que ocurran acontecimientos que varíen el planteamiento inicial, como por ejemplo, cambios sociales, económicos, financieros, necesidades del promotor e, incluso, circunstancias derivadas de los agentes atmosféricos (por ejemplo, lluvias torrenciales, nevadas o vientos), cuyas consecuencias influyen directamente en el coste y duración final del proceso de ejecución y hacen que el mantenimiento de la consistencia de los datos relativos al proyecto sea de especial dificultad.
- La documentación del proyecto es muy extensa y compleja, ya que contiene gran cantidad de documentos de diferente naturaleza que definen la información necesaria para ejecutarlo [10, 12]. Las partes más relevantes de la documentación del proyecto se corresponden con los objetivos principales de cualquier proyecto: concluirlo en tiempo, con el presupuesto establecido, atendiendo a las especificaciones técnicas y cumpliendo los requisitos de calidad. En la actualidad, cada uno de estos documentos tiene un formato distinto y se elabora con la ayuda de distintas herramientas software, lo que dificulta sobremanera su intercambio y la interoperabilidad entre partes responsables de la ejecución y gestión del proyecto.
- En la industria de la construcción, y más concretamente en edificación, cada proyecto es diferente, único y singular [15, 39, 164, 203]. Desde el punto de vista del diseño es difícil encontrar dos edificios iguales y, aun suponiendo dos proyectos idénticos, existe otra serie de circunstancias de tipo operativo que diferencian la forma de ejecutarlos, lo que hace que sean procesos de

ejecución distintos. Por ejemplo, no es igual el planteamiento de un proyecto de vivienda unifamiliar, cuando el emplazamiento es en la costa, que si ese mismo proyecto se construye en la montaña; y, a su vez en cualquiera de estos dos planteamientos, tampoco será igual si el solar se encuentra en el centro de la ciudad o en el extrarradio.

Sin embargo, es cierto que, aunque cada proyecto es único, se pueden encontrar muchas similitudes entre ellos en el proceso de producción, ya que el producto final del sistema productivo, el edificio, no es más que una suma de procesos, que individualmente sí son repetitivos. A lo citado sobre el emplazamiento, hemos de añadir otras cuestiones como los recursos que puedan ser utilizados, así como los medios auxiliares, tanto por disponibilidad, como por los posibles accesos al lugar de producción.

Hemos de mencionar también en este apartado, por tener relación con lo anteriormente expuesto, por una parte las relaciones con distintos proveedores y, por otra, la mano de obra. Ambas, además de depender de la ubicación, están supeditadas a su vez al momento de la obra y a los trabajos que se estén realizando, no siendo constantes en el tiempo. Los distintos oficios aparecen y desaparecen en el transcurso del proceso muchas veces. La cantidad de datos generada es ingente y el análisis certero de los mismos una necesidad para los responsables de tomar decisiones en la organización o empresa.

- El fenómeno de la subcontratación, que acrecienta la temporalidad de los trabajadores en una industria cuyas instalaciones son cien por cien temporales, dando lugar a una gran cantidad de contratos de corta duración. Teniendo en cuenta que se tiende a empresas altamente especializadas en la industria de la construcción, según un estudio realizado por la Universidad de Granada [167] la mayor parte de las obras se realizan con porcentajes superiores al 50 % de subcontratación en distintos niveles, lo que dificulta la transmisión y la comunicación de información. Además, esta circunstancia provoca que la cadena de mando se vaya difuminando, siendo un factor decisivo en el éxito del proyecto que la cadena esté perfectamente clara y establecida. La presencia de distintas empresas en la ejecución complica tanto la captación de datos como su intercambio entre partes.

En general, como consecuencia de estas singulares características, la gestión del proceso de construcción es compleja, ya que implica una gran cantidad de participantes que generan e intercambian una

gran cantidad de datos durante todo el proceso de construcción [221]. Además, estos datos tienen formatos muy diversos y están almacenados en distintos *almacenes*, en muchos casos en papel. Esta circunstancia hace que la gestión de los datos resulte muy complicada [13].

El ciclo de vida de un proyecto de construcción se puede dividir en tres fases principales: diseño, ejecución y mantenimiento. Aunque cada fase se gestiona de manera independiente y tiene sus propios datos, muchos de ellos se comparten a través de las diferentes fases. La gestión efectiva de dicha información es esencial para alcanzar con éxito los objetivos del proyecto; concluirlo en tiempo, dentro de presupuesto establecido, atendiendo a las especificaciones técnicas y cumpliendo los requisitos de calidad. Por lo tanto, el almacenamiento integrado de los datos es fundamental para facilitar el acceso a éstos, con el objetivo de dar soporte a la toma de decisiones durante todo el ciclo de vida del proyecto [5, 127]. La mayoría de la información se intercambia en documentos desestructurados, debido a la falta de estandarización tanto en los documentos como en los datos [216], complicando la tarea de almacenamiento [10, 38]. Además, desde el punto de vista de la gestión de los datos, este hecho produce otra serie de problemas, tales como, dificultad para la recuperación de datos, escasa interoperabilidad entre sistemas y dificultad para reutilizar la información [10, 98].

Desafortunadamente, la información no sólo se intercambia a través de los documentos, sino que muchos datos se intercambian de forma verbal entre los distintos participantes. Por lo tanto, las experiencias y el conocimiento que se generan durante el proceso de construcción se pierden o permanecen en la mente de los ingenieros y no se almacenan en ningún sistema [52, 98]. Como hemos comentado previamente, el personal es muy variable, por lo que, cuando los ingenieros terminan los proyectos o dejan la empresa, se llevan el conocimiento con ellos. Este conocimiento es un activo muy valioso que podría beneficiar a situaciones similares en futuros proyectos de la empresa. Si este conocimiento estuviera correctamente almacenado, se podría reutilizar y compartir con el fin de mejorar los procesos de construcción, así como de disminuir la duración y el coste para solucionar los problemas [130].

Para afrontar de manera eficaz la gestión de los datos del proyecto, el uso de las Tecnologías de la Información y la Telecomunicación ha sido fundamental en los últimos años, ya que permiten una mejor colaboración, coordinación e intercambio de información entre los distintos participantes que intervienen en el proceso. A pesar del reconocido potencial de las TIC en otros sectores, en el ámbito de la construcción la aplicación de estas tecnologías se ha producido de forma mucho más paulatina [8, 148, 208]. Además de por las características propias

del sector de la construcción que se han detallado en esta sección, existen otras, como la falta de conocimiento y habilidades para usar las TIC [29, 164, 208, 232], el carácter tradicional de la industria de la construcción [101], y la falta de conocimiento respecto a la rentabilidad a corto plazo de la inversión [208].

A pesar de estas dificultades, en los últimos años se han diseñado y desarrollado muchas propuestas que utilizan las TIC para facilitar la gestión de los proyectos mediante el adecuado acceso a los datos. A continuación analizamos las principales tecnologías que permiten dicha gestión.

2.2 LOS PILARES DE LAS TIC PARA LA GESTIÓN DE PROYECTOS DE CONSTRUCCIÓN

Como hemos comentado, la fragmentación, la falta de integración, y la complejidad de los procesos y actividades son inconvenientes que están presentes en los proyectos de construcción. Estas características tienen un impacto negativo en la consecución del proyecto y, por tanto, afectan de forma determinante en el logro de los objetivos principales de éste.

En este difícil contexto, mejorar la colaboración y la comunicación entre las distintas partes y participantes es un tema crucial, ya que mejora la productividad, además de maximizar la eficiencia y la eficacia. Hoy en día, el acceso a la información precisa y actualizada se ha convertido en una necesidad en el entorno de cualquier empresa y, en este sentido, se han desarrollado muchas herramientas que permiten la gestión de los datos de manera adecuada, como se refleja en la literatura del ámbito de los Sistemas de Información. A continuación, analizamos las principales tecnologías que se han utilizado en la gestión de los proyectos de construcción con el fin de permitir la gestión adecuada de los datos.

2.2.1 *Web, Web Semántica y Computación en la Nube*

Para dar soporte a la mencionada necesidad de colaboración y comunicación en el entorno de las empresas de construcción, los sistemas informáticos permiten compartir, de manera efectiva y en tiempo, la información relativa a los proyectos entre los miembros de la empresa que están geográficamente dispersos.

En la implementación de dichos sistemas, Internet es una tecnología esencial para proporcionar un entorno de trabajo colaborativo. En este sentido, el uso de la tecnología web para la gestión de los proyectos en construcción se puede observar desde que comenzaron a utilizarse las TIC hasta hoy día, ya que la mayoría de las aplicaciones actuales dependen de la tecnología web, tanto en la etapa de diseño, como en la etapa de ejecución [164]. Esta tecnología proporciona una plataforma adecuada que permite la gestión e intercambio de información de forma efectiva [52]. Además, el uso de esta tecnología aporta otros beneficios, tales como: la reducción del número de documentos en papel, la mejora del acceso a los datos permitiendo compartir documentos entre participantes en diferentes localizaciones, la disminución de discrepancias y malentendidos derivados de las distintas versiones de los documentos y la posibilidad de registrar los datos en formato multimedia.

Desafortunadamente, en la mayoría de las ocasiones, la interoperabilidad entre sistemas y el intercambio de información entre los distintos participantes sólo está disponible a nivel de archivos, archivos que normalmente presentan formatos privados con escasa información semántica [204]. Muchos sistemas de gestión basados en la web están diseñados para la gestión de documentos, siendo simples repositorios comunes para documentos de diferente naturaleza. El principal problema de estos sistemas es la cantidad de tiempo que hay que destinar para localizar la información que se necesita para tomar decisiones, ya que ésta se encuentra oculta y dispersa en el enredo de información contenida en los documentos [51]. Además, la diversidad de formatos en los distintos documentos que se necesitan a lo largo del proceso (planos, cálculos, planificación, presupuesto) también dificulta el acceso a los mismos puesto que, para poder consultarlos, los usuarios tienen que acceder a través de aplicaciones adecuadas en cada caso.

Esta diversidad conduce al inconveniente más acuciante en las herramientas de colaboración basadas en la web: la ausencia de formatos estándar [114] y la falta de representación semántica, tanto de conocimientos como de procesos [89, 256].

En este sentido, la web semántica [30] ofrece tecnologías que contribuyen a la solución del problema de interoperabilidad en tres niveles: físico, sintáctico y semántico. A diferencia de la web tradicional, la web semántica tiene como objetivo el intercambio de "significados". Este intercambio de conocimiento entre las personas y los ordenadores se realiza a través del uso de ontologías [88, 229].

Dentro del marco de la web semántica, las ontologías son las que permiten a las aplicaciones describir el vocabulario y el conocimiento

para permitir la comunicación. El uso de ontologías compartidas entre distintas aplicaciones facilita el intercambio de datos e información a nivel semántico y, por tanto, permite la interoperabilidad semántica [90]. La interoperabilidad semántica proporciona la integración en un nivel superior, permitiendo que los datos sean compartidos y que se puedan entender a través de distintas disciplinas de diseño y sistemas software heterogéneos como, por ejemplo, los sistemas de modelado de información de construcción (en inglés Building Information Modeling, BIM) y los sistemas de información geográfica (en inglés Geographic Information Systems, GIS) [133, 134], que repasaremos más adelante. Por desgracia, no existe disponible una ontología única que abarque todo el dominio de Arquitectura, Ingeniería y Construcción (AIC, AEC en inglés) [91, 133, 152]. Afortunadamente, y en reconocimiento a la importancia de las ontologías para alcanzar el trabajo colaborativo, se puede observar un incremento en el número de estudios sobre ontologías, como veremos a lo largo de la revisión de la literatura.

Un concepto nuevo, que surge a partir de la web semántica es el uso de *datos enlazados o interconectados* (en inglés, linked data) [74, 133]. Este concepto ayuda a superar las dificultades de interoperabilidad mencionadas para mejorar el intercambio de información en el dominio AEC. Tim Berners-Lee [31] propone un enfoque que trata de interconectar los datos estructurados distribuidos en la web aprovechando los protocolos abiertos y estándares del W3C [240]. Los enlaces se establecen en el nivel de información (datos) y no en el nivel de infraestructura (sistema), prestando atención a las similitudes de conceptos de la información. Esta tecnología se puede adaptar fácilmente a la infraestructura de información existente (bases de datos relacionales, almacenes de datos, etc.), con el objetivo de vincular o interconectar los datos que se quieren compartir [74].

Otra tecnología que se ha establecido recientemente alrededor de Internet es la Computación en la Nube (en inglés, Cloud Computing). La nube, computación en la nube o informática en la nube son algunos de los términos para referirse a esta tecnología. El Instituto Nacional de Estándares y Tecnología [186] americano define este concepto como "Un modelo tecnológico que permite el acceso en red ubicuo, apropiado y bajo demanda, a un conjunto compartido de recursos de computación configurables (por ejemplo: redes, servidores, almacenamiento, aplicaciones y servicios), que pueden ser rápidamente asignados y liberados con un esfuerzo de gestión reducido o con interacción mínima con el proveedor del servicio". Esta tecnología permite la comunicación e intercambio de datos e información de manera sencilla, eficiente, efectiva y precisa, independientemente de la ubicación y del tiempo. Por estas razones, puede aportar grandes beneficios para la gestión de los datos

en el ámbito de la Construcción [97, 128, 201]. En los últimos años, se han desarrollado una gran cantidad de aplicaciones específicas para el sector de la construcción haciendo uso de esta emergente tecnología, tales como, las relacionadas con la tecnología BIM antes mencionada. Por ejemplo, Autodesk BIM 360 [17], BIMServer [228] y BIM9 [32] son claros ejemplos de aplicaciones en la nube. En [63], Chong et al. revisan un total de cuarenta y dos aplicaciones, basadas en la computación en la nube, para la gestión de los proyectos.

Al igual que en los sistemas tradicionales de gestión, la interoperabilidad es una necesidad básica que se debe tener en cuenta desde el diseño de los nuevos servicios de datos basados en la nube, ya que si se diseñan de la misma forma que los sistemas tradicionales, existirán los mismos problemas de intercambio de datos [74].

Para lograr con éxito una eficiente y eficaz colaboración, la integración de sistemas es un requisito importante [214], y es de especial interés en relación a los sistemas de información.

2.2.2 *Sistemas de Información*

Los sistemas de información permiten recopilar, procesar, almacenar y distribuir la información con el objetivo de dar soporte a las empresas en su actividad diaria y en la toma de decisiones. En la actualidad, los sistemas de información son el núcleo de los sistemas informáticos de las empresas, también en la mayoría de las empresas de construcción, que cada vez son más conscientes de la necesidad de disponer sistemas de información que les permitan obtener beneficio de la información que se genera en sus operaciones diarias.

La calidad de un sistema de información de una empresa se basa en buena medida en el uso de modelos adecuados para representar la información, en la presencia de almacenes integrados que faciliten el almacenamiento y la recuperación de datos, y en el uso de técnicas de análisis de datos apropiadas para dar soporte a la toma de decisiones. Afortunadamente:

- En el ámbito de la construcción hay una creciente preocupación por la necesidad de modelos de datos adecuados que permitan la representación de la información de manera transversal y compartida a través de todo el proceso de construcción. En este sentido, la tecnología BIM es el ejemplo más claro.
- Los Sistemas de Información Geográfica (GIS) son cada vez más frecuentes en el ámbito de la construcción, ya que ofrecen la re-

ferenciación espacial de información y su posterior tratamiento, visualización y análisis.

- Los *almacenes de datos* (en inglés, Data Warehouses DW), junto con las tecnologías convencionales de bases de datos, hacen posible el diseño y la construcción de repositorios que permiten la gestión integrada de la información y cada vez están más presentes en el ámbito de la construcción.
- Las técnicas de *Extracción de Conocimiento en Bases de Datos* (en inglés, *Knowledge Discovery in Databases, KDD*) o, dentro de estas, las relativas a *Minería de datos* (en inglés, *Data Mining*), permiten explorar grandes cantidades de datos con el fin de encontrar información útil y valiosa para la empresa. Estas técnicas son cada vez más utilizadas por los profesionales de la construcción con el fin de obtener una mejor comprensión del proceso de construcción y mejorar el funcionamiento de la empresa.

A continuación revisamos las propuestas que se encuentran en la literatura basadas en estos pilares de los sistemas de información.

2.2.2.1 Modelado de los datos a través de BIM

BIM es una metodología de trabajo integrada que se apoya, no en una, sino en múltiples herramientas basadas en tecnologías de la información, proporcionando un marco de colaboración dentro de un entorno multidisciplinar como es el dominio de AEC. Por tanto, la principal ventaja que aporta BIM es la mejora de la comunicación y comprensión del proyecto a través de la información relacionada con el modelo 3D (como por ejemplo, geometría, relaciones espaciales, información geográfica, cantidades o propiedades de los componentes del edificio, entre otras) durante todo el ciclo de vida del proyecto [29, 230].

En la actualidad, el uso de BIM está más extendido en las primeras fases del proyecto, siendo su uso considerablemente menor en las fases de ejecución y de mantenimiento. En la fase de diseño, la tecnología BIM permite: integración de diseño, preconstrucción virtual, identificación y análisis de los posibles conflictos de diseño, simulación, estimación de costes, programación, acceso distribuido, recuperación y mantenimiento de los datos de construcción. A lo largo de los últimos años, la tecnología BIM ha evolucionado para integrar diferentes aspectos que son necesarios para el resto de etapas del proyecto [14, 83, 100]. Sin embargo, en la actualidad, su uso se limita a un repositorio estático de información [74, 87].

El modelado BIM añade nuevas dimensiones de información al modelo 3D, conforme se incorpora la dimensión tiempo (4D), los costes (5D), la sostenibilidad (6D) y la operación y mantenimiento (7D).

Una vez que el modelo está completo, los datos del edificio se pueden añadir a otros para racionalizar el proceso de construcción, como por ejemplo, los datos procedentes de GIS [185] o los datos procedentes de tecnologías que permiten el seguimiento del proceso [34, 107, 136], ofreciendo beneficios a la gestión de los proyectos de construcción.

Desafortunadamente, aunque se espera que el modelado BIM revolucione la forma de intercambiar la información de los proyectos de construcción, todavía se encuentra en una fase muy temprana [43]. En la actualidad existen barreras que frenan su implantación tales como, la falta de interoperabilidad entre aplicaciones para permitir el flujo de información, el cambio considerable que significa en la forma de trabajar de los usuarios, la falta de conocimiento especializado de éstos y los requisitos en cuanto a recursos informáticos que sean capaces de soportar los modelos BIM [4, 87, 145].

2.2.2.2 Sistemas de Información Geográfica

Los sistemas de información geográfica (GIS) han sido ampliamente utilizados en la última década [79]. Estos sistemas aportan un conjunto de herramientas que integra y relaciona diversos componentes (hardware, software y procesos), permitiendo la captura, el almacenamiento, la manipulación, el análisis y la actualización de datos que están vinculados a una referencia espacial. Su gestión comprende tanto los datos espaciales como los atributos de la información, de modo que pueden ser extraídos, analizados y visualizados [213].

Los sistemas de información geográfica son un buen complemento para ampliar las capacidades de representación de los sistemas basados en CAD y BIM, ganando en la suma, ya que una implementación en GIS nunca será tan detallada, ni tan rica, semánticamente como un modelo BIM. Además, se pueden usar formatos de CAD o BIM para crear un sistema de referencias geográficas más completo [79]. La integración de BIM y GIS puede aportar muchos beneficios para la gestión de los proyectos, ya que el modelado BIM se centra en el desarrollo de objetos con un nivel de detalle muy elevado en la geometría y en los espacios interiores, mientras que la tecnología GIS permite el análisis de los objetos existentes en el entorno de una manera más abstracta [134]. Para conseguir que esta integración sea satisfactoria, los datos se tienen que intercambiar de un sistema a otro de manera efectiva

y esto no es tarea fácil ya que los esquemas de cada sistema están diseñados y creados para objetivos diferentes [55]. Como veremos más adelante, para conseguir la integración total entre estas dos tecnologías, es necesario proporcionar interoperabilidad en el nivel semántico.

Recientemente, se ha observado un incremento en la utilización de herramientas que permiten la simulación y la visualización de proyectos, ya que proporcionan a los usuarios la oportunidad de entender el proceso con antelación a su construcción, con el fin de predecir futuros acontecimientos sin tener que hacer ensayos reales o sin interferir en las operaciones que ya están en ejecución [28]. Estas herramientas son de gran utilidad para los responsables en tomar decisiones en el ámbito de la construcción, ya que, a diferencia de otros sectores, el proceso de construcción no es trasladable a un laboratorio, no se puede construir y destruir para comparar resultados. Aun pudiéndose, no se podría afirmar en muchos casos con rotundidad cuál sería la mejor solución, pues habría cambiado el tiempo, el lugar, los recursos, la mano de obra, o lo que es aún más habitual, todos ellos.

2.2.2.3 Almacenes de datos

Como se ha comentado con anterioridad, a lo largo de todo el proceso de construcción se generan grandes cantidades de datos, por lo que es necesario disponer de un almacén de datos consolidados que permita un acceso fácil y amigable para los usuarios con el fin de dar soporte para la toma de decisiones. En el ámbito de las TIC existen tecnologías que permiten la construcción de tales repositorios, destacando entre ellas la de los almacenes de datos [141]. Los almacenes de datos (del inglés, *Data Warehouse*, DW) se caracterizan por integrar la información que proviene de distintas fuentes, de forma que permiten mantener una visión coherente e integrada de la información interna a la empresa. Según Inmon [118], un almacén de datos es una colección no volátil de datos orientados a un tema, integrados y variables en el tiempo para ser usada como apoyo a la toma de decisiones.

La construcción de almacenes de datos implica en sí misma la existencia de funciones que den soporte a la creación y mantenimiento de acceso a los usuarios finales para obtener datos completos y consistentes de la empresa. Aquí es donde surge el concepto de procesamiento analítico en línea (del inglés, *On-Line Analytical Processing*, OLAP). La tecnología OLAP [187] es una categoría de tecnología software que permite a los analistas, directivos y ejecutivos acceder a los datos de forma rápida, consistente e interactiva a través de una amplia variedad de vistas de la información, que han sido obtenidas a partir de datos

operacionales para reflejar la dimensionalidad real de la empresa como la entiende el usuario.

La capacidad de análisis OLAP, combinada con la capacidad de almacenamiento de datos, proporciona una oportunidad para mejorar la gestión de los datos en el ámbito de la gestión de proyectos de construcción. Esta combinación permite realizar consultas analíticas complejas y ad-hoc basadas en datos históricos en un tiempo reducido de ejecución, para facilitar el proceso de toma de decisiones.

2.2.2.4 Sistemas de Información Inteligentes

Como complemento de las capacidades analíticas y de resumen que proporcionan las herramientas que acabamos de ver, resultan muy interesantes las técnicas de Minería de datos (del inglés, *Data Mining*, DM) que permiten extraer el conocimiento de grandes cantidades de datos, proporcionando un análisis de los datos desde otra perspectiva. Existen muchas técnicas de minería de datos que involucran áreas como la Inteligencia Artificial, el Aprendizaje Automático, la Estadística y los Sistemas de Bases de Datos [215].

La comunidad de la industria de la construcción está adoptando cada vez más el uso de sistemas de información inteligentes junto con las mencionadas técnicas de minería de datos. En este sentido, técnicas como el razonamiento basado en casos (del inglés, *Case-Based Reasoning*, CBR) y la lógica difusa (*Fuzzy Logic*) presentan grandes beneficios para la gestión de los datos.

El proceso de CBR [2] no es más que otro paradigma de resolución de consiste en la recuperación de casos previos similares al problema actual que se pretende resolver para luego configurar una solución a dicho problema adaptando las soluciones utilizadas en esos casos similares.

Por último, debemos cita la lógica difusa, propuesta por Zadeh [251], que empezó a ganar popularidad en el área de construcción durante la década de los ochenta [18]. La aplicación de técnicas difusas proporciona una buena herramienta para gestionar la incertidumbre y la subjetividad que surgen durante el proceso de construcción. La fortaleza de la lógica difusa para modelar la incertidumbre en los datos se puede combinar con otras técnicas como las redes neuronales (*Neural Networks*, NN), más orientadas al reconocimiento de patrones y el aprendizaje automático [50].

2.2.3 Tecnologías para el seguimiento y localización

Las tecnologías para el seguimiento y localización (del inglés, *Tracking and location Technologies*, TT) permiten la adquisición de los datos que se generan durante el proceso de construcción de manera automática y en tiempo real [41, 124], mejorando de esta forma la integración de la información y la colaboración y comunicación entre los distintos participantes [208]. En los últimos años se ha investigado mucho sobre estas tecnologías y sus aplicaciones en el sector de la construcción [103, 182].

El sistema de posicionamiento global (del inglés, *Global Positioning System*, GPS) es un ejemplo claro de estas tecnologías, que permite determinar la posición de un objeto con mucha precisión. Este sistema funciona mediante una red de satélites en órbita y un receptor, que recibe las señales de los satélites indicando la posición mediante el método de triangulación inversa. Sin embargo, debido a las características del sector de la construcción, el GPS se debe complementar con otras tecnologías ya que los recursos en la obra se encuentran tanto en espacios interiores como exteriores [165] y las condiciones atmosféricas y la ubicación de los satélites sobre el receptor influyen en la precisión de la posición resultante [42].

También hay que mencionar la tecnología de identificación por radiofrecuencia (del inglés, *Radio Frequency Identification*, RFID), que engloba a una rama de las tecnologías de identificación automática en el que las ondas de radio se utilizan para capturar y transmitir datos mediante desde etiquetas o transpondedores. La tecnología RFID permite la captura de datos de manera muy eficiente porque los lectores se pueden fijar en cualquier soporte dentro del rango de lectura, no requieren visión directa entre el emisor y el receptor y, además, cada lector puede escanear varias etiquetas al mismo tiempo [191]. Desafortunadamente, el rango de comunicación de esta tecnología es reducido, por lo que resulta inadecuado para obras de construcción de gran extensión. Para conseguir un seguimiento completo en tiempo real del proceso de construcción, se puede combinar esta tecnología con GPS [231].

Otro tipo de tecnología de radio que permite comunicación en distancias cortas es la tecnología de banda ultra ancha (del inglés, *Ultra WideBand*, UWB), que utiliza un ancho de banda mayor y una frecuencia por encima de 500 MHz. UWB puede generar localizaciones en 2D muy precisas a través de tres o más receptores y, en 3D con cuatro o más. Estos sistemas pueden localizar e identificar múltiples objetivos al mismo tiempo, por lo que resultan de interés para el sector de la construcción [247].

Las tecnologías de visión y láser están captando mayor atención en las obras de gran extensión y en aquellas que están muy congestionadas, ya que estas tecnologías no necesitan etiquetas. La tecnología más popular en el ámbito de la construcción es LADAR (del inglés, *Laser Detection and Ranging*) que permite determinar la distancia desde el emisor láser a un objeto o superficie utilizando un haz láser pulsado. La distancia al objeto se determina midiendo el tiempo de retraso entre la emisión del pulso y su detección a través de la señal reflejada, consiguiendo datos en 3D muy precisos con pocas exploraciones. Como veremos más adelante, la mayoría de las aplicaciones que utilizan esta tecnología se dedican a la construcción de modelos del estado actual de la ejecución de la obra (en inglés, “as-built”).

En los últimos años, se ha producido un incremento en cuanto a la captura y transmisión de información gráfica relativa al proceso de ejecución debido a la facilidad de adquirir cámaras de alta resolución, al acceso a dispositivos de almacenamiento masivo a precio muy económico y a la disponibilidad de conexión a Internet en las obras [107]. Las fotografías se han convertido en un medio de comunicación muy utilizado, ya que contienen información real sobre la geometría y apariencia del proyecto de una forma económica y rápida [82]. Cuando se requiere información en tres dimensiones, se puede utilizar la tecnología analizada en el párrafo anterior, el láser [33, 136].

Todas estas tecnologías permiten el seguimiento de las operaciones del proceso de construcción de una manera eficiente y precisa, que puede aportar información en tiempo real de tareas específicas del proyecto [106, 107], así como dar soporte para la toma de decisiones a través de los datos adquiridos con estas herramientas [34, 248].

2.3 MANEJO DE DATOS EN LA GESTIÓN DE PROYECTOS DE CONSTRUCCIÓN

En la sección anterior hemos analizado las propuestas de la literatura desde el punto de vista de las tecnologías que permiten una gestión de los datos adecuada. En esta sección, vamos a analizar las propuestas que utilizan estas tecnologías desde el punto de vista de las principales tareas dentro del ámbito de la gestión de proyectos de construcción. Para este propósito, en primer lugar analizaremos las propuestas que se centran en la gestión de la información y conocimiento desde un punto de vista más general y, a continuación, procederemos con el análisis de las propuestas específicas de cada una de las principales tareas.

2.3.1 *Información general y gestión del conocimiento*

Como hemos mencionado anteriormente, una de las principales razones por las que el sector de la construcción ha sido más lento que otros sectores en aprovechar las ventajas de las TIC es la inadecuada e ineficiente coordinación y comunicación de la información de los proyectos construcción. En la mayoría de los casos, el intercambio de información se produce de forma verbal o a través de documentos de texto, por lo que la información no se almacena, a pesar de la importancia que podrían tener estos datos para la gestión del proyecto o para futuros proyectos si estuvieran correctamente almacenados.

Desafortunadamente, la información se encuentra almacenada dentro de complejos documentos desestructurados [10, 12, 40, 216] y con diversos formatos. Las fuentes de datos son muy diversas: documentos de texto (como contratos, especificaciones, presupuesto, pedidos, albaranes, actas, etc.), datos multimedia semi-estructurados (por ejemplo, documentación gráfica en 2D o 3D, archivos de audio o video) y datos estructurados para aplicaciones específicas que necesitan estructuras de datos más complejas (como por ejemplo, documentos de planificación basados en red).

Los datos tienen que estar estructurados para poder acceder a ellos y poder reutilizarlos y enriquecerlos durante todo el ciclo de vida del proyecto. Para este propósito, hay varias propuestas en la literatura, por ejemplo, organizando de manera automática los documentos en base a una estructura de clasificación preestablecida [39, 40, 98, 236] o en base a similitudes semánticas [10, 11, 198] utilizando técnicas de extracción de conocimiento. Sin embargo, la comprensión del contexto y la extracción de conceptos y relaciones semánticas de un conjunto de documentos de texto siguen siendo uno de los principales retos de la Computación, en particular en dominios tan específicos como la gestión de documentos de construcción [96].

Además de estas propuestas de clasificación de documentos, hay propuestas que tratan de organizar los datos contenidos en los documentos. En este sentido, los autores Chan and Leung [51] recuperan los datos a partir del documento original, reorganizan la información en función de tareas o usuarios específicos y permiten el acceso a la información a través de una página web.

Disponer de los datos de manera estructurada permite, entre otras cosas, tener un repositorio de datos integrados y consolidados en forma de almacén de datos, en el que se puede acceder a los datos de manera amigable para facilitar la toma de decisiones. En esta línea, se pueden encontrar los trabajos de los siguientes autores en la literatura:

- Chau et al. [53] desarrollan un sistema de ayuda a la toma de decisiones que utiliza el modelo multidimensional para facilitar la gestión de los datos referentes al inventario.
- Ahmad et al. [7] desarrollan un sistema para determinar el solar más adecuado para la construcción de casas residenciales. Los autores consideran tanto factores espaciales como de negocio, incorporando capacidades GIS y una herramienta analítica al almacén de datos.
- Fan et al. [94] proponen un almacén de datos para equipos vinculado a un sistema de soporte a la toma de decisiones basado en la web, que integra datos de diversas fuentes y asesora a los usuarios para obtener información valiosa sobre la gestión de equipos.
- Rujirayanyong y Shi [207] desarrollan un DW orientado a proyectos para gestionar los datos de la empresa, proporcionando una herramienta robusta para recolectar, almacenar y utilizar los datos históricos.

Además de la tecnología de almacenes de datos, varios autores proponen la implementación de sistemas de gestión de conocimiento, que hacen posible el intercambio y reutilización de la información.

- Tserng y Lin [225] señalan que durante la fase de ejecución de los proyectos, una forma eficaz de mejorar la gestión de la construcción es compartir experiencias entre los ingenieros. El sistema que proponen, basado en la web, permite a los usuarios acceder de forma fácil al conocimiento disponible de una cierta actividad deseada. También proporciona enlaces a otras actividades similares de proyectos anteriores.
- Lin et al. [161] proponen un sistema basado en mapas de conocimiento que permite el intercambio de conocimiento entre proyectos. El sistema permite capturar y gestionar el conocimiento durante la fase de ejecución, a través de la web, con el objetivo de beneficiar proyectos posteriores.
- Kivrak et al. [142] desarrollan un sistema basado en web que permite la captura de conocimiento tácito (por ejemplo, recomendaciones de expertos), así como el conocimiento explícito (por ejemplo, documentos, informes) que pueden ser útiles para futuros proyectos.
- Lin [158] presenta un enfoque que usa mapas basados en personas (people-based maps, PBMs) que permite capturar y representar la experiencia de los ingenieros y el conocimiento técnico

durante la fase de construcción utilizando la tecnología web. Este sistema permite el intercambio dinámico de las experiencias e incorpora servicios para poder reutilizarlas en proyectos posteriores.

- Tan et al. [220] proponen una metodología para capturar y reutilizar el conocimiento a través de la tecnología basada en la web. El sistema pretende reducir el problema de comunicación entre las distintas oficinas de la empresa que se encuentran dispersas geográficamente, facilitando el intercambio tanto conocimiento tácito y explícito en el momento oportuno.
- Lin y Lee [160] proponen una metodología nueva y práctica para capturar y representar el conocimiento del proyecto de construcción mediante el uso de comunidades de práctica basadas en proyectos (del inglés, *project-based communities of practice*, COPs) y la tecnología web, proporcionando foros para que los participantes se ayuden mutuamente en la solución de problemas.
- Dib et al. [79] presentan un modelo de base de datos integrado con GIS que permite el uso efectivo de las tecnologías informáticas para la comunicación, la documentación del proyecto y el intercambio de conocimientos durante todo el ciclo de vida del proyecto.
- Lin [159] y Deshpande et al. [78] introducen una metodología nueva y práctica para capturar y representar el conocimiento de los proyectos de construcción, tanto en la fase de diseño como de ejecución mediante el uso de BIM.
- Antony Chettupuzha y Haas [12] proponen una metodología que, mediante la utilización de técnicas de minería de datos, identifica los documentos que pueden interrumpir un proceso dando lugar a retrasos e incrementos en los costes.
- Finalmente, muchas propuestas presentan herramientas basadas en el uso de la web semántica y la web social para la gestión de la información y el conocimiento [88, 89, 91, 157, 190, 232, 256].

En la actualidad, la mayoría de las propuestas de gestión de conocimiento en las empresas de construcción utilizan las TIC [15]. Como se refleja en la literatura, hay muchas propuestas que abordan la gestión del conocimiento con la tecnología web, con el objetivo principal de facilitar la reutilización de éste, tanto en los proyectos que se están ejecutando como en posteriores. Sin embargo, para lograr todos los beneficios que aportan estos sistemas, la empresa tiene que crear una

cultura de intercambio de conocimiento que anime a todo el personal a compartir su propio conocimiento.

En esta sección, hemos revisado algunas propuestas importantes para la gestión de información y conocimiento en el dominio de la construcción desde un punto de vista general. A continuación, examinamos las propuestas específicas para cada una de las principales tareas en el ámbito de la gestión de proyectos, comenzando por la gestión de costes.

2.3.2 *Costes*

En las últimas décadas, se ha investigado mucho en relación a los factores que afectan a los costes, ya que las empresas de construcción se enfrentan a un reto importante para controlar los costes del proyecto desde la fase de diseño hasta la terminación de éste.

En la gestión de los proyectos, los costes son un tema fundamental ya que éstos están presentes durante todo el ciclo de vida del proyecto. En las primeras etapas, la estimación de costes es una tarea fundamental para el éxito del proyecto, ya que una buena estimación puede evitar sobrecostes en las siguientes etapas del proyecto [84]. En la fase de ejecución, donde se producen mayores sobrecostes, el control es fundamental para evitar desviaciones sobre los costes estimados en las primeras fases del proyecto. Muchas propuestas se han centrado en la identificación de factores que provocan un incremento en los costes del proyecto [123, 212, 223, 237]. Entre las principales causas, podemos encontrar: mala gestión y control de la obra, falta de información, complejidad del diseño, cambios en los objetivos, mucha competencia en la fase de contratación, lentitud en la toma de decisiones, y modificaciones de los clientes.

La estimación de costes del proyecto es el proceso que consiste en la determinación del coste para cada una de las unidades de obra necesarias para ejecutar el proyecto. Este proceso, se realiza en un periodo corto de tiempo y en las primeras etapas del proyecto, cuando todavía no se tiene pleno conocimiento de éste y se desconoce información que es relevante para realizar una buena estimación de costes, como por ejemplo, los recursos [125, 149, 219]. En este proceso de estimación, el acceso a datos históricos es esencial para complementar la información del mercado y la contenida en bases de precios, con el fin de conseguir una estimación más precisa [65]. Una incorrecta estimación de costes puede tener graves consecuencias en los objetivos del proyecto, ya que es la base de otras muchas tareas en la gestión de los proyectos

[20, 179]. Debido a la importancia de esta tarea, en la literatura se pueden encontrar muchas propuestas que utilizan tecnologías, tales como almacenes de datos, OLAP, CBR y BIM.

Para conseguir una estimación de costes más precisa, se requiere de un procesamiento analítico que permite extraer información útil de grandes cantidades de datos históricos almacenados en una base de datos. Para este propósito, las herramientas de análisis OLAP permiten realizar análisis de grandes cantidades de datos en un tiempo muy reducido, a partir de los datos almacenados:

- Moon et al. [181] desarrollan un sistema de gestión de costes que utiliza de manera eficiente los datos de costes de proyectos anteriores almacenados en un DW, que ayuda a los usuarios a entender las incertidumbres del proceso de estimación de costes y proporciona un método para obtener una planificación de costes más real.
- Nie y Froese [183] desarrollan dos cubos OLAP para las tareas de gestión de costes y asignación de recursos humanos. Para ello, el diseño del cubo cuenta con cinco dimensiones: tiempo, participante, tarea, producto y tipo de coste.
- Cho et al. [62] proponen una aplicación OLAP que integra datos de costes, planificación y ejecución. Los autores se centran en los datos de ejecución que están relacionados con las tareas de control, tales como, estimación de costes, cantidades, control de costes, control de planificación, certificaciones y mediciones de ejecución.

En la literatura, también se pueden encontrar trabajos significativos que utilizan técnicas de CBR, incluso, junto con algoritmos genéticos (GA):

- Chou [64] presenta un enfoque basado en la web que utiliza técnicas de CBR para la estimación de costes.
- Ji et al. [125, 126] desarrollan un modelo CBR para la estimación de costes de construcción. El modelo utiliza el concepto de distancia euclídea para medir la similitud y los algoritmos genéticos para la asignación del peso del atributo.
- Kim y Kim [139] proponen un modelo de estimación de costes que gestiona la incertidumbre eligiendo el caso correcto mediante un algoritmo genético. Este trabajo analiza la oportunidad de utilizar un algoritmo genético como alternativa para generar el peso de cada atributo que es necesario para la evaluación de la

similitud y la recuperación de casos similares, con la intención de obtener un error mínimo en la predicción.

- Du y Bormann [85] proponen un nuevo algoritmo para mejorar la medida de similitud, que es fundamental para la adecuada recuperación de casos cuando se utiliza la técnica de CBR.

En los últimos años, se puede observar un incremento en el uso de la tecnología BIM para asistir en la elaboración de mediciones y estimación de costes a través de todo el ciclo de vida del proyecto. BIM permite generar, de manera automática, las cantidades y medidas a partir del modelo 3D, reduciendo el proceso de estimación. Además, la estimación de costes se actualiza automáticamente cuando se produce alguna modificación en el modelo BIM.

- Cheung et al. [58] proporciona un método para la elaboración dinámica y adaptativa de la estimación de costes desde la fase de diseño.
- Lawrence et al. [149] proponen un enfoque para la creación y almacenamiento de las estimaciones de costes enlazando la información relativa a los costes con BIM.
- Lee et al. [152] proponen un proceso de inferencia basado en ontologías para la estimación de costes que permite la búsqueda de la unidad de trabajo apropiada, de manera automática, usando datos BIM.
- Niknam y Karshenas [185] presentan una aplicación para la preparación de estimaciones de costes basada en semántica, que combina información de: una base de conocimiento BIM, una base de conocimiento de descompuestos y servicios de web semántica de proveedores de materiales.

Además de estas tecnologías, se pueden encontrar otras propuestas en la literatura que resultan de interés:

- Staub-French et al. [219] proponen una ontología que permite a los estimadores almacenar los cálculos de los costes de construcción de una manera más integrada y consistente que en las herramientas tradicionales.
- Bansal y Pal [23] proponen una metodología que utiliza las funcionalidades de GIS para almacenar los datos espaciales y facilitarlos para la elaboración de la estimación.
- Benjaoran [29] desarrolla un sistema de control de costes basado en la web que permite monitorizar el progreso de los costes

del proyecto y realizan una comparación en tiempo real con el presupuesto.

- Chou [65] presenta un enfoque de simulación de costes, basado en las opiniones de gestores de los proyectos. Este enfoque ofrece una herramienta para la evaluación de los costes de construcción y la incertidumbre durante el proceso.
- Kim et al. [138] combinan dos métodos diferentes para reducir la incertidumbre asociada a la falta de información del proyecto en las primeras etapas: método basado en datos históricos y método de estimación basado en descompuestos. El método propuesto reconoce el papel del estimador en el proceso de estimación, proporcionando estimaciones transparentes para facilitar la aceptación de los usuarios, y haciendo un uso eficiente de los datos históricos.
- Chi et al. [60] investigan herramientas de data mining y desarrollan un marco de análisis basado en el uso de éstas técnicas para el análisis y predicción de resultado de proyectos que tienen muchos atributos .
- Williams y Gong [237] proponen el uso de técnicas de minería de texto, datos numéricos y una combinación de clasificadores para la estimación de costes de construcción. Las descripciones de texto de los proyectos se pueden combinar con datos numéricos con el objetivo de mejorar la predicción de sobrecostes.

Como se refleja en la literatura, el proceso de estimación es una tarea fundamental en la gestión de costes. Para este propósito, las propuestas que permiten el acceso a datos históricos para elaborar estimaciones de costes más precisas son las que han captado mayor atención. Una adecuada gestión de los costes, tanto en el proceso de estimación como en el de control, determinará el éxito del proyecto.

Otra tarea fundamental, y que está relacionada con el éxito de la gestión de los costes, es el control de los tiempos de las unidades de obra. En la siguiente sección analizamos las propuestas que se pueden encontrar en la literatura en relación con el uso de las TIC para mejorar la tarea de planificación.

2.3.3 *Planificación*

Entre los objetivos principales de una empresa constructora cuando se enfrenta a un proyecto, se encuentra el de entregar éste dentro del plazo

establecido. La planificación consiste en decidir cuándo debe empezar y terminar cada una de las unidades de obra y los recursos que son necesarios para ejecutarla. Para ello, el equipo del proyecto analiza el emplazamiento de la obra, estudia los documentos y los planos y desarrolla un plan para organizar las distintas actividades [233]. Este procedimiento, que en la mayoría de las empresas se sigue elaborando manualmente, supone una carga considerable para el equipo de trabajo debido a la gran cantidad de información que debe manejar en un breve periodo de tiempo. Además, el método tradicional, basado en la toma de decisiones fruto de la experiencia de los expertos, tiene una serie de limitaciones a la hora de analizar situaciones hipotéticas debido a la incapacidad de incorporar adecuadamente datos actualizados, que aportan información muy valiosa para la planificación [28].

Las desventajas que presentan los sistemas tradicionales, junto con el avance de las tecnologías, han impulsado el desarrollo de nuevas formas de planificación que permiten visualizar y simular el proceso de construcción al completo [117, 155, 156, 233]. Estas propuestas permiten a los usuarios realizar planificaciones más efectivas mediante la predicción de los resultados de los proyectos.

Durante la última década, se ha avanzado en el desarrollo de modelos BIM en cuatro dimensiones, que integran los modelos 3D con el tiempo, con el fin de generar de forma visual el proceso de construcción al detalle, permitiendo realizar planificaciones pormenorizadas, concisas y rigurosas. Afortunadamente, hay un aumento considerable en el número de propuestas que se basan en BIM para mejorar la tarea de planificación [162, 180, 218, 224]. Otra tecnología que resulta muy útil para apoyar la tarea de planificación es la tecnología GIS y así se refleja en la literatura [21, 24, 135, 192]. Algunos trabajos recientes integran las tecnologías de BIM y GIS, mediante la web semántica, para gestionar el proceso de planificación durante las fases iniciales del proyecto [133, 134]. Además, la combinación de estas dos tecnologías permite gestionar otro tipo de decisiones que también son necesarias para la planificación, como por ejemplo, identificar la mejor localización de las grúas [120] y la gestión de las instalaciones provisionales de obra [132].

Al igual que en la tarea de gestión de costes, varios trabajos han desarrollado herramientas que facilitan la tarea de planificación en base a datos históricos. La generación de planificaciones basadas en las experiencias aprendidas, en proyectos de construcción similares que se ejecutaron con éxito, es muy beneficiosa para los gerentes de proyecto, ya que permite estimar duraciones más realistas. En este sentido, Zayed y Wakil [255] presentan un proceso de extracción de conocimiento que prepara, utiliza, analiza y extrae patrones ocultos

de grandes cantidades de datos de proyectos. La propuesta incluye un enfoque difuso para predecir la duración de las actividades del proceso de construcción. Mikulanova et al. [177] presentan un sistema basado en conocimiento, que integra BIM y CBR y permite la generación automática de planificaciones utilizando experiencias de proyectos ejecutados anteriormente, que son similares al actual.

La productividad del trabajo es información fundamental para la estimación de la planificación de un proyecto de construcción. El método actual de estimación de plazos y de programación se basa en varias fuentes para obtener los valores de productividad, incluyendo la opinión personal del estimador, datos de productividad publicados y datos históricos del proyecto. En este sentido, Song y AbouRizk [217] proponen un enfoque sistemático para medir la productividad, mediante la recopilación de datos históricos.

Debido a las características del sector de la construcción, la duración de las actividades está sujeta a gran incertidumbre, que puede derivar de diversos motivos: desviación respecto a la estimación inicial, falta de disponibilidad de los recursos, retraso en la recepción de los materiales, condiciones climáticas adversas, huelgas, ausencia de trabajadores, etc. Para hacer frente a estos imponderables, los métodos tradicionales de planificación pueden no ser los más apropiados. En los últimos años, el enfoque de *construcción sin pérdidas* (Lean Construction en inglés) se está extendiendo en la gestión de proyectos para minimizar los efectos de los factores de riesgo en relación al tiempo [154]. Se trata de una nueva filosofía que pretende reducir al máximo el tiempo invertido en las actividades que no agregan valor al producto final, es decir, reducir las pérdidas en las actividades de construcción [122]. Koskela [144] estableció las bases teóricas de este concepto en la industria de la construcción, cuyo origen era la industria automovilística. En la práctica, el sistema de control *Sistema de Último Planificador* (Last Planner System en inglés) [19] mejora el cumplimiento de los plazos y costos, mediante el aumento de los compromisos entre los miembros del equipo, optimizando la coordinación e incrementando la capacidad de prever situaciones que pudieran afectar al desarrollo del proyecto.

Los proyectos de construcción se componen de cientos de actividades y es muy frecuente que se produzcan solapamientos entre ellas en la misma zona de trabajo, produciendo problemas en el proceso. Este problema ha captado la atención de investigadores. Así, por ejemplo, Moon et al. [180] proponen un método de optimización que minimiza estas actividades para mejorar el proceso de ejecución mediante la aplicación de la lógica difusa y analiza los riesgos para la planificación de estas actividades. Además, para la reducción de estas actividades se utiliza un enfoque basado en algoritmos genéticos.

Como se puede observar, las propuestas para facilitar la tarea de planificación de los proyectos se han centrado principalmente en dos enfoques. En primer lugar, las propuestas que permiten una planificación más precisa a través del acceso a los datos de proyectos anteriores y, en segundo lugar, las propuestas que permiten visualizar y simular el proceso de construcción, con el objetivo de que los usuarios puedan anticipar la ejecución del proyecto y hacer planificaciones más realistas.

2.3.4 *Gestión de riesgos*

En el ámbito de los proyectos de construcción, los riesgos se pueden definir como la probabilidad de que un evento perjudicial ocurra en el proyecto [20]. Los riesgos en el sector de la construcción pueden ser de diferente naturaleza [254]:

- Naturales, donde la ocurrencia del riesgo no es segura y es difícil de controlar, como por ejemplo: malas condiciones climáticas, fuertes lluvias o inundaciones.
- Financieros, que es uno de los riesgos al que se enfrentan la mayoría de las empresas. Algunos ejemplos de riesgos financieros son: costes laborales, fluctuaciones de precios, sobrecostes debido a la malas estimaciones.
- Sociales, que incluyen riesgos personales, inadecuada comunicación o malas relaciones entre las empresas que están involucradas en el proceso.
- Técnicos, como por ejemplo: mano de obra no cualificada, mal diseño del proyecto, malas condiciones del emplazamiento, etc.

La industria de la construcción está expuesta a más riesgos que otras industrias debido al entorno de trabajo. El término “gestión de riesgos” se puede definir como la tarea que clasifica, analiza y responde a los riesgos que existen o pueden existir durante el ciclo de vida del proyecto [131]. Por tanto, el proceso de gestión de riesgos incluye tres fases: identificación, cuantificación y control de los riesgos. En la fase de identificación, el objetivo principal es conocer las diversas incertidumbres que pueden ocurrir durante todo el ciclo de vida del proyecto, con el fin de facilitar el proceso de toma de decisiones. Kang et al. [131] presentan un sistema de simulación que combina la información de planificación con la información de los riesgos, de modo que los usuarios pueden visualizar los procesos y localizar los posibles riesgos.

Los repositorios que contienen información de riesgos son fundamentales para la fase de identificación de riesgos, porque ayudan a generar modelos de gestión de riesgos más realistas que permiten disminuir los riesgos en futuros proyectos [16]. Serpella et al. [210] señala que la gestión de riesgos en los proyectos de construcción sigue siendo poco eficaz y que la causa principal de esta situación es la falta de intercambio de conocimiento. Afortunadamente, cada vez hay más propuestas que utilizan la información de proyectos anteriores para la gestión de riesgos:

- Knight y Fayek [143] describen un modelo para predecir los posibles sobrecostos en la fase de diseño de los proyectos. Los datos utilizados para construir el modelo corresponden a proyectos anteriores y a información obtenida de los gestores del proyecto. El modelo utiliza la lógica difusa para relacionar las características del proyecto con la ocurrencia de los riesgos, con el fin de determinar el impacto en los costes en la fase de diseño del proyecto.
- Baloi y Price [20] desarrollan un entorno difuso para soportar la toma de decisiones que modela, analiza y gestiona de manera automática los riesgos que afectan al rendimiento de los costes, para permitir tomar las medidas adecuadas con el fin de reducir el impacto en los costes.
- Dikmen et al. [81] proponen un enfoque para la gestión de riesgos basado en el aprendizaje, en el que la información relacionada con los riesgos se define, almacena y actualiza en forma de una base de datos de lecciones aprendidas. Por tanto, los encargados en tomar decisiones se basan en experiencias anteriores para tomar decisiones más reales.
- Han et al. [115] desarrollan un modelo de gestión de riesgos basado en la tecnología web, que asocia las principales decisiones en cada etapa, utilizando información de casos anteriores similares.
- Tserng et al. [116] proponen un entorno de gestión de riesgos basados en ontologías para mejorar esta tarea mediante la optimización del flujo de la gestión de riesgos y la reutilización del conocimiento.
- Nieto-Morote y Ruz-Vila [184] presentan una metodología para el análisis de riesgos que están asociados a situaciones complicadas en las que la información para evaluar los riesgos es difícil de cuantificar, incompleta o no se puede obtener. La metodolo-

gía permite a los usuarios hacer sus juicios mediante términos lingüísticos que están asociados a un número difuso traapezoidal.

- Yildiz et al. [250] proponen una herramienta para la evaluación de riesgos de los proyectos de construcción internacionales que considera la causa entre las variables relacionadas con el riesgo, los impactos de los riesgos y las lecciones aprendidas de proyectos anteriores.
- Zhong y li [260] proponen un modelo basado en ontologías y semántica para la formalización del conocimiento de riesgos de construcción.

Tradicionalmente, la gestión de riesgos se ha centrado en las tareas de seguridad, costes y planificación, pero su área de aplicación se ha ampliado para incluir otras tareas, como por ejemplo, proceso de ofertas, estudios de viabilidad, estudios de mercado, evaluaciones de ejecución y gestión de manejo de diversas circunstancias que se producen en las distintas fases del ciclo de vida del proyecto [115].

La creciente complejidad y el carácter dinámico de los proyectos de construcción han plagado a la industria de la construcción con riesgos considerables [249]. La mala gestión de estos riesgos, provoca una mala ejecución del proyecto ocasionando un aumento en los costes y en los tiempos de ejecución de las actividades. Por lo tanto, existe una relación directa entre la gestión eficaz de los distintos riesgos y el éxito del proyecto, ya que los riesgos son evaluados por el efecto que éstos tienen sobre los objetivos del proyecto [122].

