

LENGOAIA ETA SISTEMA INFORMATIKOAK SAILA
DEPARTAMENTO DE LENGUAJES Y SISTEMAS INFORMÁTICOS



Universidad del País Vasco Euskal Herriko Unibertsitatea

Análisis y mejora de los marcos actuales de desarrollo y evaluación de los Trabajos Fin de Grado mediante el uso de las TIC

MEMORIA

que para optar al Grado de Doctor en Informática presenta

Mikel Villamañe Gironés

Donostia, Octubre de 2016

LENGOAIA ETA SISTEMA INFORMATIKOAK SAILA
DEPARTAMENTO DE LENGUAJES Y SISTEMAS INFORMÁTICOS



Universidad del País Vasco Euskal Herriko Unibertsitatea

Análisis y mejora de los marcos actuales de desarrollo y evaluación de los Trabajos Fin de Grado mediante el uso de las TIC

Memoria que presenta Mikel Villamañe Gironés para optar al Grado de Doctor en Informática por la Universidad del País Vasco (UPV/EHU), bajo la dirección de Ainhoa Izaro Álvarez Arana y María Begoña del Pilar Ferrero Martín.

Donostia, Octubre de 2016

A la memoria de mi madre
A mi familia

Agradecimientos

A Ainhoa y a Bego, sin vuestro esfuerzo y dedicación este trabajo nunca se hubiera terminado. Gracias por ayudarme a cerrar esta etapa.

A Mikel Larrañaga. Tendrían que inventar la figura del tercer director sólo para que la pudieras ocupar en esta tesis. Mila esker egin duzun lanagatik.

A todos los miembros del grupo GaLan. Gracias por vuestro apoyo durante estos años.

Al profesorado de la sección de Bilbao del departamento de Lenguajes y Sistemas Informáticos. Durante estos años habéis estado apoyándome y colaborando siempre que lo he necesitado. Mila esker guztioi!!

Mención especial a Anaje Armendariz y Maider Azanza que, sin ningún tipo de coacción ni amenaza por mi parte, han rellenado las rúbricas en todas las evaluaciones de las que han formado parte. Y han sido unas cuantas. La primera vez os dije que os invitaba a un café. A estas alturas os debo litros y litros.

Al resto de profesorado y alumnado que ha participado en todas las encuestas y estudios que forman parte de este trabajo. Sin vuestra colaboración no habría podido realizarse.

A todas las personas que durante estos años han pasado por mi vida, tanto profesional como personalmente. Gracias por el apoyo y por los buenos momentos vividos. Han sido muchos años y la memoria ya comienza a fallar, así que para evitar olvidarme de alguien, daros todos (y todas) por incluidos (e incluidas) en este agradecimiento.

A mi familia y amigos, que hace tiempo aprendieron que hay cosas que es mejor no preguntar.

Y por último a ti lector o lectora, que te has tomado la molestia de coger esta tesis y dedicarle parte de tu tiempo. Al menos, para leer los agradecimientos. Gracias.

ÍNDICE DE CONTENIDOS

| | |
|---|-----------|
| Resumen | V |
| CAPÍTULO 1. Introducción..... | 1 |
| 1.1 Motivación, contexto y objetivo general..... | 1 |
| 1.1.1 Contexto | 2 |
| 1.1.2 Objetivo general | 3 |
| 1.2 Problemática asociada a la realización y evaluación del TFG | 4 |
| 1.3 Objetivos específicos | 10 |
| 1.4 Resumen y estructura de la memoria | 13 |
| CAPÍTULO 2. Antecedentes y situación actual..... | 15 |
| 2.1 Metodologías para la realización y evaluación del TFG | 15 |
| 2.1.1 Agencia para la Calidad del Sistema Universitario de Cataluña | 16 |
| 2.1.2 Universidad Autónoma de Barcelona | 20 |
| 2.1.3 Universidad de Castilla - La Mancha | 21 |
| 2.1.4 Universidad de Oviedo | 23 |
| 2.1.5 Universidad Tecnológica Nanyang (Singapur)..... | 25 |
| 2.1.6 Universidad de Ciudad del Cabo (Sudáfrica) | 26 |
| 2.1.7 Universidad de Victoria (Australia) | 27 |
| 2.1.8 Universidad de Western Australia (Australia)..... | 28 |
| 2.1.9 Resumen | 29 |
| 2.2 Métodos de evaluación | 30 |
| 2.2.1 Rúbricas de evaluación..... | 31 |
| 2.2.2 Uso de rúbricas en evaluación formativa..... | 34 |
| 2.2.3 Uso de rúbricas en evaluación sumativa..... | 35 |

| | | |
|---------------------------------|---|-----------|
| 2.3 | Las TIC en la realización y evaluación de los TFG..... | 39 |
| 2.3.1 | Requisitos funcionales..... | 39 |
| 2.3.2 | Análisis de herramientas..... | 42 |
| 2.4 | Propuesta de metodología y plataforma informática de soporte..... | 57 |
| 2.4.1 | Hitos de evaluación..... | 58 |
| 2.4.2 | Instrumentos de evaluación..... | 59 |
| 2.4.3 | Plataforma informática de soporte..... | 60 |
| CAPÍTULO 3. TraMeA..... | | 63 |
| 3.1 | InterMod..... | 63 |
| 3.1.1 | Objetivos de usuario..... | 64 |
| 3.1.2 | Proceso de desarrollo de InterMod..... | 64 |
| 3.2 | Análisis general..... | 65 |
| 3.2.1 | Roles de usuario..... | 66 |
| 3.2.2 | Modelo del sistema..... | 67 |
| 3.2.3 | Objetivos de Usuario iniciales..... | 67 |
| 3.3 | Refinamiento de Objetivos de Usuario..... | 73 |
| 3.4 | Diseño y arquitectura..... | 79 |
| 3.5 | Conclusiones..... | 82 |
| CAPÍTULO 4. AdESMuS..... | | 85 |
| 4.1 | Análisis general..... | 85 |
| 4.1.1 | Roles de usuario..... | 86 |
| 4.1.2 | Modelo del sistema..... | 86 |
| 4.1.3 | Objetivos de usuario iniciales..... | 87 |
| 4.2 | Refinamiento de objetivos de usuario..... | 88 |
| 4.3 | Diseño y arquitectura..... | 90 |
| 4.4 | Procesos de ajuste..... | 94 |
| 4.5 | Ontologías de AdESMuS..... | 101 |
| 4.5.1 | Ejemplo de evaluación: el TFG..... | 102 |
| 4.5.2 | Ont_Universidad: Datos administrativos..... | 104 |
| 4.5.3 | Ont_Definición: Definición de la evaluación..... | 108 |
| 4.5.4 | Ont_Evaluación: Realización de la evaluación..... | 120 |

| | | |
|--|--|------------|
| 4.6 | Internacionalización | 125 |
| 4.7 | Una implementación concreta | 127 |
| 4.8 | Conclusiones | 129 |
| CAPÍTULO 5. Definición de un conjunto de rúbricas | | 133 |
| 5.1 | Introducción | 133 |
| 5.2 | Obtención de una versión preliminar del conjunto de rúbricas | 134 |
| 5.3 | Evaluación de las dimensiones de las rúbricas por parte del alumnado | 137 |
| 5.4 | Evaluación de la adecuación de cada tipo de rúbrica..... | 139 |
| 5.5 | Ajuste del conjunto de rúbricas | 142 |
| 5.5.1 | Conjunto de datos | 143 |
| 5.5.2 | Obtención del TAM | 144 |
| 5.5.3 | Obtención del REM..... | 146 |
| 5.6 | Conclusiones | 151 |
| CAPÍTULO 6. Evaluación..... | | 155 |
| 6.1 | Evaluación de los aspectos de evaluación de la metodología FPM | 156 |
| 6.1.1 | Adecuación de los elementos evaluables..... | 156 |
| 6.1.2 | Beneficios del uso de rúbricas de evaluación | 157 |
| 6.2 | Evaluación de TraMeA | 159 |
| 6.2.1 | Fase 1: Preparación estudio..... | 160 |
| 6.2.2 | Fase 2: Recogida de datos..... | 162 |
| 6.2.3 | Fase 3: Análisis de datos..... | 166 |
| 6.2.4 | Participantes | 172 |
| 6.2.5 | Resultados de la evaluación..... | 173 |
| 6.3 | Conclusiones | 200 |
| CAPÍTULO 7. Conclusiones y líneas futuras..... | | 201 |
| 7.1 | Resumen | 201 |
| 7.2 | Resultados y contribuciones | 204 |
| 7.2.1 | OE—1: Asegurar una calidad mínima en todos los | |

| | |
|---|-----|
| TFG..... | 204 |
| 7.2.2 OE—2: Aumentar y mejorar la retroalimentación que recibe el alumnado..... | 205 |
| 7.2.3 OE—3: Incorporar nuevos canales y posibilidades de comunicación..... | 205 |
| 7.2.4 OE—4: Establecer un proceso de evaluación continuo..... | 205 |
| 7.2.5 OE—5: Establecer un método de evaluación equitativo..... | 206 |
| 7.2.6 OE—6: Desarrollar las herramientas informáticas necesarias..... | 206 |
| 7.3 Líneas futuras..... | 208 |
| 7.3.1 Uso de AdESMuS en nuevos ámbitos..... | 208 |
| 7.3.2 Incorporación de aspectos de recomendación..... | 209 |
| 7.3.3 Incorporación de aspectos de visualización..... | 210 |
| 7.3.4 Integración de AdESMuS con otras herramientas..... | 211 |
| 7.3.5 Enriquecimiento del proceso de ajuste..... | 211 |
| 7.3.6 Otras extensiones..... | 212 |
| BIBLIOGRAFÍA..... | 215 |
| ÍNDICE DE FIGURAS..... | 227 |
| ÍNDICE DE TABLAS..... | 231 |
| ANEXO A - ACRÓNIMOS..... | 235 |
| ANEXO B - ENCUESTA ALUMNADO..... | 237 |
| ANEXO C - ENCUESTA PROFESORADO..... | 247 |
| ANEXO D - RÚBRICAS DE EVALUACIÓN..... | 261 |

Resumen

Este trabajo de tesis aborda la problemática inherente al TFG, planteando como objetivo principal la mejora de los procesos de realización y evaluación del mismo, dando soporte a los actores implicados en dichos procesos (estudiantes, supervisores y evaluadores) mediante el uso de herramientas que se benefician de las ventajas que pueden aportar las TIC.

La revisión de la bibliografía sobre el tema y las encuestas realizadas a personas que habían estado involucradas en algún TFG han permitido identificar los aspectos más problemáticos de la realización y evaluación de los TFG y determinar el conjunto de objetivos específicos a considerar para alcanzar el objetivo general de este trabajo de tesis. Con este fin, se han presentado la metodología FPM (*Final-year Project Mentoring*) y la plataforma TraMeA (*Tracking Mentoring and Assessment*) que le da soporte.

Previamente a la construcción de TraMeA, se han identificado los requisitos funcionales que debería cumplir una herramienta para alcanzar los objetivos específicos definidos en este trabajo.

La funcionalidad de TraMeA relacionada con la evaluación se ha generalizado dando lugar al sistema autónomo AdESMuS (*Adaptable Evaluation System using Multiple Sources*) que permite también dar respuesta a las necesidades de mejora del procesos de evaluación en entornos de aprendizaje combinados.

En cuanto a la evaluación, en primer lugar se han tratado los aspectos de evaluación definidos en la metodología FPM: la propuesta de evaluación continua y las rúbricas definidas para ello. A continuación se ha evaluado la plataforma TraMeA realizando un análisis de usabilidad. Finalmente, la funcionalidad ofrecida por el sistema AdESMuS se ha probado mediante su ejecución en un entorno real en combinación con TraMeA.

1

Introducción

Este capítulo describe la motivación y el punto de partida de este trabajo de tesis, así como su objetivo general. A continuación se analiza la problemática asociada a la realización y evaluación de proyectos y se definen los objetivos específicos a satisfacer para resolver los problemas detectados. El capítulo concluye describiendo la estructura general de la memoria.

1.1 Motivación, contexto y objetivo general

Con la implantación del Espacio Europeo de Educación Superior (EEES) conocido como Plan Bolonia (European Ministers of Education, 1999), todos los planes de estudio conducentes a la obtención de un título universitario se han transformado en Grados y culminan con la realización y defensa de un Trabajo Fin de Grado (TFG) que esté orientado a la evaluación de las competencias asociadas al título (*RD 1393/2007, 2007*).

Esta situación supone una novedad para muchas áreas cuyas titulaciones no contemplaban nada similar, sin embargo no es nueva para el campo de las ingenierías donde los planes de estudio previos a la implantación del EEES ya incluían la realización y evaluación de un Proyecto Fin de Carrera (PFC) que permitiera al alumnado aplicar todos los conocimientos adquiridos durante su formación.

Considerando la similitud entre los antiguos PFC y los nuevos TFG, al menos en el campo de las ingenierías, partimos de la hipótesis de que el proceso de

elaboración de ambos sea análogo y por lo tanto presenten el mismo tipo de problemática (Escudero & Hernández, 2012).

La experiencia acumulada durante los años que han estado vigentes los planes de estudio que requerían la elaboración de un PFC, ha puesto de manifiesto una serie de problemas que se han ido repitiendo a lo largo de los años, independientemente de la titulación y de la Universidad.

El TFG tiene como objetivo que el alumnado demuestre su dominio de las competencias asociadas al título, pero la evaluación del nivel de dominio concreto de cada competencia no es sencilla (M. J. Bonilla et al., 2012; Rullan et al., 2010; Valderrama et al., 2010). Además, cabe destacar la carencia, en muchos casos, de unos requisitos mínimos de calidad y de unos criterios que garanticen uniformidad y objetividad en el nivel de exigencia y en la evaluación del trabajo del alumnado (Dutson et al., 1997; Vijayalakshmi et al., 2013).

Otro problema de gran importancia que se repite año tras año, es el considerable número de proyectos que sufren un gran retraso en su elaboración, llegando en algunos casos, incluso a no finalizarse nunca (Sánchez et al., 2015).

En base a la experiencia en la aplicación de las Tecnologías de la Información y la Comunicación (TIC) en el ámbito de la educación, partimos de la hipótesis de que el empleo de las TIC en la realización y evaluación de los TFG puede ayudar a solventar la problemática detectada.

1.1.1 Contexto

Este trabajo se ha desarrollado en el seno del grupo de investigación en **Entornos de Aprendizaje Adaptativos** (GaLan¹) de la Universidad del País Vasco UPV/EHU (Fernández-Castro et al., 2012). La actividad investigadora del grupo GaLan se enmarca en el área de las herramientas de ayuda al aprendizaje, orientándose hacia los entornos combinados que, haciendo uso de las TIC, integran aprendizaje presencial y no presencial (*blended learning*). En este ámbito, su objetivo general es mejorar los sistemas de apoyo al aprendizaje desde los puntos de vista del profesorado y del alumnado.

En los últimos años, a través de diferentes proyectos de investigación y tesis doctorales, el grupo ha dirigido su trabajo hacia varias líneas de investigación relacionadas con la mejora de los entornos de aprendizaje. En este ámbito, se han desarrollado propuestas de plataformas educativas (A. Álvarez, 2010; A. Álvarez et al., 2013, 2014) orientadas al *blended learning*, que incluyen aspectos de

¹ <http://galan.ehu.es>

recomendación y analítica del aprendizaje (*learning analytics*) (Martín, 2014; Ruiz et al., 2016; Villamañe et al., 2015). Estas propuestas continúan ampliándose y completándose con el estudio e incorporación de las interacciones presenciales entre alumnado y profesorado (Ruiz et al., 2015).

También se ha trabajado en temas relacionados con la evaluación cognitiva de los estudiantes a través de ejercicios de distinto tipo tales como cuestiones de respuesta abierta en lenguaje natural (Zipitria, 2011) y realización de tareas procedimentales en entornos virtuales (Aguirre et al., 2012, 2014; Ferrero, 2004), entre otros.

Por otro lado, se han obtenido resultados interesantes en el área de las metodologías ágiles para el desarrollo de sistemas interactivos (Losada, 2014) y en el estudio de la usabilidad de software educativo (López-Gil et al., 2016). Estos resultados se utilizarán como base para el desarrollo y evaluación de las herramientas que se propongan en este trabajo de tesis.

De manera complementaria a estos trabajos, recientemente el grupo ha liderado un Proyecto de Innovación Educativa (PIE) en el que se ha analizado el proceso de adaptación de los antiguos Proyectos Fin de Carrera a los nuevos Trabajos Fin de Grado.

Este proyecto de tesis se integra en las líneas de investigación del grupo con el objetivo enriquecer los resultados alcanzados en trabajos previos, ampliándolos para satisfacer las necesidades inherentes a la realización y evaluación de los TFG.

1.1.2 Objetivo general

En este contexto, el objetivo general de la investigación desarrollada en este trabajo de tesis se puede resumir de la siguiente manera:

«Analizar los problemas que pueda plantear tanto para el profesorado como para el alumnado la realización y evaluación de un TFG y establecer una metodología que evite o disminuya la aparición de dichos problemas y minimice su impacto. Se abordará la definición de los procesos, documentación y herramientas informáticas que den soporte a dicha metodología y la evaluación tanto de la metodología, como de los elementos definidos y los resultados de su aplicación».

Para abordar este objetivo, inicialmente se realizará un análisis de los problemas que puede plantear, tanto para el profesorado como para el alumnado, la realización y evaluación de un TFG.

A continuación se realizará una propuesta de metodología para la realización y evaluación de los TFG y se abordará la definición de los procesos, documentación y

herramientas informáticas que den soporte a dicha metodología. Para ello, se aprovecharán las ventajas que aportan las TIC para evitar o disminuir la aparición de los problemas previamente detectados, minimizando su impacto durante la realización y evaluación del TFG.

Por último, se evaluará tanto la metodología como todos los elementos definidos para darle soporte y los resultados de su aplicación.

1.2 Problemática asociada a la realización y evaluación del TFG

Partiendo del objetivo general establecido para este trabajo de tesis (ver sección 1.1.2) y con el fin de determinar cuáles son las principales dificultades a las que hacen frente alumnado y profesorado durante la realización y evaluación del TFG, se ha realizado una revisión bibliográfica que se presenta en el capítulo 2. Este estudio bibliográfico se ha complementado con la realización de una encuesta a los actores implicados (alumnado y profesorado), cuyo contenido y análisis detallado se adjunta en el ANEXO B y el ANEXO C.

De los resultados tanto del estudio bibliográfico como del análisis de las respuestas a las encuestas, se han identificado un conjunto de problemas asociados a la realización y la evaluación de los TFG (Villamañe, Ferrero, Álvarez, et al., 2014). Problemas, a los que la metodología para la realización y evaluación del TFG que se desea definir, debe dar respuesta.

Alcance. Falta de criterios para definir el alcance de los TFG

Existen varios estudios que reflejan la preocupación por la ausencia de unos criterios preestablecidos que determinen de manera objetiva si la propuesta inicial de un trabajo, el llamado alcance del proyecto, tiene la entidad necesaria para demostrar el dominio de las competencias asociadas y por lo tanto para constituir un TFG (Fraile et al., 2009; Jawitz et al., 2002). Además, siendo la falta de motivación uno de los principales problemas durante la realización de un proyecto, trabajar en una idea propia puede aumentar dicha motivación y generar una mayor dedicación al proyecto (McMillan & Hearn, 2008). Por esta razón, se debe promover que sea el alumnado quien proponga sus propias ideas para realizar su TFG. De este modo, se espera que el incremento en la motivación se vea reflejado en la reducción del número de trabajos que se prolongan en exceso en el tiempo así como de los ratios de abandono.

Los resultados recogidos en la encuesta corroboran la preocupación por la falta de unos criterios que establezcan el alcance del proyecto. El 80% del profesorado

reconoce que en su titulación no existe ningún criterio para determinar la calidad de una propuesta más allá de su opinión personal (Fig. 1-1→a). Sin embargo, el 76% considera que es necesario que dichos criterios existan y se apliquen para valorar todas las propuestas de la titulación (Fig. 1-1→b).

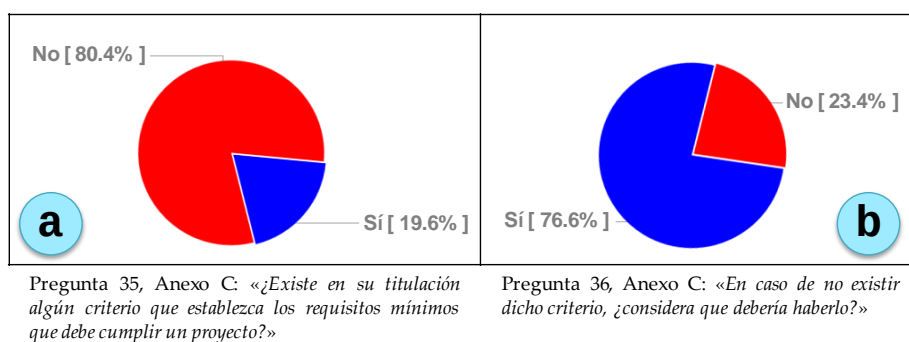


Fig. 1-1.- Resultados de las preguntas 35 (a) y 36 (b) al profesorado

Esta ausencia de criterios preocupa sobre todo al profesorado con menor experiencia en la supervisión de proyectos, ya que desconocen el nivel de exigencia del tribunal que evaluará el TFG y temen que el alumnado bajo su supervisión se encuentre con problemas si el tribunal considera que el punto de partida del proyecto no tenía la entidad suficiente. Este temor se corrobora con los datos recogidos en la encuesta, donde el 73% del profesorado afirma que en alguna ocasión se han evaluado proyectos cuya idea de partida no tenía una calidad o complejidad adecuada a la titulación (Fig. 1-2).

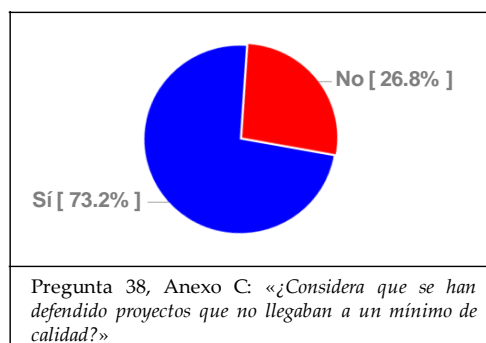


Fig. 1-2.- Resultados de la pregunta 38 al profesorado

La existencia de un proceso para evaluar las propuestas de alcance de los proyectos, dando el visto bueno a aquellas que tengan una calidad mínima y que sean apropiadas a la titulación, permite conocer desde el principio los objetivos concretos a alcanzar. De este modo, el alumnado, además de tener la seguridad de que el trabajo que está realizando es viable, puede percibir y cuantificar sus avances.

Existen titulaciones donde todas las propuestas de proyecto surgen del profesorado, pero en aquellas titulaciones que lo permiten, es interesante fomentar que sea el propio alumnado quien defina el proyecto que quiere realizar y por lo tanto su alcance. Aunque en el caso de las titulaciones de Informática impartidas en la UPV/EHU, existe esta opción, sólo el 25% de las propuestas surgen del alumnado (Fig. 1-3→a). Esto puede ser debido a que, según se recoge en las respuestas, el 55% del alumnado no conoce los requisitos que debe cumplir una idea de TFG para ser válida en su titulación (Fig. 1-3→b).

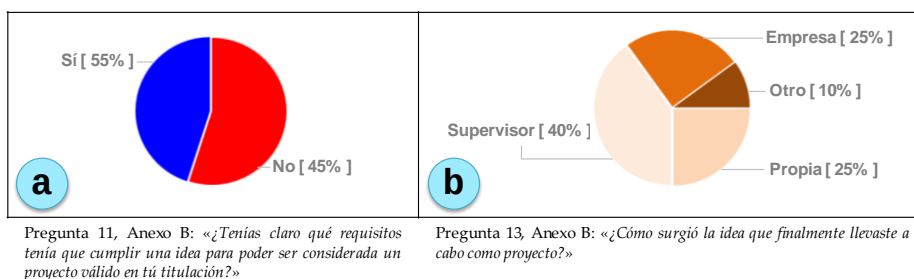


Fig. 1-3.- Resultados de las preguntas 11 (a) y 13 (b) al alumnado

Los criterios para determinar si el alcance de un TFG es adecuado, deben estar previamente definidos y consensuados por la mayor cantidad de profesorado posible. Estos criterios, además, tienen que ser públicos para que tanto profesorado como alumnado conozcan los requisitos que debe satisfacer una idea de trabajo antes de plantearla.

De esta forma, el alumnado puede sentirse más orientado y dispuesto a proponer sus propias ideas de trabajo, y así fomentar sus capacidades de innovación, dando como resultado un incremento del número de TFG realizados a partir de ideas del propio alumnado.

Comunicación. Escasez de canales de comunicación entre actores

Varios autores hacen mención al problema del aislamiento o la incomunicación del alumnado como uno de los principales problemas durante la realización de un proyecto (Heinze & Heinze, 2009; Malik et al., 2009; Sabah, 2013).

Según se puede extraer del análisis de las encuestas realizadas, tanto el alumnado como el profesorado se sienten insatisfechos con los actuales canales de comunicación.

Un 30% del alumnado reclama un mayor apoyo y soporte por parte de los supervisores. El profesorado, por su parte, se queja de que un elevado porcentaje del alumnado sólo informa del estado de su trabajo cuando se les requiere explícitamente y el 87% afirma que sin la colaboración activa del alumnado no dispone de vías para conocer el estado del TFG, lo que dificulta la tarea de supervisión del mismo (Fig. 1-4→a). El 76% del profesorado afirma que le gustaría disponer de algún medio para conocer el estado del TFG sin tener que depender de la voluntad de informar del alumnado (Fig. 1-4→b).

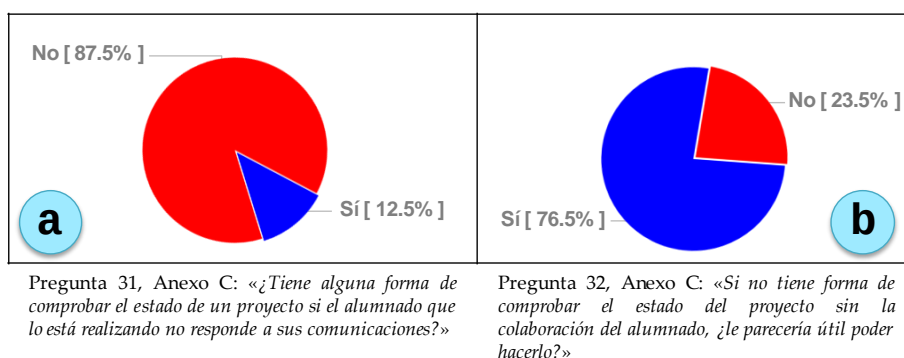


Fig. 1-4.- Resultados a las preguntas 31 (a) y 32 (b) al profesorado

Seguimiento. Falta de un proceso continuo de evaluación

El estudio bibliográfico también indica que el proceso de evaluación del TFG preocupa a todos los actores (M. I. Bonilla & Martín, 2012; Ferrer et al., 2012; Sánchez et al., 2015; Stewart & Nesbitt, 2005; Valderrama et al., 2009). Las respuestas a la encuesta confirman esta preocupación.

El profesorado expresa su malestar por realizar la evaluación de todo un proceso que conlleva varios meses de trabajo, exclusivamente al final del mismo. Esta forma de evaluar hace imposible corregir posibles errores o desviaciones que se produzcan en las primeras etapas. Es por ello que se plantea la necesidad de llevar a cabo un proceso continuo de seguimiento y revisión. Un proceso continuo, en el que el profesorado revise regularmente el avance del alumnado, y ayude a mejorar la elaboración global del TFG. El 80% del alumnado (Fig. 1-5→a) y el 66% del profesorado (Fig. 1-5→b) consideran que este sistema de evaluación sería beneficioso y ayudaría a mejorar la calidad de los trabajos realizados.

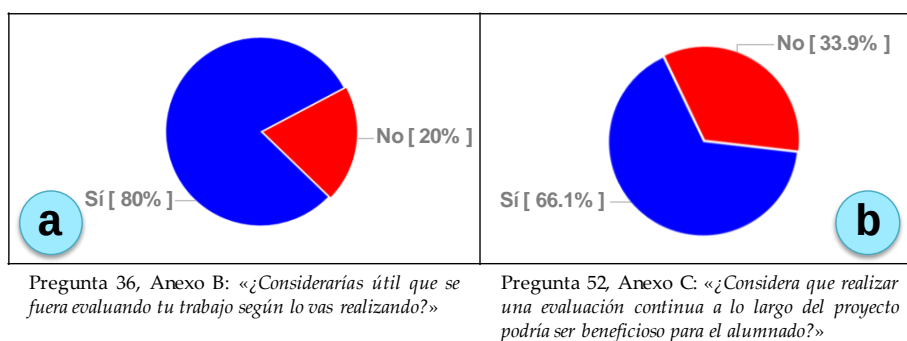


Fig. 1-5.- Resultados a las preguntas 36 (a) al alumnado y 52 (b) al profesorado

Este proceso continuo permitiría al alumnado corregir de forma más inmediata los problemas que puedan surgir, evitando las consecuencias de posponer su resolución. Además, si a este seguimiento se le incluye retroalimentación, la carencia de soporte referida por el alumnado se reduce. Este tipo de proceso también beneficia al profesorado haciendo que la labor de supervisión, al estar distribuida en el tiempo, sea más sencilla y efectiva.

Evaluación. Falta de criterios de evaluación objetivos y públicos

Tanto alumnado como profesorado consideran que la evaluación es un proceso subjetivo y variable debido a que no existen unos criterios públicos preestablecidos para valorar los TFG (Chan, 2001; Fraile et al., 2009; Jawitz et al., 2002).

Del análisis de las respuestas de la encuesta relacionadas con el proceso de evaluación, también se pueden sintetizar una serie de resultados en la misma línea. El 53% del alumnado afirma no conocer los criterios de evaluación antes de ser evaluado (Fig. 1-6→a), y lo que es peor, el 40% dice seguir sin conocerlos tras haber sido evaluado (Fig. 1-6→b).

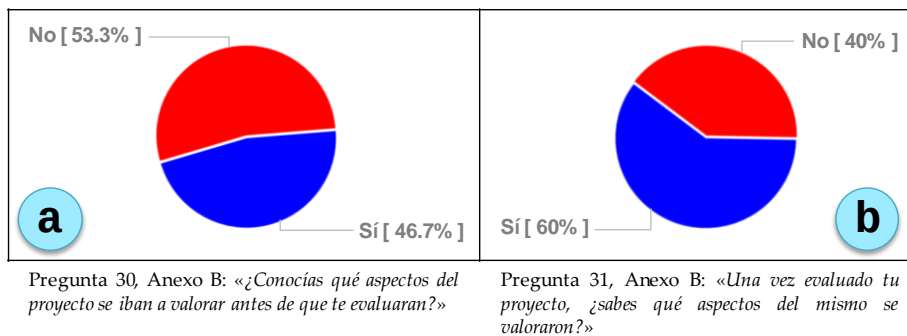
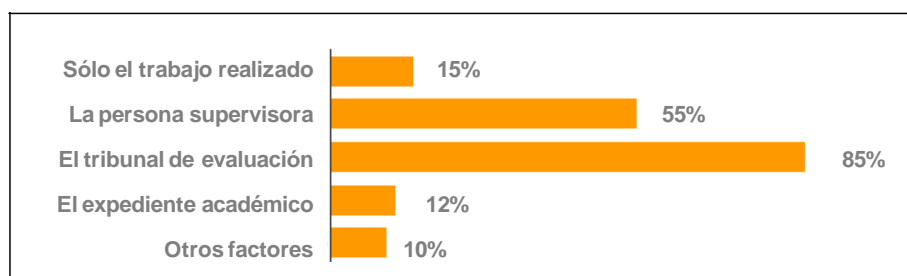


Fig. 1-6.- Resultados a las preguntas 30 (a) y 31 (b) al alumnado

Por lo tanto, desde el punto de vista del alumnado el proceso de evaluación se percibe como algo opaco, donde no se sabe muy bien qué se valora. De hecho, de acuerdo a las respuestas recogidas a la pregunta 34 del ANEXO B (ver Fig. 1-7), sólo el 15% de los participantes en la encuesta considera que a la hora de evaluar su proyecto, el trabajo realizado influye en la nota. Sin embargo, el 85% opina que la nota obtenida depende de quienes integren el tribunal y el 55% cree que la persona ha supervisado el proyecto influye en la calificación.



Pregunta 34, Anexo B: «¿Qué factores consideras que influyen en la evaluación de un proyecto? Marca todas las opciones que consideres»

Fig. 1-7.- Resultados a la pregunta 34 al alumnado

Con el objetivo de minimizar los problemas durante la evaluación, se propone establecer un conjunto de criterios públicos para la evaluación del TFG. De este modo, tanto el alumnado como el profesorado podrían conocer a priori los criterios bajo los que se va a evaluar el trabajo desarrollado y, además, se reduciría la subjetividad de la evaluación.

Cuantificar cada particularidad de los diferentes tipos de trabajos para elaborar un proceso realmente objetivo es inabordable. Sin embargo, establecer una serie de criterios que puedan aplicarse a todos los trabajos, en este caso del área de la Ingeniería Informática, reduciría considerablemente su subjetividad.

Además, aplicando los mismos criterios a todos los trabajos, cabe esperar que aquellos que sean similares obtengan calificaciones también similares, lo que terminaría con la sensación actual de falta de homogeneidad y equidad que perciben tanto el 70% del alumnado (Fig. 1-8→a) como el 71% del profesorado (Fig. 1-8→b).

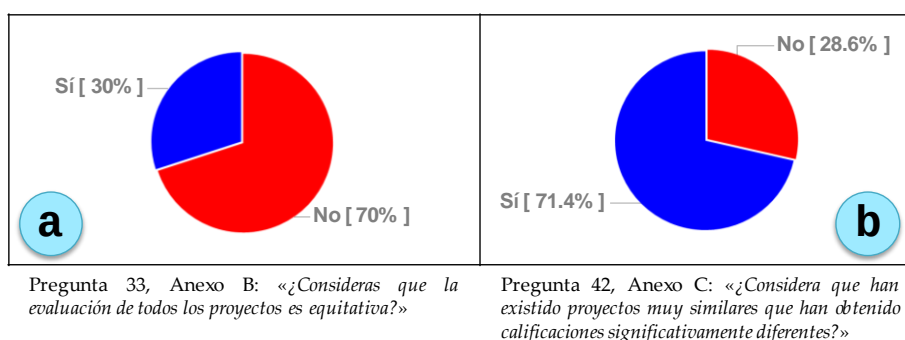


Fig. 1-8.- Resultados a las preguntas 33 (a) al alumnado y 42 (b) al profesorado

1.3 Objetivos específicos

El estudio realizado ha permitido la identificación de los principales problemas a resolver, determinando el conjunto de objetivos específicos a cubrir para alcanzar el objetivo general de esta tesis

OE-1. ASEGURAR UNA CALIDAD MÍNIMA EN TODOS LOS TFG

Tal y como refleja el problema de *Alcance*, en ocasiones se han llegado a defender, e incluso a aprobar, proyectos que a juicio del tribunal no tenían la entidad necesaria para ser considerados válidos. Además, en aquellos casos donde la normativa lo permite, el alumnado suele ser reacio a proponer sus propias ideas para realizar el proyecto debido al desconocimiento de las características que debe cumplir un proyecto para ser considerado válido en su titulación

Para dar solución a estos problemas, la metodología tiene que incorporar un proceso de validación de propuestas de TFG, que permita valorar cada idea antes

de que sea puesta en marcha. De este modo, se asegura que los TFG que finalmente se desarrollen tendrán la entidad necesaria, evitando problemas el día de su evaluación. Además, este proceso de validación de propuestas tiene que ir acompañado de la definición y publicación de los criterios que determinan los mínimos que debe cumplir una propuesta para ser considerada válida. Esto permite al alumnado conocer qué características tiene que cumplir una idea de TFG para ser válida y fomenta que presente sus propias propuestas.

OE-2. AUMENTAR Y MEJORAR LA RETROALIMENTACIÓN QUE RECIBE EL ALUMNADO

Actualmente, el alumnado demanda una mayor atención por parte del profesorado durante el proceso de realización del proyecto (problema de *Comunicación*). En muchos casos, la única retroalimentación que recibe por parte de la persona que supervisa el proyecto es la corrección de la documentación final del mismo, lo cual no facilita la corrección de errores cometidos en etapas tempranas (problema de *Evaluación*). Además, generar retroalimentación adecuada puede ser una tarea muy costosa y dado que la cantidad y la calidad de la retroalimentación proporcionada por la persona encargada de la supervisión dependen de su predisposición, pueden surgir problemas en épocas de sobrecarga de trabajo del profesorado.

Para evitar estos problemas es preciso establecer mecanismos de retroalimentación que aseguren al alumnado que recibirá retroalimentación en momentos clave a lo largo de la realización del TFG y garanticen una calidad mínima de la misma. Estos mecanismos de retroalimentación, además, no deberían suponer una excesiva sobrecarga de trabajo para los actores implicados.

OE-3. INCORPORAR NUEVOS CANALES Y POSIBILIDADES DE COMUNICACIÓN

El problema de *Comunicación* detectado, recoge la preocupación del profesorado porque en la mayoría de los casos no disponen de medios o herramientas para conocer el avance del trabajo sin necesidad de consultárselo al alumnado. Esto dificulta la supervisión del proyecto en aquellos casos en los que el alumnado no colabora activamente en la comunicación de sus avances. Además, como la realización de un proyecto es, en la mayoría de los casos, un proceso individual, el alumnado en ocasiones se siente aislado no teniendo posibilidad de compartir preocupaciones y dudas con otras personas en su misma situación.

La incorporación de nuevos canales y posibilidades de comunicación entre alumnado y profesorado que aporten mayores posibilidades respecto al correo electrónico y las tutorías presenciales, incrementará las opciones de controlar los avances del TFG por parte del profesorado, mejorando la supervisión del mismo. Asimismo, se deben establecer mecanismos para fomentar la comunicación entre el

alumnado, dándoles la posibilidad de colaborar resolviendo dudas y apoyándose mutuamente.

OE-4. ESTABLECER UN PROCESO DE EVALUACIÓN CONTINUO

El problema de *Seguimiento* hace patente que realizar una única evaluación final de un trabajo que supone meses de esfuerzo no satisface ni al alumnado, ni al profesorado.

Por tanto, es necesario definir la evaluación como un proceso continuo con varios hitos. En cada uno de esos hitos se evaluará parte del trabajo realizado por el alumnado, permitiendo de este modo solucionar de manera temprana los posibles errores que se estén cometiendo. Además, esta evaluación continua también obligará a una mayor interacción entre profesorado y alumnado, incrementando el soporte que estos últimos reciben por parte de las personas supervisoras de los TFG y solucionando parte del problema de *Comunicación* detectado.

OE-5. ESTABLECER UN MÉTODO DE EVALUACIÓN EQUITATIVO

A la hora de evaluar al alumnado, el problema de *Evaluación* establece que los criterios de evaluación deberían estar fijados, ser iguales para todos los TFG y conocerse a priori para reducir la impresión de opacidad y subjetividad del proceso.

Se debe establecer un método de evaluación que permita definir y aplicar criterios de evaluación objetivos para todos los TFG. De este modo, haciendo públicos dichos criterios, tanto el alumnado como el profesorado conocerán qué aspectos del trabajo van a ser considerados para calificarlo. Esto facilitará el trabajo del profesorado, sobre todo las primeras veces que forme parte de un tribunal de evaluación. Asimismo, también permitirá al alumnado autoevaluar su propio trabajo y conocer de manera aproximada qué calificación obtendrá antes de someterse a la evaluación real, pudiendo así mejorarlo si lo considera necesario.

OE-6. DESARROLLAR LAS HERRAMIENTAS INFORMÁTICAS NECESARIAS

La definición de una metodología para la realización y evaluación de los TFG establece los procedimientos a seguir y los métodos a utilizar para dar soporte a los procesos de realización, seguimiento y evaluación de los TFG.