2.3.5 Seguridad

La seguridad, relacionada estrechamente con la gestión de riesgos, es otra de las tareas fundamentales en el ámbito de la gestión de proyectos de construcción. Los accidentes causan tragedias humanas, desmotivan a los trabajadores, interrumpen el proceso de construcción generando retrasos en la planificación, etc. Además, los accidentes afectan a los costes, productividad y al prestigio de la empresa constructora. El contexto en el que se llevan a cabo los trabajos y las características del sector son muy importantes, ya que dificultan, en la mayoría de los casos, llevar a cabo una adecuada gestión de riesgos y prevención de accidentes.

La identificación de los riesgos de accidentes que se pueden dar en el lugar de trabajo es una de las bases para el éxito de la gestión de la seguridad. Carter y Smith [45] indican que el proceso de

identificación de riesgos en los proyectos de construcción está lejos del modelo ideal e identifican las causas que impiden esta situación: barreras en relación a la información y conocimiento (por ejemplo, falta de intercambio de información, falta de recursos en proyectos pequeños y falta de confianza en el conocimiento tácito), barreras en el proceso (por ejemplo, falta de métodos normalizados, uso de equipos no seguros, falta de aprendizaje adecuado, falta de organización de las actividades, condiciones de trabajo no seguras, no utilizar el equipo de protección individual y la actitud indiferente ante la seguridad). La identificación de riesgos comienza en la fase de diseño y continua hasta la terminación del proyecto. En la fase de diseño, se pueden incorporar consideraciones en relación a la seguridad, involucrando a los diseñadores para que evalúen los riesgos desde dicha fase. Este proceso se conoce como Prevención a través del Diseño (Prevention through Design PtD) [166, 199, 222, 258, 259].

Una de las principales barreras es la falta de comunicación a través de todo el proyecto. Desafortunadamente, las empresas no comparten los datos referentes a los accidentes que se producen en las obras de construcción con otras empresas e, incluso, con el personal de la misma empresa. Para mejorar esta situación, en la literatura se pueden encontrar muchas propuestas que se dedican a la gestión de esta tarea:

- Rivas et al. [205] evalúan diversas técnicas de minería de datos (tales como redes bayesianas, reglas de decisión, árboles de clasificación, regresión logística y máquinas de vectores de soporte (en inglés, support vector machines)), con el fin de reducir los índices de accidentes de trabajo, permitiendo que las medidas preventivas se concentren en las zonas de mayor riesgo.
- Kim et al. [137] proponen un sistema de recuperación de información automático que busca y proporciona accidentes similares con el fin de prevenir futuros accidentes. El sistema de recuperación extrae los objetos BIM y mediante la combinación de estos objetos con un sistema de información de gestión de proyectos, permite consultar en base a actividades, condiciones de trabajo y laborales)
- Fan y Li [95] describen un enfoque basado en la aplicación de técnicas de minería de texto que permite recuperar casos similares, de manera automática, de un repositorio de accidentes de construcción utilizando el modelo de espacio vectorial. Comparando las similitudes y diferencias entre el caso objeto de estudio y los casos recuperados, se pueden llevar a cabo una serie de acciones para evitar la ocurrencia de accidentes.

- Wu et al. [243] analizan los accidentes por la caída de objetos en obras de construcción y proponen un modelo integrado de gestión de la información. El modelo se divide en tres pasos. En primer lugar, analiza e identifica la frecuencia de un caso de accidente en función a casos previos similares. A continuación, mediante una red de sensores identifica la información necesaria para prevenir este tipo de accidentes y, por último, desarrolla un modelo de gestión integrado.
- Le et al.[150] proponen un sistema de red social, que utiliza web semántica y ontologías, para la coordinación e intercambio de información y conocimiento sobre seguridad y salud en la industria de la construcción.

El análisis de accidentes a través de datos históricos puede proporcionar valiosa información para la planificación de la seguridad. Sin embargo, la información que se pueda extraer no es suficiente para predecir cuándo y dónde se van a producir los accidentes. La ocurrencia de accidentes se puede reducir, pero no se pueden eliminar completamente debido a las características del proceso de construcción. Por este motivo, el análisis de datos históricos se puede complementar con otros enfoques, tales como, diseños virtuales y simuladores del proceso, que proporcionan una visión más real del proceso de construcción y, por tanto facilita a identificar los lugares o actividades que son más susceptibles de presentar riesgos para los trabajadores [259]. Para este propósito, se han desarrollado propuestas que utilizan tecnologías como BIM [199, 257, 258, 259], tecnologías de seguimiento [44, 49, 56, 103, 112, 113, 151, 182, 242, 246], simulación y visualización [66, 109] junto con GIS [22, 166].

Sin embargo, algunos riesgos son inevitables, incluso en los diseños más seguros, debido a la naturaleza de los trabajos de construcción. Estos riesgos tendrán que ser gestionados en la fase de ejecución del proyecto. Por esta razón, en los últimos años, muchas propuestas se han centrado en la aplicación de tecnologías que permiten la adquisición de datos en tiempo real de la obra. Estas herramientas permiten la monitorización de actividades que pueden mejorar la seguridad y salud, por ejemplo, detectando acciones de los trabajadores que no son seguras, de modo que se les puede advertir para mejorar su comportamiento y, por tanto, reduciendo el número de acciones no seguras que se producen en la obra [113]. Sin embargo, estas herramientas no evalúan los datos para detectar los actos o condiciones no seguras [209].

Además de su aplicación para mejorar la seguridad y salud en las obras de construcción, las propuestas que utilizan las tecnologías de

localización y seguimiento se centran fundamentalmente en la monitorización del proceso de construcción, como veremos en la siguiente sección.

2.3.6 *Monitorización del proceso*

Durante la fase de ejecución de los proyectos es fundamental llevar a cabo un buen control de todos los procesos, por lo que son necesarios procedimientos que ayuden en el seguimiento del proceso [111]. El control consiste en la monitorización constante del proceso de construcción conforme se lleva a cabo (en inglés, *as-built*) y en la comparación con la planificación establecida (en inglés, *as-planned*), de modo que se puedan realizar acciones correctivas cuando se detecte alguna desviación y, por tanto, terminar el proyecto con éxito [36, 105, 247]. Los métodos tradicionales para realizar el control del proceso requieren gran cantidad de tiempo y tienden a admitir errores, ya que los datos que son el soporte para la toma de decisiones son introducidos por los usuarios de la propia obra a través de diversas aplicaciones. Por consiguiente, con este método de control, existe un espacio de tiempo considerable entre el proceso de toma de datos en la propia obra y la disposición de éstos en los sistemas para poder tomar decisiones, reduciendo la eficacia y generando confusión y falta de datos [235].

Con la creciente complejidad de los proyectos de construcción y el aumento de las restricciones en relación al coste y tiempo, la recogida de datos *in situ* se está convirtiendo cada vez más importante. Estos datos son fundamentales para controlar el proceso y dar soporte a la toma de decisiones, con el objetivo de poder detectar cualquier desviación que se produzca durante la ejecución del proyecto. Esta cuestión se ha abordado en la literatura desde la perspectiva de diversos problemas:

- Cheung et al. [59] desarrollan un sistema basado en web que permite agilizar el proceso de seguimiento de la ejecución del proyecto, desde la entrada de los datos hasta la presentación de éstos. El sistema incluye ocho categorías que se deben considerar: personal, coste, tiempo, calidad, seguridad y salud, medio ambiente, satisfacción del cliente y comunicación.
- Wang et al. [235] presentan un sistema para mejorar la adquisición de información de la obra de construcción y proporcionar una plataforma de intercambio entre todos los participantes mediante la tecnología web y RFID.

- Elbeltagi y Dawood [92] desarrollan un sistema de visualización basado en BIM y GIS para evaluar la ejecución y facilitar el seguimiento del proceso de construcción.
- Caldas et al. [42], Park et al. [191], Razavi y Haas [202] y, Pradhananga y Teizer [193] presentan enfoques integrados para la identificación y localización automática de recursos de construcción. En el tercer trabajo, los autores razonamiento de creencias evidencial y técnicas de Computación Flexible (en inglés, Soft Computing) para tratar la imprecisión, la vaguedad y la ambigüedad de los datos adquiridos.

En los últimos años, la combinación de BIM con las tecnologías de seguimiento ha captado la atención de los investigadores para la tarea de monitorización del proceso. La combinación de estas tecnologías permite comparar el progreso de ejecución, mediante las tecnologías de seguimiento, con respecto a lo planificado, a través del modelo BIM.

- Irizarry et al. [121] integran BIM con GIS para hacer el seguimiento del estado de la cadena de suministro y dar señales de alerta que permitirían mejorar el proceso de suministro de materiales.
- Bosché et al. [34] analizan un enfoque que permite hacer un seguimiento del estado de la ejecución mediante BIM vs scanner. La propuesta aprovecha la información inicial del modelo, mediante BIM, para la generación de modelos del estado real de ejecución.
- El seguimiento del proceso de ejecución en relación a lo planificado se ha llevado a cabo a través de diferentes tecnologías para el seguimiento, tales como, laser scanning [33, 136, 226], RFID [35, 124], imágenes digitales [82, 106, 107, 111, 191] y UWB [165, 211].

La mayor parte de los trabajos de investigación dedicados al seguimiento del proceso de construcción automatizado, se han centrado en el seguimiento de los trabajos que comprenden la estructura del edificio [34]. Sin embargo, la aplicación de estas técnicas en otro tipo de trabajos que no son tan precisos como la estructura del edificio presenta diferencias considerables entre los datos del diseño y la ejecución. Por ejemplo, en las instalaciones de fontanería o electricidad, donde la instalación de los elementos es más flexible y presenta desviaciones respecto a los ejes [34, 211].

2.3.7 *Gestión de la calidad*

Por último, pero no menos importante, nos centramos en la tarea de gestión de la calidad, que está estrechamente relacionada con la monitorización del proceso. Dentro de la tarea de gestión de calidad, el control de calidad asegura que los productos y servicios cumplen con los requisitos establecidos mediante la comparación de los resultados obtenidos y deseados, con el fin de tomar decisiones que corrijan las posibles diferencias.

El control de calidad es una tarea que cada vez preocupa más a los responsables de los proyectos, ya que los defectos o fallos pueden tener consecuencias en los objetivos finales del proyecto, tales como, incremento de costes y tiempo. Esta tarea requiere la gestión de grandes cantidades de datos, ya que implica varios procesos: inspección y pruebas, informes de no conformidad y acciones correctivas [61]. En la actualidad, en la mayoría de los proyectos, el control de calidad se lleva a cabo de una manera tradicional; los ingenieros inspeccionan la obra y hacen anotaciones en documentos en soporte papel, complicando la gestión de la calidad de manera eficiente [189]. Como consecuencia de este tedioso método, en algunos casos, las inspecciones se reducen, resultando en un mayor número de defectos que afectan directamente en el coste y en los plazos establecidos [146].

La implementación de sistemas de gestión de calidad que ayudan en el proceso de adquisición, acceso, intercambio y reutilización de los datos de calidad a nivel de proyecto y de la empresa se ha afrontado en la literatura a través de diferentes tecnologías:

- Chin et al. [61] presentan un sistema de gestión basado en la web que facilita la colaboración y la comunicación entre los participantes, la gestión eficiente de la documentación y el registro de información de calidad, el seguimiento de todo el proceso de control de calidad y el almacenamiento de la información de cómo se ejecuta el proyecto.
- Lam y Ng [147] presentan un sistema de gestión de calidad basado en la web para el intercambio eficaz de información.
- Wang [234] propone un sistema de inspección de calidad que integra la tecnología web junto con RFID, que mejora la adquisición de datos en tiempo real y el intercambio de información entre los distintos participantes, facilitando la monitorización y el control.

- Kim et al. [140] proponen un sistema de inspección de calidad y de gestión de defectos, que permite recopilar datos de los defectos en la obra en tiempo real, utilizando tabletas y conexión inalámbrica a Internet y también permite gestionar eficazmente los estados y resultados de las acciones correctivas realizadas.
- Cheng et al. [57] proponen un enfoque para mejorar la calidad de las obras de construcción utilizando algoritmos genéticos. Éste proporciona conocimiento de casos similares en proyectos anteriores, con el objetivo de poder establecer estrategias de prevención en cuanto a los defectos de construcción.

Algunas propuestas, además de ayudar en el proceso de adquisición de datos e inspección, permiten analizar automáticamente los datos, como por ejemplo:

- Akinci et al.[9] esbozan un proceso de adquisición y actualización de información de diseño, identificando metas de inspección, planificación de las inspecciones, adquisición y análisis de los datos de cómo se está ejecutando y la detección y gestión de defectos.
- Zhu y Brilakis [261] presentan un enfoque para evaluar de manera automática, mediante el análisis de imágenes, el impacto de dos defectos en la superficie del hormigón.
- Park et al. [189] presentan un marco conceptual para la gestión de defectos de construcción, vinculando ontologías y realidad aumentada con BIM. El sistema permite buscar y recuperar información de defectos del proyecto o de actividades específicas, con el fin de ayudar en la gestión de defectos en la obra.
- Chen y Luo [54] proponen una solución integrada, en el entorno de trabajo BIM, para mejorar los procesos actuales de gestión de la calidad. La información recogida en la obra se incorpora en el modelo y se procesa, con el fin de comparar los datos de ensayos de calidad con los datos de referencia.
- Kwon et al. [146] presentan un sistema de gestión de defectos del hormigón armado utilizando BIM, un sistema de coincidencia basado en imágenes y un sistema de realidad aumentada para prevenir defectos. El sistema permite ahorrar tiempo de inspección y, por tanto, reducir los costes relacionados con la repetición de trabajos.

2.4 ANÁLISIS Y MOTIVACIÓN DE NUESTRA INVESTIGACIÓN

Como se ha podido ver a lo largo de la revisión de la literatura, no hay duda de que el sector de la construcción es una industria muy intensa en cuanto a la producción de información, ya que se generan e intercambian grandes cantidades de datos durante todo el ciclo de vida del proyecto. Una buena gestión de estos datos es vital para asegurar el éxito de los objetivos del proyecto. Además, el acceso a la información adecuada en el momento adecuado es fundamental tanto para el funcionamiento diario de la empresas como para dar soporte a la toma de decisiones.

En esta sección, en primer lugar, se presentan las conclusiones generales del estado actual de las propuestas que utilizan las TIC para la gestión de los datos en el ámbito de construcción. Además, se presenta de manera gráfica la distribución de tecnologías por tareas de las propuestas analizadas y una tabla que resume el objetivo de cada uno de los trabajos. En segundo lugar, en base a estas conclusiones, detallamos los retos que todavía se tienen que alcanzar para conseguir un tratamiento de los datos automático a lo largo del ciclo de vida del proyecto para dar soporte a la toma de decisiones.

2.4.1 *Análisis del estado actual*

El uso de las TIC, aunque más lento que en otras áreas debido a las particularidades del sector de la construcción, ha aportado grandes beneficios al problema de la gestión de datos en los proyectos de construcción. En la revisión de la literatura hemos visto propuestas generales que permiten la colaboración y el intercambio de información entre los distintos participantes de un proyecto, y propuestas más específicas para facilitar la gestión de las principales tareas de la gestión de los proyectos; costes, planificación, gestión de riesgos, seguridad, monitorización del proceso y control de calidad.

En estas propuestas se puede observar que la tecnología web ha sido la que ha proporcionado la infraestructura para el desarrollo de sistemas de trabajo distribuidos, incluyendo el uso de de conceptos de la web semántica para facilitar el intercambio y recuperación de información en un entorno de interoperabilidad.

Los sistemas de información y las soluciones creadas dentro del área de la Inteligencia de Negocio también han tenido un impacto en el dominio de la construcción. En el contexto del modelado de datos, un número significativo de propuestas se basan en el uso de BIM, cuyo

objetivo es la gestión de datos de manera integrada a través de la creación de un único modelo, que incorpora los datos relevantes de las distintas tareas del proyecto. Para este propósito, los datos capturados a través de las tecnologías de seguimiento y localización, que permiten la monitorización del proceso de manera precisa y actualizada, pueden ser incorporados al modelo BIM y otras tecnologías. Los datos están destinados a ser gestionados y compartidos de manera centralizada.

En cuanto al almacenamiento, consulta y análisis de los datos, las tecnologías que han captado más atención, en las empresas de construcción, son la tecnología de almacenes de datos y la minería de datos. Estas tecnologías proporcionan almacenes integrados de datos históricos que pueden ser consultados a través herramientas OLAP o a través de técnicas inteligentes que permiten encontrar información útil para mejorar el funcionamiento de la empresa. Por último, también se han analizado propuestas en el área de los sistemas de información inteligentes, más allá de la minería de datos, tales como el razonamiento basado en casos o enfoques basados en lógica difusa.

La Figura 1 resume gráficamente la distribución de las propuestas analizadas en esta sección en relación a las principales tecnologías que se han aplicado para las distintas tareas de gestión en los proyectos de construcción:

- Como se puede observar en la gráfica, en la tarea de gestión de costes se han utilizado prácticamente todas las tecnologías. Sin embargo, las tecnologías que cuentan con un mayor número de propuestas son DW, BIM y CBR. Esta diversidad, también se puede observar en la tarea de planificación, aunque en este caso las tareas que destacan son BIM y GIS.
- Tanto en la tarea de gestión de riesgos como en la de control de calidad la diversidad de tecnologías es menor. Para la gestión de riesgos destacan las propuestas basadas en lógica difusa y ontologías, mientras que para la tarea de gestión de calidad son las tecnologías de seguimiento y web.
- Para la gestión de la información y conocimiento se puede observar que la tecnología web es la que proporciona la infraestructura para afrontar el intercambio y recuperación de información.
- Por último, tanto en la tarea de seguridad como en la de monitorización del proceso se puede observar que las tecnologías que juegan un papel fundamental son BIM (con ayuda de otras tecnologías como GIS) y las tecnologías para el seguimiento.

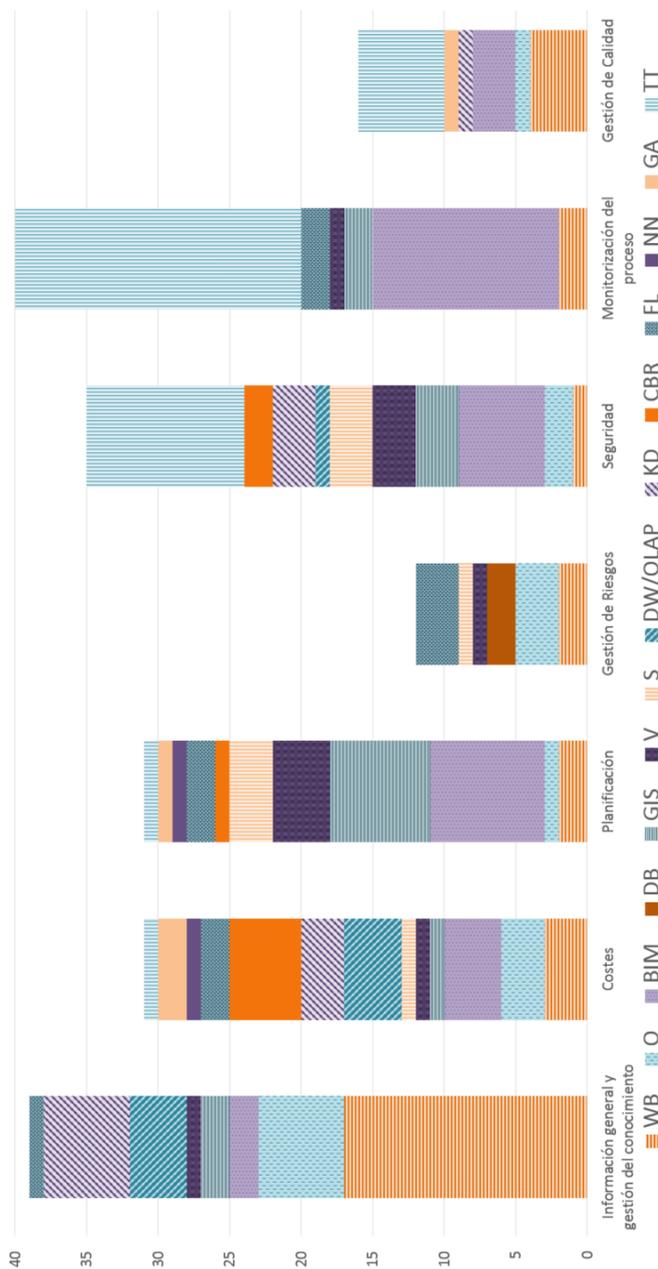


Figura 1: Distribución de tecnologías por tareas del ámbito de la gestión de proyectos

Nota: Web (WB), Ontologías (O), Modelado de información de construcción (BIM), Bases de datos (DB), Sistemas de información geográfica (GIS), Virtual (V), Simulación (S), Almacenes de datos (DW), Procesamiento analítico en línea (OLAP), Extracción de conocimiento (KD), Razonamiento basado en casos (CBR), Lógica Difusa (FL), Redes neuronales (NN), Algoritmos genéticos (GA), Tecnologías de seguimiento (TT).

La Tabla 1 resume las propuestas analizadas, indicando para cada referencia, la tarea de la gestión de proyectos a la que se dedica, las principales tecnologías que utiliza, y el objetivo principal que buscan los autores. Como se puede observar, la lista de propuestas es muy amplia y los objetivos alcanzados son significativos para cada una de las tareas.

Aunque estos logros son muchos, el carácter tradicional y la complejidad del dominio de la construcción ha condicionado la aplicación más lenta de las TIC para la gestión de los proyectos de construcción. Esto hace que la industria de la construcción se siga enfrentando a grandes desafíos en relación a importantes tareas para la gestión de datos en todos sus ámbitos: adquisición, almacenamiento, procesamiento, visualización y consulta.

Referencia	Tarea	Tecnología	Objetivo
[219]	C	O	Representar e intercambiar el conocimiento
[181]	C	OLAP/DW	Proporcionar información útil para la toma de decisiones estratégicas
[23]	C	V/GIS	Mejorar la precisión de la estimación de costes
[183]	C	OLAP/DW	Proporcionar información útil para la toma de decisiones estratégicas
[29]	C	WB	Mejorar la eficiencia de los procesos de la empresa
[64]	C	WB/CBR	Mejorar la recuperación de proyectos anteriores similares, en las fases iniciales del proyecto para ayudar a la toma de decisiones
[139]	C	CBR/GA	Mejorar la precisión de la estimación de costes
[125]	C	CBR	Mejorar la precisión de la estimación de costes
[65]	C	S	Aumentar la precisión en la simulación de costes
[126]	C	GA/CBR	Mejorar la precisión de la estimación de costes
[138]	C	KD	Mejorar la precisión de la estimación de costes y proporcionar transparencia en el proceso
[58]	C	BIM	Incorporar la estimación de costes desde la fase de diseño del proyecto
[60]	C	KD	Predecir la ejecución de los proyectos de construcción
[62]	C	OLAP	Acceder a todos los datos del proyecto desde diferentes perspectivas
[237]	C	KD	Mejorar la predicción de sobre costes
[149]	C	BIM	Disminuir el tiempo de actualización de las estimación cuando se produce un cambio en el diseño y generar estimaciones más precisas
[85]	C	NN/CBR	Ahorrar tiempo a los estimadores
[152]	C	BIM/O	Deducir automáticamente la mejor unidad de trabajo según las condiciones del trabajo
[168]	C	OLAP/DW/FL	Proporcionar información útil para la toma de decisiones estratégicas
[185]	C	BIM/WB/O	Facilitar la búsqueda, acceso, y la combinación de la información necesaria para la estimación de costes
[143]	C/GR	FL	Permitir la evaluación de las características del proyecto y de los eventos de riesgo de manera subjetiva. Un modelo más realista en el contexto del diseño

Continúa en la página siguiente.

Referencia	Tarea	Tecnología	Objetivo
[39]	IG/GC	KD	Mejorar la organización de la información y el acceso a los sistemas de información de gestión del proyecto de construcción
[53]	IG/GC	DW	Proporcionar información útil para la toma de decisiones estratégicas
[225]	IG/GC	WB	Capturar y reutilizar el conocimiento de los proyectos
[51]	IG/GC	WB	Mejorar la gestión de la información
[7]	IG/GC	DW/OLAP/GIS	Proporcionar información útil para la toma de decisiones estratégicas
[157]	IG/GC	WB/O	Proporcionar capacidades de búsqueda e indexación adecuados
[88]	IG/GC	O/WB	Representar y compartir el conocimiento
[40]	IG/GC	KD	Integrar los documentos de proyectos en sistemas de información
[161]	IG/GC	WB	Mejorar la gestión del conocimiento
[94]	IG/GC	DW/WB	Proporcionar información útil para la toma de decisiones estratégicas
[207]	IG/GC	DW	Proporcionar información útil para la toma de decisiones estratégicas
[98]	IG/GC	WB	Gestionar con éxito los documentos del proyecto
[216]	IG/GC	KD	Descubrir conocimiento novedoso de grandes bases de datos de construcción
[142]	IG/GC	WB	Mejorar la gestión del conocimiento
[158]	IG/GC	WB	Capturar y reutilizar el conocimiento
[52]	IG/GC	WB	Facilitar la gestión y comunicación de la información de construcción
[89]	IG/GC	WB/O	Representar y compartir el conocimiento
[91]	IG/GC	O	Representar y compartir el conocimiento
[256]	IG/GC	WB	Mejorar la gestión del conocimiento
[220]	IG/GC	WB	Capturar y reutilizar el conocimiento
[160]	IG/GC	WB	Mejorar la gestión del conocimiento
[10]	IG/GC	WB	Gestionar con éxito los documentos de los proyectos
[171]	IG/GC	FL	Integrar los datos de los proyectos de diversas fuentes a posteriori
[190]	IG/GC	WB/O	Mejorar la gestión del conocimiento
[79]	IG/GC	GIS	Mejorar la comunicación, documentación y intercambio de conocimientos través del ciclo de vida del proyecto
[236]	IG/GC	KD	Recuperar documentos de bases de datos grandes
[198]	IG/GC	KD	Encontrar relaciones semánticas entre los documentos almacenados
[159]	IG/GC	WB/BIM	Capturar y representar el conocimiento
[78]	IG/GC	BIM/V	Capturar y reutilizar el conocimiento
[232]	IG/GC	O	Mejorar la gestión del proceso a través de la integración, el procesamiento y el aprovechamiento de la información
[12]	IG	KD	Gestionar el flujo de la información de una forma organizada
[233]	P	V	Probar distintas estrategias de ejecución, lo más real posible, antes del inicio del proceso de construcción
[192]	P	V/GIS	Comunicar eficientemente la información referente a la planificación y proceso

Continúa en la página siguiente.

Referencia	Tarea	Tecnología	Objetivo
[117]	P	V/S	Ayudar a los planificadores de proyectos a comprender mejor el proceso de construcción y predecir posibles errores en las etapas iniciales
[217]	P	NN/S	Medir y predecir la productividad laboral
[21]	P	GIS	Facilitar una mejor comprensión de la secuencia de construcción, así como visualizar la planificación
[24]	P	GIS	Facilitar una mejor comprensión de la secuencia de construcción
[48]	P	FL	Generar la planificación del proyecto mediante modelo difusos e incorporar restricciones en elementos como materiales, tiempo y costes
[155]	P	V/S	Ayudar a los encargados de hacer la planificación a comprender mejor el proceso de construcción y predecir posibles errores en las etapas iniciales
[177]	P	BIM/CBR	Generación automática de planificaciones de obra utilizando experiencias de proyectos similares ejecutados con éxito anteriormente. Evaluación de alternativas de construcción
[255]	P	FL	Predecir la duración de las tareas de las unidades de obra
[120]	P	GIS/BIM	Identificar lugares posibles para la localización de las grúas
[218]	P	BIM	Permitir a los usuarios establecer planificaciones eficaces y facilitar la toma de decisiones mediante la predicción de los resultados del proyecto
[180]	P	GA/BIM	Encontrar una planificación con el menor número de actividades solapadas de modos que se mejore la ejecución del proyecto
[224]	P	WB/BIM	Facilidad y eficiencia para actualizar la información de la planificación
[135]	P	GIS/TT	Mejorar el proceso de toma de decisiones sobre el diseño y la planificación de las operaciones de pre-construcción
[162]	P	BIM	Generar programaciones para los proyectos de construcción mediante la simulación de la programación con el modelo BIM
[133]	P	BIM/GIS/WB/O	Permitir la interoperabilidad semántica entre el edificio y los datos espaciales
[132]	P	GIS/BIM	Integrar datos de fuentes heterogéneas
[59]	MP	WB	Ayudar a los gestores del proyecto en la monitorización y evaluación del proceso de ejecución
[42]	MP	TT	Mejorar el seguimiento y localización de materiales en las obras
[235]	MP	WB/TT	Mejorar la eficiencia y la eficacia de la adquisición de datos in situ y el intercambio de información entre los participantes
[124]	MP	TT	Mejorar los métodos tradicionales de seguimiento de progreso de la construcción
[35]	MP	TT/BIM	Seguimiento automático, análisis y visualización del progreso para mejorar la toma de decisiones
[105]	MP	TT/BIM	Representar las desviaciones de progreso a través de la superposición del modelo 4D as-planned y fotografías
[33]	MP/QM	BIM/TT	Seguimiento automático, análisis y visualización del progreso para mejorar la toma de decisiones

Continúa en la página siguiente.

Referencia	Tarea	Tecnología	Objetivo
[92]	MP	BIM/V/GIS	Facilitar el seguimiento y monitorización del progreso de la construcción. Comunicación eficaz de la información del progreso en tiempo real
[191]	MP	TT	Mejorar el seguimiento y localización de materiales en las obras
[106]	MP	BIM/TT	Mejorar la identificación, el procesamiento y la comunicación de las discrepancias de progreso
[202]	MP	FL/TT	Gestionar la imprecisión, vaguedad y ambigüedad de los datos adquiridos
[226]	MP	BIM/TT	Mejorar los métodos tradicionales de seguimiento de progreso de la construcción
[127]	MP	TT/FL	Seleccionar una tecnología adecuada para el seguimiento de los materiales de construcción
[193]	MP	TT	Seguimiento y análisis de los proceso de la obra en relación a los equipos
[121]	MP	GIS/BIM	Mejorar el seguimiento y localización de la cadena
[136]	MP	BIM/TT	Medir el proceso de ejecución
[182]	MP/S	TT	Información en tiempo real para la prevención dinámica
[34]	MP	BIM/TT	Mejorar la evaluación as-built de los proyectos de construcción
[82]	MP	BIM/TT	Mejorar la monitorización del progreso mediante la automatización de reconocimiento de materiales
[107]	MP	BIM/TT	Seguimiento automático, análisis y visualización del progreso para mejorar la toma de decisiones
[211]	MP	BIM/TT	Seguimiento automático del progreso de las actividades
[111]	MP	TT/BIM	Representar las desviaciones de progreso a través de la superposición del modelo 4D as-planned y fotografías
[163]	MP	TT	Facilitar el análisis de grandes cantidades de videos
[61]	GC	TT/WB	Mejorar la eficiencia y eficacia de la adquisición automática de datos e intercambio de información entre los participantes
[147]	GC	WB	Mejorar la comunicación, la coordinación y la calidad
[9]	GC	TT	Mejorar los procesos de control de calidad tradicionales
[234]	GC	TT/WB	Mejorar la eficiencia y eficacia de la adquisición automática de datos e intercambio de información entre los participantes
[140]	GC	WB/TT	Mejorar la eficiencia y la eficacia de la adquisición de datos en la obra
[261]	GC	TT	Evaluar de forma automática el impacto de dos defectos en la superficie de hormigón
[189]	GC	BIM/O	Reducir y prevenir la aparición de defectos durante el proceso de construcción
[146]	GC	BIM/TT	Reducir y mejorar la carga de trabajo de los ingenieros en la obra
[54]	GC	BIM	Garantizar la coherencia de la información y facilitar la tarea de gestión de la calidad
[57]	GC	GA/KD	Proporcionar conocimiento relacionado con defectos en casos similares anteriores para la prevención de defectos

Continúa en la página siguiente.

Referencia	Tarea	Tecnología	Objetivo
[20]	GR	FL	Gestionar la incertidumbre de los factores de riesgo que afectan el rendimiento de costes de construcción
[115]	GR	WB	Mejorar el acceso a la información para tomar mejores decisiones
[81]	GR	DB	Almacenar información relacionada con los riesgos
[116]	GR	O	Ayudar en la identificación, análisis y respuesta de riesgos. Para aumentar la eficacia de flujo de trabajo de gestión de riesgos
[184]	GR	FL	Hacer frente a criterios subjetivo
[131]	GR	V/S	Analizar gráficamente los grados de los riesgos
[260]	GR	WB/O	Representar el conocimiento relacionado con los riesgos de manera interpretable por los ordenadores. Permitir gestionar y reutilizar el conocimiento de una forma más flexible, dinámica y semántica
[250]	GR	O/DB	Almacenar la información relacionada con los riesgos
[103]	S	TT	Monitorización en tiempo real de la posición de los trabajadores y equipos, en las obras de construcción
[49]	S	TT	Prevenir los accidentes de colisión entre trabajadores y equipos pesados
[242]	S	TT	Información en tiempo real para el seguimiento de los accidentes
[166]	S	GIS	Mejorar el análisis de la información y la comprensión de los diversos factores de la industria de la construcción para mitigar los accidentes
[170]	S	OLAP/DW	Facilitar el análisis de los indicadores de seguridad para extraer patrones útiles de grandes cantidades de datos para mejorar el proceso de toma de decisiones
[22]	S	GIS	Facilitar la comprensión de la secuencia del proceso de construcción y predecir los lugares y actividades que son potenciales de accidentes
[151]	S	TT	Proporcionar localización precisa de la ejecución en las obras
[44]	S	TT	Alertar de los riesgos sobre la cabeza
[205]	S	KD	extraer información de los accidentes de trabajo de una base de datos
[66]	S	V/S	Ayudar a los planificadores de proyectos a comprender mejor el proceso de construcción y predecir posibles errores en las etapas iniciales
[246]	S	TT	Mejorar la seguridad en la obra
[259]	S	BIM/S	Detectar los riesgos de seguridad y sugerir medidas preventivas a los usuarios
[95]	S	KD/CBR	Recuperación efectiva de los casos históricos
[243]	S	TT	Información en tiempo real para la prevención dinámica
[109]	S	V/S	Simular e identificar los problemas de seguridad de los procesos de construcción
[112]	S	TT	Detectar automáticamente las acciones inseguras de los trabajadores
[56]	S	TT/V	Monitorizar automáticamente y visualizar la información en relación a la seguridad y actividades en la obra
[137]	S	BIM/CBR	Recuperación de accidentes similares
[199]	S	BIM	Planificar las medidas de seguridad para hacer frente a los peligros de obras de construcción desde el inicio del proyecto

Continúa en la página siguiente.

Referencia	Tarea	Tecnología	Objetivo
[150]	S	WB/O	Mejorar la comunicación y la representación de la información de seguridad
[113]	S	TT	Detectar automáticamente las acciones inseguras de los trabajadores
[258]	S	BIM	Identificar los riesgos de caídas más potenciales basados en la planificación de la obra
[55]		GIS/BIM/KD	Semi-automatizar el mapeo de esquemas de BIM y 3D GIS
[257]	S	BIM/O	Formalizar el conocimiento de seguridad en la construcción y conectarlo con BIM para el análisis de riesgos automatizado, y programación de la tarea de seguridad
[209]	S	TT	Supervisar la seguridad y salud en las obras reales, e identificar las posibles soluciones

Tabla 1: Hallazgos de las TIC en relación a las tareas de gestión en construcción

Nota: Información General y Gestión del Conocimiento (IG/GC), Costes (C), Planificación (P), Gestión de riesgos (GR), Seguridad (S), Monitorización del proceso (MP), Gestión de la calidad (GC), Web (WB), Modelado de información de construcción (BIM), Sistemas de información geográfica (GIS), Virtual (V), Simulación (S), Almacenes de datos (DW), Procesamiento Analítico en línea (OLAP), Extracción de conocimiento (KD), Razonamiento basado en casos (CBR), Algoritmos genéticos (GA), Redes neuronales (RN), Lógica difusa (FL), Tecnologías de Seguimiento (TT).

2.4.2 Necesidad de nuevos enfoques para la gestión de los datos

Los sistemas de información y las tecnologías de bases de datos proporcionan una manera eficiente de manejar los datos, pero para poder aprovechar sus beneficios, los datos tienen que estar integrados y unificados. Los sistemas de información son fiables cuando pueden garantizar, entre otras cuestiones, coherencia y ausencia de redundancia de datos, y esto sólo se puede conseguir si los datos están estructurados [52]. Desafortunadamente, como hemos mencionado anteriormente, la información en el área de la construcción está muy desestructurada debido, entre otros motivos, a la diversidad de herramientas que se utilizan para gestionar los datos y a la diversidad de participantes involucrados en el proceso. Este problema está presente tanto en los datos de los proyectos actuales como en los datos históricos.

El **problema de la integración de datos** es un reto en los proyectos que se están llevando a cabo, donde los datos relativos al diseño coexisten con los datos relativos a la ejecución. En cuanto a la integración de datos, la tecnología más prometedora es BIM ya que facilita la colaboración e integración de los datos desde la fase de diseño hasta la terminación de éste. En los últimos años, BIM está soportando una gran cantidad de tareas para el diseño y ejecución, tales como: mediciones, simulación, planificación, instalaciones y detección de conflictos, entre otras. Sin embargo, en la actualidad el uso de BIM en la gestión de

proyectos no es tan extendido como se desea, debido a las limitaciones de interoperabilidad, falta de conocimiento específico y necesidad de cambio en cuanto al método de trabajo. Además, para incrementar las capacidades de BIM es necesario la integración con otras tecnologías, como por ejemplo GIS [134]. Esta integración genera herramientas más sofisticadas pero, al mismo tiempo, se potencian los problemas de interoperabilidad e intercambio de información [133]. En cualquier caso, es evidente que los esfuerzos de generación de datos en el entorno BIM pueden ser inútiles si no van acompañados de un **proceso de estandarización**. De lo contrario, los beneficios serán parciales y la deseada interoperabilidad será muy limitada.

Desafortunadamente, la realidad es que sigue habiendo una falta de interoperabilidad y varios estándares compitiendo para la gestión de los datos [28]. En la ausencia de dichos estándares de referencia, los diferentes participantes del ámbito de la construcción desarrollan y aplican sus propias reglas. Como resultado, la documentación del proyecto es muy heterogénea: irregular, ambigua y desestructurada. Esta heterogeneidad se puede observar en uno de los documentos más importantes del proyecto, el presupuesto, que es diferente tanto en la estructura como en la redacción y definición de las actividades. Estos hechos impiden el acceso, intercambio y búsqueda de información de una manera automática e integrada. Además, dificultan la tarea de comparación de proyectos similares que han sido elaborados por diferentes profesionales [196]. En consecuencia, la gestión de la información de los proyectos en base a los presupuestos es ineficiente.

Como se ha anticipado en la introducción, la resolución de este problema es uno de los objetivos a los que nos enfrentamos en esta memoria, ya que es de especial interés por la importancia de los datos contenidos en este documento y por la relación que tienen éstos con otras tareas en el ámbito de la gestión de los proyectos de edificación.

Además del enfoque de estandarización, son necesarias otras tecnologías para hacer frente a la integración de información y los problemas de interoperabilidad para lograr con éxito la colaboración eficiente y eficaz. La **integración de sistemas** es un requisito previo muy importante [214]. Para alcanzar este objetivo, se ha utilizado el formato IFC, que facilita la interoperabilidad entre software en el dominio de la construcción. Sin embargo, existen problemas prácticos durante el proceso de intercambio de información entre los distintos sistemas, como por ejemplo, la pérdida o cambio de información [132]. Además, IFC por sí mismo no es suficiente para permitir la interoperabilidad con los sistemas que no son del entorno AEC, o con sistemas que producen datos de forma dinámica durante la fase de mantenimiento (por ejemplo, sensores) [74, 134].

La web semántica, ofrece una tecnología importante que puede contribuir a resolver el problema de interoperabilidad mencionado anteriormente. Sin embargo, la web semántica se está empezando a implantar en el ámbito de la construcción, y resulta difícil encontrar la forma adecuada de utilizar esta tecnología de forma eficiente sin emplear gran cantidad de tiempo [133]. Aunque las ontologías juegan un papel fundamental en la integración de la información y para permitir la **interoperabilidad semántica**, hasta donde nuestro conocimiento alcanza, no hay una ontología completa que abarque todas las tareas de la gestión de proyectos [91, 152]. A pesar de las contribuciones de estos enfoques, la integración de información sigue siendo un reto importante en el ámbito de la construcción [134].

En cuanto a la integración de los datos que se generan durante el proceso de ejecución, en los últimos años, la investigación se ha centrado en facilitar el proceso de adquisición de datos de la obra mediante el uso de tecnologías de seguimiento, que permiten la recolección de datos en tiempo real. Estas tecnologías han adquirido un nivel bastante aceptable en cuanto al rendimiento, sin embargo, la captura de los datos de la obra todavía tiene que afrontar importantes dificultades para poder detectar las operaciones del proceso. Por ejemplo, falta de detalle en el modelo de referencia, gran desglose de actividades en la programación, oclusiones estáticas y dinámicas, difícil comprensión de la escena, falta de precisión y reconocimiento simultáneo [111, 209]. Estos inconvenientes hacen que la mayoría de los datos de la obra se tengan que introducir manualmente y, por tanto, esta situación está lejos de ser una **adquisición automática de datos**. Como resultado, los datos se generan a través de diversas fuentes, resultando en una mala compatibilidad de los datos.

Los datos que se generan en los proyectos que se están ejecutando son importantes, no sólo para alimentar las bases de datos operacionales, sino también para alimentar repositorios integrados de datos históricos. La construcción de estos repositorios será más fácil cuando la normalización y la adquisición automática de datos sea una realidad. Sin embargo, los datos históricos que se tienen en la actualidad generados con las metodologías convencionales tienen que ser manejados de alguna manera y, no es una tarea fácil. En primer lugar, como se ha mencionado, los datos en la mayoría de los casos están desestructurados en diversos formatos; por otro lado, aunque compartan formato, cada profesional involucrado en el proceso utiliza su propio lenguaje y forma de organizar los datos. La razón es que la comunidad de profesionales que intervienen en el proceso del proyecto es muy diversa y, esa diversidad se refleja claramente en el lenguaje utilizado para describir los proyectos, a pesar de ser un lenguaje muy técnico.

El documento del presupuesto es un claro ejemplo de este problema, ya que cada presupuesto es muy diferente en su redacción y en la estructura en función del profesional que lo elabora, incluso perteneciendo a la misma empresa o al mismo estudio. En este sentido, el uso de técnicas de Minería de datos y métodos de clasificación inteligentes puede ayudar a automatizar el proceso de **extracción de información en documentos existentes** con el fin de ser incorporados en un almacén centralizado.

El desarrollo de la investigación en este sentido, es de vital importancia, ya que, a pesar de los esfuerzos necesarios de estandarización, el tipo de documento actual no va a desaparecer en un futuro próximo debido a la naturaleza tradicional de la industria de la construcción. Por estos motivos, nuestro interés se centra especialmente en la extracción y clasificación automática de los datos contenidos en el documento del presupuesto, con el fin de afrontar con garantías el objetivo de incorporar estos datos en un almacén común de manera integrada para dar soporte a la toma de decisiones.

El procesamiento de datos, que va de la mano de la adquisición de datos, también debe ser afrontado. La cantidad de datos disponibles ha aumentado gradualmente con la creciente complejidad de los proyectos de construcción y el gran número de participantes que intervienen en ellos. Por otra parte, a medida que se avanza en el reto de la adquisición automática de datos, el problema es aún mayor. Hacer frente a este problema de *big data* requiere más avances tecnológicos para gestionar, analizar, visualizar y extraer información útil a partir de grandes volúmenes de datos. Sin embargo, el procesamiento y la reducción de datos, para producir conclusiones significativas y la integración de datos procedentes de diversas fuentes, sigue siendo un importante obstáculo, como por ejemplo, para el desarrollo de una solución que permita el seguimiento del proceso de manera automatizada para la gestión de las obras de construcción [211, 214] o, herramientas que permitan el acceso y gestión de los datos de manera estructurada con el fin de dar soporte a la toma de decisiones.

Con los objetivos que se han planteado en esta memoria, es posible la construcción de repositorios integrados que se nutran no sólo de nuevos documentos bien estructurados, sino también de antiguos documentos cuyos datos se clasifican de manera automática en una estructura de referencia. Además, no sólo el manejo de los datos es necesario, sino que la extracción de información a partir de grandes conjuntos de datos se está convirtiendo en una necesidad cada vez más importante en cualquier entorno empresarial y, en concreto, en el ámbito de la construcción. Estas razones nos han motivado a afrontar el último objetivo de esta memoria, el desarrollo de herramientas que

permitan, por un lado, el manejo de los datos de manera estructurada e integrada, y por otro lado, realizar consultas flexibles sobre los datos almacenados.

Finalmente, las técnicas de **visualización** están aumentando y se están convirtiendo imprescindibles en las distintas tareas de gestión, sobre todo en las relativas a seguridad, planificación y seguimiento del proceso, y calidad. A pesar de la viabilidad de las técnicas de visualización, sigue siendo un reto; el manejo eficiente de las grandes cantidades de datos que se recogen de manera automática en la obra; y la complejidad computacional para analizar con detalle las desviaciones que se producen entre lo que se está ejecutando realmente y lo que se tenía que haber ejecutado [108]. Por otra parte, para proporcionar la información al completo del estado real del proceso de construcción, se necesita extraer otro tipo de información además de la geométrica, tales como, información relativa a los materiales e información de las relaciones o conexiones espaciales entre los distintos elementos [82]. En general, el uso de este tipo de herramientas es limitado en la actualidad, ya que no permiten la visualización de elementos que van más allá de los estructurales, tales como, elementos dinámicos o temporales [54].

La revisión de la literatura y el análisis de las propuestas nos ha permitido identificar los retos para garantizar el tratamiento de los datos de manera integrada durante el proceso de gestión de los proyectos. En los siguientes capítulos de esta memoria, presentamos las propuestas que abordan los problemas de: integración de datos, adquisición de datos automática y consulta de uno de los documentos más importantes en el ámbito de la gestión de proyectos, el presupuesto.

MODELO FORMAL PARA LA ESTRUCTURA DEL PRESUPUESTO

RESUMEN: En este capítulo se propone un modelo formal para el almacenamiento integral y estructurado de las descripciones de partida incluidas en el presupuesto. Además, en base al modelo propuesto e integrando conocimiento experto, se presenta una estructura de referencia para proyectos españoles la cual permite organizar la información del presupuesto para su posterior procesamiento y/o toma de decisiones.

Para ello, en primer lugar analizamos el procedimiento que siguen los profesionales de la construcción para la elaboración del documento del presupuesto. A continuación, presentamos la formalización del modelo y, por último, proponemos una instancia de ese modelo para ser usada como estructura de referencia.

ÍNDICE

3.1	Introducción	58
3.2	Descripción del problema	59
3.2.1	El proceso de elaboración del presupuesto	59
3.2.2	Problema de la ausencia de un modelo común	62
3.2.2.1	Ejemplo ilustrativo	63
3.3	Formalización del modelo de datos	65
3.3.1	Descripción de partida	66
3.3.2	Estructura de descomposición del trabajo	67
3.3.3	Relación de Parentesco	68
3.3.4	Elementos de la Estructura de Descomposición del Trabajo	70
3.4	Una estructura de referencia basada en conocimiento experto	72
3.4.1	Nivel L ₂ : Capítulos	73
3.4.2	Nivel L ₃ : Subcapítulos	75
3.4.3	Nivel L ₄ : Descripciones de Partidas	77
3.5	Conclusiones	77

3.1 INTRODUCCIÓN

La gestión de proyectos de construcción es una tarea compleja, independientemente del tamaño o tipología de proyecto, ya que implica la manipulación de grandes cantidades de datos que se generan durante todo el ciclo de vida del proyecto. Como hemos visto con anterioridad en esta memoria, en las últimas décadas se han hecho grandes esfuerzos para facilitar la gestión de los datos mediante el uso de las Tecnologías de la Información y la Comunicación. Sin embargo, el análisis de la literatura sigue indicando la presencia de grandes retos abiertos en relación con la gestión eficiente y automática de los datos asociados a los proyectos.

Desde un punto de vista general en el ámbito de los sistemas de información, un requisito fundamental para lograr el tratamiento automatizado de los datos con garantías, es que éstos tienen que estar estructurados y unificados en un almacén de datos centralizado. Desafortunadamente, los datos de los proyectos de edificación no están convenientemente estructurados debido a varios motivos, entre los que destacan: la diversidad de herramientas que se utilizan para elaborar los proyectos, la diversidad de participantes involucrados en el proceso, y la falta de estándares de referencia para organizar dicha información, entre otros. Como resultado, la documentación del proyecto es irregular, ambigua y desestructurada.

Esta heterogeneidad se refleja notablemente en uno de los documentos más importantes del proyecto, el presupuesto, que sirve de guía de referencia en todo el proceso de construcción. La información contenida en este documento aporta información de gran interés en el ámbito de la gestión del proyecto, ya que, además de aportar la información económica de la obra, establece la división y definición de las distintas actividades que se tienen que realizar para completar el objetivo del proyecto. Sin embargo, como ya hemos comentado, el actual proceso que se sigue para la elaboración de este documento, junto con la falta de estándares que definan una estructura común, hace que cada presupuesto sea diferente tanto en su estructura como en la redacción y definición de las actividades. Todo ello provoca que la gestión de la información de los proyectos en base a los presupuestos sea ineficiente y complica la toma de decisiones durante el proceso de construcción.

Durante la toma de decisiones, es necesario consultar la información contenida en el documento del presupuesto y relacionarla con experiencias anteriores; para ello, toda la información debiera estar *bien* almacenada. Sin embargo, al no existir una estructura común

para dicho documento, esa información suele almacenarse en ficheros independientes, en muchos casos semi-estructurados e inconexos.

Por tanto, la falta de una estructura común que permita reducir la heterogeneidad existente en el documento del presupuesto, así como el almacenamiento integral de dicha información plantean un reto interesante de cara al tratamiento automatizado de los datos en el ámbito de los proyectos de edificación.

En primer lugar, en la sección 3.2 se analiza el procedimiento que siguen los profesionales de la construcción para la elaboración del documento del presupuesto. A continuación, en la sección 3.3 se propone un modelo formal para organizar los datos asociados al presupuesto, habilitando de esta manera su almacenamiento estructurado en un sistema de información. Por último, en la sección 3.4, proponemos una instancia de ese modelo para ser usada como estructura de referencia.

3.2 DESCRIPCIÓN DEL PROBLEMA

Para establecer la dimensión del problema al que nos enfrentamos, se va a explicar en esta sección el procedimiento que siguen los profesionales de la construcción para la elaboración de uno de los documentos del proyecto, el presupuesto. Posteriormente, mediante ejemplos reales, se identificarán los problemas con los que se enfrentan los responsables de la gestión de proyectos cuando necesitan acceder y consultar los datos almacenados, problemas fundamentalmente debidos a la falta de un modelo de datos común y universal.

3.2.1 *El proceso de elaboración del presupuesto*

En la actualidad, el arquitecto es el profesional responsable de la elaboración de los proyectos de obra que tengan por objeto la construcción de edificios para los usos indicados en el *grupo a* del *apartado 1* del artículo 2 de la LOE [153]. Entre los documentos que han de componer un proyecto, está el mencionado presupuesto, del que forman parte las mediciones. El documento de mediciones, que según el Código Técnico de la Edificación [1]: “desarrolla las partidas, agrupadas en capítulos, conteniendo toda las descripciones técnicas necesarias para su especificación y valoración”, es el documento que, completado con la valoración de la ejecución material de todas las partidas, da lugar al presupuesto.