Teniendo en cuenta la experiencia previa del grupo GaLan en el área de las herramientas de ayuda al aprendizaje haciendo uso de las TIC, se desarrollarán las herramientas informáticas necesarias para soportar los procedimientos y métodos establecidos en la metodología. Además, se deben tener en cuenta también los objetivos específicos OE-2 y OE-3 de este trabajo de tesis y dichas herramientas

deben facilitar la retroalimentación e incorporar nuevos canales y posibilidades de comunicación para el alumnado y el profesorado.

1.4 Resumen y estructura de la memoria

En este capítulo se ha presentado el objetivo principal de este trabajo de tesis que consiste en analizar el actual marco de realización y evaluación de los Trabajos Fin de Grado y proponer mejoras al mismo aprovechando a su vez las ventajas que aporta el uso de las TIC.

Tras analizar el marco actual de realización y evaluación, se ha identificado un conjunto de problemas y se han definido los objetivos específicos a cubrir por este trabajo.

En el segundo capítulo se realiza una revisión de los antecedentes y desarrollos actuales relacionados con este trabajo de tesis, concluyendo con la propuesta de la metodología FPM (*Final-year Project Mentoring*) para la realización y evaluación del TFG. Asimismo también se perfila la plataforma que le da soporte: TraMeA (*Tracking, Mentoring and Assessment*). El capítulo 3 presenta el análisis y diseño de TraMeA y finaliza con la definición de un sistema que proporciona servicios de evaluación, AdESMuS (*Adaptable Evaluation System using Multiple Sources*), cuyo análisis, diseño y desarrollo se describe en el capítulo 4. El capítulo 5 aborda la generación del conjunto de rúbricas de evaluación utilizado en la aplicación de la metodología FPM a los distintos Grados en Ingeniería Informática que se imparten en la UPV/EHU. La evaluación de la metodología FPM, de la plataforma que le da soporte y del sistema de evaluación AdESMuS, se realiza en el capítulo 6. Finalmente el capítulo 7 resume las principales contribuciones realizadas y las líneas futuras del trabajo presentado.

2

Antecedentes y situación actual

En este capítulo se presenta un estudio bibliográfico de los distintos aspectos relacionados con el objetivo principal de este trabajo de tesis. Por un lado se analizan las soluciones aportadas a problemáticas similares por otros autores en la bibliografía relacionada y se estudian los distintos métodos de evaluación empleados en la evaluación de TFG. Por otro lado, se definen los requisitos que deben cumplir las herramientas informáticas a utilizar durante la realización y evaluación de los TFG. Una vez establecidos dichos requisitos se presenta una revisión de distintas herramientas y se analiza hasta qué punto satisfacen los requisitos. El capítulo finaliza con la propuesta de la metodología FPM para la realización y evaluación de TFG, y de los instrumentos de evaluación y herramientas informáticas a utilizar como apoyo a dicha metodología.

2.1 Metodologías para la realización y evaluación del TFG

En este apartado se describen las principales propuestas de metodologías para la realización y evaluación del TFG que se han encontrado en la bibliografía. Teniendo en cuenta los objetivos específicos definidos para este trabajo de tesis (ver sección 1.3), se han seleccionado únicamente propuestas que planteen realizar una evaluación continua.

De este modo, las propuestas analizadas satisfacen el objetivo de *Establecer un proceso de evaluación continuo* (OE-4) Además, la evaluación y seguimiento continuos del alumnado mejoran el proceso de supervisión y garantiza la calidad del producto

final (Escudero & Hernández, 2012), por lo que estas propuestas también satisfacen el objetivo de *Asegurar una calidad mínima en todos los TFG* (OE-1).

Para cada una de las propuestas se detalla en qué momentos de la realización del TFG se realiza la evaluación (hitos), qué elementos se evalúan, la influencia en la nota final de dichos elementos y quiénes son las personas encargadas de su evaluación.

Asimismo, también se analizan los instrumentos de evaluación contemplados en cada una de las propuestas y su utilidad para satisfacer los objetivos específicos de *Aumentar y mejorar la retroalimentación que recibe el alumnado* (OE-2) y de *Establecer un método de evaluación equitativo* (OE-5).

Dado que este trabajo de tesis aborda la realización y evaluación de los TFG en cualquier disciplina, se han estudiado propuestas en ámbitos como la Ingeniería, la Administración de Empresas y el Derecho, para analizar las similitudes y diferencias entre los procesos de realización y evaluación del TFG distintas áreas.

A continuación se presentan las distintas propuestas estudiadas indicando la institución en la que se han definido. Para cada una de ellas se presentan en primer lugar los hitos de evaluación que define y, a continuación, los instrumentos de evaluación que utiliza.

El apartado finaliza con un resumen que recoge los principales aspectos del estudio bibliográfico llevado a cabo.

2.1.1 Agencia para la Calidad del Sistema Universitario de Cataluña

La Agencia para la Calidad del Sistema Universitario de Cataluña publicó una guía para la evaluación de competencias en los Trabajos Fin de Grado y de Máster en las Ingenierías (Valderrama, 2009) y otra para la evaluación de competencias en los Trabajos Fin de Grado en el ámbito de las Ciencias Sociales y Jurídicas (Mateo, 2009).

Hitos de evaluación

La guía para la evaluación de competencias en los Trabajos Fin de Grado y de Máster en las Ingenierías propone un método de evaluación continuo con 3 hitos de evaluación (ver Tabla 2-1): al inicio del trabajo; durante la realización (*Seguimiento*) y al final del TFG.

Tabla 2-1.- Evaluación propuesta por (Valderrama et al., 2009)

| Hitos de evaluación | Elementos evaluables | Peso | Responsables evaluación |
|---------------------|---------------------------------------|------|--|
| Inicio | Informe técnico y presentación oral | -- | Persona supervisora / miembros alumnado / personas ajenas al TFG |
| Seguimiento | Informe técnico | -- | Persona supervisora |
| Final | Documentación final y defensa pública | -- | Tribunal de evaluación |

Los pesos correspondientes a la evaluación realizada en cada uno de esos hitos no están especificados en la guía, permitiendo así su adaptación a las características propias de cualquier titulación. Además, esta metodología incorpora a la calificación final un informe emitido por la persona supervisora en el que se evalúa el proceso de seguimiento llevado a cabo durante el TFG.

La evaluación del primer hito (*Inicio*), se basa en un informe técnico y en una exposición oral sobre los objetivos del TFG y la estrategia a seguir para satisfacerlos. Aunque esta evaluación la realiza la persona supervisora, la propuesta recomienda permitir que el resto del alumnado que está realizando el TFG y otras personas expertas en la materia y ajenas al TFG participen también en la valoración de las presentaciones orales. En el hito de *Seguimiento*, la persona evaluadora califica un informe elaborado por el o la estudiante, en el que se recoge la evolución del TFG y las decisiones tomadas. En caso de considerarlo necesario, también se puede incorporar a la evaluación una entrevista personal. En el último de los hitos de evaluación, los componentes del tribunal de evaluación deben valorar la documentación final y una defensa pública.

Por su parte, la guía para la evaluación de competencias en los Trabajos Fin de Grado en el ámbito de las Ciencias Sociales y Jurídicas establece que la evaluación se realice en 5 hitos y especifica el peso que debe tener la evaluación de cada uno de ellos en la calificación final (ver Tabla 2-2).

Tabla 2-2.- Evaluación de los TFG del área de Ciencias Sociales y Jurídicas (Mateo, 2009)

| Hitos de evaluación | Elementos evaluables | Peso | Responsables evaluación |
|----------------------------------|--------------------------------|-------------|--------------------------------|
| Elección del tema | Informe | 10% | Persona supervisora |
| Planificación | Informe y presentación pública | 10% | Persona supervisora / alumnado |
| Desarrollo | Informe y presentación pública | 30% | Persona supervisora / alumnado |
| Entrega y presentación pública | Informe y presentación pública | 30% | Comisión experta |
| Evaluación, mejora y prospectiva | Informe | 20% | Persona supervisora |

En esta propuesta se incorpora un hito de evaluación en el momento que se elige el tema sobre el que se va a realizar el TFG. Con este hito se consigue reconducir el trabajo cuando la calidad del tema seleccionado no sea la adecuada, garantizando así, un mínimo de calidad a la finalización del mismo.

En todos los hitos se evalúa un informe escrito elaborado por el o la estudiante, pero además, en tres de ellos (*Planificación, Desarrollo y Entrega y presentación pública*), se añade a la evaluación una exposición oral. La guía establece que el 70% de la calificación lo determina la persona supervisora valorando los distintos informes escritos, pero sugiere que se permita al alumnado participar en la evaluación de las presentaciones orales que se realicen asociadas a los hitos de *Planificación y Desarrollo*. El 30% restante de la calificación lo decide una comisión experta en la materia del TFG, que evalúa tanto el informe escrito como la presentación pública correspondientes al hito *Entrega y presentación pública*.

Instrumentos de evaluación

La guía para la evaluación de competencias en los Trabajos Fin de Grado y de Máster en las Ingenierías propone la definición de entre 5 y 10 indicadores para cada uno de los elementos que se evalúan en los hitos. Los indicadores deben ser definidos para cada titulación concreta teniendo en cuenta las competencias asociadas al título correspondiente.

Asimismo, la guía también establece que la valoración de los indicadores se realice puntuando de 0 a 3 su nivel de cumplimiento e incorpora un apartado de comentarios en el que la persona evaluadora puede añadir comentarios o matizaciones a la evaluación.

En la Tabla 2-3 se muestra un ejemplo de posibles indicadores para el elemento evaluable correspondiente al informe que se evalúa al inicio del trabajo.

Tabla 2-3.- Ejemplo de indicadores para un elemento evaluable. Adaptado de (Valderrama et al., 2010)

| Indicadores | Puntos (0 a 3) |
|---|----------------|
| Identifica las partes fundamentales del proyecto, dibujando un diagrama a nivel de bloques que las describa y que visualiza las relaciones entre ellas. | |
| Identifica los conocimientos implicados en la base de la resolución del proyecto, tanto sean propios de la disciplina como ajenos. | |
| Valora la importancia relativa de cada una de las partes del proyecto y de los conocimientos implicados. | |
| Da una solución realizable al proyecto empleado, considerando las herramientas propias de la Ingeniería. | |
| Incluye la planificación del proyecto en forma de un GANTT o similar. | |
| Comentarios: | |

Este sistema de evaluación, a pesar de definir los criterios que se tienen en cuenta a la hora de valorar el TFG, no detalla cómo se diferencian los distintos niveles de cumplimiento de cada criterio. Esto es, el alumnado sigue sin conocer la diferencia entre un nivel 3 de cumplimiento y un nivel 2.

Además, el apartado que permite a la persona evaluadora matizar las puntuaciones otorgadas introduce subjetividad en el método de evaluación, ya que cede parte de la calificación a la opinión de dicha persona sin basarla en criterios pre-establecidos.

Tomando como base esta metodología, otros autores han realizado una propuesta de indicadores para valorar las competencias transversales de los TFG (Rullan et al., 2010). Esta propuesta define los indicadores para valorar las competencias transversales con independencia de la disciplina, lo que permite su aplicación en cualquier ámbito. Sin embargo, esta propuesta no es suficientemente completa, ya que además de las competencias transversales, hay que valorar las competencias técnicas específicas de cada disciplina.

La guía para la evaluación de competencias en los Trabajos Fin de Grado en el ámbito de las Ciencias Sociales y Jurídicas sin embargo, es mucho más específica ya que establece directamente los criterios a tener en cuenta para valorar tanto las

competencias transversales como las competencias técnicas correspondientes a su disciplina.

Esta guía establece una serie de criterios cuyo cumplimiento debe ser valorado por la persona evaluadora mediante los niveles D (De manera nada satisfactoria), C (De manera poco satisfactoria), B (De manera satisfactoria) y A (De manera muy satisfactoria). Para poder aprobar el TFG es condición indispensable alcanzar, de media, el nivel C en todos los elementos que forman parte de la evaluación.

Sin embargo, la guía no incluye ninguna indicación que establezca las diferencias entre cada uno de los niveles más allá de la opinión personal de la persona que está realizando la evaluación. Por lo tanto, este método de evaluación, no permite conocer al detalle qué condiciones debe cumplir un trabajo para obtener una calificación determinada.

2.1.2 Universidad Autónoma de Barcelona

En la Universidad Autónoma de Barcelona han realizado una propuesta de metodología orientada al grado en Administración y Dirección de Empresas, pero suficientemente general como para poder ser adaptada a otras titulaciones (Vilardell Riera, 2010).

Hitos de evaluación

Esta metodología define la realización de un TFG como una serie de actividades a lo largo del curso en las que el alumnado debe tomar parte. La elaboración de un trabajo en grupo o cooperativo es una más de las actividades que se deben realizar.

La elaboración del TFG se plantea como un trabajo a realizar por un grupo de 3 personas. El trabajo realizado se evalúa al finalizarlo mediante un único hito de evaluación, tal y como se muestra en la Tabla 2-4. A pesar de incluir un único hito de evaluación, esta metodología se ha incluido en esta revisión bibliográfica porque plantea realizar a lo largo del curso una evaluación formativa informal de las actividades definidas en la metodología. Por lo tanto, se puede considerar que esta metodología realiza una evaluación continua del TFG a pesar de que la evaluación oficial se realice únicamente al final.

Tabla 2-4.- Evaluación propuesta por (Vilardell Riera, 2010)

| Hitos de evaluación | Elementos evaluables | Peso | Responsables evaluación |
|---------------------|---|-------|--|
| Final del TFG | Documentación y presentación oral | 60% | Comisión evaluadora formada por 2 personas |
| | Informe sobre el desarrollo del TFG | 32.5% | Persona supervisora |
| | Presentación oral | 2.5% | Alumnado del curso |
| | Asistencia y participación en las presentaciones del resto del alumnado | 5% | Persona moderadora |

En esta metodología las competencias asociadas al TFG se evalúan a través de un sistema que combina evaluación formativa, sumativa y entre iguales. La comisión evaluadora realiza una evaluación sumativa de la documentación y de la presentación oral, mientras que la persona supervisora realiza una evaluación formativa de la documentación asociada al desarrollo del trabajo. La presentación oral la evalúa el resto del alumnado del curso en unas jornadas en las que cada estudiante presenta su trabajo y la persona que modera las presentaciones se encarga de valorar tanto la asistencia como la participación del alumnado.

Instrumentos de evaluación

Para realizar la evaluación de los distintos elementos propuestos en esta metodología, existen una serie de pautas, pero no se utiliza ningún instrumento que garantice una aplicación de criterios homogénea. Las personas encargadas de realizar la evaluación usan su propia experiencia y su opinión personal para emitir una calificación.

Por lo tanto, esta metodología no aporta soluciones a los problemas referentes a la falta de criterios claros y públicos para evaluar los TFG y tampoco favorece la consecución del objetivo específico de *Establecer un método de evaluación equitativo* (OE-5).

2.1.3 Universidad de Castilla - La Mancha

En el ámbito de la Administración y Dirección de Empresas, en la Facultad de Ciencias Sociales de Talavera de la Reina han definido una metodología para la evaluación de los TFG (M. I. Bonilla & Martín, 2012).

Hitos de evaluación

En esta propuesta, se definen 3 hitos de evaluación (ver Tabla 2-5): *Inicio*, *Progreso* y *Final*.

Tabla 2-5.- Evaluación propuesta por (M. I. Bonilla & Martín, 2012)

| Hitos de evaluación | Elementos evaluables | Peso | Responsables evaluación |
|----------------------------|---|-------------|--|
| Inicio | Informe y exposición oral | 10% | Persona supervisora / miembros alumnado / personas ajenas al TFG |
| Progreso | Informe | 20% | Persona supervisora / Autoevaluación |
| Final | Documentación final y presentación oral | 70% | Tribunal de evaluación |

En el hito de *Inicio* se valora un informe inicial que contiene una descripción del trabajo a realizar y la exposición oral del mismo. Aunque en principio la valoración de este elemento la realiza la persona supervisora, la metodología incluye la posibilidad de que participen en la evaluación otros miembros del alumnado así como miembros del profesorado y profesionales expertos en la materia. En el hito de *Progreso* se evalúa un informe y se tiene en cuenta el trabajo desarrollado y la constancia demostrada por el o la estudiante. La persona supervisora realiza la evaluación en este hito, aunque también se contempla que sea el propio estudiante quien aporte un informe de autoevaluación que forme parte de la calificación. Por último, en el hito *Final* los componentes del tribunal de evaluación tienen en cuenta tanto la documentación final como la presentación oral realizada.

Instrumentos de evaluación

Para realizar la evaluación de los elementos evaluables, existen una serie de indicadores asociados a las competencias de los TFG en el área, que cada persona evaluadora debe calificar con una nota entre 0 y 10 en base a su opinión personal.

Este método de evaluación deja en manos de cada persona que participa en la evaluación, la decisión de establecer las condiciones que debe cumplir un trabajo para asignarle una puntuación determinada en cada indicador. Esto puede generar desconocimiento entre el alumnado sobre la calidad que debe presentar su trabajo para lograr una nota determinada y dar lugar a distintos niveles de exigencia en función de la persona que esté evaluando el trabajo

2.1.4 Universidad de Oviedo

En la Universidad de Oviedo han definido una metodología para los Trabajos Fin de Grado en Derecho (M. Álvarez & Pascual, 2012).

Hitos de evaluación

Esta metodología define 5 hitos, cada uno de ellos con un elemento evaluable a excepción del último, en el que además de la documentación final también se evalúa una defensa oral (ver Tabla 2-6).

Tabla 2-6.- Evaluación propuesta por (M. Álvarez & Pascual, 2012)

| Hitos de evaluación | Elementos evaluables | Peso | Responsables evaluación |
|----------------------------|------------------------------------|-------------|--------------------------------|
| Elección del tema | Informe | 3% | Persona supervisora |
| Planificación | Informe | 7% | Persona supervisora |
| Desarrollo | Informe | 15% | Persona supervisora |
| Seguimiento | Informe | 25% | Persona supervisora |
| Final | Documentación final y defensa oral | 50% | Tribunal de evaluación |

Esta metodología incorpora un primer hito consistente en la evaluación por parte de la persona supervisora del tema elegido por el o la estudiante para la realización de su TFG. Esto permite que si el tema seleccionado no tiene la calidad suficiente, el trabajo pueda ser reconducido desde el principio.

La persona supervisora realiza la evaluación correspondiente a los cuatro primeros hitos definidos y un tribunal realiza la evaluación del hito final. De esta manera, un 50% de la calificación depende de la persona supervisora y el otro 50% del tribunal de evaluación.

Instrumentos de evaluación

Para obtener la calificación de cada uno de los elementos evaluables definidos, se usan cuatro rúbricas de evaluación, definidas a partir del trabajo de (Valderrama et al., 2010) y adaptando los criterios al ámbito del derecho.

Cada una de las rúbricas utilizadas en esta metodología evalúa una competencia distinta a través de una serie de indicadores. Para cada indicador se definen cuatro niveles de cumplimiento y una descripción explícita de las condiciones que debe cumplir el trabajo evaluado para alcanzar cada nivel de valoración.

Tabla 2-7.- Ejemplo de rúbrica para la evaluación de una competencia. Adaptada de (M. Álvarez & Pascual, 2012)

| Competencia: Capacidad para buscar, obtener y seleccionar información | | | | |
|---|---|--|--|--|
| | No se cumple 0 | Mínimo 1 | Medio 2 | Excelente 3 |
| Accede de manera correcta a bases de datos doctrinales, de legislación y jurisprudencia | No acude a bases de datos | Utiliza fundamentalmente motores de búsqueda de Internet | Demuestra que accede a bases de datos especializadas | Amplio uso de bases de datos y revistas especializadas |
| Selecciona la doctrina, la legislación y la jurisprudencia más relevante | Selección deficiente de información | Recopilación indiscriminada de información | Recopilación básica de información relevante | Selección exhaustiva y ordenada de información relevante |
| Sistematiza las líneas doctrinales y jurisprudenciales relacionadas con el objeto del trabajo | No hay aportación de líneas jurisprudenciales y doctrinales | Reproduce la información obtenida sin interpretar | Identifica y analiza exclusivamente la principal línea doctrinal y jurisprudencial | Identifica y analiza adecuadamente diversas líneas doctrinales y jurisprudenciales |
| Correcto uso de las citas doctrinales, legislativas y jurisprudenciales | No hace citas | Las citas son escasas o manifiestamente intrascendentes | Hay suficientes citas jurisprudenciales | Combina adecuadamente citas jurisprudenciales y doctrinales relevantes |

En cada uno de los hitos de evaluación de esta metodología sólo se evalúan algunos de los indicadores de determinadas rúbricas. Así por ejemplo, en el hito de *Planificación* se evalúan sólo tres indicadores, cada uno de ellos perteneciente a una rúbrica distinta. La calificación de cada hito viene dada por el valor medio de los niveles de cumplimiento obtenido en base a los indicadores a evaluar en ese hito. Dicha calificación se encuentra en la escala [0,3], por lo que hay que trasladarla a la escala [0,10] tal y como establece el sistema universitario español.

A pesar de la definición detallada de las rúbricas de evaluación que propone esta metodología, el hecho de que en cada hito de evaluación sólo se utilicen parte de los indicadores de ciertas rúbricas, complica su uso. Además, para la obtención de la calificación de cada hito se obtiene el valor medio de los niveles de cumplimiento de los distintos indicadores, lo que impide priorizar unos indicadores sobre otros. Priorización que puede resultar interesante cuando se deseen evaluar indicadores cuyo cumplimiento no sea obligado, pero sea indicativo de la excelencia de un TFG.

2.1.5 Universidad Tecnológica Nanyang (Singapur)

En la Universidad Tecnológica Nanyang en Singapur proponen una metodología de evaluación para los proyectos de Ingeniería Eléctrica (Teo & Ho, 1998).

Hitos de evaluación

La metodología propuesta se basa en 5 hitos de evaluación. En cada uno de estos hitos participa una única persona que puede ser la persona supervisora o la persona a la que denominan moderadora, que evalúa el trabajo una vez finalizado (ver Tabla 2-8).

Tabla 2-8.- Evaluación propuesta por (Teo & Ho, 1998)

| Hitos de evaluación | Elementos evaluables | Peso | Responsables evaluación |
|------------------------------|---|------|-------------------------|
| Informe inicial | Informe técnico | 15% | Persona supervisora |
| Documentación final | Informe técnico | 50% | Persona supervisora |
| Presentación oral | Presentación oral | 5% | Persona supervisora |
| Documentación y demostración | Documentación final y demostración práctica | 25% | Persona moderadora |
| Presentación oral | Presentación oral | 5% | Persona moderadora |

Los elementos evaluables definidos en esta propuesta que evalúa la persona supervisora son: un primer informe sobre el trabajo que se va a desarrollar; la versión final de la documentación y la presentación oral. El conjunto de estos tres elementos supone un 70% de la nota final.

Por su parte, la persona moderadora evalúa por un lado la documentación final del proyecto junto con una demostración práctica; y por otro lado la presentación oral.

Instrumentos de evaluación

La evaluación de los cinco elementos evaluables se realiza asignando un valor numérico entre el 1 y 8 a cada uno de los criterios definidos para los distintos elementos.

Así por ejemplo, los criterios a emplear para evaluar el informe inicial son: el establecimiento de objetivos y el plan de trabajo; el entusiasmo y responsabilidad mostrados por el o la estudiante; el contenido del informe; los logros y resultados alcanzados hasta el momento y el plan de acciones para el futuro.

Cada uno de estos criterios va acompañado de una pequeña descripción de los aspectos a tener en cuenta en su evaluación, pero el valor numérico concreto a asignar a cada uno de los criterios se decide en base únicamente a su opinión personal.

Este método de evaluación provoca que el alumnado, a pesar de conocer los criterios que van a usarse durante la evaluación, desconozca el nivel de exigencia de los mismos.

2.1.6 Universidad de Ciudad del Cabo (Sudáfrica)

En la Facultad de Ingeniería y Urbanismo de la Universidad de Ciudad del Cabo en Sudáfrica se imparten 5 programas de ingeniería distintos y en todos ellos existe un proyecto final que el alumnado debe superar para finalizar sus estudios (Jawitz et al., 2002). Cada uno de estos programas sigue una metodología distinta, por lo que a continuación se presenta un resumen de todas ellas.

Hitos de evaluación

En la Fig. 2-1 se recoge un resumen del proceso de evaluación de los distintos tipos de proyecto que se pueden realizar dentro de cada uno de los programas de ingeniería (denominados A, B, C, D y E).

| Programme—Title | Class size | Students work | Assessment item | | | | | | | | |
|-----------------------|------------|---------------|-----------------|-----------------|--------------------|--------------|--------------|---------|------|--------|--------|
| | | | Prop-osal | Literat. Review | Work plan & method | Super's mark | Draft report | Seminar | Oral | Poster | Report |
| A—Engineering Project | Med | in pairs | 5% | | | 5% | | 5% | 10% | 10% | 65% |
| A—Engineering Design | Med | indiv | | | | | | | | | 65% |
| B—Thesis | Med | group | | 10% | 30% | | 10% | | 5% | 10% | 40% |
| B—Design Project | Med | indiv | | | | | | | | 25% | 25% |
| C—Thesis Project | Large | group | | | | | | | | DP | DP |
| D—Laboratory Project | Small | indiv | | | | 10% | | 10% | | 10% | 70% |
| D—Design Project | Small | group | | | | | | | | | 100% |
| E—Individual | Large | indiv | | | | | | | | DP | DP |
| E—Research Project | | | | | | | | | | | |
| E—Design Project | Large | group | | | | | | | | 15% | 85% |

Note: DP—required but not included in summative assessment

Fig. 2-1.- Resumen evaluación proyectos (Jawitz et al., 2002)

Como se puede observar, dependiendo del programa y del tipo de proyecto a realizar el método de evaluación propuesto varía. Hay algunos proyectos en los que únicamente se valora el informe final (D-Design Project) y otros en los que se valoran hasta seis elementos distintos (A-Engineering Project).

También existen proyectos (por ejemplo, C-Thesis Project) en los que es necesario entregar elementos que no forman parte de la evaluación sumativa y por

lo tanto no forman parte de la calificación final (indicado mediante las siglas DP en la figura).

Entre los elementos que se tienen en cuenta en la valoración de los distintos tipos de proyecto están: la propuesta inicial del proyecto (Proposal), la revisión bibliográfica asociada (Literat. Review), la opinión de la persona supervisora sobre el proceso de desarrollo del proyecto (Super's mark) o la elaboración de un poster (Poster). El único elemento que está presente en la evaluación de todos los tipos de proyecto es la documentación final (Report), cuyo peso en la nota final varía desde un 25% hasta un 100% dependiendo del tipo de proyecto y del programa en el que se realice.

Instrumentos de evaluación

La evaluación de los elementos evaluables en los distintos programas se realiza, de manera general, en base a la opinión personal de las personas encargadas de la evaluación. En algunos de los programas existen listas de criterios a tener en cuenta para valorar cada elemento, pero su uso no es obligatorio.

Por lo tanto, el alumnado no dispone de unos criterios claros y equitativos que le permitan conocer cómo se va a evaluar su proyecto y aseguren que todos los proyectos serán evaluados por igual.

2.1.7 Universidad de Victoria (Australia)

La Universidad de Victoria en Australia dispone de una metodología para la evaluación de los proyectos que se realizan en el ámbito de la Ingeniería Eléctrica (Ku & Goh, 2010).

Hitos de evaluación

Los proyectos de Ingeniería Eléctrica de la Universidad de Victoria se evalúan a través de seis elementos evaluables (ver Fig. 2-2), que incluyen un estudio sobre la viabilidad del proyecto, cuatro presentaciones orales en distintos momentos del curso académico y una documentación final.

| Project contract | Weighting (%) |
|---|---------------|
| Feasibility study | 5 |
| Mid of first semester presentation | 10 |
| End of first semester presentation | 10 |
| Mid of second semester presentation | 10 |
| End of second semester presentation and demonstration | 40 |
| Final project report | 25 |
| Total marks | 100 |

Fig. 2-2.- Evaluación de los proyectos final de carrera de Ingeniería Eléctrica en la Universidad de Victoria (Ku & Goh, 2010)

El tribunal de evaluación evalúa únicamente la última presentación, lo que supone un 40% de la nota final del proyecto. El resto de elementos son calificados únicamente por la persona supervisora del proyecto.

Instrumentos de evaluación

Para realizar la evaluación de los distintos elementos, existe una descripción de los criterios que deben tenerse en cuenta para cada elemento evaluable, y en base a esos criterios cada persona emite una valoración usando su opinión personal.

A pesar de que la existencia de unos criterios permita conocer qué se va a valorar del trabajo realizado, el hecho de que cada persona emita su valoración en base a su opinión personal provoca que no se pueda saber cuál es el nivel de exigencia requerido para obtener una calificación determinada.

2.1.8 Universidad de Western Australia (Australia)

En la Universidad de Western Australia han definido una metodología de evaluación para los proyectos que se realizan en su Escuela de Ingeniería Mecánica².

Hitos de evaluación

Esa propuesta presenta tres hitos de evaluación en los que se evalúan 4 elementos evaluables (ver Tabla 2-9): la propuesta inicial del proyecto; un informe de progreso intermedio; la presentación de un artículo de investigación y la documentación final.

² <https://uims.research.uwa.edu.au/>

Tabla 2-9.- Evaluación de los proyectos en la Escuela de Ingeniería Mecánica de la Universidad de Western Australia³

| Hitos de evaluación | Elementos evaluables | Peso | Responsables evaluación |
|----------------------------|-----------------------------|-------------|--------------------------------|
| Inicio | Propuesta inicial | 10% | Persona supervisora |
| Seguimiento | Informe de progreso | 5% | Persona supervisora |
| Final | Presentación artículo | 10% | Dos miembros del profesorado |
| | Documentación final | 75% | Panel de personas expertas |

La persona supervisora realiza la evaluación de los dos primeros elementos. La presentación del artículo de investigación la evalúan dos miembros del profesorado entre los que no está la persona que realiza la supervisión. La documentación final, por su parte, es evaluada por un panel de personas expertas en la materia del proyecto.

Instrumentos de evaluación

La evaluación de los distintos elementos evaluables definidos en la metodología se realiza en base a la opinión personal de las personas encargadas de su evaluación.

Por lo tanto, esta metodología no aporta soluciones al problema de falta de criterios de evaluación objetivos y públicos.

2.1.9 Resumen

A lo largo de esta sección se han analizado distintas metodologías de realización y evaluación del TFG, detallando los hitos de evaluación que presenta cada una de ellas y los elementos que se tienen en cuenta en su evaluación.

Todas las metodologías analizadas establecen un mínimo de tres hitos de evaluación: uno al comienzo del trabajo, otro durante el desarrollo y un último hito al finalizar. Sin embargo, algunas metodologías establecen más hitos intermedios e incluso un hito dedicado exclusivamente a evaluar si la temática seleccionada para el trabajo es adecuada o no.

³ <https://uims.research.uwa.edu.au/>

El porcentaje de la calificación final del TFG correspondiente a cada hito varía considerablemente entre las distintas metodologías. También varían la persona encargada de evaluar cada uno de los hitos y los elementos a evaluar en los mismos.

Si bien la mayoría de evaluaciones se realizan en base a informes escritos, también se valoran presentaciones orales e incluso alguna de las metodologías valora aspectos como la participación en ciertas actividades.

La mayoría de las metodologías estudiadas aportan una serie de indicaciones de los aspectos que deben tenerse en cuenta a la hora de realizar la evaluación, pero dejan en manos de las personas evaluadoras la decisión sobre qué nivel de exigencia establecer para cada puntuación.

2.2 Métodos de evaluación

Los métodos de evaluación se pueden englobar en dos aproximaciones en función del nivel de detalle con el que se definen los criterios de evaluación (Swartz et al., 1999): aproximación holística y aproximación analítica (ver Fig. 2-3).



Fig. 2-3.- Clasificación de métodos de evaluación. Adaptada de (Hunter et al., 1996)

En la aproximación holística, los criterios de evaluación se describen de manera superficial, el objeto de evaluación se evalúa como un todo y se emite una valoración global del mismo. Sin embargo, en la aproximación analítica, los criterios de evaluación se detallan en mayor profundidad. Se evalúan las distintas partes que componen el objeto de evaluación recibiendo cada una de esas partes una valoración individual (Hunter et al., 1996; Swartz et al., 1999).

El método de evaluación usado tradicionalmente en la evaluación de los TFG es la «Impresión general» (Fig. 2-3→a), que se enmarca dentro de la aproximación holística. En este método de evaluación, la persona evaluadora emite una valoración personal sobre el resultado final del TFG, teniendo en cuenta su experiencia y los criterios de evaluación que considere adecuados.

El hecho de que sólo se evalúe el resultado final y que la deliberación del tribunal no esté sujeta a ningún tipo de pautas o criterios, puede ser el origen de algunos de los problemas identificados (falta de un proceso continuo de evaluación y falta de criterios de evaluación objetivos y públicos, ver sección 1.2). Para evitar estos problemas, algunos autores proclaman que explicitar los criterios de evaluación en los TFG puede proporcionar soluciones a esta problemática (Jawitz et al., 2002; Littlefair & Gossman, 2008).

Los criterios se pueden explicitar, por ejemplo, utilizando listas de control (*checklists*), escalas de clasificación (*rating scales*) o rúbricas. Sin embargo, entre todas estas opciones, las rúbricas son el instrumento más completo y que mejor permite expresar explícitamente las expectativas asociadas a cada calificación (Dandis, 2014).

Por este motivo, el uso de rúbricas de evaluación se está extendiendo (Brookhart, 1999; Hafner & Hafner, 2003; Jonsson & Svingby, 2007; Stevens et al., 2012) y se han seleccionado como instrumento de evaluación en la metodología FPM.

Atendiendo al nivel de detalle con el que se definen las rúbricas de evaluación, éstas pueden situarse dentro de la aproximación holística o de la aproximación analítica (Moskal, 2000). En la aproximación holística se encuentran las rúbricas denominadas holísticas (Fig. 2-3→b), mientras que en la aproximación analítica se distinguen dos tipos de rúbricas: las analíticas (Fig. 2-3→c) y las atomísticas. (Fig. 2-3→d).

A continuación se detallan los distintos tipos de rúbrica junto con sus componentes y su uso tanto en evaluación formativa como en evaluación sumativa.

2.2.1 Rúbricas de evaluación

La utilización de rúbricas de evaluación reduce la ansiedad de los estudiantes ante la evaluación (Panadero & Jonsson, 2013) y la cantidad de quejas sobre la calificación recibida (Barney et al., 2012). También establecen pautas para otorgar una valoración justa y consistente, por lo que aumenta la objetividad en la evaluación (Peat, 2006; Stevens et al., 2012; Valderrama et al., 2009).

El uso de rúbricas también permite al profesorado abordar los cuatro aspectos fundamentales relacionados con la evaluación del alumnado (Stewart & Nesbitt, 2005):

- Explicitar los criterios precisos de calificación.
- Cuantificar las calificaciones para cada criterio.
- Validar la nota final.
- Mantener la uniformidad en la calificación entre el alumnado.

Independientemente del tipo de rúbrica, su definición implica establecer los siguientes cuatro elementos (Stevens et al., 2012):

- Descripción de la tarea: Una breve descripción de la actividad o del comportamiento general que se va a evaluar.
- Escala: Los diferentes niveles de valoración de la tarea. Se deben definir tantos niveles de valoración como distinciones significativas en la descripción sean capaces de realizar los creadores de la misma (Brookhart, 1999).
- Dimensiones: Cada una de las partes involucradas en la tarea o de los criterios de evaluación a tener en cuenta. A través de la asignación de un peso distinto para cada dimensión, permite ver la importancia relativa de los distintos elementos de la tarea.
- Descripción de cada nivel de valoración: expectativas específicas para cada dimensión respecto a la escala definida.

Rúbricas holísticas

Las rúbricas holísticas presentan una única dimensión que engloba todos los criterios que deben tenerse en cuenta en la evaluación. En este tipo de rúbricas se valora el trabajo realizado de manera global, sin especificar el nivel de desempeño alcanzado por el estudiante en cada uno de los criterios. En la Tabla 2-10 se muestra un extracto de una rúbrica holística para la evaluación de un informe escrito.

Tabla 2-10.- Extracto de una rúbrica holística

| Excelente | Avanzado | Medio | Básico | Insuficiente |
|---|---|---|--|--|
| El informe cubre todos los aspectos del tema de manera profunda y detallada. La redacción es correcta ... | El informe cubre todos los aspectos del tema a tratar, pero no entra en detalles. La redacción es correcta, sin faltas de ortografía... | El informe cubre los aspectos básicos del tema a tratar. La redacción es correcta, pero la sintaxis es mejorable... | Algunas partes del informe son mejorables. El contenido del mismo es superficial y la redacción presenta varios errores... | El informe no se ha realizado o no presenta el suficiente nivel de detalle en ninguno de sus aspectos... |

Las rúbricas holísticas son muy sencillas de utilizar. La persona encargada de la evaluación sólo tiene que valorar el trabajo de manera global e indicar cuál es el nivel de valoración cuya descripción encaja mejor con el trabajo realizado por el

estudiante y asignarle una puntuación numérica en los casos en los que se desee obtener una calificación a partir de la rúbrica.

Rúbricas analíticas

En las rúbricas pertenecientes a la aproximación analítica se tienen en consideración distintas partes del trabajo o distintos aspectos a evaluar y se genera una dimensión para cada uno de ellos. El nivel de detalle de la definición de cada dimensión es decisión de las personas que construyan la rúbrica, pudiendo generar una dimensión para cada uno de los criterios de evaluación o generar las dimensiones agrupando varios criterios. En función de dicha decisión se distingue entre las denominadas rúbricas analíticas y las rúbricas atomísticas.

El número de niveles de valoración típico de las rúbricas con aproximación analítica es 3, pero puede incrementarse o disminuirse en función de las necesidades y de la capacidad de las personas que definan la rúbrica para especificar descripciones significativamente distintas para cada nivel (Stevens et al., 2012).

En las rúbricas analíticas, los criterios de evaluación se agrupan para formar dimensiones más generales y se establece una única descripción por cada nivel de valoración combinando todos los criterios implicados. Por ejemplo, en el extracto de rúbrica analítica para un informe escrito que se muestra en la Tabla 2-11, se observa cómo para la dimensión «Redacción y aspecto visual» se tienen en cuenta distintos criterios como «Puntuación», «Ortografía» o «Uso del lenguaje formal». Combinando dichos factores, se genera una descripción distinta para cada uno los niveles de valoración de la rúbrica.

Tabla 2-11.- Extracto de una rúbrica analítica

| | Excelente | Avanzado | Medio | Básico | Insuficiente |
|-----------------------------------|---|---|--|--|--|
| ... | | ... | ... | ... | ... |
| Redacción y aspecto visual | El trabajo está correctamente puntuado, no presenta faltas de ortografía y emplea un lenguaje adecuado, lo usa correctamente... | El trabajo está correctamente puntuado, no presenta faltas de ortografía y emplea un lenguaje adecuado, pero no lo usa correctamente... | El trabajo está correctamente puntuado, presenta menos de tres faltas ortográficas y emplea un lenguaje adecuado, pero no lo usa correctamente | El trabajo está correctamente puntuado en su mayoría, presenta menos de tres faltas ortográficas y el lenguaje empleado es inadecuado... | El trabajo no está correctamente puntuado, presenta más de tres faltas de ortografía y el lenguaje empleado es inadecuado... |
| Contenido | | ... | ... | ... | ... |

En el caso de las rúbricas atomísticas se define una dimensión para cada uno de los criterios a tener en cuenta en la evaluación (Hunter et al., 1996; Nilson, 2010). En la Tabla 2-12 se muestra el extracto de una rúbrica atomística correspondiente a los criterios relativos a la redacción y al aspecto visual del trabajo evaluado. En ella se observa cómo los criterios «Puntuación», «Ortografía» y «Uso de lenguaje formal» entre otros, se corresponden con las dimensiones y cada uno de ellos tiene una descripción explícita, asociada a cada uno de los niveles de valoración.