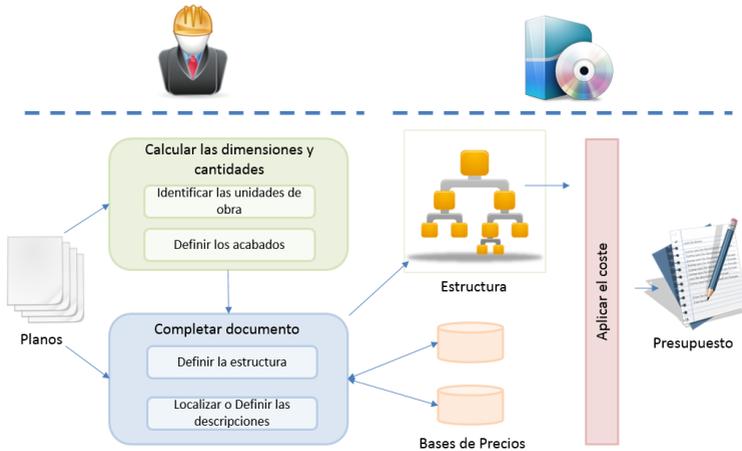


Figura 2: Esquema del proceso de elaboración del presupuesto

En la Figura 2 se puede observar el proceso actual de elaboración del presupuesto. A continuación describimos las etapas del proceso:

1. *Identificación de unidades de obra*: una vez que se tiene la documentación gráfica del proyecto (planos) se procede, como inicio de esta fase, a la identificación de las diferentes unidades de obra que son necesarias para ejecutar el proyecto, así como a la cuantificación de cada una de ellas, lo que se denomina mediciones. En esta etapa, además de los planos, que definen geoméricamente la obra que se va a construir, se debe consultar el resto de los documentos que definen y completan el proyecto, las memorias y pliegos de condiciones con sus especificaciones técnicas, ya que el conjunto de todos estos documentos aporta la información necesaria para la correcta elaboración del documento del presupuesto. En estos documentos, se encuentra información detallada en relación a la forma, materiales, medios auxiliares, cumplimiento de normativa técnica, etc., con los que se ejecutarán las distintas unidades de obra. En general, esta primera etapa consiste principalmente en el estudio del proyecto con el máximo nivel de detalle, desde el punto de vista de su ejecución, ya que hay que identificar todas las unidades de obra necesarias para, posteriormente, cuantificarlas.
2. *Elaboración del documento*: una vez identificadas y conocidas todas las unidades de obra que se tienen que realizar, el arquitecto define la estructura que ha de tener el documento del presu-

puesto, descomponiéndolo en diferentes niveles hasta llegar al nivel más bajo donde se sitúan las descripciones de partida. Hoy en día, para la elaboración de este documento existen muchas herramientas comerciales diseñadas para facilitar la realización del mismo. Como por ejemplo, se pueden citar entre otras, Arquimedes [75], Premeti [194], o - la más utilizada - Presto [196]. A diferencia de las dos primeras que sólo contienen una única base de precios, Presto proporciona soporte para ser utilizado por diversas bases de precios realizadas por otros proveedores, tanto pertenecientes a organismos públicos como a entes privados.

En la Tabla 2 se pueden observar algunas de las bases de precios más utilizadas por los arquitectos para elaborar el presupuesto del proyecto. Estas bases de precios contienen miles de descripciones de partidas con sus precios correspondientes, agrupadas en distintos niveles y descritas según el criterio de un equipo de profesionales, designado por el organismo que las ha desarrollado. Además, cada base de precios tiene su propio sistema de codificación, en función de la estructura desarrollada, no existiendo tampoco una estructura común para agrupar las distintas unidades de obra en los correspondientes capítulos.

Con la ayuda de estas bases de precios, gracias a los datos y experiencias de proyectos anteriores y, por supuesto, mediante su propio conocimiento y criterio, el arquitecto completa cada uno de los niveles de la estructura que él mismo ha definido anteriormente. En este paso, el arquitecto puede copiar directamente la descripción de alguna partida de las bases de precios que haya consultado, de proyectos anteriores de análogas características o, también, redactarla con sus propios vocabulario y estilo gramatical.

El proceso que se sigue para la elaboración de este documento, junto con la falta de estándares para elaborarlo, así como la entrada en concurso de los criterios específicos del autor del proyecto, hace que cada presupuesto sea diferente tanto en su estructura como en la descripción de las diferentes unidades de obra. De hecho, esta diversidad es habitual incluso para un mismo arquitecto en diferentes proyectos. Esta heterogeneidad dificulta la inserción de esta valiosa información en un almacén de datos y complica el acceso de los usuarios a los datos necesarios en cada momento. En la siguiente sección se detallan los problemas debidos a la falta de un modelo de referencia.

Bases de precios	URL
Colegio de Arquitectos Técnicos y Aparejadores de Guadalajara	[67]
Base de Costes de la Construcción de Andalucía	[129]
Base de Precios de la Construcción de la Comunidad de Madrid	[27]
Fundación de Estudios para la Calidad en la Edificación de Asturias	[25]
Instituto de Tecnología de la Construcción de Cataluña	[26]
Base de Precios de la Construcción de la Junta de Extremadura	[68]
Precios de Edificación y Obra Civil en España	[195]
Base de Precios de la Construcción de Castilla y León	[119]

Tabla 2: Bases de precios

3.2.2 Problema de la ausencia de un modelo común

En la sección anterior se ha descrito el procedimiento para la elaboración del presupuesto del proyecto y se han identificado las principales causas que provocan la heterogeneidad en este documento, incluso para proyectos de similares características.

La causa fundamental para ello es la falta de estándares universales para la organización de la información del presupuesto. En algunos países existen estándares locales para organizar la información de los proyectos de construcción, como por ejemplo, *MasterFormat*[70], que se utiliza para la mayoría de los proyectos de construcción en Estados Unidos y Canadá. Este estándar, desarrollado por el *Construction Specifications Institute*[72] y *Construction Specifications Canada*[69], presenta una división jerárquica con códigos y títulos para organizar la información de construcción, como por ejemplo: manuales, costes, planos, anotaciones y especificaciones técnicas, etc. La estructura, hasta el año 2004, constaba de 16 divisiones principales; sin embargo, en la última versión, esta estructura cuenta con un total de 50, debido a la reestructuración y creación de nuevas divisiones.

Además de *MasterFormat*, otra iniciativa que se utiliza en estos dos países para estructurar la información de los costes es *UniFormat*[71], desarrollado por los mismos organismos. La división de niveles es menos extensa y es más específica en cuanto a los elementos constructivos del proyecto. De manera similar, la norma Alemana DIN 276-1[99] y la inglesa NRM2[206], estructuran los diferentes trabajos de construcción mediante una estructura jerárquica.

Por desgracia, en España, no existen estándares a los que remitirse a la hora de organizar la información de los proyectos de forma es-

estructurada. En general, dicha información suele estructurarse de forma jerárquica, contemplando distintos niveles de detalle de las actividades a ejecutar en un proyecto, pero carentes de estructura común. Es decir, no existe una estructura definida ni para la organización de la información ni para la denominación de los distintos niveles de detalle de ésta.

Esta carencia de estructura común, como ya hemos mencionado, provoca que la gestión de la información de los proyectos sea ineficiente, tanto por la falta de estructura en los documentos como por la derivada falta de apoyo de software basado en el almacenamiento de dicha información. Al no tener una estructura común, la información se almacena en ficheros independientes de tal forma que, por ejemplo, llevar a cabo una consulta asociada al coste de una actividad para un proyecto, donde la información asociada a los presupuestos no está almacenada en un repositorio común, requiere más tiempo del deseado y está sujeta a la presencia de inconsistencias. Algo tan simple para la toma de decisiones como es hacer una consulta sobre *el coste medio de una cierta actividad para los proyectos de viviendas unifamiliares* supone buscar en cada uno de los ficheros que contienen los presupuestos, localizar en cada uno la actividad deseada y llevar a cabo la consulta.

3.2.2.1 Ejemplo ilustrativo

A continuación, mediante varios ejemplos, se van a detallar los inconvenientes a los que se enfrentan los responsables de la gestión de proyectos a la hora de acceder a los datos del presupuesto para tomar decisiones. Para este propósito, se han seleccionado tres proyectos de edificación de viviendas unifamiliares con similares características. En la Tabla 3 se puede observar la estructura de primer nivel presente en el presupuesto de cada uno de éstos proyectos.

Como se puede observar, pese a tratarse de obras de similar complejidad, la descomposición que se plantea tiene distintos niveles de granularidad y utiliza un lenguaje distinto para describir estos niveles. Por ejemplo, el primer proyecto nombra al primer capítulo como *Movimiento del terreno*, el segundo proyecto, *Acondicionamiento del terreno* y el tercero, *Movimiento de tierras*.

Estas diferencias se producen por las razones que se han explicado en la sección anterior. El arquitecto de cada uno de estos proyectos ha definido la estructura en base a la estructura de alguna base de precios o creándola él mismo según su criterio. Como se ha comentado anteriormente, cada base de precios utiliza un lenguaje de descripción distinto. Por ejemplo, la base de precios *Preoc* [195] denomina a dicho

Proyecto 1	Proyecto 2	Proyecto 3
Movimiento del terreno	Demoliciones	Movimiento de tierras
Red horizontal de saneamiento	Acondicionamiento del terreno	Red horizontal de saneamiento
Cimentaciones	Estructuras	Cimentaciones
Estructuras	Fachadas	Estructuras
Cerramientos	Instalaciones	Cerramientos
Particiones interiores	Impermeabilización y aislamiento	Particiones interiores
Cubiertas	Cubiertas	Cubiertas
Aislamientos	Revestimientos	Aislamientos
Impermeabilizaciones	Señalización y equipamiento	Impermeabilizaciones
Revestimientos	Urbanización	Revestimientos
Carpintería interior	Varios	Alicatados y chapados
Carpintería exterior		Pavimentos
Cerrajería		Carpintería interior
Vidriería		Carpintería exterior
Falsos techos		Cerrajería
Pinturas		Vidriería
Electricidad		Falsos techos
Fontanería		Pinturas
Calefacción		Electricidad
Climatización		Fontanería
Transportes		Calefacción
Telecomunicaciones		
Protección contra incendios		
Instalaciones especiales		
Urbanización		
Varios		

Tabla 3: Estructuración en el primer nivel de organización del documento de presupuesto correspondiente a tres proyectos de edificación de viviendas unifamiliares

capítulo *movimiento de tierras* mientras que la base de precios *Centro* [67] lo nombra como *acondicionamiento del terreno*. Además, el arquitecto es libre de utilizar su propio vocabulario.

A continuación, para seguir profundizando en el problema, nos vamos a centrar en un capítulo muy específico dentro del proyecto, el relacionado con los mencionados *trabajos del terreno*. Para este propósito, se han seleccionado tres actividades que aparecen en los tres proyectos de ejemplo: *Desbroce y limpieza a mano*. Como se puede observar en la Tabla 4, esta actividad se encuentra localizada dentro del mismo capítulo, aunque se utilizan distintas descripciones lingüísticas para referirse a ella. Sin embargo, en cada proyecto, el camino a seguir para localizar la actividad es distinto como se puede observar en la Tabla.

A continuación, se presenta otro ejemplo donde la misma actividad se encuentra localizada en capítulos que se dedican a actividades de diferente naturaleza. Como se puede observar en la Tabla 5, la acti-

Proyecto	Capítulo	Subcapítulo	Descripción de partida
Proyecto 1	Movimiento del terreno		Desbroce y limpieza a mano
Proyecto 2	Acond. del terreno	Limpieza y desbroce	Desbroce y limpieza del terreno a mano
Proyecto 3	Mov. de tierras	Acond. y prep. del terreno	Desbroce y limpieza del terreno, con medio manuales

Tabla 4: Ejemplo de descripción de partida de trabajos del terreno en los tres proyectos de ejemplo

vidad: “*Geotextil danofelt py200*” se encuentra dentro del capítulo de *acondicionamiento del terreno* en el proyecto 2 y dentro del de *impermeabilizaciones* en el proyecto 3. En el proyecto 1 no se encuentra esta misma actividad, aunque en el caso de estar, lo lógico sería que se localizara dentro del capítulo de *impermeabilizaciones* al igual que en el proyecto 3.

Proyecto	Capítulo	Subcapítulo	Descripción de partida
Proyecto 2	Acond. del terreno	Geotextiles	Geotextil DANOFELT PY-200
Proyecto 3	Impermeabilizaciones	Drenajes	Geotextil DANOFELT PY 200

Tabla 5: Ejemplo de descripción de partida de geotextil en los tres proyectos de ejemplo

Esta diversidad dificulta el acceso integrado a la información para dar soporte el proceso de toma de decisiones de futuros proyectos. Integrando datos de todos los proyectos de la empresa, los usuarios podrán resolver consultas más ambiciosas y útiles, como por ejemplo: *El coste medio del capítulo movimiento de tierras de los proyectos de un determinado ámbito geográfico*. Y esto lo podrán hacer independientemente de si los proyectos que nutren la solución de la citada consulta incluyen tal cual descrito dicho trabajo de la obra o lo tienen descrito de otra manera.

Para hacer un aporte en este sentido es conveniente proponer una estructura de referencia para el documento de presupuesto. Para ello, como parte de los objetivos de esta memoria, primero formalizamos un modelo de datos para dicho documento y, después, en base a dicho modelo proponemos una estructura de referencia para ser usada en proyectos de construcción en España.

3.3 FORMALIZACIÓN DEL MODELO DE DATOS

El procedimiento más utilizado por los expertos de la construcción para organizar las actividades que conforman un proyecto, a lo largo

del tiempo, es la *Estructura de Descomposición del Trabajo* (del inglés, Work Breakdown Structure, WBS). Por medio de una *estructura de descomposición del trabajo*, las actividades se clasifican y agrupan en forma de árbol jerárquico de conceptos, que se descomponen sucesivamente en otros conceptos inferiores aportando mayor nivel de detalle. Este tipo de estructuras responde a la estrategia de diseño descendente y, en el ámbito de la construcción, carecen de formalización.

En lo que sigue de este capítulo, se procede a formalizar esta estructura desde un punto de vista de minería de textos que, como veremos posteriormente en la memoria, será la base para afrontar tareas de clasificación automática de los datos contenidos en una estructura de descomposición del trabajo.

El nivel más bajo de la jerarquía corresponde al de las descripciones de partida.

3.3.1 Descripción de partida

Las descripciones de partida son secuencias de palabras que describen las actividades que se tienen que hacer para cumplir el objetivo del proyecto. Formalmente,

Definición 1 Sea T el conjunto de términos que pueden ser utilizados para describir actividades en el presupuesto. Una descripción de partida (WD) es una lista ordenada de términos $\delta = \{t_1, \dots, t_k\}$, $t_i \in T$ que detalla una actividad en el proyecto.

Según el problema en el que se considere, puede interesar tener una forma preprocesada de dicha lista de términos, que llamaremos δ^s . Por ejemplo, como los términos en una descripción δ pueden aparecer más de una vez, podemos considerar que δ^s es la lista de términos ordenada que se obtiene considerando sólo la primera ocurrencia de cada término en la descripción de partida. Según la tarea que se quiera realizar con las descripciones de partidas, el preprocesamiento de los datos puede ser diferente.

Ejemplo 1 Vamos a considerar un proyecto de una vivienda unifamiliar, donde una de las tareas que se tienen que realizar es: Excavación, en apertura de caja, en terreno blando, realizada con medios mecánicos, incluso perfil hasta una profundidad máxima de 50 cm. Medida en perfil natural.

Asociada a esta tarea, podemos considerar la siguiente descripción de partida:

$\delta_1 = \{\text{excavación, apertura, caja, terreno, blando, hacer, medio, mecánico, incluso, perfil, profundidad, máximo, medir, perfil, natural}\}$

cuya forma corta es:

$\delta_1^s = \{\text{excavación, apertura, caja, terreno, blando, hacer, medio, mecánico, incluso, perfil, profundidad, máximo, medir, natural}\}$

Más adelante, en la sección 4.2.1, se detalla el proceso para obtener la descripción de partida a partir del texto que describe la actividad.

3.3.2 Estructura de descomposición del trabajo

Una vez que hemos descrito las hojas, podemos pasar a la formalización de la estructura completa. Como hemos comentado anteriormente, una estructura de descomposición del trabajo, es una descomposición jerárquica, por medio de la cual las actividades se clasifican y agrupan, descomponiéndose sucesivamente en otras inferiores que aportan mayor nivel de detalle.

Es decir, una estructura de descomposición de trabajo es una jerarquía en árbol donde cada nivel inferior representa una definición más detallada de la actividad, hasta llegar al nivel de las hojas que se corresponden con las descripciones de partidas del proyecto. Formalmente,

Definición 2 *Una estructura de descomposición del trabajo, es una descomposición jerárquica, en forma de árbol, con n niveles, $\omega = \{L_1, \dots, L_n\}$, donde:*

- *El nivel L_n es un conjunto de WD, $L_n = \{\delta_1, \dots, \delta_m\}$.*
- *Cada nivel $L_i, i = 1 \dots n - 1$ es un conjunto de nodos $\{GT_i^1, \dots, GT_{SIZE_i}^1\}$, donde cada GT_j^i se usa para agrupar nodos del nivel L_{i+1} .*
- *Adicionalmente:*

$$m \geq SIZE_{n-1}$$

$$SIZE_{i+1} \geq SIZE_i$$

$$SIZE_1 = 1$$

Ejemplo 2 Vamos a considerar de nuevo el proyecto de la vivienda unifamiliar. Para este proyecto, podemos considerar la siguiente estructura de descomposición del trabajo:

- Proyecto_{VivUni} = {L₁, L₂, L₃, L₄}
- L₁ = {GT₁¹ = Todo}
- L₂ = {GT₁² = Movimiento de tierras, GT₂² = Cimentaciones, GT₃² = Estructura, GT₄² = Albañilería, GT₅² = Instalaciones, ...}
- L₃ = {GT₁³ = Actuaciones previas, GT₂³ = Excavación, ..., GT₁₀³ = Encofrados, ..., GT₂₀³ = Divisiones interiores, ..., GT₂₅³ = Sanitarios, ...}
- L₄ = {δ₁ = {excavación, apertura, caja, terreno, blando, hacer, medio, mecánico, incluso, perfil, profundidad, profundidad, máximo, medir, perfil, natural}, δ₂₀ = {encofrado, desencofrado, zapata, zuncho, tablero, madera, pino}, ..., δ₉₀ = {tabique, ladrillo, hueco, recibir, pegamento, especial, soporte, banda, elástico, incluso, limpiar, medio, auxiliar}, ..., δ₂₂₅ = {instalar, lavabo, suspendido, tanque, empotrado, sistema, doble, descarga, incluso, fijación, conexión, tabique, ayuda, albañilería, ...}}

En la Figura 3 se ilustra la estructura de descomposición del trabajo del proyecto del ejemplo 2.

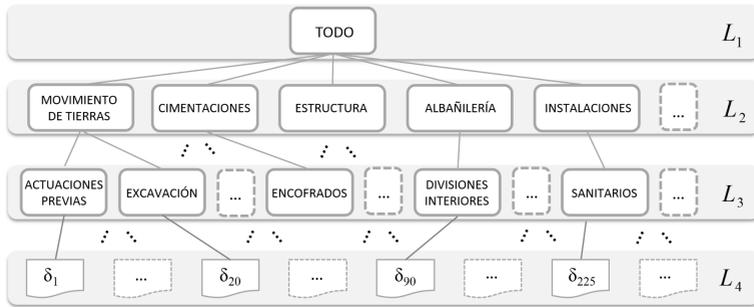


Figura 3: Estructura de descomposición del trabajo del ejemplo 2

3.3.3 Relación de Parentesco

Entre los nodos de los niveles de la estructura de descomposición de trabajo descrita anteriormente, se establece una relación de parentesco que define las relaciones entre niveles en la jerarquía. Formalmente,

Definición 3 Sea $\omega = \{L_1, \dots, L_n\}$ una estructura de descomposición del trabajo. Para cada dos niveles consecutivos L_i, L_{i+1} , se define una relación de parentesco KS^i , $KS^i \subseteq L_i \times L_{i+1}$, tal que:

- Si $i \in [1 \dots n - 2]$, $\langle GT_j^i, GT_k^{i+1} \rangle \in KS^i$, si, y sólo si, GT_k^{i+1} es un hijo de GT_j^i en ω .
- Si $i = n - 1$, $\langle GT_j^{n-1}, \delta_k \rangle \in KS^{n-1}$ si, y sólo si, δ_k es una descripción de partida hija de GT_j^{n-1} en ω .

Es decir, en los niveles superiores, la relación de parentesco establece relaciones padre-hijo entre nodos de agrupamiento, mientras que en el penúltimo nivel, se establece esta relación entre nodos de agrupamiento de ese penúltimo nivel y las hojas de la estructura, es decir, las descripciones de partida.

La estructura jerárquica en árbol requiere que cada nodo sólo tenga un padre, a excepción de la raíz, GT_1^1 , que no tiene ninguno. Por consiguiente, la siguiente restricción se debe satisfacer al plantear las relaciones de parentesco:

Propiedad 3.1 (Padre único) $(\langle a, b \rangle \in KS^i \wedge \langle c, b \rangle \in KS^i) \rightarrow a = c$

Ejemplo 3 Vamos a considerar de nuevo la estructura de descomposición del trabajo del ejemplo anterior correspondiente a la vivienda unifamiliar. Por ejemplo:

$$\begin{aligned} \langle \text{Todo, Preparación del terreno} \rangle &\in KS^1 \\ \langle \text{Todo, Albañilería} \rangle &\in KS^1 \\ &\dots \\ \langle \text{Preparación del terreno, Excavación} \rangle &\in KS^2 \\ \langle \text{Albañilería, Divisiones interiores} \rangle &\in KS^2 \\ &\dots \\ \langle \text{Excavación, } \delta_1 \rangle &\in KS^3 \\ \langle \text{Divisiones interiores, } \delta_{90} \rangle &\in KS^3 \end{aligned}$$

Como hemos comentado, la restricción de padre único evita la asignación de dos padres diferentes para el mismo hijo. En este ejemplo, si tenemos:

$$\langle \text{preparación del terreno, actuaciones previas} \rangle \in KS^2$$

entonces,

$$\forall x \neq \text{preparación del terreno, } \langle x, \text{actuaciones previas} \rangle \notin KS^2$$

La figura 4 muestra el modelo de datos jerárquico formalizado.

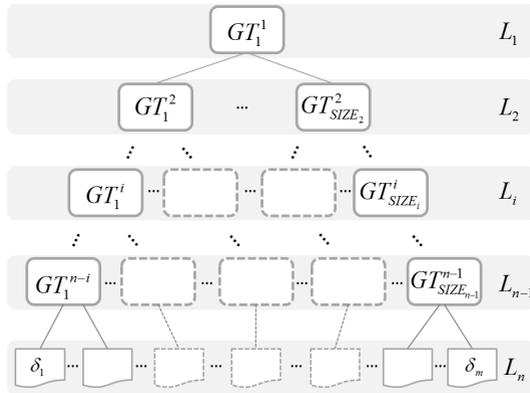


Figura 4: La estructura de descomposición de trabajo (WBS)

3.3.4 Elementos de la Estructura de Descomposición del Trabajo

En la estructura de descomposición del trabajo previamente definida, vamos a diferenciar dos elementos: por un lado, vamos a considerar una subestructura para agrupar actividades y, por otro, el conjunto conjunto de descripciones de partida.

Definición 4 (Estructura de Agrupación de Actividades) Sea $\omega = \{L_1, \dots, L_n\}$ una estructura de descomposición del trabajo asociada a un determinado proyecto. La Estructura de Agrupación de Actividades de ω , G_ω , es el subárbol de ω compuesto por los niveles $\{L_1, \dots, L_{n-1}\}$.

Definición 5 (Conjunto de descripciones de partida) Sea $\omega = \{L_1, \dots, L_n\}$ una estructura de descomposición del trabajo asociada a un determinado proyecto. El conjunto de descripciones de partida de ω , S_ω , es el conjunto de hojas del nivel L_n .

Ejemplo 4 Para la estructura de descomposición del trabajo del Ejemplo 2, tenemos que $G_\omega = L_1, L_2, L_3$ y $S_\omega = L_4$.

La Figura 5 ilustra los dos elementos de la estructura de descomposición de trabajo que hemos definido previamente. En la figura, los enlaces entre nodos de niveles consecutivos representan relaciones padre-hijo, definidas a través de la *relación de parentesco*.

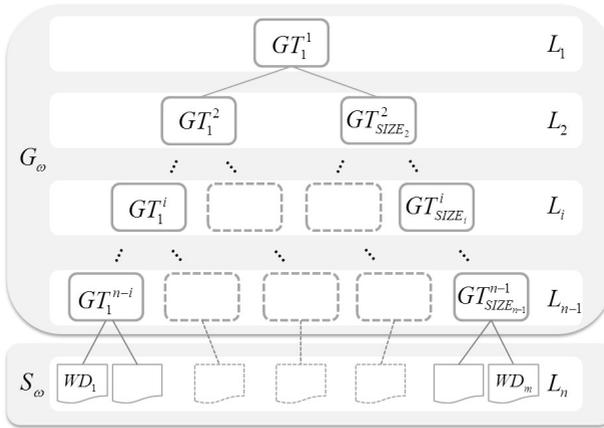


Figura 5: Elemento de la estructura de descomposición de trabajo (WBS)

Los dos elementos considerados en esta sección, responden a dos motivaciones distintas a la hora de informar sobre el proyecto:

- Por un lado, la estructura de agrupación de actividades permite establecer distintos conglomerados de tareas, con distintos niveles de granularidad.
- Por otro lado, el conjunto de descripciones de partidas aporta información detallada de las tareas que hay que realizar para completar el proyecto.

En este sentido, a la hora de plantear el almacenamiento de los datos de una estructura de descomposición del trabajo, como veremos, la estructura de agrupación de actividades será considerada para definir y organizar estructuras de almacenamiento, mientras que el conjunto de descripciones de partidas servirá de enlace final con los datos que se van a almacenar en relación con las tareas del proyecto.

Por este motivo, para utilizar la estructura de agrupación de actividades como soporte para el almacenamiento de los datos del presupuesto, es necesario llevar a cabo los siguientes objetivos:

- Establecer una estructura de agrupación de actividades de referencia G_ω^* , que será la estructura común que se utilizará en el sistema de información para soportar el almacenamiento de los datos contenidos en el documento del presupuesto. Este objetivo se abordará en la siguiente sección.

- Desarrollar un método de clasificación que extraiga cada descripción de partida $\delta \in S_\omega$ de un determinado presupuesto de entrada y automáticamente la clasifique como una de las hojas de la estructura de referencia G_ω^* , colocándola en el nodo adecuado GT_j^{n-1} de dicha estructura de referencia. Es decir, el método de clasificación deberá establecer adecuadamente la pareja $\langle GT_j^{n-1}, \delta \rangle$ al KS^{n-1} . Este objetivo se aborda en el siguiente capítulo de esta memoria.

En relación al primer objetivo, la existencia de una estructura de referencia para la elaboración del documento del presupuesto facilitaría la elección de la estructura de agrupación de actividades G_ω^* . Sin embargo, en la ausencia de estándares generales, es necesario definir una estructura de agrupación de actividades que se ajuste al uso de los expertos que utilizarán el sistema desarrollado. A continuación, se presenta una estructura de referencia en base al modelo formal que se ha presentado en esta sección.

3.4 UNA ESTRUCTURA DE REFERENCIA BASADA EN CONOCIMIENTO EXPERTO

Como se ha comentado en varias ocasiones en esta memoria, no existe un estándar de referencia en España para estructurar la información contenida en el documento del presupuesto ni para nombrar cada uno de los niveles de esta estructura. En la ausencia de dicho estándar, en este trabajo de investigación se propone una estructura de referencia que ha sido contrastada con un panel de expertos compuesto por ingenieros de edificación de la Universidad de Granada. Nuestro objetivo se centra en los proyectos de edificación; por lo tanto, la estructura contempla las actividades propias de este tipo de proyectos y que suponen una actividad real en el proceso de ejecución del proyecto.

La propuesta es una instancia del modelo de datos presentado en la primera parte de este capítulo y considera aspectos generales que tienen en común la mayoría de los presupuestos en España, tales como utilizar la mencionada estructura jerárquica y ordenar los elementos del primer nivel de la jerarquía según el proceso de ejecución del proyecto.

La estructura propuesta establece que la estructura de descomposición de trabajo organiza la información del documento en una jerarquía de cuatro niveles como se puede observar en la Figura 6. La estructura de agrupación de actividades G_ω^* de referencia está formada por tres

niveles, L_1 , L_2 y L_3 , que se denominan *Proyecto*, *Capítulo* y *Subcapítulo* respectivamente. El nivel donde se encuentra el conjunto de las descripciones de partidas S_ω corresponde al L_4 .

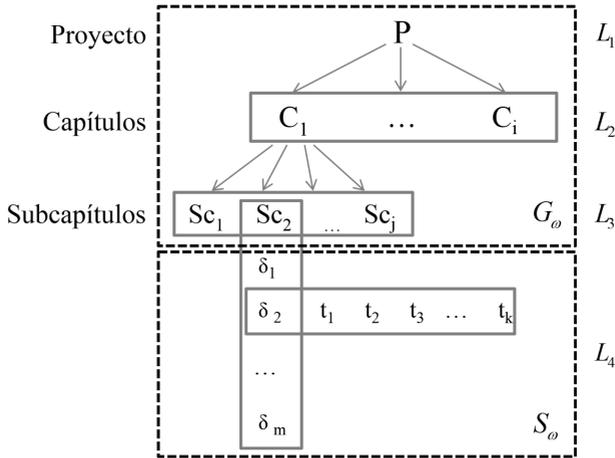


Figura 6: Propuesta de estructura jerárquica para los presupuestos en proyectos de edificación en España

En las siguientes subsecciones, se detallan los distintos elementos de la estructura de referencia propuesta. En la sección 3.4.1 se describe la división en capítulos, en la sección 3.4.2 se muestra la división en subcapítulos y en la sección 3.4.3 se detalla el nivel de las descripciones de partidas.

3.4.1 Nivel L_2 : Capítulos

En primer lugar se ha establecido la división de *capítulos*. Como veremos a continuación, se ha optado por una división de capítulos no muy extensa, porque, una estructura con muchos capítulos resulta inconveniente para organizar la información ya que aumenta de forma considerable la incertidumbre a la hora de insertar convenientemente las descripciones de partida. Para el orden de los capítulos se ha considerado el orden cronológico de las distintas actividades del proyecto desde un punto de vista general.

Según estas consideraciones, la división de capítulos propuesta para la estructura de referencia es la siguiente:

- *Trabajos previos:* Aparecerán todas las partidas necesarias para empezar la obra. Puede ser necesario la demolición de alguna construcción existente o la conservación.
- *Acondicionamiento del terreno:* En este capítulo se recogen las partidas de obra necesarias para adaptar el terreno a la configuración final que debe tener y que se especifique en el proyecto. En obras de edificación, este capítulo no tiene excesiva incidencia en el coste final del proyecto, pudiendo ser más importante en obras de urbanización y en obras civiles.
- *Red horizontal de saneamiento:* Son todas las partidas necesarias para la recogida y eliminación de las aguas usadas, fecales y pluviales de una construcción. Las que se encuentran en este capítulo, tienen carácter de horizontalidad, a diferencia de las verticales (bajantes, desagües) que se engloban en el capítulo de fontanería o cubiertas.
- *Cimentaciones:* Abarcará todas las partidas de obra que se refieren a los elementos de transmisión de cargas o solicitaciones del edificio o construcción al terreno.
- *Estructuras:* Lo constituye el conjunto de elementos, realizados con diferentes materiales, que tienen por objeto formar el armazón resistente en las construcciones, recibiendo las solicitaciones que junto con su propio peso, transmiten a la cimentación.
- *Albañilería:* Aparecerán todas las partidas referentes a cerramientos de viviendas tanto por el interior como por el exterior y trabajos relacionados con estas actividades.
- *Cubiertas:* Constituye el cierre superior de un edificio y estará ejecutado por el elemento estructural de sustentación, y el material de cubrimiento o cobertura.
- *Instalaciones:* Este capítulo abarca todas las instalaciones que se pueden instalar en obras de edificación a excepción de la red horizontal de saneamiento que debido a su pronta ejecución en la obra se ha separado del resto.
- *Aislamientos e impermeabilizaciones:* Abarcará las partidas necesarias para aislar el edificio tanto térmico como acústicamente y las necesarias para impermeabilizar las zonas correspondientes.
- *Revestimientos:* Este capítulo comprende aquellos trabajos necesarios para el recubrimiento de superficies verticales y horizontales, que servirán de apoyo para otros acabados posteriores como pinturas, aplacados o solados.

- *Carpintería interior (madera)*: Se incluirán en este capítulo, las diferentes unidades de obra que tiene por objeto cerrar los huecos. El material de todas las partidas será la madera.
- *Carpintería exterior y cerrajería*: Se incluirán en este capítulo, las diferentes unidades de obra que tiene por objeto cerrar los huecos exteriores y elementos de cerrajería tales como barandillas, pasamanos, celosías, etc.
- *Vidriería y elaborados sintéticos*: Se incluirán las partidas cuyo material principal es el vidrio y que sirven para cerrar espacios y dejar pasar la luz.
- *Pinturas*: Todas las partidas cuya finalidad es la de embellecer y acabar los paramentos y divisiones, así como otros elementos que sea necesario darle una terminación, para protegerlos de las oxidaciones o de los agentes externos agresivos.
- *Equipamiento, decoración y varios*: Aquellas partidas que sirven para el equipamiento y decoración del edificio. Es muy variado, ya que abarca partidas tan diferentes como, buzones, señalizaciones, mobiliario, jardín, etc.

3.4.2 Nivel L₃: Subcapítulos

A continuación, se ha establecido la división de cada uno de estos capítulos en un nivel inferior, los *subcapítulos*, que contienen actividades más específicas. Dependiendo de la naturaleza del capítulo, se ha establecido el número de subcapítulos. Se ha optado por no dividir estos subcapítulos en otro nivel inferior, ya que esta división tiene más sentido en las bases de precios donde se tienen que almacenar grandes cantidades de descripciones de partidas muy diferentes, sin embargo, para gestionar la información del presupuesto de un proyecto es más intuitivo y manejable disponer de sólo dos niveles. Por consiguiente, dentro de cada subcapítulo se disponen las diferentes *descripciones de partidas*.

Por ejemplo, para el capítulo 1 (*trabajos previos*) se ha establecido una división de 3 subcapítulos: consolidaciones, derribos y, cargas y transportes. Sin embargo, para el segundo capítulo (*Acondicionamiento del Terreno*) se ha establecido una división más extensa debido a la diversidad de actividades: preparación del terreno, explanación, excavación, refino, relleno, compactación, carga, transporte y canon.

C1	Trabajos previos	SC3	Cubiertas planas no transitables	
	SC1	Consolidaciones	SC4	Cubiertas planas transitables
	SC2	Derribos	SC5	Bajantes-Canalones
	SC3	Cargas y Transportes	SC6	Remates Varios
C2	Acondicionamiento del Terreno	C8	Instalaciones	
	SC1	Preparación del Terreno	SC1	Fontanería
	SC2	Explanación	SC2	Aparatos Sanitarios
	SC3	Excavación	SC3	Eléctrica e Iluminación
	SC4	Refino	SC4	Telecomunicaciones-Informática
	SC5	Relleno	SC5	Calefacción - ACS
	SC6	Compactación	SC6	AA Ventilación
	SC7	Carga	SC7	Gas
	SC8	Transporte	SC8	Elevación
	SC9	Canon	SC9	Protección
C3	Red Horizontal de Saneamiento	C9	Aislamientos e Impermeabilizaciones	
	SC1	Acometidas	SC1	Aislamientos
	SC2	Arquetas y Pozos	SC2	Impermeabilizaciones
	SC3	Sistemas de depuración	C10	Revestimientos
	SC4	Tuberías y Colectores	SC1	Falsos techos
	SC5	Sumideros	SC2	Paramentos
	SC6	Sistemas drenantes	SC3	Pavimentos
C4	Cimentaciones	C11	Carpintería de madera	
	SC1	Armadura	SC1	Puertas
	SC2	Cimentaciones especiales	SC2	Armarios
	SC3	Encofrados	SC3	Ventanas
	SC4	Hormigones	SC4	Escaleras-Barandillas-Pasamanos
	SC5	Soleras y Losas	SC5	Persiana-Capialzado-Celosía
C5	Estructuras	C12	Carpintería exterior y cerrajería	
	SC1	Estructura de acero	SC1	Carpintería exterior
	SC2	Estructura de hormigón	SC2	Cierre-Protección
	SC3	Estructura de hormigón prefabricado	SC3	Varios
	SC4	Estructura de madera	C13	Vidriería y elaborados sintéticos
	SC5	Varios	SC1	Vidrios
C6	Albañilería	SC2	Vidrios especiales y varios	
	SC1	Cantería	C14	Pinturas
	SC2	Cerramientos	SC1	Pinturas
	SC3	Divisiones Interiores	SC2	Tratamientos
	SC4	Recibidos	SC3	Pinturas especiales
	SC5	Prefabricados, ventilación y varios	C15	Equipamiento, decoración y varios
	SC6	Ayudas	SC1	Equipamiento
C7	Cubiertas	SC2	Piscina	
	SC1	Formación de pendiente	SC3	Jardín-Riego
	SC2	Cubiertas inclinadas		

Tabla 6: División de capítulos y subcapítulos propuesta

En la Tabla 6 se muestra la estructura completa de agrupación de actividades (G_{ω}^*) propuesta, donde se puede observar la división de subcapítulos para cada capítulo.

3.4.3 Nivel L_4 : *Descripciones de Partidas*

El último nivel definido en la estructura propuesta debiera ser el de las diferentes descripciones de partida L_4 , que se corresponde con el conjunto de descripciones de partidas S_{ω} definido anteriormente. Con este nivel se completa la estructura de referencia.

Sin embargo, para este nivel, no se aporta una instanciación de referencia puesto que, como veremos más adelante en esta memoria, se nutre de los datos de los distintos proyectos.

3.5 CONCLUSIONES

En este capítulo se ha analizado en detalle el proceso de elaboración del documento del presupuesto y se han identificado las principales causas que originan la diversidad y las irregularidades que se observan en relación con este documento, tanto en la forma de organizar las actividades como en el marco lingüístico utilizado para definir las.

Asimismo, mediante la ilustración de varios ejemplos reales, se han identificado los inconvenientes a los que se enfrentan los responsables en tomar decisiones debido a la ausencia de un modelo de datos común. Como hemos visto, un inconveniente principal es la dificultad de consultar esos datos de manera integrada, algo que, sin la ayuda de un almacén integrado y herramientas de análisis, requiere por parte de los usuarios una gran cantidad de tiempo para localizar la información deseada.

Como primer objetivo de los planteados en esta memoria, habilitar una estructura de referencia para el almacenamiento integrado de los datos de los proyectos, se ha formalizado primero un modelo de datos general jerárquico para luego proponer una instancia de dicho modelo que sirva para el caso de los proyectos de edificación en España. El modelo de datos formal propuesto se basa en la estructura de descomposición de trabajo, comúnmente aceptada en el ámbito nacional e internacional para la organización de las actividades del presupuesto. La instancia propuesta, ha sido validada por un panel de

expertos de la Universidad de Granada y se ajusta a las directrices que suelen seguir los responsables en elaborar dicho documento.

Una vez que se ha establecido la estructura de referencia, el segundo objetivo al que nos enfrentamos, en el siguiente capítulo, es el desarrollo de un método de clasificación que extraiga cada una de las descripciones de partida del *conjunto de descripciones* de un determinado presupuesto y, de manera automática, las clasifique en el capítulo y subcapítulo correspondiente de la estructura de referencia G_{ω} .

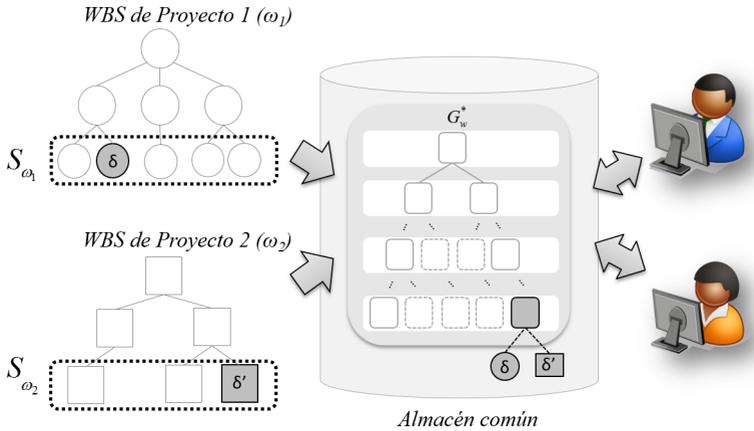


Figura 7: Proceso de clasificación

La Figura 7 ilustra el proceso de clasificación que proponemos, donde se pueden observar dos descripciones de partida de naturaleza similar pertenecientes al S_{ω_1} y S_{ω_2} , y que están situadas en diferentes grupos de actividades en sus estructuras originales ω_1 y ω_2 . Sin embargo, estas dos descripciones pueden ser colocadas en el mismo grupo de actividades de la estructura de referencia G_{ω}^* mediante el método de clasificación que proponemos en el siguiente capítulo.

MECANISMO DE CLASIFICACIÓN AUTOMÁTICO

RESUMEN: En este capítulo se presenta un mecanismo de clasificación de las descripciones de partida contenidas en el documento del presupuesto que permite afrontar con garantías nuestro objetivo de incorporar éstos datos en la estructura de referencia propuesta en el capítulo anterior, independientemente de la herramienta con la que ha sido elaborado y del marco lingüístico que utiliza, para su posterior procesamiento y/o toma de decisiones.

En primer lugar, detallamos la construcción de un modelo de conocimiento, acorde al modelo formal propuesto en el capítulo anterior, que más adelante es usado para clasificar descripciones de partida. Posteriormente describimos el proceso de clasificación que proponemos y se presenta un ejemplo para ilustrar dicho proceso. Finalmente, presentamos una experimentación y una evaluación del modelo de clasificación propuesto.

ÍNDICE

4.1	Introducción	80
4.2	Extracción de conocimiento para la clasificación	81
4.2.1	Preprocesamiento del texto	82
4.2.2	Cálculo de información relativa a los términos	83
4.2.2.1	Formalización de las medidas	85
4.3	Proceso de clasificación	88
4.3.1	Método de agregación multi-criterio	89
4.3.2	Ejemplo ilustrativo	91
4.4	Experimentación	96
4.4.1	Datos de aprendizaje y prueba	96
4.4.2	Ajuste de los pesos del operador OWA	98
4.4.3	Resultados detallados con la configuración OWA seleccionada	101
4.5	Evaluación	105
4.6	Conclusiones	109

4.1 INTRODUCCIÓN

Como venimos señalando en esta memoria, en la gestión de los proyectos de edificación, el acceso a la información oportuna es fundamental para garantizar que el proceso de toma de decisiones se realice con éxito. Esta información permite a los responsables de la organización tener un conocimiento más amplio del proceso, con el fin de adoptar mejores estrategias en la gestión de proyectos. Además, debido a las exigentes restricciones de plazos durante todo el ciclo de vida del proyecto, el tiempo disponible para tomar decisiones es muy limitado, pese a que, en la mayoría de las ocasiones, el resultado de estas decisiones puede influir directamente en el éxito del proyecto.

En el capítulo anterior se ha presentado un modelo de datos para la organización de los datos contenidos en el documento del presupuesto. Adicionalmente, basándonos en este modelo, se ha especificado una estructura de referencia para proyectos de edificación en España, con el fin de facilitar la construcción de un repositorio común de datos estructurados. Para nutrir dicho repositorio, es interesante desarrollar un modelo de extracción, que permita almacenar automáticamente dicha información en base a su semántica, desprendiéndonos de las particularidades del lenguaje o de la herramienta que se ha utilizado para elaborarlos.

Para este propósito, en este capítulo se plantea la construcción de un modelo de conocimiento junto con una serie de procedimientos que permiten clasificar en la estructura de referencia las descripciones de partida que aparecen en los presupuestos. En el modelo propuesto se consideran los términos utilizados para definir las descripciones de partida. Gracias a esta propuesta, se puede nutrir automáticamente un repositorio común con los datos contenidos en el documento del presupuesto de manera integrada y unificada de acuerdo con la estructura de referencia propuesta en el capítulo anterior.

En primer lugar, en la sección 4.2 se detalla la construcción de un modelo de conocimiento acorde al modelo formal propuesto en el capítulo anterior que más adelante es usado para clasificar descripciones de partidas. Posteriormente, en la sección 4.3, se describe el proceso de clasificación propuesto y se presenta un ejemplo para ilustrar dicho proceso. En la sección 4.4 se aportan los datos, tanto de la fase de aprendizaje como de prueba y la experimentación realizada. Finalmente, en la sección 4.5 se presenta una evaluación del modelo de clasificación propuesto.

4.2 EXTRACCIÓN DE CONOCIMIENTO PARA LA CLASIFICACIÓN

El modelo de extracción de conocimiento que se propone en este capítulo tiene como objetivo establecer de qué manera puede considerarse la aparición de términos relevantes en las descripciones de partidas para llevar a cabo su correcta clasificación en la estructura de referencia propuesta en el capítulo anterior.

Para establecer la lista de términos del modelo de conocimiento mencionado, proponemos la siguiente metodología:

1. En primer lugar, se selecciona una lista lo más exhaustiva posible de términos a partir de un conjunto lo más amplio posible de las diferentes descripciones de partida que pueden aparecer en el documento del presupuesto. Para este propósito, las bases de precios son una buena fuente para obtener este vocabulario, ya que proporcionan una lista bastante extensa de las distintas descripciones de partidas que pueden aparecer en los presupuestos. Para ello, se crea un conjunto de *descripciones de partidas* S_{ω}^* como el conjunto de todas las descripciones pertenecientes a las bases de precios consideradas en este primer paso. Como veremos en la sección 4.4.1, en nuestro caso, hemos seleccionado dos de las bases de precios más utilizadas en España como apoyo a la elaboración del documento del presupuesto.
2. A continuación, se coloca cada descripción del conjunto S_{ω}^* en su correspondiente nodo de la estructura de referencia G_{ω}^* , de modo que se obtiene un conjunto de datos de entrenamiento para el posterior aprendizaje supervisado. Esta correspondencia no es una tarea fácil porque la estructura de referencia G_{ω}^* no se corresponde con la estructura de agrupación de partidas original de las bases de precios y, por tanto, se debe realizar con la ayuda de un experto o de un panel de expertos.
3. Una vez que el conjunto S_{ω}^* se ha clasificado correctamente en la estructura de agrupación de actividades G_{ω}^* , se procede a aprender cómo clasificar correctamente las nuevas descripciones de partida que se extraen de los presupuestos de proyectos reales y que queremos almacenar en nuestro repositorio común. Como veremos más adelante, este análisis (y el método de clasificación) se basan en la idea de que las descripciones de partida descritas con términos similares tienen que ser colocadas en el mismo nodo o grupo de tareas.

En la sección 4.2.1 se detalla el preprocesamiento del texto contenido en las descripciones de partidas con el fin de obtener y unificar los términos. En la sección 4.2.2 se determina información discriminante, asociada a los distintos términos, para el posterior proceso de clasificación.

4.2.1 *Preprocesamiento del texto*

El análisis mencionado en la sección anterior comienza con la aplicación de un preprocesamiento automático de datos que extrae y unifica los términos de las descripciones de partida pertenecientes al conjunto de aprendizaje S_{ω}^* . Este proceso se compone de las siguientes etapas secuenciales:

- Preprocesamiento Sintáctico. Esta primera etapa consiste en una limpieza de carácter sintáctico, realizado con la herramienta Freeling [188] y que se divide en los siguientes pasos:

- Tokenización y lematización: es el proceso que consiste en dividir, en forma de lista, las cadenas de texto y hallar el lema correspondiente.

Por ejemplo, la descripción de partida “Desbroce y limpieza del terreno con medios manuales.”, se convierte en: “desbrozar”, “y”, “limpiar”, “de”, “el”, “terreno”, “con”, “medio”, “manual”, “.”.

- Eliminación de palabras vacías (Stopwords) y signos de puntuación y números. Se consideran palabras vacías a aquellas que no aportan información relevante para la tarea de clasificación, tales como preposiciones, conjunciones, pronombres, etc.

Continuando con el ejemplo anterior, el resultado de este proceso es el siguiente: “desbrozar”, “limpiar”, “terreno”, “medio”, “manual”.

- Preprocesamiento Semántico. Esta segunda etapa consiste en la detección y agrupación de sinónimos. Cada grupo de sinónimos tendrá un representante canónico, que será el término que presente al grupo de sinónimos. El representante canónico del grupo sustituirá todas las apariciones de los términos del grupo de sinónimos del texto original. Para realizar la detección de sinónimos, se realizan las siguientes tareas:

- Generación de conjuntos de sinónimos: los términos se agrupan en base a su significado obteniendo conjuntos de sinónimos. Durante esta etapa se utilizan diccionarios electrónicos junto con conocimiento experto para enriquecer los conjuntos de sinónimos con términos adicionales que no aparecen en el texto original. Aunque existen generadores de sinónimos electrónicos, como por ejemplo *WordNet* [197], en nuestro caso no se han podido utilizar por dos razones. En primer lugar, porque no hay disponible una estructura de *WordNet* en castellano y, en segundo lugar, debido a la especificidad del dominio. Por tanto, se ha tenido que crear la base de conocimiento a partir de conocimiento experto para el dominio de los proyectos de edificación, que se podrá ir enriqueciendo con la inserción de nuevos términos a los grupos de sinónimos.

Por ejemplo, a partir de este proceso, la palabra “jácena” se convierte en “viga”, “hgon” en “hormigón” y “partición” en “división”.

- Selección del representante canónico: para cada sinónimo se determina su representante canónico, es decir el término que va a representar a todo el conjunto de sinónimos.

En el ejemplo anterior, las palabras “hormigón”, “viga” y “división” son los representantes canónicos de sus correspondientes conjuntos de sinónimos.

A partir de este preprocesamiento del texto, se obtiene el conjunto de términos T unificado que se utiliza para definir las descripciones de partidas.

4.2.2 Cálculo de información relativa a los términos

Una vez obtenido el conjunto de términos T unificado, es importante determinar cuáles de ellos son discriminantes para el proceso de clasificación y en qué medida. En nuestra propuesta, vamos a trabajar con un modelo de agregación multi-criterio basado en tres dimensiones:

- Una dimensión está relacionada con la importancia de cada término en base a la frecuencia de aparición en un cierto grupo de tareas (nodo GT), en relación a la frecuencia con la que este mismo término aparece en los nodos *hermanos* de la estructura. Como veremos más adelante, los valores de esta dimensión se calculan en función del vocabulario del conjunto de descrip-

ciones de partida de S_{ω}^* que el experto ha clasificado en cada grupo de tareas GT.

- Otra dimensión se centra en la importancia de cada término en relación a la posición que ocupa éste en la descripción de partida de entrada. Como hemos comentado, el objetivo de las descripciones de partidas es describir de manera textual los diferentes trabajos que se tienen que realizar en el proyecto. Normalmente, los términos que aparecen en primeras posiciones describen la naturaleza del trabajo, y por tanto, suelen ser decisivos para indicar el grupo de tareas GT donde la partida debe ser colocada.

Ejemplo 5 *Vamos a considerar dos descripciones de partida del ejemplo 2, donde el mismo término aparece en posiciones diferentes:*

$\delta_{90}^s = \{ \text{tabique, ladrillo, hueco, recibir, pegamento, especial, soporte, banda, elástico, incluso, limpiar, medio, auxiliar} \}$

$\delta_{225}^s = \{ \text{instalar, lavabo, suspendido, tanque, empotrado, sistema, doble, descarga, incluso, fijación, conexión, tabique, ayuda, albañilería} \}$

Como se puede observar, el término *tabique* aparece en las dos descripciones. En δ_{90}^s , la presencia del término en primera posición facilita claramente la identificación del grupo de tareas donde se debe colocar. En este caso, en el grupo de tareas relacionado con los trabajos de albañilería (GT_4^2). Sin embargo, en la segunda descripción, el mismo término aparece en las última posiciones y no ofrece información relevante para colocar la partida en el grupo de tareas correcto. En este caso, esta partida se debe colocar en el grupo de tareas en relación a las instalaciones (GT_5^2).

Por consiguiente, en nuestra propuesta se asigna más peso a los términos localizados en las primeras posiciones de la WD de entrada, mientras que el peso se reduce conforme nos adentramos en los términos en últimas posiciones. Utilizamos valores normalizados en $[0, 1]$.

- La tercera dimensión refleja la opinión del experto en relación a la importancia de la aparición de un cierto término para indicar que la partida debe ser colocada en un cierto grupo de tareas GT. Esta dimensión complementa a las dos anteriores y permite equilibrar decisiones erróneas que pueden ser sugeridas por las otras dos dimensiones, que se basan exclusivamente en el análisis automático del texto. Por ejemplo, puede ocurrir que términos, que parecen no tener relevancia en cuanto a las

dimensiones de frecuencia y posición, sean muy relevantes para determinar la clasificación de una partida en un cierto grupo de tareas GT.

4.2.2.1 Formalización de las medidas

A continuación se presentan las definiciones formales y las ecuaciones necesarias para el cálculo de cada una de las medidas asociadas a las dimensiones anteriores.

Definición 6 (Importancia basada en la Frecuencia) Sea $\omega = \{L_1, \dots, L_n\}$ una estructura de descomposición del trabajo. La Importancia basada en la frecuencia (I_f) de un término $t \in T$ en un grupo de tareas GT_j^i es el cociente de la frecuencia del término t en la unión de partidas descendientes de GT_j^i y su frecuencia en partidas descendientes de su padre GT_p^{i-1} . Se calcula mediante la siguiente ecuación:

$$I_f^{GT_j^i}(t) = \frac{F^{GT_j^i}(t)}{F^{GT_p^{i-1}}(t)} \quad (4.1)$$

donde la frecuencia $F^{GT}(t)$ es el número de apariciones de un término t en las partidas descendientes de GT , y $\langle GT_p^{i-1}, GT_j^i \rangle \in KS^{i-1}$, $i = 2, \dots, n - 1$.