Tabla 2-12.- Extracto rúbrica atomística

| | Nivel alto | Nivel medio | Nivel bajo |
|-------------------------------|--------------------------------------|---|-----------------------------------|
| ... | ... | .. | ... |
| Puntuación | Correcta | Correcta en su mayoría | Incorrecta |
| Ortografía | Sin faltas ortográficas | Dos o tres faltas | Más de tres faltas |
| Uso de lenguaje formal | Uso correcto y adecuado del lenguaje | Lenguaje adecuado, pero usado incorrectamente | Uso inadecuado del lenguaje |
| Consistencia y estilo | Estilo consistente y adecuado | Estilo consistente pero inadecuado | Estilo inconsistente e inadecuado |
| Sintaxis y gramática | Correcta | Correcta en su mayoría | Incorrecta |
| ... | ... | ... | ... |

2.2.2 Uso de rúbricas en evaluación formativa

Debido a sus características, las rúbricas de evaluación están especialmente orientadas a su uso en evaluaciones de tipo formativo donde lo más importante es la retroalimentación que se proporciona al estudiante para que mejore sus resultados (Sadler, 1989).

Sin embargo, las rúbricas holísticas presentan varios problemas de cara a proporcionar una evaluación formativa al alumnado. La retroalimentación generada por este tipo de rúbricas es bastante limitada, ya que el alumnado no puede conocer hasta qué punto ha cumplido las expectativas en cada uno de los criterios a tener en cuenta. Si un estudiante recibe una valoración media, ¿significa que ha alcanzado un nivel medio en todos los criterios de evaluación? o ¿significa que en unos

criterios ha alcanzado un nivel alto, mientras que en otros su nivel ha sido bajo, compensándose unos con otros?

Las rúbricas holísticas también presentan dificultades a la hora de usarlas como herramienta de autoevaluación por parte del alumnado. Si el trabajo realizado no encaja exactamente con ninguna de las descripciones, la valoración final del mismo dependerá en gran medida de la opinión personal y del criterio de la persona evaluadora. De este modo, la autoevaluación pierde su utilidad como herramienta que permite a los estudiantes prever la calificación que recibirá su trabajo.

Por lo tanto, el uso de rúbricas de tipo holístico en evaluaciones formativas presenta un gran nivel de subjetividad y la retroalimentación que proporciona es bastante limitada.

Por su parte, las rúbricas englobadas en la aproximación analítica, tanto las atomísticas como las analíticas, permiten al alumnado recibir una retroalimentación mucho más detallada que las de tipo holístico, lo que las convierte en una herramienta apropiada para realizar una evaluación formativa (Panadero & Jonsson, 2013).

Al usar rúbricas de tipo analítico la persona evaluadora está explicitando el nivel de desempeño alcanzado por el alumnado en cada una de las partes o cada uno de los criterios de evaluación del trabajo evaluado. De este modo, el alumnado conoce las debilidades y fortalezas del trabajo realizado y puede tomar las medidas necesarias para consolidar sus fortalezas y enmendar sus debilidades.

Asimismo, este tipo de rúbrica permite definir con mayor nivel de detalle cada criterio o parte a evaluar, por lo tanto, se reduce la subjetividad y se aumenta su efectividad como herramienta de autoevaluación.

2.2.3 Uso de rúbricas en evaluación sumativa

A pesar de su idoneidad como herramientas de evaluación formativa, en ocasiones las rúbricas también se usan en evaluaciones de tipo sumativo. En la evaluación sumativa es necesario obtener una valoración final, habitualmente numérica, que permita clasificar el trabajo realizado por el estudiante (Taras, 2005).

En el caso de las rúbricas holísticas, el cálculo de una calificación numérica es muy sencillo, ya que la persona evaluadora lo asigna directamente en función de su criterio teniendo en cuenta la descripción de los distintos niveles de valoración y si éstos están acompañados de el rango de valores numéricos aplicable a dicho nivel (ver Tabla 2-10).

Tabla 2-13.- Extracto de una rúbrica holística con sugerencia de valores numéricos para los niveles

| Excelente 9 - 10 | Avanzado 8 - 8,9 | Medio 7 - 7,9 | Básico 5- 6,9 | Insuficiente 0 - 4,9 |
|---|---|---|--|--|
| El informe cubre todos los aspectos del tema de manera profunda y detallada. La redacción es correcta ... | El informe cubre todos los aspectos del tema a tratar, pero no entra en detalles. La redacción es correcta, sin faltas de ortografía... | El informe cubre los aspectos básicos del tema a tratar. La redacción es correcta, pero la sintaxis es mejorable... | Algunas partes del informe son mejorables. El contenido del mismo es superficial y la redacción presenta varios errores... | El informe no se ha realizado o no presenta el suficiente nivel de detalle en ninguno de sus aspectos... |

Para poder calcular la calificación numérica generada por una rúbrica de evaluación de tipo analítico, es necesario asignar un valor numérico a cada nivel de valoración y un peso porcentual a cada una de las dimensiones.

Para asignar tanto los valores numéricos como los pesos porcentuales existen varias alternativas en función de las características de la evaluación a llevar a cabo y del comportamiento que se desee que presente la rúbrica. Realizar esta asignación de manera que se consigan los resultados esperados no es tarea sencilla (Allen & Knight, 2009; Peeters et al., 2010; Salinas & Erochko, 2015; Villamañe, Álvarez, et al., 2016b).

Por ejemplo, los niveles de valoración de la rúbrica pueden estar distribuidos de manera equitativa o no. Cuando los niveles de valoración se distribuyen de manera equitativa, el nivel de valoración que se encuentra en la mitad de la escala se corresponde realmente con la valoración de un trabajo medio. Sin embargo, lo habitual en las rúbricas de evaluación es que la distribución de los niveles de valoración no sea equitativa y existan más niveles por encima de lo que se considera una evaluación media que por debajo (Trice, 2000). Esto es debido a que normalmente se suele dejar el nivel de valoración más bajo para aquellos casos en los que no se cumplen las expectativas, y el resto de niveles se usan para precisar y detallar hasta qué punto se cumplen.

Para asignar un valor numérico a una rúbrica con niveles de valoración distribuidos equitativamente, los valores numéricos asociados a los niveles deben ser secuenciales, comenzando típicamente en 0 ó en 1 para el nivel de valoración más bajo e incrementándolo para los superiores.

En la Tabla 2-14 se muestra el extracto de una rúbrica cuyos niveles son equitativos junto con los valores numéricos asociados a dichos niveles.

Tabla 2-14.- Extracto de rúbrica con niveles de valoración equitativos

| | Nivel alto 3 | Nivel medio 2 | Nivel bajo 1 |
|-------------------|-------------------------|--------------------------|-------------------------|
| ... | ... | .. | ... |
| Puntuación | Correcta | Correcta en su mayoría | Incorrecta |
| Ortografía | Sin faltas ortográficas | Dos o tres faltas | Más de tres faltas |
| ... | ... | ... | ... |

Cuando los niveles de valoración no se reparten de manera equitativa, se asigna a cada nivel de valoración un valor numérico pero no tiene por qué respetarse la equidistancia entre los valores de los distintos niveles.

En la Tabla 2-15 se muestra el extracto de una rúbrica cuyos niveles de valoración no se distribuyen de manera equitativa ya que una valoración media del trabajo se correspondería con el nivel denominado «Básico», dejando un único nivel por debajo y tres por encima. En este caso, los valores numéricos asignados a cada nivel no son equidistantes y son establecidos directamente por la persona encargada de diseñar la rúbrica.

Tabla 2-15.- Extracto de una rúbrica con niveles de valoración no equidistantes

| | Excelente 10 | Avanzado 9 | Medio 7,5 | Básico 5 | Insuficiente 3 |
|-----------------------------------|---|--|---|--|---|
| ... | | ... | ... | ... | ... |
| Redacción y aspecto visual | El trabajo está correctamente puntuado, no presenta faltas de ortografía y el lenguaje empleado es adecuado y se usa correctamente... | El trabajo está correctamente puntuado, no presenta faltas de ortografía y el lenguaje empleado es adecuado, pero no se usa correctamente... | El trabajo está correctamente puntuado, presenta menos de tres faltas ortográficas y el lenguaje empleado es adecuado, pero no se usa correctamente | El trabajo está correctamente puntuado en su mayoría, presenta menos de tres faltas ortográficas y el lenguaje empleado es inadecuado... | El trabajo está incorrectamente puntuado, presenta más de tres faltas de ortografía y el lenguaje empleado es inadecuado... |
| Contenido | | ... | ... | ... | ... |

A la hora de calcular la calificación numérica generada por la rúbrica, la ecuación (2-1) se usa en aquellas situaciones en las que la distribución de los niveles de valoración es equitativa.

$$NotaRúbrica = \frac{\sum W_i \times (g_i - min)}{\sum W_i \times (max - min)} \quad (2-1)$$

donde g_i representa el valor numérico asociado al nivel de valoración seleccionado para la dimensión i ; W_i es el peso asociado a la dimensión i ; e min y max delimitan el rango de valores asociados a los niveles de valoración de la rúbrica.

Sin embargo, esta fórmula no es válida en aquellos casos en los que los niveles de valoración no están distribuidos de manera equitativa. Por ejemplo, si en la rúbrica cuyo extracto se muestra en la Tabla 2-15, suponemos cuatro dimensiones, todas ellas con el mismo peso asociado (25%) y la persona que realiza la evaluación decide seleccionar el nivel «Medio» en todas las dimensiones, la fórmula obtendría una nota de 0,64 (en la escala [0,1]) como se muestra en la aplicación de la ecuación (2-1):

$$NotaRúbrica = \frac{\sum 0,25 \times (7,5 - 3)}{\sum 0,25 \times (10 - 3)} = 0,64$$

Sin embargo, la nota otorgada por la rúbrica debería ser de 0,75 ya que ese, trasladado a la escala [0,10], es el valor numérico asociado al nivel «Medio» que se ha seleccionado para todas las dimensiones.

Para solucionarlo, en aquellos casos en los que los niveles de valoración no están distribuidos de manera equitativa, la fórmula para obtener una calificación numérica es la representada en la ecuación (2-2).

$$NotaRúbrica = \frac{\sum W_i \times g_i}{\sum W_i \times max} \quad (2-2)$$

donde al igual que en la ecuación (2-1), g_i representa el valor numérico asociado al nivel de valoración seleccionado para la dimensión i ; W_i es el peso asociado a la dimensión i ; y max representa el mayor de los valores asociados a los niveles de valoración de la rúbrica.

Al aplicar esta ecuación al supuesto expuesto anteriormente, la nota obtenida, trasladada a la escala [0,10], sí se corresponde con el valor numérico del nivel de valoración seleccionado para todas las dimensiones (7,5), tal y como se muestra a continuación.

$$NotaRúbrica = \frac{\sum 0,25 \times 7,5}{\sum 0,25 \times 10} = 0,75$$

En aquellos casos en los que se considere que todas las dimensiones tienen la misma importancia y no se desee realizar una ponderación entre ellas, en ambas ecuaciones se puede asignar a todas las dimensiones el mismo valor para el elemento W_i o eliminarlo.

Ambas ecuaciones calculan una nota numérica en el rango [0, 1] que posteriormente habrá que trasladar a la escala numérica con la que se desee trabajar, en el caso español, típicamente [0, 10].

2.3 Las TIC en la realización y evaluación de los TFG

La experiencia del grupo GaLan en el desarrollo de herramientas informáticas y en la aplicación de las TIC en entornos educativos, le permite analizar y establecer las características técnicas necesarias en la consecución de los objetivos planteados en este trabajo de tesis (Villamañe, Ferrero, & Álvarez, 2014).

A continuación se detallan los requisitos funcionales que una herramienta debería cumplir para dar respuesta a la problemática descrita en el capítulo anterior (ver sección 1.2) y satisfacer los objetivos específicos definidos. También se analizan distintas herramientas existentes para ver hasta qué punto satisfacen dichos requisitos y se pueden emplear en la consecución del objetivo específico OE-6 (*Desarrollar las herramientas informáticas necesarias*).

2.3.1 Requisitos funcionales

Durante la realización y evaluación de un TFG existen numerosos aspectos que pueden repercutir en el mismo (ver sección 1.2). Con el fin de minimizar la aparición de problemas y disminuir su impacto, a partir del objetivo general (ver sección 1.1.2) de este trabajo de tesis, se han determinado una serie de objetivos específicos a satisfacer (ver sección 1.3).

Partiendo de dichos objetivos específicos, en este apartado se establecen los requisitos funcionales que las herramientas informáticas que den soporte a la realización y evaluación de los TFG deberían cumplir.

Los requisitos funcionales establecidos a partir del objetivo específico OE-1 (*Asegurar una calidad mínima en todos los TFG*) y del objetivo OE-5 (*Establecer un método de evaluación equitativo*) de este trabajo, hacen referencia a la necesidad de definir y hacer públicos criterios de evaluación equitativos.

Para la definición de criterios públicos de evaluación y que su aplicación sea lo más objetiva posible, independientemente de la persona encargada de aplicarlos, en este trabajo se ha optado por el uso de rúbricas de evaluación (ver sección 2.2). Por lo tanto, los requisitos extraídos de estos objetivos específicos hacen referencia a la definición y gestión de las rúbricas de evaluación.

Además, el objetivo específico OE-4 (*Establecer un proceso de evaluación continuo*) hace que la evaluación del TFG tenga que realizarse como la combinación de sucesivas evaluaciones, lo que influye en los requisitos que deben establecerse para definir la evaluación.

Los requisitos funcionales establecidos a partir de los objetivos específicos OE-1, OE-5 y OE-4 son:

REQ.1 *Definición y gestión de rúbricas.* Para poder trabajar con rúbricas es necesario que éstas se puedan definir, modificar y borrar cuando sea necesario. Además, la calificación final debe poder definirse como la combinación de varias evaluaciones, cada una de las cuales debe tener una rúbrica asociada y un peso determinado en el cálculo de la nota final. Asimismo, también es necesario poder indicar las personas encargadas de utilizar cada rúbrica.

REQ.2 *Consulta de criterios.* Es necesario que los criterios de evaluación establecidos a través de las rúbricas sean públicos y puedan ser consultados por cualquier persona.

REQ.3 *Uso de rúbricas.* La persona o personas encargadas de realizar la evaluación deben poder usar las rúbricas, y los resultados de las mismas deben combinarse para generar la calificación final. Una vez realizada la evaluación, su resultado debe estar accesible para todas las personas implicadas.

Los objetivos específicos OE-2 (*Aumentar y mejorar la retroalimentación que recibe el alumnado*) y OE-3 (*Incorporar nuevos canales y posibilidades de comunicación*), tienen que poder ser satisfechos sin incrementar significativamente la carga de trabajo de la persona supervisora.

Por lo tanto, se definen requisitos funcionales orientados a facilitar el trabajo de la persona supervisora ofreciéndole la posibilidad de realizar un seguimiento continuo del trabajo y de organizar la documentación. Esto repercutirá en la posibilidad de proporcionar retroalimentación y mejorar su calidad. Los requisitos a cumplir para satisfacer estos objetivos específicos son:

REQ.4 *Planificación temporal compartida.* Los estudiantes implicados en un TFG deben poder definir su planificación temporal e ir indicando los avances que van realizando a lo largo del tiempo. Además, las personas encargadas de la supervisión deben poder acceder a dicha planificación y consultar su estado en cualquier momento. De este modo, se pretende motivar al alumnado a

cumplir esa planificación que tiene establecida y a la que puede acceder la persona supervisora en cualquier momento.

REQ.5 *Gestión de la documentación.* Para facilitar la labor de supervisores y componentes del tribunal, es necesario que la documentación generada por el alumnado esté organizada y sea fácil de localizar. Para ello, es necesario que los documentos estén asociados a la fase del TFG que corresponden, facilitando así la búsqueda.

REQ.6 *Acceso a la documentación.* Debe controlarse quién tiene acceso a qué documentación, especialmente en el caso de los supervisores y de las personas encargadas de evaluar las versiones finales de la misma.

REQ.7 *Gestión de versiones.* Para facilitar la tarea de documentación es necesario proporcionar la gestión de las diferentes versiones de los elementos que se vayan generando y modificando a lo largo del tiempo.

REQ.8 *Retroalimentación.* Es necesario que la persona supervisora pueda proporcionar una retroalimentación sobre el trabajo realizado, de manera que quede almacenada para poder ser consultada en un futuro. Además, cuando dicha retroalimentación se corresponda con elementos concretos, deberá quedar asociada a los mismos para mayor claridad y comodidad a la hora de consultarla. El alumnado por su parte debe poder solicitar retroalimentación, cuando lo considere necesario, sobre aspectos generales o sobre algún elemento concreto de la documentación generada.

Además, la comunicación entre estudiantes también puede ser provechosa para quien empieza su TFG, ya que permite que pueda aprender de la experiencia de aquellas personas que ya se encuentran en las fases más avanzadas de la realización del TFG y evite caer en los mismos errores. Por ello, se define el siguiente requisito que también favorece el cumplimiento del objetivo específico OE-3:

REQ.9 *Realizar consultas en base a intereses comunes.* Esto facilitaría la comunicación entre estudiantes que estén realizando el TFG sobre una temática común, o entre todos los estudiantes que estén realizando su TFG para consultas sobre la documentación, trámites y similares.

Además de los requisitos establecidos a partir de los objetivos específicos definidos, se establece un último requisito, necesario para el correcto funcionamiento de cualquier herramienta informática.

REQ.10 *Administración.* Es necesario poder definir todos los datos administrativos. También debe controlarse el rol a ejercer por cada persona usuaria y restringir el acceso a los datos correspondientes.

2.3.2 Análisis de herramientas

Existe una gran variedad de herramientas disponibles que permiten satisfacer parte de los requisitos identificados en el apartado anterior. En este apartado se presenta un estudio de estas herramientas, analizando hasta qué nivel cubren los requisitos funcionales identificados en la sección anterior.

Las herramientas analizadas se han agrupado en categorías en función de sus características y sus funcionalidades: herramientas usadas actualmente en la UPV/EHU; herramientas de gestión de documentación; herramientas de comunicación; herramientas de gestión de proyectos y herramientas de trabajo con rúbricas.

Al finalizar la descripción de las herramientas de cada categoría, se presenta una tabla indicando si las herramientas descritas en esa categoría satisfacen («SI») o no (casillas sombreadas) los requisitos funcionales establecidos en el apartado anterior (sección 2.3.1). En caso de que satisfagan el requisito parcialmente o a través de la instalación de algún complemento, se indica mediante el símbolo «★» detallándose a continuación las limitaciones encontradas.

Para finalizar este apartado se presenta un resumen con las principales conclusiones del análisis de herramientas realizado.

2.3.2.1 Herramientas usadas actualmente en la UPV/EHU

En este apartado se analizarán las distintas plataformas que usa actualmente la UPV/EHU y que pueden resultar de utilidad a la hora de realizar y evaluar los TFG.

Análisis

Moodle⁴

Moodle es un sistema de gestión de aprendizaje (*Learning Management System, LMS*), o plataforma web de aprendizaje, usada en la UPV/EHU (bajo la denominación eGela) y una de las más usadas globalmente.

Moodle está enfocado a dar soporte virtual a la docencia permitiendo una gestión sencilla de los distintos cursos, las asignaturas y los participantes (profesorado y alumnado) en cada asignatura.

Así mismo proporciona herramientas que facilitan la creación y evaluación de contenidos. Moodle también cuenta con canales de comunicación tales como foros

⁴ <https://moodle.org/>

dentro de cada asignatura o la posibilidad de enviar mensajes entre los distintos usuarios.

Moodle también cuenta con una gran comunidad de usuarios que desarrollan complementos que se pueden instalar incorporando nuevas funcionalidades a la plataforma.

Sin embargo en la UPV/EHU, la gestión y administración del sistema se realiza de manera unificada, lo que dificulta la instalación de nuevos complementos e impide la creación de cursos que no se correspondan con la oferta oficial de la Universidad.

Sistema de Gestión Académica Universitaria Renovada (G.A.U.R.)⁵

Es el sistema de gestión usado en la UPV/EHU para todos los temas relacionados con la docencia, desde la asignación docente de las materias a impartir hasta la evaluación del alumnado. El sistema cuenta con distintos roles de usuario, de manera que lo usan tanto el profesorado como el alumnado.

G.A.U.R. incorpora un módulo específico para la gestión de los TFG al que tienen acceso tanto los supervisores como los estudiantes. Sin embargo, la funcionalidad de la que dispone actualmente dicho módulo es muy reducida, centrándose en la parte de oferta y asignación de TFG a supervisores y estudiantes, pero sin incorporar ningún protocolo o criterio de evaluación de dichas propuestas.

Sí que cuenta con un apartado denominado “Tutorías” donde es posible establecer las fechas de las tutorías a realizar entre supervisores y estudiantes, lo que puede considerarse como un canal de comunicación entre ambos.

ehubox⁶

Ehubox es el sistema de sincronización y almacenamiento en la nube que la UPV/EHU proporciona a su personal.

El espacio de almacenamiento disponible para cada persona usuaria es de 50GB, y dispone de clientes para los principales sistemas operativos. El principal problema de ehubox es que sólo está disponible para profesorado y personal investigador de la UPV/EHU. El alumnado no tiene acceso al mismo, lo que imposibilita su uso como herramienta de gestión de la documentación en la realización de los TFG.

⁵ <http://gestion.ehu.eus/gaur>

⁶ <https://ehubox.ehu.eus>

Archivo Digital para la Docencia y la Investigación (ADDI)⁷

ADDI es un depósito de documentos digitales, cuyo objetivo es organizar, archivar, preservar y difundir la producción intelectual generada en la UPV/EHU, entre la que se encuentran los TFG.

Actualmente ADDI está siendo usado por varios centros de la UPV/EHU como almacén de la documentación generada a lo largo de la realización del TFG. El alumnado es el encargado de depositar los documentos y la persona encargada de la supervisión debe validar cada uno de los documentos almacenados. Sin embargo, en el caso de que la persona encargada de la supervisión rechace un documento, el sistema no almacena las distintas versiones, ni la retroalimentación proporcionada sobre el mismo. ADDI sólo permite el acceso a las versiones de los documentos que hayan recibido el visto bueno de la persona supervisora.

Comparativa

La Tabla 2-16 recoge la comparativa correspondiente a las distintas herramientas que se utilizan actualmente en la UPV/EHU.

Tabla 2-16.- Comparativa herramientas usadas actualmente en la UPV/EHU

| | REQ.1 | REQ.2 | REQ.3 | REQ.4 | REQ.5 | REQ.6 | REQ.7 | REQ.8 | REQ.9 | REQ.10 |
|----------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| Moodle | ★ | ★ | ★ | ★ | ★ | ★ | ★ | SI | ★ | ★ |
| G.A.U.R | | | | | | | | ★ | | SI |
| ehubox | | | | | ★ | | ★ | | | |
| ADDI | | | | | ★ | SI | | ★ | | SI |

Entre las herramientas usadas actualmente en la UPV/EHU destaca Moodle, que mediante la instalación de distintos complementos es capaz de satisfacer varios de los requisitos establecidos como los relacionados con la gestión de las rúbricas de evaluación (REQ.1, REQ.2 y REQ.3), la planificación temporal compartida (REQ.4) y la gestión de documentación y de versiones (REQ.5, REQ.6 y REQ.7).

Sin embargo, la administración del sistema (REQ.10) que ofrece Moodle no es completamente satisfactoria. Para poder especificar la información de cada TFG, del alumnado que participa en su realización y del profesorado que participa en su supervisión y evaluación, es necesario definir un curso en la plataforma por cada

⁷ <https://addi.ehu.es>

TFG. Esto impide la comunicación entre el alumnado de distintos cursos y la compartición de recursos entre ellos, lo que impediría la posibilidad de realizar consultas en base a intereses comunes (REQ.9).

Además, las políticas de gestión de la plataforma educativa en cada centro o universidad puede imposibilitar la creación de dichos cursos. Tal y como ocurre en la UPV/EHU donde sólo se generan cursos correspondientes a las asignaturas que se imparten oficialmente, lo que posibilita la creación de un único curso para la asignatura Trabajo Fin de Grado, pero no uno para cada estudiante.

El sistema G.A.U.R. es el núcleo de la gestión del profesorado, alumnado y evaluaciones en la UPV/EHU, por lo que satisface el requisito relacionado con la administración del sistema (REQ.10). Este sistema también proporciona la opción de emitir una retroalimentación acompañando a cada evaluación. Sin embargo, no permite que el alumnado solicite retroalimentación sobre los elementos que considere oportuno. Sólo se puede emitir retroalimentación en los momentos que estén establecidos por defecto y que formen parte de la evaluación. Por lo tanto, la opción de retroalimentación que proporciona G.A.U.R. no se corresponde totalmente con retroalimentación que se desea poder proporcionar (REQ.8).

Por su naturaleza de herramienta de sincronización y almacenamiento en la nube, ehubox satisface los requisitos relacionados con la gestión de la documentación (REQ.5) y la gestión de versiones (REQ.7). Sin embargo, ehubox sólo está disponible para el profesorado, lo que impide su uso por parte del alumnado.

La herramienta ADDI importa la configuración de personas usuarias establecida en G.A.U.R. para cada TFG (REQ.10), lo que permite controlar quién tiene acceso a cada documento almacenado, quién debe supervisarlos y quién evaluarlos (REQ.6). Sin embargo, a pesar de almacenar la documentación y ser fácil de encontrar, ADDI no permite organizarla por fases (REQ.5). Asimismo, permite a las personas evaluadoras emitir retroalimentación sobre la documentación, pero dicha retroalimentación se envía por correo electrónico y no queda almacenada en el sistema para una posterior consulta (REQ.8).

2.3.2.2 Herramientas de gestión de documentación

Bajo este epígrafe se detallan herramientas que se pueden usar para la gestión de la documentación, permitiendo de este modo un control de todo el trabajo realizado y de sus distintas versiones.

Todas las herramientas presentadas en esta categoría permiten almacenar en la nube cualquier tipo de archivo y recuperarlo en la misma máquina en la que fue generado o en otra distinta. La organización de la documentación generada depende de la persona que la genera que debería crear una estructura de directorios

adecuada y asegurarse de almacenar cada elemento de la documentación en el lugar que le corresponda.

Además, para cada TFG, habría que configurar los permisos necesarios para que sólo pudieran acceder a la documentación las personas implicadas en el mismo.

Todas las herramientas aquí mencionadas comparten muchas características. Entre otras muchas, todas son multiplataforma, permiten controlar el acceso a la documentación a determinadas personas y recuperar versiones antiguas del contenido almacenado.

En este análisis sólo se remarcarán aquellas características particulares que pueden afectar de manera positiva o negativa a la gestión de la documentación generada durante la elaboración del TFG.

Análisis

Git⁸

Git es un software para el control de versiones que trabaja con repositorios distribuidos de manera que cada usuario tiene un repositorio local y entre todos los implicados comparten un repositorio remoto.

Su principal desventaja proviene de la complejidad de su uso para personas no expertas. Además, en el entorno de la realización del TFG, en general, no va a haber varias personas trabajando sobre la misma documentación; el estudiante creará la documentación y la persona supervisora accederá a ella, pero no se necesita llevar un control estricto de la misma.

Subversion⁹

Subversion también es un software de gestión de versiones, pero en este caso se trabaja siempre sobre un único repositorio que se encontrará alojado en un servidor remoto.

Su uso es más sencillo que el de Git debido a que trabaja con un único repositorio, pero esa es también su desventaja ya que en aquellos momentos en los que no se cuenta con acceso al repositorio hay que trabajar con copias del mismo. Dichas copias quedan fuera del control automático y hay que gestionarlas de manera manual al recuperar la conexión con el repositorio central.

⁸ <http://git-scm.com/>

⁹ <https://subversion.apache.org/>

Google Drive¹⁰

Google Drive es el servicio de almacenamiento en la nube de ficheros que ofrece Google. Ofrece una capacidad de 15GB de manera gratuita y la posibilidad de contratar distintos tipos de planes para aumentar dicha capacidad.

El principal problema que presenta Google Drive es el relativo a la privacidad de los documentos alojados en el mismo. Al usar el servicio se está aceptando la política de privacidad del mismo, que incluye otorgar permiso a Google y sus colaboradores para usar, alojar, almacenar, reproducir, modificar, crear obras derivadas, comunicar, publicar, ejecutar o mostrar públicamente y distribuir dicho contenido¹¹. Lo cual no es adecuado para los TFG, ya que en muchos casos se preferirá mantener los derechos sobre todo lo generado.

Dropbox¹²

Dropbox es uno de los servicios de almacenamiento en red más extendidos a nivel mundial. Ofrece una capacidad de 5GB de manera gratuita y la posibilidad de aumentar la capacidad hasta 1TB mediante un pago anual.

Aunque la política de privacidad de Dropbox es mucho más restrictiva que la de Google Drive, el hecho de alojar documentación en Dropbox supone otorgar permiso para que analicen dicha información¹³. En un TFG donde la confidencialidad sea importante, esto puede suponer un problema.

Comparativa

En relación a las herramientas de gestión de documentación, en la Tabla 2-17 se observa que todas satisfacen los mismos requisitos.

Las únicas diferencias remarcables son las relativas a su facilidad de uso y las políticas de privacidad y propiedad intelectual que aplican a los contenidos.

¹⁰ https://www.google.com/intl/es_es/drive/

¹¹ <https://www.google.com/policies/terms/>

¹² <https://www.dropbox.com/>

¹³ https://www.dropbox.com/es_ES/privacy

Tabla 2-17.- Comparativa herramientas de gestión de documentación

| | REQ.1 | REQ.2 | REQ.3 | REQ.4 | REQ.5 | REQ.6 | REQ.7 | REQ.8 | REQ.9 | REQ.10 |
|---------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| Git | | | | | SI | SI | SI | | | SI |
| Subversion | | | | | SI | SI | SI | | | SI |
| Google Drive | | | | | SI | SI | SI | | | SI |
| Dropbox | | | | | SI | SI | SI | | | SI |

2.3.2.3 Herramientas de comunicación

En esta categoría se agrupan todas aquellas herramientas que pueden ser usadas para permitir la comunicación entre los distintos participantes en un TFG.

Análisis

Correo electrónico

El correo electrónico es la herramienta de comunicación por excelencia en el mundo digital. Permite el envío de mensajes de un remitente a uno o varios destinatarios.

Sin embargo, para poder enviar un correo electrónico es necesario conocer la dirección del destinatario, lo que no permitiría que el alumnado que comienza la realización de su TFG pueda realizar consultas a otros estudiantes si no les conoce previamente.

Además, todas las comunicaciones son privadas entre remitente y destinatario con lo que no hay forma de poder consultar las respuestas a dudas o problemas similares planteados previamente por otras personas.

Redes sociales y servicios de mensajería

Las redes sociales como Twitter¹⁴ o Facebook¹⁵ y los servicios de mensajería como Whatsapp¹⁶ o Telegram¹⁷ tienen como principal objetivo la formación de redes de usuarios que comparten intereses. En este caso, el interés común sobre el que giraría la comunicación sería el TFG.

¹⁴ <https://twitter.com/>

¹⁵ <https://www.facebook.com/>

¹⁶ <https://www.whatsapp.com/>

¹⁷ <https://telegram.org/>

El uso de redes sociales y servicios de mensajería para la comunicación de aspectos relacionados con la realización de los TFG, permitiría al alumnado recibir, de manera inmediata, la información que la persona encargada de la supervisión desee transmitir e incluso formar un debate en ese mismo momento, por ejemplo, usando las etiquetas de Twitter.

Además, las dos redes sociales contempladas tienen la posibilidad de realizar comunicaciones directas entre usuarios, simulando e incluso a veces replicando el funcionamiento del correo electrónico tradicional.

La principal desventaja de este tipo de herramientas es la configuración de las mismas ya que habría que generar distintos grupos o subredes para poder organizar a los estudiantes por sus temas de interés y que no reciban las comunicaciones de temas o personas que no les interesan.

Además, también está el problema de la privacidad de los datos personales del alumnado. Por ejemplo, para poder comunicarse a través de Whatsapp es necesario conocer el número de teléfono de la otra persona, por lo que todo el alumnado debería estar dispuesto a dar a conocer su número de teléfono.

Foros, blogs y wikis

Los foros, blogs y wikis son herramientas de comunicación muy útiles, pero carecen de inmediatez y agilidad. Este tipo de herramientas requieren que sea el alumnado el que, de manera activa, busque la información que le interesa y solicite estar al día de las novedades a través de la suscripción a los distintos contenidos que les puedan resultar de interés.

Slack¹⁸

Slack es una herramienta pensada para la comunicación interna de grupos de trabajo agrupando características de correo electrónico, foros, chats y redes sociales entre otras muchas.

En Slack se pueden formar grupos públicos o privados a los que los usuarios pueden apuntarse o ser invitados (en el caso de los grupos privados). Además cuenta con clientes para las principales plataformas móviles lo que posibilita que los usuarios reciban la información de sus grupos en cualquier lugar y puedan participar en los debates de manera sencilla.

Slack también destaca por su potente herramienta de búsqueda que permite buscar palabras o elementos concretos tanto en los mensajes como en el interior de los archivos que se hayan intercambiado en sus canales.

¹⁸ <https://slack.com/>

Comparativa

Entre las herramientas de comunicación examinadas (ver Tabla 2-18), Slack es la más completa ya que permite: proporcionar retroalimentación asociada a elementos concretos y almacenarla en el sistema (REQ.8); generar foros de comunicación en base a intereses comunes (REQ.9) y administrar fácilmente qué personas deben poder comunicarse por estar involucradas en la realización, supervisión y evaluación del TFG (REQ.10).

Tabla 2-18.- Comparativa herramientas de comunicación

| | REQ.1 | REQ.2 | REQ.3 | REQ.4 | REQ.5 | REQ.6 | REQ.7 | REQ.8 | REQ.9 | REQ.10 |
|---|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| Correo electrónico | | | | | | | | SI | | |
| Redes sociales y servicios de mensajería | | | | | | | | SI | SI | ★ |
| Foros, blogs y wikis | | | | | | | | SI | SI | ★ |
| Slack | | | | | | | | SI | SI | SI |

Las redes sociales, los servicios de mensajería, los foros, los blogs y las wikis satisfacen los requisitos relacionados con la retroalimentación (REQ.8) y la comunicación (REQ.9). Sin embargo, presentan problemas para configurar qué personas deben poder comunicarse entre sí (REQ.10) y para garantizar la privacidad de los datos de dichas personas.

2.3.2.4 Herramientas de gestión de proyectos

La mayoría de herramientas de esta categoría están principalmente pensadas para la gestión de proyectos a nivel empresarial, por lo que todas ellas ofrecen una serie de funcionalidades típicas de la gestión de proyectos: permiten la definición de distintos proyectos, la definición de las tareas y subtareas del proyecto, de las personas implicadas en el mismo, etc. En este apartado se mencionan aquellos aspectos que diferencian a cada una de las herramientas revisadas.

Análisis

Microsoft Project Server¹⁹

Microsoft Project Server es la aplicación para la gestión de proyectos empresariales ofrecida por Microsoft.

Permite su integración con otras herramientas de Microsoft como la herramienta de colaboración y trabajo en equipo SharePoint y la herramienta de comunicación Lync para aumentar sus funcionalidades.

Microsoft Project Server es una herramienta de pago enfocada a la gestión de proyectos empresariales, por lo que carece de aspectos relacionados con la evaluación.

JIRA Software²⁰

JIRA Software es una herramienta de gestión de proyectos creada y distribuida por la empresa Atlassian.

JIRA Software está especialmente orientada a la gestión de proyectos que utilicen las metodologías ágiles en su gestión y desarrollo.

Las funcionalidades proporcionadas en su versión más básica pueden ser ampliadas a través de la instalación de distintos complementos que permiten por ejemplo, utilizar software de gestión de versiones como Git.

JIRA Software es una herramienta de pago y al estar dirigida a la gestión de proyectos en el mundo empresarial, no incorpora ninguna funcionalidad relacionada con la evaluación de dichos proyectos.

project-open²¹

project-open[, también denominado jpo[, es una aplicación web para la gestión de proyectos que a través de distintos módulos incorpora también funcionalidades financieras y de ERP (*Enterprise Resource Planning*) lo cual la convierte en una herramienta de gestión muy potente, pero enfocada principalmente a la gestión de proyectos empresariales.

Al no estar enfocado al mundo educativo, no incorpora ningún módulo que permita trabajar con rúbricas de evaluación o evaluar los proyectos mediante otros criterios. Así mismo, tampoco incorpora ningún sistema que permita ofrecer retroalimentación sobre los elementos del proyecto generados.

¹⁹ <https://products.office.com/es-es/project/enterprise-project-server>

²⁰ <https://es.atlassian.com/software/jira>

²¹ <http://www.project-open.com/>

]po[permite asociar distintos tipos de documentos a cada proyecto, pero no hacerlo a cada una de las tareas del proyecto, por lo que la organización de la documentación no es todo lo precisa que sería deseable. Sin embargo, permite la integración de distintos sistemas de gestión de versiones como Subversion o CVS para gestionar las distintas versiones de la documentación generadas.

OpenProject²²

OpenProject es una herramienta web de gestión de proyectos que también está muy orientada al mundo empresarial por lo que entre sus características no se encuentran aquellas más relacionadas con los proyectos académicos como pueden ser la retroalimentación o la evaluación.

Sí que dispone de diversos canales de comunicación como foros o la posibilidad de gestionar la convocatoria de reuniones desde la propia aplicación.

Para la gestión de la documentación, OpenProject proporciona un complemento que permite asociar documentos al proyecto, y gestionar sus versiones mediante la integración con sistemas de control de versiones como Git y Subversion. También permite la inclusión de documentos asociados a las tareas concretas del proyecto.

Redmine²³

Al igual que los dos sistemas anteriores, Redmine es una aplicación web para la gestión de proyectos. Redmine también incorpora herramientas de comunicación entre distintos usuarios del sistema a través de sus foros y la posibilidad de controlar las versiones de la documentación generada a través de software de terceros como Git, Subversion, Cvs, Mercurial y otros.

Al igual que OpenProject, permite asociar documentos a tareas concretas, lo que permite una gestión de la documentación más detallada tanto para la persona supervisora como para el alumnado.

Redmine tampoco incorpora ningún módulo que permita gestionar la retroalimentación proporcionada por las personas supervisoras, ni la evaluación de los proyectos.

Technical Project²⁴

Technical Project es un complemento para Moodle que permite la definición de un proyecto con sus tareas y su planificación temporal en el contexto de un curso definido en la plataforma.

²² <https://www.openproject.org/>

²³ <http://www.redmine.org/>

²⁴ https://moodle.org/plugins/view.php?plugin=mod_techproject

Technical Project depende de la infraestructura y de la gestión de cursos y usuarios que permite realizar el sistema Moodle, lo que dificulta su uso en un entorno donde cada estudiante tiene su propia planificación, que a veces estará compartida con otros estudiantes en el caso de un proyecto compartido. Además, el profesorado debería poder acceder sólo a las planificaciones de aquellos estudiantes a los que está supervisando.

Comparativa

Tal y como se aprecia en la Tabla 2-19, todas las herramientas de gestión de proyectos son muy similares. Pero destaca Redmine por ser gratuita, de código abierto y ofrecer compatibilidad con Git, Subversion y la posibilidad de realizar la identificación de usuarios mediante el protocolo LDAP. Este protocolo es el usado en la UPV/EHU tanto por alumnado como por profesorado por lo que no haría falta la creación de nuevos usuarios.

Tabla 2-19.- Comparativa herramientas de gestión de proyectos

| | REQ.1 | REQ.2 | REQ.3 | REQ.4 | REQ.5 | REQ.6 | REQ.7 | REQ.8 | REQ.9 | REQ.10 |
|---------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| Microsoft Project Server | | | | SI | SI | SI | SI | SI | SI | SI |
| JIRA Software | | | | SI | SI | SI | SI | SI | SI | SI |
|]po[| | | | SI | SI | SI | SI | SI | SI | SI |
| OpenProject | | | | SI | SI | SI | SI | SI | SI | SI |
| Redmine | | | | SI | SI | SI | SI | SI | SI | SI |
| Technical Project | | | | ★ | SI | SI | SI | SI | ★ | SI |

En el caso de Technical Project, al ser un complemento de la plataforma Moodle, el poder tener una planificación temporal compartida (REQ.4) y ofrecer la posibilidad de realizar consultas en base a intereses comunes (REQ.9), vienen dadas por el propio Moodle. Por lo tanto, también presenta el problema relativo a la creación de un curso por cada TFG, lo que en primer lugar depende de la política de creación de cursos del centro o de la universidad y en segundo lugar afecta a la posibilidad de comunicación entre iguales.

2.3.2.5 *Herramientas de trabajo con rúbricas*

En este apartado se analizarán herramientas existentes actualmente y que se pueden usar para la creación, gestión y uso de rúbricas.

Se debe tener en cuenta que el uso de las rúbricas va a consistir en la evaluación del TFG y que ésta va a realizarse usando varias rúbricas que deberán ser completadas por los distintos componentes del tribunal de evaluación.

Análisis

Complemento de rúbricas de Moodle²⁵

Es el complemento oficial de la plataforma Moodle que posibilita la evaluación a través de rúbricas.

Permite definir las rúbricas personalizándolas para cada elemento concreto. La gestión de las distintas rúbricas de una asignatura y cómo se combinan para obtener la nota final, se realiza a través del libro de calificaciones de Moodle. Sólo el profesorado perteneciente al curso concreto donde se ha definido la rúbrica pueden usarla para realizar la evaluación.

Evalcomix²⁶

Evalcomix es un servicio web para la evaluación en contextos formativos que permite su integración en plataformas como Moodle.

Si se usa de manera independiente de Moodle, no permite definir que la nota final depende de distintas rúbricas, cada una de ellas con un peso porcentual, ni permite diferenciar quién tiene que rellenar cada rúbrica.

Si se usa integrado en Moodle, esta carencia se puede cubrir definiendo distintas tareas evaluables, cada una de ellas con su propia rúbrica. Además, gracias al módulo calificador que incorpora Moodle, se puede realizar la fórmula matemática que calcule una única nota final en base a la nota obtenida en cada una de las tareas.

²⁵ <https://docs.moodle.org/22/en/Rubrics>

²⁶ <http://evalcomix.uca.es/>

Calc²⁷²⁸

Calc es el software de cálculo de LibreOffice / OpenOffice.

Calc sí permite realizar el cálculo de la nota final teniendo en cuenta distintas rúbricas y el peso porcentual de cada una de ellas, pudiendo automatizar dichos cálculos.

Sin embargo, no proporciona ningún tipo de gestión de usuarios para indicar quién tiene la responsabilidad y los permisos de rellenar cada rúbrica. A pesar de no permitir una gestión de usuarios propiamente dicha, sí que se puede importar de manera sencilla en un único fichero las evaluaciones realizadas por cada usuario y automatizar el cálculo final.

eRubrica²⁹

Es una herramienta web para la evaluación por competencias que permite la definición de rúbricas. Se puede definir quién es el encargado de la evaluación, pero no distingue los roles de estudiante y docente, ni realiza ninguna gestión a nivel de asignatura. A la hora de definir las rúbricas, se puede elegir entre distintas opciones de niveles de valoración, pero no se puede personalizar.

Rubistar³⁰

Es un portal web que permite únicamente la definición de rúbricas de evaluación, no permite la gestión de su uso en la evaluación. Para definir las rúbricas proporciona una serie de plantillas aportadas por la comunidad de usuarios y que pueden ser adaptadas a las necesidades y deseos del profesorado, pero siempre manteniendo la estructura que viene definida por defecto en la que hay cuatro niveles de valoración. Las rúbricas creadas pueden imprimirse o exportarse en formato PDF.

iRubric³¹

iRubric es una herramienta, integrada en el LMS RCampus, que permite la definición y el uso de rúbricas de evaluación. IRubric permite la reutilización de rúbricas definidas previamente y que hayan sido hechas públicas por sus autores. También permite indicar qué docentes o estudiantes deben usar cada rúbrica. La posibilidad de definir que la nota de una asignatura está formada por distintas

²⁷ <https://es.libreoffice.org/caracteristicas/calc/>

²⁸ <http://www.openoffice.org/es/producto/calc.html>

²⁹ <http://erubrica.uma.es/>

³⁰ <http://rubistar.4teachers.org/index.php>

³¹ <http://www.rcampus.com/indexrubric.cfm>

rúbricas, viene dada gracias al sistema de calificaciones de la plataforma RCampus, no por el sistema IRubric.

teAchnology General Rubric Generator³²

Es un portal web que permite la creación de rúbricas en base a una serie de plantillas. Sólo permite la definición e impresión de las rúbricas, no permite su uso desde la propia plataforma, ni su integración en otras herramientas.

Comparativa

La Tabla 2-20 recoge la comparativa de las herramientas de trabajo con rúbricas que se han analizado.

Tabla 2-20.- Comparativa herramientas de trabajo con rúbricas

| | REQ.1 | REQ.2 | REQ.3 | REQ.4 | REQ.5 | REQ.6 | REQ.7 | REQ.8 | REQ.9 | REQ.10 |
|--|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|
| Complemento de rúbricas de Moodle | ★ | SI | SI | | | | | | | ★ |
| Evalcomix | ★ | SI | SI | | | | | | | ★ |
| Calc | SI | SI | SI | | | | | | | |
| eRubrica | | SI | SI | | | | | | | ★ |
| Rubistar | ★ | SI | | | | | | | | |
| iRubric | ★ | SI | SI | | | | | | | SI |
| teAchnology GRG | ★ | SI | | | | | | | | |

De las herramientas analizadas, la única que por sí misma permite definir la evaluación como la combinación de distintas rúbricas, asignando a cada una un peso distinto en la calificación final (REQ.1), es la herramienta de cálculo Calc de los paquetes LibreOffice y OpenOffice. Sin embargo, esta herramienta no facilita la administración de la evaluación (REQ.10) al no permitir definir de manera sencilla las personas que deben completar cada rúbrica y obligar a crear una hoja de cálculo distinta para cada TFG.

El resto de herramientas analizadas, sólo permiten definir la evaluación como una combinación de evaluaciones si son integradas en una plataforma de aprendizaje, caso del complemento de rúbricas de Moodle, Evalcomix e iRubric. O

³² http://www.teach-nology.com/web_tools/rubrics/general/

se limitan a permitir la definición de distintas rúbricas, pero para usarlas hay que imprimirlas y hacer los cálculos para obtener la calificación de manera manual (Rubistar y teAchnology GRG).

2.3.2.6 Resumen

En el análisis de las herramientas realizado, se ha comprobado que no existe ninguna que satisfaga completamente todos los requisitos funcionales establecidos.

En general, las herramientas analizadas en cada una de las categorías son bastante similares y satisfacen más o menos los mismos requisitos. Así por ejemplo, todas las herramientas de gestión de la documentación analizadas ofrecen las funcionalidades necesarias para satisfacer los requisitos REQ.5, REQ.6, REQ.7 y REQ.10, pero sólo esos.

Atendiendo al número de requisitos funcionales completamente satisfechos, las herramientas de gestión de proyectos obtienen los mejores resultados. La mayoría de las herramientas de esta categoría satisfacen completamente todos los requisitos funcionales a excepción de los relacionados con la evaluación (REQ.1, REQ.2 y REQ.3). Esto se debe a que estas herramientas están orientadas a la gestión de proyectos empresariales, en los que no se precisa la evaluación, ni formativa, ni sumativa del trabajo realizado. Entre las herramientas de gestión de proyectos analizadas, destaca Redmine por sus posibilidades de ampliación mediante la instalación de complementos.

Por otro lado, si analizamos los requisitos que se cumplen, aunque sea de manera parcial, la herramienta que mayor número de requisitos satisface es Moodle.

2.4 Propuesta de metodología y plataforma informática de soporte

En este apartado se presenta la metodología FPM (*Final-year Project Mentoring*) definida para dar respuesta a los objetivos específicos de este trabajo de tesis. Para ello se detallan los hitos de evaluación establecidos en la metodología y los instrumentos de evaluación propuestos (Villamañe et al., 2014). Además, se presenta una plataforma para dar soporte a la metodología y que en combinación con ella, permite satisfacer los objetivos específicos establecidos en este trabajo de tesis (ver sección 1.3).

2.4.1 Hitos de evaluación

La metodología FPM adapta en dos aspectos la guía definida en (Valderrama et al., 2009). Por un lado, la ajusta a las necesidades específicas de los grados en informática. Por otro lado, da una mayor relevancia al proceso de evaluación continua de los trabajos (OE-4), usando criterios públicos y objetivos (OE-5) y promoviendo los comentarios de retroalimentación por parte del supervisor en los momentos clave durante la realización del TFG (OE-2).

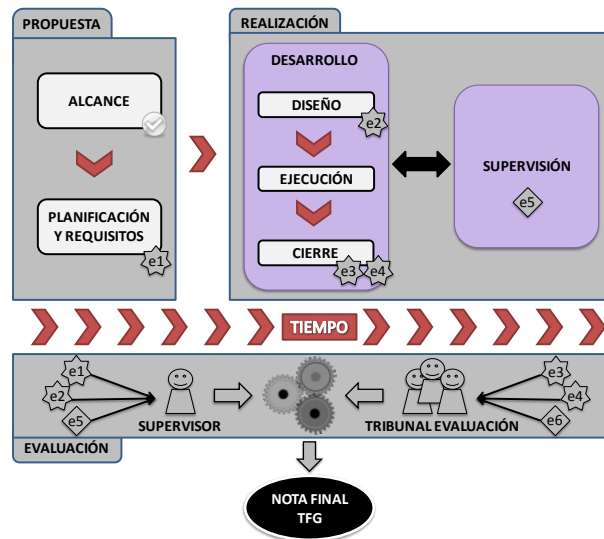


Fig. 2-4.- Metodología FPM

La metodología FPM establece tres hitos (Propuesta, Realización y Evaluación) y define 6 elementos a tener en cuenta para evaluar el TFG. Cuatro de estos elementos evaluables están basados en elementos generados por parte del alumnado y los otros dos son responsabilidad del profesorado encargado de la supervisión y de la evaluación.

En el hito de Propuesta, la idea de TFG a realizar (también llamada alcance) debe recibir el visto bueno del comité definido para tal fin y formado por parte del profesorado. De este modo, se asegura que el trabajo que se va a realizar cumple unos requisitos u objetivos mínimos (OE-1). Una vez obtenido el visto bueno a la propuesta, se realizan la planificación y captura de requisitos del TFG y se genera el primero de los elementos evaluables (*Eva_Requisitos*, Fig. 2-4→e1).

En el hito de Realización, el alumnado trabaja sobre el TFG, realizando en primer lugar un diseño del trabajo a llevar a cabo y generando el segundo de los elementos

evaluables (*Eva_Diseño*, Fig. 2-4→e2). A continuación, se realiza el TFG propiamente dicho y, una vez esté terminado, se procede a su cierre finalizando los dos últimos elementos evaluables definidos por la metodología: la documentación escrita que se ha ido elaborando desde el comienzo del TFG (*Eva_Memoria*, Fig. 2-4→e3) y la presentación y defensa oral que se realiza ante el tribunal de evaluación (*Eva_Presentación*, (Fig. 2-4→e4).

En el hito de Realización, la labor de la persona encargada de la supervisión del TFG es de vital importancia para asegurar su buen desarrollo y su cierre de manera correcta. Esta labor finaliza con la elaboración de un informe sobre la actitud del estudiante, su constancia, interés de mejora y demás cuestiones relacionadas con su comportamiento y profesionalidad. Dicho informe (*Eva_Actitud*, Fig. 2-4→e5), es otro de los elementos que forman parte de la evaluación del TFG.

Finalmente, el hito de Evaluación define el protocolo de evaluación que considera los distintos elementos evaluables generados desde el comienzo del TFG hasta el día de su defensa ante el tribunal de evaluación. Los elementos *Eva_Requisitos*, *Eva_Diseño* y *Eva_Actitud* son evaluados por la persona encargada de la supervisión del TFG en el momento en el que son generados. Por su parte, los elementos *Eva_Memoria* y *Eva_Presentación* son evaluados por las personas integrantes del tribunal. Asimismo, se ha definido un último elemento de evaluación, que se corresponde con la complejidad técnica del TFG (*Eva_Complejidad*, Fig. 2-4→e6), y que también es evaluado por las personas que forman el tribunal.

Dado que la normativa de evaluación del TFG es diferente en cada universidad e incluso en cada centro, el calendario de evaluación y el proceso de cálculo de la nota final del TFG definidos en la metodología FPM son flexibles.

Por ejemplo, la normativa de la UPV/EHU sobre elaboración y defensa del Trabajo Fin de Grado («Disposición General 2012/2805», 2012) establece que sea el tribunal de evaluación el encargado de determinar la calificación final del TFG. Por lo tanto, para el cálculo de dicha nota la persona encargada de la supervisión debe comunicar al tribunal el resultado de las evaluaciones que ha realizado y éste, teniéndolas en cuenta, establecerá la calificación final.

2.4.2 Instrumentos de evaluación

La metodología FPM define un proceso de evaluación continua que aplica un conjunto de criterios públicos y objetivos pre-establecidos para cada elemento evaluable generado en el proceso. Además, se promueve la retroalimentación por parte del supervisor en los momentos clave de la realización del TFG.

Para dar soporte al método de evaluación definido en la metodología FPM, se ha optado por el uso de rúbricas de evaluación. Sin embargo, dado que las rúbricas de evaluación de tipo holístico no son las más indicadas para lograr una buena evaluación formativa y una retroalimentación de calidad (ver sección 2.2.2), en este trabajo se ha decidido prescindir de este tipo de rúbricas y centrarse en las rúbricas pertenecientes a la aproximación analítica: las rúbricas analíticas y las rúbricas atomísticas.

En lo referente a la objetividad, como señala (Moskal, 2000), la probabilidad de que dos personas de manera independiente asignen la misma calificación a una respuesta dada, se incrementa cuando se dispone de una descripción de las características que debe cumplir la respuesta para recibir una determinada puntuación. En ese aspecto, las rúbricas permiten clarificar los niveles de calidad del trabajo realizado aumentando la objetividad en la evaluación (Rosenberry, 2010).

Esta clarificación de los niveles de calidad convierte a las rúbricas en una potente herramienta de autoevaluación y retroalimentación (Jonsson & Svingby, 2007). Además, las rúbricas aportan al alumnado argumentaciones sobre su calificación, haciendo el protocolo de evaluación más transparente e inteligible (Valderrama et al., 2009), lo que aumenta la auto-eficacia y fomenta la auto-regulación entre los estudiantes (Huba & Freed, 2000; Jackson & Larkin, 2002). Por lo tanto, el uso de rúbricas de evaluación facilita la comunicación entre los diferentes participantes en el proceso de realización (Wehlburg, 2006), ayudando a aumentar y mejorar la retroalimentación que recibe el alumnado (OE-2) y sirviendo de apoyo para una mejor comunicación entre las personas supervisoras y el alumnado (OE-3).

Existen trabajos previos en el desarrollo de rúbricas para la evaluación de TFG (Sánchez et al., 2015; Valderrama et al., 2009). Sin embargo, a pesar de que estos trabajos establecen rúbricas para la evaluación de los TFG, se centran principalmente en la evaluación de competencias transversales (J. L. Sanchez et al., 2011) o no abordan los aspectos técnicos (F. Sanchez et al., 2014). Esto puede ocasionar que un porcentaje importante de la nota siga siendo poco clara para el alumnado (F. Sanchez et al., 2014).

Por ello, en este trabajo se propone definir un conjunto de rúbricas adecuado a la metodología FPM. Esta definición partirá de los trabajos previos analizados, con el objetivo de aportarles un mayor nivel de especificidad e incluir aspectos técnicos.

2.4.3 Plataforma informática de soporte

Finalmente, para completar este trabajo se plantea la necesidad de proporcionar una plataforma informática que facilite la aplicación de la misma (OE-6).

En la sección 2.3.1 se han identificado los requisitos funcionales que debería satisfacer tal plataforma y se ha concluido que no existe ninguna que satisfaga completamente todos los requisitos funcionales establecidos. Además, hay que tener en cuenta que se debe evitar que el uso de herramientas informáticas suponga una mayor carga de trabajo para los usuarios, para lo que es necesario garantizar su usabilidad. Por este motivo, no se plantea como opción el uso de dos o más herramientas distintas que no estén integradas en una única plataforma permitiendo trabajar con un único usuario y disponiendo de toda la información centralizada en el mismo sitio.

Dado que ninguna de las herramientas analizadas cumple todos los requisitos establecidos, se ha optado por la opción de seleccionar una de las herramientas analizadas y desarrollar las funcionalidades que le permitan cubrir los requisitos funcionales que no satisface de manera deseada. Desde este punto de vista, y en base a las características explicadas anteriormente, las herramientas más adecuadas para tomar como punto de partida son Redmine y Moodle.

En el caso de Redmine, y añadiéndole el complemento que ofrece para interactuar con Slack, sólo es necesario desarrollar un complemento que permita la definición y gestión de la evaluación de un TFG en base a rúbricas. Dicho complemento también debe permitir asignar a las personas encargadas de realizar la evaluación en cada TFG. Los complementos en Redmine son en gran parte autocontenidos y sólo es necesario acceder a las definiciones de usuarios que ofrece el sistema para poder indicar qué personas están involucradas en la evaluación de cada TFG

Para Moodle, es necesario desarrollar un complemento que permita la gestión de usuarios de modo que se puedan definir y asignar distintos roles tanto a las planificaciones como a las rúbricas y evaluaciones dentro de un mismo curso. Además, el modificar la gestión de usuarios y roles en la plataforma implica modificaciones en todos los complementos que se utilicen para poder asignar los permisos correspondientes a los distintos usuarios.

Teniendo en cuenta las modificaciones que implica cada una de ellas, en este trabajo de tesis se ha optado por utilizar Redmine como base de TraMeA (*Tracking, Mentoring and Assessment*), la plataforma a desarrollar para dar soporte a la metodología FPM, cuyo diseño y desarrollo se detalla en el siguiente capítulo.

3

TraMeA

En este capítulo se detalla el análisis y diseño de TraMeA (*Tracking, Mentoring and Assessment*), la plataforma que permite dar soporte a la metodología FPM. Para ello, el capítulo comienza con una descripción de InterMod, la metodología de desarrollo utilizada en el desarrollo de TraMeA. A continuación, se relatan los procesos de análisis y diseño de TraMeA y se continúa con la definición del sistema AdESMuS surgido a partir de las experiencias adquiridas durante el diseño y desarrollo de TraMeA. El capítulo finaliza con un apartado de conclusiones.

3.1 InterMod

InterMod (Losada, 2014; Losada et al., 2013) es una metodología ágil para el desarrollo de software interactivo de calidad que integra aspectos de metodologías ágiles, desarrollo dirigido por modelos y diseño centrado en el usuario. InterMod ha sido usada con éxito para el desarrollo de sistemas en entornos educativos e integrada con técnicas de usabilidad (Losada et al., 2012), lo que la convierte en una herramienta adecuada para el desarrollo de TraMeA.

InterMod organiza el desarrollo de un proyecto como una serie de iteraciones en las que partiendo de los denominados objetivos de usuario, estos se van refinando a través de la ejecución de una serie de actividades hasta producir el software final.

A continuación se describen los objetivos de usuario y el proceso de desarrollo que establece InterMod.

3.1.1 Objetivos de usuario

InterMod define como Objetivo de Usuario (*User Objective* o UO) cada necesidad de una persona que vaya a utilizar el sistema y que puede ser satisfecha mediante una o más funcionalidades. Un UO puede corresponderse con uno o varios requisitos funcionales y con uno o varios requisitos no funcionales.

Los UO se formalizan describiendo el comportamiento de la persona (acciones a realizar en el sistema) y el comportamiento del sistema (su reacción) en términos de interfaz gráfica. Al limitarse a describir el comportamiento de la interfaz gráfica, estas descripciones incluyen sólo una parte de los requisitos funcionales. A lo largo del proceso definido por InterMod se profundiza en los UO para extraer todos los requisitos funcionales y poder satisfacerlos.

InterMod define dos tipos de UO: directos e indirectos. Los UO directos son los objetivos que surgen directamente a partir de peticiones realizadas por las personas que van a utilizar el sistema. Los UO indirectos surgen por necesidades internas durante el proceso de desarrollo, no son peticiones directas. Los UO indirectos también aparecen tras la división o fusión de otros UO, tanto directos como indirectos.

3.1.2 Proceso de desarrollo de InterMod

Como puede observarse en Fig. 3-1, InterMod organiza el desarrollo de un proyecto en una etapa inicial de análisis del proyecto seguido de la ejecución iterativa de una serie de pasos.

En la etapa inicial, se analiza el sistema a desarrollar y a través de entrevistas con las personas usuarias se establecen los UO directos (o iniciales) y se toman las principales decisiones de diseño. Las decisiones de diseño como el tipo de dispositivos en los que tiene que funcionar la aplicación, las preferencias y necesidades sobre la gama de colores o los aspectos de accesibilidad que debe cumplir, se recogen en el *Modelo del sistema*.

Tras la etapa inicial de análisis, InterMod define tres pasos que se ejecutan de manera secuencial y cuya ejecución se repite de manera iterativa hasta completar el desarrollo de la aplicación.

En cada iteración, se comprueba cuáles de los UO definidos ya han sido finalizados y se actualiza la lista incluyendo los nuevos UO directos que han podido surgir a partir de nuevas peticiones de las personas que van a utilizar el sistema o los UO indirectos creados por fusión o división de UO ya existentes.

El uso de encuestas online en lugar de entrevistas personales, permite ampliar fácilmente el número de personas consultadas, recoger la misma información y además aporta comodidad porque la encuesta se puede completar en cualquier momento, sin necesidad de ceñirse a una cita concreta.

A partir de las necesidades expresadas en las encuestas, se confeccionó la lista de UO iniciales y se identificaron las distintas formas de trabajar con TraMeA que dieron lugar a diferentes roles de usuario.

En los siguientes apartados se definen los roles de usuario establecidos, se detalla el modelo del sistema y se especifican los UO iniciales identificados.

3.2.1 Roles de usuario

Los usuarios finales de TraMeA son el alumnado y el profesorado universitario. Dependiendo del funcionamiento establecido en cada centro o universidad, el personal de administración y servicios (PAS) puede ser el encargado de gestionar los aspectos relacionados con la administración del sistema. Sin embargo, en muchos casos dichas tareas deben ser realizadas por el propio profesorado, por lo que se ha decidido incluir al PAS dentro del profesorado y tener en cuenta que el profesorado puede ejercer distintos roles en función de sus responsabilidades en cada TFG y en la administración del sistema.

En el marco de los TFG se pueden identificar 6 roles diferentes:

- Persona usuaria: cualquiera que utiliza TraMeA.
- Persona administradora: se encarga de las tareas de administración del sistema.
- Estudiante: persona que desarrolla el TFG. Además, en ocasiones, el/la estudiante también puede proponer el alcance del TFG.
- Docente: se encarga de dar el visto bueno al alcance propuesto para un TFG. También puede definir propuestas de TFG.
- Persona supervisora: persona que supervisa el proceso de realización de un TFG y se encarga de evaluar parte del mismo.
- Componente tribunal: persona que usa TraMeA para evaluar el TFG desarrollado por un estudiante.

Tanto el alumnado como el profesorado ejercen el rol Persona usuaria. El resto de roles, a excepción del rol Estudiante, son ejercidos exclusivamente por el profesorado.

3.2.2 Modelo del sistema

TraMeA se ha desarrollado tomando como base el sistema Redmine y sus complementos (ver sección 2.4.3), por lo que su aspecto gráfico ha estado predeterminado por las características técnicas y los elementos visuales presentes en Redmine. En la Fig. 3-2 puede observarse el aspecto gráfico, por defecto, que presenta el sistema Redmine, utilizado como punto de partida para el desarrollo de TraMeA.

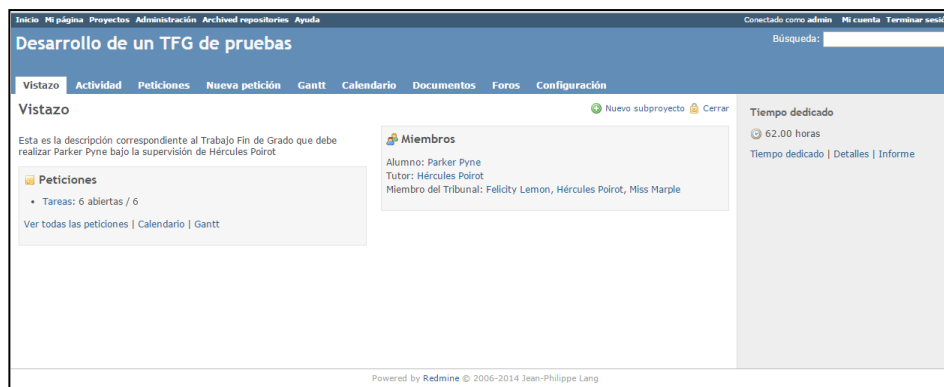


Fig. 3-2.- Interfaz gráfica de Redmine

Se ha considerado que todos los usuarios (independientemente de su rol) presentarían las mismas necesidades de accesibilidad, por lo que el sistema se ha desarrollado teniendo en cuenta las recomendaciones generales sobre accesibilidad y usabilidad de aplicaciones (Chisholm et al., 2001; Nielsen, 2000).

3.2.3 Objetivos de Usuario iniciales

Partiendo de los resultados de las encuestas (ver ANEXO B y ANEXO C) se han definido una serie de UO para TraMeA que se presentan a continuación agrupados por el rol con el que se relacionan. Para una mayor claridad se ha enriquecido la nomenclatura establecida por InterMod para los UO con un código que representa el rol de usuario al que están asociados: *U* - Persona usuaria; *A* - Persona administradora; *E* - Estudiante; *T* - Persona supervisora; *D* - Docente y *Tr* - Componente tribunal.

Rol Persona usuaria

- UO01-U: Realizar consultas en base a intereses. Los usuarios de TraMeA deben poder comunicarse entre ellos para tratar intereses particulares o comunes.

Rol Persona administradora

- UO01-A: Gestionar criterios alcance. La persona encargada de la administración de TraMeA debe poder gestionar (crear, modificar y eliminar) los criterios para valorar una propuesta de alcance.
- UO02-A: Gestionar criterios evaluación. La persona encargada de administrar el sistema debe poder gestionar (definir, modificar y eliminar) los criterios de evaluación de los Trabajos Fin de Grado.

Rol Estudiante

- UO01-E: Conocer criterios alcance. El alumnado debe poder conocer los criterios que determinan si el alcance de una propuesta es adecuado para que sea considerado válido para la realización de un TFG.
- UO02-E: Realizar propuestas alcance. Los y las estudiantes tienen que poder realizar propuestas de TFG indicando qué componentes del profesorado desean que se las valoren (si la elección es posible) y pudiendo consultar posteriormente la evaluación que éstos realicen. En caso de que la evaluación sea negativa tienen que poder enviar una nueva versión de la misma propuesta u otra completamente diferente.
- UO03-E: Solicitar retroalimentación. El alumnado tiene que poder solicitar y recibir retroalimentación sobre un elemento concreto. Además debe poder gestionar el flujo de retroalimentación que se genere sobre dicho elemento.
- UO04-E: Actualizar estado TFG. Los y las estudiantes deben tener la opción de actualizar el estado del TFG.
- UO05-E: Conocer criterios evaluación. Los criterios de evaluación de los TFG deben ser públicos y el alumnado tiene que poder consultarlos cuando desee.
- UO06-E: Solicitar evaluación elemento. El alumnado debe poder solicitar la evaluación de aquellos elementos evaluables que haya finalizado.
- UO07-E: Consultar evaluación. Los y las estudiantes deben poder consultar las evaluaciones que hayan recibido.

Rol Docente

- UO01-D: Usar criterios alcance. El profesorado debe poder consultar las propuestas de alcance recibidas y utilizar los criterios preestablecidos para valorarla. Esta valoración junto con los comentarios que los revisores consideren oportuno generar deben almacenarse para que puedan ser consultados posteriormente.

Rol Persona supervisora

- UO01-T: Conocer estado TFG. Permitir que la persona encargada de la supervisión de un TFG pueda consultar el estado del mismo.
- UO02-T: Generar retroalimentación. La persona encargada de la supervisión tiene que recibir las solicitudes de retroalimentación, ser capaz de responderlas y poder consultar en cualquier momento la retroalimentación proporcionada.
- UO03-T: Evaluar elemento: Las personas encargadas de la supervisión deben poder usar los criterios de evaluación definidos para valorar los distintos elementos evaluables del TFG cuando lo solicite el o la estudiante que lo está realizando.

Rol Componente tribunal

- UO01-Tr: Evaluar TFG: Los componentes del tribunal deben poder evaluar los Trabajos Fin de Grado finalizados que les asignen, usando para ello los criterios establecidos. También deben poder consultar las evaluaciones que hubieran realizado previamente.

La Tabla 3-1 muestra la correspondencia entre los UO extraídos de las encuestas y los requisitos funcionales establecidos en la sección 2.3.1: REQ.1 (Definición y gestión de rúbricas); REQ.2 (Consulta de criterios); REQ.3 (Uso de rúbricas); REQ.4 (Planificación temporal compartida); REQ.5 (Gestión de la documentación); REQ.6 (Acceso a la documentación); REQ.7 (Gestión de versiones); REQ.8 (Retroalimentación); REQ.9 (Realizar consultas en base a intereses comunes) y REQ.10 (Administración).

Tabla 3-1.- Correspondencia de los UO extraídos de las encuestas con los requisitos

| Código UO | Descripción | Requisito relacionado |
|-----------|--|-----------------------|
| U001-U | Realizar consultas en base a intereses | REQ.9 |
| U001-A | Gestionar criterios alcance | REQ.1 |
| U002-A | Gestionar criterios evaluación | REQ.1 |
| U001-E | Conocer criterios alcance | REQ.2 |
| U002-E | Realizar propuestas alcance | REQ.8 |
| U003-E | Solicitar retroalimentación | REQ.8 |
| U004-E | Actualizar estado TFG | REQ.4 |
| U005-E | Conocer criterios evaluación | REQ.2 |
| U006-E | Solicitar evaluación elemento | REQ.8 |
| U007-E | Consultar evaluación | REQ.3 |
| U001-D | Usar criterios alcance | REQ.3 |
| U001-T | Conocer estado TFG | REQ.4 |
| U002-T | Generar retroalimentación | REQ.8 |
| U003-T | Evaluar elemento | REQ.3 |
| U001-Tr | Evaluar TFG | REQ.3 |

Además de los UO extraídos a través de las encuestas, existen otros UO que se corresponden con los aspectos tradicionales de administración de un sistema y de gestión durante la elaboración de un TFG y que también son necesarios para satisfacer los requisitos establecidos. Estos UO también se han recogido, asociándolos al rol de usuario que los desempeña en el sistema.

Rol Persona usuaria

- U002-U: Acceder al sistema. Los usuarios deben poder acceder a las funcionalidades del sistema que les correspondan y sólo a esas.
- U003-U: Gestionar datos personales. Los usuarios deben poder gestionar los datos que TraMeA almacena sobre ellos.

Rol Persona administradora

- U003-A: Iniciar y finalizar TFG. La persona encargada de la administración del sistema debe poder asignar un Trabajo Fin de Grado a un estudiante o a un grupo de estudiantes e indicar la

persona o personas encargadas de su supervisión. También debe poder modificar o eliminar dicha información. Asimismo también debe poder indicar la finalización del TFG y asignar los componentes del tribunal encargado de su evaluación.

Rol Estudiante

- UO08-E: Gestionar documentación TFG. El alumnado debe poder gestionar toda la documentación y los elementos (código, imágenes, videos, presentaciones, etc.) que genere a lo largo de la realización de su TFG.
- UO09-E: Solicitar evaluación TFG. Cuando el alumnado considere que ha finalizado la realización y la documentación de su TFG debe poder solicitar la defensa del mismo.

Rol Persona supervisora

- UO04-T: Autorizar evaluación TFG. Cuando el alumando solicite la defensa de su TFG, las personas encargadas de su supervisión deben poder dar el visto bueno a dicha defensa.

Rol Componente tribunal

- UO02-Tr: Consultar documentación TFG. Una vez autorizada la defensa de un TFG, los componentes del tribunal encargados de su evaluación deben poder acceder a la documentación del mismo.

La Tabla 3-2 recoge la relación entre los UO surgidos de los aspectos de administración y gestión de los TFG y los requisitos.

Tabla 3-2.- Correspondencia de los UO de administración y gestión con los requisitos

| Código UO | Descripción | Requisito relacionado |
|------------------|-----------------------------|------------------------------|
| UO02-U | Acceder al sistema | REQ.10 |
| UO03-U | Gestionar datos personales | REQ.10 |
| UO03-A | Iniciar y finalizar TFG | REQ.10 |
| UO08-E | Gestionar documentación TFG | REQ.5 REQ.7 |
| UO09-E | Solicitar evaluación TFG | REQ.8 |
| UO04-T | Autorizar evaluación TFG | REQ.10 |
| UO02-Tr | Consultar documentación TFG | REQ.6 |

La Tabla 3-3 recoge la lista completa de los UO iniciales agrupados según los distintos roles presentes en el sistema.

Tabla 3-3.- Lista completa de Objetivos de Usuario iniciales

| Rol | Código UO | Descripción |
|------------------------|--------------------------|--|
| Persona usuaria | U001-U | Realizar consultas en base a intereses |
| | U002-U | Acceder al sistema |
| | U003-U | Gestionar datos personales |
| Persona administradora | U001-A | Gestionar criterios alcance |
| | U002-A | Gestionar criterios evaluación |
| | U003-A | Iniciar y finalizar TFG |
| Estudiante | U001-E | Conocer criterios de alcance |
| | U002-E | Realizar propuestas |
| | U003-E | Solicitar retroalimentación |
| | U004-E | Actualizar estado TFG |
| | U005-E | Conocer criterios evaluación |
| | U006-E | Solicitar evaluación elemento |
| | U007-E | Consultar evaluación |
| | U008-E | Gestionar documentación TFG |
| U009-E | Solicitar evaluación TFG | |
| Docente | U001-D | Usar criterios alcance |
| Persona supervisora | U001-T | Conocer estado TFG |
| | U002-T | Generar retroalimentación |
| | U003-T | Evaluar elemento |
| | U004-T | Autorizar evaluación TFG |
| Componente tribunal | U001-Tr | Evaluar TFG |
| | U002-Tr | Consultar documentación TFG |

A continuación se detalla el proceso seguido para el análisis de TraMeA a través de las sucesivas iteraciones definidas en InterMod.

3.3 Refinamiento de Objetivos de Usuario

Con la lista completa de UO iniciales de TraMeA presentada en la Tabla 3-3, se procedió a la ejecución de las sucesivas iteraciones del proceso establecido en InterMod para refinar los distintos UO y generar UO indirectos. En el caso de algunos UO, su sencillez provoca que no sea necesario refinarlos y por lo tanto no generen UO indirectos.

En la Tabla 3-4 se muestran los UO indirectos generados a lo largo del proceso para cada uno de los UO originales pertenecientes a los roles Persona usuaria y Persona administradora.

Tabla 3-4.- Objetivos de Usuario indirectos obtenidos en las iteraciones para los roles Persona usuaria y Persona administradora

| Código UO | Código UO Indirecto | Descripción |
|------------------|----------------------------|-----------------------------------|
| U001-U | U001-U1 | Buscar información temática |
| | U001-U2 | Realizar preguntas/comentarios |
| | U001-U3 | Responder preguntas/comentarios |
| U002-U | --- | Acceder al sistema |
| U003-U | U003-U1 | Consultar sus datos |
| | U003-U2 | Modificar sus datos |
| | U003-U3 | Solicitar la baja del sistema |
| U001-A | U001-A1 | Añadir criterios de alcance |
| | U001-A2 | Modificar criterios de alcance |
| | U001-A3 | Eliminar criterios de alcance |
| U002-A | U002-A1 | Añadir criterios de evaluación |
| | U002-A2 | Modificar criterios de evaluación |
| | U002-A3 | Eliminar criterios de evaluación |
| U003-A | U003-A1 | Crear TFG y personas asociadas |
| | U003-A2 | Finalizar TFG |
| | U003-A3 | Asignar tribunal de evaluación |

En la Tabla 3-5 se recogen los UO indirectos generados para el rol Estudiante a lo largo de las iteraciones.

Tabla 3-5.- Objetivos de Usuario indirectos obtenidos en las iteraciones para el rol Estudiante

| Código UO | Código UO Indirecto | Descripción |
|------------------|----------------------------|-------------------------------|
| U001-E | --- | Conocer criterios de alcance |
| U002-E | --- | Realizar propuesta |
| U003-E | U003-E1 | Solicitar retroalimentación |
| | U003-E2 | Consultar retroalimentación |
| U004-E | --- | Actualizar estado TFG |
| U005-E | --- | Conocer criterios evaluación |
| U006-E | --- | Solicitar evaluación elemento |
| U007-E | --- | Consultar evaluación |
| U008-E | U008-E1 | Consultar documentación TFG |
| | U008-E2 | Crear nuevo documento |
| | U008-E3 | Eliminar documento |
| U009-E | --- | Solicitar evaluación TFG |

La Tabla 3-6 recoge los UO indirectos surgidos en las sucesivas iteraciones del proceso, correspondientes a los roles Docente, Persona supervisora y Componente tribunal.

Tabla 3-6.- Objetivos de Usuario indirectos obtenidos en las iteraciones para los roles Docente, Persona supervisora y Componente tribunal

| Código UO | Código UO Indirecto | Descripción |
|------------------|----------------------------|---|
| U001-D | U001-D1 | Ver criterios de alcance |
| | U001-D2 | Evaluar usando los criterios de alcance |
| | U001-D3 | Consultar evaluación realizada |
| U001-T | U001-T1 | Consultar planificación |
| | U001-T2 | Consultar actividad estudiante |
| | U001-T3 | Consultar detalles tarea |
| U002-T | U002-T1 | Consultar documento |
| | U002-T2 | Dar retroalimentación |
| | U002-T3 | Consultar retroalimentación |
| U003-T | U003-T1 | Ver criterios evaluación elemento |
| | U003-T2 | Evaluar elemento |
| | U003-T3 | Consultar evaluación de elemento |
| U004-T | --- | Autorizar evaluación TFG |
| U001-Tr | U001-Tr1 | Ver criterios evaluación TFG |
| | U001-Tr2 | Evaluar TFG |
| | U001-Tr3 | Consultar evaluación de TFG |
| U002-Tr | U002-Tr1 | Consultar planificación |
| | U002-Tr2 | Consultar actividad estudiante |
| | U002-Tr3 | Consultar detalles tarea |
| | U002-Tr4 | Consultar documento |

Durante el refinamiento de los UO a través de las iteraciones también se puede producir la fusión entre UO similares. Por ejemplo, el U001-A1 y el U002-A1 (ver Tabla 3-4) consisten en añadir los criterios que se van a usar para evaluar el alcance y los distintos elementos evaluables del TFG respectivamente. En ambos casos la funcionalidad consiste en añadir al sistema los criterios de evaluación, siendo la única diferencia el texto que describe cada criterio. Por lo tanto, desde el punto de vista del diseño y desarrollo, deberían ser considerados una única funcionalidad. Por ello, estos UO se pueden fusionar creando un único UO que es el que finalmente se ha diseñado e implementado (UOF01 en Tabla 3-7).

CAPÍTULO 3

También se pueden fusionar UO de distintos roles cuyo funcionamiento sea similar. Por ejemplo, tanto el UO03-E1 (Tabla 3-5) como el UO02-T3 (Tabla 3-6) permiten la consulta de la retroalimentación generada previamente, por lo que se han fusionado en el UOF06 (Tabla 3-7).

Entre la Tabla 3-7 y la Tabla 3-8 se muestran los distintos UO indirectos que se han fusionado y los nuevos UO a los que ha dado lugar dicha fusión.