Ejemplo 6 Vamos a considerar el término *tabique* como ejemplo para detallar el proceso de cálculo de la importancia basada en frecuencia. Mediante el preprocesamiento del texto que se ha detallado en la sección anterior, se han unificado los términos de las descripciones de partidas que previamente el experto había clasificado en cada nodo de la estructura de referencia (según el segundo paso de la metodología presentada en la sección 4.2). Para cada término en cada uno de los nodos de la estructura de referencia se calcula el número de apariciones.

En la Tabla 7 se muestra el número de apariciones del término *tabique* para cada subcapítulo del capítulo de *albañilería*. El número total de apariciones para este capítulo es de 109, por tanto, la frecuencia de este término para el capítulo de *albañilería* es el cociente del número de apariciones de cada subcapítulo y el total de apariciones del capítulo. En la tabla, se muestran los resultados obtenidos para la *importancia basada en frecuencia* en cada uno de los subcapítulos de este capítulo.

A continuación, se calcula la *importancia basada en frecuencia* en el nivel superior de la estructura de referencia del mismo modo. Para

C6						
	SC1	SC2	SC3	SC4	SC5	SC6
Nºapariciones	0	20	87	2	0	0
I _f (tabique)	0,00	0,18	0,80	0,02	0,00	0,00

Tabla 7: Número de apariciones del término *tabique* en los subcapítulos del capítulo de *albañilería*.

ello, es necesario conocer el número de apariciones de este término en el resto de capítulos, en este caso el número total de apariciones es de 176. En la Tabla 8 se muestra el número de apariciones para cada capítulo y el valor obtenido para la *importancia basada en frecuencia* en cada uno de ellos.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
Nºapariciones	13	0	0	0	0	109	1	3	28	2	0	1	19	0	0
I _f (tabique)	0,07	0,00	0,00	0,00	0,00	0,62	0,01	0,02	0,16	0,01	0,00	0,01	0,11	0,00	0,00

Tabla 8: Número de apariciones del término *tabique* en cada uno de los capítulos de la estructura de referencia.

Definición 7 (Importancia basada en la Posición) Sea $\omega = \{L_1, \dots, L_n\}$ una estructura de descomposición del trabajo. La importancia basada en la posición I_p de un término $t \in T$ en una descripción de partida δ se calcula mediante la siguiente ecuación:

$$I_p^\delta(t) = 1 - \frac{P_{\delta^s}(t)}{|\delta^s|} \quad (4.2)$$

donde $P_{\delta^s}(t) \in [0, |\delta^s| - 1]$ es la posición del término t en δ^s y $|\delta^s|$ es el cardinal de δ^s .

Una vez que se han definido las importancias de un término en relación a la frecuencia y posición, basadas en el análisis automático del texto, nuestra propuesta considera una tercera dimensión, la *importancia del experto*. El experto que diseña el sistema utiliza esta tercera dimensión para incorporar directamente su opinión en relación a como la aparición de un término en una descripción de partida determina la clasificación de la partida en un cierto grupo de tareas GT.

Definición 8 (Importancia del Experto) Sea $\omega = \{L_1, \dots, L_n\}$ una estructura de descomposición del trabajo. La importancia del experto I_e de un término $t \in T$ en un grupo de tareas GT_j^i se calcula como sigue:

- En el caso que $i = n - 1$: el experto asigna directamente para cada término t un valor del intervalo $[0, 1]$ indicando su opinión sobre la importancia de aparición del término t para tomar la decisión de colocar la partida en GT_j^{n-1} .
- Una vez que se han asignado los valores correspondientes a la importancia del experto para cada uno de los términos en el nivel más bajo de la estructura de referencia, se procede al cálculo de éstos en el resto de niveles. En este punto, nuestra propuesta consiste en la agregación de estos valores teniendo en cuenta la relación de parentesco entre los nodos de la estructura. Se ha optado por esta opción por dos motivos: en primer lugar, para simplificar el trabajo del experto, ya que no es tarea fácil la asignación de los valores para cada uno de los términos en cada nodo y, en segundo lugar, para evitar inconsistencias.

$$I_e^{GT_j^i}(t) = \max_{GT_l^{i+1} | < GT_j^i, GT_l^{i+1} > \in KSi} (I_e^{GT_l^{i+1}}(t)) \quad (4.3)$$

A partir del cálculo de cada una de las dimensiones propuestas en las definiciones anteriores (6, 7 y 8), se han obtenido tres valores (I_f , I_p y I_e) para cada término de cada nodo de la estructura de referencia. Con el fin de tener un valor representativo de cada término para cada nodo de la estructura, agregamos estos tres valores en un único valor de importancia general que denotaremos como I :

Definición 9 (Importancia general) Sea $\omega = \{L_1, \dots, L_n\}$ una estructura de descomposición del trabajo y δ una descripción de partida. La importancia I de un término $t \in T$ para clasificar la partida en un grupo de tareas GT_j^i es la agregación de los valores obtenidos de las tres dimensiones, mediante la siguiente ecuación:

$$I_\delta^{GT_j^i}(t) = \text{Agg}(I_e^{GT_j^i}(t), I_p^\delta(t), I_f^{GT_j^i}(t)) \quad (4.4)$$

donde Agg es un operador de agregación. Utilizamos el término de agregación en el sentido de resumir los valores en un único valor representativo. En la sección 4.3.1 se detalla el operador de agregación que proponemos para nuestra propuesta.

Las importancias de los términos para cada grupo de tareas de la estructura de conocimiento propuesta se calculan de abajo a arriba, es decir desde el nivel L_n al L_1 :

- Una vez que se ha realizado el preprocesamiento y se ha obtenido la lista de términos, se calcula la información en relación a la frecuencia, posición y experto para los grupos de tareas del nivel más bajo.
- Los valores de las importancias en los niveles superiores se calculan, mediante agregación, cuando son necesarios.

En esta sección se ha definido el modelo de conocimiento basado en tres importancias que nos indican cómo la aparición de un término en una descripción de partida de entrada influye en la decisión para determinar su asignación en un cierto grupo de tareas.

4.3 PROCESO DE CLASIFICACIÓN

En esta sección se detalla cómo se pueden utilizar los valores de las importancias definidas en la sección anterior con el objetivo de seleccionar el mejor grupo de tareas para una partida dada. Como veremos, el proceso de clasificación se lleva a cabo desde el primer nivel de la jerarquía hasta el último, con el fin de simplificar los cálculos. De este modo, la partida se clasificará en el nivel más alto y posteriormente en los niveles inferiores, tal y como se ilustra en la Figura 8.

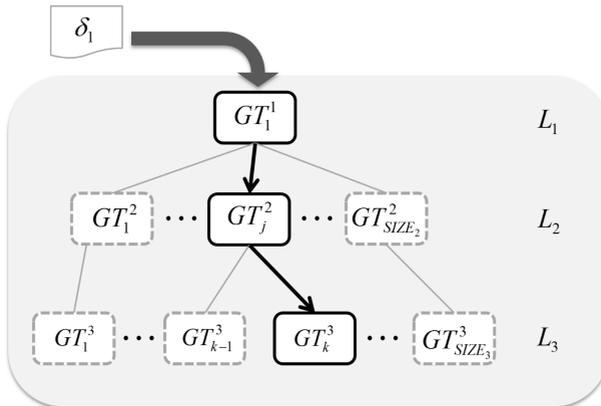


Figura 8: Proceso de clasificación

4.3.1 Método de agregación multi-criterio

Las partidas que se van a insertar en la estructura son consideradas como listas de términos ordenados $\delta = \{t_1, \dots, t_n\}$ donde t_k corresponde con el k -ésimo término de la descripción. Para obtener este conjunto, las descripciones de partida de entrada siguen un pre-procesamiento de datos similar al que hemos definido anteriormente en la sección 4.2.1 para el conjunto de partidas de aprendizaje. De esta forma, se garantiza que cada partida sólo contiene términos que pertenecen al vocabulario de referencia T .

Como se ha explicado en la sección anterior, para cada uno de estos términos se tendrá un valor de Importancia agregado I que resume los valores de las distintas dimensiones de importancia (*frecuencia*, *posición* y *experto*) para cada GT_j^i . Por lo tanto, necesitamos agregar los valores de importancia I de cada uno de los términos de la partida considerada para poder clasificarla en un grupo de tareas determinado. Este proceso de agregación se calcula mediante la siguiente función:

Definición 10 (Idoneidad de clasificación) *La idoneidad de clasificación de una descripción de partida δ en el grupo de tareas del nivel GT_j^i es el valor que se obtiene mediante la agregación de todas las importancias generales I de sus términos:*

$$\text{Idoneidad}^{GT_j^i}(\delta) = \text{Agg}_{(t \in \delta^s)}(I_{\delta}^{GT_j^i}(t)) \quad (4.5)$$

donde Agg es un operador de agregación.

Este valor es determinante en el proceso de clasificación porque la partida se clasificará en el grupo de tareas donde se obtenga el valor de idoneidad más alto.

Como se ha comentado anteriormente, el proceso de clasificación se lleva a cabo desde el nivel más alto de la jerarquía hasta el más bajo. De este modo, la descripción de partida se clasificará en un nodo del nivel más alto y posteriormente en los hijos de este nodo. A continuación se detalla la función de clasificación:

Definición 11 (Función de clasificación) *La función de clasificación de una descripción de partida δ entre los hijos de un grupo de tareas GT_j^i , se define como sigue:*

$$F^{GT_j^i}(\delta) = GT_{l}^{i+1}, \text{ tal que,}$$

- $\langle GT_j^i, GT_l^{i+1} \rangle \in KS^i$
- $\nexists \langle GT_j^i, GT_k^{i+1} \rangle \in KS^i \mid Idoneidad^{GT_k^{i+1}}(\delta) > Idoneidad^{GT_l^{i+1}}(\delta)$
- $\forall \langle GT_j^i, GT_k^{i+1} \rangle \in KS^i \mid Idoneidad^{GT_k^{i+1}}(\delta) = Idoneidad^{GT_l^{i+1}}(\delta) \rightarrow k \geq l$

Como se puede observar, la segunda regla se utiliza para resolver empates.

En el proceso de clasificación, se necesitan agregar varios valores en los cálculos para determinar la clasificación de una partida en un cierto grupo de tareas GT ($Idoneidad^{GT}(\delta)$). En primer lugar, se tienen que agregar las importancias basadas en *frecuencia*, *posición* y *experto* para cada uno de los términos (ecuación 4.4) y, en segundo lugar, se tienen que agregar las *importancias generales* obtenidas para cada término con el fin de obtener el valor de *Idoneidad* de la descripción de partida (ecuación 4.5).

Nótese que utilizamos el término *agregar* en el sentido de resumir los valores en un único valor representativo. Para realizar esta tarea existen muchos operadores en la literatura, por ejemplo, operadores como el máximo, mínimo, media, etc. Entre estos, un operador muy conocido y ampliamente utilizado es el operador de media ordenada ponderada (del inglés, Ordered Weighted Averaging, OWA). Este operador, propuesto por Yager [244, 245], proporciona una agregación parametrizada muy versátil que incluye el máximo, mínimo y la media. La capacidad que tiene para modelar diferentes agregaciones ha hecho que se haya utilizado en múltiples áreas de aplicación. Se define como sigue:

Definición 12 (Operador OWA) : Un operador OWA de dimensión n es una función $OWA : \mathbb{R}^n \rightarrow \mathbb{R}$ que tiene asociado un vector de pesos W de dimensión n con $w_j \in [0, 1]$ y $\sum_{j=1}^n w_j = 1$ tal que:

$$OWA(a_1, \dots, a_n) = \sum_{j=1}^n w_j b_j \quad (4.6)$$

donde b_j es el j -ésimo elemento más grande de la colección $a_1 \dots a_n$.

Un aspecto fundamental de los operadores OWA es el paso de la reordenación. Los pesos están asociados con una posición ordenada en lugar de con los elementos en sí.

Ejemplo 7 Por ejemplo, si $W = [0.6, 0.3, 0.1]$ y los valores de importancias para un término dado t son los siguientes: $I_f = 0.65$ $I_p = 0.75$ $I_e = 0.8$, entonces $I^{GT_i}(t) = 0.6 \cdot 0.8 + 0.3 \cdot 0.75 + 0.1 \cdot 0.65 = 0.77$.

La principal ventaja del operador OWA es que, dependiendo del vector de pesos que se utilice, ofrece diferente comportamiento entre los operadores de disyunción y conjunción lógicos. Esta característica permite gran flexibilidad para elegir el tipo de reglas de agregación que se utilizará en cada problema. Por esta razón, se ha convertido en un operador muy aceptado y utilizado por la comunidad investigadora [93].

Como este operador es parametrizable, para su uso, se tiene que establecer el vector de pesos. Como hemos comentado, este hecho proporciona un grado de libertad para ajustar mejor el proceso de agregación. En nuestro modelo de clasificación, como hemos comentado, se necesitan dos operadores OWA:

- El primer operador se utiliza para la agregación de las tres importancias de cada uno de los términos (ecuación 4.4). Por lo tanto, se necesita un vector de pesos con tres valores.
- El segundo operador se utiliza para la agregación de las *importancias generales* obtenidas para cada término con el fin de obtener el valor de *idoneidad* de la descripción de partida (ecuación 4.5). Por lo tanto, se necesita un vector de pesos con tantos valores como términos.

En la sección 4.4.2 se detalla la experimentación que se llevado a cabo con el objetivo de localizar una buena configuración de los operadores OWA para nuestro modelo de clasificación con la estructura de referencia propuesta en el anterior capítulo.

4.3.2 Ejemplo ilustrativo

En esta sección se presenta un ejemplo del proceso de clasificación desde el inicio hasta la clasificación de la partida en un determinado grupo de tareas.

Ejemplo 8 Consideremos de nuevo la siguiente descripción de partida de entrada:

Excavación, en apertura de caja, en terreno blando, realizada con medios mecánicos, incluso perfil hasta una profundidad máxima de 50 cm. Medida en perfil natural

Se va a considerar también la siguiente estructura de agrupación de actividades G_{ω}^* compuesta por tres niveles, L_1, L_2 y L_3 :

- $L_1 = GT_1^1 = \text{Todo}$
- $L_2 = GT_1^2, GT_2^2, GT_3^2, GT_4^2$
- $L_3 = \dots, GT_4^3, GT_5^3, GT_6^3, \dots$

El proceso de clasificación es el siguiente:

1. *Extracción de los términos de la partida:* Primero, se realiza el preprocesamiento sintáctico y semántico detallado en la sección 4.2.1. Después del preprocesamiento, se obtiene:

$\delta_1 = \{\text{excavación, apertura, caja, terreno, blando, hacer, medio, mecánico, incluso, perfil, profundidad, máximo, medir, perfil, natural}\}$

Cuya forma corta es:

$\delta_1^s = \{\text{excavación, apertura, caja, terreno, blando, hacer, medio, mecánico, incluso, perfil, profundidad, máximo, medir, natural}\}$

2. *Obtención de los valores de importancia:* una vez que el preprocesamiento se ha llevado a cabo, se calculan las importancias para cada término y para cada GT_j^2 del nivel L_2 . Los valores de las importancias *frecuencia* y *experto* no dependen de la descripción de partida de entrada; de hecho, éstas se establecen durante la fase de aprendizaje, detallado en la sección 4.2. Estos dos valores se completan con la dimensión basada en la *posición*, que sí depende de la partida de entrada y se calcula mediante la ecuación 4.2.

En la Tabla 9 se muestran los valores de las *importancias* para cada uno de los términos de nuestra partida de ejemplo, para el grupo de tareas del nivel L_2 .

Por ejemplo, para el término *excavación*, se puede observar como los valores correspondientes a la dimensión basada en la *posición* son iguales para cada grupo de tareas del nivel L_2 , ya que depende de la posición del término en la descripción que se pretende clasificar. Sin embargo, los valores para las importancias de *frecuencia* y *experto* son diferentes para cada grupo de tareas. En este caso, el experto ha establecido que el término

		GT ₁ ²			GT ₂ ²			GT ₃ ²			GT ₄ ²		
		I _e	I _f	I _p	I _e	I _f	I _p	I _e	I _f	I _p	I _e	I _f	I _p
t1	Excavación	0,50	0,02	1,00	0,98	0,58	1,00	0,60	0,23	1,00	0,30	0,00	1,00
t2	Apertura	0,50	0,01	0,93	0,70	0,26	0,93	0,20	0,01	0,93	0,50	0,05	0,93
t3	Caja	0,50	0,00	0,86	0,00	0,00	0,86	0,00	0,00	0,86	0,40	0,01	0,86
t4	Blando	0,00	0,00	0,79	0,00	0,36	0,79	0,00	0,00	0,79	0,00	0,00	0,79
t5	Terreno	0,00	0,02	0,71	0,90	0,72	0,71	0,30	0,05	0,71	0,50	0,12	0,71
t6	Hacer	0,00	0,00	0,64	0,00	0,00	0,64	0,00	0,00	0,64	0,30	0,03	0,64
t7	Medio	0,70	0,02	0,57	0,50	0,17	0,57	0,20	0,02	0,57	0,70	0,02	0,57
t8	Mecánico	0,50	0,00	0,50	0,70	0,01	0,50	0,00	0,00	0,50	0,70	0,07	0,50
t9	Incluso	0,00	0,00	0,43	0,10	0,97	0,43	0,00	0,00	0,43	0,40	0,01	0,43
t10	Perfil	0,70	0,07	0,36	0,70	0,10	0,36	0,80	0,09	0,36	0,50	0,00	0,36
t11	Profundidad	0,00	0,00	0,29	0,00	0,00	0,29	0,20	0,00	0,29	0,50	0,01	0,29
t12	Máximo	0,70	0,02	0,21	0,40	0,01	0,21	0,10	0,01	0,21	0,70	0,01	0,21
t13	Medir	0,00	0,00	0,14	1,00	0,20	0,14	0,20	0,01	0,14	0,70	0,10	0,14
t14	Natural	0,70	0,00	0,07	0,00	0,01	0,07	0,00	0,00	0,07	0,70	0,01	0,07

Tabla 9: Cálculos de $F^{\text{Todo}}(\delta_1)$ (Importancias)

excavación es más representativo del grupo de tareas GT_2^2 que del resto de grupos de tareas, ya que es donde se encuentra el valor más alto (0.98). En cuanto a la importancia basada en frecuencia, se puede observar como este término, en base al modelo de conocimiento realizado en la fase de aprendizaje, es más frecuente en el GT_2^2 que en el resto de grupo de tareas, de hecho, en el GT_4^2 no aparece ya que el valor es 0.

3. *Obtención de los valores agregados correspondientes a las importancias generales*: los valores de *importancia* obtenidos para cada uno de los términos se agregan en un único valor. En la Tabla 10 se muestran los valores obtenidos mediante la ecuación 4.4 utilizando un operador OWA con el siguiente vector de pesos $W = [0.33, 0.33, 0.33]$. Continuando con el ejemplo del paso anterior, para el término *excavación* los valores de las tres dimensiones son: $I_e = 0.50$ $I_f = 0.02$ $I_p = 1$, entonces $I^{GT_1^2}(t) = 1 \cdot 0.33 + 0.50 \cdot 0.33 + 0.02 \cdot 0.33 = 0.51$. Este vector de pesos W sólo se utiliza para ilustrar el ejemplo. En este caso, el operador OWA actúa como un operador de media. Como ya hemos indicado, en la sección 4.4.2 se detalla la experimentación que se ha llevado a cabo con el fin de encontrar un buen vector de pesos para los conjuntos de datos y la estructura de referencia propuesta en el capítulo anterior.
4. *Cálculo del valor de Idoneidad*: una vez que se ha obtenido la importancia general para cada término de la partida de entrada, se agregan estas importancias generales, mediante la ecuación 4.5, y así, poder obtener el valor de *idoneidad*. En la Tabla 10 se

		GT_1^2	GT_2^2	GT_3^2	GT_4^2
		I_δ	I_δ	I_δ	I_δ
t1	Excavación	0,51	0,85	0,61	0,43
t2	Apertura	0,48	0,63	0,38	0,49
t3	Caja	0,45	0,29	0,29	0,42
t4	Blando	0,26	0,38	0,26	0,26
t5	Terreno	0,24	0,78	0,35	0,44
t6	Hacer	0,21	0,21	0,21	0,32
t7	Medio	0,43	0,41	0,26	0,43
t8	Mecánico	0,33	0,40	0,17	0,42
t9	Incluso	0,14	0,50	0,14	0,28
t10	Perfil	0,38	0,39	0,42	0,29
t11	Profundidad	0,10	0,10	0,16	0,27
t12	Máximo	0,31	0,21	0,11	0,31
t13	Medir	0,05	0,45	0,12	0,31
t14	Natural	0,26	0,03	0,02	0,26
Idoneidad 1		0,51	0,85	0,61	0,49

Tabla 10: Cálculos de $F^{T^{odo}}(\delta_1)$ (Importancia general e idoneidad)

pueden observar los valores de idoneidad obtenidos utilizando un operador OWA con el vector de pesos $W = [1, 0, 0, \dots, 0]$ que, en este caso, actúa como un operador máximo. Para la descripción de ejemplo, y según los valores de importancia generales para cada término, el valor de idoneidad se calcula como sigue: $\text{Idoneidad}^{GT_1^2} = 0.51 \cdot 1 + 0.48 \cdot 0 + 0.45 \cdot 0 + 0.43 \cdot 0 + 0.38 \cdot 0 + 0.33 \cdot 0 + 0.31 \cdot 0 + 0.26 \cdot 0 + 0.26 \cdot 0 + 0.24 \cdot 0 + 0.21 \cdot 0 + 0.14 \cdot 0 + 0.10 \cdot 0 + 0.05 \cdot 0 = 0.51$.

Al igual que en el tercer apartado, este vector de pesos se ha utilizado para ilustrar el ejemplo, y en la sección 4.4.2 se detallará la experimentación realizada para encontrar un buen vector de pesos para el caso de la estructura de referencia propuesta en el capítulo 2.

5. *Clasificación de la partida*: finalmente, se clasifica la descripción de partida de entrada en el grupo de tareas donde se haya obtenido el valor de idoneidad más alto. Como se puede observar en la tabla, el valor de idoneidad más alto es 0.85 para el GT_2^2 . Por lo tanto, δ_1 se clasificará en este grupo de tareas.

Una vez que la descripción de partida se ha clasificado en un grupo de tareas del segundo nivel GT_1^2 , ésta se debe clasificar en un grupo de tareas perteneciente a los hijos de este GT. Vamos a asumir que $\{GT_4^3, GT_5^3, GT_6^3\}$ son hijos de este GT_2^2 . El proceso para clasificar la WD en el tercer nivel es el mismo que se ha detallado, pero sólo desde el segundo al quinto paso.

		GT ₄ ³			GT ₅ ³			GT ₆ ³		
		I _e	I _f	I _p	I _e	I _f	I _p	I _e	I _f	I _p
t1	Excavación	0,00	0,00	1,00	0,00	0,00	1,00	0,98	0,86	1,00
t2	Apertura	0,00	0,00	0,93	0,00	0,00	0,93	0,70	0,85	0,93
t3	Caja	0,00	0,00	0,86	0,00	0,00	0,86	0,00	0,00	0,86
t4	Blando	0,00	0,00	0,79	0,00	0,00	0,79	0,00	0,00	0,79
t5	Terreno	0,40	0,01	0,71	0,00	0,00	0,71	0,40	0,69	0,71
t6	Hacer	0,00	0,00	0,64	0,00	0,00	0,64	0,00	0,00	0,64
t7	Medio	0,20	0,09	0,57	0,20	0,03	0,57	0,50	0,68	0,57
t8	Mecánico	0,00	0,00	0,50	0,00	0,00	0,50	0,00	0,00	0,50
t9	Incluso	0,10	0,09	0,43	0,10	0,01	0,43	0,10	0,53	0,43
t10	Perfil	0,00	0,00	0,36	0,00	0,00	0,36	0,80	1,00	0,36
t11	Profundidad	0,00	0,00	0,29	0,00	0,00	0,29	0,00	0,00	0,29
t12	Máximo	0,00	0,00	0,21	0,10	0,20	0,21	0,00	0,00	0,21
t13	Medir	0,00	0,00	0,14	0,00	0,00	0,14	0,30	0,17	0,14
t14	Natural	0,00	0,00	0,07	0,00	0,00	0,07	0,00	0,00	0,07

Tabla 11: Cálculos de $FGT_2^3(\delta_1)$ (Importancias)

La Tabla 11 muestra los valores de *importancia* para cada término en cada GT de acuerdo al segundo paso del proceso. La Tabla 12 muestra los valores de las *importancias generales* obtenidas en el tercer paso y la *idoneidad* para cada GT_j^3 mediante el cuarto paso. Finalmente, la descripción de partida de entrada se clasifica en GT_6^3 porque es el GT donde se ha obtenido el valor más alto. Para los cálculos, se han utilizado los mismos valores para los vectores de pesos de los operadores OWA que en el nivel L_2 .

En el capítulo anterior se ha formalizado la estructura de agrupación de actividades para soportar el almacenamiento de la información contenida en el documento del presupuesto en España. En las secciones anteriores de este capítulo, se ha presentado un modelo de conocimiento y una metodología que permite la clasificación de las descripciones de partidas en una estructura de referencia. Veamos ahora, una experimentación completa sobre cómo aplicar el método de clasificación sobre la estructura de referencia propuesta en el capítulo anterior.

		GT_4^3	GT_5^3	GT_6^3
		I_δ	I_δ	I_δ
t1	Excavación	0,33	0,33	0,95
t2	Apertura	0,31	0,31	0,83
t3	Caja	0,29	0,29	0,29
t4	Blando	0,26	0,26	0,26
t5	Terreno	0,37	0,24	0,60
t6	Hacer	0,21	0,21	0,21
t7	Medio	0,29	0,27	0,58
t8	Mecánico	0,17	0,17	0,17
t9	Incluso	0,21	0,18	0,35
t10	Perfil	0,12	0,12	0,72
t11	Profundidad	0,10	0,10	0,10
t12	Máximo	0,07	0,17	0,07
t13	Medir	0,05	0,05	0,20
t14	Natural	0,02	0,02	0,02
Idoneidad 1		0,37	0,33	0,95

Tabla 12: Cálculos de $F^{GT^2}(\delta_1)$ (Importancia general e idoneidad)

4.4 EXPERIMENTACIÓN

En esta sección se detalla la experimentación realizada con el fin de evaluar el método de clasificación. Para ello, en primer lugar se proporciona información sobre los datos que se han utilizado en la experimentación, tanto de la fase de aprendizaje como de la fase de prueba. Posteriormente, la segunda sección se centra en la experimentación que se ha realizado con el fin de encontrar una buena configuración para los vectores de pesos del operador OWA para nuestros datos. Finalmente, la última sección presenta los resultados de clasificación obtenidos cuando se usa la configuración del operador OWA seleccionada en el método de clasificación.

4.4.1 Datos de aprendizaje y prueba

Para la experimentación del método de clasificación propuesto, se utiliza la estructura de referencia que se ha propuesto en el capítulo anterior. Por lo tanto, en base a esta estructura se realiza la fase de aprendizaje (sección 4.2) y será ésta la estructura de referencia donde se clasificarán las descripciones de partida de prueba. A continuación, se detallan los datos correspondientes a la fase de aprendizaje:

- En primer lugar, se extraen todas las descripciones de partida de dos de las bases de precios más conocidas para la elaboración del documento del presupuesto: CENTRO [67] y PREOC [195]. El objetivo de estas dos bases de precios es facilitar la elaboración de los presupuestos y mediciones de los proyectos. Ambas bases de precios cuentan con un larga trayectoria en el sector de la construcción. La primera de ellas la edita el Colegio Oficial de Aparejadores, Arquitectos Técnicos e Ingenieros de Edificación de Guadalajara y se publica anualmente desde el año 1984. La segunda base de precios la edita la sociedad anónima Atayo s.a y se publica desde el año 1992.

La unión de todas estas descripciones constituye el conjunto de descripciones de partida (S_{ω}^*) para la fase de aprendizaje. Este conjunto contiene 7.029 descripciones diferentes con un vocabulario T compuesto de 7.327 términos.

- En segundo lugar, el experto ha establecido la relación de parentesco en el nivel de subcapítulo, mediante la colocación de cada partida en su correspondiente subcapítulo, tal y como se detalló en la sección 4.2.
- Finalmente, la estructura de conocimiento se completa con el cálculo de las importancias de los términos, como se describió en la sección 4.2.2.

En la Tabla 18 se puede observar información más detallada correspondiente a esta fase de aprendizaje, así como el número de términos y el número de descripciones de partidas para cada capítulo (L_2) y subcapítulo (L_3) de la estructura de referencia.

Para la fase de prueba, se han considerado los siguientes conjuntos de datos:

- Un conjunto de descripciones de partida muy extenso extraídas de dos bases de precios diferentes a las utilizadas en la fase de aprendizaje: base de precios de la Junta de Andalucía (BCCA [129]) y base de precios de la Junta de Extremadura (EXT [68]). El número total de descripciones de partidas diferentes es de 12730.
- Un conjunto de descripciones de partidas extraídas de 50 presupuestos de proyectos reales. Estos documentos han sido elaborados por diferentes arquitectos y para diferentes tipologías de edificación. Por lo tanto, en estas descripciones de partida se encontrarán más irregularidades debido a la intervención del

arquitecto. El número total de descripciones de partidas para este conjunto es de 9669.

Como se puede observar, el proceso de prueba en nuestra experimentación no se ha realizado sobre el mismo conjunto de datos utilizado en la fase de aprendizaje, sino que se ha considerado un nuevo conjunto de datos muy extenso. En primer lugar, el conjunto de prueba está compuesto por dos bases de precios que nos garantizan que en la experimentación haya descripciones de partidas de todos los capítulos y subcapítulos. En segundo lugar, está compuesto por un conjunto de descripciones de partidas que presentan más irregularidades debido a las diferentes tipologías de los proyectos y a la diversidad de las descripciones de partidas que se pueden encontrar debido al proceso de elaboración del documento del presupuesto.

A continuación, nos centramos en la experimentación para encontrar una buena configuración para los vectores de pesos del operador OWA en el proceso de clasificación que hemos propuesto.

4.4.2 Ajuste de los pesos del operador OWA

Como hemos visto en este capítulo, el modelo de clasificación propuesto utiliza dos operadores OWA para agregar las importancias en la estructura de conocimiento. El primero, (OWA_1) que agrega las tres *importancias* (frecuencia, posición y experto) de cada término en un valor general y , el segundo (OWA_2), para agregar las *importancias generales* (resultado del OWA_1) de los términos en un valor que llamamos, valor de *idoneidad*, para una descripción de partida dada.

Para el primer operador (OWA_1), se necesita establecer un vector de pesos con tres valores, mientras que para el segundo (OWA_2) necesitaríamos definir un vector de pesos de igual número de elementos que términos que aparecen en la descripción de partida. Como el número de términos en las descripciones no es constante, se va a determinar un vector de pesos para un número prefijado de términos. En base a experimentos previos, hemos considerado un vector de diez pesos. Por tanto, se considerarán los diez términos que hayan obtenido los valores de importancia más altos.

Asimismo, para definir los pesos de los dos OWA (tres para el primero y diez para el segundo) hemos realizado una experimentación con nuestros datos con un doble objetivo, encontrar un equilibrio entre unos pesos que proporcionen unos buenos resultados de clasificación además de delimitar el número de combinaciones puesto que éste es muy elevado. En base a esta experimentación, se han establecido, para

ambos vectores, combinaciones de pesos con un decimal de precisión y orden descendente. En particular, para el caso del vector de pesos del (OWA_1) se han establecido aquellas combinaciones que incluyen las tres dimensiones (frecuencia, posición y experto), es decir, valores mayores que 0 para cada elemento del vector. La Tabla 13 muestra todas las posibles combinaciones para este vector de pesos según el criterio establecido. Sin embargo, para el vector de pesos del (OWA_2) se han considerado todas las posibles combinaciones de los diez pesos con un decimal de precisión y orden descendente, es decir, valores iguales a 0 también se consideran. El número de combinaciones para este vector es de 42.

Debido al proceso de clasificación propuesto, para obtener el *valor de idoneidad* de una descripción de partida es necesario agregar las *importancias*, frecuencia, posición y experto, de cada uno de sus términos mediante el operador OWA_1 y, a continuación, agregar el resultado con el OWA_2 . Por tanto, el número total de combinaciones posibles para los vectores de pesos es 336.

Vectores de pesos
0.8,0.1,0.1
0.7,0.2,0.1
0.6,0.3,0.1
0.6,0.2,0.2
0.5,0.4,0.1
0.5,0.3,0.2
0.4,0.4,0.2
0.4,0.3,0.3

Tabla 13: Combinaciones para el vector de pesos del OWA_1 que agrega los valores de las tres importancias

Como se ha explicado anteriormente, el proceso de clasificación empieza por el primer nivel de la jerarquía y continúa con los niveles inferiores, con el fin de simplificar los cálculos. Es decir, las descripciones de partida de entrada se clasificarán en los capítulos y posteriormente en un subcapítulo del capítulo seleccionado. En la experimentación, con el fin de localizar la mejor configuración para los vectores de pesos del operador OWA , se ha seguido el mismo proceso. En primer lugar, se buscan los dos vectores de pesos adecuados para el nivel de capítulo y, a continuación, se buscan los dos vectores adecuados para el nivel de subcapítulo.

Por lo tanto, para cada una de estas combinaciones y para cada uno de los conjuntos de test se obtienen los resultados de clasificación a los que conducen para el nivel de capítulo y subcapítulo. Para el nivel de capítulo, la medida de precisión cuantifica el porcentaje de partidas del conjunto de prueba que se han clasificado correctamente, mientras que, para el nivel de subcapítulo, la medida de precisión cuantifica el porcentaje de partidas clasificadas correctamente en su capítulo y que han sido correctamente clasificadas en su subcapítulo.

Para establecer el vector de pesos más adecuado para nuestros conjuntos de datos, se calcula la media ponderada de los resultados para cada conjunto según el número de partidas en cada uno de ellos. En la Tabla 14 se pueden observar los cinco mejores vectores y su correspondiente porcentaje de precisión.

	Precisión	Vectores de pesos		
Nivel Capítulo	94,18	0.5,0.4,0.1	0.5,0.2,0.1,0.1,0.1,0,0,0,0,0	
	94,17	0.5,0.4,0.1	0.6,0.1,0.1,0.1,0.1,0.1,0,0,0,0,0	
	94,07	0.5,0.4,0.1	0.5,0.1,0.1,0.1,0.1,0.1,0.1,0,0,0,0,0	
	93,93	0.6,0.3,0.1	0.6,0.1,0.1,0.1,0.1,0.1,0,0,0,0,0	
	93,89	0.6,0.3,0.1	0.5,0.2,0.1,0.1,0.1,0,0,0,0,0	
Nivel Subcapítulo	87,84	0.5,0.4,0.1	0.5,0.2,0.2,0.1,0,0,0,0,0,0	
	87,81	0.5,0.4,0.1	0.4,0.3,0.2,0.1,0,0,0,0,0,0	
	87,81	0.5,0.4,0.1	0.4,0.3,0.1,0.1,0.1,0,0,0,0,0	
	87,77	0.5,0.4,0.1	0.5,0.2,0.1,0.1,0.1,0,0,0,0,0	
	87,76	0.5,0.4,0.1	0.5,0.3,0.1,0.1,0,0,0,0,0,0	

Tabla 14: Los cinco mejores resultados para el nivel de capítulo y subcapítulo para ambos conjuntos de test

Como se puede observar en la Tabla 14, el vector $W = [0.5, 0.4, 0.1]$ es una buena elección para agregar las tres *importancias* de cada término en un único valor, y el vector $W = [0.5, 0.2, 0.1, 0.1, 0.1]$ para agregar las *importancias generales* en un único valor de *idoneidad*.

Del mismo modo, se procede a buscar el mejor vector de pesos para el proceso de clasificación en el nivel de subcapítulo. En este caso, la mejor opción también es el vector $W = [0.5, 0.4, 0.1]$ para agregar las tres *importancias* de cada término en un único valor, y el vector $W = [0.5, 0.2, 0.2, 0.1]$ para agregar las *importancias generales* en un único valor *idoneidad*.

4.4.3 Resultados detallados con la configuración OWA seleccionada

Para concluir nuestra descripción de la experimentación que hemos realizado con nuestro modelo, en esta sección se presentan con mayor detalle los resultados obtenidos en el conjunto de prueba con los vectores de pesos de los operadores OWA seleccionados en la sección anterior. Los resultados se dividen para los niveles de capítulo y subcapítulo.

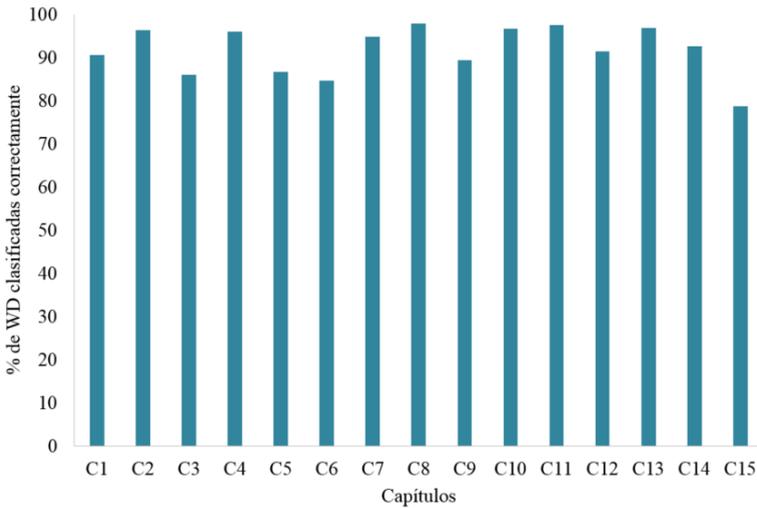


Figura 9: Resultados por capítulos en el nivel L_2 (capítulos)

En las Figuras 9 y 10 se pueden observar los resultados obtenidos para el nivel de capítulo y subcapítulo del conjunto de prueba correspondiente a las bases de precios. Como se puede observar en la Figura 9, en la mayoría de los capítulos las partidas se han clasificado correctamente con una precisión por encima del 90% y, sólo en cuatro se ha obtenido una precisión por debajo de este porcentaje (capítulos 3, 5, 6 y 15). En el nivel de subcapítulo, cinco han obtenido un resultado por encima del 90%, mientras que la mayoría se encuentran entre el 80% y el 90%. En este nivel, sólo tres de ellos han obtenido un resultado inferior.

Como se puede observar, los resultados obtenidos en el nivel de subcapítulo son más bajos que en el nivel de capítulo, lo que se explica porque en este nivel las descripciones de partida comparten más términos y es más complicado determinar la clasificación entre esos

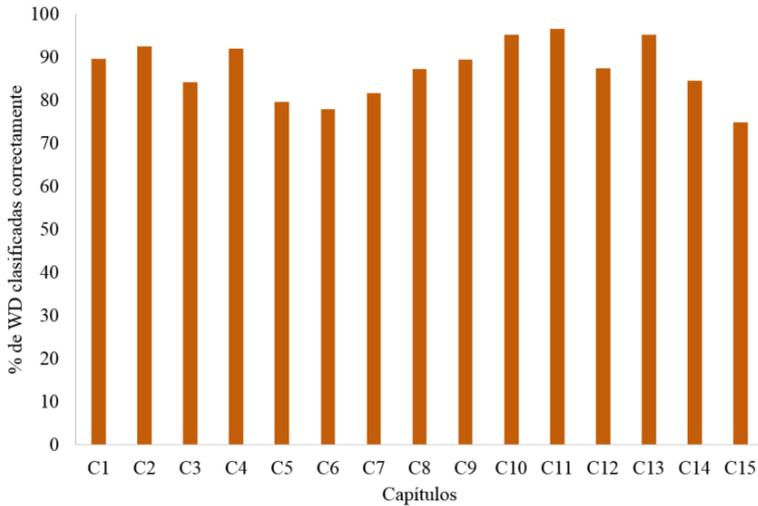


Figura 10: Resultados por subcapítulos en el nivel L_2 (capítulos)

subcapítulos. También, si aparecen términos nuevos o mal escritos que no están incluidos en nuestra lista de términos, la clasificación en el nivel de subcapítulo es más complicada, porque la naturaleza de los trabajos es muy similar. En el nivel de capítulo, las diferencias en cuanto a la naturaleza de los trabajos entre los diferentes capítulos son mayores.

En general, los resultados más bajos se obtienen por las siguientes razones:

- Capítulos o subcapítulos que utilizan un vocabulario similar para describir los trabajos. Por ejemplo, el capítulo 3, correspondiente a los trabajos de *saneamiento*, tiene muchos términos en común con el capítulo 8, correspondientes a los trabajos de *instalaciones*. Entre los subcapítulos de un mismo capítulo, es todavía más frecuente que éstos compartan gran cantidad de términos. Así por ejemplo, el capítulo 6 (*Albañilería*) tiene un total de 1353 términos y la distribución de términos para cada subcapítulo es muy elevada, tal y como se puede observar en la Tabla 18.
- Capítulos o subcapítulos que contienen descripciones de partidas de entrada descritas con términos diferentes a las consideradas en la fase de aprendizaje o que presentan errores tipográficos.

- Capítulos o subcapítulos que han tenido menor número de descripciones en la fase de aprendizaje y algunos de ellos tienen gran cantidad de términos. Es el caso de los capítulos 5,6,9 y 15.

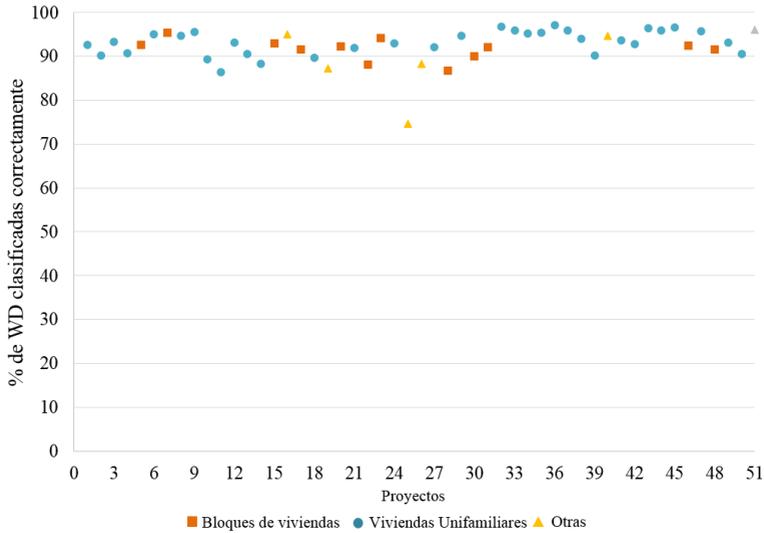


Figura 11: Resultados para el conjunto de proyectos en el nivel L₂ (capítulos)

Las Figuras 11 y 12 muestran los resultados detallados para el nivel de capítulo y subcapítulo para el conjunto de prueba correspondiente a los proyectos reales, respectivamente. En ambas figuras se han utilizado diferentes marcadores para distinguir las diferentes tipologías de proyectos; viviendas unifamiliares, bloques de viviendas y otros tipos de edificación, tales como: rehabilitación, administrativos o industriales. Dentro de esta última categoría, los resultados de las descripciones de partidas pertenecientes a las bases de precios se muestran en color gris claro.

Como se puede observar en la Figura 11, para el nivel de capítulo la mayoría de los proyectos obtienen un resultado por encima del 90% y sólo uno obtiene un resultado más bajo del 80%. En concreto, este proyecto corresponde a la categoría de *otros tipos de tipologías* donde las descripciones de partidas son más singulares y dificultan su clasificación de manera automática.

En el nivel de subcapítulo, la mayoría de los proyectos han obtenido resultado entre el 80% y el 90%. Al igual que en el caso anterior, los resultados son más bajos debido a que es más difícil distinguir los matices lingüísticos en partidas correspondientes a distintos grupos de

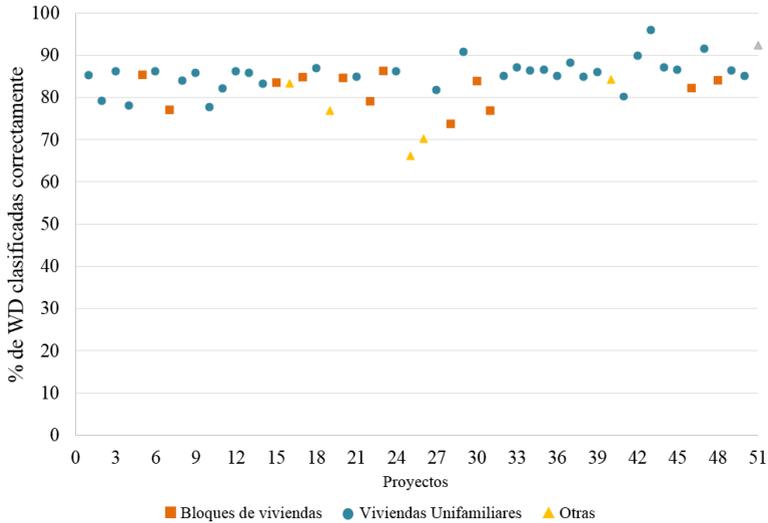


Figura 12: Resultados para el conjunto de proyectos en el nivel L₃ (subcapítulos)

este nivel. Dentro del pequeño grupo de proyectos que han obtenido resultado por debajo del 80%, se encuentran de nuevo los proyectos pertenecientes a *otros tipos de tipologías*.

En nuestra experimentación, también se ha analizado el comportamiento de las distintas importancias en el desempeño del operador OWA, mediante el análisis de las importancias (*frecuencia, posición y experto*) que han sido más influyentes en el proceso de clasificación. De esta manera, pretendemos comprobar si las tres *importancias* propuestas son útiles para el proceso de clasificación.

Para llevar a cabo este análisis, se ha calculado el porcentaje de veces que cada una de las importancias ha alcanzado la primera, segunda y tercera posición en el proceso de agregación del operador OWA.

En la Tabla 15 se pueden observar los resultados obtenidos con los valores establecidos previamente para el operador OWA. Las filas de las tablas no suman el 100% debido a los empates que se producen. Como se puede observar, las tres importancias juegan un papel muy importante en el proceso de clasificación. En el nivel de capítulo, las importancias de posición y experto son más destacadas, mientras que en el nivel de subcapítulo la importancia de la frecuencia obtiene más relevancia, en detrimento de la importancia del experto.

		I _e	I _f	I _p
Nivel de capítulo	1ª posición	44,91	31,93	39,99
	2ª posición	36,45	14,18	32,74
	3ª posición	18,63	53,90	27,27
		I _e	I _f	I _p
Nivel de subcapítulo	1ª posición	29,91	39,99	56,81
	2ª posición	32,34	22,34	20,55
	3ª posición	37,75	37,67	22,64

Tabla 15: Influencia de las *Importancias* en el proceso de clasificación

En la siguiente sección se presenta la evaluación de los resultados obtenidos para ambos conjuntos de prueba en el proceso de clasificación con otras medidas diferentes que son muy utilizadas en la literatura.

4.5 EVALUACIÓN

En esta sección, se plantea una evaluación del clasificador propuesto en base a la medida *F-measure* (ecuación 4.7), definida como la media armónica de las conocidas medidas *Precision* y *Recall*.

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (4.7)$$

En este sentido, dado un grupo de tareas GT_j^i , la medida *Precision* calcula el porcentaje de asignaciones correctas entre todas las descripciones asignadas al GT_j^i en el proceso de clasificación. La medida *Recall* representa el porcentaje de asignaciones correctas en GT_j^i entre todas las descripciones de partidas que deberían ser asignadas al GT_j^i .

Sea TP_j^i el número de verdaderos positivos (*true positive*), es decir, aquellas partidas que son de GT_j^i y el clasificador ha clasificado en GT_j^i .

Sea FP_j^i el número de falsos positivos (*false positive*), es decir, aquellas partidas que no son de GT_j^i y el clasificador las ha clasificado en GT_j^i .

Sea FN_j^i el número de falsos negativos (*false negative*), es decir, aquellas partidas que no son de GT_j^i pero el clasificador las ha clasificado como pertenecientes a dicho GT_j^i .

Definimos Pr_j^i (ecuación 4.8) y Re_j^i (ecuación 4.9) como sigue:

Definición 13 (Precision) : *porcentaje de casos clasificados que son relevantes.*

$$Pr_j^i = \frac{TP_j^i}{TP_j^i + FP_j^i} \quad (4.8)$$

Definición 14 (Recall) : *porcentaje de casos relevantes clasificados entre todos los casos relevantes.*

$$Re_j^i = \frac{TP_j^i}{TP_j^i + FN_j^i} \quad (4.9)$$

Lo ideal es obtener valores altos para las medidas de *Precision* y *Recall*, es decir, que la mayoría de las descripciones clasificadas son relevantes, y la mayoría de las descripciones relevantes son clasificadas.

	Precision	Recall	F-measure
C1	0,94	0,91	0,93
C2	0,98	0,96	0,97
C3	0,96	0,86	0,91
C4	0,92	0,96	0,94
C5	0,85	0,87	0,86
C6	0,92	0,84	0,87
C7	0,87	0,94	0,90
C8	0,95	0,98	0,96
C9	0,96	0,89	0,93
C10	0,94	0,97	0,96
C11	0,91	0,98	0,94
C12	0,95	0,91	0,93
C13	0,99	0,97	0,98
C14	0,95	0,92	0,93
C15	0,91	0,78	0,84

Tabla 16: *Precision, Recall y F-measure* para cada Capítulo

En las Tablas 16 y 17 se muestran los valores obtenidos para las medidas *Precision, Recall* y *F-measure* para cada capítulo y subcapítulo, respectivamente. Los resultados obtenidos para el nivel de subcapítulo, se basan en los resultados del nivel de capítulo. Es decir, los valores para *Precision, Recall* y *F-measure* en el nivel de subcapítulo corresponden

a las descripciones de partidas que han sido clasificadas correctamente en el nivel de capítulo.

Como se puede observar en la medida *F-measure* en las Tablas 16 y 17, se ha obtenido una buena relación entre las medidas *Precision* y *Recall*. En el caso de los capítulos (Tabla 16), la mayoría de los valores superan el 0.9, excepto en tres capítulos donde la medida *F-measure* se encuentra alrededor del 0.85. De manera similar ocurre en el caso de los subcapítulos (Tabla 17), donde en la mayoría de los casos, se han obtenido valores entre 0.8 y 0.9 para la medida *F-measure*. En cuatro subcapítulos, se han obtenido valores para esta medida que están por debajo de 0.6. Estos resultados se deben a que algunos subcapítulos no incluyen muchos términos discriminantes y por lo tanto, comparten términos con el resto de subcapítulos pertenecientes al mismo capítulo. Este hecho afecta a los valores obtenidos para la medida *Recall*.

Subcapítulo	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>
C1 SC1	0,95	0,90	0,92
C1 SC2	0,95	0,92	0,93
C1 SC3	0,95	0,64	0,77
C2 SC1	0,98	0,79	0,88
C2 SC2	1,00	0,75	0,86
C2 SC3	0,93	1,00	0,96
C2 SC4	1,00	1,00	1,00
C2 SC5	0,91	0,95	0,93
C2 SC6	1,00	0,54	0,70
C2 SC7	1,00	0,29	0,44
C2 SC8	1,00	0,92	0,96
C2 SC9	0,75	1,00	0,86
C3 SC1	1,00	0,73	0,85
C3 SC2	0,92	0,90	0,91
C3 SC3	1,00	0,81	0,90
C3 SC4	0,92	0,81	0,86
C3 SC5	0,97	0,87	0,92
C3 SC6	0,95	0,79	0,87
C4 SC1	0,88	0,94	0,91
C4 SC2	0,80	0,82	0,81
C4 SC3	0,92	0,99	0,95
C4 SC4	0,90	0,95	0,92
C4 SC5	0,87	0,84	0,86

Continúa en la página siguiente.

Subcapítulo	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>
C5 SC1	0,76	0,91	0,83
C5 SC2	0,84	0,79	0,81
C5 SC3	0,69	0,84	0,76
C5 SC4	0,95	0,90	0,93
C5 SC5	0,83	0,37	0,51
C6 SC1	0,78	0,83	0,80
C6 SC2	0,83	0,96	0,89
C6 SC3	0,97	0,85	0,91
C6 SC4	0,96	0,74	0,83
C6 SC5	0,76	0,47	0,58
C6 SC6	0,99	0,93	0,96
C7 SC1	0,78	0,76	0,77
C7 SC2	0,82	0,95	0,88
C7 SC3	0,76	0,65	0,70
C7 SC4	0,74	0,63	0,68
C7 SC5	0,57	0,82	0,68
C7 SC6	0,73	0,82	0,77
C8 SC1	0,80	0,85	0,82
C8 SC2	0,95	0,93	0,94
C8 SC3	0,87	0,93	0,90
C8 SC4	0,79	0,77	0,78
C8 SC5	0,85	0,84	0,84
C8 SC6	0,82	0,80	0,81
C8 SC7	0,95	0,45	0,61
C8 SC8	0,95	0,90	0,92
C8 SC9	0,78	0,88	0,82
C9 SC1	0,94	0,86	0,90
C9 SC2	0,98	0,92	0,95
C10 SC1	0,93	0,98	0,96
C10 SC2	0,95	0,92	0,94
C10 SC3	0,92	0,97	0,94
C11 SC1	0,92	0,98	0,95
C11 SC2	0,90	0,98	0,94
C11 SC3	0,85	0,99	0,92
C11 SC4	0,88	0,81	0,84
C11 SC5	0,95	0,84	0,89

Continúa en la página siguiente.