Tabla 3-7.- Lista de Objetivos de Usuario indirectos fusionados (Parte I)

| Código UO Fusionado | Descripción | Código UO Indirecto | Descripción |
|----------------------------|-----------------------------|----------------------------|-----------------------------------|
| UOF01 | Añadir criterios | U001-A1 | Añadir criterios de alcance |
| | | U002-A1 | Añadir criterios de evaluación |
| UOF02 | Modificar criterios | U001-A2 | Modificar criterios de alcance |
| | | U002-A2 | Modificar criterios de evaluación |
| UOF03 | Eliminar criterios | U001-A3 | Eliminar criterios de alcance |
| | | U002-A3 | Eliminar criterios de evaluación |
| UOF04 | Ver criterios | U001-E | Conocer criterios de alcance |
| | | U005-E | Conocer criterios evaluación |
| | | U001-D1 | Ver criterios de alcance |
| | | U003-T1 | Ver criterios evaluación elemento |
| | | U001-Tr1 | Ver criterios evaluación TFG |
| UOF05 | Solicitar evaluación | U002-E | Realizar propuesta |
| | | U003-E1 | Solicitar retroalimentación |
| | | U006-E | Solicitar evaluación elemento |
| | | U009-E | Solicitar evaluación TFG |
| UOF06 | Consultar retroalimentación | U003-E1 | Consultar retroalimentación |
| | | U002-T3 | Consultar retroalimentación |

Tabla 3-8.- Lista de Objetivos de Usuario indirectos fusionados (Parte II)

| Código UO Fusionado | Descripción | Código UO Indirecto | Descripción |
|----------------------------|--------------------------------|----------------------------|---|
| UOF07 | Consultar planificación | U001-T1 | Consultar planificación |
| | | U002-Tr1 | Consultar planificación |
| UOF08 | Consultar actividad estudiante | U001-T2 | Consultar actividad estudiante |
| | | U002-Tr2 | Consultar actividad estudiante |
| UOF09 | Consultar detalles tarea | U001-T3 | Consultar detalles tarea |
| | | U002-Tr3 | Consultar detalles tarea |
| UOF10 | Consultar documento | U008-E1 | Consultar documentación TFG |
| | | U002-T1 | Consultar documento |
| | | U002-Tr4 | Consultar documento |
| UOF11 | Evaluar | U001-D2 | Evaluar usando los criterios de alcance |
| | | U003-T2 | Evaluar elemento |
| | | U001-Tr2 | Evaluar TFG |
| UOF12 | Consultar evaluación | U007-E | Consultar evaluación |
| | | U001-D3 | Consultar evaluación realizada |
| | | U003-T3 | Consultar evaluación de elemento |
| | | U001-Tr3 | Consultar evaluación de TFG |

La Tabla 3-9 recoge a modo de resumen la lista final de los UO de TraMeA tras las divisiones y fusiones que se han producido a lo largo del proceso de refinamiento.

Tabla 3-9.- Objetivos de Usuario finales de TraMeA

| Código UO | Descripción |
|------------------|---------------------------------|
| U001-U1 | Buscar información temática |
| U001-U2 | Realizar preguntas/comentarios |
| U001-U3 | Responder preguntas/comentarios |
| U002-U | Acceder al sistema |
| U003-U1 | Consultar sus datos |
| U003-U2 | Modificar sus datos |
| U003-U3 | Solicitar la baja del sistema |
| U003-A1 | Crear TFG y personas asociadas |
| U003-A2 | Finalizar TFG |
| U003-A3 | Asignar tribunal de evaluación |
| U004-E | Actualizar estado TFG |
| U008-E2 | Crear nuevo documento |
| U008-E3 | Eliminar documento |
| U002-T2 | Dar retroalimentación |
| U004-T | Autorizar evaluación TFG |
| UOF01 | Añadir criterios |
| UOF02 | Modificar criterios |
| UOF03 | Eliminar criterios |
| UOF04 | Ver criterios |
| UOF05 | Solicitar evaluación |
| UOF06 | Consultar retroalimentación |
| UOF07 | Consultar planificación |
| UOF08 | Consultar actividad estudiante |
| UOF09 | Consultar detalles tarea |
| UOF10 | Consultar documento |
| UOF11 | Evaluar |
| UOF12 | Consultar evaluación |

3.4 Diseño y arquitectura

TraMeA se basa en Redmine (ver sección 2.4.3), un software ya existente, por lo que es posible que algunos de los UO obtenidos se encuentren directamente satisfechos gracias a las funcionalidades aportadas por dicho software. Por lo tanto, antes de proceder al diseño e implementación de los UO definidos siguiendo las pautas establecidas por InterMod, se comprobó si alguno de ellos se satisfacían directamente con las funcionalidades aportadas por Redmine.

Tras analizar las características y funcionalidades ofrecidas por Redmine, se identificaron los UO que satisface y cuyo análisis a través del proceso de iteraciones de InterMod se puede dar por finalizado. A su vez, se identificaron los UO todavía pendientes de resolver y que por lo tanto, habría que desarrollar a lo largo de las sucesivas iteraciones de InterMod. La Tabla 3-10 presenta los UO de TraMeA, indicando para cada uno, si se satisface a través de las funcionalidades aportadas por Redmine (✓) o si debe desarrollarse siguiendo las pautas marcadas por InterMod (✗).

Tabla 3-10.- Objetivos de Usuario de TraMeA en relación a las funcionalidades proporcionadas por Redmine

| Código UO | Descripción | Satisfecho por Redmine |
|------------------|---------------------------------|-------------------------------|
| U001-U1 | Buscar información temática | ✓ |
| U001-U2 | Realizar preguntas/comentarios | ✓ |
| U001-U3 | Responder preguntas/comentarios | ✓ |
| U002-U | Acceder al sistema | ✓ |
| U003-U1 | Consultar sus datos | ✓ |
| U003-U2 | Modificar sus datos | ✓ |
| U003-U3 | Solicitar la baja del sistema | ✓ |
| U003-A1 | Crear TFG y personas asociadas | ✓ |
| U003-A2 | Finalizar TFG | ✓ |
| U003-A3 | Asignar tribunal de evaluación | ✓ |
| U004-E | Actualizar estado TFG | ✓ |
| U008-E2 | Crear nuevo documento | ✓ |
| U008-E3 | Eliminar documento | ✓ |
| U002-T2 | Dar retroalimentación | ✓ |
| U004-T | Autorizar evaluación TFG | ✓ |
| UOF01 | Añadir criterios | ✗ |
| UOF02 | Modificar criterios | ✗ |
| UOF03 | Eliminar criterios | ✗ |
| UOF04 | Ver criterios | ✗ |
| UOF05 | Solicitar evaluación | ✗ |
| UOF06 | Consultar retroalimentación | ✓ |
| UOF07 | Consultar planificación | ✓ |
| UOF08 | Consultar actividad estudiante | ✓ |
| UOF09 | Consultar detalles tarea | ✓ |
| UOF10 | Consultar documento | ✓ |
| UOF11 | Evaluar | ✗ |
| UOF12 | Consultar evaluación | ✗ |

Como se puede observar en la Tabla 3-10 todos los UO pendientes de satisfacer están relacionados con el proceso de evaluación. Por lo tanto, para completar el desarrollo de TraMeA es necesario definir un módulo que se encargue de todos los aspectos relacionados con la evaluación y que se comuniquen con Redmine (ver Fig. 3-3).

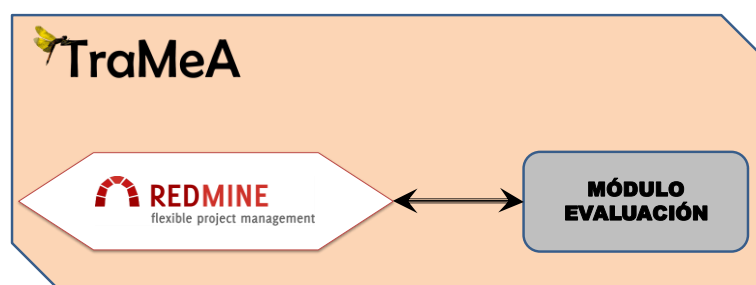


Fig. 3-3.- Arquitectura inicial de TraMeA

Al abordar el desarrollo los UO relacionados con la evaluación, se decidió no limitar el ámbito de aplicación del sistema de evaluación que se está planteando a los TFG y facilitar, así, su reutilización en cualquier otra asignatura. Sin embargo, la arquitectura inicial propuesta para TraMeA (Fig. 3-3) no es adecuada para la mayoría de las asignaturas, en las que el uso de Redmine y sus funcionalidades relativas a la gestión de proyectos carece de sentido y no aporta ninguna ventaja.

Por esta razón se decidió ampliar las características iniciales del módulo de evaluación y desarrollar un sistema independiente que pueda ofrecer soporte a distintas tareas de evaluación, interactuando para ello con sistemas de propósito específico como Redmine o con entornos de apoyo al aprendizaje tales como MAgAdI (A. Álvarez, 2010) o Moodle³³ (ver Fig. 3-4). Este sistema independiente se ha denominado AdESMuS (*Adaptable Evaluation System using Multiple Sources*).

AdESMuS es un sistema que soporta múltiples configuraciones para adaptarse a la evaluación de diferentes asignaturas teniendo en cuenta diversos aspectos tales como: los elementos que se desean evaluar, la forma en que se van a evaluar, la influencia de cada elemento en la nota final o quiénes van a ser las personas encargadas de la evaluación.

³³ <http://moodle.org>

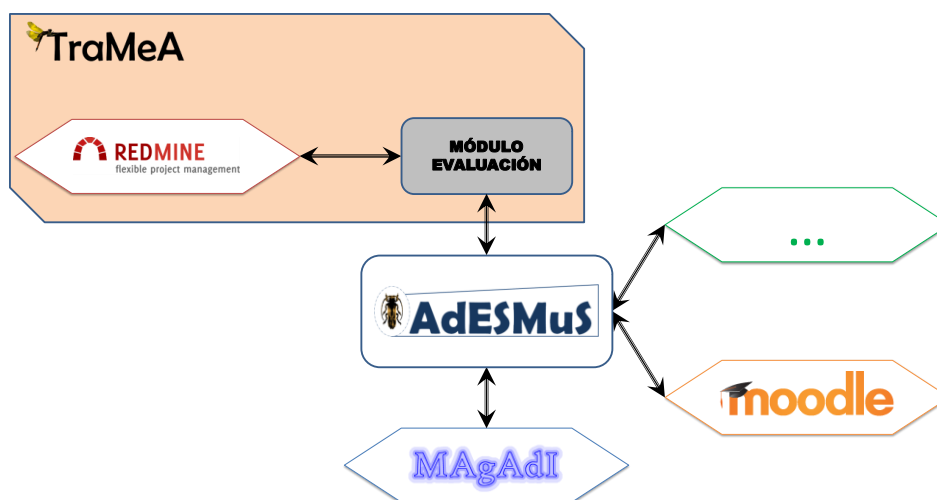


Fig. 3-4.- Arquitectura final de TraMeA con AdESMuS

De este modo, AdESMuS puede usarse en combinación con otros sistemas cuando la evaluación de las asignaturas no requiera considerar los aspectos relacionados con la realización y evaluación de TFG; o en combinación con Redmine para cubrir todos los requisitos establecidos para la realización, supervisión y evaluación de los TFG. En este último caso, el módulo de evaluación de TraMeA ejercerá de intermediario entre Redmine y los servicios de evaluación ofrecidos por AdESMuS. También será el encargado de satisfacer el UO UOF05.-«Solicitar evaluación» que, a pesar de ser necesario en TraMeA para que los estudiantes puedan solicitar la evaluación de sus avances en el TFG, en AdESMuS no tiene cabida ya que lo habitual en el resto de asignaturas es que sea el profesorado quien establezca qué se evalúa y cuándo.

3.5 Conclusiones

En este capítulo se ha presentado TraMeA, un sistema que da soporte a la metodología FPM. A partir de las encuestas realizadas tanto a docentes como a estudiantes se han extraído las funcionalidades que tiene que ofrecer TraMeA y, siguiendo la metodología InterMod, se han identificado los roles de usuario del sistema y se han definido los Objetivos de Usuario correspondientes.

Para construir TraMeA se ha reutilizado un software ya existente, Redmine, y se ha analizado cuáles de los UO definidos están satisfechos directamente por Redmine y cuáles quedan pendientes. Los UO pendientes son todos los

relacionados con la evaluación, por lo que resulta imprescindible desarrollar un módulo que proporcione las funcionalidades necesarias para evaluar los TFG.

Dadas las posibilidades de reutilización en otros ámbitos de un sistema de evaluación como el necesario para TraMeA, se ha decidido implementar un sistema de evaluación externo, que soporte más funcionalidades de las estrictamente necesarias para evaluar los TFG y que sea accesible no sólo desde Redmine, sino también desde otras herramientas de ayuda al aprendizaje. Este nuevo sistema se ha denominado AdESMuS y se describe en detalle en el siguiente capítulo.

4

AdESMuS

AdESMuS (*Adaptable Evaluation System using Multiple Sources*) es un sistema configurable que permite la evaluación de Trabajos Fin de Grado. Además, AdESMuS soporta múltiples configuraciones para adaptarse a diferentes asignaturas teniendo en cuenta diversos aspectos tales como: los elementos que se desean evaluar, la forma en que van a evaluarse, la influencia de cada elemento en la nota final o quiénes van a ser las personas evaluadoras.

En este capítulo se presenta AdESMuS, para cuyo desarrollo se ha seguido la metodología InterMod (ver sección 3.1). Primero se describe el análisis general realizado, donde se identifican los roles y los objetivos de usuario iniciales junto con el modelo del sistema. A continuación se presenta el refinamiento de los objetivos de usuario y el diseño del sistema que incluye la arquitectura del mismo. Posteriormente, se describen las ontologías definidas para las bases de conocimiento incluidas en la arquitectura (BC) y aspectos relacionados con la internacionalización y la perspectiva de género. A continuación se expone la implementación concreta que se ha realizado del sistema. El capítulo finaliza con las conclusiones extraídas del proceso de definición, diseño e implementación de AdESMuS.

4.1 Análisis general

La metodología InterMod (ver sección 3.1) establece el análisis general del sistema como la primera etapa del proceso de desarrollo de un sistema.

En los siguientes apartados se define el proceso seguido durante el análisis general para obtener los roles de usuario, el modelo del sistema y los UO iniciales correspondientes a AdESMuS.

4.1.1 Roles de usuario

A pesar de que AdESMuS surge a partir de TraMeA, ampliar sus funcionalidades para poder utilizarlo en cualquier asignatura conlleva la adaptación de los roles de usuario definidos en TraMeA.

Cuatro de los roles de usuario coinciden parcialmente con los definidos para TraMeA (Persona usuaria, Persona administradora, Docente y Estudiante), pero es necesario adaptar su definición a las funcionalidades ampliadas que ofrece AdESMuS.

El rol Persona supervisora desaparece al no existir dicha figura en las asignaturas convencionales. Además, dado que la evaluación no la lleva a cabo un tribunal, se genera un nuevo rol (Persona evaluadora) que engloba tanto a profesorado como alumnado para poder dar soporte a métodos de evaluación tales como la autoevaluación o la evaluación entre iguales (*peer review*).

Así pues, en AdESMuS se definen los 5 roles de usuario siguientes:

- Persona administradora: se encarga de las tareas de administración del sistema.
- Docente: se encarga de la docencia de las asignaturas que se evalúan en AdESMuS.
- Estudiante: persona evaluada mediante AdESMuS en alguna de las asignaturas que cursa.
- Persona evaluadora: persona que usa AdESMuS para evaluar a estudiantes. Puede ser docente o estudiante en el caso de que se esté realizando una coevaluación o autoevaluación.
- Persona usuaria: cualquiera que utiliza AdESMuS.

4.1.2 Modelo del sistema

AdESMuS se plantea como un sistema con capacidad para funcionar tanto de manera autónoma como combinado con otros sistemas. Para facilitar la interacción con otros sistemas, AdESMuS se limitará a proporcionar los servicios necesarios para definir y llevar a cabo la evaluación.

Por lo tanto, los sistemas que hagan uso de los servicios de AdESMuS tendrán la responsabilidad de proporcionar una interfaz gráfica que cumpla unos requisitos mínimos de usabilidad y accesibilidad.

4.1.3 Objetivos de usuario iniciales

En este apartado se recogen los UO iniciales definidos para el sistema AdESMuS. Para facilitar la lectura y evitar confusiones, los UO de AdESMuS se han nombrado añadiendo el prefijo *Ad* a la nomenclatura usada hasta ahora. Asimismo, para denotar los UO correspondientes al nuevo rol de usuario Persona evaluadora, se utilizará el sufijo *Eval*.

AdESMuS debe satisfacer como mínimo los UO que se habían definido para TraMeA y que estaban pendientes de desarrollo (ver UO marcados como ***** en la Tabla 3-10) a excepción del UO UOF05.-«Solicitar evaluación» (ver sección 3.4).

Además, existen algunos UO relacionados con el acceso al sistema y la gestión de los datos personales que en TraMeA estaban cubiertos gracias a Redmine, pero que AdESMuS necesita cubrir para lograr su autonomía.

En la Tabla 4-1 se recogen los UO que AdESMuS hereda de TraMeA.

Tabla 4-1.- Objetivos de Usuario heredados de TraMeA

| Rol | Código UO | Descripción | UO de TraMeA relacionados |
|--------------------|------------------|--|----------------------------------|
| Persona usuaria | Ad_UO01-U | Acceder al sistema | UO02-U |
| | Ad_UO02-U | Gestionar sus datos | UO03-U1, UO03-U2, UO03-U3 |
| Docente | Ad_UO01-D | Gestionar las formas de evaluación de los grupos | UOF01, UOF02, UOF03, UOF04 |
| Persona evaluadora | Ad_UO01-Eval | Evaluar | UOF10 |
| | Ad_UO02-Eval | Consultar las evaluaciones realizadas | UOF11 |
| Estudiante | Ad_UO01-Est | Consultar los resultados de las evaluaciones | |

Además, aparecen nuevos UO (ver Tabla 4-2) que surgen de las necesidades de administración para permitir la gestión y evaluación de cualquier asignatura.

Tabla 4-2.- Objetivos de Usuario de AdESMuS

| Rol | Código UO | Descripción |
|------------------------|-----------|---|
| Persona administradora | Ad_U001-A | Gestionar centros |
| | Ad_U002-A | Gestionar departamentos |
| | Ad_U003-A | Gestionar usuarios |
| | Ad_U004-A | Gestionar titulaciones |
| | Ad_U005-A | Gestionar asignaturas |
| | Ad_U006-A | Gestionar grupos |
| | Ad_U007-A | Gestionar la matrícula de los estudiantes en las distintas titulaciones, asignaturas y grupos |

4.2 Refinamiento de objetivos de usuario

Una vez identificados los UO directos, a lo largo de las sucesivas iteraciones definidas por el proceso de InterMod, éstos se dividen dando lugar a los UO indirectos. La división de los UO directos en UO indirectos permite refinar los requisitos funcionales de los usuarios para facilitar su desarrollo.

A continuación se detallan los UO indirectos obtenidos del refinamiento de los UO directos recogidos en las Tabla 4-1 y Tabla 4-2.

La Tabla 4-3 refleja la división en UO indirectos de los UO pertenecientes al rol Persona usuaria. En dicha tabla se observa cómo, por ejemplo, se refina el UO inicial Ad_UO02-U.- «Gestionar sus datos» en objetivos más específicos tales como la consulta o modificación de los datos.

Tabla 4-3.- División en Objetivos de Usuario Indirectos de los UO de Persona usuaria

| Código UO | Código UO Indirecto | Descripción |
|-----------|---------------------|-------------------------------|
| Ad_U001-U | Ad_U001-U1 | Acceder al sistema |
| | Ad_U001-U2 | Recuperar contraseña |
| Ad_U002-U | Ad_U002-U1 | Consultar sus datos |
| | Ad_U002-U2 | Modificar sus datos |
| | Ad_U002-U3 | Solicitar la baja del sistema |

Para el rol Docente sólo existe un UO (Ad_UO01-D.- «Gestionar las formas de evaluación de los grupos»), cuya división en UO indirectos se detalla en la Tabla 4-4.

Tabla 4-4.- División en Objetivos de Usuario Indirectos de los UO de Docente

| Código UO | Código UO Indirecto | Descripción |
|------------------|----------------------------|--|
| Ad_UO01-D | Ad_UO01-D1 | Añadir competencias a evaluar |
| | Ad_UO01-D2 | Modificar competencias a evaluar |
| | Ad_UO01-D3 | Eliminar competencias a evaluar |
| | Ad_UO01-D4 | Añadir elemento evaluable |
| | Ad_UO01-D5 | Modificar elemento evaluable |
| | Ad_UO01-D6 | Eliminar elemento evaluable |
| | Ad_UO01-D7 | Gestionar relación entre competencias y elementos evaluables |
| | Ad_UO01-D8 | Definir forma de evaluación |
| | Ad_UO01-D9 | Definir fuente de evaluación |
| | Ad_UO01-D10 | Definir rúbrica de evaluación |
| | Ad_UO01-D11 | Obtener listado de estudiantes a evaluar en un grupo |
| | Ad_UO01-D12 | Asignar forma de evaluación a estudiante |

El rol Persona evaluadora posee dos UO directos cuya división en UO indirectos se recoge en la Tabla 4-5.

Tabla 4-5.- División en Objetivos de Usuario Indirectos de los UO de Persona evaluadora

| Código UO | Código UO Indirecto | Descripción |
|------------------|----------------------------|---|
| Ad_UO01-Eval | Ad_UO01-Eval1 | Listar evaluaciones pendientes |
| | Ad_UO01-Eval2 | Listar estudiantes pendientes en una evaluación |
| | Ad_UO01-Eval3 | Evaluar un estudiante |
| Ad_UO01-Eval | Ad_UO01-Eval1 | Buscar evaluaciones |
| | Ad_UO01-Eval2 | Consultar evaluación |

En la Tabla 4-6 se detalla la división en UO indirectos del único UO del rol Estudiante (Ad_UO01-Est.- «Consultar los resultados de las evaluaciones»).

Tabla 4-6.- División de Ad_UO01-Est en Objetivos de Usuario Indirectos

| Código UO | Código UO Indirecto | Descripción |
|------------------|----------------------------|-------------------------------|
| Ad_UO01-Est | Ad_UO01-Est1 | Listar evaluaciones recibidas |
| | Ad_UO01-Est2 | Consultar detalles evaluación |

En el caso de los objetivos de usuario del rol Persona administradora, su división en UO indirectos es muy similar ya que todos son relativos a la gestión de los distintos elementos necesarios para administrar la estructura organizativa. La Tabla 4-7 detalla, a modo de ejemplo, los UO indirectos derivados del UO Ad_UO01-A.- «Gestionar Centros».

Tabla 4-7.- División de Ad_UO01-A en Objetivos de Usuario Indirectos

| Código UO | Código UO Indirecto | Descripción |
|------------------|----------------------------|-------------------------|
| Ad_UO01-A | Ad_UO01-A1 | Añadir centro |
| | Ad_UO01-A2 | Consultar datos centro |
| | Ad_UO01-A3 | Modificar datos centro |
| | Ad_UO01-A4 | Eliminar centro |
| | Ad_UO01-A5 | Obtener listado centros |

4.3 Diseño y arquitectura

La metodología ágil InterMod que se ha utilizado para desarrollar AdESMuS, al igual que la mayoría de metodologías ágiles, se caracteriza por una intensiva comunicación entre todos los participantes, una retroalimentación continua y frecuentes pruebas y evaluaciones.

Sin embargo, InterMod no aborda una de las principales carencias de las metodologías ágiles (Mishra & Mishra, 2011; Stojanovic et al., 2003): la poca importancia que dan a las decisiones sobre el diseño y la arquitectura del sistema general. Estas metodologías se centran en la generación de código a través de iteraciones que producen pequeños incrementos funcionales en el sistema, sin establecer una arquitectura que soporte el sistema completo. Esto puede ocasionar

problemas de escalabilidad y mantenimiento al desarrollar sistemas de cierta complejidad (Pressman, 2005).

Para paliar esa carencia, durante la aplicación de InterMod, se han analizado las funcionalidades que ofrece AdESMuS y se ha definido una arquitectura modular que facilita su escalabilidad y garantiza su independencia de otras herramientas con las que podría integrarse (ver Fig. 4-1).

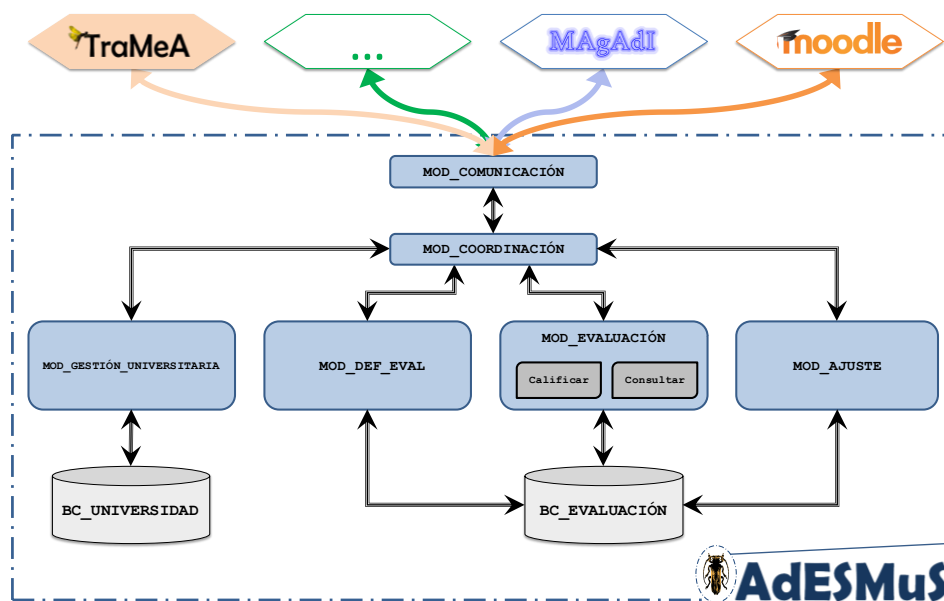


Fig. 4-1.- Arquitectura modular de AdESMuS

AdESMuS se divide en 6 módulos y dos bases de conocimiento. La estructura conceptual de las dos bases de conocimiento (BC_Universidad y BC_Evaluación) se define mediante tres ontologías. Ont_Universidad permite representar el conocimiento necesario para la gestión del sistema, Ont_Definición representa la estructura que permite definir la evaluación de una asignatura, mientras que la parte de la realización de las evaluaciones se representa en Ont_Evaluación.

En lo que respecta a los módulos, Mod_Comunicación ejerce de interfaz hacia el exterior y Mod_Coordinación actúa como interfaz entre el resto de los módulos del sistema. Los cuatro restantes, Mod_Gestión_Universitaria, Mod_Def_Eval, Mod_Evaluación y Mod_Ajuste, son módulos funcionales internos y tienen acceso a las bases de conocimiento.

Para conseguir que el diseño de AdESMuS sea lo más modular y escalable posible, se ha decidido que cada funcionalidad expresada como UO, sólo pueda tener acceso a una base de conocimiento. De este modo, se incrementa la seguridad del sistema, ya que al establecerse límites en los accesos, se minimiza el acoplamiento y se maximiza la cohesión de los módulos siguiendo los principios básicos de desarrollo de software orientado a objetos (Martin, 2003; Meyer, 1997). Por lo tanto, aquellos UO que precisen acceder a las dos bases de conocimiento de AdESMuS deberán dividirse en tantos UO indirectos como sea preciso para garantizar esta premisa.

Por ejemplo, para dar respuesta al Ad_UO01-D8.- «Definir forma de evaluación» (Tabla 4-4) es necesario acceder a BC_Evaluación para almacenar en ella información sobre cómo se va a evaluar a cada estudiante. Pero también es necesario acceder a BC_Universidad para guardar en qué asignaturas es posible usar esa forma de evaluación y sus respectivos elementos. Por lo tanto, este UO debe dividirse en UO indirectos más específicos, de modo que cada uno de ellos precise acceder únicamente a una base de conocimiento.

A través de las distintas iteraciones establecidas por InterMod, los UO originales (Tabla 4-1 y Tabla 4-2) se han dividido dando lugar, finalmente, a un listado de UO indirectos, cada uno de los cuales representa una funcionalidad específica que precisa acceso a una única base de conocimiento. Cada uno de los UO indirectos finales se aloja en uno de los módulos internos definidos en el sistema. El reparto de los UO indirectos entre los distintos módulos funcionales internos de AdESMuS se realiza en función de la base de conocimiento a la que necesitan acceder.

Además, debido a la gran cantidad de UO que acceden a BC_Evaluación, éstos se agrupan en distintos módulos teniendo en cuenta las características de la funcionalidad que representan: definición de los aspectos relativos a la evaluación por un lado; realización de la evaluación y consulta de los resultados por otro y, por último, el cálculo del modelo óptimo de evaluación.

A continuación se describe brevemente cada uno de los módulos de AdESMuS.

Módulos de AdESMuS

El módulo de comunicación de AdESMuS (Mod_Comunicación) ejerce de nexo de unión entre los módulos internos y la interfaz gráfica de cada tipo de usuario o sistema educativo. Mod_Comunicación proporciona los servicios necesarios para conectar AdESMuS con TraMeA u otros sistemas de apoyo al aprendizaje tales como Moodle o MAgAdI (A. Álvarez et al., 2013, 2014). Mod_Comunicación presenta el conjunto completo de servicios o funcionalidades que ofrece AdESMuS,

funcionando como nexo de unión entre la parte interna de la arquitectura y los sistemas externos que hagan uso de las funcionalidades de AdESMuS.

Mod_Comunicación es el responsable de ofrecer una respuesta a todos los objetivos de usuario que cualquier usuario o sistema externo pudiera definir. Es el único módulo de AdESMuS accesible desde el exterior del sistema, incrementando, de este modo, la seguridad del mismo.

El módulo de coordinación (Mod_Coordinación) es el encargado de gestionar la comunicación entre los módulos internos de AdESMuS, minimizando sus dependencias y proporcionando a Mod_Comunicación los datos y funcionalidades que se desean ofrecer al exterior. Es el encargado de recibir las peticiones de Mod_Comunicación y redirigirlas al resto de módulos internos en función de la base de conocimiento a la que se deba acceder para darles respuesta.

Si para atender una petición, Mod_Coordinación necesitara acceder tanto a BC_Universidad como a BC_Evaluación, gestionaría las llamadas a los UO correspondientes en cada módulo y sería el responsable de administrar el trasvase de información entre los módulos para componer la respuesta a la petición y devolverla a través de Mod_Comunicación.

Gracias a esta organización, tanto los módulos existentes en AdESMuS como cualquier otro que se quiera añadir en un futuro sólo necesitarán acceder a Mod_Coordinación, evitando dependencias con el resto de módulos.

El módulo Mod_Gestión_Universitaria permite la definición y gestión de todos los elementos pertenecientes a la base de conocimiento que contiene los datos administrativos necesarios para la gestión del sistema (BC_Universidad). Entre esos datos se encuentran por ejemplo la información sobre las asignaturas, las titulaciones a las que pertenecen, los docentes responsables o los estudiantes que cursan cada asignatura.

El módulo de definición de formas de evaluación (Mod_Def_Eval) permite al docente gestionar todo el proceso de definición de la evaluación de una asignatura concreta. A través de este módulo, los docentes responsables de los grupos pueden definir las formas de evaluar las asignaturas, que se almacenarán en BC_Evaluación.

El módulo de evaluación (Mod_Evaluación) se divide a su vez en dos sub-módulos de acuerdo la funcionalidad que se desee abordar: realizar la evaluación (sub-módulo Evaluar) o consultar los resultados (sub-módulo Consultar).

El sub-módulo Evaluar da respuesta a los UO indirectos involucrados en la realización de las evaluaciones que tengan asignadas tanto docentes como estudiantes (en el caso de la coevaluación). El sub-módulo Consultar se encarga de

satisfacer los UO relacionados con la consulta de las evaluaciones, tanto de los estudiantes como de los docentes.

Finalmente, el módulo `Mod_Ajuste` facilita la labor de configurar el sistema de evaluación de una asignatura. La evaluación de una asignatura puede definirse como la combinación de una serie de evaluaciones parciales, cada una de ellas con un porcentaje de influencia distinto en la calificación final. Configurar el sistema de evaluación estableciendo el porcentaje de influencia más adecuado en la calificación final para cada una de las evaluaciones parciales, no es sencillo. Sobre todo en el caso de docentes noveles o nuevas asignaturas. Además, si en la evaluación se utilizan rúbricas, también hay que determinar los valores numéricos correspondientes a cada nivel de valoración y el peso de cada dimensión en la nota calculada por la rúbrica, lo que es una tarea compleja (ver sección 2.2.3).

Para facilitar esta labor, `Mod_Ajuste` permite ejecutar a demanda un conjunto de procesos que calculan de manera automática el porcentaje de influencia óptimo de cada elemento que forma parte de una evaluación. Si se utilizan rúbricas, `Mod_Ajuste` también calcula cuáles son los valores correspondientes a los niveles de valoración y el peso más adecuado para cada dimensión.

`Mod_Ajuste` también se encarga de actualizar `BC_Evaluación` para almacenar los nuevos porcentajes de influencia, valores y pesos ideales de cada elemento. A continuación se detallan los procesos de ajuste definidos.

4.4 Procesos de ajuste

Para establecer los porcentajes de influencia, valores y pesos más adecuados para cada elemento que forma parte de una evaluación, el módulo de ajuste (`Mod_Ajuste`) ejecuta dos procesos distintos.

Por un lado, `Mod_Ajuste` puede calcular la influencia que cada tarea (o parte en la que se divide la evaluación) debe tener en el cálculo de la nota final. De este proceso se obtiene el Modelo Ajustado de Tareas (*Task Adjusted Model*, TAM).

Por otro lado, cuando la calificación de alguna tarea se realiza utilizando rúbricas de evaluación, `Mod_Ajuste` permite determinar el Modelo de Elementos de las Rúbricas (*Rubric Element Model*, REM). El REM establece la combinación más adecuada entre el peso de cada dimensión de las rúbricas que forman parte de la evaluación y los valores numéricos que se deberían asociar a los distintos niveles de valoración de las mismas.

Al modelo obtenido por el módulo de ajuste unificando el TAM y el REM para la evaluación de una asignatura se le ha denominado Modelo de Evaluación Multi-Fuente (*Multi-Source Evaluation Model*, MSEM).

Para que el módulo de ajuste pueda llevar a cabo ambos procesos (obtención del TAM y obtención del REM) es necesario recopilar datos relativos a las evaluaciones. Por un lado, el profesorado (en base a su opinión personal) asigna la calificación correspondiente a cada una de las tareas (Fig. 4-2→a). Por otro lado, el profesorado evalúa las tareas usando las rúbricas de evaluación correspondientes (Fig. 4-2→b). Por último, debe indicar qué nota final considera que debería obtener cada estudiante en la asignatura (Fig. 4-2→c).

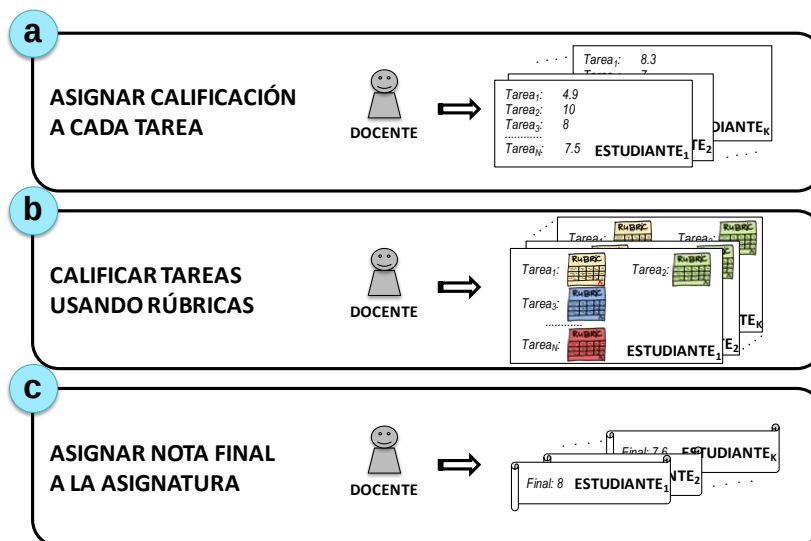


Fig. 4-2.- Recolección de datos necesarios para los procesos de ajuste

Mediante la combinación de las calificaciones recogidas en estos tres pasos se obtienen los datos necesarios para calcular el TAM y el REM, y por lo tanto el MSEM, a través de los procesos que se detallan a continuación.

Proceso de obtención del Task Adjusted Model (TAM)

Para obtener el TAM, que establece la influencia que cada tarea debe tener en la evaluación final, es necesario contar con las notas asignadas por el profesorado en base a su opinión personal. Tanto las correspondientes a cada tarea (Fig. 4-2→a) como las notas finales de la asignatura (Fig. 4-2→c).

A partir de dichos datos, establecer la influencia de cada tarea en la nota final es un problema de regresión, donde la variable dependiente es la nota final del estudiante y las variables independientes son las notas asignadas por los docentes a cada tarea.

Como todas las tareas que se consideran en la evaluación de una asignatura reflejan aspectos que se deben satisfacer, el porcentaje de influencia de todas las tareas tiene que ser mayor o igual a cero. Asociar a una tarea un porcentaje de influencia igual a cero indica que dicha tarea no influye realmente en el cálculo de la nota final del estudiante. Sin embargo, un porcentaje de influencia negativo carece de sentido ya que significa que se están evaluando aspectos que no deben cumplirse, que se están puntuando de manera positiva los aspectos negativos del trabajo. Esto es, que cuanto mayor sea la calificación del estudiante en esos aspectos, menor será su nota final.

Para asegurar que los porcentajes de influencia obtenidos en el TAM son coherentes, el proceso de obtención del TAM usa el algoritmo de mínimos cuadrados con coeficientes positivos de Lawson-Hanson (*Non-negative least squares*, NNLS) (Lawson & Hanson, 1995) tal y como se muestra en la Fig. 4-3.

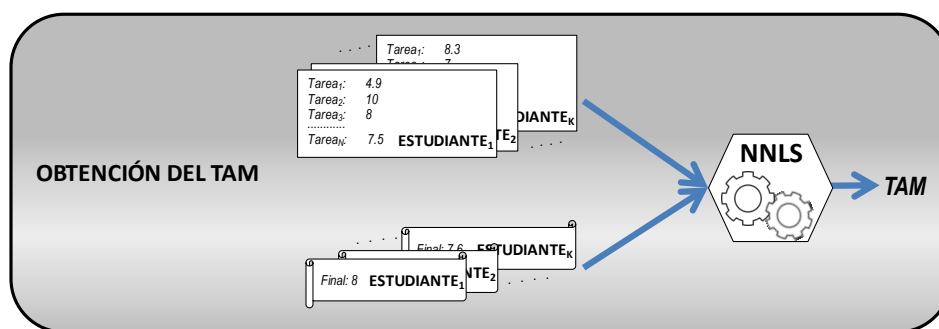


Fig. 4-3.- Proceso de obtención del TAM

Proceso de obtención del Rubric Element Model (REM).

Para todas aquellas tareas cuya evaluación se realice a través de rúbricas, es necesario calcular el Modelo de Elementos de las Rúbricas (REM).

El REM está compuesto de dos elementos. El primero es el Modelo de niveles de la Escala (*Scale-level Model*, SM) que establece el valor numérico más adecuado para cada nivel de valoración de las rúbricas. El segundo elemento es el peso más adecuado para cada dimensión presente en las rúbricas.

Cuando en la evaluación de una asignatura se utilizan rúbricas en varias tareas, el REM calculado por `Mod_Ajuste` usa el mismo SM para todas ellas. Por lo tanto,

el número de niveles de valoración debe ser el mismo en todas las rúbricas. Así se consigue coherencia a lo largo de toda la evaluación, garantizándose que las rúbricas usadas son similares y que los niveles de valoración están asociados al mismo valor numérico.

El proceso de obtención del REM también es un problema de regresión donde la variable dependiente es la nota proporcionada por las rúbricas y las variables independientes son los niveles de valoración indicados por las personas evaluadoras para cada una de las dimensiones. Para poder llevar a cabo este proceso es necesario contar con todos los datos recogidos (ver Fig. 4-2): las notas otorgadas por los evaluadores a cada tarea en base a su opinión personal, los datos correspondientes a la evaluación de las tareas usando las rúbricas y la nota final otorgada por los expertos a cada estudiante en la asignatura.

Para obtener los valores más adecuados para el REM, *Mod_Ajuste* ejecuta un proceso dividido en cuatro pasos (Ver Fig. 4-4).