Subcapítulo	<i>Precision</i>	<i>Recall</i>	<i>F-measure</i>
C12 SC1	0,91	0,95	0,93
C12 SC2	0,91	0,82	0,86
C12 SC3	0,92	0,39	0,55
C13 SC1	0,97	1,00	0,98
C13 SC2	0,96	0,89	0,93
C14 SC1	0,87	0,95	0,91
C14 SC2	0,71	0,69	0,70
C14 SC3	0,95	0,44	0,60
C15 SC1	0,85	0,73	0,78
C15 SC2	0,88	0,52	0,65
C15 SC3	0,97	0,94	0,95

Tabla 17: *Precision*, *Recall* y *F-measure* para cada subcapítulo

4.6 CONCLUSIONES

En este capítulo se ha presentado una metodología que permite la clasificación de las descripciones de partida contenidas en un documento de presupuesto dado en la estructura de referencia propuesta en el capítulo anterior.

La base de la metodología son los términos que aparecen en las descripciones de partida para definir las diferentes actividades del proyecto. Previamente, se ha construido un modelo de conocimiento en el que se realiza un análisis de los términos con el fin de establecer la importancia de un determinado término cuando aparece en una descripción para determinar la colocación de ésta en un nodo determinado de la estructura de referencia. En base a este análisis, se ha propuesto el uso de un modelo de agregación multi-criterio basado en tres *importancias*: *frecuencia*, *posición* y *experto*. En la experimentación realizada en este capítulo, se ha analizado el comportamiento de las distintas *importancias* en el desempeño del proceso de clasificación y se ha podido comprobar como las tres importancias juegan un papel fundamental en el proceso de clasificación.

Con el fin de evaluar nuestra propuesta, se ha realizado una experimentación con dos conjuntos de datos muy extensos (un total de 22399 descripciones de partidas), distintos a los utilizados en la fase de aprendizaje. El primer conjunto está compuesto por descripciones de partidas pertenecientes a dos bases de precios y, el segundo con-

junto está compuesto de descripciones de partidas procedentes de proyectos reales que han sido realizados por diferentes arquitectos y correspondientes a diferentes tipologías de edificación.

Los resultados obtenidos muestran como nuestra metodología es capaz de clasificar las descripciones de partidas de los dos conjuntos de datos de prueba con un alto porcentaje de acierto, independientemente del marco lingüístico utilizado para definir las descripciones de partida. El nivel de desempeño permite incorporar la metodología en un módulo de adquisición de datos para nutrir un repositorio común de los datos contenidos en el presupuesto.

Para ello, en el siguiente capítulo, se presenta una herramienta que permite la gestión de los datos contenidos en el documento del presupuesto de manera automática e integrada. Esta herramienta se compone de tres módulos, que permiten la adquisición, clasificación, edición y recuperación de los datos de manera fácil e intuitiva.

	Nº D	Nº T		Nº D	Nº T
C1 Trabajos previos	620	558	SC3 Cubiertas planas no transitables	51	717
SC1 Consolidaciones	45	485	SC4 Cubiertas planas transitables	33	701
SC2 Derribos	560	62	SC5 Bajantes-Canalones	50	1093
SC3 Cargas y Transportes	15	458	SC6 Remates Varios	104	804
C2 Acondicionamiento del Terreno	226	229	C8 Instalaciones	1768	3797
SC1 Preparación del Terreno	19	187	SC1 Fontanería	170	3201
SC2 Explanación	3	208	SC2 Aparatos sanitarios	199	2917
SC3 Excavación	152	144	SC3 Eléctrica e Iluminación	192	2176
SC4 Refino	4	202	SC4 Telecomunicaciones-Informática	153	2460
SC5 Relleno	15	139	SC5 Calefacción y ACS	261	2448
SC6 Compactación	4	201	SC6 AA Ventilación	297	2946
SC7 Carga	8	192	SC7 Gas	51	3451
SC8 Transporte	18	192	SC8 Elevación	90	3538
SC9 Canon	3	219	SC9 Protección	351	2420
C3 Red Horizontal de Saneamiento	134	543	C9 Aislamientos e Impermeabilizaciones	264	1302
SC1 Acometidas	4	490	SC1 Aislamientos	189	584
SC2 Arquetas y Pozos	46	294	SC2 Impermeabilizaciones	75	396
SC3 Sistemas de depuración	13	485	C10 Revestimientos	611	2106
SC4 Tuberías y Colectores	43	345	SC1 Falsos techos	100	1564
SC5 Sumideros	16	415	SC2 Paramentos	311	1186
SC6 Sistemas drenantes	12	300	SC3 Pavimentos	200	572
C4 Cimentaciones	461	588	C11 Carpintería de madera	263	565
SC1 Armadura	11	482	SC1 Puertas	145	325
SC2 Cimentaciones especiales	36	191	SC2 Armarios	45	380
SC3 Encofrados	19	545	SC3 Ventanas	43	321
SC4 Hormigones	75	473	SC4 Escaleras-Barandillas-Pasamanos	11	462
SC5 Soleras y Losas	20	423	SC5 Persiana-Capialzado-Celosía	19	408
C5 Estructura	323	890	C12 Carpintería exterior y cerrajería	796	1251
SC1 Estructura de acero	70	536	SC1 Carpintería exterior	547	453
SC2 Estructura de hormigón	185	497	SC2 Cierre-Protección	200	530
SC3 Estructura de hormigón prefabricado	19	584	SC3 Varios	49	801
SC4 Estructura de madera	29	743	C13 Vidriería y elaborados sintéticos	241	471
SC5 Varios	20	593	SC1 Vidrios	104	277
C6 Albañilería	434	1353	SC2 Vidrios especiales y varios	137	68
SC1 Cantería	70	902	C14 Pinturas	266	891
SC2 Cerramientos	202	625	SC1 Pinturas	185	284
SC3 Divisiones Interiores	81	1015	SC2 Tratamientos	43	627
SC4 Recibidos	20	1134	SC3 Pinturas especiales	38	379
SC5 Prefabricados, ventilación y varios	43	897	C15 Equipamiento, decoración y varios	537	1479
SC6 Ayudas	18	1194	SC1 Equipamiento	235	691
C7 Cubiertas	385	1168	SC2 Piscina	48	1129
SC1 Formación de pendiente	56	798	SC3 Jardín-Riego	254	821
SC2 Cubiertas inclinadas	91	623			

Tabla 18: Número de términos y número de descripciones de partidas para cada capítulo (L2) y subcapítulo (L3) de la estructura de referencia

APLICACIÓN PARA LA ADQUISICIÓN DE DATOS Y GESTIÓN DE COSTES

RESUMEN: En este capítulo se presentan dos aplicaciones que ponen en valor el modelo y la técnica de clasificación propuestas en los dos capítulos anteriores.

En primer lugar, se presenta la aplicación *SIGPE* (Sistema Inteligente de Gestión de Proyectos de Edificación) que permite la adquisición, gestión y recuperación de los datos contenidos en el presupuesto de manera estructurada e integrada desde cualquier lugar o dispositivo. En segundo lugar, se propone una herramienta flexible para el análisis de los costes mediante el modelo multidimensional, que incorpora conceptos de la Teoría de los Conjuntos Difusos.

ÍNDICE

5.1	Introducción	114
5.2	Sistema de adquisición y edición	115
5.2.1	Módulo de adquisición	116
5.2.1.1	Importación del documento	116
5.2.1.2	Proceso de clasificación	119
5.2.2	Módulo de edición	122
5.2.3	Módulo de recuperación	124
5.2.4	Implementación de la aplicación <i>SIGPE</i>	126
5.2.4.1	Arquitectura del sistema on-line	127
5.2.4.1.1	Servidor	127
5.2.4.1.2	Cliente	128
5.2.5	Evaluación de la aplicación	129
5.2.5.1	Resultados	129
5.2.5.2	Cuestionario	132
5.3	Un herramienta flexible para el análisis de costes	135
5.3.1	Estructura multidimensional sobre datos de costes	136
5.3.1.1	Conceptos previos modelo multidimensional difuso	137
5.3.1.2	Propuesta de estructura multidimensional	139
5.3.1.2.1	Dimensiones	139
5.3.1.2.2	Medidas	145
5.3.1.2.3	Cubo de datos	146
5.3.1.3	Consultas	146
5.3.2	Implementación de la aplicación	150
5.4	Conclusiones	151

5.1 INTRODUCCIÓN

Como hemos indicado desde el inicio de esta memoria, el acceso a los datos de manera integral es fundamental para mejorar el proceso de toma de decisiones en cualquier organización. Por desgracia, en la práctica actual dentro del ámbito de la edificación, dicho acceso no es tarea fácil debido principalmente a la desestructuración de los datos causada por la forma de funcionar tradicional de esta industria y la falta de estándares de referencia a la hora de estructurar la información. Además, la cantidad de datos que se debe manejar, así como la diversidad de personas que intervienen en el proceso de construcción, hacen que el desarrollo de herramientas integrales de gestión de datos sea un reto.

Inicialmente, para solventar los inconvenientes mencionados, se implantaron sistemas para la gestión centralizada de documentos, facilitando así el acceso a todos los participantes de la empresa. Sin embargo, estos sistemas se basan en repositorios estáticos que almacenan gran cantidad de documentos con el propósito de permitir el intercambio de y el acceso a los documentos que, anteriormente, se encontraban dispersos en diferentes sistemas o en papel. A pesar de la gran utilidad que pueden tener los datos contenidos en estos documentos, no se ha prestado atención a la estructuración de dichos datos con el fin de aprovechar las ventajas que éstos pueden aportar para futuros proyectos.

Hoy en día, sigue existiendo la necesidad de herramientas que integren los datos de manera estructurada y que permitan el acceso y gestión de éstos con el fin de dar soporte a la toma de decisiones. Además, no sólo el manejo de los datos es necesario, sino que la extracción de información a partir de grandes conjuntos de datos se está convirtiendo en una necesidad cada vez más importante en cualquier entorno empresarial. Y herramientas como las presentadas en los dos capítulos anteriores hacen posible la construcción de repositorios integrados que se nutran no sólo de nuevos documentos bien estructurados, sino también de antiguos documentos cuyas descripciones de partidas sean bien clasificadas de manera automática en las estructuras de referencia.

En este capítulo de la Tesis, presentamos dos aplicaciones para poner en valor el modelo y la técnica de clasificación propuestas en los dos capítulos anteriores.

Por un lado, se presenta una aplicación interactiva *SIGPE* (Sistema Inteligente de Gestión de Proyectos de Edificación) que permite gestionar los datos contenidos en el documento del presupuesto de una

manera estructurada e integrada desde cualquier lugar o dispositivo. La aplicación está compuesta de tres módulos:

- El **módulo de adquisición de datos** permite la importación y clasificación de la información contenida en el presupuesto independientemente de la herramienta utilizada para la elaboración del documento, y de la estructura y marco lingüístico utilizado. El corazón de este módulo es el método de clasificación automático presentado en el capítulo anterior.
- El **módulo de edición de datos** permite la edición de los resultados obtenidos tras el proceso de clasificación, con el fin de ubicar cada descripción de partida en su capítulo y subcapítulo correspondiente.
- El **módulo de recuperación de datos** permite acceder a los datos almacenados en la estructura de referencia de manera integrada para dar soporte a la toma de decisiones. En esta versión de prueba de concepto, se incorporan algunas funcionalidades básicas de consulta.

Por otro lado, se muestra cómo se puede utilizar un repositorio integrado para nutrir de datos hipercubos sobre los que realizar procesamiento analítico en línea (OLAP). En este sentido, se propone un modelo multidimensional de datos para el análisis de costes, que incorpora conceptos de la Teoría de Conjuntos Difusos para facilitar la consulta flexible de datos y que se ha implementado con la herramienta Linguistic F-Cube Factory [47].

5.2 SISTEMA DE ADQUISICIÓN Y EDICIÓN DE PRESUPUESTOS DE PROYECTOS DE EDIFICACIÓN

En esta sección se presenta el sistema *SIGPE* para la adquisición y edición de los datos contenidos en el documento del presupuesto de una manera estructurada e integrada. Para ello, en la sección 5.2.1 se detalla el módulo de adquisición de datos que se divide en dos procesos, por un lado, la importación de los datos y, en segundo lugar, la clasificación de éstos en la estructura de referencia mediante el proceso de clasificación descrito en el capítulo anterior. Posteriormente, en la sección 5.2.2 se muestra el módulo que permite a los usuarios editar y validar los resultados obtenidos en el primer módulo. A continuación, en la sección 5.2.3 se presenta el módulo de recuperación

que permite acceder a los datos almacenados en la estructura de referencia de manera integrada. Finalmente, en las secciones 5.2.4 y 5.2.5 se detallan los datos de la implementación y se describe el proceso de evaluación que se ha realizado, respectivamente.

5.2.1 *Módulo de adquisición*

El módulo de adquisición permite la captura y clasificación de la información contenida en el presupuesto independientemente de la herramienta utilizada para la elaboración de éste, y de la estructura y marco lingüístico utilizado. Dentro de este módulo se diferencian dos procesos: por un lado, el proceso que permite importar el documento del presupuesto y, por otro lado, el proceso que permite la clasificación de las descripciones de partidas en la estructura de referencia de nuestro almacén de datos común.

5.2.1.1 **Importación del documento**

Como se ha mencionado en esta memoria, los responsables de la elaboración del presupuesto utilizan diferentes herramientas comerciales disponibles en el mercado para la preparación de dicho documento. Sin embargo, en la práctica actual, aunque se utilizan estas herramientas para generar el documento, los usuarios no gestionan los datos durante el resto del ciclo de vida del proyecto a través de estas herramientas. El formato más común para llevar a cabo la tarea de gestión y control de costes entre los distintos participantes es la hoja de cálculo. Todas las herramientas permiten exportar los presupuestos elaborados a hojas de cálculo, aunque cada una de éstas exporta los datos con distintas estructuras.

Sin embargo, la hoja de cálculo no es la herramienta más adecuada para llevar a cabo las tareas relacionadas con la gestión de costes en el proceso de la construcción. Las hojas de cálculo dificultan las tareas de reutilización y consulta sobre la información contenida en estos documentos, además de otros inconvenientes mencionados anteriormente, tales como duplicidad de documentos o pérdida de información.

En los capítulos anteriores hemos motivado la necesidad de disponer la información estructurada para poder reutilizarla de manera automática con el fin de ofrecer soporte para el proceso de toma de decisiones. También hemos puesto de manifiesto que pretendemos integrar tanto los presupuestos de los proyectos actuales, así como los presupuestos de proyectos históricos. Por lo tanto, es necesario proporcionar un

mecanismo que configure la entrada de los datos a nuestra aplicación, de modo que se puedan importar los datos directamente desde los formatos establecidos por las distintas herramientas que se utilizan para elaborar los presupuestos. Por tanto, este mecanismo proporcionará las descripciones de partida, que son la base para el posterior proceso de clasificación.

En las Figuras 13 y 14 se puede observar un ejemplo de la diferencia entre las dos estructuras de metadatos, cuando se exporta el presupuesto en formato Excel [176], que usan dos de las herramientas más utilizadas para la elaboración del presupuesto, Cype [75] y Presto [196]. Como se puede observar en las figuras, la organización de los datos que aparecen en las columnas coincide en estos dos formatos, aunque se denominen de manera diferente. La primera indica el código, que dependerá de la base de precios que se haya utilizado o del criterio de la persona que lo ha elaborado. El resto de columnas identifican el nivel de la estructura, la unidad de medida, el resumen de cada nodo de la estructura, cantidad, precio e importe. Sin embargo, la información que se incluye en cada fila difiere según el formato de origen. Por ejemplo, las descripciones de partida, que son la base de nuestra propuesta, aparecen en campos diferentes. En el presupuesto de la Figura 13 se puede observar cómo la descripción de partida aparece en la misma fila que el tipo *Partida*, mientras que en el presupuesto de la Figura 14 en la fila de *Partida* aparece un resumen de ésta. La descripción detallada de la partida aparece en la fila inferior.

La aplicación que hemos desarrollado permite la configuración de los datos de entrada y se basa en el uso de esquemas. Un esquema es un fichero XML que proporciona la información necesaria a la aplicación para *entender* los ficheros de salida de las herramientas que, a su vez, son los ficheros de entrada para nuestra herramienta de adquisición y edición de presupuestos. Se admiten dos tipos de esquemas:

- Esquema para Excel. En nuestro caso, la mayoría de los documentos de entrada están en este formato. Por tanto, los esquemas indicarán a la aplicación las columnas y desplazamientos necesarios para localizar los distintos datos contenidos en el documento.
- Esquema para XML. Se tienen que configurar las distintas etiquetas que proporcionan la información que se requiere de un presupuesto en formato XML.

Por lo tanto, estos esquemas permiten extraer las descripciones de partida del presupuesto, independientemente de la estructura del

Obra: REFORMA Y AMPLIACION DE VIVIENDA UNIFAMILIAR									
Presupuesto							% C.I. 3		
Código	Tipo	Ud	Resume	Cantida	Precio (€)	Importe			
MEDICION	Capitulo		REFORMA Y AMPLIACION DE VIVIENDA UNIFAMILIAR		67.207,56	67.207,56			
D	Capitulo		Demoliciones		4.611,57	4.611,57			
DE	Capitulo		Estructuras		2.786,53	2.786,53			
DEH	Capitulo		Hormigón		2.786,53	2.786,53			
DEH022	Partida	m ²	Corte de forjado unidireccional de hormigón armado con viguetas prefabricadas de hormigón, entrevigado de bovedillas cerámicas o de hormigón y capa de compresión de hormigón, con sierra con disco diamantado, y carga manual de escombros sobre camión o contenedor.	4,200	663,46	2.786,53			
				Uds.	Largo	Ancho	Alto	Parcial	Subtotal
			HUECO ESCALERAS	1	4,200	1,000		4,200	4,200
			DEH					2.786,53	2.786,53
			DE					2.786,53	2.786,53
DQ	Capitulo		Cubiertas		1.646,69	1.646,69			
DQC	Capitulo		Capa de cobertura		1.030,58	1.030,58			
DQC040	Partida	m ²	Arranque de cobertura de teja cerámica curva y elementos de fijación, colocada con mortero a menos de 20 m de altura, en cubierta inclinada con una pendiente media del 30%, con medios manuales y carga manual de escombros sobre camión o contenedor.	103,472	9,96	1.030,58			
				Uds.	Largo	Ancho	Alto	Parcial	Subtotal
			PAÑOS DE CUBIERTA	1	8,920	6,230		55,572	
				1	8,920	5,370		47,900	103,472
			DQC					1.030,58	1.030,58

Figura 13: Ejemplo de presupuesto elaborado con la herramienta Cype y exportado a formato Excel

Vivienda unifamiliar en Pozuelo						
Presupuesto						
Código	Nat	Ud	Resumen	CanPres	PrPres	ImpPres
01	Capitulo		MOVIMIENTO DE TIERRAS	1	2.518,67	2.518,67
E02AM010	Partida	m2	Desbroce y limpieza de terreno a máquina	61,52	0,45	27,68
			Desbroce y limpieza superficial del terreno por medios mecánicos, sin carga ni transporte al vertedero y con p.p. de medios auxiliares.			
O01OA070	Mano de obra	h.	Peón ordinario	0,005	13,09	0,07
M05PND10	Maquinaria	h.	Pala cargadora neumáticos 85 cv/1,2m3	0,010	38,00	0,38
			E02AM010	61,52	0,45	27,68
E02EM030	Partida	m3	Excavación vaciado a máquina terreno compacto	66,38	12,14	805,85
			Excavación en zanjas, en terrenos compactos, por medios mecánicos, con extracción de tierras a los bordes, sin carga ni transporte al vertedero y con p.p. de medios auxiliares.			
O01OA070	Mano de obra	h.	Peón ordinario	0,125	13,09	1,64
M05EN030	Maquinaria	h.	Excav.hidráulica neumáticos 100 cv	0,250	42,00	10,50
			E02EM030	66,38	12,14	805,85

Figura 14: Ejemplo de presupuesto elaborado con la herramienta Presto y exportado a formato Excel

documento original para organizar la información. Estas descripciones de partida, son la base del proceso de clasificación.

El usuario, a través de la definición de los esquemas, indica a SIGPE cómo interpretar el fichero de entrada. Como resultado de este proceso, la información contenida en el documento del presupuesto del fichero de entrada se importa de manera automática en nuestro sistema, con

el fin de permitir el proceso de clasificación que se presenta en la siguiente sección.

5.2.1.2 Proceso de clasificación

Una vez que la aplicación es capaz de importar las descripciones de partida del documento del presupuesto, el siguiente objetivo dentro de este módulo es la clasificación de éstas en la estructura de referencia propuesta en el capítulo 3.

Como se ha descrito en el capítulo 4, el proceso de clasificación se basa en los términos que aparecen en las descripciones de partida, para determinar la clasificación en un capítulo y subcapítulo de la estructura de referencia. Para este propósito, se ha presentado en dicho capítulo nuestra propuesta de agregación multi-criterio basada en tres dimensiones (*frecuencia, posición y experto*). La aplicación de este método proporciona para cada partida un *valor de idoneidad* en el intervalo $[0, 1]$ para cada capítulo y subcapítulo de la estructura de referencia. Este valor establece el *grado de compatibilidad* entre la descripción de partida y cada uno de los capítulos y subcapítulos de la estructura de referencia. La descripción de la partida se clasificará en el capítulo y subcapítulo donde se obtengan el resultado más alto, tal y como se ha explicado en la sección 4.3.1.

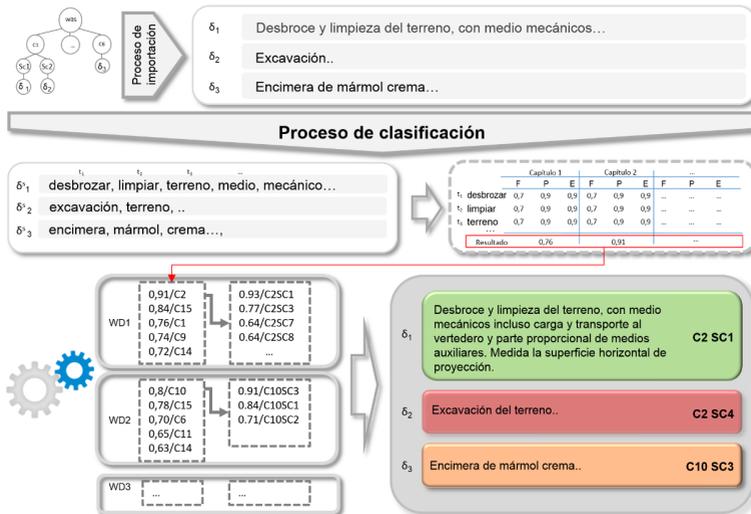


Figura 15: Ejemplo del proceso del módulo de adquisición

La Figura 15 ilustra el proceso del módulo de adquisición. En la parte superior izquierda se puede observar nuestro presupuesto de ejemplo con la estructura de descomposición de trabajo WBS. En el nivel más bajo de la estructura se encuentran las descripciones de partidas (δ_1, δ_2 y δ_3) que se pretenden clasificar en nuestro almacén común. Estas descripciones de partidas están descritas y organizadas en su estructura de origen según el criterio de la persona que ha elaborado el documento.

Mediante el proceso de adquisición, se extrae el conjunto de las descripciones de partida de dicho presupuesto, como se puede observar en la parte superior derecha. A continuación, una vez extraídas las descripciones de partida, se aplica el proceso de clasificación que, mediante una serie de cálculos, permite la clasificación de cada una de éstas descripciones de partida en la estructura de referencia.

Como se puede observar, la descripción de partida δ_1 se clasifica en el Capítulo C2, ya que es el capítulo donde se ha obtenido el *resultado* más alto (0,91). Una vez que la descripción se ha clasificado en un capítulo, se procede a la clasificación de ésta en uno de los subcapítulos pertenecientes a dicho capítulo. En este caso, el subcapítulo SC1 ha obtenido el *resultado* más alto (0,93). Como se puede observar en la Figura 15, la aplicación muestra el resultado de δ_1 en color verde, mientras que para la segunda δ_2 se muestra en color rojo y en naranja para la tercera δ_3 .

Esta escala de colores representa la fiabilidad con la que el clasificador propuesto determina la clasificación de las descripciones de partidas en la estructura de referencia. Esta funcionalidad, como veremos en la siguiente sección, permite a los usuarios localizar las descripciones de partidas en las que hay mayor riesgo de error tras su clasificación en un capítulo o subcapítulo dados, con el fin de facilitar la tarea de edición y validación de los resultados proporcionados por el clasificador.

A continuación, se muestran los pasos a seguir para la inserción de un presupuesto en SIGPE a través del módulo de adquisición que proponemos:

1. Rellenar los campos que se muestran en la pantalla con la información del nuevo proyecto y adjuntar el fichero del presupuesto (Figura 16).
2. A continuación, la aplicación extrae de manera automática las descripciones de partidas contenidos en el fichero de entrada y presenta al usuario la información del presupuesto (Figura 17).

ADQUISICIÓN EDICIÓN RECUPERACIÓN

Nuevo proyecto

VIV.UNIF.0840

ESPAÑA ANDALUCIA

GRANADA GRANADA

EDIFICACIÓN 2015-10-13

Vivienda Unifamiliar aislada en calle San Jerónimo nº 7

Documento del presupuesto (P.ej. .xlsx)

Seleccionar archivo | WBS_VivUnif0840.xlsx

ACEPTAR

Tratamiento Inteligente de Datos en Proyectos de Edificación
© Copyright 2015 - María Martínez Rojas

Figura 16: Pantalla para insertar un nuevo proyecto en la aplicación

ADQUISICIÓN EDICIÓN RECUPERACIÓN

VIV.UNIF.0840

PROCESO DE CLASIFICACIÓN

Resumen del proyecto

244 partidas pendientes de clasificar

País: ESPAÑA
Comunidad autónoma: ANDALUCIA
Ciudad: GRANADA
Población: GRANADA
Descripción: Vivienda Unifamiliar aislada en calle San Jerónimo nº 7

Tratamiento Inteligente de Datos en Proyectos de Edificación
© Copyright 2015 - María Martínez Rojas

Figura 17: Pantalla que muestra la información del proyecto y permite iniciar el proceso de clasificación

3. A continuación se procede a la clasificación de las descripciones importadas en la estructura de referencia. En la Figura 18 se

muestra la pantalla que aporta la distribución de estas descripciones de partidas en los capítulos de la estructura de referencia.



Figura 18: Pantalla que muestra la información del proyecto y la distribución de las descripciones de partida

5.2.2 Módulo de edición

Como se ha visto en el capítulo anterior, la metodología de clasificación propuesta consigue un buen porcentaje de aciertos. Sin embargo, para tener un almacén de datos consolidado, es necesario que todas las descripciones de partidas se encuentren ubicadas en su capítulo y subcapítulo correspondiente. De este modo, se garantiza que durante el proceso de consulta, se recuperen todos los datos adecuados.

Gracias a este módulo de edición, los usuarios pueden identificar de forma sencilla las descripciones de partidas que pueden no haber sido clasificadas correctamente durante el proceso de clasificación, con el fin de ubicarlas, en su caso, en su correspondiente capítulo y/o subcapítulo. Para facilitar al usuario la tarea de identificación de descripciones de partidas que pueden no estar correctamente clasificadas, se propone un enfoque visual que proporciona a los usuarios la fiabilidad con la que el clasificador determina la ubicación de la descripción de partida en la estructura de referencia. Para ello, se propone una escala de color que se establece a partir del *grado de compatibilidad* que se obtiene en el

proceso de clasificación. Con este enfoque, el usuario puede identificar “a simple vista” las descripciones de partida que pueden estar mal clasificadas y que, por tanto, deberá modificar.

La escala de color propuesta se compone de cinco colores: verde oscuro, verde claro, naranja, rojo claro y rojo oscuro. El verde corresponde con los resultados para las descripciones de partida que el clasificador está más seguro de que están bien clasificadas, mientras que el resultado de las de color rojo es más incierto.

La escala de colores se ha establecido en base a los resultados obtenidos en la experimentación del capítulo anterior para el conjunto de prueba de los proyectos reales. Para cada una de las descripciones de partida de este conjunto, se obtiene un *grado de compatibilidad*, el cual nos permite establecer unos intervalos para determinar la escala de color que muestren al usuario de manera lo más certera posible los resultados de los nuevos proyectos que se van a incorporar en el almacén común.

Los intervalos se han establecido con la idea de que las descripciones de partida que aparecen en color verde proporcionen al usuario un alto grado de fiabilidad. De este modo, es muy poco frecuente que una descripción de partida en color verde esté mal clasificada en la estructura de referencia. Sin embargo, para poder garantizar este criterio, es probable que descripciones de partida que hayan sido clasificadas correctamente, aparezcan en tonos naranjas o rojos. Consideramos que es preferible para el usuario, cuando se enfrenta a la tarea de edición y validación, identificarle las descripciones que hayan sido clasificadas de manera errónea, a pesar de que para alcanzar este objetivo, alguna descripción de partida bien clasificada aparezca en colores naranjas o rojos.

En la Figura 19 se muestra una captura de pantalla donde se pueden observar los diferentes colores que representan la fiabilidad de los resultados proporcionados por el clasificador.

En este módulo, el usuario deberá validar las descripciones de partida que estén correctamente clasificadas y, modificará el capítulo y/o subcapítulo de las descripciones mal clasificadas. Para ello, la aplicación dispone de un menú desplegable para cada partida que permite editar el resultado obtenido en el proceso de clasificación tanto para los capítulos como para los subcapítulos. En la Figura 20 se puede observar dicho menú.

Una vez que el usuario ha concluido la fase de validación y edición a través de este módulo, todas las descripciones del presupuesto

ADQUISICIÓN EDICIÓN RECUPERACIÓN			
C2-ACONDICIONAMIENTO DEL TERRENO	8845.71€	Contiene 4 subcapítulos	
C3-RED HORIZONTAL DE SANEAMIENTO	3448.66€	Contiene 5 subcapítulos	
C4-CIMENTACIONES	50773.54€	Contiene 4 subcapítulos	
C5-ESTRUCTURA	28550.66€	Contiene 2 subcapítulos	
C6-ALBAÑILERIA	66079.49€	Contiene 5 subcapítulos	
C7-CUBIERTAS	6889.86€	Contiene 2 subcapítulos	
C8-INSTALACIONES	31948.04€	Contiene 7 subcapítulos	
ELECTRICA E ILUMINACIÓN 6,403.78 € Contiene 23 partidas			
ACERO CORR. PREFOR. B 500 S	C 0.64	P 1.46	T 0.93
Acero corrugado B 500 S, preformado en taller y colocado en obra. Según EHE y CTE-SE-A.	✓		
	Selecione un capítulo		
Gastos tramitación contratación suministro eléctrico	C 1.00	P 52.53	T 52.53
Ud. Gastos tramitación contratación por Kie con la Compañía para el suministro al edificio desde sus redes de distribución, incluido derechos de acometida, enganche y verificación en la contratación de la póliza de abono.	✓		
	Selecione un capítulo		
PLUZ SENCILLO SIMÓN 75	C 9.00	P 29.00	T 261.00
Punto de luz sencillo realizado con tubo PVC corrugado de M 20/gp5 y conductor rígido de 1.5 mm2 de Cu, y aislamiento VV 750 V, incluyendo caja de registro, caja de mecanismo universal con tornillos, interruptor unipolar Simón serie 75, instalado.	✓		
	Selecione un capítulo		
PLUZ CONMUTADO SIMÓN 75	C 6.00	P 39.39	T 236.34
Punto conmutado sencillo realizado con tubo PVC corrugado de M 20/gp5 y conductor rígido de 1.5 mm2 de Cu, y aislamiento VV 750 V, incluyendo caja de registro, cajas de mecanismo universal con tornillos, conmutadores Simón serie 75, instalado.	✓		
	Selecione un capítulo		
P.DOUBLE CONMUTAD. SIMÓN 75	C 2.00	P 66.07	T 132.14
Punto doble conmutador realizado con tubo PVC corrugado de M 20/gp5 y conductor rígido de 1.5 mm2 de Cu, y aislamiento VV 750 V, incluyendo caja de registro, cajas de mecanismo universal con tornillos, conmutadores Simón serie 75, instalado.	✓		

Figura 19: Pantalla con los capítulos desplegados

estarán almacenadas correctamente en la estructura de referencia. A continuación, se presenta el módulo de consulta.

5.2.3 Módulo de recuperación

Una vez que todas las descripciones de partida se han clasificado en su correspondiente capítulo y subcapítulo, gracias a este módulo se permite el acceso a los datos de todos los presupuestos que han sido almacenados en el repositorio común. En la Figura 21 se puede observar la pantalla que muestra los proyectos que se han almacenado en el repositorio común.

El hecho de que la aplicación esté desarrollada de manera interactiva y en línea aporta importantes beneficios para los usuarios, ya que podrán acceder a los datos en cualquier momento, lugar y dispositivo, con el único requisito de disponer una conexión a Internet.

The screenshot displays a software interface for managing construction projects. It features a list of items with a dropdown menu for modification. The items are categorized into sections like VENTANAS, ARMARIOS, and EQUIPAMIENTO. The dropdown menu is currently open, showing options C12-Carpintería Exterior y Cerrajería, C13-Vidriería, C14-Pinturas, and C15-Equipamiento. The interface also shows a total price of 1,048.27 € for the selected items.

Item Description	C	P	T
PREPERCO PINO 70x35 mm.P/1 HOJA	1.00	12.51	12.51
PUERTA CHAPA LISA 70x200 GALV.	1.00	72.20	72.20
PUERTA 0.80x2.00 40/14 STD	1.00	148.58	148.58
VENTANAS			1,048.27 €
BARN.MADERA INT.BRILLANT 2 MAN.	98.80	10.61	1,048.27
ARMARIOS			
FTE.ARM.CORR.LISO PINO PIPINTAR			
C12-CARPINTERÍA EXTERIOR Y CERRAJERÍA			47403.82€
C13-VIDRIERÍA			3607.33€
C14-PINTURAS			13762.04€
C15-EQUIPAMIENTO			4787.37€

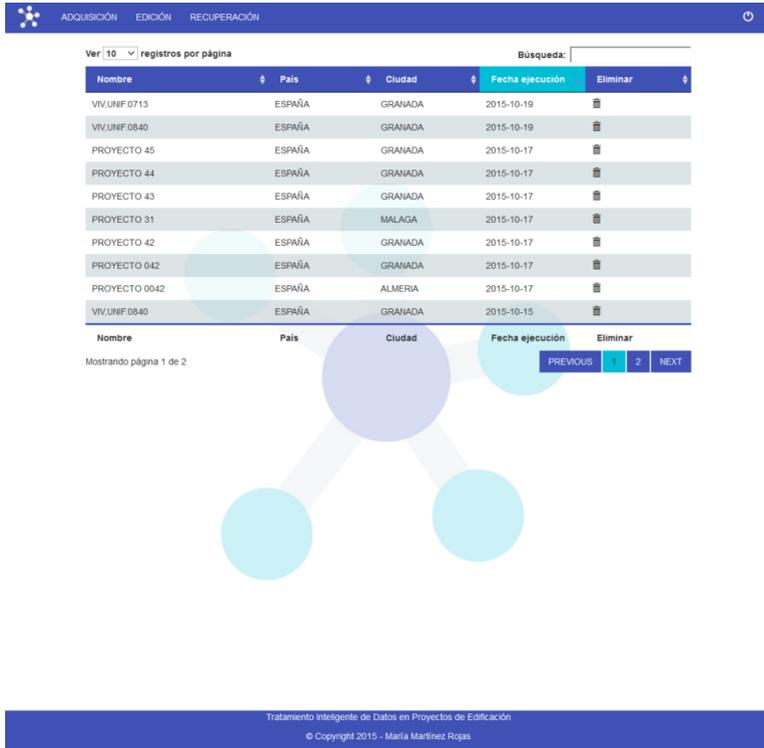
© Copyright 2015 - María Martínez Rojas

Figura 20: Pantalla con menú desplegable para modificar el resultado

La mayoría de las aplicaciones que se utilizan hoy en día para la gestión de los proyectos no son accesibles a través de la web, por lo que los usuarios deben desplazarse hasta la oficina para poder consultar los datos. En el caso de obras de pequeña envergadura, en las que no se suele instalar una oficina en la propia obra, el acceso a los datos se demora hasta que el usuario vuelva a la oficina. Estos inconvenientes generan retrasos en la ejecución de los proyectos, afectando al éxito del proyecto.

En este prototipo se permite un acceso rápido a datos estructurados mediante herramientas de filtrado, búsqueda y consulta sencillas, que permiten resolver consultas como el precio de actividades para realizar una estimación de costes o la información del presupuesto desde la misma obra.

En la Figura 22 se puede ver la pantalla principal del módulo de consulta, en la que permite al usuario realizar consultas sobre los datos almacenados en el repositorio común de proyectos de edificación.



Nombre	País	Ciudad	Fecha ejecución	Eliminar
VIV.LINF.0713	ESPAÑA	GRANADA	2015-10-19	
VIV.LINF.0840	ESPAÑA	GRANADA	2015-10-19	
PROYECTO 45	ESPAÑA	GRANADA	2015-10-17	
PROYECTO 44	ESPAÑA	GRANADA	2015-10-17	
PROYECTO 43	ESPAÑA	GRANADA	2015-10-17	
PROYECTO 31	ESPAÑA	MALAGA	2015-10-17	
PROYECTO 42	ESPAÑA	GRANADA	2015-10-17	
PROYECTO 042	ESPAÑA	GRANADA	2015-10-17	
PROYECTO 0042	ESPAÑA	ALMERIA	2015-10-17	
VIV.LINF.0840	ESPAÑA	GRANADA	2015-10-15	

Mostrando página 1 de 2

PREVIOUS 2 NEXT

Tratamiento Inteligente de Datos en Proyectos de Edificación
© Copyright 2015 - María Martínez Rojas

Figura 21: Pantalla con la lista de proyectos almacenados

Un acceso más sofisticado a los datos se plantea en la aplicación para el análisis de costes que mostramos más adelante en este capítulo (sección 5.3). Por ejemplo, consultas que involucren datos asociados al tiempo, tipo de proyecto, tamaño de empresa, etc. Este tipo de consultas requieren el manejo de grandes cantidades de datos y pueden tomar ventaja de mecanismos de consulta flexible. Como veremos más adelante en este capítulo, para ello proponemos el uso de un modelo multidimensional con relaciones de parentesco difusas.

5.2.4 Implementación de la aplicación SIGPE

En esta sección, en primer lugar, se detalla la arquitectura de la aplicación para la adquisición y gestión de los datos contenidos en el presupuesto. A continuación, se presentan las tecnologías empleadas en el desarrollo de la misma.

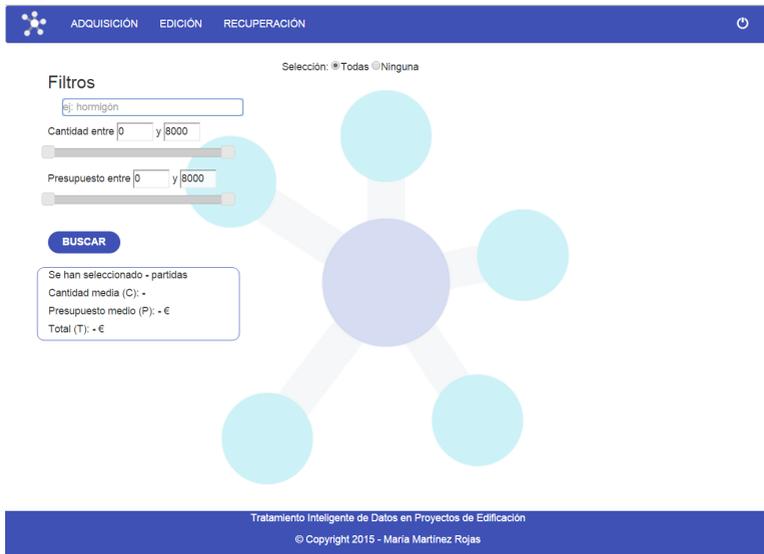


Figura 22: Pantalla de consulta

5.2.4.1 Arquitectura del sistema on-line

La aplicación se ha desarrollado siguiendo la arquitectura que se ilustra en la Figura 23. La arquitectura cliente/servidor de dos capas incluye los siguientes elementos: interfaz de usuario, un servidor y un servidor web que permite la comunicación entre los dos primeros.

A continuación se detallan las distintas tecnologías utilizadas en el servidor y en el cliente.

5.2.4.1.1 Servidor

En la capa del servidor se han utilizado las siguientes tecnologías para dar soporte a las distintas funcionalidades:

- PHP [200]. Se utiliza para el procesamiento de datos y comunicación con la base de datos, además, nos sirve para la lectura y escritura de ficheros.
- HTML5 [239]. Se utiliza para estructurar toda la información que se muestra a los usuarios.
- CSS [238]. Se utiliza para proporcionar formato al estilo de la página y hacer amigable la interfaz de usuario.

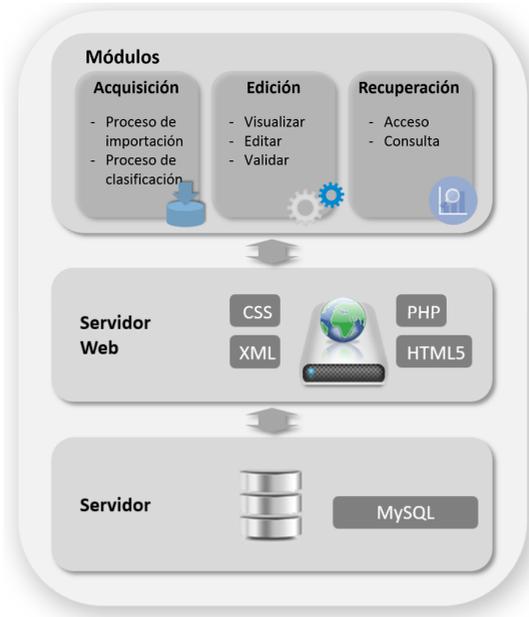


Figura 23: La arquitectura cliente-servidor de la aplicación *SIGPE*

- XML [80]. Se utiliza para la configuración de los esquemas.

Para gestionar los datos de los proyectos, se ha utilizado un sistema de gestión de base de datos relacional (MySQL) [86]. Esta base de datos también gestiona los datos recuperados de los documentos.

5.2.4.1.2 Cliente

El cliente del sistema ha sido implementado para que no se requieran grandes recursos para su ejecución, de modo que el usuario sólo necesite tener instalado un navegador con acceso a la red donde se encuentre el servidor. La interfaz permite al usuario acceder a las distintas funcionalidades a través de menús.

Las tecnologías utilizadas para el sistema basado en la web son software de código abierto y están disponibles a través de Internet. El sistema es compatible con los principales navegadores: Internet Explorer 9 o superior, Mozilla Firefox 35 o superior y Chrome 40 o superiores. Para la comunicación con el servidor, se ha realizado JavaScript [110] y AJAX [241], de esta forma se han realizado llamadas

asíncronas al servidor para obtener los datos necesarios. El navegador debe soportar esta tecnología para poder ejecutar la aplicación.

5.2.5 *Evaluación de la aplicación*

Para evaluar la aplicación se propone un test de dos vías. En primer lugar, analizando la interacción de usuarios reales y, en segundo lugar, mediante un cuestionario específico sobre la aplicación. Para alcanzar estos objetivos, se ha contado con un grupo de expertos relacionados con la gestión de proyectos de construcción.

A continuación, se detallan los pasos requeridos a los usuarios con el fin de evaluar correctamente la herramienta:

- **Paso 1:** Los expertos acceden a la aplicación mediante un usuario y clave personal.
- **Paso 2:** Cada usuario importa cinco documentos de presupuesto de cinco proyectos diferentes mediante el módulo de adquisición que se ha definido en este capítulo. Para cada uno de los proyectos, el usuario tiene que proporcionar la información requerida para identificar el proyecto que está cargando en la aplicación.
- **Paso 3:** Una vez que los presupuestos han sido importados y clasificados, los usuarios validan o editan los resultados de clasificación proporcionados por la aplicación, a través del módulo de edición. Durante la ejecución de este paso, el sistema almacena todos los cambios que el usuario realiza para cada proyecto, con el fin de valorar la bondad de los resultados proporcionados por la aplicación y los proporcionados por el usuario.
- **Paso 4:** Finalmente, los usuarios acceden y consultan datos a través del módulo de recuperación.

5.2.5.1 **Resultados**

Mediante la interacción de los usuarios con la aplicación según el procedimiento descrito en la sección anterior, podemos evaluar de forma alternativa a la realizada en el capítulo 4 la bondad del proceso de clasificación propuesto.

La Tabla 19 muestra los resultados detallados correspondientes a cada uno de los proyectos que los usuarios expertos han insertado en

la aplicación. Para cada proyecto se detalla el número de partidas y el porcentaje de acierto obtenido tanto en capítulo como en subcapítulo.

El porcentaje de acierto medio para el nivel de capítulo es del 92,4 %, mientras que para el nivel de subcapítulo es del 91,09 %. El porcentaje de subcapítulo se calcula sobre el porcentaje de descripciones de partidas que se han clasificado correctamente en el nivel de capítulo, es decir, del 92,4 %, el 91,09 % se ha clasificado correctamente en el subcapítulo correspondiente. La desviación típica para el nivel de capítulo es de 3,04, mientras que para el nivel de subcapítulo es de 4,87.

En la experimentación realizada en el capítulo 4, los resultados obtenidos para el conjunto de prueba correspondiente a los proyectos reales son similares a los obtenidos en esta evaluación para el nivel de capítulo. Sin embargo, los resultados para el nivel de subcapítulo son sensiblemente mejores a los obtenidos en el capítulo anterior. En la experimentación de los usuarios, la totalidad de los proyectos insertados por los usuarios corresponden a proyectos de edificación para uso residencial.

Proyecto	Nº WD	% Acierto Capítulo	% Acierto Subcapítulo
1	139	95,68	84,96
2	139	94,24	91,60
3	154	94,81	86,30
4	421	94,54	89,45
5	624	90,54	84,07
6	132	93,94	89,52
7	384	94,01	90,03
8	215	97,21	90,43
9	190	89,47	89,41
10	204	91,18	95,16
11	88	89,77	75,95
12	99	91,92	92,31
13	37	97,30	88,89
14	152	90,13	81,75
15	187	96,79	89,50
16	193	97,41	89,89
17	272	87,50	89,08
18	419	97,85	91,71
19	121	95,87	83,62

Continúa en la página siguiente.

Proyecto	Nº WD	% Acierto Capítulo	% Acierto Subcapítulo
20	268	88,06	92,37
21	211	91,94	94,33
22	158	90,51	96,50
23	187	93,05	90,23
24	150	90,67	92,65
25	135	93,33	98,41
26	51	88,24	93,33
27	182	89,01	91,98
28	162	89,51	95,86
29	125	89,60	93,75
30	95	90,53	89,53
31	201	87,56	85,80
32	183	94,54	91,33
33	510	90,39	86,98
34	210	95,71	92,04
35	935	88,56	87,92
36	138	94,93	98,47
37	97	96,91	95,74
38	81	90,12	93,15
39	94	88,30	97,59
40	98	92,86	97,80
41	129	94,57	95,90
42	133	96,24	89,06
43	99	89,90	96,63
44	137	96,35	90,91
45	295	95,25	98,93
46	257	93,77	97,93
47	137	95,62	92,37
48	176	94,32	95,78
49	252	95,63	98,76
50	264	92,80	99,59

Tabla 19: Resultados obtenidos para el nivel de capítulo y subcapítulo mediante la interacción de los usuarios con la aplicación SIGPE

Se ha realizado un test global de diferencia de medias. El intervalo de confianza para la media al 95 % establece el límite inferior en 91.4 y

el límite superior al 93.5, es decir, el porcentaje de acierto se encuentra dentro de dicho intervalo.

Se ha realizado el análisis de varianza de un factor por capítulos y el valor de significancia obtenido es de 0.000. Esto quiere decir que efectivamente aceptamos que hay diferencia estadísticamente significativa entre los resultados, es decir, decidimos rechazar la hipótesis de igualdad de medias y concluimos que hay diferencias entre los capítulos.

Con el fin de conocer entre qué grupos se ha producido la diferencia se realiza el test de Scheffé, que permite encontrar diferencias estadísticamente significativas para las combinaciones. Los resultados de dicho test muestran dos subconjuntos homogéneos; el primero se compone de todos los capítulos con un valor de significancia de 0.175, y el segundo corresponde al capítulo 1 con un valor de significación de 1 ya que, evidentemente no presenta diferencias consigo mismo. La razón es que el número de descripciones de partida en este capítulo es considerablemente menor al resto de capítulos.

Asimismo, se ha realizado el análisis de varianza de un factor por proyectos. El valor de significancia obtenido es de 0.304. Esto quiere decir que no hay diferencia estadísticamente significativa entre los resultados de los proyectos.

5.2.5.2 Cuestionario

El segundo objetivo que se pretende con la evaluación de la aplicación es conocer el nivel de satisfacción de los usuarios en relación con su usabilidad y utilidad de cada uno de los módulos después de utilizar la aplicación *SIGPE*.

Para este propósito, los usuarios deben completar el cuestionario que se les proporciona (Tabla 20). Éste cuenta con un total de 16 preguntas divididas en cuatro apartados. Los tres primeros corresponden a preguntas sobre los distintos módulos de la aplicación y la última en relación a la aplicación en general.

Para cada una de las preguntas los usuarios tienen que marcar la casilla que mejor represente su evaluación en una escala de 1 a 5, en la que el 1 corresponde a una respuesta de *Nada adecuado* y 5 corresponde a una respuesta *Totalmente adecuado*.

En la Tabla 21 se muestran los resultados del cuestionario, la media aritmética obtenida (μ) y la desviación típica (σ) de las puntuaciones obtenidas por los usuarios expertos.

Adquisición de los datos					
1	Facilidad para importar un presupuesto	1	2	3	4 5
2	Tiempo empleado en el proceso de adquisición	1	2	3	4 5
3	Sencillez de interacción con la interfaz de adquisición de datos	1	2	3	4 5
4	Satisfacción general del módulo de adquisición	1	2	3	4 5
Edición de los datos					
5	Facilidad para editar los resultados proporcionados por el sistema	1	2	3	4 5
6	Utilidad de la escala de colores	1	2	3	4 5
7	Facilidad para acceder a los proyectos almacenados	1	2	3	4 5
8	Sencillez de interacción con la interfaz de edición de datos	1	2	3	4 5
9	Satisfacción general del módulo de edición	1	2	3	4 5
Consulta de los datos					
10	Facilidad para consultar los datos almacenados	1	2	3	4 5
11	Utilidad de los datos consultados en la fase de estimación de costes	1	2	3	4 5
12	Tiempo de respuesta para resolver la consulta	1	2	3	4 5
13	Sencillez de interacción con la interfaz de recuperación de datos	1	2	3	4 5
14	Satisfacción general del módulo de consulta	1	2	3	4 5
Uso General					
15	Amigabilidad de la interfaz	1	2	3	4 5
16	Facilidad de uso	1	2	3	4 5

Tabla 20: Cuestionario sobre la aplicación SIGPE

		μ	σ
Adquisición de los datos			
1	Facilidad para importar un presupuesto	4.3	0.67
2	Tiempo empleado en el proceso de adquisición	3.6	0.70
3	Sencillez de interacción con la interfaz de adquisición de datos	4.4	0.84
4	Satisfacción general del módulo de adquisición	4.1	0.74
Edición de los datos			
5	Facilidad para editar los resultados proporcionados por el sistema	4.0	0.94
6	Utilidad de la escala de colores	4.8	0.42
7	Facilidad para acceder a los proyectos almacenados	3.8	0.79
8	Sencillez de interacción con la interfaz de edición de datos	4.1	0.74
9	Satisfacción general del módulo de edición	4.1	0.74
Consulta de los datos			
10	Facilidad para consultar los datos almacenados	3.3	0.82
11	Utilidad de los datos consultados en la fase de estimación de costes	4.8	0.42
12	Tiempo de respuesta para resolver la consulta	3.8	0.63
13	Sencillez de interacción con la interfaz de recuperación de datos	4.2	0.79
14	Satisfacción general del módulo de consulta	3.8	0.63
Uso General			
15	Amigabilidad de la interfaz	4.1	0.74
16	Facilidad de uso	4.0	0.67

Tabla 21: Media aritmética (μ) y desviación típica (σ) de las puntuaciones obtenidas por los usuarios expertos sobre el cuestionario de la Tabla 20.

En general, los resultados obtenidos de las encuestas realizadas por los expertos sobre la usabilidad de la aplicación *SIGPE* son satisfactorios (una valoración media alrededor de 4 sobre 5). A continuación se detallan los resultados para cada uno de los módulos:

- En particular, para el módulo de adquisición de datos, los expertos muestran una tendencia satisfactoria en cuanto a la facilidad y sencillez para adquirir datos de proyectos (valoración media entorno a 4.2 sobre 5). Sin embargo, el tiempo empleado en el proceso de adquisición, es el que los expertos consideran peor valorado en el cuestionario (valoración media de 3.6 sobre 5).
- Respecto al módulo de edición de datos, la tendencia de facilidad para editar resultados y facilidad para acceder a proyectos almacenados, así como amigabilidad de la interfaz está valorada alrededor de 4. Es destacable la valoración obtenida respecto a la utilidad de la escala de colores, la cual los expertos están realmente satisfechos con esta utilidad (valoración media de 4.8 sobre 5).
- Finalmente, respecto a la consulta de datos, la tendencia de satisfacción que muestra este módulo es positiva (valoración media alrededor de 4), destacando la utilidad de los datos consultados en la fase de estimación de costes (valoración media de 4.8 sobre 5). Este último dato pone de manifiesto uno de los problemas abordados en esta Tesis, la heterogeneidad de los datos, la cual se reduce con el modelo propuesto, siendo muy útil la consulta sobre dichos datos de forma integral.

5.3 UN HERRAMIENTA FLEXIBLE PARA EL ANÁLISIS DE COSTES

Durante todo el ciclo de vida del proyecto, los usuarios necesitan acceder de manera rápida y flexible a la información contenida en el documento del presupuesto con el fin de dar soporte a otras tareas del ámbito de la gestión de proyectos. Por ejemplo, en las primeras etapas, para dar soporte al proceso de estimación, y durante la ejecución de la obra, para facilitar el control de los costes. En la práctica actual, el acceso a esta información requiere gran cantidad de tiempo porque la información está distribuida en diversos sistemas.

Gracias a los objetivos alcanzados en los capítulos anteriores, los datos contenidos en el documento del presupuesto se almacenan de manera centralizada en una estructura de referencia, permitiendo así

su acceso integrado. Sin embargo, no sólo el manejo de los datos es necesario, sino que la extracción de información a partir de grandes conjuntos de datos se está convirtiendo en una necesidad cada vez más importante en cualquier entorno empresarial.

En esta sección, se aborda esta necesidad mediante la herramienta Linguistic F-Cube Factory [47], que posibilita la recuperación de información novedosa que se encuentra oculta entre la gran cantidad de datos almacenados. Esta herramienta facilita al usuario las tareas de creación y gestión de cubos de datos de forma transparente. Además, permite realizar análisis incorporando conceptos de la Teoría de Conjuntos Difusos [251] que facilita la consulta flexible de los datos. A la hora de afrontar esta necesidad, consideramos que es de especial interés reproducir en el mundo virtual lo que sucede en el proceso de comunicación entre las personas. Los seres humanos por naturaleza tenemos una capacidad muy alta para trabajar con información imprecisa y, por tanto, es importante que el modelo sea capaz de gestionar la imperfección de los datos de manera adecuada con el fin de proporcionar a los usuarios la información de manera más intuitiva para que puedan tomar decisiones más correctas.

En el ámbito de la lógica difusa, los conceptos imprecisos se pueden representar mediante variables lingüísticas [252]. Este concepto juega un rol esencial en las aplicaciones de lógica difusa debido a que asemeja la forma en la cual los humanos perciben, razonan y se comunican. El uso de palabras puede ser visto como una forma de síntesis de datos numéricos que explota la tolerancia a la imprecisión para cumplir con los principios de la Computación Flexible [253].

En esta sección, en primer lugar, se presentan los principales conceptos de la estructura multidimensional basada en el modelo presentado por Molina et al. [77, 178], para luego presentar la propuesta de modelo que permite el análisis de costes en base a dicho modelo. Finalmente, se presentan consultas que se pueden realizar con dicha propuesta.

5.3.1 *Estructura multidimensional sobre datos de costes*

En esta sección se presenta nuestra propuesta de estructura multidimensional basada en el modelo presentado por Molina et al. [77, 178]. Este modelo consta de dimensiones en las que los miembros de las jerarquías se definen a través de una relación de parentesco difusa. El uso de este tipo de jerarquías permite el modelado de las dimensiones en base a conceptos difusos.

Antes de definir las dimensiones y las medidas que proporcionan el contexto para el análisis de los costes en el ámbito de la gestión de proyectos de edificación, se van a definir los principales conceptos del modelo.

5.3.1.1 Conceptos previos modelo multidimensional difuso

Comenzaremos con la definición de la estructura de las dimensiones y las jerarquías que podemos definir en ellas. Tras éstas, definiremos qué serán los hechos para concluir con la estructura que tendrán los cubos de datos acorde a la propuesta de estructura multidimensional de Molina et al [77, 178].

Definición 15 Una dimensión sobre el dominio es una tupla $d = (l, \leq_d, l_{\perp}, l_{\top})$ donde $l = \{l_1, \dots, l_m\}$ tal que cada l_i es un conjunto $l_i = \{c_{i1}, \dots, c_{im_i}\}$ tal que $l_i \cap l_j = \emptyset$ si $i \neq j$, \leq_d es una relación de orden parcial tal que $l_i \leq_d l_k$ si $\forall c_{ij} \in l_i \Rightarrow \exists c_{kp} \in l_k / c_{ij} \subseteq c_{kp}$. l_{\perp} y l_{\top} son dos elementos de l tal que $\forall l_i \in l, l_{\perp} \leq_d l_i$ y $l_i \leq_d l_{\top}$.

A cada elemento l_i se le denominaremos nivel. Los niveles especiales l_{\perp} y l_{\top} son denominados *nivel base* y *nivel superior* respectivamente. La relación de orden parcial \leq en una dimensión es la que establece la jerárquica entre los niveles. Un nivel l_i podrá tener niveles inferiores que estén directamente conectados en la jerarquía. Es decir, niveles menores que l_i y que no exista otro nivel menor que l_i y que a su vez mayor que estos. El conjunto de los niveles que cumplan esta condición para un nivel, lo denominaremos conjunto de hijos del nivel.

Definición 16 Para cada l_i , el conjunto

$$H_i = \{l_j / l_j \neq l_i \wedge l_j \leq_d l_i \wedge \neg \exists l_k l_j \leq_d l_k \leq_d l_i\} \quad (5.1)$$

Se denominará conjunto de hijos del nivel l_i .

En cualquier dimensión, dada la definición que hemos presentado del conjunto de hijos, tendríamos que para el nivel base su conjunto de hijos sería siempre el conjunto vacío. En el caso de las jerarquías difusas, un elemento puede estar relacionado con más de un elemento del nivel superior y el grado de esta relación toma valores en el intervalo $[0,1]$. Para establecer este grado de pertenencia, utilizaremos la relación de parentesco.

Definición 17 Dado l_i

$$P_i = \{l_j / l_i \neq l_j \wedge l_i \leq_d l_j \wedge \neg \exists l_k l_j \leq_d l_k \leq_d l_i\} \quad (5.2)$$

y se denominará conjunto de padres del nivel l_i .

En el caso del nivel superior de la jerarquía estaríamos en una situación análoga al conjunto de hijos para el nivel base. Es decir, el conjunto de padres será el conjunto vacío.

Definición 18 Para cada par de niveles l_i y l_j tal que $l_j \in H_i$, tendremos la relación

$$\mu_{ij} : l_i \times l_j \rightarrow [0, 1] \quad (5.3)$$

A esta relación la denominaremos relación de parentesco

Mediante esta relación podremos definir el grado de inclusión de los elementos de un nivel en sus niveles padre. Si esta función sólo toma como posibles valores 0 ó 1 y cada elemento de un nivel sólo toma este segundo para un valor de un nivel padre, nos encontramos en el caso de una relación de jerarquía precisa entre los dos niveles. Este concepto de jerarquía se puede suavizar si permitimos que la función entre los niveles tome valores en todo el intervalo $[0,1]$ y que se puedan establecer más de un valor mayor que cero para un valor con diferentes valores de un nivel padre. En este caso, estaríamos hablando de una relación de jerarquía difusa entre los niveles. Mediante esta relación podemos manejar la imprecisión entre los elementos en las jerarquías.

Las variables del dominio que queremos analizar definirán los hechos del cubo de datos. Éstos serán los que en mayor medida acotarán el dominio de análisis del cubo de datos que definamos, limitando qué medidas podemos obtener.

Definición 19 Considerando un conjunto de atributos A_1, \dots, A_n con dominios D_1, \dots, D_n , denominaremos hecho a cualquier par (h, α) , donde $h = (x_1, \dots, x_n)$ con $x_i \in D_i \forall i = 1, \dots, n$, es decir, cualquier n -tupla definida sobre los dominios de los atributos que interesa estudiar, y $\alpha \in [0, 1]$.

El valor α sería el que controlaría la influencia del hecho en el posterior análisis. Cuanto más cercano a 1 sea este valor, mayor será la influencia.

La estructura de un cubo de datos difuso se define como sigue:

Definición 20 *Un cubo de datos es una tupla $C = (D, l_b, F, A, H)$ tal que $D = (d_1, \dots, d_n)$ es un conjunto de dimensiones, $l_b = (l_{1b}, \dots, l_{nb})$ es un conjunto de niveles tal que l_{ib} pertenece a d_i , $F = R \cup \emptyset$ donde R es el conjunto de hechos y \emptyset es un símbolo especial, A es una aplicación definida como $A : l_{1b} \times \dots \times l_{nb} \rightarrow F$, que para cada conjunto de valores de las dimensiones devuelve el hecho relacionado con estas coordenadas, y H es un objeto de tipo historia.*

F-Cube Factory implementa las operaciones OLAP usuales: *slice* and *dice* relacionada con la selección en los datos en un cubo, *roll-up* y *drill-down* que se encargan de cambiar la granularidad en el cubo navegando a través de los distintos niveles de abstracción de las jerarquías en las dimensiones, y la operación *pivot* para obtener representaciones alternativas de los datos. Cuando se trata con datos numéricos, se dispone de las funciones de agregación más comunes como el máximo y el mínimo, la media o la suma, entre otras. Todas estas operaciones se encuentran disponibles también para trabajar con hechos y dimensiones difusas.

Una vez definidos los conceptos de la estructura multidimensional difusa, en la siguiente sección presentamos nuestra propuesta de estructura multidimensional para el análisis de los costes en el ámbito de los proyectos de edificación.

5.3.1.2 Propuesta de estructura multidimensional

En esta sección se presenta la estructura del modelo multidimensional propuesto y las dimensiones que proporcionan el contexto para el análisis de los costes desde diversas perspectivas. Como se puede observar en la Figura 24, el modelo cuenta con un total de siete dimensiones: *proyecto*, *actividad*, *empresa*, *tiempo*, *localización*, *promotor* y *tipología de construcción*. A continuación, se detallan cada una de estas dimensiones.

5.3.1.2.1 Dimensiones

Como se ha comentado, las dimensiones establecen el contexto del análisis sobre los hechos del cubo y permiten la exploración de los datos desde varias perspectivas. Para acceder a los hechos con diferentes niveles de detalle, se pueden definir jerarquías sobre las dimensiones, estableciendo los distintos niveles de granularidad.

Como veremos, algunos conceptos de los miembros de las jerarquías resulta de especial interés modelarlos mediante etiquetas lingüísticas,

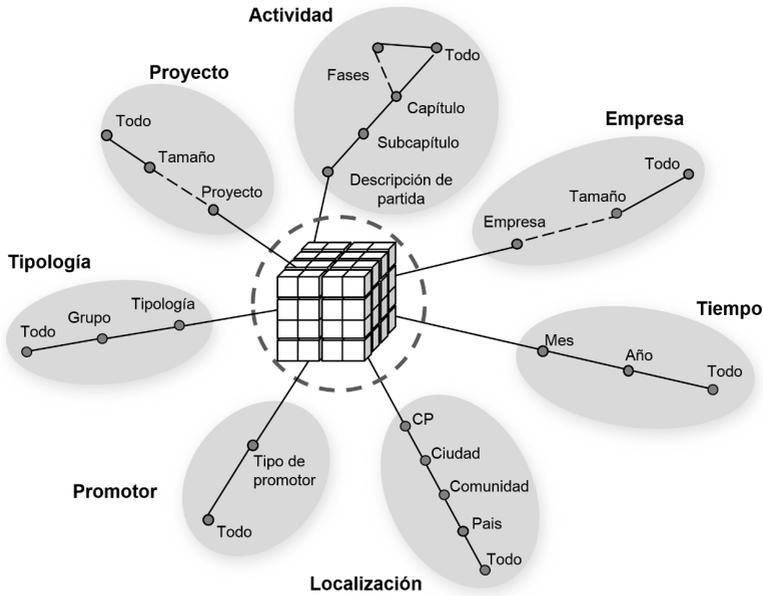


Figura 24: Cubo para el análisis de costes

ya que permiten manejar estos conceptos de manera intuitiva y más cercana a la manera en la que los usuarios utilizan esos conceptos. En el modelo propuesto, las etiquetas lingüísticas y las funciones de pertenencia correspondientes se han utilizado para modelar la relación de parentesco de algunas dimensiones del cubo.

1. Proyecto

El cubo propuesto se ha diseñado para ser utilizado en el entorno de una empresa constructora que desea analizar los datos de costes para los distintos proyectos que ha ejecutado, por lo tanto, una de las variables será el *proyecto*. En el nivel base de esta dimensión se encuentran cada uno de los proyectos almacenados. Sobre este nivel, se define otro que clasifica los proyectos según su tamaño: pequeños, medianos y grandes. La clasificación de estos grupos se realiza a partir de la cuantía del presupuesto del proyecto.

En esta dimensión, el uso de etiquetas lingüísticas resulta interesante, ya que los usuarios definen el tamaño del proyecto en base a estos términos en lugar de por la cantidad del presupuesto de la obra. En la Figura 25 se puede observar un ejemplo de función de pertenencia de las etiquetas que definen el "ta-

maño del proyecto". La función se puede ajustar en base a las preferencias de los usuarios.

La estructura de la dimensión es la siguiente:

$Proyecto = (\{Proyecto, Tamaño, Todo\}, \leq_{Proyecto}, Proyecto, Todo)$.

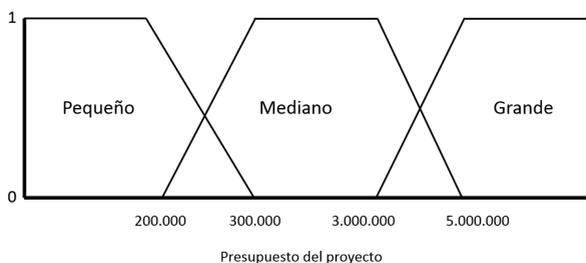


Figura 25: Valores y agrupación del nivel *tamaño* en la dimensión *proyecto*

2. Tiempo

La dimensión *tiempo* es fundamental para situar los proyectos según la fecha en la que se ejecutaron y tener una perspectiva histórica de los datos. Por ejemplo, los datos históricos resultan muy útiles para analizar la tendencia de los costes de las actividades, ya que uno de los factores que afectan a los costes de construcción es la fluctuación de los precios a lo largo del tiempo debido a la inflación de materiales, equipos y costes laborales [175].

Como se puede observar en la Figura 24, en el nivel de la base de la jerarquía se encuentra el mes de comienzo del proyecto y sobre este nivel se encuentra el año. Con la definición de la dimensión *tiempo* en el cubo se pueden resolver consultas, tales como, "evolución del coste de una cierta actividad en la última década". La estructura de la dimensión es la siguiente:

$Tiempo = (\{Mes, Año, Todo\}, \leq_{Tiempo}, Mes, Todo)$.

3. Localización

Otra característica que modelamos es la localización, ya que el coste de las actividades depende de la localización geográfica donde se ejecuta el proyecto. Por ejemplo, el coste de una actividad para un proyecto situado en el centro de una ciudad suele ser más cara que en una zona más retirada, debido a: salarios más elevados, costes asociados a las restricciones de acceso, espacio limitado para instalaciones y almacenamiento

de materiales, así como medidas de seguridad adicionales [73]. Esta dimensión permite a los usuarios analizar la información según la localización del proyecto, por ejemplo, “Coste medio de una cierta actividad en proyectos realizados en Granada en la última década”. Como se puede ver en la Figura 24 la jerarquía de la dimensión es la siguiente: código postal, ciudad, localidad y país. Por tanto, la estructura de la dimensión es la siguiente:

Localización = ({*Ciudad, Localidad, Pais, Todo*}, $\leq_{\text{Localización}}$, *Ciudad, Todo*).

4. Promotor

Según la LOE [153], el promotor es: “cualquier persona física o jurídica, pública o privada que, individual o colectivamente, decide, impulsa, programa y financia, con recursos propios o ajenos, las obras de edificación para sí o para su posterior enajenación, entrega o cesión a terceros bajo cualquier título”. Para el análisis de costes desde una perspectiva más general, puede resultar interesante el análisis según la tipología del promotor, por ejemplo, analizar si el hecho de que el promotor sea privado o público afecta a los costes del proyecto. La estructura de la dimensión es la siguiente:

Promotor = ({*Tipo de Promotor, Todo*}, \leq_{Promotor} , *Tipo de Promotor, Todo*).

5. Tipología

EL coste de un proyecto depende de las características técnicas y del objetivo de éste [73]. Aunque cada proyecto es único, las actividades que se tienen que realizar para llevar a cabo el proyecto pueden ser comunes incluso para aquellos proyectos que el diseño es totalmente diferente. Sin embargo, algunas características pueden producir una variación en el precio de la actividad, como por ejemplo, la cantidad que se tiene que ejecutar. Por ejemplo, el coste del m^3 de hormigón para la construcción de una vivienda unifamiliar no será el mismo que para la construcción de un hospital que requerirá muchos más m^3 . Seguramente, los responsables de la ejecución del proyecto del hospital conseguirán un precio más económico para esta unidad de obra.

En este sentido, parece interesante distinguir por tipologías de obras. Como se puede observar en la Figura 24, sobre el nivel de la base hemos definido el nivel *grupo* que clasifica las tipologías de los proyectos según su uso: residencial, industrial y servicios. De esta forma, los usuarios finales podrán consultar los datos

de manera más específica, como por ejemplo, “beneficio medio, por capítulos, dependiendo de la tipología de construcción”. La estructura de la dimensión es la siguiente:

$Tipología = (\{Tipología, Grupo, Todo\}, \leq_{Tipología}, Tipología, Todo)$.

A continuación, se pueden observar los valores de los miembros de esta jerarquía:

- Tipología: unifamiliar, adosada, pareada, bloque, nave, administrativo, comercial, hospital, docente etc.
- Grupo: residencial, industrial y servicios

6. Empresa

En el sector de la construcción es muy frecuente que la empresa constructora responsable de la ejecución del proyecto no realice todas las actividades del proyecto y que subcontrate estas actividades a empresas más especializadas. En esta dimensión se modela la información relativa a estas empresas. En el nivel base se encuentran las distintas empresas y, en el nivel superior, se establece una clasificación según el tamaño de éstas. Para establecer dicha clasificación, se utilizan etiquetas lingüísticas que se ajustan mejor a la forma en la que los usuarios describen esta característica de las empresas. En la Figura 26 se puede observar un ejemplo donde se han definido cuatro etiquetas en función del número de empleados de la empresa: micro, pequeña, mediana y grande. Al igual que en los casos anteriores, las etiquetas se pueden ajustar según las preferencias de los usuarios. En la práctica actual, la clasificación del tamaño de la empresa se realiza con intervalos precisos (crisp). Esta dimensión proporciona información que puede ser muy útil en el proceso de contratación de empresas para realizar las actividades del proyecto. Por ejemplo, “beneficio de un cierto capítulo según el tamaño de la empresa que ha realizado las actividades de ese capítulo en proyectos de viviendas unifamiliares en la provincia de Granada”. La estructura de la dimensión es la siguiente:

$Empresa = (\{Empresa, Tamaño, Todo\}, \leq_{Empresa}, Empresa, Todo)$.

7. Actividad

Como se ha comentado en varias ocasiones en esta memoria, el proyecto de construcción se divide en actividades siguiendo el concepto de estructura de descomposición de trabajo. Cada nivel inferior representa información más detallada de las actividades del proyecto. Por lo tanto, el nivel base de esta dimensión

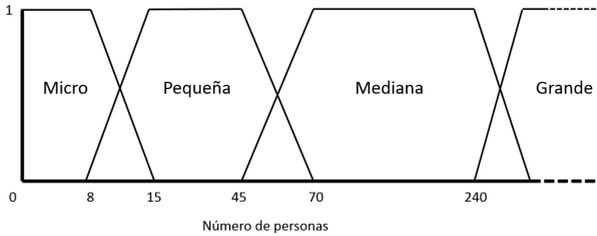


Figura 26: Valores y agrupación del nivel *tamaño* en la dimensión *empresa*

son las descripciones de partida. Sobre este nivel, se definen los niveles correspondientes a los subcapítulos y capítulos. De esta forma, el usuario puede resolver consultas, tales como, “Beneficio según los capítulos del proyecto para proyectos de viviendas unifamiliares de tamaño medio”.

La definición de la jerarquía de esta dimensión es posible gracias a, por un lado, la propuesta de formalización del documento del presupuesto para establecer los niveles de la jerarquía y, por otro lado, el mecanismo de clasificación que permite nutrir los distintos niveles de la estructura de una manera integrada.

Un ejemplo de la jerarquía de la dimensión de actividades es la siguiente:

- WD: Excavación, en apertura de caja, en terreno blando, realizada con medios mecánicos, incluso perfil hasta una profundidad máxima de 50 cm. Medida en perfil natural
- Subcapítulo: Excavación
- Capítulo: Preparación del terreno

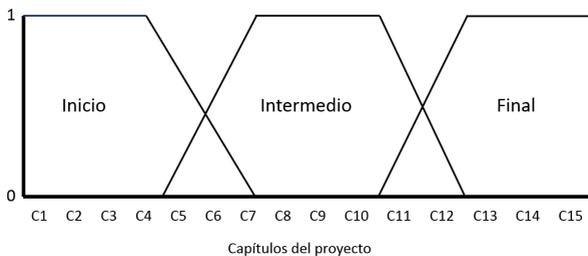


Figura 27: Valores y agrupación del nivel *fases* en la dimensión *actividad*

En esta dimensión se define un nivel adicional, el de las *fases de la obra*, que se define utilizando las siguientes etiquetas lingüísticas:

inicio, intermedio y final. Aunque la relación de parentesco entre los niveles de la jerarquía son precisos, es muy frecuente en el ámbito de la construcción dividir el proceso de ejecución del proyecto en estas tres fases. En la Figura 27 se muestran un ejemplo de las funciones de pertenencia de las *Fases* de la obra y en la Tabla 22 se muestran los valores que establecen la relación de parentesco entre el nivel de capítulo y las fases de la obra en la jerarquía de la dimensión.

Capítulos	Fases		
	Inicio	Intermedio	Final
C1	1	0	0
C2	1	0	0
C3	1	0	0
C4	1	0	0
C5	0.6	0.4	0
C6	0.4	0.6	0
C7	0	1	0
C8	0	1	0
C9	0	1	0
C10	0	0.8	0.2
C11	0	0.3	0.7
C12	0	0.1	0.9
C13	0	0	1
C14	0	0	1
C15	0	0	1

Tabla 22: Relación de parentesco de las fases en la dimensión Actividad

La estructura de la dimensión es la siguiente:

$Actividad = (\{Descripción de Partida, Subcapítulo, Capítulo, Fases, Todo\}, \leq_{WBS}, Descripción de partida, Todo).$

5.3.1.2.2 Medidas

En el modelo de datos multidimensional, como se ha comentado, las medidas o los hechos son el objeto del análisis. La mayoría de los hechos son numéricos, continuamente valuados y aditivos. La razón de que sean así es la siguiente: las consultas a la tabla de hechos necesitarán a su vez consultas de cientos, miles o incluso millones de registros para construir el conjunto respuesta.

En la Tabla 23 se muestran algunas medidas de ejemplo que se pueden analizar con la estructura multidimensional propuesta y que son de interés en el ámbito de la gestión de costes.

Medida	Acrónimo	Descripción
Coste estimado	CE	El coste estimado corresponde con el coste asignado en el presupuesto del proyecto.
Coste de ejecución	CEX	El coste de ejecución corresponde con el coste real de la ejecución de la actividad.
Desviación de coste	DC	La desviación de coste es la diferencia entre el coste estimado y el coste ejecutado.
Beneficio	B	El beneficio es la diferencia entre el coste total según proyecto y el coste de lo realmente ejecutado.

Tabla 23: Medidas sobre el cubo de costes

5.3.1.2.3 Cubo de datos

Una vez que se han definido las estructuras para las distintas dimensiones y los hechos, se puede establecer la estructura del cubo de datos:

$$C_{\text{Coste}} = \{\text{Tiempo, Proyecto, Promotor, Tipología, Localización, Actividad, Empresa}\}, \text{Coste estimado, Coste de ejecución, desviación de coste, Beneficio} \cup \emptyset, \Omega, A$$

donde A es la relación que asocia cada hecho con los correspondientes valores de la base de cada una de las dimensiones.

A continuación se presentan algunos ejemplos de consultas que se pueden realizar con la estructura multidimensional presentada en esta sección.

5.3.1.3 Consultas

En esta sección se proporcionan algunos ejemplos de consultas que se pueden realizar gracias a la estructura multidimensional propuesta para el análisis de costes.

Como se ha comentado anteriormente, gracias a la estructura de referencia y al método de clasificación que se han propuesto en esta memoria, se ha podido definir la jerarquía de la dimensión y nutrir el cubo de datos con las actividades de los proyectos de manera estructurada. Además, gracias al módulo de adquisición presentado en este capítulo, no sólo se extraen las descripciones de partida, sino que también se extrae la información relacionada con cada una de éstas, tales como, coste del presupuesto y cantidad. En la fase de ejecución del proyecto, se incorporará el coste real de ejecución para cada una de las actividades.

Las consultas presentadas en esta sección se realizan sobre el conjunto de descripciones de partida de los proyectos de la experimentación del capítulo anterior, ya que presenta más diversidad en cuanto a las tipologías de edificación. El objetivo que se pretende alcanzar en esta sección es ilustrar diferentes tipos de consulta que se pueden realizar con la estructura multidimensional propuesta en esta sección. Como veremos a continuación, Linguistic F-Cube Factory incluye mecanismos que permiten ocultar la complejidad de los resultados y facilitan al usuario la comprensión de los resultados de las consultas. Para ello, presenta los resultados por medio de: combinación de colores mediante gradiente y barras de anchura variable.

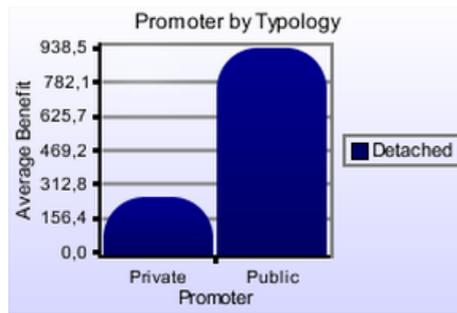


Figura 28: Beneficio medio en los proyectos de viviendas unifamiliares en función de la tipología del promotor - Crisp

- En la Figura 28 se puede observar el resultado de la consulta: *“Beneficio medio en los proyectos de viviendas unifamiliares en función de la tipología del promotor”*. En este caso, los resultados no presentan imprecisión porque las dimensiones involucradas en la consulta no representan conceptos difusos.
- En la Figura 29 se puede observar el resultado de la consulta *“Beneficio medio en la fase inicial en función del tamaño de la empresa”*. Como se puede observar, las *empresas grandes* obtienen mayor

beneficio en la fase inicial de la obra, mientras que, los otros tipos de empresa obtienen un resultado más bajo y similar entre ellos. Este tipo de consulta resulta interesante para los responsables en el proceso de contratación de empresas para que realicen las diferentes actividades del proyecto.

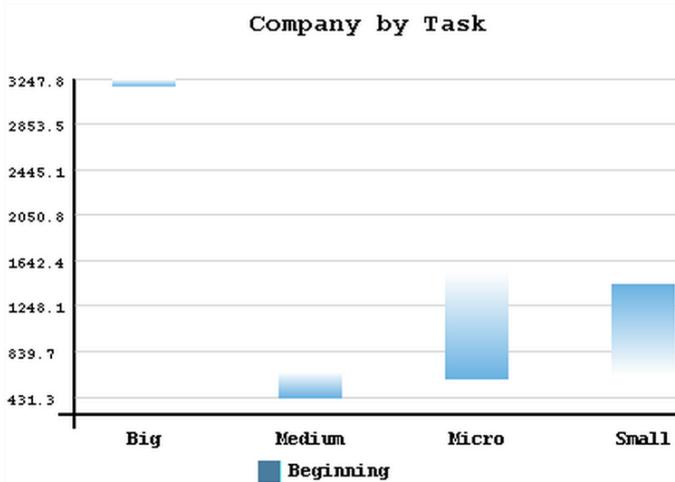


Figura 29: Beneficio medio en la fase inicial en función del tamaño de la empresa

- Otra consulta interesante para los responsables en tomar decisiones es la siguiente, *“Desviación de coste máximo en cada fase según las tipologías de los proyectos”*. La desviación de coste es la diferencia entre el coste estimado en la fase de diseño y el coste real para una actividad dada en la fase de ejecución. Como resultado de la consulta, en la Figura 30 se puede observar que la desviación máxima se produce en la fase intermedia en la tipología de *servicios*. La siguiente desviación corresponde con la tipología de proyectos *residenciales* en la fase intermedia. Como se puede observar en la figura, en los proyectos de tipo industrial la desviación de coste que se produce es menor a la de las otras tipologías.
- En la Figura 31 se puede observar el resultado de la consulta: *“Desviación máxima de coste en proyectos de gran tamaño según las fases de la obra”*. Como se puede observar, en la fase intermedia es dónde se produce la desviación de coste mayor y, además, donde presenta mayor imprecisión.

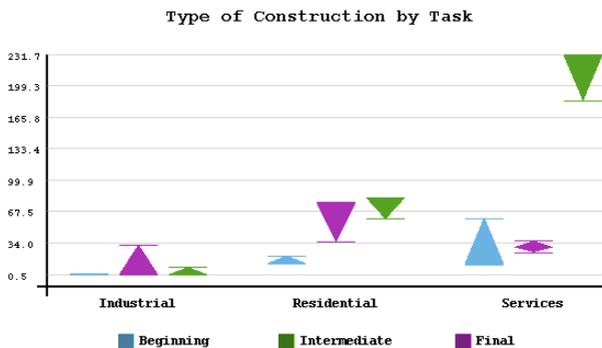


Figura 30: Desviación de coste máximo en cada fase según las tipologías de los proyectos

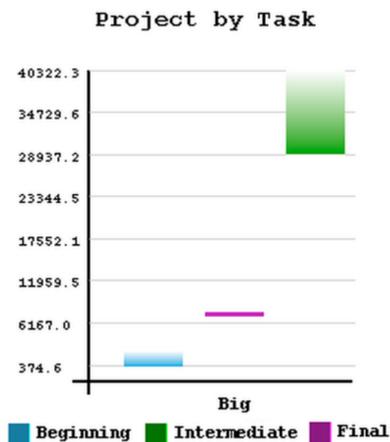


Figura 31: Desviación máxima de coste en proyectos de gran tamaño según las fases de la obra

- A partir de la consulta anterior, los usuarios pueden ampliar el nivel de detalle sobre los capítulos que presentan una desviación de coste mayor, mediante la siguiente consulta: “Desviación de coste media en proyectos de gran tamaño según los capítulos del presupuesto.”. En la Figura 32 se pueden observar los capítulos donde se

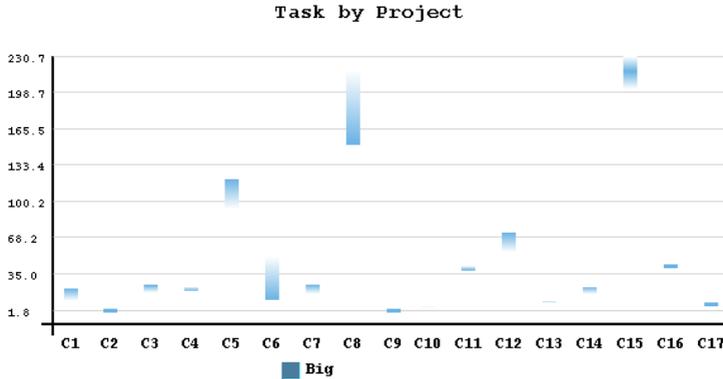


Figura 32: Desviación de coste media en proyectos de gran tamaño según los capítulos del presupuesto

5.3.2 Implementación de la aplicación

Como hemos comentado anteriormente, el modelo multidimensional ha sido implementado utilizando la herramienta Linguistic F-Cube Factory [47], que es una herramienta consolidada para el análisis de cubos OLAP.

Linguistic F-Cube Factory es una plataforma web basada en un diseño cliente-servidor, por lo que el usuario no necesita instalar software en su equipo. El usuario solamente necesita un navegador web y una conexión a la red de trabajo donde se encuentre el servidor.

La parte del servidor es la encargada de soportar la componente más pesada del sistema y para ello está compuesta por dos módulos principales sobre los que se asocian otros módulos adicionales. El primero de los módulos principales es el que se encarga de la gestión de los cubos de datos y la conexión con la base de datos. Una de las funcionalidades del módulo es el soporte a las consultas. Se incluye además la posibilidad de trabajar con vistas de usuario.

El otro módulo principal se ocupa de las funciones de agregación usadas en las consultas. Dicho módulo interactúa con el anterior cuando queremos cambiar el nivel de detalle de un cubo de datos. Existen dos tipos de agregaciones implementadas: las usuales para cubos de datos convencionales o crisp y las difusas que se aplican sobre cubos de datos difusos.

5.4 CONCLUSIONES

En este capítulo se han presentado dos herramientas basadas en un diseño cliente-servidor que permiten a los usuarios gestionar la información relativa a los costes de un proyecto para dar soporte al proceso de toma de decisiones de manera integral.

La primera aplicación *SIGPE*, permite la adquisición y gestión de los datos contenidos en este documento de manera estructurada e integrada. Las funcionalidades de la herramienta se han dividido en los siguientes tres módulos:

- El primero, el módulo de adquisición, permite la inserción del presupuesto independientemente de la herramienta que se ha utilizado para elaborarlo. Mediante este módulo y, de manera transparente y sencilla para el usuario, los datos contenidos en el presupuesto se almacenen en la estructura de referencia.
- El segundo, el módulo de edición, permite la validación y edición de los resultados obtenidos en el proceso de clasificación. Para realizar esta tarea, la herramienta facilita al usuario mediante una escala de colores la identificación de las descripciones de partidas que el clasificador no tiene seguridad de haber clasificado correctamente.
- Finalmente, el módulo de recuperación permite acceder a los datos que han sido almacenados en la estructura de referencia de manera integrada.

La aplicación se ha evaluado a través de dos vías, por un lado, mediante la interacción de los usuarios con la aplicación y, por otro, mediante el cuestionario que han completado dichos usuarios. En la primera evaluación, los resultados obtenidos a partir de la interacción de los usuarios ratifican la capacidad de la metodología propuesta para clasificar las descripciones de partidas en la estructura de referencia con un alto porcentaje de acierto. En la segunda evaluación, los resultados obtenidos en las encuestas realizadas a los usuarios expertos muestran resultados muy satisfactorios sobre la usabilidad y la utilidad de *SIGPE*.

La segunda herramienta, nos ha permitido comprobar cómo el hecho de disponer los datos de los presupuestos de manera estructurada en un repositorio común, gracias a las propuestas de los capítulos anteriores, permite nutrir de datos hipercubos sobre los que realizar procesamiento analítico en línea (OLAP) para extraer conocimiento útil para el proceso de toma de decisiones. Se ha propuesto una estructura de datos multidimensional, que permite el análisis de los costes de

una manera más flexible e intuitiva gracias a la incorporación de conceptos de la Teoría de Conjuntos Difusos proporcionando información estratégica para el funcionamiento de la empresa.

CONCLUSIONES Y TRABAJOS FUTUROS

RESUMEN: En este capítulo se presentan las conclusiones obtenidas tras el trabajo realizado y la descripción de las futuras líneas de investigación abiertas a raíz del presente trabajo.

ÍNDICE

6.1	Versión en Castellano	153
6.1.1	Conclusiones	153
6.1.2	Trabajos Futuros	158
6.2	English Version	162
6.2.1	Conclusions	162
6.2.2	Future Work	166
6.3	Publicaciones científicas	169

6.1 VERSIÓN EN CASTELLANO

6.1.1 Conclusiones

El objetivo general que ha marcado la realización de esta Tesis ha sido la propuesta de herramientas que permitan la construcción de un Sistema de Información para el tratamiento inteligente de los datos contenidos en el documento del presupuesto. Para ello, se han planteado tres objetivos principales: por un lado, proponer una estructura de referencia que pueda ser utilizada para la construcción de almacenes integrados de datos contenidos en este documento. Por otro lado, proponer una técnica para automatizar la adquisición de los datos y la clasificación de éstos en la estructura de referencia y, por último, la puesta en valor de dicha propuesta mediante una herramienta que permite la adquisición y edición de documentos de presupuesto y una propuesta de hipercubo de datos para la gestión analítica de costes en proyectos de edificación.

En la consecución de este objetivo principal, este trabajo de investigación realiza aportaciones como las siguientes:

- En primer lugar, se ha realizado un análisis exhaustivo del tratamiento de datos mediante las Tecnologías de la Información y la Comunicación en el sector de la construcción, que se puede encontrar en el **capítulo 2** y que ha permitido identificar las características propias del área que dificultan la gestión de los datos en los proyectos de edificación y señalar y motivar la necesidad de una gestión integral inteligente de los datos en dichos proyectos.

Para llevar a cabo ese análisis, se ha realizado una completa revisión bibliográfica de las distintas propuestas existentes en la literatura desde dos puntos de vista.

- En primer lugar, se han analizado las principales tecnologías que se han utilizado en la gestión de los proyectos para permitir un manejo adecuado de los datos.
- En segundo lugar, se han analizado las propuestas que se pueden encontrar en la literatura especializada desde el punto de vista de las principales tareas dentro del ámbito de la gestión de proyectos de construcción.

Con este análisis hemos identificado los retos a explorar en relación a las principales tareas del ciclo de vida de los datos; adquisición, almacenamiento, procesamiento y consulta.

- En segundo lugar, en el **capítulo 3**, se ha analizado el proceso de elaboración del documento del presupuesto y se han determinado las principales causas que originan la diversidad y heterogeneidad de los datos en este documento, dificultando el acceso oportuno a los datos.

En este capítulo, se ha propuesto un modelo de datos jerárquico general para organizar los datos contenidos en el documento del presupuesto. El modelo propuesto se basa en la estructura de descomposición del trabajo, comúnmente aceptada en el ámbito nacional e internacional para la organización de las actividades del presupuesto.

En el mismo capítulo, para solventar la falta de estándares de referencia para estructurar los datos de este documento, se ha establecido una instancia concreta de dicho modelo, basada en conocimiento experto en el ámbito de la Edificación en España, y que es utilizada como estructura de referencia en el resto de la memoria. Esta propuesta, si bien ha sido necesaria para el desarrollo de posteriores objetivos de este trabajo de investigación, entendemos que es un aporte significativo en el área de la

Ingeniería de Edificación, puesto que permite organizar la información del presupuesto de proyectos de edificación españoles de manera estructurada y homogénea.

- Para poder construir un almacén integrado de datos de presupuestos no sólo es necesaria la estructura de referencia propuesta, que debe guiar la elaboración de proyectos en el futuro, sino también incorporar todo el conocimiento contenido en documentos de presupuesto ya elaborados y disponibles en las organizaciones. Estos documentos carecen de la estructuración y homogeneidad deseadas y, por este motivo, en tercer lugar y en el **capítulo 4**, se ha presentado un método para la clasificación automática de descripciones de partida en una estructura de referencia común.
 - El método propuesto se basa en el aprendizaje de un modelo de conocimiento que permite utilizar la aparición de términos en las descripciones de partidas para ubicarlas de manera adecuada en la estructura de referencia. La metodología que se ha desarrollado utiliza técnicas de preprocesamiento sintáctico y semántico de datos para la extracción y unificación de los términos que aparecen en las descripciones de partida. Este modelo se ha establecido en base a un proceso de aprendizaje con un conjunto de datos muy utilizados en la elaboración del documento del presupuesto. Concretamente, se han utilizado 7029 descripciones de partidas diferentes correspondientes a las dos bases de precios más usadas en proyectos de edificación españoles, a saber, PREOC [195] y CENTRO [67].
 - Al mismo tiempo, ese modelo se utiliza mediante un método de agregación multicriterio basada en tres dimensiones para la ubicación efectiva de las descripciones de partida: frecuencia, posición y conocimiento experto. Estas dimensiones han establecido, para cada término, la importancia de éste en cada uno de los nodos de la estructura de referencia. La agregación de estas dimensiones con el método propuesto permite la clasificación de una descripción de partida en un cierto nodo de la estructura de referencia. El proceso de agregación que se ha propuesto se fundamenta en el uso de operadores OWA, dada su flexibilidad y potencia bien conocidas en la literatura de la Computación Flexible.
 - Con el objetivo de evaluar la metodología de clasificación propuesta, se ha realizado una experimentación con dos

conjuntos de datos diferentes a los utilizados en la fase de aprendizaje. Por un lado, se han utilizado las bases de precios EXT [68] y BCCA [129] con un total de 12730 descripciones de partida. Y por otro, 9669 descripciones de partida procedentes de proyectos de edificación reales. Se ha probado la técnica sobre estos datos (un total de 22399 descripciones de partida) y se han utilizado diferentes medidas para evaluar la bondad de nuestra metodología para la clasificación de descripciones de partida. La experimentación llevada a cabo ha sido muy ambiciosa pues se ha tenido que realizar con una muestra suficiente de vectores de configuración de los operadores OWA para localizar una instanciación adecuada de los mismos para el proceso de clasificación propuesto.

Los resultados obtenidos demuestran que el mecanismo propuesto cumple con nuestro objetivo de permitir la incorporación de las descripciones de partida en un almacén común de manera integrada.

- Finalmente, en el **capítulo 5**, se han presentado dos herramientas que ponen en valor el modelo y la técnica de clasificación propuestos en los capítulos anteriores.
 - La primera herramienta nos ha permitido mostrar la utilidad de nuestras propuestas para la construcción de un almacén integrado de datos de documentos de presupuesto en el ámbito de los proyectos de edificación. En este sentido, la aplicación desarrollada permite realizar la adquisición, edición y recuperación de los datos contenidos en los presupuestos de proyectos de edificación, de acuerdo con la estructura de referencia y la metodología de clasificación propuestas. La aplicación incorpora tres módulos principales:
 - * El módulo de adquisición importa los datos contenidos en el presupuesto según una descripción de formato previamente establecida y que permite importar datos independientemente de la herramienta con la que ha sido elaborado el documento de presupuesto. Este módulo, haciendo uso del modelo de conocimiento y método de clasificación propuesto en el capítulo 4, permite reestructurar el proyecto importado en base a la estructura de referencia.

- * El módulo de edición permite al usuario interactuar con el sistema para verificar y, en su caso, corregir los resultados del proceso de clasificación previo. El sistema utiliza un código de colores para llamar la atención del usuario sobre decisiones de clasificación más controvertidas. De esta forma, el usuario puede ubicar descripciones de partida en los nodos de la estructura de referencia de manera sencilla.
 - * Por último, un sencillo módulo de recuperación, permite acceder a los datos almacenados en la estructura de referencia de manera integrada de forma fácil e intuitiva.
- La segunda herramienta nos ha permitido mostrar la utilidad de disponer de almacenes de datos organizados según la estructura de referencia propuesta para dar soporte a la toma de decisiones. Para ello, hemos propuesto un modelo de datos multidimensional para la gestión de datos relativos a costes. El modelo propuesto utiliza herramientas de Computación Flexible para la representación y recuperación de los datos, que lo hacen más potente y lo acercan a la forma natural de expresarse los usuarios.

En este sentido, se ha diseñado una estructura multidimensional considerando siete dimensiones que son de interés en el ámbito del análisis de costes: proyecto, localización, tipología, promotor, tiempo, empresa y actividad. Esta última dimensión, imprescindible en el análisis de costes, se ha podido definir gracias a la estructura de referencia propuesta en el capítulo 3.

La estructura multidimensional se ha implementado mediante la herramienta Linguistic F-Cube Factory [47], la cual permite el diseño de dimensiones en base a conceptos difusos. Esta implementación facilita al usuario la realización de consultas en las que se involucran conceptos imprecisos relacionados con datos de los proyectos de edificación de manera más intuitiva.

Estas aportaciones nos han permitido alcanzar el objetivo principal propuesto al inicio de este trabajo de investigación y, en su conjunto, entendemos que son una contribución sólida en el tratamiento inteligente de datos en el ámbito de la Construcción y, al mismo tiempo, abren una línea de investigación futura ambiciosa y prometedora.

6.1.2 *Trabajos Futuros*

A continuación presentaremos las líneas que, en nuestra opinión, pueden aportar beneficios a la gestión de los datos en el ámbito de la gestión de los proyectos y que afrontaremos en el futuro.

Con respecto a la estructura de referencia para la organización de documentos de presupuesto en proyectos de edificación, a parte de los capítulos actualmente contemplados en ésta relacionados con actividades reales en el proceso de ejecución de los proyectos de edificación, se plantea en el futuro la extensión de la estructura para incluir nuevos capítulos *transversales* como por ejemplo, gestión de residuos, control de calidad y, los relativos a seguridad y salud.

Respecto al método de clasificación de partidas presentado en el capítulo 4 se plantean los siguientes trabajos futuros:

- Extender las capacidades del modelo propuesto para que se pueda utilizar en otros países. Para ello, nos planteamos estudiar el documento del presupuesto en otros idiomas y analizar nuestra metodología en este nuevo contexto.
- La experimentación que hemos presentado en este trabajo, ha considerado distintas configuraciones para los operadores OWA utilizados, si bien las configuraciones consideradas han estado restringidas por una heurística de precisión y de orden explicadas en la memoria. Sin embargo, se plantea ahora realizar una experimentación todavía más completa para la configuración de dichos operadores OWA, que considere todo el abanico posible de configuraciones y que, mediante técnicas de optimización, nos permita localizar el mejor vector de pesos para los datos de los proyectos de edificación. En este sentido, estamos considerando el uso de algoritmos de búsqueda y optimización evolutivos.
- En el mismo sentido de mejorar el método de clasificación, se está considerando estudiar configuraciones de la metodología de clasificación propuesta que discriminen en función del tipo de edificación al que corresponden las partidas que se pretenden clasificar.
- Considerar evaluaciones lingüísticas, en lugar de las evaluaciones numéricas aportadas por el experto con respecto a la dimensión que considera conocimiento experto.

Con respecto a la aplicación para la adquisición y edición de presupuestos de proyectos de edificación, nos planteamos los siguientes trabajos futuros:

- Desarrollar módulos que enriquezcan el sistema con los siguientes tipos de datos:
 - Como se ha comentado, el documento del presupuesto establece la descomposición de actividades que son la base de otras muchas tareas en la gestión de los proyectos de construcción. Por lo tanto, una vez que esta información está estructurada, se le puede integrar información muy útil con el fin de dar soporte a otras tareas de la gestión de proyectos. Por ejemplo, se puede añadir información relativa a los costes de ejecución, planificación, estado de ejecución etc.
 - Incorporar datos procedentes de sistemas ubicuos, generados a partir de diversas tecnologías, como por ejemplo, mediante el uso de redes de sensores en tiempo real y computación móvil (mobile computing). Estos sistemas proporcionarán datos útiles para el sistema de información ampliando las opciones de consulta.
- Mejorar la amigabilidad e interacción con los usuarios. Para ello, nos planteamos los siguientes trabajos futuros:
 - Incorporar la retroalimentación de los usuarios para adaptar dinámicamente la herramienta de clasificación, a partir de un análisis de la interacción llevada a cabo por los usuarios con el módulo de edición.
 - Explorar nuevas reglas para establecer la escala de color que proporciona al usuario la fiabilidad con la que el clasificador establece la clasificación de las descripciones de partida. Para ello nos planteamos el uso de sistemas de extracción automática de reglas en conjuntos de datos e incorporarlas en nuestro sistema mediante nuevas escalas de color.

Con respecto a la elaboración de herramientas para la explotación del sistema de información de datos de proyectos de edificación, nos planteamos los siguientes trabajos futuros:

- Desarrollar nuevos cubos de datos para dar soporte a otras tareas importantes en la gestión de los proyectos de edificación, tales como la planificación, gestión de riesgos y la seguridad. Por

ejemplo, el análisis de las desviaciones que se pueden producir en las actividades de la obra en base a experiencias anteriores puede ayudar a prevenir retrasos en los proyectos actuales. Del mismo modo, la información relativa a la gestión de los diferentes riesgos asociados a dichas actividades puede aportar valiosa información para el proceso de toma de decisiones.

- Incorporar técnicas que permitan realizar consultas flexibles sobre el propio sistema en términos de recuperación. Por ejemplo, técnicas de similitud de cadenas que nos permitan recuperar descripciones de partida similares a una descripción dada. También, podríamos contemplar el uso de nubes de palabras (en inglés, tag cloud), considerando los términos de las descripciones, con el objetivo de facilitar la búsqueda de información de manera visual sobre los niveles de la estructura de referencia. Por último, incorporar el concepto de consultas bipolares (en inglés, bipolar queries) que ofrecen mayor flexibilidad en la recuperación de información, permitiendo establecer criterios de consulta de una manera mas cercana al razonamiento humano [76].
- Explorar otras soluciones de consulta amigables para los usuarios, como por ejemplo, las basadas en el uso del lenguaje natural y sensibles al contexto en la interacción con el sistema [37, 102, 227]. En este sentido, se plantea la posibilidad de proponer herramientas que permitan construir, de manera automática, descripciones lingüísticas sobre la ejecución de los proyectos o con información relevante para la toma de decisiones. En este sentido, una línea de especial interés es la relacionada con la descripción lingüística de series de datos temporales, tan presentes en los proyectos de edificación, y que se muestra muy activa en el ámbito de los sistemas de Generación de Lenguaje Natural y el resumen de datos en Computación Flexible [46].

Por último, nos planteamos estudiar la integración de los datos de nuestro repositorio de datos históricos, con las herramientas del entorno BIM. Como hemos comentado, BIM es un modelo nD de datos único que incorpora a través del modelo 3D los datos requeridos en las diferentes fases del ciclo de vida del proyecto, entre ellos, los relacionados con los costes (BIM 5D). En este sentido, la aportación de los datos correspondientes a proyectos históricos enriquecería el proceso de estimación de costes.

La interoperabilidad en el entorno BIM es un obstáculo para su completa implantación en el ciclo de vida de los proyectos, ya que dificulta la integración de información y recuperación [3]. En este sentido,

en la actualidad, los principales esfuerzos se están produciendo en cuanto al desarrollo de representaciones de conocimiento más flexibles que permitan capturar la semántica de los modelos, para facilitar la recuperación de información e integrar BIM con otros sistemas [104].

6.2 ENGLISH VERSION

6.2.1 *Conclusions*

The general objective that has encouraged the development of this thesis has been the proposal of tools that allow the construction of an Information System for the intelligent processing of data contained in the bill of quantities document. To do this, three main objectives have been posed. Firstly, to propose a reference structure that can be used for the construction of integrated repositories of data contained in this document. Secondly, to propose a methodology to automate data acquisition and the classification of these data in the reference structure. Finally, the development of such proposals by means of, on the one hand, a tool that allows us to acquire and edit data and, on the other hand, a proposal of a data cube for analytical cost management in building projects.

After achieving these main objectives, this doctoral dissertation provides the followings contributions:

- Firstly, in **Chapter 2**, a complete analysis regarding the data processing by the use of the information and communications technologies in the construction sector has been conducted. This analysis has enabled the identification of construction characteristics that make data management in the building projects difficult and motivate the need for an intelligent integrated data management in such projects.

To accomplish this analysis, a complete review of the proposals in the literature from two perspectives has been carried out. To this end, we have reviewed the main ICT tools that have been used in project management research to allow proper data handling, and then, we have reviewed ICT-based proposals in relation to the project management tasks that they aim to improve.

By means of this analysis, we have identified the challenges in relation to the main tasks of the data life cycle: acquisition, storage, processing, and query.

- Secondly, in **Chapter 3** the elaboration process of the bill of quantities document has been analyzed. Then, we have identified the main causes of the diversity and heterogeneity of the data in this document which make the timely access to data difficult.

In addition, in this chapter, a general hierarchical data model for organizing data contained in bill of quantities documents has been proposed. The proposed model is based on the work breakdown structure, commonly accepted at national and international level to organize the activities in the bill of quantities document.

To solve the lack of reference standards for structuring the information in this document, a particular instance of the model has been proposed in this chapter. This instance is based in expert knowledge in the building domain and is used as a reference structure in this memory.

This proposal, which has been necessary to develop further objectives of this research, is also, in our opinion, a significant contribution in the area of Building Engineering, as it allows the organization of the information contained in bills of quantities from Spanish building projects in a structured and homogeneous way.

- To build an integrated repository of bill of quantities data, not only as a reference structure that guides the development of future projects is necessary, but also, to incorporate all the knowledge contained in already elaborated bill of quantities documents, available in the organizations. These documents lack the desired structure and homogeneity. As a result, in **Chapter 4**, a method to automatically classify work descriptions in a reference structure has been proposed.