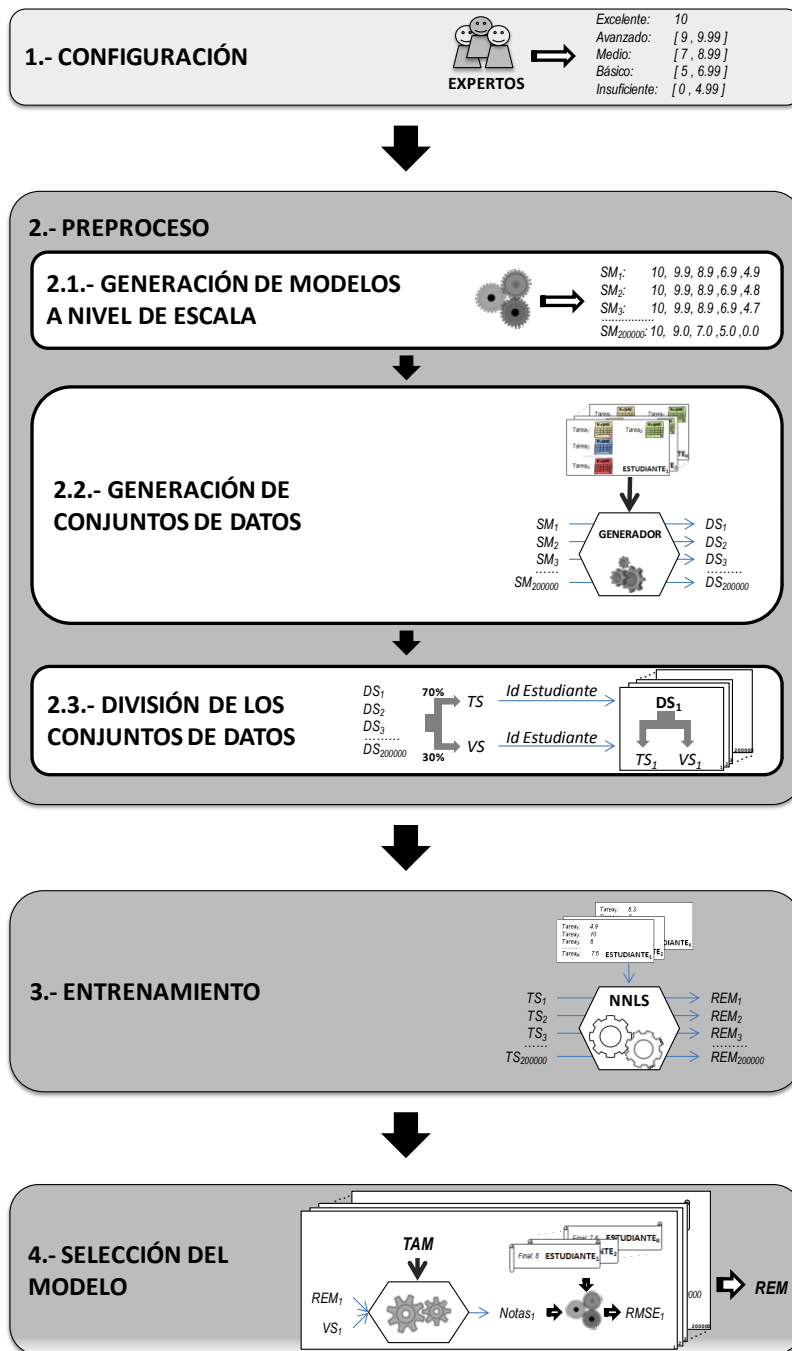


Fig. 4-4.- Proceso de obtención del REM

En el paso de «Configuración», se asigna un rango de posibles valores numéricos a cada nivel de valoración de la rúbrica. Dicha asignación la realiza el grupo de expertos que haya definido las rúbricas, que, por lo tanto, son los más apropiados para indicar el valor numérico aproximado correspondiente a cada nivel de valoración.

Por ejemplo, en un sistema en el que las calificaciones estén en el rango [0,10] y existan 5 niveles de valoración, podría asignarse al nivel superior la nota más alta: el 10; al siguiente nivel en la escala, el rango [9, 9.9] y así sucesivamente hasta el nivel de valoración más bajo. De este modo se podría generar una correspondencia entre los niveles de valoración y las categorías de calificación usadas tradicionalmente en el sistema educativo español (ver Tabla 4-8)

Tabla 4-8.- Posible asignación de valores numéricos

| Nivel de valoración de la rúbrica | Calificación (categoría) | Rango de valores |
|--|---------------------------------|-------------------------|
| Excelente | Matrícula de Honor | 10 |
| Avanzado | Sobresaliente | [9, 9.9] |
| Medio | Notable | [7, 8.9] |
| Básico | Aprobado | [5, 6.9] |
| Insuficiente | Suspenso | [0, 4.9] |

A partir de los rangos establecidos, en el paso de «Preproceso», se generan los Modelos de niveles de la Escala (SM_i). Para generar los distintos SM, se tienen en cuenta todas las posibles combinaciones de valores numéricos para cada nivel de valoración. Las combinaciones se generan comenzando desde el valor numérico superior de cada rango y decrementando una décima a uno de los rangos en cada combinación.

Por ejemplo, para los rangos indicados en la Tabla 4-8, parte de los SM_i generados serían los mostrados en la Tabla 4-9. El primero de los modelos (SM₁) tiene los valores correspondientes al límite superior de los rangos definidos. Para generar el siguiente modelo (SM₂), se resta una décima al valor correspondiente al nivel más bajo («Insuficiente») y se mantiene el resto. Volviendo a decrementar una décima al valor del nivel más bajo y manteniendo el resto de valores se van generando los distintos modelos. Una vez alcanzado el límite inferior del rango para un nivel («Insuficiente» en SM₄₀), se resta una décima en el nivel anterior («Básico») y se vuelven a inicializar los valores del resto de niveles al límite superior del correspondiente rango (SM₄₁). La generación de modelos continúa volviendo a decrementar una décima al valor del nivel más bajo (SM₄₂) y repitiendo el proceso

hasta alcanzar el límite inferior para los rangos definidos en todos los niveles de valoración (SM_{200000}).

Tabla 4-9.- Ejemplo de distintos SM para el rango [0, 10]

| | Excelente | Avanzado | Medio | Básico | Insuficiente |
|----------------------------|-----------|----------|-------|--------|--------------|
| SM₁ | 10 | 9.9 | 8.9 | 6.9 | 4.9 |
| SM₂ | 10 | 9.9 | 8.9 | 6.9 | 4.8 |
| SM₃ | 10 | 9.9 | 8.9 | 6.9 | 4.7 |
| ... | ... | ... | ... | ... | ... |
| SM₄₀ | 10 | 9.9 | 8.9 | 6.9 | 0 |
| SM₄₁ | 10 | 9.9 | 8.9 | 6.8 | 4.9 |
| SM₄₂ | 10 | 9.9 | 8.9 | 6.8 | 4.8 |
| ... | ... | ... | ... | ... | ... |
| SM₂₀₀₀₀₀ | 10 | 9 | 7 | 5 | 0 |

Una vez que se han obtenido los distintos SM y teniendo en cuenta las evaluaciones realizadas por los evaluadores utilizando las rúbricas, se genera un conjunto de datos (*dataset*, DS) para cada SM (paso 2.2 en Fig. 4-4). En cada DS_i se sustituye el nivel de valoración elegido por los evaluadores para cada dimensión por el valor numérico asignado a dicho nivel en el correspondiente SM_i (ver Tabla 4-9).

A continuación cada conjunto de datos se divide en datos de entrenamiento (*training set*, TS) y datos de validación (*validation set*, VS) en función de los identificadores de los estudiantes y siguiendo una proporción de 70 / 30 (paso 2.3 en Fig. 4-4).

Para garantizar que los modelos obtenidos se puedan comparar adecuadamente, es importante que el conjunto de entrenamiento sea el mismo para todos ellos. Para ello, la división en base a los identificadores se efectúa de manera aleatoria para TS_i y VS_i, y se respeta dicha división para el resto de TS_i y VS_i. De esta manera, se garantiza que todos los TS_i están formados por datos de los mismos estudiantes y que por lo tanto, los modelos obtenidos a partir de ellos son directamente comparables. Lo mismo sucede a la hora de validar los modelos obtenidos. Como

todos los VS_i contienen datos de los mismos alumnos, se pueden comparar los resultados de la validación de los distintos modelos.

Tras el paso de «Preproceso» tiene lugar el paso de «Entrenamiento». En este paso, a partir de cada conjunto de datos de entrenamiento (TS_i) y teniendo en cuenta las notas asignadas por los evaluadores a las tareas en base a su opinión personal, se obtiene un REM_i . En cada REM_i está reflejado el peso que cada dimensión de las rúbricas involucradas en el proceso de evaluación debería tener atendiendo a los valores numéricos asignados a los niveles de valoración en el correspondiente SM_i a partir del cual se generó el TS_i .

Al igual que sucede con las tareas o partes de la evaluación a la hora de calcular el TAM, las dimensiones de las rúbricas se corresponden con aspectos que debe cumplir el trabajo de los estudiantes. Por lo tanto, pesos negativos para las dimensiones carecen de sentido porque indicarían que cuanto mayor sea la valoración en esa dimensión, menor debe ser la calificación. Por esta razón, para calcular los REM_i (paso 3 en Fig. 4-4) se ha utilizado el algoritmo NNLS.

Al finalizar el paso de «Entrenamiento» se tienen tantos REM_i como SM_i se generaron en el paso 2.1 del proceso. Cada uno de los distintos REM_i contiene el peso más adecuado para cada una de las dimensiones de las distintas rúbricas que forman parte del proceso de evaluación teniendo en cuenta los valores numéricos establecidos por el correspondiente SM_i .

Para seleccionar el REM más adecuado entre todos los generados, el proceso finaliza con el paso «Selección del modelo» (paso 4 en Fig. 4-4). En este paso cada REM_i se aplica a los datos del correspondiente VS_i y, usando el TAM de la evaluación de la asignatura, se calculan las notas finales que dicho REM_i genera para los estudiantes del conjunto de datos de validación ($Notas_i$). Comparando estas notas con las otorgadas por las personas evaluadoras en base a su opinión personal, se calcula la Raíz del Error Cuadrático Medio (*Root-Mean-Squared Error*, RMSE) (Chai & Draxler, 2014) que estima cómo de adecuado es el correspondiente REM_i . El REM_i que produce el menor RMSE es el que presenta un comportamiento más parecido al deseado por los expertos y por lo tanto, será el REM seleccionado finalmente por el proceso.

4.5 Ontologías de AdESMuS

Para determinar la estructura de las ontologías de AdESMuS, se realizaron una serie de entrevistas con el profesorado. Las entrevistas se centraron en el proceso de evaluación de los TFG, ya que éstos son el núcleo de este trabajo de tesis y

presentan una casuística lo suficientemente diversa como para ilustrar las necesidades del sistema.

A continuación se resume la información extraída de las entrevistas sobre el proceso de evaluación de los TFG y se detallan las tres ontologías de AdESMuS que permiten representar esta información: Ont_Universidad, Ont_Definición y Ont_Evaluación. Además, se presentan diversos ejemplos de instanciación de las mismas.

4.5.1 Ejemplo de evaluación: el TFG

El Grado en Ingeniería Informática de Gestión y Sistemas de Información (GIIGySI) se imparte en dos centros de la UPV/EHU: la Escuela Universitaria de Ingeniería Técnica Industrial de Bilbao (EUITI) y la Escuela Universitaria de Ingeniería de Vitoria-Gasteiz (EUI).

En ambos casos la asignatura «*Trabajo Fin de Grado*» es una asignatura obligatoria correspondiente al cuarto curso. En el caso de la EUITI, el TFG se puede realizar en euskera o castellano indistintamente, por lo que la asignatura cuenta con dos grupos docentes distintos: el grupo 16 para el alumnado que realiza su TFG en castellano y el grupo 46 para los que lo realizan en euskera. En la EUI sólo se puede realizar el TFG en castellano, por lo que sólo existe el grupo docente de castellano (también denominado grupo 16).

Actualmente la evaluación de esta asignatura la realiza un tribunal formado por 3 docentes que emiten una valoración sobre un documento técnico y una presentación oral realizados por el alumnado.

Los responsables de la asignatura en ambos centros desean probar diferentes formas de evaluación para comparar los resultados que se obtienen con cada una de ellas y extraer criterios que garanticen la homogeneidad y transparencia en la calificación de los TFG.

Por un lado está la forma de evaluación que se ha denominado «tradicional», en la que se realiza una única valoración global sobre un documento técnico y una presentación oral. Por tanto, sólo hay un elemento evaluable, el conjunto del documento y la presentación, que se corresponde con el 100% de la evaluación. Las personas encargadas de realizar esa evaluación son los 3 docentes nombrados al efecto, y la nota se determina de manera «manual», sin utilizar ningún tipo de herramienta para su cálculo (ver Tabla 4-10).

Por otro lado, se desea probar la forma de evaluación que propone la metodología FPM presentada en este trabajo de tesis (ver sección 2.4) y a la que se ha denominado «continua». En esta forma de evaluación se definen los 6 elementos

evaluables presentados en 2.4.1: *Eva_Requisitos*, *Eva_Diseño*, *Eva_Actitud*, *Eva_Memoria*, *Eva_Presentación* y *Eva_Complejidad*. Asimismo, se siguen las recomendaciones de la metodología FPM sobre la persona más indicada para evaluar cada uno de estos elementos y se ha asignado un porcentaje de influencia en la nota final a cada uno de ellos.

De este modo, la persona supervisora, deberá evaluar los siguientes elementos evaluables, con los porcentajes de influencia que se indican para cada uno: *Eva_Requisitos* un 10%; *Eva_Diseño* un 15% y *Eva_Actitud* un 5%. Por otro lado, los elementos evaluables que deberán evaluar individualmente cada una de las 3 personas que conforman el tribunal son *Eva_Memoria* con una influencia del 45%; *Eva_Presentación* con un 15% y *Eva_Complejidad* con un 10%.

Para valorar los distintos elementos evaluables se desea usar rúbricas de evaluación, tanto de tipo analítico como de tipo atomístico (ver sección 2.2.1). De este modo, la fuente de obtención de la nota para los elementos evaluables será siempre alguno de estos tipos de rúbricas. Las personas encargadas de la evaluación podrán decidir qué tipo de rúbrica desean usar para cada elemento evaluable.

La Tabla 4-10 muestra un resumen de las dos formas de evaluación que se desean experimentar en la asignatura «*Trabajo Fin de Grado*» y que AdESMuS debe soportar.

Tabla 4-10.- Posibles formas de evaluación

| Forma de evaluación | Elemento evaluable | Porcentaje | Personas encargadas evaluación | Fuente de obtención de la nota |
|---------------------|--|------------|---------------------------------|--|
| Tradicional | Conjunto de documento técnico y presentación | 100% | Tribunal (3 docentes) | Manual |
| Continua | Eva_Requisitos | 10% | Persona supervisora (1 docente) | Rúbrica Analítica / Rúbrica Atomística |
| | Eva_Diseño | 12% | Persona supervisora (1 docente) | Rúbrica Analítica / Rúbrica Atomística |
| | Eva_Memoria | 53% | Tribunal (3 docentes) | Rúbrica Analítica / Rúbrica Atomística |
| | Eva_Actitud | 7% | Persona supervisora (1 docente) | Rúbrica Analítica / Rúbrica Atomística |
| | Eva_Presentación | 18% | Tribunal (3 docentes) | Rúbrica Analítica / Rúbrica Atomística |
| | Eva_Complejidad | 10% | Tribunal (3 docentes) | Rúbrica Analítica / Rúbrica Atomística |

4.5.2 Ont_Universidad: Datos administrativos

Ont_Universidad (cuya representación UML se muestra en la Fig. 4-5) recoge la información necesaria para definir las asignaturas que se desean evaluar con el sistema y su entorno: titulación y centro en el que se imparten, estudiantes matriculados y docentes responsables de definir la evaluación entre otros datos.

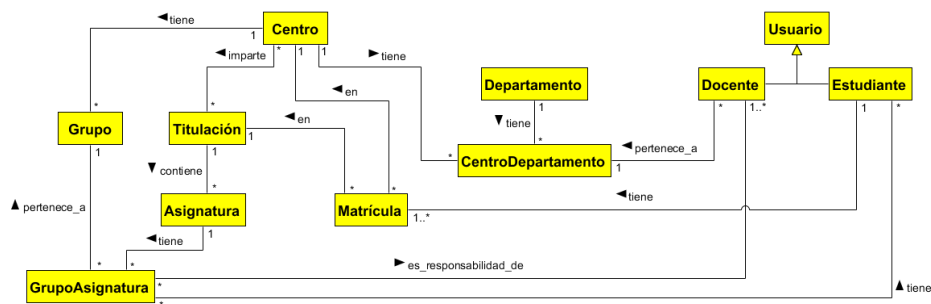


Fig. 4-5.- Representación UML de la ontología Ont_Universidad

La clase centro (ver Fig. 4-6→a) permite representar información sobre los centros educativos que usan el sistema AdESMuS e incluye una referencia a las titulaciones que se imparten y a los distintos grupos existentes en ese centro (clase Grupo, Fig. 4-6→c). Cada titulación (clase Titulación, Fig. 4-6→b), almacena, entre otros datos, su nombre y las asignaturas que la componen.

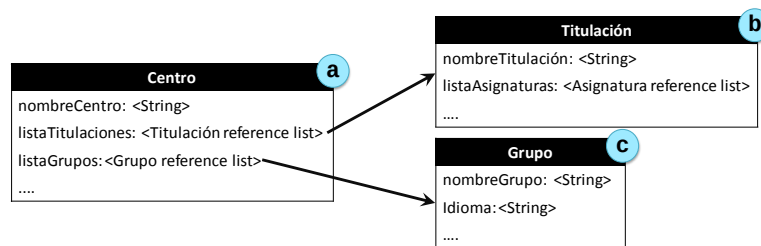


Fig. 4-6.- Definición de las clases Centro, Titulación y Grupo

La clase Asignatura (ver Fig. 4-7→a), identifica cada una de las asignaturas que se imparten en una titulación; y la clase GrupoAsignatura (ver Fig. 4-7→b) representa cada grupo de una asignatura concreta. Un GrupoAsignatura tiene uno o varios docentes responsables (atributo *listaResponsables*), y una serie de estudiantes (atributo *listaEstudiantes*) que se encuentran cursando la asignatura en él. Ambos datos se representan en forma de lista de referencias a objetos pertenecientes a las clases Docente y Estudiante respectivamente. Las clases Docente y Estudiante especializan la clase usuario, que representa de manera genérica cualquier tipo de usuario de AdESMuS (ver Fig. 4-5).

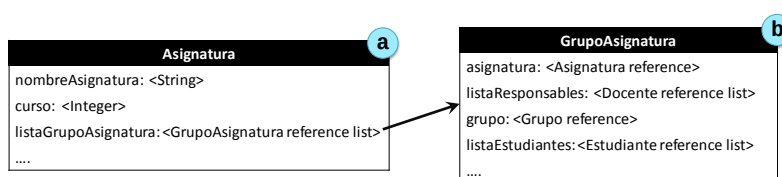


Fig. 4-7.- Definición de las clases Asignatura y GrupoAsignatura

La clase `Departamento` representa a los distintos departamentos que existen en la Universidad; y la clase `CentroDepartamento`, permite conocer qué departamentos imparten docencia en cada centro y a qué centro y departamento está adscrito cada docente.

Finalmente, la clase `Matrícula` permite representar la matrícula que realiza un estudiante para estudiar una titulación en un centro concreto.

Ejemplo de instanciación

A continuación se detalla un ejemplo con las instancias necesarias para definir la estructura administrativa básica para que AdESMuS soporte el funcionamiento descrito en la sección 4.5.1.

La Fig. 4-8 muestra los objetos correspondientes a las clases `Centro`, `Titulación` y `Grupo` necesarios para soportar el funcionamiento presentado en la sección 4.1. Los objetos [centro@1] y [centro@2] (Fig. 4-8→a), representan a la EUI de Vitoria-Gasteiz y la EUITI de Bilbao respectivamente. Ambos objetos incluyen en su atributo *listaTitulaciones* una referencia al objeto [titulacion@1] (ver Fig. 4-8→b), correspondiente a la titulación del Grado en Ingeniería Informática de Gestión y Sistemas de Información que se imparte en ambos centros.

La clase `Grupo` representa los grupos docentes de un centro concreto, por lo que aunque existan grupos con la misma nomenclatura en distintos centros, se representan con objetos distintos. Así por ejemplo, el grupo 16 de castellano se imparte en los dos centros y en cada uno se representa con una instancia diferente [grupo@1] en Vitoria (Fig. 4-8→c) y [grupo@3] en Bilbao (Fig. 4-8→d). En la EUITI de Bilbao se imparte además en euskera, por lo que en las referencias de su atributo *listaGrupos* se encuentra también el grupo 46, representado por el objeto [grupo@2] (Fig. 4-8→d).

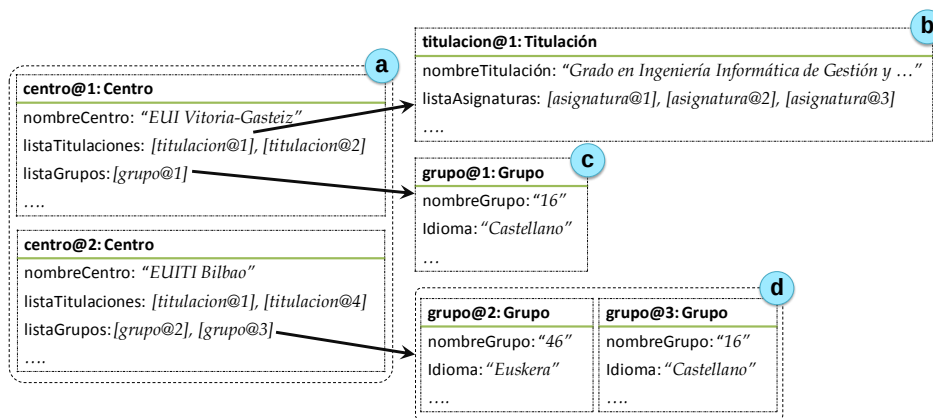


Fig. 4-8.- Ejemplo de objetos de las clases Centro, Titulación y Grupo

El atributo *listaAsignaturas* del objeto [titulacion@1], incluye las asignaturas definidas en la titulación, entre otras, el objeto [asignatura@1] correspondiente a la asignatura «Trabajo Fin de Grado» (ver Fig. 4-9→a).

En el atributo *listaGrupoAsignatura* del objeto [asignatura@1] se encuentran referenciados los tres objetos de la clase GrupoAsignatura existentes, [grupoasignatura@1] (Fig. 4-9→b), correspondiente al objeto [grupo@1] de la EUI de Vitoria-Gasteiz así como [grupoasignatura@2] y [grupoasignatura@3] (Fig. 4-9→c), relacionados con los objetos [grupo@2] y [grupo@3] de la EUITI de Bilbao. Además cada objeto GrupoAsignatura referencia a las personas encargadas de la asignatura en ese grupo (objetos de la clase Docente), y a los estudiantes que están cursando la asignatura en ese grupo (objetos de la clase Estudiante).

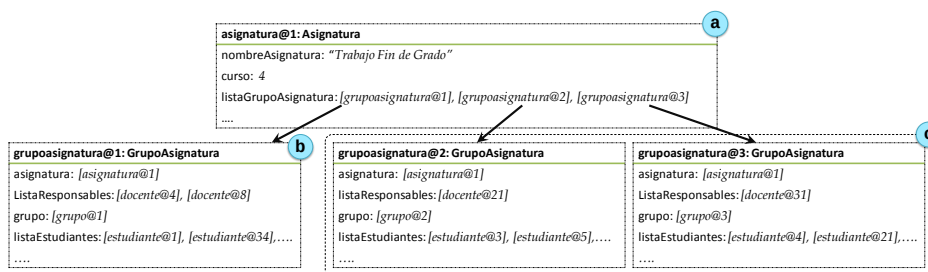


Fig. 4-9.- Ejemplo de objetos de las clases Asignatura y GrupoAsignatura

4.5.3 Ont_Definición: Definición de la evaluación

Ont_Definición (representada usando UML en la Fig. 4-10) especifica la estructura necesaria para definir la evaluación de una asignatura. Aunque las clases *Asignatura* y *GrupoAsignatura* pertenecen a Ont_Universidad (ver sección 4.5.2), se han incluido en esta figura para una mejor comprensión de la ontología Ont_Definición, cuyas clases se describen a continuación.

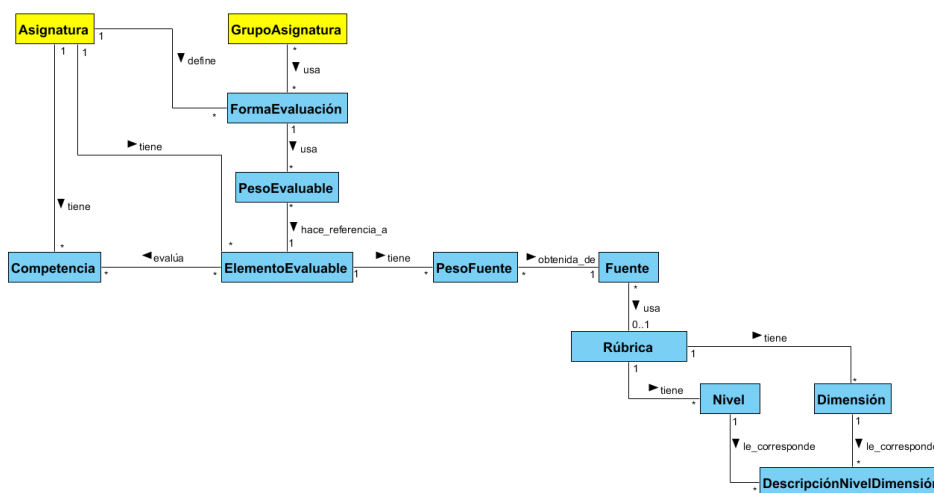


Fig. 4-10.- Representación UML de la ontología Ont_Definición

Clases para la definición general de la evaluación

Ont_Definición permite representar las distintas formas posibles de evaluar una asignatura (*FormaEvaluación*, Fig. 4-11→a). Los docentes responsables de cada *GrupoAsignatura* pueden elegir qué *FormaEvaluación* desean usar, así como los elementos (exámenes, entregables, prácticas, etc.) que se pueden tener en cuenta en la evaluación (*ElementoEvaluable*, Fig. 4-11→c). Cada *ElementoEvaluable* en una asignatura hace referencia a las competencias (clase *Competencia*, Fig. 4-11→d) que se están evaluando con dicho elemento.

Un mismo *ElementoEvaluable* se puede utilizar en distintas formas de evaluación, teniendo en cada una de ellas un porcentaje de influencia distinto en la nota final del estudiante. Por eso se define la clase *PesoEvaluable* (Fig. 4-11→b), que relaciona los *ElementoEvaluables* con las formas de evaluación, indicando el porcentaje de influencia correspondiente en cada caso. *PesoEvaluable* tiene dos atributos para indicar el porcentaje de influencia: *porcentajeReal* y *porcentajeIdea1*. El atributo *porcentajeReal* representa el porcentaje indicado por

el docente encargado de la definición y se usará para el cálculo de la nota; *porcentajeIdeal* se actualizará con el porcentaje de influencia óptimo (TAM, ver sección 4.4) tras su cálculo por *Mod_Ajuste*.

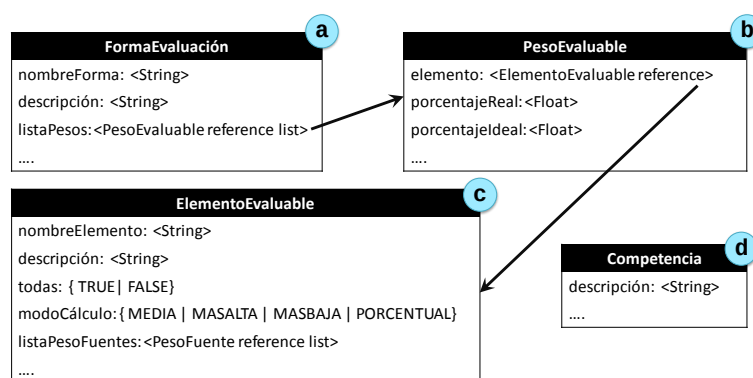


Fig. 4-11.- Definición de las clases FormaEvaluación, PesoEvaluable, ElementoEvaluable y Competencia

AdESMuS permite que cada `ElementoEvaluable` se evalúe de distintas maneras (clase `Fuente`, Fig. 4-12→b). La evaluación de cada elemento se puede realizar mediante rúbricas de evaluación, de manera manual o a través de un sistema externo. Además, la clase `ElementoEvaluable` incluye entre sus atributos el atributo *todas* que indica si es obligatorio evaluar ese elemento usando todas las `Fuentes` que se relacionan con él, o si basta con usar sólo una de ellas. En caso de que sea obligatorio usar todas las `Fuentes`, el atributo *modoCálculo* indica cómo se calcula la calificación del `ElementoEvaluable`. Este modo de cálculo puede consistir en calcular la media aritmética de la nota obtenida usando cada una de las `Fuentes` (opción *MEDIA*), en seleccionar la nota más alta (opción *MASALTA*), la más baja (opción *MASBAJA*), o en realizar un cálculo asignando un peso porcentual distinto a cada una de las `Fuentes` (opción *PORCENTUAL*).

Para permitir los distintos modos de cálculo de la nota, la relación entre un `ElementoEvaluable` y sus `Fuentes` se realiza a través del atributo *listaPesoFuentes* que referencia elementos de la clase `PesoFuente` (Fig. 4-12→a). Si el modo de cálculo de la nota del `ElementoEvaluable` es *PORCENTUAL*, en el atributo *porcentajeReal* de la clase `PesoFuente` se indica el porcentaje de influencia asignado por el docente y en el atributo *porcentajeIdeal*, el porcentaje óptimo actualizado calculado a través de *Mod_Ajuste*.

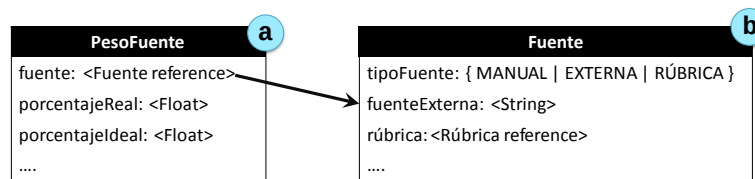


Fig. 4-12.- Definición de las clases `PesoFuente` y `Fuente`

Actualmente, están definidas tres formas de obtener la calificación recogidas en el atributo `tipoFuente` de la clase `Fuente`:

- **MANUAL**: la persona evaluadora introduce la nota directamente, sin que quede reflejado ningún tipo de evidencia de la evaluación.
- **EXTERNA**: la nota se obtiene a través de un LMS externo o algún sistema similar. En este caso, el atributo `fuenteExterna` indica de dónde proviene dicha evaluación.
- **RÚBRICA**: existe una rúbrica de evaluación asociada. El atributo `rúbrica` referencia a la rúbrica de evaluación correspondiente.

Ejemplo de instanciación para la definición general de la evaluación

En el ejemplo de evaluación presentado en la sección 4.5.1, se establecen dos formas de evaluación para el Trabajo Fin de Grado: la «tradicional» y la «continua».

La Fig. 4-13 muestra las instancias necesarias para definir la forma de evaluación «tradicional» (Ver Fig. 4-13→a). En esta `FormaEvaluación` se realiza una única valoración sobre el conjunto formado por un documento técnico y una presentación oral (Fig. 4-13→c). Por otro lado, se indica que el porcentaje de influencia de [elementoevaluable@1] es del 100% ([pesoevaluable@1], Fig. 4-13→b). Al incluir la referencia a [pesoevaluable@1] en el atributo `listaPesos` de [formaevaluacion@1] se establece la relación que representa que en la `FormaEvaluación` «tradicional», el conjunto del documento técnico y la presentación oral supone el 100% de la nota.

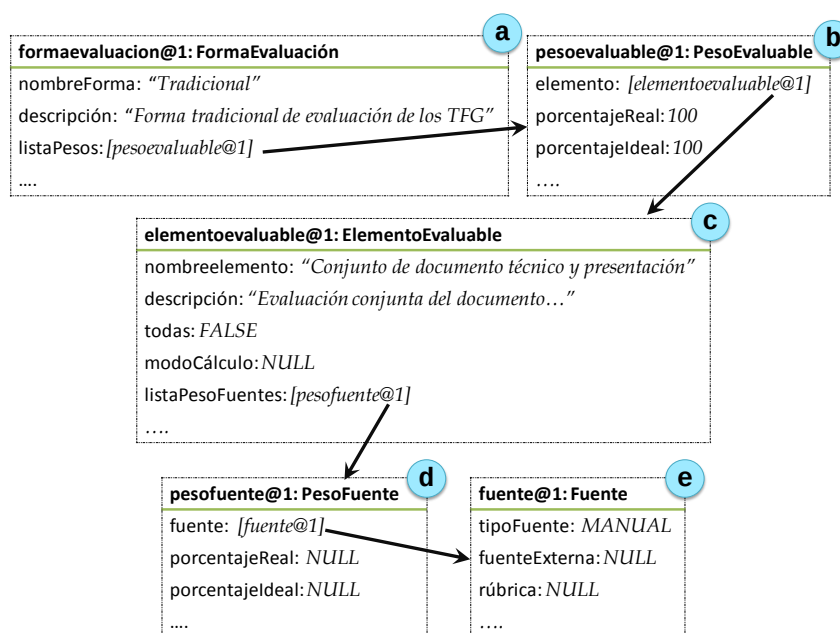


Fig. 4-13.- Ejemplo de objetos de las clases FormaEvaluación, PesoEvaluable, ElementoEvaluable, PesoFuente y Fuente

En la *FormaEvaluación* «tradicional» la evaluación se realiza sin almacenar ningún tipo de evidencia sobre la misma, por lo que se establece que el *tipoFuente* es *MANUAL* (Fig. 4-13→e). Esto es, las personas encargadas de evaluar ese *ElementoEvaluable*, introducirán directamente en AdESMuS la valoración numérica que consideren adecuada.

En el ejemplo concreto que se está tratando sólo hay una *Fuente* (Fig. 4-13→e) para el *ElementoEvaluable*, por lo que no es necesario establecer porcentaje de influencia en el objeto *PesoFuente* que lo referencia (Fig. 4-13→d). Para establecer la relación entre la *Fuente* de origen de la calificación y el *ElementoEvaluable*, se incluye el objeto `[pesofuente@1]` en el atributo `listaPesoFuentes` de `[elementoevaluable@1]`.

Para la *FormaEvaluación* «continua» es necesaria la definición de más objetos., tal y como se muestra en la Fig. 4-14.

Para dar soporte completo a la *FormaEvaluación* «continua» (Fig. 4-14→a) es necesaria la definición de un objeto de la clase *ElementoEvaluable* por cada uno de los seis elementos que se quieren valorar. En la Fig. 4-14→b se representan dos de los seis elementos que se valoran en esta forma de evaluación.

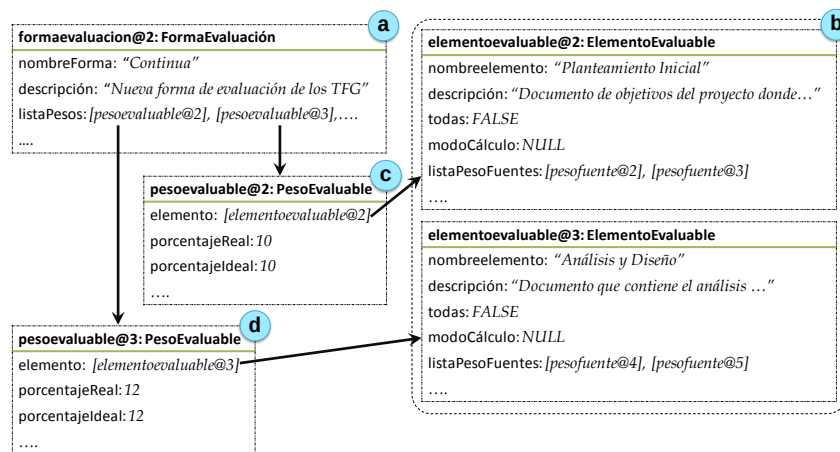


Fig. 4-14.- Ejemplo de objetos de las clases FormaEvaluación, ElementoEvaluable y PesoEvaluable

Cada uno de los `ElementoEvaluables` se relaciona con la `FormaEvaluación` a través de la clase `PesoEvaluable` donde se indica el porcentaje de influencia de ese `ElementoEvaluable` en la nota final calculada según esa `FormaEvaluación`. En la Fig. 4-14 se muestran los objetos `PesoEvaluable` correspondientes a los `ElementoEvaluables` «Planteamiento Inicial» (Fig. 4-14→c) con un peso del 10% en la nota final y «Análisis y Diseño» con un peso del 12% (Fig. 4-14→d).

En la `FormaEvaluación` «continua», cada uno de los `ElementoEvaluables` tiene dos posibles Fuentes de origen de la nota, y es la persona encargada de la evaluación quien decide qué Fuente usar en cada caso: una rúbrica de evaluación de tipo analítico o una de tipo atomístico. Para ello, los objetos que simbolizan a los `ElementoEvaluables` (Fig. 4-14→b) tienen en su atributo `listaPesoFuentes` dos referencias a objetos de tipo `PesoFuente`, uno por cada tipo de rúbrica y mediante el valor `FALSE` en su atributo `todas`, indican que sólo se debe usar una de las Fuentes.

En la Fig. 4-15 se observa que el objeto `ElementoEvaluable` (Fig. 4-15→a) se relaciona con dos objetos de tipo `PesoFuente` (Fig. 4-15→b); cada uno de los cuales hace referencia a una Fuente (Fig. 4-15→c), y en este caso, ambos objetos indican que representan rúbricas de evaluación (opción `RÚBRICA` en el atributo `tipoFuente`). Además, la opción `FALSE` en el atributo `todas` del objeto `ElementoEvaluable` indica que para obtener la calificación de ese elemento es suficiente con rellenar sólo una de las rúbricas.

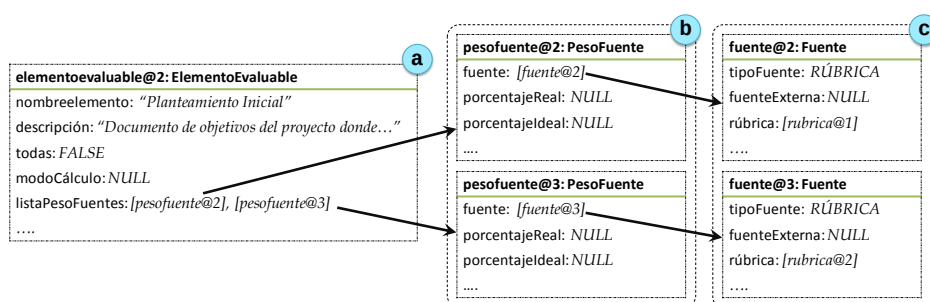


Fig. 4-15.- Ejemplo de objetos de las clases ElementoEvaluable, PesoFuente y Fuente

Clases para la definición de la evaluación mediante rúbricas

En aquellos casos en los que el proceso de evaluación se basa en rúbricas, la ontología de evaluación incluye un conjunto de clases que permiten representar la información relacionada con las rúbricas.

De manera general, una rúbrica de tipo analítico está formada por N dimensiones y M niveles de valoración (Ver sección 2.2.1 y Tabla 4-11). Las dimensiones reflejan cada uno de los aspectos del objeto de evaluación que se desean evaluar y los niveles de valoración conforman la escala de calificaciones que se usa para valorar las dimensiones. Además, cada dimensión tiene asignado un peso porcentual en la calificación final calculada con la rúbrica y cada nivel de valoración se corresponde con un valor numérico para poder obtener una calificación final numérica.

Tabla 4-11.- Estructura general de una rúbrica de evaluación

| | Nivel 1 10 | | Nivel M 0 |
|--|----------------------------|------|----------------------------|
| Dimensión 1 Peso D1 % | Descripción _{1,1} | | Descripción _{1,M} |
| | | | |
| Dimensión N Peso DN % | Descripción _{N,1} | ... | Descripción _{N,M} |

Para facilitar el uso de la rúbrica, las dimensiones se acompañan de una descripción para cada nivel de valoración, indicando los criterios que debe cumplir el objeto de evaluación para que en esa dimensión se le asigne dicho nivel de valoración.