The proposed method is based on the learning of a knowledge model that allows us to use the appearance of terms in the work descriptions to appropriately place them in the reference structure. The developed methodology uses syntactic and semantic data preprocessing techniques for the extraction and unification of the terms used in the descriptions. This model has been established on the basis of a learning process with a widely used dataset used for the elaboration of the bill of quantities document. In particular, 7029 different work descriptions have been used corresponding to two of the most commonly used price databases in building projects in Spain, namely, PREOC [195] and CENTRO [67].

This knowledge model is used by a multi criteria aggregation method based on three dimensions to the effective location of the work descriptions: frequency, position, and expert knowledge.

These dimensions have established, for each term, its importance in each node of the reference structure. The aggregation of these dimensions by means of the proposed method allows the classification of a work description in a given node of the reference structure. The proposed aggregation method is based on the use of OWA operators given their well-known flexibility and power in the flexible computing literature.

With the aim of assessing the proposed classification methodology, a complete experimentation with two different data sets to those used in the learning phase was conducted. The first set is composed of two price databases, EXT [68] and BCCA [129], with a total of 12730 work descriptions. And the second set is composed of 9669 work descriptions corresponding to real building projects. The technique has been tested with these data (a total of 22399 work descriptions) and various measures have been used to evaluate the goodness of our methodology for classifying work descriptions.

The experimentation has been very ambitious because it had to be performed with a wide sample of weighting vectors of OWA operators in order to find a good setting for the proposed classification process.

The results show that the proposed methodology meets with guarantees our objective of incorporating work descriptions in a common repository in an integrated manner.

- Finally, in **Chapter 5**, two applications that highlight the significance of the model and the classification method proposed in the previous chapter have been proposed.
 - The first application has allowed us to show the usefulness of our proposals for the construction of an integrated repository of bill of quantities documents of the building projects domain. In this sense, the developed application allows the acquisition, edition, and retrieval of data regarding the reference structure and the classification methodology proposed. The application includes three main modules:
 - * The acquisition module imports the data contained in the bill of quantities document according to a previous description format which allows us to import data regardless of the tool that has been used to develop the document. This module enables the restructuring of the imported data in the reference structure

by means of the knowledge model and classification method proposed in Chapter 4

- * The editing module allows users to interact with the system to verify and, if necessary, to correct the results of the previous classification process. The system uses a color code to attract the attention of the user in controversial classification decisions. Thus, users can easily locate the work descriptions in the reference structure.
 - * Finally, an easy retrieval module which provides access to data stored in the reference structure in an integrated, easy, and intuitive manner.
- The second application has allowed us to show the usefulness of having organized data warehouses according to the proposed reference structure. To do this, we have proposed a multidimensional data model for managing cost data. The proposed model uses flexible computing tools for data representation and retrieval which make the model more powerful and closer to the way in which users express themselves. In this sense, a multidimensional structure has been designed considering seven interesting dimensions in the cost analysis domain: project, location, typology, promoter, time, company, and tasks. This latter dimension, which is essential in the cost analysis, has been defined thanks to the proposed reference structure in chapter 3. The multidimensional model has been implemented using Linguistic F-Cube Factory [47] which allows the design of dimensions based on fuzzy concepts. This implementation allows users to make queries involving imprecise concepts regarding building project data in an intuitive way.

These contributions have enabled us to achieve the main objective planned at the beginning of this thesis dissertation and we consider that they are a solid contribution in intelligent data processing in the construction domain. At the same time, these contributions open a future, ambitious, and promising research line.

6.2.2 *Future Work*

Below, we present some of the lines which, according to our opinion, can benefit data handling in the project management domain, and which we will face in the future.

With respect to the reference structure for the organization of bill of quantities documents of building projects, in addition to the chapters currently considered in relation to real activities in the execution process of building projects, we pose the extension of the structure to include new chapters, such as, waste management, quality control, and those relating to health and safety.

Regarding the proposed classification method in chapter 4, we pose the following future work:

- Extending the capabilities of the proposed model in order to be able its use in other countries. To this end, we plan to study the bill of quantities document in other languages and analyze our methodology in this new context.
- The experiments we have presented in this Thesis have considered various settings for OWA operators. The configurations considered have been restricted by a precision and order heuristic as explained in the memory. However, we aim to perform a more complete experimentation for OWA configuration, which considers all the possible range of settings by means of optimization techniques that allow us to locate the best weighting vector for building project data. In this regard, we are considering the use of evolutionary optimization algorithms.
- Also for improving our classification method, we are considering to study different configurations in our classification methodology that can discriminate work descriptions depending on the type of building.
- Taking into account linguistic evaluations, instead of the numerical ones provided by the expert in the dimension considered to be expert knowledge.

According to our application for the bill of quantities acquisition and editing in building projects, we propose the following future work:

- To develop modules that incorporate data in real time and ubiquitously. As we have mentioned, the bill of quantities document provides the breakdown of activities that are the basis for many other tasks in the management of construction projects. Therefo-

re, once this information is structured, you can integrate other useful information in order to support other tasks in project management. For example, you can add information regarding costs, planning, execution status, etc. Thus, firstly, the amount of time to store this information would be reduced and, secondly, query options of the tool would be expanded.

- To incorporate feedback from users to dynamically adapt our classification tool, from an analysis of the interaction of the users with the editing module.
- To explore new rules for setting the color scale which shows in a visual way the reliability with which the classifier establishes the classification of the work descriptions. To do this, we consider the use of automatic extraction rule systems from datasets and propose to incorporate them into our system through new color scales.

As regards the development of tools for the exploitation of the integrated information system in construction projects, we consider the following future work:

- To develop new data cubes to support other important tasks in the management of building projects, such as planning, risk management, and security. For example, the analysis of the deviations that can occur in the activities of the work based on past experiences can help prevent delays in current projects. Similarly, information relating to the management of the various risks associated with such activities can provide us with valuable information for the decision-making process.
- To incorporate other techniques allowing flexible querying in terms of retrieval on the system. For example, “string similarity” techniques that allow to retrieve similar work descriptions from a given description. We may also consider the use of “tag cloud”, considering the terms of work descriptions, in order to facilitate the search for information on the different levels of the reference structure in a visual way. Furthermore, to integrate the concept of “bipolar queries” whose offer greater flexibility in the information retrieval task allowing users to query in a more closely manner to human reasoning [76].
- To explore other user friendly query solutions, such as those based on the use of natural language and context sensitive in the interaction with the system [37, 102, 227]. In this sense, we consider the use of tools to build automatically linguistic descriptions on the basis of relevant information to the decision-

making process in the project execution. In this sense a special line of interest is related to the linguistic description of time series data, which is present in building projects, and is very active in the field of Natural Language Generation systems and in the data in Flexible Computation [46].

Finally, we pose to study the data integration from our repository with BIM tools environment. As mentioned, BIM is a nD data model that incorporates required data in the different phases of the project life cycle, including related costs (BIM 5D) through the 3D model. In this sense, data from historical projects might enrich the estimating costs process.

Interoperability in the BIM environment is an obstacle to its full implementation in the projects life cycle since it makes difficult information integration and retrieval [3]. Hence, currently the main efforts are taking in the development of more flexible knowledge representations that capture the semantics of models in order to facilitate information retrieval and to integrate BIM with other systems [104].

6.3 PUBLICACIONES CIENTÍFICAS

Parte de los resultados obtenidos durante el desarrollo de la Tesis doctoral se han incluido en las siguientes publicaciones:

- Martínez-Rojas, M., Marín, N. y Vila, M.A.,[173], “An Approach for the Automatic Classification of Work Descriptions in Construction Projects”. *Computer-Aided Civil and Infrastructure Engineering*, 2015, (In press).
- Martínez-Rojas, M., Marín, N. y Vila, M.A.,[174], “The Role of Information Technologies to Address Data Handling in Construction Project Management”. *Journal of Computing in Civil Engineering*, 2015, p. 04015064. DOI: 10.1061/(ASCE)CP.1943-5487.
- Martínez-Rojas, M.; Marín, N.; Molina, C. y Vila, M.A.,[168], “Cost Analysis in Construction Projects using Fuzzy OLAP Cubes”. En: *IEEE International Conference on Fuzzy Systems, FUZZIEEE 2015, Istanbul, Turkey, 2015*.
- Martínez-Rojas, M., Marín, N. y Vila, M.A., [171], “A preliminary approach to classify work descriptions in construction projects”. En *IFSA World Congress and NAFIPS Annual Meeting (IFSA/-NAFIPS)*, 2013 Joint, pp. 1090-1095, 2013. DOI: 10.1109/IFSA-NAFIPS.2013.6608552.
- Martínez-Rojas, M., Marín, N. y Vila, M.A., [170], “Applying ICTs to prevent accidents in construction sector”. En: *International Congress on Safety and Labour Market*, 2013.
- Martínez-Rojas, M.; Marín, N. y Vila, M.A.,[172], “Aplicación de las TICs en el Ámbito de la Construcción”, *Iniciación a la Investigación*, 2013, 1(5), pp. 1-9. ISSN 1988-415X.
- Martínez-Rojas, M., Marín, N. y Vila, M.A.,[169], “Sistemas de Data Warehousing en el Ámbito de la Construcción”. En: *Actas de las II Jornadas Andaluzas de Informática*, , 2013.

MANUAL DE LA APLICACIÓN SIGPE

En este apéndice se presenta el manual de la aplicación *SIGPE* (Sistema Inteligente de Gestión de Proyectos de Edificación) para la adquisición y edición de presupuestos de proyectos de edificación propuesta en el capítulo 5.

A modo de ejemplo, se ha seleccionado un presupuesto correspondiente a un proyecto de edificación real de una vivienda unifamiliar aislada. El documento del presupuesto ha sido elaborado con la herramienta *Presto*, que es una de las más usadas en el contexto de la construcción y ofrece al usuario gran libertad para organizar el documento. Por ejemplo, *Presto* permite la incorporación de descripciones de partidas de diferentes bases de precios, la edición de éstas o la redacción de nuevas descripciones. El proyecto de ejemplo usado en este manual presenta una estructura para organizar las descripciones que difiere considerablemente de la estructura de referencia en la que se clasificarán las descripciones de partida de manera automática, poniendo de manifiesto la utilidad real de *SIGPE* de integrar descripciones de partidas independientemente de la estructura original del documento.

En la Tabla 24 se puede observar la estructura de descomposición del trabajo (WBS) del proyecto de ejemplo. Según el modelo formal definido en el capítulo 3, la estructura del proyecto se compone de cuatro niveles, los tres primeros niveles, L_1 , L_2 y L_3 , se denominan *Proyecto*, *Capítulo* y *Subcapítulo* respectivamente. El nivel donde se encuentra el conjunto de las descripciones de partidas corresponde al L_4 .

La Figura 33 muestra la apariencia de la pantalla principal de *SIGPE*. En esta pantalla, el usuario deberá introducir su nombre de usuario y su contraseña, con el fin de poder acceder a la aplicación. Para poder acceder a *SIGPE* es necesario que el usuario esté registrado en el sistema. Este registro, sólo puede hacerlo el administrador del sistema por las siguientes razones:

- Permitir acceso sólo a los usuarios autorizados.
- Permitir a cada usuario el acceso a funciones específicas del sistema, en función de la responsabilidad del usuario.

CAPÍTULOS (L ₂)	SUBCAPÍTULOS (L ₃)
ACONDICIONAMIENTO Y CIMENTACIÓN	CONTENCIONES
	MOVIMIENTO DE TIERRAS
	CIMENTACIONES SUPERFICIALES
ESTRUCTURAS	ESTRUCTURAS.HORMIGÓN ARMADO
	ESTRUCTURAS.HORMIGÓN PRETENSADO
FACHADAS Y PARTICIONES	ACRISTALAMIENTOS
	CERRAJERÍA
	DEFENSAS
	FÁBRICAS
	PUERTAS CARPINTERÍA
	REMATES
INSTALACIONES	VENTANAS CARPÍNERÍA
	AUDIOVISUALES
	CLIMATIZACIÓN
	DEPÓSITOS
	ELECTRICIDAD
	FONTANERÍA
	ILUMINACIÓN
PROTECCIÓN	
AISLAMIENTOS	SANEAMIENTO
	IMPERMEABILIZACIÓN
CUBIERTAS	AISLAMIENTO
	AZOTEAS
REVESTIMIENTOS	ALICATADOS
	ENFOCADOS
	GUARNECIDOS Y ENLUCIDOS
	PINTURAS
	SOLADOS
SEÑALIZACIÓN Y EQUIPAMIENTO	FALSOS TECHOS
	MOBILIARIO
URBANIZACIÓN	URBANIZACIÓN
VARIOS	VARIOS

Tabla 24: Estructura de descomposición de trabajo (WBS) del proyecto de ejemplo correspondiente a una vivienda unifamiliar aislada.

- Validar la aplicación mediante la interacción de los usuarios. Más adelante, el análisis de la interacción de cada uno de los usuarios puede aportar información interesante para adaptar algunos parámetros del proceso de adquisición de manera personalizada a cada usuario.

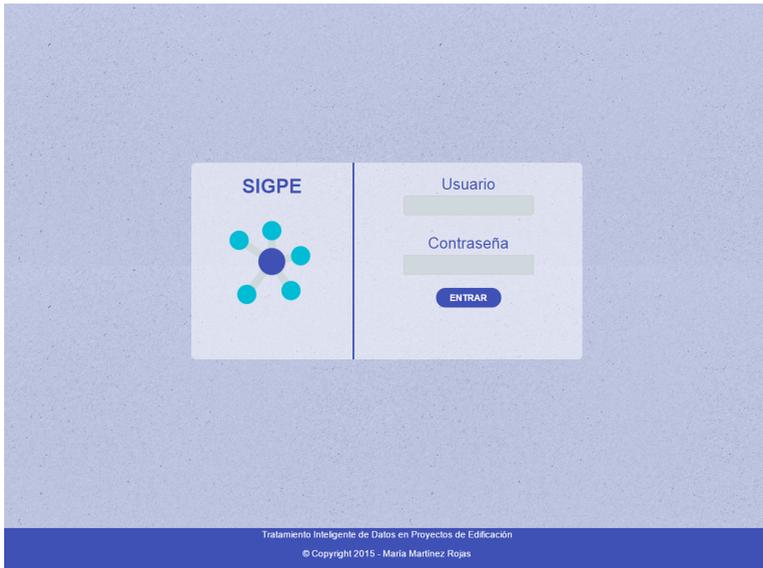


Figura 33: Inicio de sesión en SIGPE.

Una vez que el usuario se ha identificado correctamente y ha accedido a SIGPE, éste puede acceder a las funcionalidades de la aplicación, bien a través del menú de la parte superior o bien a través de los botones “ir” correspondientes a cada uno de los módulos: Adquisición, Edición y Recuperación. Para cerrar la sesión, se puede realizar mediante el botón de la derecha de la barra de menú.

A continuación, en las siguientes secciones se ilustran ejemplos de uso de las distintas funcionalidades de SIGPE sobre el proyecto de ejemplo. En la sección A.1 se procede a la inserción del presupuesto de ejemplo (se puede acceder mediante el botón de *adquisición*), así como a la clasificación de las partidas del presupuesto de ejemplo en la estructura de referencia. En la sección A.2 se ilustra cómo se interactúa con SIGPE en la tarea de edición de los resultados de clasificación. Finalmente, en la sección A.3 se muestran ejemplos de consulta que se pueden llevar a cabo en SIGPE.



Figura 34: Ventana principal en *SIGPE*.

A.1 INSERCIÓN DE UN PRESUPUESTO

La inserción de un nuevo proyecto con un documento de presupuesto se puede realizar a través del botón “Adquisición” de la barra de menú de la parte superior. La Figura 35 muestra el aspecto de éste módulo.

El usuario deberá rellenar los campos asociados al nuevo proyecto, tales como, nombre, país, ciudad, tipología y una breve descripción del proyecto así como la localización del proyecto mediante la ubicación en el mapa. Ésta última se puede fijar fácilmente interactuando con el mapa, simplemente arrastrando el icono de localización a la localización deseada.

Asimismo, se puede adjuntar el fichero del presupuesto seleccionando un archivo con la información de éste. En este caso, el presupuesto del proyecto de ejemplo se ha elaborado con la herramienta *Presto* y se ha exportado en formato .xls (Excel) para insertarlo en *SIGPE*.

Una vez seleccionado el archivo Excel, e introducidos los datos del nuevo proyecto, al pulsar el botón “Aceptar” se crea automáticamente un nuevo proyecto en *SIGPE* con dichos datos junto con las descripciones de partidas del presupuesto procedentes de la herramienta *Presto*.

Durante el proceso de creación del nuevo proyecto se muestra una barra de estado con el progreso de éste. Al finalizar, *SIGPE* muestra un resumen del proyecto donde aparece el número total de partidas que se han extraído del presupuesto de ejemplo (Figura 36).

Figura 35: Insercción de un nuevo proyecto en *SIGPE*

Las partidas extraídas del documento del presupuesto, procedentes de *Presto*, aún no se han clasificado en la estructura de referencia. Este proceso se puede realizar ahora, mediante el botón “Proceso de clasificación” (Figura 36), o posteriormente, mediante el módulo de Edición (accesible desde el menú Edición), tal y como veremos en la sección A.2. De esta forma *SIGPE* da la elección al usuario si continuar con la revisión de la clasificación de partidas ahora o continuar en otro momento.

Si el usuario decide clasificar ahora las partidas en la estructura de referencia, bastará con pulsar el botón “Proceso de clasificación” y *SIGPE* comenzará el proceso de clasificación. Durante este proceso se muestra una barra de estado con el avance del proceso de clasificación. Además, se muestran el número total de partidas a clasificar (en este ejemplo 244) y la partida por la que *SIGPE* va clasificando. El proceso de clasificación es eficiente, en el caso del ejemplo anterior, *SIGPE* tarda unos 20 segundos en clasificar las 244 partidas en la estructura de referencia. Una vez finalizado el proceso de clasificación, *SIGPE* muestra un resumen del proyecto con los capítulos de la estructura de

Figura 36: Información del proyecto una vez importado documento del presupuesto.

referencia junto con el número de partidas clasificadas en cada capítulo (Figura 37).

A.2 EDICIÓN DE UN PRESUPUESTO

Una vez que las descripciones de partidas han sido clasificadas en la estructura de referencia, es necesario validar el resultado proporcionado por el clasificador, con el fin de ubicar de forma definitiva las descripciones de partida en su lugar correspondiente. Este módulo se encarga de esa tarea, al cual se puede acceder desde el menú “Edición”.

SIGPE muestra una tabla con la lista de todos proyectos asociados a cada usuario (Figura 38) para su edición, tanto clasificación como borrado. En la tabla aparecen remarcados con el logo de *SIGPE* aquellos proyectos que el usuario ha insertado pero aún no ha clasificado. Los proyectos se pueden ordenar en orden ascendente o descendente para cada uno de los campos que aparecen en el encabezado de la tabla, tales como, nombre, país, ciudad o fecha de ejecución con el objetivo de facilitar la interacción con el usuario. Además, mediante la ventana de búsqueda que aparece en la parte superior derecha, se puede buscar proyectos mediante la inserción de texto. Asimismo, con los botones



Figura 37: Información del proyecto y la distribución de las descripciones de partida en la estructura de referencia.

“Anterior” y “Siguiete” se puede navegar por la tabla de manera ascendente y descendente.

Si se selecciona un proyecto de los listados en la tabla que se muestra en la Figura 38, se procede al proceso de revisión y edición de los resultados de clasificación. Por ejemplo, la Figura 39 muestra el listado de capítulos del proyecto de ejemplo (VIV.UNIF 0840) junto con el coste total de cada capítulo así como el número de subcapítulos contenidos en cada capítulo según la estructura de referencia y el proceso de clasificación llevado a cabo por SIGPE. A este proceso se podría haber llegado el caso de haber seleccionado el botón “Ver partidas” cuando terminó el proceso de clasificación comentado en la sección anterior. Nótese que en este ejemplo no aparece el capítulo 1 (demoliciones C₁), ya que el ejemplo de proyecto mostrado se trata de una obra de edificación de nueva planta que no necesita demolición.

En el margen izquierdo se pueden observar cinco círculos que, como se ha detallado en la sección 5.2.2, proporcionan a los usuarios la fiabilidad, de forma visual, con la que el clasificador determina la ubicación de la descripción de partida en la estructura de referencia. Los colores verdosos representan mayor confianza en la clasificación, mientras que los colores rojizos representan menor confianza en los resultados de clasificación.

Nombre	País	Ciudad	Fecha ejecución	Eliminar
VIV.LINF.0713	ESPAÑA	GRANADA	2015-10-19	
VIV.LINF.0840	ESPAÑA	GRANADA	2015-10-19	
PROYECTO 45	ESPAÑA	GRANADA	2015-10-17	
PROYECTO 44	ESPAÑA	GRANADA	2015-10-17	
PROYECTO 43	ESPAÑA	GRANADA	2015-10-17	
PROYECTO 31	ESPAÑA	MALAGA	2015-10-17	
PROYECTO 42	ESPAÑA	GRANADA	2015-10-17	
PROYECTO 042	ESPAÑA	GRANADA	2015-10-17	
PROYECTO 0042	ESPAÑA	ALMERIA	2015-10-17	
VIV.LINF.0840	ESPAÑA	GRANADA	2015-10-15	

Mostrando página 1 de 2

PREVIOUS 2 NEXT

Tratamiento Inteligente de Datos en Proyectos de Edificación
© Copyright 2015 - María Martínez Rojas

Figura 38: Lista de proyectos almacenados para un usuario

Para acceder a las descripciones de partidas de un capítulo basta con hacer click sobre dicho capítulo. En la Figura 40 se puede observar un ejemplo de las descripciones de partidas del capítulo 8 (Instalaciones) pintadas de color, acorde a la escala de color mencionada anteriormente.

Para cada una de las descripciones de partida, *SIGPE* aporta la siguiente información que se ha extraído del documento original:

- Resumen de descripción de partida
- Descripción de partida
- Cantidad (C)
- Coste de presupuesto (P)
- Total (T)

ADQUISICIÓN EDICIÓN RECUPERACIÓN	
C2-ACONDICIONAMIENTO DEL TERRENO	8845.71€ Contiene 4 subcapítulos
C3-RED HORIZONTAL DE SANEAMIENTO	3448.66€ Contiene 5 subcapítulos
C4-CIMENTACIONES	50773.54€ Contiene 4 subcapítulos
C5-ESTRUCTURA	28550.66€ Contiene 2 subcapítulos
C6-ALBAÑILERÍA	66079.49€ Contiene 5 subcapítulos
C7-CUBIERTAS	6889.86€ Contiene 2 subcapítulos
C8-INSTALACIONES	31948.04€ Contiene 7 subcapítulos
C9-AISLAMIENTO E IMPERMEABILIZACIONES	11948.82€ Contiene 2 subcapítulos
C10-REVESTIMIENTOS	57060.22€ Contiene 3 subcapítulos
C11-CARPINTERÍA DE MADERA	11088.61€ Contiene 4 subcapítulos
C12-CARPINTERÍA EXTERIOR Y CERRAJERÍA	47403.82€ Contiene 3 subcapítulos
C13-VIDRIERÍA	3607.33€ Contiene 1 subcapítulo
C14-PINTURAS	13762.04€ Contiene 1 subcapítulo
C15-EQUIPAMIENTO	4787.37€ Contiene 3 subcapítulos

Tratamiento Inteligente de Datos en Proyectos de Edificación
© Copyright 2015 - María Martínez Rojas

Figura 39: Pantalla con la lista de capítulos para el proyecto de ejemplo

Debajo de esta información aparece un menú desplegable que permite al usuario interactuar de manera sencilla con *SIGPE* en el sentido de editar el resultado obtenido en el proceso de clasificación tanto para los capítulos como para los subcapítulos (Figura 41).

A modo de ejemplo, en la Figura 42, se muestra una interacción sobre una partida que *SIGPE* ha clasificado en un capítulo erróneamente. Concretamente, la descripción de partida “*Barnizado de carpintería de madera interior o exterior con dos manos de barniz sintético brillante*”, la cual aparece en color naranja, ha sido clasificada en el subcapítulo de *ventanas* perteneciente al capítulo de *carpintería de madera* (Figura 42a). El resultado proporcionado por el clasificador no es correcto, ya que esta descripción debería estar almacenada en el subcapítulo de *tratamientos* correspondiente al capítulo de *pinturas*. En la Figura 42b, se muestra cómo se utiliza el menú desplegable para seleccionar el capítulo *C14 - Pinturas*, que es donde debería haberse clasificado la descripción de partida. A continuación, en el tercer paso (Figura 42c), se utiliza el menú desplegable y se selecciona el subcapítulo *C14SC2 -*

ADQUISICIÓN EDICIÓN RECUPERACIÓN			
C2-ACONDICIONAMIENTO DEL TERRENO	8845.71€	Contiene 4 subcapítulos	
C3-RED HORIZONTAL DE SANEAMIENTO	3448.66€	Contiene 5 subcapítulos	
C4-CIMENTACIONES	50773.54€	Contiene 4 subcapítulos	
C5-ESTRUCTURA	28550.66€	Contiene 2 subcapítulos	
C6-ALBAÑILERIA	66079.49€	Contiene 5 subcapítulos	
C7-CUBIERTAS	6889.86€	Contiene 2 subcapítulos	
C8-INSTALACIONES	31948.04€	Contiene 7 subcapítulos	
ELECTRICA E ILUMINACIÓN	6.403.78 €	Contiene 23 partidas	
ACERO CORR. PREFOR. B 500 S	C 0.64 P 1.46 T 0.93	<input checked="" type="checkbox"/> Acero corrugado B 500 S, preformado en taller y colocado en obra. Según EHE y CTE-SE-A. Seleccione un capítulo	
Gastos tramitación contratación suministro eléctrico	C 1.00 P 52.53 T 52.53	<input checked="" type="checkbox"/> Ud. Gastos tramitación contratación por Kie con la Compañía para el suministro al edificio desde sus redes de distribución, incluido derechos de acometida, enganche y verificación en la contratación de la póliza de abono. Seleccione un capítulo	
PLUZ SENCILLO SIMÓN 75	C 9.00 P 29.00 T 261.00	<input checked="" type="checkbox"/> Punto de luz sencillo realizado con tubo PVC corrugado de M 20(gp5) y conductor rígido de 1.5 mm2 de Cu, y aislamiento VV 750 V, incluyendo caja de registro, caja de mecanismo universal con tornillos, interruptor unipolar Simón serie 75, instalado. Seleccione un capítulo	
PLUZ CONMUTADO SIMÓN 75	C 6.00 P 39.39 T 236.34	<input checked="" type="checkbox"/> Punto conmutado sencillo realizado con tubo PVC corrugado de M 20(gp5) y conductor rígido de 1.5 mm2 de Cu, y aislamiento VV 750 V, incluyendo caja de registro, cajas de mecanismo universal con tornillos, conmutadores Simón serie 75, instalado. Seleccione un capítulo	
P.DOBLE CONMUTAD. SIMÓN 75	C 2.00 P 66.07 T 132.14	<input checked="" type="checkbox"/> Punto doble conmutado realizado con tubo PVC corrugado de M 20(gp5) y conductor rígido de 1.5 mm2 de Cu, y aislamiento VV 750 V, incluyendo caja de registro, cajas de mecanismo universal con tornillos, conmutadores Simón serie 75, instalado. Seleccione un capítulo	

Figura 40: Capítulo 8 desplegado

Tratamientos. Posteriormente, y de forma automática, la descripción de partida se clasifica en el capítulo y subcapítulo que se ha seleccionado y, como se puede observar en la Figura 42d, el fondo de la descripción de partida se ha cambiado a color blanco. Esto permite al usuario identificar las partidas que ya ha revisado, con el fin de agilizar la tarea de revisión. Para las partidas que están clasificadas correctamente, el usuario podrá ir validando el resultado del clasificador mediante el botón de “verificar” ubicado justo encima de los menús desplegables. Cuando se valida una partida, el color del fondo también cambia a color blanco. De este modo, el usuario puede realizar la tarea de validación en diferentes momentos en el tiempo, ya que este botón le permite tener conocimiento hasta que descripción de partida ha revisado.

Una vez que el usuario ha concluido la fase de validación a través del menú de edición, el presupuesto está almacenado en la estructura de referencia.

Item	C	P	T
PREPERCO PINO 70x35 mm.P/1 HOJA	1.00	12.51	12.51
PUERTA CHAPA LISA 70x200 GALV.	1.00	72.20	72.20
PUERTA 0.80x2.00 40/14 STD	1.00	148.58	148.58
VENTANAS	1.048.27 € Contiene 1 partida		
BARN.LADERA INT.BRILLANT2 MAN.	98.80	10.61	1.048.27
ARMARIOS	C12-Carpintería Exterior y Cerrajería		
FTE.ARM.CORR.LISO PINO PIPINTAR	C13-Vidriería		
	C14-Pinturas		
	C15-Equipamiento		

C12-CARPINTERÍA EXTERIOR Y CERRAJERÍA	47403.82€ Contiene 3 subcapítulos
C13-VIDRIERÍA	3607.33€ Contiene 1 subcapítulo
C14-PINTURAS	13762.04€ Contiene 1 subcapítulo
C15-EQUIPAMIENTO	4787.37€ Contiene 3 subcapítulos

Tratamiento Inteligente de Datos en Proyectos de Edificación
 © Copyright 2015 - María Martínez Rojas

Figura 41: Pantalla con menú desplegable para modificar el resultado

A.3 RECUPERACIÓN/CONSULTA

Mediante este módulo se pueden realizar consultas sobre todos los datos almacenados, por todos los usuarios, en el almacén de proyectos. Al tener los datos estructurados, puede resultar interesante consultar datos no solo de los proyectos de un usuario, sino de todos los proyectos almacenados en *SIGPE*.

La Figura 43 muestra la apariencia de este módulo, al cual se puede acceder a través del menú “Recuperación”.

En la parte de la izquierda se encuentran los tres tipos de filtros que se pueden aplicar para localizar la información deseada. El primero es el filtro por texto, que permite buscar descripciones de partidas que contengan los términos que el usuario desee. Como complemento a este filtro, se pueden aplicar los dos siguientes: cantidad y presupuesto. Este tipo de acceso ofrece soporte a los usuarios que realizan las

C11-CARPINTERÍA DE MADERA		10040.34€	
		Contiene 3 subcapítulos	
VENTANAS		1,048.27€ Contiene 1 partida	
BARNI.MADERA INT.BRILLANT.2 MAN.	C	P	T
Barnizado de carpintería de madera interior o exterior con dos manos de barniz sintético brillante.	98.80	10.61	1,048.27
	<input checked="" type="checkbox"/>		
Seleccione un capítulo			

(a)

C11-CARPINTERÍA DE MADERA		10040.34€	
		Contiene 3 subcapítulos	
VENTANAS		1,048.27€ Contiene 1 partida	
BARNI.MADERA INT.BRILLANT.2 MAN.	C	P	T
Barnizado de carpintería de madera interior o exterior con dos manos de barniz sintético brillante.	98.80	10.61	1,048.27
	<input checked="" type="checkbox"/>		
Seleccione un capítulo			
ARMARIOS	C12-Carpintería Exterior y Cerrajería		
	C13-Vidriería		
	C14-Pinturas		
FTE.ARM.CORR.LISO PINO P/PINTAR	C15-Equipamiento		
Frente de armario empotrado corredero, con hojas y maleteros lisos de pino del país para pintar de 11 mm. de espesor (A/MLM) con prececo de pino de 70x35 mm. con galce o cerco visto de pino rojo 70x35 mm. Incluye: pintura exterior mate de 2L, barnizado de pino 70x35...			

(b)

C11-CARPINTERÍA DE MADERA		10040.34€	
		Contiene 3 subcapítulos	
VENTANAS		1,048.27€ Contiene 1 partida	
BARNI.MADERA INT.BRILLANT.2 MAN.	C	P	T
Barnizado de carpintería de madera interior o exterior con dos manos de barniz sintético brillante.	98.80	10.61	1,048.27
	<input checked="" type="checkbox"/>		
Seleccione un capítulo			
ARMARIOS	C14-Pinturas		
	C14 SC1 Pinturas		
	C14 SC2 Tratamientos		
FTE.ARM.CORR.LISO PINO P/PINTAR	C14 SC3 Pinturas Especiales		
Frente de armario empotrado corredero, con hojas y maleteros lisos de pino del país para pintar de 11 mm. de espesor (A/MLM) con prececo de pino de 70x35 mm. con galce o cerco visto de pino rojo 70x35 mm. Incluye: pintura exterior mate de 2L, barnizado de pino 70x35...			

(c)

C14-PINTURAS		14810.31€	
		Contiene 2 subcapítulos	
TRATAMIENTOS		1,048.27€ Contiene 1 partida	
BARNI.MADERA INT.BRILLANT.2 MAN.	C	P	T
Barnizado de carpintería de madera interior o exterior con dos manos de barniz sintético brillante.	98.80	10.61	1,048.27
	<input checked="" type="checkbox"/>		
Seleccione un capítulo			

(d)

Figura 42: Procedimiento para la modificación de una descripción de partida

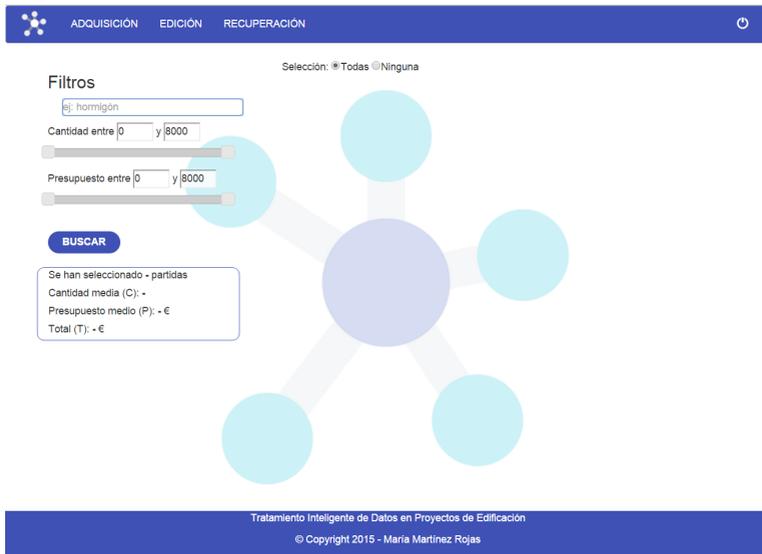


Figura 43: Apariencia del módulo de “Recuperación”

estimaciones de costes correspondientes a nuevos proyectos. De manera más rápida pueden buscar y consultar información relacionada con las descripciones de partida que han sido almacenadas por los distintos usuarios de la empresa.

Por ejemplo, si se quiere consultar el coste de la descripción de partida *“Membrana drenante Danodren H-25 plus de polietileno de alta densidad modulado, fijada al muro mediante rosetas Danodren y clavos de acero, con los nódulos contra el muro y solapes de 12 cm...”*, podemos introducir los siguientes términos en el campo de texto: *“Membrana, drenante, danodren”*. Una vez que se han establecido los filtros, se selecciona el botón *“Buscar”*. Como se puede observar en la Figura 44, en la parte de la derecha aparecen las descripciones que contienen los términos que se han solicitado. Además, SIGPE ofrece la opción de filtrar los resultados, de manera que se pueden visualizar todas o ninguna descripción de partida de las resultantes en el proceso de búsqueda. Para ello, simplemente bastaría con seleccionar una de las opciones de filtrado *todas* o *ninguna*.

En la parte de la izquierda, también se muestran algunos cálculos (cantidad media, coste medio y presupuesto medio) en base a las descripciones de partidas que han sido seleccionadas.

Este tipo de consulta permite acceder a los datos de todos los proyectos almacenados en el repositorio común de una manera rápida y

sencilla. Aunque este tipo de consultas pueden resultar simples, en la práctica actual no es tarea fácil resolverlas ya que, como se ha comentado en varias ocasiones en esta memoria, la información de los proyectos se encuentra dispersa en diferentes sistemas operacionales, ficheros o, incluso, en soporte papel. Para poder realizar una consulta similar a las que se pueden realizar en este módulo, el usuario necesita gran cantidad de tiempo porque tiene que localizar la información en los diversos documentos de los proyectos.

The screenshot shows the SIGPE application interface with the following elements:

- Navigation Bar:** ADQUISICIÓN, EDICIÓN, RECUPERACIÓN.
- Filtros (Filters):**
 - Search term: membrana drenante danodren
 - Quantity range: 0 to 8000
 - Budget range: 0 to 8000
 - SEARCH button
- Summary Box:**
 - Se han seleccionado 2 partidas
 - Cantidad media (C): 91.28
 - Presupuesto medio (P): 9.11 €
 - Total (T): 831.51 €
- Search Results Table:**

Selección: Todas Ninguna

	C	P	T
<input checked="" type="checkbox"/> MEM.DRENANTE P.E.A.D. VERT.H-25 PLUS Membrana drenante Danodren H-25 plus de polietileno de alta densidad nodulado, fijada al muro mediante rosetas Danodren y clavos de acero, con los nódulos contra el muro y solapes de 12 cm., lprotección del borde superior con perfil angular, sin incluir el tubo de drenaje inferior, ni el relleno ni la excavación de la zanja.	105.00	9.11	956.55
<input checked="" type="checkbox"/> MEM.DRENANTE P.E.A.D. VERT.H-25 PLUS Membrana drenante Danodren H-25 plus de polietileno de alta densidad nodulado, fijada al muro mediante rosetas Danodren y clavos de acero, con los nódulos contra el muro y solapes de 12 cm., lprotección del borde superior con perfil angular, sin incluir el tubo de drenaje inferior, ni el relleno ni la excavación de la zanja.	77.55	9.11	706.48
<input type="checkbox"/> IMPERMEAB. MURO MEMBRANA 4.8MM + LAMINA ABOTONA DE IMPERMEABILIZACION DE MURO FORMADA POR LAMINA DE MEMBRANA ASFALTICA TIPO MORTER-PLAS 4.8 FP DOBLE ARMADURA DE POLIETILENO CONTRAPEADAS A CUBREJUNTA Y SOLDADAS, GEOTEXTILTERRAN 1000, LAMINA DRENANTE ABOTONADA DRENTEX PROTECT PLUS, INCLUSO CAPAS DE MORTERO DE REGULARIZACION Y PROTECCION DE 2 cm. DE ESPESOR CON MORTERO M-4 (1:6)Y P.P. DE	564.33	19.80	11173.73

Figura 44: Ilustración de un ejemplo de búsqueda en SIGPE.

BIBLIOGRAFÍA

- [1] *Código técnico de la Edificación*. BOE-A-2006-5515, **2006**.
- [2] AAMODT, A. y PLAZA, E.: «Case-based reasoning; Foundational issues, methodological variations, and system approaches». *AI Communications*, **1994**, 7(1), pp. 39–59.
- [3] ABANDA, F.H.; TAH, J.H.M. y KEIVANI, R.: «Trends in built environment semantic Web applications: Where are we today?» *Expert Systems with Applications*, **2013**, 40(14), pp. 5563 – 5577. ISSN 0957-4174. doi: <http://dx.doi.org/10.1016/j.eswa.2013.04.027>.
- [4] ABANDA, F.H.; VIDALAKIS, C.; OTI, A.H. y TAH, J.H.M.: «A critical analysis of Building Information Modelling systems used in construction projects». *Advances in Engineering Software*, **2015**, 90, pp. 183 – 201. ISSN 0965-9978. doi: <http://dx.doi.org/10.1016/j.advengsoft.2015.08.009>.
- [5] ABUDAYYEH, O.; TEMEL, B.; AL-TABTABAI, H. y HURLEY, B.: «An Intranet-based cost control system». *Advances in Engineering Software*, **2001**, 32(2), pp. 87 – 94. ISSN 0965-9978. doi: [http://dx.doi.org/10.1016/S0965-9978\(00\)00094-6](http://dx.doi.org/10.1016/S0965-9978(00)00094-6).
- [6] ADRIAANSE, A.; VOORDIJK, H. y DEWULF, G.: «Adoption and Use of Interorganizational ICT in a Construction Project». *Journal of Construction Engineering and Management*, **2010**, 136(9), pp. 1003–1014. doi: 10.1061/(ASCE)CO.1943-7862.0000201.
- [7] AHMAD, I.; AZHAR, S. y LUKAUSKIS, P.: «Development of a decision support system using data warehousing to assist builders/developers in site selection». *Automation in Construction*, **2004**, 13(4), pp. 525 – 542. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2004.03.001>.
- [8] AHUJA, V.; YANG, J. y SHANKAR, R.: «Benchmarking Framework to Measure Extent of ICT Adoption for Building Project Management». *Journal of Construction Engineering and Management*, **2010**, 136(5), pp. 538–545. doi: 10.1061/(ASCE)CO.1943-7862.0000155.
- [9] AKINCI, B.; BOUKAMP, F.; GORDON, C.; HUBER, D.; LYONS, C. y PARK, K.: «A formalism for utilization of sensor systems and integrated project models for active construction quality control».

- Automation in Construction*, **2006**, 15(2), pp. 124 – 138. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2005.01.008>.
- [10] AL QADY, M. y KANDIL, A.: «Document Discourse for Managing Construction Project Documents». *Journal of Computing in Civil Engineering*, **2013**, 27(5), pp. 466–475. doi: 10.1061/(ASCE)CP.1943-5487.0000201.
- [11] AL QADY, M. y KANDIL, A.: «Automatic Classification of Project Documents on the Basis of Text Content». *Journal of Computing in Civil Engineering*, **2015**, 29(3), p. 04014043. doi: 10.1061/(ASCE)CP.1943-5487.0000338.
- [12] ANTONY-CHETTUPUZZHA, A.J. y HAAS, C.T.: «Algorithm for Determining the Criticality of Documents within a Construction Information System». *Journal of Computing in Civil Engineering*, **2015**, p. 04015039.
- [13] ANUMBA, C.J. y WANG, X.: *Mobile and Pervasive Computing in Construction*. Wiley, **2012**. ISBN 9781118422274.
- [14] AOUD, G.; LEE, A. y WU, S.: «From 3D to nD modelling». *ITcon - Special Issue From 3D to nD modelling*, **2005**, 10(0), pp. 15 – 16.
- [15] ARRIAGADA, R. y ALARCÓN, L.: «La organización en sitio y las oportunidades de gestionar conocimiento en las empresas de construcción». *Revista de la construcción*, **2011**, 10, pp. 86 – 98. ISSN 0718-915X.
- [16] ATKINSON, R.; CRAWFORD, L. y WARD, S.: «Fundamental uncertainties in projects and the scope of project management». *International Journal of Project Management*, **2006**, 24(8), pp. 687 – 698. ISSN 0263-7863. doi: <http://dx.doi.org/10.1016/j.ijproman.2006.09.011>.
- [17] AUTODESK: «BIM 360», **2015**.
<http://www.autodesk.com/products/bim-360/overview>
- [18] BAKHT, M.N. y EL-DIRABY, T.E.: «Synthesis of Decision-Making Research in Construction». *Journal of Construction Engineering and Management*, **2015**, p. 04015027.
- [19] BALLARD, G.: *The Last Planner system of production control*. Tesis doctoral, School of Civil Engineering. The University of Birmingham, **2000**.
- [20] BALOI, D. y PRICE, A.D.F.: «Modelling global risk factors affecting construction cost performance». *International Journal of Project Management*, **2003**, 21(4), pp. 261 – 269. ISSN 0263-7863. doi: [http://dx.doi.org/10.1016/S0263-7863\(02\)00017-0](http://dx.doi.org/10.1016/S0263-7863(02)00017-0).

- [21] BANSAL, V. y PAL, M.: «Generating, Evaluating, and Visualizing Construction Schedule with Geographic Information Systems». *Journal of Computing in Civil Engineering*, **2008**, 22(4), pp. 233-242. doi: 10.1061/(ASCE)0887-3801(2008)22:4(233).
- [22] BANSAL, V.K.: «Application of geographic information systems in construction safety planning». *International Journal of Project Management*, **2011**, 29(1), pp. 66 – 77. ISSN 0263-7863. doi: <http://dx.doi.org/10.1016/j.ijproman.2010.01.007>.
- [23] BANSAL, V.K. y PAL, M.: «Potential of geographic information systems in building cost estimation and visualization». *Automation in Construction*, **2007**, 16(3), pp. 311 – 322. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2006.07.002>.
- [24] BANSAL, V.K. y PAL, M.: «Construction schedule review in GIS with a navigable 3D animation of project activities». *International Journal of Project Management*, **2009**, 27(5), pp. 532 – 542. ISSN 0263-7863. doi: <http://dx.doi.org/10.1016/j.ijproman.2008.07.004>.
- [25] BASE DE DATOS ASTURIANA DE LA CONSTRUCCIÓN: «Base de precios». <http://www.fecea.org/home.html>, **2015**. [Online; accessed 05-October-2015].
- [26] BASE DE DATOS BEDEC: «Base de precios». www.itec.es, **2015**. [Online; accessed 05-October-2015].
- [27] BASE DE PRECIOS DE LA CONSTRUCCIÓN DE LA COMUNIDAD DE MADRID: «Base de precios». <http://www.madrid.org/bdccm/>, **2007**. [Online; accessed 05-October-2015].
- [28] BEHZADAN, A.H.; MENASSA, C.C.; TISHMAN, J.L. y PRADHAN, A.R.: «Enabling real time simulation of architecture, engineering, construction, and facility management (AEC/FM) systems: A review of formalism, model architecture, and data representation». *Journal of Information Technology in Construction*, **2015**, 20, pp. 1-23.
- [29] BENJAORAN, V.: «A cost control system development: A collaborative approach for small and medium-sized contractors». *International Journal of Project Management*, **2009**, 27(3), pp. 270 – 277. ISSN 0263-7863. doi: <http://dx.doi.org/10.1016/j.ijproman.2008.02.004>.
- [30] BERNERS-LEE, T.; HENDLER, J. y LASSILA, O.: «The semantic web». *Scientific American*, **2001**, 284(5), pp. 34-43. doi: 10.1038/scientificamericano501-34.

- [31] BERNERS-LEE, T.: «Design Note: Linked Data». <http://www.w3.org/DesignIssues/LinkedData.html>. [Online; accessed 19-October-2015].
- [32] BIM 9; 2015. <http://bim9.com>
- [33] BOSCHÉ, F.: «Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction». *Advanced Engineering Informatics*, 2010, 24(1), pp. 107 – 118. ISSN 1474-0346. doi: <http://dx.doi.org/10.1016/j.aei.2009.08.006>.
- [34] BOSCHÉ, F.; GUILLEMET, A.; TURKAN, Y.; HAAS, C.T. y HAAS, R.: «Tracking the Built Status of MEP Works: Assessing the Value of a Scan-vs-BIM System». *Journal of Computing in Civil Engineering*, 2014, 28(4), p. 05014004. doi: 10.1061/(ASCE)CP.1943-5487.0000343.
- [35] BOSCHÉ, F.; HAAS, C.T. y AKINCI, B.: «Automated Recognition of 3D CAD Objects in Site Laser Scans for Project 3D Status Visualization and Performance Control». *Journal of Computing in Civil Engineering*, 2009, 23(6), pp. 311–318. doi: 10.1061/(ASCE)0887-3801(2009)23:6(311).
- [36] BRAUN, A.; BORRMANN, A.; TUTTAS, S. y STILLA, U.: «Towards automated construction progress monitoring using BIM-based point cloud processing». *eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2014*, 2014, p. 101.
- [37] CADENAS, J. T.; MARÍN, N. y VILA, M. A.: «Context-Aware Fuzzy Databases». *Appl. Soft Comput.*, 2014, 25, pp. 215–233. doi: 10.1016/j.asoc.2014.09.020.
- [38] CALDAS, C.; SOIBELMAN, L. y HAN, J.: «Automated Classification of Construction Project Documents». *Journal of Computing in Civil Engineering*, 2002, 16(4), pp. 234–243. doi: 10.1061/(ASCE)0887-3801(2002)16:4(234).
- [39] CALDAS, C.H. y SOIBELMAN, L.: «Automating hierarchical document classification for construction management information systems». *Automation in Construction*, 2003, 12(4), pp. 395 – 406. ISSN 0926-5805. doi: [http://dx.doi.org/10.1016/S0926-5805\(03\)00004-9](http://dx.doi.org/10.1016/S0926-5805(03)00004-9).
- [40] CALDAS, C.H.; SOIBELMAN, L. y GASSER, L.: «Methodology for the integration of project documents in model-based information systems». *Journal of Computing in Civil Engineering*, 2005, 19(1), pp. 25–33.

- [41] CALDAS, C.H; TORRENT, D.G y HAAS, C.T.: «Integration of automated data collection technologies for real-time field materials management». En: *Proceedings of the 21st International Symposium on Automation and Robotics in Construction, Jeju, Korea, , 2004*.
- [42] CALDAS, C.H.; TORRENT, D.G. y HAAS, C.T.: «Using global positioning system to improve materials-locating processes on industrial projects». *Journal of Construction Engineering and Management*, **2006**.
- [43] CAO, D.; WANG, G.; LI, H.; SKITMORE, M.; HUANG, T. y ZHANG, W.: «Practices and effectiveness of building information modelling in construction projects in China». *Automation in Construction*, **2015**, 49, Part A(0), pp. 113 – 122. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2014.10.014>.
- [44] CARBONARI, A.; GIRETTI, A. y NATICCHIA, B.: «A proactive system for real-time safety management in construction sites». *Automation in Construction*, **2011**, 20(6), pp. 686 – 698. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2011.04.019>.
- [45] CARTER, G. y SMITH, S.: «Safety Hazard Identification on Construction Projects». *Journal of Construction Engineering and Management*, **2006**, 132(2), pp. 197–205. doi: 10.1061/(ASCE)0733-9364(2006)132:2(197).
- [46] CASTILLO-ORTEGA, R.; MARÍN, N. y SÁNCHEZ, D.: «Linguistic query answering on data cubes with time dimension». *Int. J. Intell. Syst.*, **2011**, 26(10), pp. 1002–1021. doi: 10.1002/int.20510. <http://dx.doi.org/10.1002/int.20510>
- [47] CASTILLO-ORTEGA, R.; MARÍN, N.; SÁNCHEZ, D. y MOLINA, C.: «Flexible Querying with Linguistic F-Cube Factory». En: *Flexible Query Answering Systems - 10th International Conference, FQAS 2013, Granada, Spain, September 18-20, 2013. Proceedings*, pp. 245–256. ISBN 978-3-642-40768-0, **2013**. doi: 10.1007/978-3-642-40769-7_22.
- [48] CASTRO-LACOUTURE, D.; SÜER, G.; GONZALEZ-JOAQUI, J. y YATES, J.: «Construction Project Scheduling with Time, Cost, and Material Restrictions Using Fuzzy Mathematical Models and Critical Path Method». *Journal of Construction Engineering and Management*, **2009**, 135(10), pp. 1096–1104. doi: 10.1061/(ASCE)0733-9364(2009)135:10(1096).
- [49] CHAE, S. y YOSHIDA, T.: «Application of RFID technology to prevention of collision accident with heavy equipment». *Automation*

- in Construction*, **2010**, 19(3), pp. 368 – 374. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2009.12.008>.
- [50] CHAN, A.; CHAN, D. y YEUNG, J.: «Overview of the Application of “Fuzzy Techniques” in Construction Management Research». *Journal of Construction Engineering and Management*, **2009**, 135(11), pp. 1241–1252. doi: 10.1061/(ASCE)CO.1943-7862.0000099.
- [51] CHAN, S. y LEUNG, N.: «Prototype Web-Based Construction Project Management System». *Journal of Construction Engineering and Management*, **2004**, 130(6), pp. 935–943. doi: 10.1061/(ASCE)0733-9364(2004)130:6(935).
- [52] CHASSIAKOS, A.P. y SAKELLAROPOULOS, S.P.: «A web-based system for managing construction information». *Advances in Engineering Software*, **2008**, 39(11), pp. 865 – 876. ISSN 0965-9978. doi: <http://dx.doi.org/10.1016/j.advengsoft.2008.05.006>.
- [53] CHAU, K.W.; CAO, YING; ANSON, M y ZHANG, JIANPING: «Application of data warehouse and Decision Support System in construction management». *Automation in Construction*, **2003**, 12(2), pp. 213 – 224. ISSN 0926-5805. doi: [http://dx.doi.org/10.1016/S0926-5805\(02\)00087-0](http://dx.doi.org/10.1016/S0926-5805(02)00087-0).
- [54] CHEN, L. y LUO, H.: «A BIM-based construction quality management model and its applications». *Automation in Construction*, **2014**, 46(0), pp. 64 – 73. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2014.05.009>.
- [55] CHENG, J.C.P.; DENG, Y. y ANUMBA, C.: «Mapping BIM schema and 3D GIS schema semi-automatically utilizing linguistic and text mining techniques». *Journal of Information Technology in Construction*, **2015**, 20, pp. 193–212.
- [56] CHENG, T. y TEIZER, J.: «Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications». *Automation in Construction*, **2013**, 34(0), pp. 3 – 15. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.10.017>.
- [57] CHENG, Y.; YU, W. y LI, Q.: «GA-based multi-level association rule mining approach for defect analysis in the construction industry». *Automation in Construction*, **2015**, 51(0), pp. 78 – 91. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2014.12.016>.
- [58] CHEUNG, F.K.T.; RIHAN, J.; TAH, J.; DUCE, D. y KURUL, E.: «Early stage multi-level cost estimation for schematic BIM models».

- Automation in Construction*, **2012**, 27(0), pp. 67 – 77. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.05.008>.
- [59] CHEUNG, S.O.; SUEN, H.C.H. y CHEUNG, K.K.W.: «PPMS: a Web-based construction Project Performance Monitoring System». *Automation in Construction*, **2004**, 13(3), pp. 361 – 376. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2003.12.001>.
- [60] CHI, S.; SUK, S.J.; KANG, Y. y MULVA, S.P.: «Development of a data mining-based analysis framework for multi-attribute construction project information». *Advanced Engineering Informatics*, **2012**, 26(3), pp. 574 – 581. ISSN 1474-0346.
- [61] CHIN, S.; KIM, K. y KIM, Y.S.: «A process-based quality management information system». *Automation in Construction*, **2004**, 13(2), pp. 241 – 259. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2003.08.010>.
- [62] CHO, D.; RUSSELL, J. y CHOI, J.: «Database Framework for Cost, Schedule, and Performance Data Integration». *Journal of Computing in Civil Engineering*, **2013**, 27(6), pp. 719–731. doi: [10.1061/\(ASCE\)CP.1943-5487.0000241](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000241).
- [63] CHONG, H.Y.; WONG, J.S. y WANG, X.: «An explanatory case study on cloud computing applications in the built environment». *Automation in Construction*, **2014**, 44(0), pp. 152 – 162. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2014.04.010>.
- [64] CHOU, J.S.: «Web-based CBR system applied to early cost budgeting for pavement maintenance project». *Expert Systems with Applications*, **2009**, 36(2, Part 2), pp. 2947 – 2960. ISSN 0957-4174. doi: <http://dx.doi.org/10.1016/j.eswa.2008.01.025>.
- [65] CHOU, J.S.: «Cost simulation in an item-based project involving construction engineering and management». *International Journal of Project Management*, **2011**, 29(6), pp. 706 – 717. ISSN 0263-7863. doi: <http://dx.doi.org/10.1016/j.ijproman.2010.07.010>.
- [66] CHUN, C.K.; LI, H. y SKITMORE, M.: «The use of virtual prototyping for hazard identification in the early design stage». *Construction Innovation*, **2012**, 12(1), pp. 29–42.
- [67] COLEGIO OFICIAL DE APAREJADORES, ARQUITECTOS TÉCNICOS E INGENIEROS DE EDIFICACIÓN DE GUADALAJARA: «Base de precios centro de la construcción». <http://www.preciocentro.com/>, **2012**. [Online; accessed 06-October-2015].

- [68] CONSEJERÍA DE FOMENTO, ORDENACIÓN DEL TERRITORIO Y TURISMO DEL GOBIERNO DE EXTREMADURA, VIVIENDA: «Base de precios de la construcción del Gobierno de Extremadura». <http://fomento.gobex.es/fomento/live/informacion-ciudadano/Arquitectura/baseprecios.html//>, 2012. [Online; accessed 7-Julio-2014].
- [69] CONSTRUCTION SPECIFICATIONS CANADA;, 2015. <https://www.csc-dcc.ca/>
- [70] CONSTRUCTION SPECIFICATIONS INSTITUTE:. «MasterFormat». <http://www.masterformat.com/>. [Online; accessed 07-Octubre-2015].
- [71] CONSTRUCTION SPECIFICATIONS INSTITUTE:. «UniFormat». <http://www.csinet.org/uniformat>. [Online; accessed 07-Octubre-2015].
- [72] CONSTRUCTION SPECIFICATIONS INSTITUTE;, 2015. <http://www.csinet.org/>
- [73] CUNNINGHAM, T.: «Factors Affecting The Cost of Building Work-An Overview», 2013.
- [74] CURRY, E.; O'DONNELL, J.; E.CORRY; HASAN, S.; KEANE, M. y O'RIAIN, S.: «Linking building data in the cloud: Integrating cross-domain building data using linked data». *Advanced Engineering Informatics*, 2013, 27(2), pp. 206 – 219. ISSN 1474-0346. doi: <http://dx.doi.org/10.1016/j.aei.2012.10.003>.
- [75] CYPE INGENIEROS: «Arquimedes. Programa de presupuestos y mediciones». <http://arquimedes.cype.es/>, 2014. [Online; accessed 24-Octubre-2014].
- [76] DELGADO, M.; MARÍN, N.; PÉREZ, Y. y VILA, M.A.: «Bipolar queries on fuzzy univalued and multivalued attributes in object databases». *Fuzzy Sets and Systems*, 2015.
- [77] DELGADO, M.; MOLINA, C.; RODRÍGUEZ-ARIZA, L.; SÁNCHEZ, D. y VILA, M.A.: «F-Cube Factory: a Fuzzy OLAP System for Supporting Imprecision». *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2007, 15(Supplement-1), pp. 59–81. doi: 10.1142/S0218488507004467.
- [78] DESHPANDE, A.; AZHAR, S. y AMIREDDY, S.: «A Framework for a BIM-based Knowledge Management System». *Procedia Engineering*, 2014, 85(0), pp. 113 – 122. ISSN 1877-7058. doi: <http://dx.doi.org/10.1016/j.proeng.2014.10.535>.

- [79] DIB, H.; ADAMO-VILLANI, N. y ISSA, R.A.: «A GIS-based Visual Information Model for Building Construction Project Management». *International Journal of Construction Management*, **2013**, 13(1), pp. 1–18. doi: 10.1080/15623599.2013.10773202.
- [80] DICK, K.: *XML: A Manager's Guide*. Addison-Wesley information technology series. Addison-Wesley, **2003**. ISBN 9780201770063.
- [81] DIKMEN, I.; BIRGONUL, M.T.; ANAC, C.; TAH, J.H.M. y AOUAD, G.: «Learning from risks: A tool for post-project risk assessment». *Automation in Construction*, **2008**, 18(1), pp. 42 – 50. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2008.04.008>.
- [82] DIMITROV, A. y GOLPARVAR-FARD, M.: «Vision-based material recognition for automated monitoring of construction progress and generating building information modeling from unordered site image collections». *Advanced Engineering Informatics*, **2014**, 28(1), pp. 37 – 49. ISSN 1474-0346. doi: <http://dx.doi.org/10.1016/j.aei.2013.11.002>.
- [83] DING, L.; ZHOU, Y. y AKINCI, B.: «Building Information Modeling BIM application framework: The process of expanding from 3D to computable nD». *Automation in Construction*, **2014**, 46(0), pp. 82 – 93. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2014.04.009>.
- [84] DOLOI, H.: «Cost Overruns and Failure in Project Management: Understanding the Roles of Key Stakeholders in Construction Projects». *Journal of Construction Engineering and Management*, **2013**, 139(3), pp. 267–279. doi: 10.1061/(ASCE)CO.1943-7862.0000621.
- [85] DU, J. y BORMANN, J.: «Improved Similarity Measure in Case-Based Reasoning with Global Sensitivity Analysis: An Example of Construction Quantity Estimating». *Journal of Computing in Civil Engineering*, **2014**, 28(6), p. 04014020. doi: 10.1061/(ASCE)CP.1943-5487.0000267.
- [86] DuBois, P.: *MySQL*. Developer's Library. Pearson Education, **2008**. ISBN 9780132704649.
- [87] EADIE, R.; BROWNE, M.; ODEYINKA, H.; MCKEOWN, C. y McNIFF, S.: «BIM implementation throughout the UK construction project lifecycle: An analysis». *Automation in Construction*, **2013**, 36(0), pp. 145 – 151. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2013.09.001>.

- [88] EL-DIRABY, T.A.; LIMA, C. y FEIS, B: «Domain taxonomy for construction concepts: toward a formal ontology for construction knowledge». *Journal of computing in civil engineering*, **2005**, 19(4), pp. 394–406.
- [89] EL-GOHARY, N.M. y EL-DIRABY, T.E.: «Dynamic knowledge-based process integration portal for collaborative construction». *Journal of Construction Engineering and Management*, **2009**, 136(3), pp. 316–328.
- [90] EL-GOHARY, N.M. y EL-DIRABY, T.E.: «Merging architectural, engineering, and construction ontologies». *Journal of Computing in Civil Engineering*, **2009**, 25(2), pp. 109–128.
- [91] EL-GOHARY, N.M. y EL-DIRABY, T.E.: «Domain ontology for processes in infrastructure and construction». *Journal of Construction Engineering and Management*, **2010**, 136(7), pp. 730–744.
- [92] ELBELTAGI, E. y DAWOOD, M.: «Integrated visualized time control system for repetitive construction projects». *Automation in Construction*, **2011**, 20(7), pp. 940 – 953. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2011.03.012>.
- [93] EMROUZNEJAD, A. y MARRA, M.: «Ordered Weighted Averaging Operators 1988–2014: A Citation-Based Literature Survey». *International Journal of Intelligent Systems*, **2014**, 29(11), pp. 994–1014.
- [94] FAN, H.; KIM, H. y ZAÏANE, O.R.: «Data warehousing for construction equipment management». *Canadian Journal of Civil Engineering*, **2006**, 33(12), pp. 1480–1489. doi: 10.1139/l06-108.
- [95] FAN, H. y LI, H.: «Retrieving similar cases for alternative dispute resolution in construction accidents using text mining techniques». *Automation in Construction*, **2013**, 34(0), pp. 85 – 91. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.10.014>.
- [96] FAN, H.; XUE, F. y LI, H.: «Project-Based As-Needed Information Retrieval from Unstructured AEC Documents». *Journal of Management in Engineering*, **2015**, 31(1), p. A4014012. doi: 10.1061/(ASCE)ME.1943-5479.0000341.
- [97] FATHI, M.S.; ABEDI, M.; RAMBAT, S.; RAWAI, S.; ZAKIYUDIN, M.Z. y otros: «Context-aware cloud computing for construction collaboration». *Journal of Cloud Computing*, **2012**, 2012, pp. 1–11.
- [98] FORCADA, N.; CASALS, M.; ROCA, X. y GANGOLELLS, M.: «Adoption of web databases for document management in SMEs of the construction sector in Spain». *Automation in Construction*, **2007**, 16(4), pp. 411 – 424. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2006.07.011>.

- [99] FÜR NORMUNG, DEUTSCHES INSTITUT: *DIN 276-1*, Germany, 2006.
- [100] FU, C.; AOUAD, G.; LEE, A.; MASHALL-PONTING, A. y WU, S.: «IFC model viewer to support nD model application». *Automation in Construction*, 2006, 15(2), pp. 178 – 185. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2005.04.002>.
- [101] GAJENDRAN, T. y BREWER, G.: «Cultural consciousness and the effective implementation of information and communication technology». *Construction Innovation*, 2012, 12(2), pp. 179–197. doi: 10.1108/14714171211215930.
- [102] GALINDO, J.: *Handbook of research on fuzzy information processing in databases*. IGI Global, 2008.
- [103] GIRETTI, A.; CARBONARI, A.; NATICCHIA, B. y DEGRASSI, M.: «Design and first development of an automated real-time safety management system for construction sites». *Journal of Civil Engineering and Management*, 2009, 15(4), pp. 325–336. doi: 10.3846/1392-3730.2009.15.325-336.
- [104] GÓMEZ-ROMERO, J.; BOBILLO, F.; ROS, M.; MOLINA-SOLANA, M.; RUIZ, M.D. y MARTÍN-BAUTISTA, M.J.: «A fuzzy extension of the semantic Building Information Model». *Automation in Construction*, 2015, 57, pp. 202 – 212. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2015.04.007>.
<http://www.sciencedirect.com/science/article/pii/S0926580515000850>
- [105] GOLPARVAR-FARD, M.; PEÑA-MORA, F.; ARBOLEDA, C.A. y LEE, S.: «Visualization of construction progress monitoring with 4D simulation model overlaid on time-lapsed photographs». *Journal of Computing in Civil Engineering*, 2009, 23(6), pp. 391–404.
- [106] GOLPARVAR-FARD, M.; PEÑA-MORA, F. y SAVARESE, S.: «Integrated Sequential As-Built and As-Planned Representation with D4AR Tools in Support of Decision-Making Tasks in the AEC/FM Industry». *Journal of Construction Engineering and Management*, 2011, 137(12), pp. 1099–1116. doi: 10.1061/(ASCE)CO.1943-7862.0000371.
- [107] GOLPARVAR-FARD, M.; PEÑA-MORA, F. y SAVARESE, S.: «Automated progress monitoring using unordered daily construction photographs and IFC-based building information models». *Journal of Computing in Civil Engineering*, 2015, 29(1), p. 04014025. doi: 10.1061/(ASCE)CP.1943-5487.0000205.

- [108] GOLPARVAR-FARD, M.; TANG, P.; CHO, Y. y SIDDIQUI, M.: *Grand Challenges in Data and Information Visualization for the Architecture, Engineering, Construction, and Facility Management Industries*. capítulo 106, pp. 849–856, **2013**. doi: 10.1061/9780784413029.106.
- [109] GUO, H.L.; LI, HENG y LI, VERA: «VP-based safety management in large-scale construction projects: A conceptual framework». *Automation in Construction*, **2013**, 34(0), pp. 16 – 24. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.10.013>.
- [110] GUTIERREZ, E.: *JavaScript: Conceptos básicos y avanzados (bibliotecas Prototype y Script.aculo.us)*. Recursos informáticos. ENI Editions, **2009**. ISBN 9782746052208.
- [111] HAN, K.K. y GOLPARVAR-FARD, M.: «Appearance-based material classification for monitoring of operation-level construction progress using 4D BIM and site photologs». *Automation in Construction*, **2015**, 53(0), pp. 44 – 57. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2015.02.007>.
- [112] HAN, S.; LEE, S. y PEÑA-MORA, F.: «Vision-Based Detection of Unsafe Actions of a Construction Worker: Case Study of Ladder Climbing». *Journal of Computing in Civil Engineering*, **2013**, 27(6), pp. 635–644. doi: 10.1061/(ASCE)CP.1943-5487.0000279.
- [113] HAN, S.; LEE, S. y PEÑA-MORA, F.: «Comparative Study of Motion Features for Similarity-Based Modeling and Classification of Unsafe Actions in Construction». *Journal of Computing in Civil Engineering*, **2014**, 28(5), p. A4014005. doi: 10.1061/(ASCE)CP.1943-5487.0000339.
- [114] HAN, S.H.; CHIN, K.H. y CHAE, M.J.: «Evaluation of CITIS as a collaborative virtual organization for construction project management». *Automation in Construction*, **2007**, 16(2), pp. 199 – 211. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2006.04.002>.
- [115] HAN, S.H.; KIM, D.Y.; H.KIM y JANG, W.S.: «A web-based integrated system for international project risk management». *Automation in Construction*, **2008**, 17(3), pp. 342 – 356. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2007.05.012>.
- [116] H.P.TSERNG; YIN, S.Y.L.; DZENG, R.J.; WOU, B.; TSAI, M.D. y CHEN, W.Y.: «A study of ontology-based risk management framework of construction projects through project life cycle». *Automation in Construction*, **2009**, 18(7), pp. 994 – 1008. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2009.05.005>.

- [117] HUANG, T.; KONG, C.W.; GUO, H.L.; BALDWIN, A. y LI, HENG: «A virtual prototyping system for simulating construction processes». *Automation in Construction*, **2007**, 16(5), pp. 576 – 585. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2006.09.007>.
- [118] INMON, W. H.: *Building the Data Warehouse*, **1996**.
- [119] INSTITUTO DE LA CONSTRUCCIÓN DE CASTILLA Y LEÓN: «Base de Precios de la construcción de Castilla y León». <http://www.basedeprecios.com/>, **2015**. [Online; accessed 05-Octubre-2015].
- [120] IRIZARRY, J. y KARAN, E.P.: «Optimizing location of tower cranes on construction sites through GIS and BIM integration». *Electronic Journal of Information Technology in Construction*, **2012**, 17, pp. 361–366.
- [121] IRIZARRY, J.; KARAN, E.P. y JALAEI, F.: «Integrating BIM and GIS to improve the visual monitoring of construction supply chain management». *Automation in Construction*, **2013**, 31(0), pp. 241 – 254. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.12.005>.
- [122] ISSA, U.H.: «Implementation of lean construction techniques for minimizing the risks effect on project construction time». *Alexandria Engineering Journal*, **2013**, 52(4), pp. 697 – 704. ISSN 1110-0168. doi: <http://dx.doi.org/10.1016/j.aej.2013.07.003>.
- [123] IYER, K.C. y JHA, K.N.: «Factors affecting cost performance: evidence from Indian construction projects». *International Journal of Project Management*, **2005**, 23(4), pp. 283 – 295. ISSN 0263-7863. doi: <http://dx.doi.org/10.1016/j.ijproman.2004.10.003>.
- [124] JANG, W.S. y SKIBNIEWSKI, M.J.: «Embedded System for Construction Asset Tracking Combining Radio and Ultrasound Signals». *Journal of Computing in Civil Engineering*, **2009**, 23(4), pp. 221–229. doi: 10.1061/(ASCE)0887-3801(2009)23:4(221).
- [125] JI, S.; PARK, M.; LEE, H.; AHN, J.; KIM, N. y SON, B.: «Military Facility Cost Estimation System Using Case-Based Reasoning in Korea». *Journal of Computing in Civil Engineering*, **2011**, 25(3), pp. 218–231. doi: 10.1061/(ASCE)CP.1943-5487.0000082.
- [126] JI, S.H.; PARK, M. y LEE, H.S.: «Cost estimation model for building projects using case-based reasoning». *Canadian Journal of Civil Engineering*, **2011**, 38(5), pp. 570–581. doi: 10.1139/l11-016.
- [127] JIANG, S.; JANG, W.S. y SKIBNIEWSKI, M.J.: «Selection of Wireless Technology for Tracking Construction Materials Using a Fuzzy

- Decision Model». *Journal of Civil Engineering and Management*, **2012**, 18(1), pp. 43–59. doi: 10.3846/13923730.2011.652157.
- [128] JIAO, Y.; WANG, Y.; ZHANG, S.; LI, Y.; YANG, B. y YUAN, L.: «A cloud approach to unified lifecycle data management in architecture, engineering, construction and facilities management: Integrating BIMs and SNS». *Advanced Engineering Informatics*, **2013**, 27(2), pp. 173 – 188. ISSN 1474-0346. doi: <http://dx.doi.org/10.1016/j.aei.2012.11.006>.
- [129] JUNTA DE ANDALUCÍA: «Base de precios de la Junta de Andalucía, howpublished = "http://www.juntadeandalucia.es/fomentoyvivienda/portal-web/web/areas/vivienda/texto/706e4686-1fbd-11e0-89b8-998a90d310ed/", note = "[Online; accessed 7-Julio-2014]"», **2010**.
- [130] KANAPECKIENE, L.; KAKLAUSKAS, A.; ZAVADSKAS, E.K. y SENTUT, M.: «Integrated knowledge management model and system for construction projects». *Engineering Applications of Artificial Intelligence*, **2010**, 23(7), pp. 1200 – 1215. ISSN 0952-1976. doi: <http://dx.doi.org/10.1016/j.engappai.2010.01.030>.
- [131] KANG, L.S.; KIM, S.K.; MOON, H.S. y KIM, H.S.: «Development of a 4D object-based system for visualizing the risk information of construction projects». *Automation in Construction*, **2013**, 31(0), pp. 186 – 203. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.11.038>.
- [132] KANG, T.W. y HONG, C.H.: «A study on software architecture for effective BIM/GIS-based facility management data integration». *Automation in Construction*, **2015**, 54(0), pp. 25 – 38. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2015.03.019>.
- [133] KARAN, E.P. y IRIZARRY, J.: «Extending BIM interoperability to preconstruction operations using geospatial analyses and semantic web services». *Automation in Construction*, **2015**, 53(0), pp. 1 – 12. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2015.02.012>.
- [134] KARAN, E.P.; IRIZARRY, J. y HAYMAKER, J.: «BIM and GIS Integration and Interoperability Based on Semantic Web Technology». *Journal of Computing in Civil Engineering*, **2015**, 0(0), p. 04015043. doi: 10.1061/(ASCE)CP.1943-5487.0000519.
- [135] KARAN, E.P.; SIVAKUMAR, R.; IRIZARRY, J. y GUHATHAKURTA, S.: «Digital Modeling of Construction Site Terrain Using Remotely Sensed Data and Geographic Information Systems Analyses».

- Journal of Construction Engineering and Management*, **2013**, 140(3), p. 04013067.
- [136] KIM, C.; SON, H. y KIM, C.: «Automated construction progress measurement using a 4D building information model and 3D data». *Automation in Construction*, **2013**, 31(0), pp. 75 – 82. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.11.041>.
- [137] KIM, H.; LEE, H.S.; PARK, M.; CHUNG, B. y HWANG, S.: «Information Retrieval Framework for Hazard Identification in Construction». *Journal of Computing in Civil Engineering*, **2013**, 29(3), p. 04014052.
- [138] KIM, H.J.; SEO, Y.C. y HYUN, C.T.: «A hybrid conceptual cost estimating model for large building projects». *Automation in Construction*, **2012**, 25(0), pp. 72 – 81. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.04.006>.
- [139] KIM, K.J. y KIM, K.: «Preliminary cost estimation model using case-based reasoning and genetic algorithms». *Journal of Computing in Civil Engineering*, **2010**, 24(6), pp. 499–505.
- [140] KIM, Y.S.; OH, S.W.; CHO, Y.K. y SEO, J.W.: «A PDA and wireless web-integrated system for quality inspection and defect management of apartment housing projects». *Automation in Construction*, **2008**, 17(2), pp. 163 – 179. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2007.03.006>.
- [141] KIMBALL, R.: *The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses*. John Wiley, **1996**. ISBN 0-471-15337-0.
- [142] KIVRAK, S.; ARSLAN, G.; DIKMEN, I. y BIRGONUL, M.: «Capturing Knowledge in Construction Projects: Knowledge Platform for Contractors». *Journal of Management in Engineering*, **2008**, 24(2), pp. 87–95. doi: 10.1061/(ASCE)0742-597X(2008)24:2(87).
- [143] KNIGHT, K. y ROBINSON FAYEK, A.: «Use of Fuzzy Logic for Predicting Design Cost Overruns on Building Projects». *Journal of Construction Engineering and Management*, **2002**, 128(6), pp. 503–512. doi: 10.1061/(ASCE)0733-9364(2002)128:6(503).
- [144] KOSKELA, L.: *Application of the new production philosophy to construction*. 72. Stanford University (Technical Report No. 72, Center for Integrated Facility Engineering, Department of Civil Engineering). Stanford, CA, **1992**.

- [145] KU, KIHONG y TAIEBAT, MOJTABA: «BIM Experiences and Expectations: The Constructors' Perspective». *International Journal of Construction Education and Research*, **2011**, 7(3), pp. 175–197. doi: 10.1080/15578771.2010.544155.
- [146] KWON, O.S.; PARK, C.S. y LIM, C.R.: «A defect management system for reinforced concrete work utilizing BIM, image-matching and augmented reality». *Automation in Construction*, **2014**, 46(0), pp. 74 – 81. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2014.05.005>.
- [147] LAM, K.C y NG, S.T.: «A cooperative Internet-facilitated quality management environment for construction». *Automation in Construction*, **2006**, 15(1), pp. 1 – 11. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2005.01.009>.
- [148] LAM, P.; WONG, F. y TSE, K.: «Effectiveness of ICT for Construction Information Exchange among Multidisciplinary Project Teams». *Journal of Computing in Civil Engineering*, **2010**, 24(4), pp. 365–376. doi: 10.1061/(ASCE)CP.1943-5487.0000038.
- [149] LAWRENCE, M.; POTTINGER, R.; STAUB-FRENCH, S. y NEPAL, M.P.: «Creating flexible mappings between Building Information Models and cost information». *Automation in Construction*, **2014**, 45(0), pp. 107 – 118. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2014.05.006>.
- [150] LE, Q.T.; LEE, D.Y. y PARK, C.S.: «A social network system for sharing construction safety and health knowledge». *Automation in Construction*, **2014**, 46(0), pp. 30 – 37. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2014.01.001>.
- [151] LEE, H.S.; LEE, K.P.; PARK, M.; BAEK, Y. y LEE, S.: «RFID-based real-time locating system for construction safety management». *Journal of Computing in Civil Engineering*, **2011**, 26(3), pp. 366–377.
- [152] LEE, S.K.; KIM, K.R. y Yu, J.H.: «BIM and ontology-based approach for building cost estimation». *Automation in Construction*, **2014**, 41(0), pp. 96 – 105. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2013.10.020>.
- [153] LEY DE ORDENACIÓN DE LA EDIFICACIÓN: «Boletín Oficial del estado». *Informe técnico páginas. 38925–38934*, Gobierno de España, **6 de noviembre de 1999**.
- [154] L.F. ALARCÓN CÁRDENAS, E. PELLICER ARMIÑANA: «Un nuevo enfoque en la gestión, la construcción sin pérdidas». *Revista de obras públicas*, **2013**, (3496), pp. 45 – 52. ISSN 0034-8619.

- [155] LI, H.; CHAN, N.; HUANG, T.; GUO, H.L.; LU, W. y SKITMORE, M.: «Optimizing construction planning schedules by virtual prototyping enabled resource analysis». *Automation in Construction*, **2009**, 18(7), pp. 912 – 918. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2009.04.002>.
- [156] LIM, T.K.; YI, C.Y.; LEE, D.E y ARDITI, D.: «Concurrent Construction Scheduling Simulation Algorithm». *Computer-Aided Civil and Infrastructure Engineering*, **2014**, 29(6), pp. 449–463. ISSN 1467-8667. doi: 10.1111/mice.12073.
- [157] LIMA, C.; EL-DIRABY, T. y STEPHENS, J.: «Ontology-based optimization of knowledge management in e-construction». *ITcon*, **2005**, 10(2005), pp. 305–327.
- [158] LIN, Y.C.: «Developing construction assistant experience management system using people-based maps». *Automation in Construction*, **2008**, 17(8), pp. 975 – 982. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2008.04.004>.
- [159] LIN, Y.C.: «Construction 3D BIM-based knowledge management system: a case study». *Journal of Civil Engineering and Management*, **2014**, 20(2), pp. 186–200.
- [160] LIN, Y.C. y LEE, H.Y.: «Developing project communities of practice-based knowledge management system in construction». *Automation in Construction*, **2012**, 22(0), pp. 422 – 432. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2011.10.004>.
- [161] LIN, Y.C; WANG, L.C. y TSENG, H.P.: «Enhancing knowledge exchange through web map-based knowledge management system in construction: Lessons learned in Taiwan». *Automation in Construction*, **2006**, 15(6), pp. 693 – 705. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2005.09.006>.
- [162] LIU, H.; AL-HUSSEIN, M. y LU, M.: «BIM-based integrated approach for detailed construction scheduling under resource constraints». *Automation in Construction*, **2015**, 53(0), pp. 29 – 43. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2015.03.008>.
- [163] LIU, K. y GOLPARVAR-FARD, M.: «Crowdsourcing construction activity analysis from jobsite video streams». *Journal of Construction Engineering and Management*, **2015**, p. 04015035.
- [164] LU, Y.; LI, Y.; SKIBNIEWSKI, M.; WU, Z.; WANG, R. y LE, Y.: «Information and Communication Technology Applications in Architecture, Engineering, and Construction Organizations: A 15-Year Review». *Journal of Management in Engineering*, **2014**, 31(1).

- [165] MAALEK, R. y SADEGHPOUR, F.: «Accuracy assessment of Ultra-Wide Band technology in tracking static resources in indoor construction scenarios». *Automation in Construction*, **2013**, 30(0), pp. 170 – 183. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.10.005>.
- [166] MANASE, D.; HEESOM, D.; OLOKE, D.; PROVERBS, D.; YOUNG, C. y LUCKHURST, D.: «A GIS analytical approach for exploiting construction health and safety information». *Journal of Information Technology in Construction*, **2011**, 16, pp. 335–356.
- [167] MARTINEZ, G.; RUBIO, M.; MENENDEZ, A. y RUBIO, J.: «Gestión de la prevención de riesgos laborales en las obras de infraestructuras de transporte». *Estudios de construcción y transportes*, **2004**, (100), pp. 8–106.
- [168] MARTÍNEZ-ROJAS, M.; MARÍN, N.; MOLINA, CARLOS y VILA, M.A.: «Cost Analysis in Construction Projects using Fuzzy OLAP Cubes». En: *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2015, Istanbul, Turkey*, , **2015**.
- [169] MARTÍNEZ-ROJAS, M.; MARÍN, N. y VILA, M.A.: «Sistemas de Data Warehousing en el Ámbito de la Construcción». En: *Actas de las II Jornadas Andaluzas de Informática*, , **2011**.
- [170] MARTÍNEZ-ROJAS, M.; RUIZ, N. MARÍN y VILA, M.A.: «Applying ICTs to prevent accidents in construction sector». En: *International Congress on Safety and Labour Market*, , **2013**.
- [171] MARTÍNEZ-ROJAS, M.; RUIZ, N. MARÍN y VILA, M.A.: «A preliminary approach to classify work descriptions in construction projects». En: *IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS), 2013 Joint*, pp. 1090–1095, **2013**. doi: 10.1109/IFSA-NAFIPS.2013.6608552.
- [172] MARTÍNEZ-ROJAS, M.; MARÍN, N. y VILA, M.A.: «Aplicación de las TICs en el Ámbito de la Construcción». *Iniciación a la Investigación*, **2013**, 1(5), pp. 1–9. ISSN 1988-415X.
- [173] MARTÍNEZ-ROJAS, M.; MARÍN, N. y VILA, M.A.: «An Approach for the Automatic Classification of Work Descriptions in Construction Projects». *Computer-Aided Civil and Infrastructure Engineering*, **2015**, 0(0). doi: {Inpress}.
- [174] MARTÍNEZ-ROJAS, M.; MARÍN, N. y VILA, M.A.: «The Role of Information Technologies to Address Data Handling in Construction Project Management». *Journal of Computing in Civil Engineering*, **2015**, 0(0), p. 04015064. doi: 10.1061/(ASCE)CP.1943-5487.0000538.

- [175] MEMON, A.H.; RAHMAN, I.A.; ABDULLAH, M.R. y AZIS, A.A.: «Factors affecting construction cost performance in project management projects: Case of MARA large projects». *International Journal of Civil Engineering and Built Environment*, **2014**, 1(1).
- [176] MICROSOFT.: «Excel». <https://products.office.com/es-es/excel>.
- [177] MIKULAKOVA, E.; KÖNIG, M.; TAUSCHER, E. y BEUCKE, K.: «Knowledge-based schedule generation and evaluation». *Advanced Engineering Informatics*, **2010**, 24(4), pp. 389–403.
- [178] MOLINA, C.; ARIZA, L. RODRÍGUEZ; SÁNCHEZ, D. y MIRANDA, M.A. VILA: «A New Fuzzy Multidimensional Model». *IEEE T. Fuzzy Systems*, **2006**, 14(6), pp. 897–912. doi: 10.1109/TFUZZ.2006.879984.
- [179] MONTEIRO, A. y MARTINS, J.P.: «A survey on modeling guidelines for quantity takeoff-oriented BIM-based design». *Automation in Construction*, **2013**, 35(0), pp. 238 – 253. ISSN 0926-5805.
- [180] MOON, H.; KIM, H.; KAMAT, V. y KANG, L.: «BIM-Based Construction Scheduling Method Using Optimization Theory for Reducing Activity Overlaps». *Journal of Computing in Civil Engineering*, **2013**, 0(0), p. 04014048. doi: 10.1061/(ASCE)CP.1943-5487.0000342.
- [181] MOON, S.W.; KIM, J.S. y KWON, K.N.: «Effectiveness of OLAP-based cost data management in construction cost estimate». *Automation in Construction*, **2007**, 16(3), pp. 336 – 344. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2006.07.008>.
- [182] NATICCHIA, B.; VACCARINI, M. y CARBONARI, A.: «A monitoring system for real-time interference control on large construction sites». *Automation in Construction*, **2013**, 29(0), pp. 148 – 160. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.09.016>.
- [183] NIE, H.H.; STAUB-FRENCH, S. y FROESE, T.: «OLAP-Integrated Project Cost Control and Manpower Analysis». *Journal of Computing in Civil Engineering*, **2007**, 21(3), pp. 164–174. doi: 10.1061/(ASCE)0887-3801(2007)21:3(164).
- [184] NIETO-MOROTE, A. y RUZ-VILA, F.: «A fuzzy approach to construction project risk assessment». *International Journal of Project Management*, **2011**, 29(2), pp. 220 – 231. ISSN 0263-7863. doi: 10.1016/j.ijproman.2010.02.002.

- [185] NIKNAM, M. y K., SAEED: «Integrating distributed sources of information for construction cost estimating using Semantic Web and Semantic Web Service technologies». *Automation in Construction*, **2015**. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2015.04.003>.
- [186] NIST (NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY): «The NIST definition of cloud computing», **2011**. <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>
- [187] OLAP COUNCIL: <http://www.olapcouncil.org/>, **2015**. [Online; accessed 16-Septiembre-2015].
- [188] PADRÓ, L.: «Analizadores Multilingües en FreeLing». *Linguamática*, **2011**, 3(2), pp. 13–20.
- [189] PARK, C.S.; LEE, D.Y.; KWON, O.S. y WANG, X.: «A framework for proactive construction defect management using BIM, augmented reality and ontology-based data collection template». *Automation in Construction*, **2013**, 33(0), pp. 61 – 71. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.09.010>.
- [190] PARK, M.; LEE, K.; LEE, H.; JIAYI, P. y YU, J.: «Ontology-based construction knowledge retrieval system». *KSCE Journal of Civil Engineering*, **2013**, 17(7), pp. 1654–1663. ISSN 1226-7988. doi: [10.1007/s12205-013-1155-6](https://doi.org/10.1007/s12205-013-1155-6).
- [191] PARK, M.W.; KOCH, C. y BRILAKIS, I.: «Three-dimensional tracking of construction resources using an on-site camera system». *Journal of Computing in Civil Engineering*, **2011**.
- [192] POKU, S. y ARDITI, D.: «Construction Scheduling and Progress Control Using Geographical Information Systems». *Journal of Computing in Civil Engineering*, **2006**, 20(5), pp. 351–360. doi: [10.1061/\(ASCE\)0887-3801\(2006\)20:5\(351\)](https://doi.org/10.1061/(ASCE)0887-3801(2006)20:5(351)).
- [193] PRADHANANGA, N. y TEIZER, J.: «Automatic spatio-temporal analysis of construction site equipment operations using GPS data». *Automation in Construction*, **2013**, 29(0), pp. 107 – 122. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.09.004>.
- [194] PREMEDI: «Programa de mediciones». <http://www.preoc.es/>, **2014**. [Online; accessed 23-Septiembre-2014].
- [195] PREOC: «Base de precios». <http://www.preoc.es/>, **2010**. [Online; accessed 7-Julio-2014].

- [196] PRESTO: «Gestión de costes de la construcción». <http://www.soft.es/>, 2015. [Online; accessed 5-Abril-2015].
- [197] PRINCETON UNIVERSITY. 2010:. «WordNet». <http://wordnet.princeton.edu>. [Online; accessed 19-October-2015].
- [198] QADY, M. AL y KANDIL, A.: «Automatic clustering of construction project documents based on textual similarity». *Automation in Construction*, 2014, 42(0), pp. 36 – 49. ISSN 0926-5805.
- [199] QI, J.; ISSA, R.; OLBINA, S. y HINZE, J.: «Use of Building Information Modeling in Design to Prevent Construction Worker Falls». *Journal of Computing in Civil Engineering*, 2014, 28(5), p. A4014008. doi: 10.1061/(ASCE)CP.1943-5487.0000365.
- [200] RASMUS LERDORF.: «php». <http://www.php.net/>.
- [201] RAWAI, N.M.; FATHI, M.S.; ABEDI, M. y RAMBAT, S.: «Cloud Computing for Green Construction Management». En: *Intelligent System Design and Engineering Applications (ISDEA), 2013 Third International Conference on*, pp. 432–435, 2013. doi: 10.1109/ISDEA.2012.107.
- [202] RAZAVI, S.N. y HAAS, C.T.: «Reliability-Based Hybrid Data Fusion Method for Adaptive Location Estimation in Construction». *Journal of Computing in Civil Engineering*, 2012, 26(1), pp. 1–10. doi: 10.1061/(ASCE)CP.1943-5487.0000101.
- [203] REZGUI, Y.; BODDY, S.; WETHERILL, M. y COOPER, G.: «Past, present and future of information and knowledge sharing in the construction industry: Towards semantic service-based e-construction?» *Computer-Aided Design*, 2011, 43(5), pp. 502 – 515. ISSN 0010-4485. doi: <http://dx.doi.org/10.1016/j.cad.2009.06.005>.
- [204] REZGUI, Y. y ZARLI, A.: «Paving the way to the vision of digital construction: A strategic roadmap». *Journal of Construction Engineering and Management-ASCE*, 2006, 132(7), pp. 767–776. ISSN 0733-9364. doi: {10.1061/(ASCE)0733-9364(2006)132:7(767)}.
- [205] RIVAS, T.; PAZ, M.; MARTÍN, J.E.; MATÍAS, J.M.; GARCÍA, J.F. y TABOADA, J.: «Explaining and predicting workplace accidents using data-mining techniques». *Reliability Engineering & System Safety*, 2011, 96(7), pp. 739 – 747. ISSN 0951-8320. doi: <http://dx.doi.org/10.1016/j.res.2011.03.006>.
- [206] ROYAL INSTITUTE OF CHARTERED SURVEYORS (RICS) QUANTITY SURVEYING AND CONSTRUCTION PROFESSIONAL GROUP: «New

- rules of measurement 2 (NRM2)», 2015.
<http://www.nrm2.org/>
- [207] RUJIRAYANYONG, T. y SHI, J.J.: «A project-oriented data warehouse for construction». *Automation in Construction*, 2006, 15(6), pp. 800 – 807. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2005.11.001>.
- [208] SARDROUD, J.M.: «Perceptions of automated data collection technology use in the construction industry». *Journal of Civil Engineering and Management*, 2015, 21(1), pp. 54–66. doi: [10.3846/13923730.2013.802734](https://doi.org/10.3846/13923730.2013.802734).
- [209] SEO, J.; HAN, S.; LEE, S. y KIM, H.: «Computer vision techniques for construction safety and health monitoring». *Advanced Engineering Informatics*, 2015, 29(2), pp. 239 – 251. ISSN 1474-0346. doi: <http://dx.doi.org/10.1016/j.aei.2015.02.001>.
- [210] SERPELLA, A.F.; FERRADA, X.; HOWARD, R. y RUBIO, L.: «Risk Management in Construction Projects: A Knowledge-based Approach». *Procedia - Social and Behavioral Sciences*, 2014, 119(0), pp. 653 – 662. ISSN 1877-0428. doi: <http://dx.doi.org/10.1016/j.sbspro.2014.03.073>.
- [211] SHAHI, A.; SAFA, M.; HAAS, C.T. y WEST, J.S.: «Data Fusion Process Management for Automated Construction Progress Estimation». *Journal of Computing in Civil Engineering*, 2014, 0(0), p. 04014098. doi: [10.1061/\(ASCE\)CP.1943-5487.0000436](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000436).
- [212] SHANE, J.; MOLENAAR, K.; ANDERSON, S. y SCHEXNAYDER, C.: «Construction Project Cost Escalation Factors». *Journal of Management in Engineering*, 2009, 25(4), pp. 221–229. doi: [10.1061/\(ASCE\)0742-597X\(2009\)25:4\(221\)](https://doi.org/10.1061/(ASCE)0742-597X(2009)25:4(221)).
- [213] SHE, T. H.; y SARSHAR, M.: «A Geographic Information System GIS Based Bridge Management System». *Computer-Aided Civil and Infrastructure Engineering*, 1999, 14(6), pp. 417–427. ISSN 1467-8667. doi: [10.1111/0885-9507.00160](https://doi.org/10.1111/0885-9507.00160).
- [214] SHEN, W.; HAO, Q.; MAK, H.; NEELAMKAVIL, J.; XIE, H.; DICKINSON, J.; THOMAS, R.; PARDASANI, A. y XUE, H.: «Systems integration and collaboration in architecture, engineering, construction, and facilities management: A review». *Advanced Engineering Informatics*, 2010, 24(2), pp. 196 – 207. ISSN 1474-0346. doi: <http://dx.doi.org/10.1016/j.aei.2009.09.001>.
- [215] SOIBELMAN, L. y KIM, H.: «Data preparation process for construction knowledge generation through knowledge discovery in

- databases». *Journal of Computing in Civil Engineering*, **2002**, 16(1), pp. 39–48.
- [216] SOIBELMAN, L.; WU, J.; CALDAS, C.; BRILAKIS, I. y LIN, K.Y.: «Management and analysis of unstructured construction data types». *Advanced Engineering Informatics*, **2008**, 22(1), pp. 15 – 27. ISSN 1474-0346. doi: <http://dx.doi.org/10.1016/j.aei.2007.08.011>.
- [217] SONG, L. y ABOURIZK, S.: «Measuring and Modeling Labor Productivity Using Historical Data». *Journal of Construction Engineering and Management*, **2008**, 134(10), pp. 786–794. doi: [10.1061/\(ASCE\)0733-9364\(2008\)134:10\(786\)](https://doi.org/10.1061/(ASCE)0733-9364(2008)134:10(786)).
- [218] SONG, S.; YANG, J. y KIM, N.: «Development of a BIM-based structural framework optimization and simulation system for building construction». *Computers in Industry*, **2012**, 63(9), pp. 895 – 912. ISSN 0166-3615. doi: <http://dx.doi.org/10.1016/j.compind.2012.08.013>.
- [219] STAUB-FRENCH, SHERYL; FISCHER, MARTIN; KUNZ, JOHN; PAULSON, BOYD y ISHII, KOS: «An ontology for relating features of building product models with construction activities to support cost estimating», **2002**.
- [220] TAN, H.; CARRILLO, P. y ANUMBA, C.: «Case Study of Knowledge Management Implementation in a Medium-Sized Construction Sector Firm». *Journal of Management in Engineering*, **2012**, 28(3), pp. 338–347. doi: [10.1061/\(ASCE\)ME.1943-5479.0000109](https://doi.org/10.1061/(ASCE)ME.1943-5479.0000109).
- [221] TIBAUT, A. y JAKOSA, D.: «Development of Knowledge Model for Construction Projects». En: Lorna Uden; Marjan Hericko y I-Hsien Ting (Eds.), *Knowledge Management in Organizations*, volumen 224 de *Lecture Notes in Business Information Processing*, pp. 248–259. Springer International Publishing, **2015**. ISBN 978-3-319-21008-7. doi: [10.1007/978-3-319-21009-4_19](https://doi.org/10.1007/978-3-319-21009-4_19).
http://dx.doi.org/10.1007/978-3-319-21009-4_19
- [222] TOOLE, T.: «Construction Site Safety Roles». *Journal of Construction Engineering and Management*, **2002**, 128(3), pp. 203–210. doi: [10.1061/\(ASCE\)0733-9364\(2002\)128:3\(203\)](https://doi.org/10.1061/(ASCE)0733-9364(2002)128:3(203)).
- [223] TROST, S. y OBERLENDER, G.: «Predicting Accuracy of Early Cost Estimates Using Factor Analysis and Multivariate Regression». *Journal of Construction Engineering and Management*, **2003**, 129(2), pp. 198–204. doi: [10.1061/\(ASCE\)0733-9364\(2003\)129:2\(198\)](https://doi.org/10.1061/(ASCE)0733-9364(2003)129:2(198)).
- [224] TSERNG, H.P.; HO, S.P. y JAN, S.H.: «Developing BIM-assisted as-built schedule management system for general contractors».

- Journal of Civil Engineering and Management*, **2014**, 20(1), pp. 47–58. doi: 10.3846/13923730.2013.851112.
- [225] TSERNG, H.P. y LIN, Y.C.: «Developing an activity-based knowledge management system for contractors». *Automation in Construction*, **2004**, 13(6), pp. 781–802.
- [226] TURKAN, Y.; BOSCHE, F.; HAAS, C.T. y HAAS, R.: «Automated progress tracking using 4D schedule and 3D sensing technologies». *Automation in Construction*, **2012**, 22(0), pp. 414 – 421. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2011.10.003>.
- [227] TURNER, R.; SRIPADA, S. y REITER, E.: «Generating approximate geographic descriptions». En: *Empirical methods in natural language generation*, pp. 121–140. Springer, **2010**.
- [228] VAN DEN HELM, P.; BÖHMS, M. y VAN BERLO, L.: «IFC-based clash detection for the open-source BIMserver». En: *Computing in civil and building engineering, proceedings of the international conference. Nottingham University Press, Nottingham, UK*, volumen 30, p. 181, **2010**.
- [229] VAN OOSTEROM, P. y ZLATANOVA, S.: *Creating Spatial Information Infrastructures: Towards the Spatial Semantic Web*. CRC Press, **2008**. ISBN 9781420070729.
- [230] VANLANDE, R.; NICOLLE, C. y CRUZ, C.: «IFC and building lifecycle management». *Automation in Construction*, **2008**, 18(1), pp. 70 – 78. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2008.05.001>.
- [231] VASENEV, A.; HARTMANN, T. y DORÉE, A.G.: «A distributed data collection and management framework for tracking construction operations». *Advanced Engineering Informatics*, **2014**, 28(2), pp. 127 – 137. ISSN 1474-0346. doi: <http://dx.doi.org/10.1016/j.aei.2014.01.003>.
- [232] VIJAMAA, E. y PELTOMAA, I.: «Intensified construction process control using information integration». *Automation in Construction*, **2014**, 39(0), pp. 126 – 133. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2013.08.015>.
- [233] WALY, A. F y THABET, W.Y.: «A Virtual Construction Environment for preconstruction planning». *Automation in Construction*, **2003**, 12(2), pp. 139 – 154. ISSN 0926-5805. doi: [http://dx.doi.org/10.1016/S0926-5805\(02\)00047-X](http://dx.doi.org/10.1016/S0926-5805(02)00047-X).

- [234] WANG, L.C.: «Enhancing construction quality inspection and management using RFID technology». *Automation in Construction*, **2008**, 17(4), pp. 467 – 479. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2007.08.005>.
- [235] WANG, L.C.; LIN, Y.C. y LIN, P.H.: «Dynamic mobile RFID-based supply chain control and management system in construction». *Advanced Engineering Informatics*, **2007**, 21(4), pp. 377 – 390. ISSN 1474-0346. doi: <http://dx.doi.org/10.1016/j.aei.2006.09.003>.
- [236] W.D., YU y J.Y., HSU: «Content-based text mining technique for retrieval of CAD documents». *Automation in Construction*, **2013**, 31(0), pp. 65 – 74. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.11.037>.
- [237] WILLIAMS, T.P. y GONG, J.: «Predicting construction cost overruns using text mining, numerical data and ensemble classifiers». *Automation in Construction*, **2014**, 43(0), pp. 23 – 29. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2014.02.014>.
- [238] WORLD WIDE WEB CONSORTIUM.: «Cascading Style Sheets». <http://www.w3c.es/>.
- [239] WORLD WIDE WEB CONSORTIUM.: «HyperText Markup Language, versión 5». <http://www.w3c.es/>.
- [240] WORLD WIDE WEB CONSORTIUM.: «Linked Data», **2015**. <http://www.w3.org/standards/semanticweb/data>
- [241] WOYCHOWSKY, E.: *Ajax: Creating Web Pages with Asynchronous JavaScript and XML*. Bruce Perens' Open Source series. Prentice Hall, **2007**. ISBN 9780132272674.
- [242] WU, W.; YANG, H.; CHEW, D.A.S.; YANG, S.; GIBB, A.G.F. y LI, Q.: «Towards an autonomous real-time tracking system of near-miss accidents on construction sites». *Automation in Construction*, **2010**, 19(2), pp. 134 – 141. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2009.11.017>.
- [243] WU, W.; YANG, H.; LI, Q. y CHEW, D.: «An integrated information management model for proactive prevention of struck-by-falling-object accidents on construction sites». *Automation in Construction*, **2013**, 34(0), pp. 67 – 74. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.10.010>.
- [244] YAGER, R.: «Ordered Weighted Averaging Aggregation Operators in Multicriteria Decisionmaking.» *IEEE Transactions on Systems, Man and Cybernetics*, **1988**, 18(1), pp. 183–190. doi: 10.1109/21.87068.

- [245] YAGER, R.; KACPRZYK, J. y BELIAKOV, G.: *Recent developments in the ordered weighted averaging operators: theory and practice*. volumen 265. Springer Science & Business Media, **2011**.
- [246] YANG, H.; CHEW, D.A.S.; WU, W.; ZHOU, Z. y LI, Q.: «Design and implementation of an identification system in construction site safety for proactive accident prevention». *Accident Analysis & Prevention*, **2012**, 48(0), pp. 193 – 203. ISSN 0001-4575. doi: <http://dx.doi.org/10.1016/j.aap.2011.06.017>.
- [247] YANG, J.: «IT supported construction management: A comprehensive review and ongoing research». En: Huang, Y and Bao, T and Wang, H (Ed.), *Construction and urban planning, pts 1-4*, volumen 671-674 de *Advanced Materials Research*, pp. 1998–2004. ISBN 978-3-03785-661-1. ISSN 1022-6680, **2013**. doi: [10.4028/www.scientific.net/AMR.671-674.1998](http://dx.doi.org/10.4028/www.scientific.net/AMR.671-674.1998).
- [248] YANG, J.; PARK, M.W.; VELA, P.A. y GOLPARVAR-FARD, M.: «Construction performance monitoring via still images, time-lapse photos, and video streams: Now, tomorrow, and the future». *Advanced Engineering Informatics*, **2015**, 29(2), pp. 211 – 224. ISSN 1474-0346. doi: <http://dx.doi.org/10.1016/j.aei.2015.01.011>.
- [249] YILDIZ, A.E.; DIKMEN, I. y BIRGONUL, M.T.: «Using Expert Opinion for Risk Assessment: A Case Study of a Construction Project Utilizing a Risk Mapping Tool». *Procedia - Social and Behavioral Sciences*, **2014**, 119(0), pp. 519 – 528. ISSN 1877-0428.
- [250] YILDIZ, A.E.; I.D.; BIRGONUL, M.T.; ERCOSKUN, K. y ALTEN, S.: «A knowledge-based risk mapping tool for cost estimation of international construction projects». *Automation in Construction*, **2014**, 43(0), pp. 144 – 155. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2014.03.010>.
- [251] ZADEH, L.A.: «Fuzzy Sets». *Information and Control*, **1965**, 8(3), pp. 338–353.
- [252] ZADEH, L.A.: «The Concept of a Linguistic Variable and its Application to Aproximate Reasoning I, II and III», **1975**. *Information Sciences*, 8, pp. 199–251, 301–357; 9, pp. 43–80.
- [253] ZADEH, L.A.: «A new direction in AI: Toward a computational theory of perceptions». *AI magazine*, **2001**, 22(1), p. 73.
- [254] ZAINI, A.A.; ENDUT, I.R. y SHEHU, Z.: «A preliminary study of the subcontractor's risk identification for the construction projects». En: *Business, Engineering and Industrial Applications (ISBEIA), 2012 IEEE Symposium on*, pp. 878–883, **2012**. doi: [10.1109/ISBEIA.2012.6423018](http://dx.doi.org/10.1109/ISBEIA.2012.6423018).

- [255] ZAYED, T. y WAKIL, E. EL: «Data Management for Construction Processes Using Fuzzy Approach». En: *Construction Research Congress 2012*, pp. 1222–1231, 2012. doi: 10.1061/9780784412329.123.
- [256] ZHANG, J. y EL-DIRABY, T.E.: «Social semantic approach to support communication in AEC». *Journal of Computing in Civil Engineering*, 2011, 26(1), pp. 90–104.
- [257] ZHANG, S.; BOUKAMP, F. y TEIZER, J.: «Ontology-based semantic modeling of construction safety knowledge: Towards automated safety planning for job hazard analysis JHA». *Automation in Construction*, 2015, 52(0), pp. 29 – 41. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2015.02.005>.
- [258] ZHANG, S.; SULANKIVI, K.; KIVINIEMI, M.; ROMO, I.; EASTMAN, C.M. y TEIZER, J.: «BIM-based fall hazard identification and prevention in construction safety planning». *Safety Science*, 2015, 72(0), pp. 31 – 45. ISSN 0925-7535. doi: <http://dx.doi.org/10.1016/j.ssci.2014.08.001>.
- [259] ZHANG, S.; TEIZER, J.; LEE, J.K.; EASTMAN, C.M. y VENUGOPAL, M.: «Building Information Modeling BIM and Safety: Automatic Safety Checking of Construction Models and Schedules». *Automation in Construction*, 2013, 29(0), pp. 183 – 195. ISSN 0926-5805. doi: <http://dx.doi.org/10.1016/j.autcon.2012.05.006>.
- [260] ZHONG, B. y LI, Y.: «An Ontological and Semantic Approach for the Construction Risk Inferring and Application». *Journal of Intelligent and Robotic Systems*, 2014, pp. 1–15. ISSN 0921-0296. doi: 10.1007/s10846-014-0107-9.
- [261] ZHU, Z. y BRILAKIS, I.: «Machine vision-based concrete surface quality assessment». *Journal of Construction Engineering and Management*, 2009, 136(2), pp. 210–218.