Para representar las rúbricas de evaluación en AdESMuS, se define la clase rúbrica (Fig. 4-16→a) que permite referenciar las dimensiones de la rúbrica (atributo *listaDimensiones*) y sus distintos niveles de valoración (atributo *listaNiveles*).

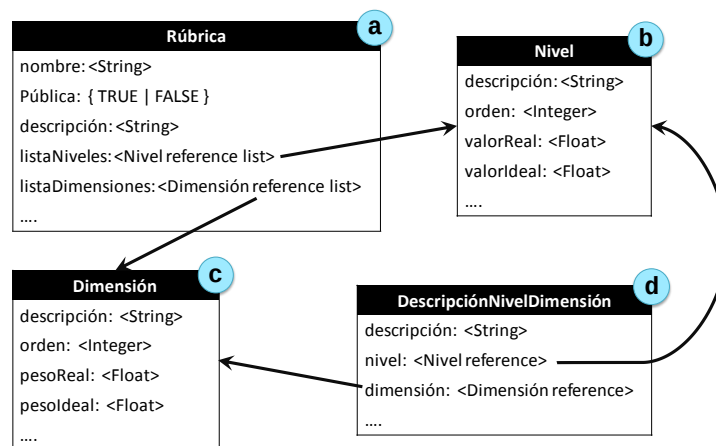


Fig. 4-16.- Definición de las clases Rúbrica, Nivel, Dimensión y DescripciónNivelDimensión

Para cada nivel (Fig. 4-16→b) se puede indicar la valoración otorgada a ese nivel de evaluación por la persona encargada de definir la rúbrica (atributo *valorReal*); y la valoración óptima calculada por Mod_Ajuste (atributo *valorIdeal*). También almacena la información necesaria para poder mostrar los niveles de evaluación ordenados en forma de escala (atributo *orden*).

Dimensión (Fig. 4-16→c) es la clase que representa las dimensiones de la rúbrica y establece el peso de la dimensión en la nota calculada por la rúbrica (atributo *pesoReal*) y el peso óptimo otorgado por el REM calculado por Mod_Ajuste a esa dimensión en función de los resultados obtenidos (atributo *pesoIdeal*). Al igual que para los niveles, también se indica el orden en el que deben mostrarse las dimensiones (atributo *orden*).

La clase DescripciónNivelDimensión (Fig. 4-16→d), representa el texto correspondiente a cada nivel de evaluación para cada dimensión (atributo *descripción*). Este texto es el que ayudará a la persona encargada de la evaluación a determinar qué nivel es el más adecuado para cada dimensión al valorar el trabajo de un estudiante.

Ejemplo de instanciación para la definición de la evaluación mediante rúbricas

Retomando el ejemplo de la FormaEvaluación «continua», el objeto Fuente (Fig. 4-15→c y Fig. 4-17→a) referencia al objeto rúbrica (Fig. 4-17→b) que representa la rúbrica analítica correspondiente al «Planteamiento Inicial» que se muestra parcialmente en la Tabla 4-12.

Tabla 4-12.- Extracto de la rúbrica analítica del «Planteamiento Inicial»

| | Excelente | Muy Adecuado | Adecuado | Básico | Inadecuado |
|----------------------------|--|---|---|--|--|
| Objetivos 30% | Explica de forma clara, concisa y concreta cada uno de los objetivos propuestos, lo que permite comprobar fácilmente si se han alcanzado. Prioriza la secuencia de su realización. | Explica de forma clara los objetivos propuestos, expresándolos en términos que admiten una sola interpretación. | Incluye una relación de objetivos, redactados de forma apropiada, aunque la claridad y/o concreción de por lo menos algunos de ellos no permite comprobar fácilmente si se han alcanzado. | Expresa los objetivos propuestos de forma excesivamente general y/o no secuenciados adecuadamente. | Explica de forma difusa y muy poco concreta, con lo que no se aprecia una relación ordenada de objetivos concretos a alcanzar en el TFG. |
| Antecedentes 30% | Indica claramente el punto de partida y realiza un estudio exhaustivo de soluciones similares y de las fuentes más relevantes para el TFG. | Indica claramente el punto de partida del TFG. Realiza comparativas con soluciones similares. Menciona fuentes relevantes relacionadas. | Indica de manera confusa el punto de partida del TFG. Realiza comparativas básicas con soluciones similares. Menciona sólo algunas de las fuentes relevantes. | Realiza comparativas con soluciones similares pero se queda en un nivel superficial. Las fuentes que menciona no son relevantes. | No realiza comparativas con soluciones similares. No menciona las fuentes bibliográficas relacionadas. |
| ... | | ... | ... | ... | ... |

Esta rúbrica posee 5 niveles de valoración (Excelente, Muy Adecuado, Adecuado, Básico e Inadecuado) y 5 dimensiones (Objetivos, Antecedentes, Planificación, Gestión de riesgos y Redacción/Presentación), cada una de ellas con un porcentaje de influencia distinto.

Para representar esta información en AdESMuS, el objeto rúbrica (ver Fig. 4-17→b) incluye las referencias a todos los niveles y dimensiones correspondientes a dicha rúbrica. Los objetos de tipo *Nivel* referenciados (atributo *listaNiveles*), representan cada uno de los niveles de la rúbrica e indican su valoración numérica para calcular la calificación (atributo *valorReal*). En la Fig. 4-17→c puede observarse la definición de uno de los cinco niveles de esta rúbrica.

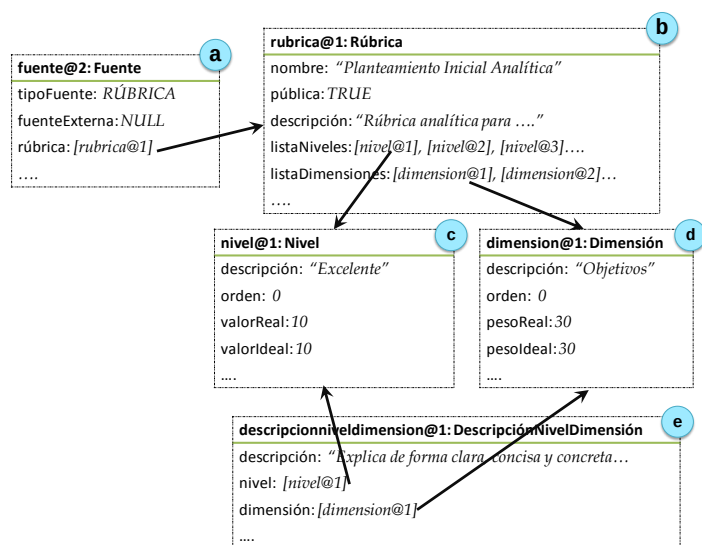


Fig. 4-17.- Ejemplo de objetos de las clases Rúbrica, Nivel, Dimensión y DescripciónNivelDimensión

Así mismo, para cada una de las 5 dimensiones de la rúbrica se define un objeto de tipo Dimensión, asignando a su atributo *pesoReal* el peso correspondiente a dicha dimensión en la nota calculada por la rúbrica (ver Fig. 4-17→d). Además, para cada una de las posibles combinaciones entre niveles y dimensiones, se define un objeto de tipo DescripciónNivelDimensión donde se establece el texto correspondiente a ese nivel y esa dimensión en la rúbrica. En la Fig. 4-17→e se muestra el objeto que hace referencia al nivel «Excelente» (Fig. 4-17→c) de la dimensión «Objetivos» (Fig. 4-17→d).

La rúbrica atomística correspondiente al «Planteamiento Inicial» (ver Tabla 4-13), tiene 3 niveles de valoración (A, B y C) y 18 dimensiones.

Tabla 4-13.- Extracto de la rúbrica atomística del «Planteamiento Inicial»

| | A | B | C |
|---|-----------|--------|-------------|
| Exposición de los objetivos del TFG (10%) | Muy clara | Normal | Confusa |
| Posibilidad de analizar la consecución de los objetivos propuestos (10%) | Alta | Baja | Inexistente |
| ... | | ... | ... |

El objeto que representa la rúbrica atomística correspondiente al «Planteamiento Inicial» (Fig. 4-18→a), agrupa los objetos correspondientes a los niveles y las dimensiones de esa rúbrica. Este objeto está a su vez referenciado por el objeto [fuente@3] (Fig. 4-18→e) para indicar que la rúbrica es una posible Fuente de evaluación del ElementoEvaluable «Planteamiento Inicial» (Fig. 4-15→a).

En este caso se definen tres objetos de la clase Nivel que representan cada uno de los niveles de valoración (ver Fig. 4-18→b) y los 18 objetos de la clase Dimensión correspondientes a las dimensiones de la rúbrica. En la Fig. 4-18→d se muestra uno de los 18 objetos de tipo Dimensión creados para dar soporte a la rúbrica atomística del «Planteamiento Inicial».

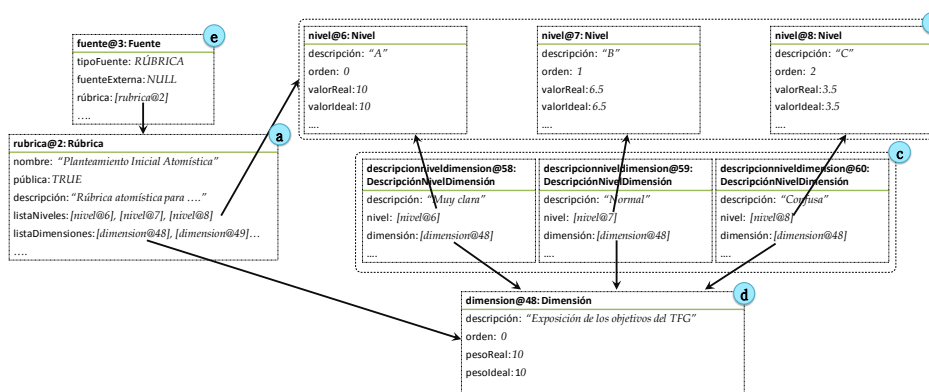


Fig. 4-18.- Ejemplo de objetos de las clases Rúbrica, Nivel, Dimensión y DescripciónNivelDimensión

Una vez definidas las dimensiones y los niveles de la rúbrica, es necesario definir un objeto DescripciónNivelDimensión por cada posible combinación entre niveles y dimensiones. En la Fig. 4-18→c se muestran los tres objetos creados para la dimensión «Exposición de los objetivos del TFG» (Fig. 4-18→d).

Mediante la definición de los objetos pertenecientes a las clases FormaEvaluación, ElementoEvaluable, PesoEvaluable, Fuente y PesoFuente se realiza la definición general del funcionamiento de la evaluación; se establecen los porcentajes de influencia de cada elemento; y se especifica la Fuente de la nota para cada uno de ellos. De manera similar, con los objetos rúbrica, Nivel, Dimensión y DescripciónNivelDimensión se definen las 12 rúbricas necesarias (6 analíticas y otras 6 atomísticas) y todos sus elementos. De este modo, AdESMuS soporta el comportamiento deseado en la FormaEvaluación «continua» permitiendo a las personas encargadas de evaluar el TFG decidir qué tipo de rúbrica usar a la hora de valorar cada uno de los ElementoEvaluable definidos.

Clases para la incorporación de la perspectiva de género

Por otro lado, en el sistema se ha incorporado la perspectiva de género para permitir y facilitar el uso de textos respetuosos con el género de las personas que lo utilicen. En el caso de AdESMuS, al redactar algunos de los elementos contenidos en la ontología Ont_Definición (tales como el texto descriptivo de las competencias o de las dimensiones de las rúbricas) es importante tener en cuenta la perspectiva de género y evitar el sexismo lingüístico en sus distintas variantes (Sánchez-Apellániz, 2009).

Especialmente habitual es el androcentrismo (Jiménez et al., 2011) que, a través del uso del término masculino como, por ejemplo «el alumno» o «el profesor», excluye a una gran parte del profesorado y del alumnado. Sin embargo, tal y como afirma Ignacio Bosque, miembro de la RAE en su informe sobre sexismo lingüístico y visibilidad de la mujer, *«hay acuerdo general entre los lingüistas en que el uso no marcado (o uso genérico) del masculino para designar los dos sexos está firmemente asentado en el sistema gramatical del español... y en que no hay razón para censurarlo.»* (Bosque, 2012), por lo que el uso de «los profesores» o «los alumnos» debe considerarse correcto y no sexista.

A pesar de esta aclaración de la RAE, en aquellas situaciones donde se desea evitar el androcentrismo, algunas recomendaciones (Dirección para la igualdad, 2010; Guerrero, 2013) proponen usar términos que engloben tanto a hombres como a mujeres, como «profesorado»; o usar los términos en su versión tanto masculina como femenina, «los profesores y las profesoras». Sin embargo, hay ocasiones en las que el uso de términos globales no es posible, como por ejemplo cuando se está haciendo referencia a una única persona o en las que el hecho de incluir explícitamente la versión femenina y masculina dificulta la lectura del texto.

AdESMuS incorpora la perspectiva de género a aquellos textos susceptibles de ser redactados usando lenguaje sexista. De este modo, cuando un usuario redacta un texto, puede indicar si la redacción es independiente del género o no; y en caso de que no lo sea, puede definir tantos textos como necesite para asegurar la incorporación de la perspectiva de género.

Por ejemplo, teniendo en cuenta la perspectiva de género, la descripción de uno de los niveles de valoración para la dimensión «Persona que sugería las mejoras al TFG» contenida en la rúbrica analítica correspondiente al «Seguimiento», se puede redactar en función del género de la persona encargada de la supervisión y del género del estudiante, tal y como se muestra en la Tabla 4-14.

Tabla 4-14.- Ejemplo de redacción incluyendo la perspectiva de género

| Género docente \ Género estudiante | Hombre | Mujer |
|---------------------------------------|----------------------------------|----------------------------------|
| Mujer | Entre la supervisora y el alumno | Entre la supervisora y la alumna |
| Hombre | Entre el supervisor y el alumno | Entre el supervisor y la alumna |

Sin tener en cuenta el género de las personas involucradas y realizando una redacción única, pero respetuosa con la perspectiva de género, el texto podría ser: «Entre la persona encargada de la supervisión y la persona encargada de la realización». Sin embargo, al atender al género tanto del docente como del estudiante, se consigue una redacción más sencilla y natural.

Para dar soporte a la incorporación de la perspectiva de género, la ontología Ont_Definición define la clase *Perspectiva* (ver Fig. 4-19) que especifica la redacción (atributo *texto*) correspondiente a un atributo (atributo *campo*) de una clase en función del género del docente y del estudiante (atributos *géneroDocente* y *géneroEstudiante*).

| Perspectiva |
|----------------------------|
| campo: <String> |
| texto: <String> |
| géneroDocente: <String> |
| géneroEstudiante: <String> |
| |

Fig. 4-19.- Definición de la clase *Perspectiva*

Las clases de la ontología incluyen dos atributos para poder relacionarlos con la clase *Perspectiva*. Por un lado, el atributo *independienteGénero* indica si existe una única redacción independiente del género de las personas implicadas. Por otro lado, el atributo *listaInclusiva* permite incluir los objetos *Perspectiva* necesarios para respetar la perspectiva de género.

Ejemplo de instanciación para la incorporación de la perspectiva de género

En la Fig. 4-20→a puede observarse un objeto de la clase *DescripciónNivelDimensión* donde la redacción de su campo *descripción* es independiente del género.

| descripcionniveldimension@58: DescripciónNivelDimensión a | descripcionniveldimension@127: DescripciónNivelDimensión b |
|--|---|
| descripción: "Muy clara" nivel: [nivel@6] dimensión: [dimension@48] independienteGénero: TRUE listaInclusiva: NULL | descripción: NULL nivel: [nivel@29] dimensión: [dimension@137] independienteGénero: FALSE listaInclusiva: [perspectiva@17], [perspectiva@18], |

Fig. 4-20.- Ejemplo de objetos de la clase DescripciónNivelDimensión incluyendo la perspectiva de género

Sin embargo, el objeto representado en la Fig. 4-20→b representa el ejemplo expuesto en la Tabla 4-14, por lo que en este caso el atributo *independienteGénero* tiene el valor *FALSE*; y *listaInclusiva*, contiene los objetos *Perspectiva* necesarios para tratar todas las combinaciones de género que afectan al atributo *descripción* del objeto (Fig. 4-21).

| | |
|---|--|
| perspectiva@17: Perspectiva campo: "descripción" texto: "Entre la supervisora y el alumno" géneroDocente: "Mujer" géneroEstudiante: "Hombre" | perspectiva@18: Perspectiva campo: "descripción" texto: "Entre la supervisora y la alumna" géneroDocente: "Mujer" géneroEstudiante: "Mujer" |
| perspectiva@19: Perspectiva campo: "descripción" texto: "Entre el supervisor y el alumno" géneroDocente: "Hombre" géneroEstudiante: "Hombre" | perspectiva@20: Perspectiva campo: "descripción" texto: "Entre el supervisor y la alumna" géneroDocente: "Hombre" géneroEstudiante: "Mujer" |

Fig. 4-21.- Ejemplo de objetos de la clase Perspectiva

4.5.4 Ont_Evaluación: Realización de la evaluación

La estructura necesaria para integrar la información relativa a la definición de la evaluación, las personas involucradas en ella y los resultados de la misma, viene definida por la ontología Ont_Evaluación.

Ont_Evaluación, cuyos componentes se muestran en verde en la Fig. 4-22, establece la estructura de información necesaria para realizar la evaluación y recopilar sus resultados. Para ello, define una serie de clases que sirven de unión entre las clases *usuario* y *Estudiante* de Ont_Universidad (representadas en amarillo

en la Fig. 4-22), y las clases Fuente, FormaEvaluación, ElementoEvaluable y DescripciónNivelDimensión de Ont_Definición (coloreadas en azul en la Fig. 4-22).

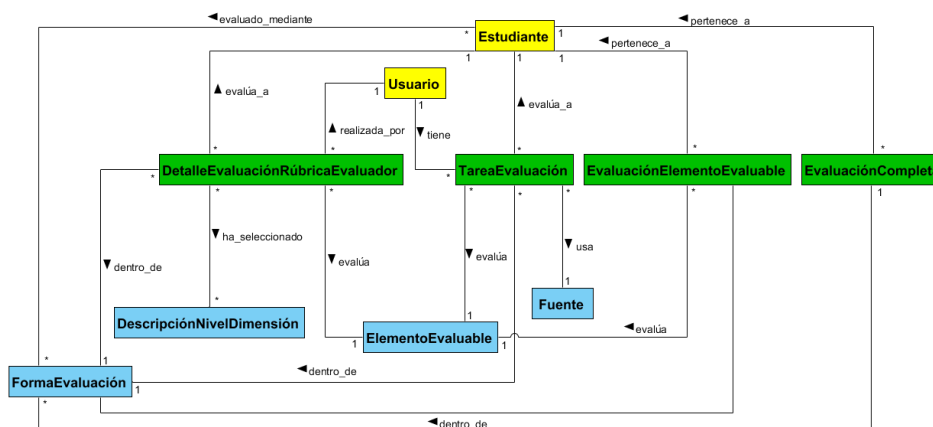


Fig. 4-22.- Representación UML de la ontología Ont_Evaluación (en verde)

Ont_Evaluación permite definir quién es la persona encargada (docente o estudiante) de evaluar cada ElementoEvaluable de cada estudiante (clase TareaEvaluación, Fig. 4-23); también permite especificar qué Fuente debe usarse para realizar la evaluación y en qué FormaEvaluación.

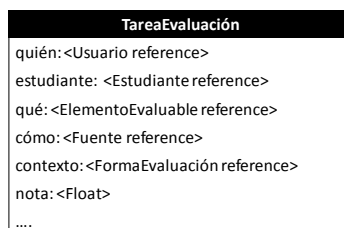


Fig. 4-23.- Definición de la clase TareaEvaluación

Cada objeto de tipo TareaEvaluación representa la evaluación de un ElementoEvaluable (atributo *qué*), para un estudiante (atributo *estudiante*), realizada por una persona (atributo *quién*), usando una Fuente concreta (atributo *cómo*), en una FormaEvaluación determinada (atributo *contexto*) y recoge el resultado de la misma en su atributo *nota*.

Sin embargo, dependiendo del proceso de evaluación definido, ese resultado puede ser parcial porque sea necesario realizar la evaluación del ElementoEvaluable usando más de una Fuente o porque pueda haber más personas implicadas en la

evaluación del `ElementoEvaluable` para el estudiante. En ese caso, será necesario definir un objeto de tipo `TareaEvaluación` para cada una de las Fuentes a usar y/o para cada una de las personas encargadas de evaluar al estudiante.

Los resultados parciales de la evaluación de cada `ElementoEvaluable` se recopilan en objetos de tipo `EvaluaciónElementoEvaluable` (Fig. 4-24→a) donde se recoge la nota resultante (atributo *notaCalculada*) de evaluar un `ElementoEvaluable` concreto, independientemente de las Fuentes usadas o de las personas evaluadoras.

La clase `EvaluaciónCompleta` (Fig. 4-24→b), recoge la nota calculada (atributo *notaCalculada*) para un estudiante teniendo en cuenta la nota obtenida en todos los `ElementoEvaluables` definidos y sus porcentajes de influencia en la nota final.

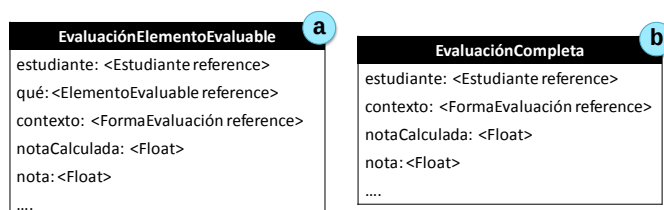


Fig. 4-24.- Definición de las clases `EvaluaciónElementoEvaluable` y `EvaluaciónCompleta`

Se permite también que el docente encargado de la asignatura asigne a un estudiante en un `ElementoEvaluable` concreto o en una `FormaEvaluación` concreta, una nota distinta a la calculada de manera automática por AdESMuS (atributo *nota*). De esta forma, el docente puede modificar la nota asignada teniendo en cuenta criterios adicionales no representados en AdESMuS o no cuantificables como puede ser el grado de participación en las clases presenciales.

Para las evaluaciones cuya Fuente de origen sea una rúbrica de evaluación, además de la nota calculada, se debe recoger qué nivel de valoración ha seleccionado la persona evaluadora para cada dimensión de la rúbrica. Para ello, se define la clase `DetalleEvaluaciónRúbricaEvaluador` (Fig. 4-25), que identifica tanto a la persona que realiza la evaluación (atributo *quién*), como al estudiante que se está evaluando (atributo *estudiante*), el elemento que está siendo evaluado (atributo *qué*) y la `FormaEvaluación` de la evaluación (atributo *contexto*). También se recogen los objetos de tipo `DescripciónNivelDimensión` seleccionados por la persona evaluadora al realizar la evaluación (atributo *listaElegidos*).

| DetalleEvaluaciónRúbricaEvaluador |
|---|
| quién: <Usuario reference> |
| estudiante: <Estudiante reference> |
| qué: <ElementoEvaluable reference> |
| contexto: <FormaEvaluación reference> |
| listaElegidos: <DescripciónNivelDimensión reference list> |
| |

Fig. 4-25.- Definición de la clase DetalleEvaluaciónRúbricaEvaluador

Ejemplo de instanciación para la realización de la evaluación

En el ejemplo que se está ilustrando en este capítulo, se han definido dos instancias de la clase FormaEvaluación para representar la evaluación «tradicional» (objeto [formaevaluacion@1], Fig. 4-13→a) y la evaluación «continua» (objeto [formaevaluacion@2], Fig. 4-14→a).

En la primera, un tribunal formado por tres docentes debe evaluar un único ElementoEvaluable (objeto [elementoevaluable@1], Fig. 4-13→c), utilizando la Fuente establecida para dicho elemento (objeto [fuente@1], Fig. 4-13→e). En la Fig. 4-26 se muestran los tres objetos de la clase TareaEvaluación, uno para cada componente del tribunal, que permiten formalizar este comportamiento.

| tareaevaluacion@1: TareaEvaluación | tareaevaluacion@2: TareaEvaluación | tareaevaluacion@3: TareaEvaluación |
|------------------------------------|------------------------------------|------------------------------------|
| quién: [usuario@27] | quién: [usuario@8] | quién: [usuario@54] |
| estudiante: [estudiante@3] | estudiante: [estudiante@3] | estudiante: [estudiante@3] |
| qué: [elementoevaluable@1] | qué: [elementoevaluable@1] | qué: [elementoevaluable@1] |
| cómo: [fuente@1] | cómo: [fuente@1] | cómo: [fuente@1] |
| contexto: [formaevaluacion@1] | contexto: [formaevaluacion@1] | contexto: [formaevaluacion@1] |
| nota: NULL | nota: NULL | nota: NULL |
| | | |

Fig. 4-26.- Ejemplo de objetos de la clase TareaEvaluación (I)

En la segunda FormaEvaluación (objeto [formaevaluacion@2], Fig. 4-14→a), se definen seis ElementoEvaluables que se deben evaluar utilizando rúbricas, permitiendo que cada persona evaluadora elija el tipo de rúbrica que desea usar para cada ElementoEvaluable.

Los cuatro primeros ElementoEvaluables los debe evaluar únicamente la persona supervisora, que podrá utilizar la Fuente que desee, eligiéndola entre las dos disponibles para cada elemento. Para realizar esta evaluación, se han generado dos objetos de tipo TareaEvaluación para cada uno de los ElementoEvaluables, uno por cada posible Fuente.

En la Fig. 4-27 se muestran los objetos TareaEvaluación definidos para que la persona que supervisa el TFG, pueda realizar la evaluación del «Planteamiento Inicial» ([elementoevaluable@2], Fig. 4-14→b), del estudiante, usando la rúbrica analítica ([fuente@2], Fig. 4-17→a) o la atomística ([fuente@3], Fig. 4-18→e).

| tareaevaluacion@9: TareaEvaluación | tareaevaluacion@10: TareaEvaluación |
|------------------------------------|-------------------------------------|
| quién: [usuario@47] | quién: [usuario@47] |
| estudiante: [estudiante@3] | estudiante: [estudiante@3] |
| qué: [elementoevaluable@2] | qué: [elementoevaluable@2] |
| cómo: [fuente@2] | cómo: [fuente@3] |
| contexto: [formaevaluacion@2] | contexto: [formaevaluacion@2] |
| nota: NULL | nota: NULL |
| | |

Fig. 4-27.- Ejemplo de objetos de la clase TareaEvaluación (II)

Los dos ElementoEvaluables restantes de la FormaEvaluación «continua», deben ser evaluados por las tres personas que componen el tribunal de evaluación. Para realizar esta evaluación, cada componente del tribunal puede usar el tipo de rúbrica que desee, por lo que es necesario generar dos objetos de tipo TareaEvaluación por cada ElementoEvaluable que tenga que evaluar cada persona.

En la Fig. 4-28 pueden observarse los objetos de tipo TareaEvaluación que permiten a cada componente del tribunal elegir qué Fuente usar ([fuente@10] ó [fuente@11]) para evaluar la memoria del TFG (objeto [elementoevaluable@6]).

| | | |
|---|---|---|
| tareaevaluacion@25: TareaEvaluación quién: [usuario@47] estudiante: [estudiante@3] qué: [elementoevaluable@6] cómo: [fuente@10] contexto: [formaevaluacion@2] nota: NULL | tareaevaluacion@26: TareaEvaluación quién: [usuario@48] estudiante: [estudiante@3] qué: [elementoevaluable@6] cómo: [fuente@10] contexto: [formaevaluacion@2] nota: NULL | tareaevaluacion@27: TareaEvaluación quién: [usuario@49] estudiante: [estudiante@3] qué: [elementoevaluable@6] cómo: [fuente@10] contexto: [formaevaluacion@2] nota: NULL |
| tareaevaluacion@28: TareaEvaluación quién: [usuario@47] estudiante: [estudiante@3] qué: [elementoevaluable@6] cómo: [fuente@11] contexto: [formaevaluacion@2] nota: NULL | tareaevaluacion@29: TareaEvaluación quién: [usuario@48] estudiante: [estudiante@3] qué: [elementoevaluable@6] cómo: [fuente@11] contexto: [formaevaluacion@2] nota: NULL | tareaevaluacion@30: TareaEvaluación quién: [usuario@49] estudiante: [estudiante@3] qué: [elementoevaluable@6] cómo: [fuente@11] contexto: [formaevaluacion@2] nota: NULL |

Fig. 4-28.- Ejemplo de objetos de la clase TareaEvaluación (III)

A medida que las personas encargadas de la evaluación introducen su valoración en AdESMuS, se actualiza el atributo *nota* de los objetos *TareaEvaluación* correspondientes. Cuando todas las personas evaluadoras han terminado la evaluación de un *ElementoEvaluable*, se genera el objeto de tipo *EvaluaciónElementoEvaluable* correspondiente al estudiante que se está evaluando (Fig. 4-29→a). Al finalizar la evaluación de todos los elementos evaluables de un estudiante en una *FormaEvaluación*, se genera el objeto de la clase *EvaluaciónCompleta* (Fig. 4-29→b) donde se calcula la nota final del estudiante en esa *FormaEvaluación*.

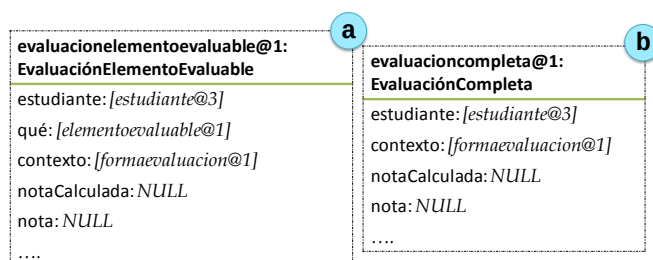


Fig. 4-29.- Ejemplo de objetos de las clases *EvaluaciónElementoEvaluable* y *EvaluaciónCompleta*

En el caso de que la evaluación se realice mediante rúbricas, cada vez que una persona evaluadora selecciona el nivel de valoración para una dimensión, se genera un objeto de tipo *DetalleEvaluaciónRúbricaEvaluador* (ver Fig. 4-30) indicando qué objetos de tipo *DescripciónNivelDimensión* se han elegido (atributo *listaElegidos*).

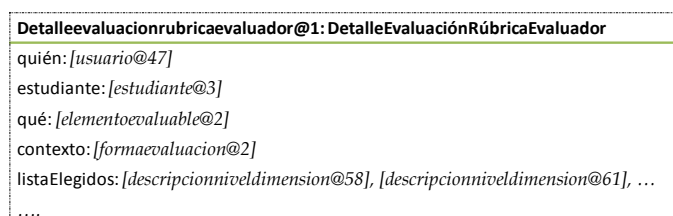


Fig. 4-30.- Ejemplo de objeto de la clase *DetalleEvaluaciónRúbricaEvaluador*

4.6 Internacionalización

AdESMuS también recoge la información necesaria para ser un sistema multilingüe y permitir su utilización en el idioma deseado por cada persona.

Para lograrlo, no precisa que se le proporcionen las traducciones de todos los atributos incluidos en las ontologías. Existen clases cuyos atributos de tipo texto no requieren internacionalización (por ejemplo el atributo *nombre* de la clase *usuario*) o que hacen referencia a elementos independientes del idioma (por ejemplo los atributos de la clase *Grupo*).

Para dar soporte a la internacionalización de AdESMuS, se ha definido la clase *Internacionalización* (Ver Fig. 4-31) cuyos atributos permiten almacenar la traducción (atributo *traducción*) a un idioma (atributo *idioma*) de un atributo concreto (atributo *campo*) de cualquiera de las clases que contengan atributos susceptibles de ser internacionalizados.

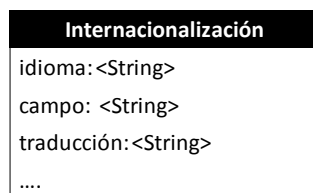


Fig. 4-31.- Definición de la clase *Internacionalización*

Cada una de las clases que poseen atributos susceptibles de ser internacionalizados se ha enriquecido con un atributo *idioma* donde se indica el idioma por defecto en el que se encuentra el contenido de los atributos; y otro atributo *listaTraducciones*, donde se referencian los objetos *Internacionalización* correspondientes a los atributos de esa clase. Por ejemplo, si para un centro concreto se desea almacenar su nombre en tres idiomas diferentes, se debe definir un objeto *Centro* y dos instancias de *Internacionalización*. En el ejemplo de la Fig. 4-32, se muestra el objeto correspondiente a la EUI de Vitoria-Gasteiz con su contenido en castellano (Fig. 4-32→a) y las referencias a los objetos *Internacionalización* que contienen la traducción al euskera y al inglés respectivamente del atributo *nombreCentro* (Fig. 4-32→b).

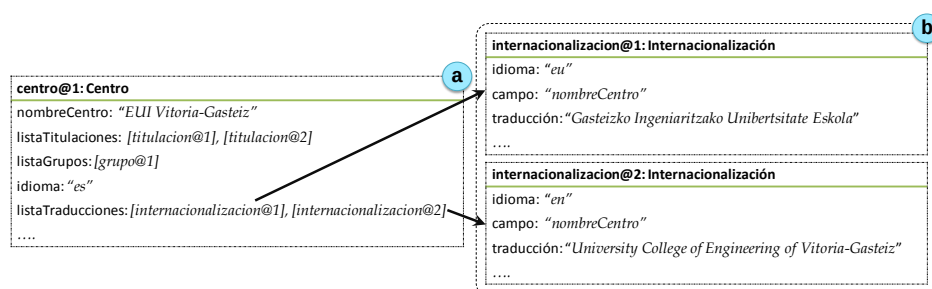


Fig. 4-32.- Ejemplo de internacionalización

Los identificadores utilizados para representar los distintos idiomas siguen la norma ISO-639 (ISO, 2002), por lo que «es» representa el castellano, «eu» el euskera y «en» el inglés.

El diseño elegido para incluir el aspecto de internacionalización en AdESMuS posibilita la inclusión de nuevos idiomas y traducciones en cualquier momento sin necesidad de modificar el diseño del sistema.

4.7 Una implementación concreta

Con el objetivo de facilitar su independencia y reutilización, AdESMuS se ha implementado en forma de API tipo REST (*Representational State Transfer*) (Fielding, 2000) y se ha desplegado en un servidor web. Este tipo de implementación convierte a AdESMuS en un proveedor de recursos sin interfaz gráfica asociada, por lo que ésta deberá ser proporcionada por los sistemas que consuman sus recursos (los clientes).

Un recurso pone a disposición del cliente una serie de funcionalidades que le permiten obtener o modificar información de las bases de conocimiento. El acceso a un recurso se realiza mediante el intercambio de mensajes a través del protocolo HTTP y los verbos GET, POST, PUT y DELETE para indicar al sistema qué se desea hacer con el recurso. La Tabla 4-15 recoge algunos ejemplos de mensajes sobre el recurso estudiantes de AdESMuS y el resultado que generan.

Tabla 4-15.- Ejemplos de mensajes para consumir recursos de AdESMuS

| Mensaje | Resultado |
|---|---|
| GET http://adesmus.com/estudiantes | Obtiene el listado de todos los estudiantes registrados |
| GET http://adesmus.com/estudiantes/123 | Obtiene los datos del estudiante con identificador 123 |
| POST http://adesmus.com/estudiantes | Añade un nuevo estudiante en el sistema. Los datos del estudiante irán en el cuerpo del mensaje |
| PUT http://adesmus.com/estudiantes/123 | Actualiza los datos del estudiante con identificador 123. Los nuevos datos irán en el cuerpo del mensaje |
| DELETE http://adesmus.com/estudiantes | Elimina todos los estudiantes del sistema |
| DELETE http://adesmus.com/estudiantes/123 | Elimina del sistema el estudiante con identificador 123 |

Siguiendo los principios REST, cada mensaje HTTP enviado desde el cliente es totalmente independiente de los anteriores, por lo que contiene los datos precisos para que AdESMuS pueda procesarlos y devolver, en formato JSON, la información correspondiente. Para que AdESMuS conozca quién le está solicitando la información y si tiene los permisos necesarios para consumir el servicio solicitado, se han utilizado tokens siguiendo el protocolo de autenticación OAuth2.

AdESMuS genera los tokens en el momento que las personas usuarias se identifican y los envía a la aplicación (el cliente) que la persona usuaria esté usando para acceder a los servicios de AdESMuS. La aplicación cliente es la encargada de gestionar los tokens y enviarlos en el cuerpo de cada petición que realice. La información cifrada en cada token contiene entre otros elementos: la fecha de expiración del token, el código identificador de la persona que se ha identificado en el sistema y el rol de usuario que le corresponde.

También se ha incorporado en la implementación de AdESMuS un control de acceso a sus funcionalidades basado en roles que permiten determinar los permisos de cada tipo de usuario para realizar acciones en la API. Cada una de las funcionalidades implementadas lleva asociada la lista de roles de usuario que pueden ejecutarla. De este modo, antes de ejecutar la operación solicitada, se comprueba si el rol de la persona que ha realizado la petición tiene permiso para ejecutar dicha operación o no. Este mecanismo de control permite reutilizar directamente el código de las funcionalidades compartidas por varios roles de usuario como, por ejemplo, la consulta de criterios de evaluación. Asimismo,

también facilita la escalabilidad futura del sistema en el caso de que haya que modificar los roles de usuario con permiso para ejecutar una funcionalidad concreta.

Para la implementación de AdESMuS se ha utilizado la infraestructura de aplicaciones (*framework*) Express³⁴. Express es un framework para el desarrollo de aplicaciones basadas en el entorno de ejecución de JavaScript Node.js³⁵. Por su parte, las bases de conocimiento de AdESMuS se han implementado usando el sistema de gestión de bases de datos MySQL.

Para realizar las pruebas y evaluación de AdESMuS se ha usado HTML, JavaScript y CSS para implementar un sistema web independiente que ejerza de cliente y consuma los recursos ofertados por aquél (ver Fig. 4-33). Este sistema ha permitido a profesorado y alumnado usar AdESMuS y evaluar su funcionamiento.



Fig. 4-33.- Interfaces del sistema web

4.8 Conclusiones

En este capítulo se ha descrito AdESMuS, un sistema que permite definir y gestionar la evaluación de una asignatura usando múltiples fuentes para calcular la calificación final de los estudiantes.

AdESMuS permite al profesorado definir los distintos elementos que se desean evaluar en una asignatura y el peso que cada elemento debe tener en la nota final. Asimismo, también recoge quién o quiénes (tanto docentes como estudiantes) son

³⁴ <http://expressjs.com>

³⁵ <https://nodejs.org>

las personas encargadas de efectuar la evaluación de cada elemento y los medios a través de los cuales debe realizarse (las fuentes). Actualmente, el sistema admite que la persona encargada de la evaluación introduzca la calificación directamente en AdESMuS; que se importe desde un sistema externo como Moodle; o que se calcule a través de rúbricas de evaluación.

Cada elemento a evaluar se puede calificar usando una o varias fuentes. Cuando existen varias fuentes de calificación para un mismo elemento, el profesorado es el encargado de definir en el sistema cómo se combinan las distintas calificaciones para obtener la nota final del elemento.

AdESMuS ofrece las funcionalidades necesarias para la definición y el uso de rúbricas como herramienta de evaluación y se ha diseñado de modo que tenga en cuenta los aspectos de internacionalización y de perspectiva de género a la hora de definir las rúbricas.

AdESMuS posee una arquitectura modular y dos bases de conocimiento con una estructura definida a través de ontologías. Esta estructura modular facilita la escalabilidad y el mantenimiento futuro del sistema, además de dotarlo de independencia, haciendo posible su uso en combinación con otras herramientas educativas.

Entre los módulos de AdESMuS, cabe destacar el módulo de ajuste (`Mod_Ajuste`) que calcula cuál es el modelo matemático más adecuado tanto para el proceso de evaluación de la asignatura en general (TAM) como para las rúbricas de evaluación involucradas en el mismo (REM).

El TAM y el REM calculados por `Mod_Ajuste` permiten conocer la influencia concreta que están teniendo en la evaluación las tareas definidas y los criterios contemplados en las rúbricas. Este módulo permite, además, identificar tareas y criterios de evaluación que realmente no estén influyendo en la calificación final o que no lo estén haciendo en la medida que se deseaba.

El conocimiento de esta información favorece la toma de las medidas correctoras que mejoren el proceso de aprendizaje y evaluación de la asignatura, beneficiando tanto al profesorado como al alumnado.

Por ejemplo, si se identifican tareas o criterios que no estén influyendo en la calificación, el proceso de evaluación se puede simplificar eliminando dichos criterios. Sin embargo, puede ocurrir que el profesorado no esté de acuerdo con los modelos obtenidos por `Mod_Ajuste`, lo que debería llevarle a reflexionar sobre su forma de calificar.

Los modelos obtenidos por `Mod_Ajuste` dependen directamente de los datos a partir de los cuales se han calculado, por lo que un conjunto distinto de datos puede

originar unos modelos diferentes. Por esta razón, los modelos obtenidos no deben utilizarse como modelos ideales, sino que deben ser usados como herramienta de apoyo y reflexión en la definición del proceso de evaluación.

Por ejemplo, si el TAM indica que alguna de las partes de la evaluación no tiene una influencia real en el cálculo de la calificación, el profesorado deberá reflexionar si realmente desea eliminar dicha parte del proceso de evaluación o si, por el contrario, lo que debe modificar es su forma de calificar para darle mayor importancia.

AdESMuS se ha diseñado e implementado como un sistema autónomo que puede comunicarse con otros sistemas y permite gestionar la evaluación de cualquier asignatura. Estas características hacen posible que AdESMuS pueda utilizarse desde TraMeA para evaluar los TFG de acuerdo a los criterios y procesos definidos en la metodología FPM.

5

Definición de un conjunto de rúbricas

La metodología FPM (sección 2.4) establece que el proceso de evaluación de los TFG se base en un conjunto de elementos evaluables, que serán evaluados mediante rúbricas de evaluación. Estas rúbricas deben ser rellenas por la persona encargada de la supervisión del TFG y por las personas que conformen el tribunal de evaluación del mismo.

Este capítulo detalla el proceso seguido para definir un conjunto de rúbricas para la evaluación de los TFG tal y como establece la metodología FPM. A continuación presenta la evaluación realizada sobre el conjunto de rúbricas obtenido. Finalmente, se describe el proceso de ajuste de los pesos de las dimensiones y de la influencia de cada rúbrica en la nota final realizado para que el conjunto de rúbricas tenga el comportamiento esperado por el profesorado.

5.1 Introducción

En general, la definición de una rúbrica de evaluación implica determinar sus dimensiones, escala, descripciones y peso de cada dimensión (ver sección 2.2.1). De acuerdo a la metodología FPM, en la evaluación del TFG influyen varios elementos evaluables, por lo que en este caso también se debe decidir la influencia de cada elemento evaluable (el peso que tiene cada rúbrica en el cálculo de la nota final del TFG).

Para definir un conjunto de rúbricas que permitan evaluar los TFG de manera eficaz es necesario contar con la colaboración del profesorado involucrado en la supervisión y evaluación del TFG y aplicar su experiencia en ambos procesos.

Para ello, con la ayuda de un grupo de expertos formado por docentes con más de 10 años de experiencia en la supervisión y evaluación de proyectos, se definieron las versiones iniciales de las rúbricas. En base a su criterio experto, se asignaron los pesos correspondientes a cada dimensión para el cálculo de la nota de la rúbrica (el REM, ver sección 4.4) y la influencia que cada rúbrica debe tener en la obtención de la nota final del TFG (el TAM del proceso, ver sección 4.4).

Debido a sus características y limitaciones, se decidió desechar las rúbricas de tipo holístico y centrarse en las de tipo analítico (ver sección 2.2.1). Dentro de las rúbricas de tipo analítico se decidió realizar una versión analítica y otra atomística de la rúbrica para cada uno de los elementos evaluables definidos por la metodología FPM y así poder analizar y evaluar sus posibles beneficios y desventajas en la evaluación.

A la hora de decidir las dimensiones a incluir en las rúbricas de evaluación, se intentó obtener un conjunto consensuado de criterios que tanto profesorado como alumnado consideraran adecuados para evaluar los TFG. Por ello, una vez que el profesorado acordó una versión preliminar del conjunto de rúbricas, se preguntó al alumnado su opinión sobre las dimensiones presentes en dichas rúbricas.

Tras obtener la conformidad de profesorado y alumnado sobre los aspectos de evaluación contemplados en las rúbricas, se recogió la opinión de ambos tipos de usuario sobre la adecuación de las rúbricas analíticas y atomísticas a la evaluación de los TFG. Por último, se realizó el proceso de ajuste tanto de los pesos de cada dimensión de las rúbricas como de la influencia de cada rúbrica en la nota final del TFG.

A continuación se detallan el proceso de obtención de la versión preliminar del conjunto de rúbricas por parte del grupo de expertos y los resultados de la evaluación de las dimensiones por parte del alumnado. Seguidamente, se analiza la opinión tanto del alumnado como del profesorado sobre el tipo de rúbrica más adecuado para evaluar los TFG. Finalmente se describe el proceso de ajuste de rúbricas llevado a cabo.

5.2 Obtención de una versión preliminar del conjunto de rúbricas

Para definir el conjunto de rúbricas para evaluar los TFG, el grupo de expertos ejecutó en cuatro ocasiones el proceso de la Fig. 5-1 hasta obtener una versión preliminar del conjunto de rúbricas de evaluación.

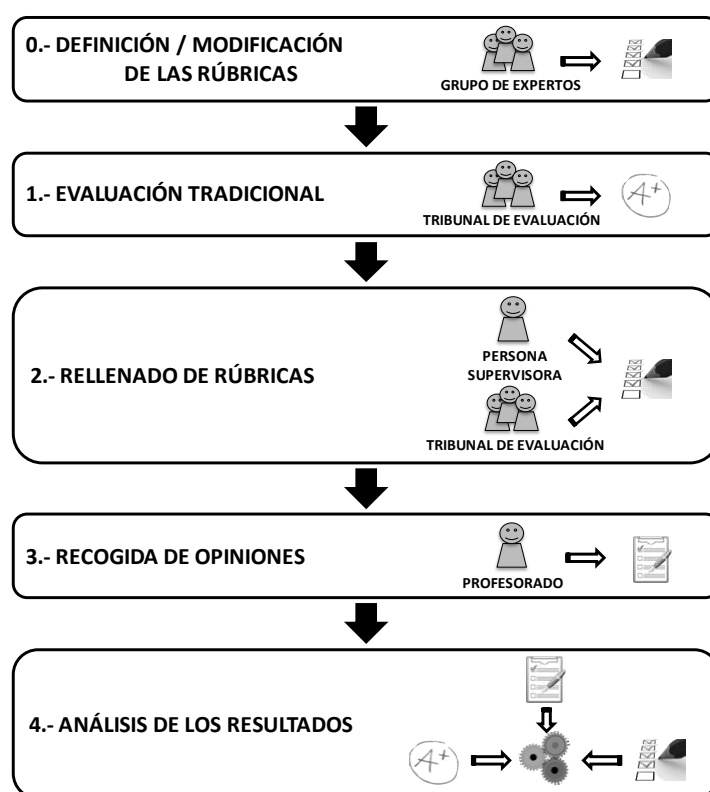


Fig. 5-1.- Proceso de definición de las rúbricas

En la primera ejecución del paso 0, los expertos definieron las versiones iniciales de las rúbricas: dimensiones junto a su peso, escala de valoración y descripciones. Además, determinaron también la influencia de cada rúbrica en la nota final del TFG.

A continuación, el tribunal de evaluación evaluó los trabajos de una convocatoria de evaluación siguiendo el método tradicional de acordar una nota entre todos los componentes del tribunal (paso 1) y después, tanto el tribunal como la persona responsable de la supervisión del TFG, rellenaron las rúbricas correspondientes (paso 2).

Una vez finalizada la evaluación de los TFG, se recogió la opinión de todo el profesorado involucrado en el uso de las rúbricas (paso 3) y se analizaron sus opiniones y los resultados tanto de la evaluación tradicional como de las notas calculadas utilizando las rúbricas (paso 4).

El proceso de la Fig. 5-1 se repitió en cuatro ocasiones para adecuar la definición y las calificaciones calculadas por las rúbricas teniendo en cuenta los resultados de los análisis realizados. En la Fig. 5-2 se muestran los datos correspondientes a cada una de las iteraciones realizadas, detallando el número de docentes implicados y de TFG evaluados en dicha iteración.

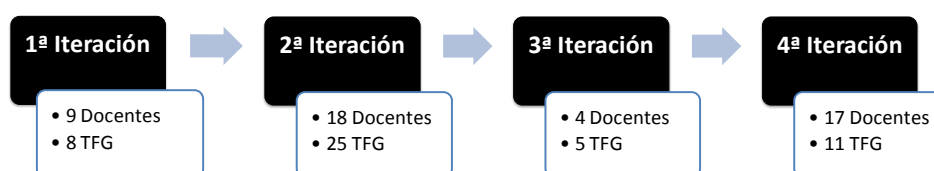


Fig. 5-2.- Iteraciones del proceso de definición de rúbricas

Como resultado del análisis de las opiniones del profesorado, en las distintas iteraciones del proceso fue necesario modificar aspectos de las rúbricas tales como la escala de valoración, algunas dimensiones y sus descripciones. Asimismo, en cada iteración se analizaron las calificaciones obtenidas mediante la evaluación tradicional y se compararon con las obtenidas utilizando las rúbricas. En base a dicho análisis, los expertos ajustaron los pesos de las dimensiones y la influencia de cada rúbrica en la nota final del TFG. Este proceso se repitió hasta que se obtuvieron unos resultados aceptables de acuerdo al grupo de expertos cuya opinión fue usada a modo de estándar de calidad.

El resultado final de este proceso fue el conjunto de rúbricas que puede consultarse en el ANEXO D con los pesos de las dimensiones y la influencia de cada rúbrica en la nota final que se muestra en la Tabla 5-1. En esta tabla pueden verse los pesos establecidos finalmente por el grupo de expertos para cada dimensión en la versión analítica de las rúbricas. Como se ha mencionado en la sección 2.2.1, cada dimensión de la versión analítica se divide en distintos criterios que dan lugar a las dimensiones de la versión atomística. Siguiendo este razonamiento, el peso asignado a cada dimensión de la versión analítica se distribuyó de manera equitativa entre todos los criterios en los que se divide dicha dimensión para la versión atomística.

Tabla 5-1.- Valores para el conjunto preliminar de rúbricas

| Rúbrica / Elemento evaluable | Influencia | Dimensión | Peso |
|------------------------------|------------|--------------------------------------|------|
| Eva_Requisitos | 10% | Objetivos | 30% |
| | | Antecedentes | 30% |
| | | Planificación | 15% |
| | | Gestión de riesgos | 15% |
| | | Redacción y presentación | 10% |
| Eva_Diseño | 15% | Captura de requisitos | 40% |
| | | Análisis y diseño | 50% |
| | | Redacción y presentación | 10% |
| Eva_Actitud | 5% | Autonomía del estudiante | 35% |
| | | Receptividad ante críticas y cambios | 25% |
| | | Interés por mejorar | 25% |
| | | Constancia | 15% |
| Eva_Memoria | 45% | Desarrollo | 30% |
| | | Calidad global y coherencia | 15% |
| | | Evaluación | 20% |
| | | Conclusiones y trabajo futuro | 15% |
| | | Bibliografía | 10% |
| | | Redacción y presentación | 10% |
| Eva_Presentación | 15% | Contenido y organización | 40% |
| | | Habilidades comunicativas | 15% |
| | | Preguntas | 35% |
| | | Uso del tiempo | 10% |
| Eva_Complejidad | 10% | Complejidad técnica del TFG | 100% |

5.3 Evaluación de las dimensiones de las rúbricas por parte del alumnado

La adecuación de las dimensiones de las rúbricas definidas por el profesorado fue sometida a la opinión del alumnado. Se solicitó a los estudiantes que indicaran, por

cada una de las dimensiones presentes en las rúbricas, si consideraban adecuado o no que formara parte de los criterios de evaluación del TFG. La encuesta se realizó sobre la versión atomística de las rúbricas, ya que combinando las dimensiones presentes de manera explícita en las rúbricas atomísticas se forman las descripciones correspondientes a las versiones analíticas.

Un total de 40 estudiantes egresados respondieron a la encuesta en la que además de responder Si/No para indicar si cada una de las dimensiones debía ser tomada en cuenta o no, disponían de un apartado para incluir cualquier comentario o sugerencia que desearan.

En la Tabla 5-2 se muestran las dimensiones que más votos negativos recibieron por parte del alumnado, señalando a qué rúbrica pertenecen y el tanto por ciento de votos negativos recibidos. De las 72 dimensiones sometidas a valoración, sólo 2 superaron el 50% de votos negativos (resaltadas en la Tabla 5-2). En opinión de los estudiantes, «la cantidad de referencias bibliográficas» y «la calidad y posibilidad de contraste de las referencias bibliográficas» no debería tenerse en cuenta a la hora de evaluar la documentación del TFG.

De las 70 dimensiones restantes, a pesar de obtener más votos positivos que negativos, hubo 10 dimensiones cuyos votos negativos superaron el 30% (ver Tabla 5-2). Por lo tanto, de las 70 dimensiones valoradas positivamente por el alumnado, 60 recibieron una valoración positiva por parte de más del 70% del alumnado encuestado.

En cuanto a la opción que tenía el alumnado de expresar su opinión o añadir cualquier comentario o sugerencia, los participantes no sugirieron ningún aspecto de valoración que no estuviera previamente contemplado en las dimensiones sometidas a evaluación.

Los resultados de la encuesta sobre la adecuación de las dimensiones realizada entre el alumnado demuestran que también consideran adecuados los criterios establecidos por el grupo de expertos. Los criterios con menor aceptación por parte del alumnado son los referentes a aspectos formales del trabajo y los únicos que despiertan una opinión negativa son los referentes a formalismos en cuanto a referencias bibliográficas. Dicha opinión negativa puede estar influenciada por la falta de experiencia del alumnado en la elaboración de documentos técnicos formales. Sin embargo, el grupo de expertos consideró que dichos formalismos son necesarios y que para motivar al alumnado a prestarles la atención adecuada, deben formar parte de la evaluación. Por lo tanto, se decidió mantener esas dimensiones en las rúbricas.

Tabla 5-2.- Dimensiones peor valoradas por el alumnado

| Rúbrica / Elemento evaluable | Dimensión | % Votos negativos |
|------------------------------|---|-------------------|
| Eva_Actitud | La relevancia de los temas tratados en las reuniones | 32 |
| | La existencia de terceras personas perjudicadas por el ritmo de trabajo | 36 |
| Eva_Memoria | La cantidad de referencias bibliográficas | 61 |
| | La calidad de las referencias y la posibilidad de contraste | 75 |
| | La cita a lo largo de la documentación de las referencias | 32 |
| | La mención a las fuentes bibliográficas o soluciones similares más relevantes | 39 |
| | El razonamiento realizado sobre la priorización de los riesgos | 32 |
| | La calidad del plan de seguimiento de los riesgos | 36 |
| Eva_Presentación | Tiempo dedicado a cada apartado | 47 |
| | El ajuste de la defensa al tiempo asignado | 32 |
| | El tono de voz usado | 39 |
| Eva_Complejidad | El número de asignaturas cuyos conocimientos están involucrados en la realización del TFG | 46 |

5.4 Evaluación de la adecuación de cada tipo de rúbrica

Una vez obtenido un conjunto de rúbricas cuyas dimensiones habían sido sometidas a la valoración tanto del profesorado como del alumnado, se realizó una evaluación para determinar la versión de rúbrica más adecuada para evaluar los TFG (analítica o atomística). En esta nueva evaluación se volvió a contar con la participación de todos los actores involucrados: estudiantes y docentes.

Para determinar qué versión de la rúbrica es más adecuada desde el punto de vista del profesorado, se solicitó a 19 docentes que rellenasen una encuesta (ver

extracto en la Tabla 5-3) para recoger su opinión tras su primera experiencia usando ambas versiones de las rúbricas para evaluar un TFG.

Tabla 5-3.- Extracto de la encuesta sobre versiones de rúbricas (profesorado)

| Pregunta | | Respuestas posibles |
|-----------|--|-------------------------------------|
| P1 | ¿Con qué tipo de rúbrica prefieres trabajar? | Analítica / Atomística / Indistinta |
| P2 | ¿Qué tipo de rúbrica consideras que evalúa a los estudiantes de manera más precisa? | Analítica / Atomística / Indistinta |
| P3 | ¿Qué tipo de rúbrica consideras que proporciona una mejor retroalimentación a los estudiantes? | Analítica / Atomística / Indistinta |
| P4 | Enumera las desventajas de cada tipo de rúbrica | Libre |
| P5 | Enumera las ventajas de cada tipo de rúbrica | Libre |
| P6 | Otros comentarios y/o sugerencias | Libre |

En cuanto a la versión de rúbrica que le resultaba más cómoda para trabajar (P1 en Tabla 5-3), los resultados obtenidos están muy igualados (ver Fig. 5-3→a): un 45% de los docentes indicaron su preferencia por trabajar con las rúbricas analíticas, mientras que a otro 45% prefirieron las atomísticas. El otro 10% de los docentes opinó que no tenía preferencia por ninguno de los dos tipos de rúbrica.

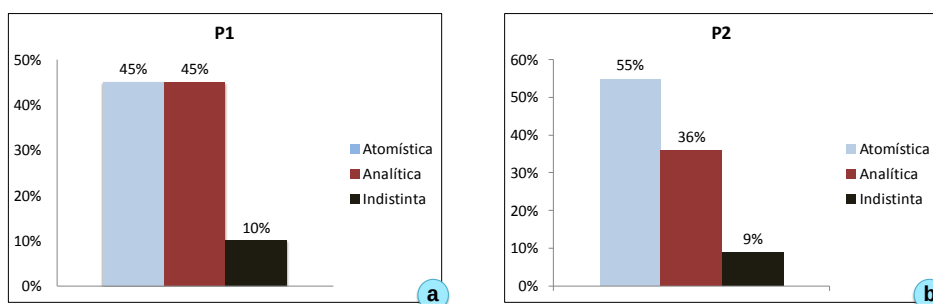


Fig. 5-3.- Resultados de las preguntas P1 (a) y P2 (b) de la encuesta sobre rúbricas (profesorado)

Sin embargo, a la pregunta sobre qué tipo de rúbrica consideran más útil para la evaluación (P2), los resultados fueron más desiguales (ver Fig. 5-3→b), el 55% se decantó por la versión atomística frente al 36% que prefirió la versión analítica. El 9% restante opinó que la utilidad de ambas versiones de rúbrica era similar.

En cuanto a la opinión sobre la calidad de la retroalimentación proporcionada por cada tipo de rúbrica (P3), el 46% opinó que las rúbricas de tipo atomístico proporcionaban una mejor retroalimentación. Un 27% indicó que la retroalimentación de las rúbricas de tipo analítico era más adecuada y el otro 27% consideró que la retroalimentación proporcionada por ambos tipos de rúbrica era similar.

En el apartado de otros comentarios o sugerencias, no se recogieron aportaciones significativas.

Desde la perspectiva del alumnado se realizó otra encuesta a 40 estudiantes que habían finalizado y presentado su Trabajo Fin de Grado y que, por lo tanto, podían valorar el proceso de evaluación desde la experiencia personal. La encuesta, cuyo extracto se muestra en la Tabla 5-4, versaba sobre su opinión acerca de las versiones analítica y atomística de las rúbricas.

Tabla 5-4.- Extracto de la encuesta sobre rúbricas (alumnado)

| Pregunta | | Respuestas posibles |
|-----------------|--|--|
| E1 | ¿En qué tipo de rúbrica consideras que quedan más claros los criterios de evaluación? | Analítica / Atomística / Indistinta |
| E2 | ¿Qué tipo de rúbrica consideras que proporciona una mejor retroalimentación a los estudiantes? | Analítica / Atomística / Indistinta |
| E3 | Otros comentarios y/o sugerencias | Libre |

Sobre la diferencia en la claridad de los criterios en las rúbricas de tipo analítico y las de tipo atomístico (E1 en Tabla 5-4), un 50% de los estudiantes opina que la rúbrica atomística explicita mejor los criterios de evaluación, frente a un 33% que considera que es más clara la rúbrica analítica y un 17% que opina que ambos tipos son igual de claros (ver Fig. 5-4→a). Asimismo, a la pregunta sobre la utilidad de las rúbricas como retroalimentación para los estudiantes (E2), el 59% se decantan por la rúbrica atomística, frente al 31% que lo hacen por la analítica y el 10% que no se decanta por ninguno de los dos tipos de rúbrica (ver Fig. 5-4→b).

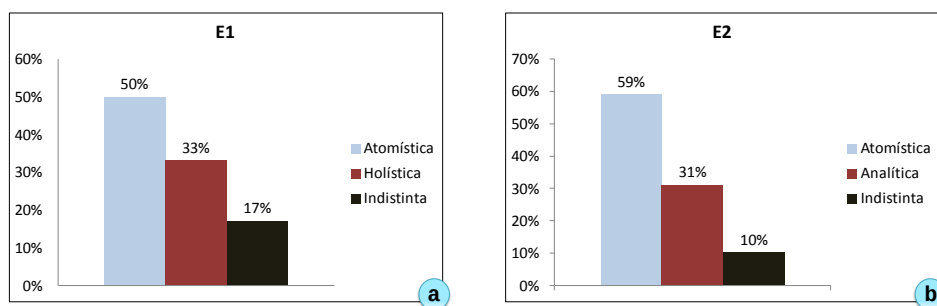


Fig. 5-4.- Resultados de las preguntas E1 (a) y E2 (b) de la encuesta sobre rúbricas (alumnado)

Entre los comentarios recogidos en la encuesta destaca la preocupación de 3 estudiantes por el hecho de que el conocimiento de las rúbricas de evaluación concretas que se van a usar para valorar el TFG, conlleve una obsesión por cumplir con cada una de las dimensiones presentes en las mismas. El propio alumnado sugiere que en vez de publicar las rúbricas completas se publiquen únicamente las dimensiones que se van a tener en cuenta, sin explicitar qué se exige para alcanzar cada nivel de valoración.

Aunque la mayoría del profesorado considera que la versión atomística de las rúbricas es más precisa y proporciona una mejor retroalimentación al alumnado, casi la mitad del profesorado encuestado prefiere trabajar con la versión analítica. Esto puede estar ocasionado por la menor cantidad de dimensiones a evaluar en la versión analítica, lo que hace que su relleno sea más rápido.

Entre el alumnado, la preferencia por la versión atomística es más evidente. La consideran más clara a la hora de conocer los criterios de evaluación y opinan que la retroalimentación proporcionada por este tipo de rúbrica es mejor que la aportada por la versión analítica.

5.5 Ajuste del conjunto de rúbricas

Una vez obtenido el conjunto de rúbricas para evaluar los TFG y tras su evaluación por parte del profesorado y del alumnado, se usaron en la evaluación de los TFG en las distintas convocatorias oficiales de defensa que fueron teniendo lugar.

Los datos recogidos en AdESMuS correspondientes a dichas evaluaciones se emplearon para calcular el MSEM (*Multi-Source Evaluation Model*) que se describe en la sección 4.4. El MSEM está compuesto por el TAM (*Task Adjusted Model*) y el REM (*Rubric Element Model*) que establecen los pesos de cada elemento evaluable en la

nota final y los pesos e influencias de cada componente de las rúbricas respectivamente. El TAM y el REM obtenidos por `Mod_Ajuste` se compararon con el TAM y el REM establecidos por el grupo de expertos (ver Tabla 5-1) para su evaluación.

A continuación se detalla el conjunto de datos utilizado por `Mod_Ajuste` para calcular el MSEM correspondiente al sistema de evaluación de los TFG usando el conjunto de rúbricas definidas. También se describe la obtención del TAM y del REM y se analizan los resultados obtenidos.

5.5.1 Conjunto de datos

A lo largo de las convocatorias oficiales de evaluación de TFG que tuvieron lugar una vez establecido el conjunto de rúbricas, se evaluaron un total de 77 TFG. Para evaluar esos TFG y recopilar todos los datos necesarios para ejecutar los procesos de ajuste se usó el sistema AdESMuS. Con el fin de evitar desviaciones, se eliminaron 9 trabajos del conjunto de datos porque su evaluación generó importantes discrepancias entre los integrantes del tribunal y no se consiguió la unanimidad sobre la calificación que debían obtener.

En la evaluación de los 68 TFG restantes, participaron 24 docentes de 3 centros distintos. 18 de los docentes ejercieron el rol de persona supervisora en alguno de los trabajos defendidos, los 6 docentes restantes participaron sólo como integrantes del tribunal. Los docentes participantes evaluaron los TFG a través de los elementos evaluables definidos por la metodología FPM (ver sección 2.4). Se usó AdESMuS tanto para indicar la nota de cada elemento en base a la opinión personal de los docentes, como para rellenar el conjunto de rúbricas en su versión analítica y en su versión atomística.

Los docentes también indicaron en el sistema la nota final que recibió el TFG. Dicha nota final se emitió usando la opinión personal de los componentes del tribunal, ya que la propuesta de evaluación presentada en este trabajo estaba en proceso de validación.

Para el cálculo del TAM y del REM, el conjunto de 68 TFG fue dividido de manera aleatoria y siguiendo una proporción 70 / 30, en un conjunto de datos de entrenamiento (46 TFG) y un conjunto de datos de validación (22 TFG). Los modelos TAM y REM se calcularon utilizando el conjunto de datos de entrenamiento; el conjunto de datos de validación se usó posteriormente para evaluar los modelos obtenidos.

5.5.2 Obtención del TAM

Utilizando el conjunto de datos de entrenamiento, se usó el módulo de ajuste de AdESMuS para calcular el TAM del proceso de evaluación de los TFG siguiendo la metodología FPM (Villamañe et al., 2015).

Modelo calculado

A través del proceso del cálculo del TAM (sección 4.4) el módulo de ajuste proporcionó un modelo indicando la influencia ideal de cada una de las partes de la evaluación, esto es, de cada uno de los elementos evaluables definidos en la metodología. En la Tabla 5-5 se recogen los valores obtenidos por *Mod_Ajuste* (TAM *Mod_Ajuste*) junto con los valores establecidos inicialmente por los expertos en base a su opinión personal (TAM experto).

Tabla 5-5.- Comparación del TAM calculado por *Mod_Ajuste* y el TAM establecido por los expertos

| | TAM experto | TAM <i>Mod_Ajuste</i> |
|-------------------------|--------------------|------------------------------|
| Eva_Requisitos | 10% | 19% |
| Eva_Diseño | 15% | 12% |
| Eva_Actitud | 5% | 22% |
| Eva_Memoria | 45% | 23% |
| Eva_Presentación | 15% | 2% |
| Eva_Complejidad | 10% | 22% |

Como se puede observar, los valores obtenidos por el módulo de ajuste varían considerablemente respecto a los definidos inicialmente por los expertos. La influencia del primer elemento evaluable (*Eva_Requisitos*) crece en nueve puntos, pasando del 10% al 19%, lo que indica que a la hora de evaluar un TFG, a la planificación y captura de requisitos se le da más importancia de la previamente considerada por el grupo de expertos. La influencia del análisis y diseño (*Eva_Diseño*) disminuye 3 puntos, pero se acerca bastante a la establecida por los expertos.

En base al modelo obtenido, la actitud del estudiante durante la realización del TFG (*Eva_Actitud*) y la complejidad del mismo (*Eva_Complejidad*) también se valoran más de lo previamente establecido. La influencia de la actitud se fijó en un 5%, pero

el modelo indica que debería ser un 22%. En el caso de la complejidad debería pasar del 10% al 22%.

En el extremo opuesto está el elemento evaluable correspondiente a la documentación final (*Eva_Memoria*), cuya influencia debería reducirse hasta el 23% desde el 45% que establecieron inicialmente los expertos.

Mención especial merece *Eva_Presentación*, para el cual los expertos habían establecido una influencia del 15%, pero que el modelo obtenido por *Mod_Ajuste* reduce al 2%. Esto es, en base a los datos de las evaluaciones realizadas, todos aquellos aspectos considerados en *Eva_Presentación* apenas influyen a la hora de asignar una calificación al TFG.

A continuación se detalla la evaluación que se realizó del TAM obtenido por el módulo de ajuste.

Evaluación del modelo

Para evaluar el TAM obtenido por *Mod_Ajuste* y comparar sus resultados con el modelo establecido por el grupo de expertos, se usaron los 22 TFG del conjunto de validación.

Partiendo de las notas asignadas por los expertos a cada uno de los elementos evaluables de la metodología FPM, se aplicaron los porcentajes de influencia indicados por cada uno de los modelos (el de los expertos y el calculado por *Mod_Ajuste*) para obtener la nota final para cada TFG. Las notas finales obtenidas por ambos modelos se compararon con las notas reales asignadas por el tribunal de evaluación. Las notas otorgadas por el tribunal fueron usadas como estándar de calidad debido a la amplia experiencia en la evaluación de proyectos de sus componentes.

Para poder comparar las notas de los TFG en base a distintos modelos es necesario establecer el margen de error admisible. Normalmente, a la hora de asignar la nota a un TFG, el tribunal tiende a redondearla al intervalo de medio punto más cercano a la calificación que considera adecuada. Por ello, es habitual encontrar notas de 7 ó de 7.5, pero no de 7.3. Teniendo en cuenta este factor, se consideró admisible un error de hasta 0.5 puntos.

Para realizar la validación de los modelos se usaron dos métricas distintas: el coeficiente de correlación de Pearson (Taylor, 1990) para analizar la relación entre la nota real y la obtenida por el modelo; y la Raíz del Error Cuadrático Medio (*Root-Mean-Squared Error*, RMSE) para cuantificar la diferencia entre ambas notas. La Tabla 5-6 recoge los resultados obtenidos.

Tabla 5-6.- Resultados validación modelos TAM

| | Correlación | RMSE |
|----------------|-------------|------|
| TAM experto | 0.93 | 0.53 |
| TAM Mod_Ajuste | 0.96 | 0.35 |

El modelo establecido por los expertos presenta una gran correlación con la nota real de los TFG (0.93), sin embargo también cuenta con un error de 0.53 en una escala de 10 puntos, lo que no está en el margen de error admisible.

El modelo obtenido por el módulo de ajuste de AdESMuS mejora ambos resultados, ya que presenta una mayor correlación y un error de sólo 0.35 puntos, valor que se encuentra dentro del margen de error considerado admisible.

Conclusiones

El módulo de ajuste de AdESMuS ha calculado un Modelo de Ajuste de Tareas (TAM) que varía significativamente del establecido inicialmente por el grupo de expertos.

No sólo varía considerablemente en el porcentaje de influencia de algunos de los elementos evaluables de la metodología FPM, si no que indica que uno de los mismos (*Eva_Presentación*) apenas es significativo a la hora de calificar un TFG.

La validación realizada sobre los modelos demuestra que el modelo calculado por *Mod_Ajuste* es mejor que el establecido por los expertos, presentando una mayor correlación con la nota real de los trabajos y un error de sólo 0.35 puntos, lo que entra dentro del margen de error considerado admisible (0.5 puntos).

5.5.3 Obtención del REM

Mod_Ajuste también calcula el Modelo de Elementos de las Rúbricas (REM) para las tareas o elementos que se evalúan mediante rúbricas. El REM establece la mejor combinación posible entre el peso que debería tener cada dimensión de las rúbricas en el cálculo de la calificación numérica y el valor numérico más adecuado para cada nivel de valoración (Modelo de niveles de la Escala, SM).

En el caso de la evaluación de los TFG siguiendo la metodología FPM, todos los elementos se evalúan a través de rúbricas de evaluación, por lo que en el cálculo del REM se tuvieron en cuenta todos ellos (Villamañe, Álvarez, et al., 2016b).

Modelo calculado

Aunque durante la evaluación de los TFG se recogieron datos tanto de la versión analítica como de la versión atomística de las rúbricas, el cálculo del REM se realizó únicamente con la versión analítica de las mismas. De este modo, si el REM establece un peso de cero para alguna dimensión de la versión analítica, se pueden desechar de una vez varias dimensiones de la versión atomística.

El REM del proceso de evaluación está compuesto por el Modelo de niveles de la Escala (SM) y el peso correspondiente a cada una de las dimensiones. Tal y como establece el proceso de obtención del REM (ver Fig. 4-4), en el primer paso del proceso los expertos adjudicaron a cada nivel de valoración de las rúbricas un rango de posibles valores numéricos (ver Tabla 5-7).

Tabla 5-7.- Posible asignación de valores numéricos

| Nivel de valoración de la rúbrica | Rango de valores |
|-----------------------------------|------------------|
| Excelente | 10 |
| Avanzado | [9, 9.99] |
| Medio | [7, 8.99] |
| Básico | [5, 6.99] |
| Insuficiente | [0, 4.99] |

A partir de los rangos de posibles valores *Mod_Ajuste* generó 200.000 SM distintos (ver sección 4.4) y seleccionó el que presentaba un mejor comportamiento. En la Tabla 5-8 se recogen los valores numéricos del SM seleccionado.

Tabla 5-8.- Valores numéricos del SM seleccionado

| | Excelente | Avanzado | Medio | Básico | Insuficiente |
|-----------------------|-----------|----------|-------|--------|--------------|
| Valores seleccionados | 10 | 9 | 7 | 5 | 0 |

Los valores numéricos del SM se corresponden con el límite inferior del rango indicado por los expertos para cada uno de los niveles de valoración de la rúbrica.

En la Tabla 5-9 se recogen los pesos establecidos por *Mod_Ajuste* para las dimensiones, junto con los indicados inicialmente por los expertos.

Tabla 5-9.- Comparación de los pesos de las dimensiones obtenidos por Mod_Ajuste y los establecidos por los expertos

| Rúbrica / Elemento evaluable | Dimensión | Peso expertos | Peso Mod_Ajuste |
|-------------------------------------|--------------------------------------|----------------------|------------------------|
| Eva_Requisitos | Objetivos | 30% | 51% |
| | Antecedentes | 30% | 26% |
| | Planificación | 15% | 0% |
| | Gestión de riesgos | 15% | 23% |
| | Redacción y presentación | 10% | 0% |
| Eva_Diseño | Captura de requisitos | 40% | 47% |
| | Análisis y diseño | 50% | 48% |
| | Redacción y presentación | 10% | 5% |
| Eva_Actitud | Autonomía del estudiante | 35% | 21% |
| | Receptividad ante críticas y cambios | 25% | 0% |
| | Interés por mejorar | 25% | 33% |
| | Constancia | 15% | 46% |
| Eva_Memoria | Desarrollo | 30% | 56% |
| | Calidad global y coherencia | 15% | 0% |
| | Evaluación | 20% | 29% |
| | Conclusiones y trabajo futuro | 15% | 0% |
| | Bibliografía | 10% | 7% |
| | Redacción y presentación | 10% | 8% |
| Eva_Presentación | Contenido y organización | 40% | 0% |
| | Habilidades comunicativas | 15% | 73% |
| | Preguntas | 35% | 3% |
| | Uso del tiempo | 10% | 24% |
| Eva_Complejidad | Complejidad técnica del TFG | 100% | 100% |

Los resultados obtenidos por Mod_Ajuste varían considerablemente de los pesos establecidos por los expertos para las distintas rúbricas.

Según el REM calculado, dos de las dimensiones establecidas en *Eva_Requisitos* («Planificación» y «Redacción y presentación»), no son significativas para el cálculo de la calificación de dicha rúbrica. Sin embargo, coincide con los expertos en que los

aspectos más relevantes de dicha rúbrica son los correspondientes a la dimensión «Objetivos», aunque su influencia varía considerablemente. La influencia de la dimensión «Antecedentes» se reduce en cuatro puntos, mientras que la de la dimensión «Gestión de riesgos» se incrementa pasando de un 15% a un 23%.

Para *Eva_Diseño*, el REM calculado es bastante similar al establecido por los expertos, con pequeñas variaciones en la influencia de las dimensiones, pero respetando el orden de importancia entre ellas. La dimensión «Análisis y diseño» es la más influyente (48%), seguida de la dimensión «Captura de requisitos» (47%) y finalizando con la dimensión «Redacción y presentación» cuya influencia es únicamente del 5%.

Según el modelo obtenido para el elemento evaluable correspondiente a la actitud del estudiante durante la elaboración del trabajo (*Eva_Actitud*), la dimensión «Receptividad ante críticas y cambios» no tiene influencia significativa sobre la nota calculada por la rúbrica, por lo que debería eliminarse de la misma. La influencia del resto de dimensiones en el modelo calculado por *Mod_Ajuste* varía considerablemente respecto a las establecidas por los expertos. La influencia de la dimensión «Autonomía del estudiante» es catorce puntos inferior a la establecida por los expertos, mientras que la influencia de las dimensiones «Interés por mejorar» y «Constancia» se incrementa en ocho y treinta y un puntos respectivamente.

En el elemento correspondiente a la documentación final del TFG (*Eva_Memoria*), el modelo obtenido por *Mod_Ajuste* elimina dos de las dimensiones presentes en la rúbrica. El REM establece que la «Calidad global y la coherencia», y las «Conclusiones y el trabajo futuro» no son relevantes para el cálculo de la calificación. En el resto de dimensiones, el cambio más destacable es el producido en la dimensión «Desarrollo», cuya influencia pasa del 30% establecido por los expertos al 56% establecido en el REM.

En el caso de la defensa oral (*Eva_Presentación*), el REM indica que la dimensión «Contenido y organización» no tiene una influencia real sobre la calificación. Además, la influencia de la dimensión «Preguntas» es únicamente del 3%. Las dimensiones realmente influyentes para el cálculo de la nota son las «Habilidades comunicativas», con una influencia del 73%; y el «Uso del tiempo», cuya influencia según el modelo obtenido es del 24%.

La complejidad técnica (*Eva_Complejidad*) se evalúa usando una única dimensión, por lo que el peso establecido por los expertos, el 100%, se mantiene en el REM calculado.

Evaluación del modelo

El REM obtenido representa la mejor combinación posible entre los pesos de las dimensiones de las rúbricas implicadas y los valores numéricos para los niveles de valoración, en función del TAM utilizado en el proceso de cálculo.

Debido a esta dependencia entre el REM y el TAM, para evaluar el REM calculado por *Mod_Ajuste* es necesario usar de nuevo el TAM y por lo tanto, evaluar el Modelo de Evaluación Multi-Fuente (MSEM) en su conjunto.

La evaluación del REM se realizó comparando la calificación obtenida por el MSEM (TAM y REM) establecido por *Mod_Ajuste* con la calificación otorgada realmente al TFG por los componentes del tribunal en base a su opinión personal. Asimismo, los resultados obtenidos por el MSEM establecido por *Mod_Ajuste* se compararon también con los resultados calculados en base a los modelos (TAM y REM) establecidos por los expertos (MSEM experto).

La Tabla 5-10 muestra la Raíz del Error Cuadrático Medio (RMSE) obtenido para cada uno de los modelos evaluados.

Tabla 5-10.- Resultados validación MSEM

| | RMSE |
|------------------------|-------------|
| MSEM experto | 0.61 |
| MSEM Mod_Ajuste | 0.57 |

Como se puede observar en la tabla, el modelo calculado por *Mod_Ajuste* reduce el error hasta los 0.57 puntos, acercándose al límite de accesibilidad.

Conclusiones

El REM calculado por *Mod_Ajuste* indica que varias de las dimensiones identificadas por los expertos no son significativas en el cálculo de la calificación numérica de la rúbrica de la que forman parte.

El modelo también indica que los valores numéricos más adecuados para el SM de las rúbricas, son los límites inferiores de los rangos establecidos por los expertos. Aunque para confirmarlo harían falta más estudios, esa correspondencia con los niveles inferiores de los rangos puede deberse al hecho de que cuando un TFG llega a defenderse, en la inmensa mayoría de los casos ha pasado por una labor de revisión y corrección por parte de la persona supervisora que asegura un mínimo de calidad en el trabajo. Por lo tanto, sólo en casos excepcionales, un TFG que llega a

defenderse delante del tribunal recibe una calificación inferior al 5, lo que en la práctica limita el rango de calificaciones a obtener utilizando la rúbrica al intervalo [5, 10].

El modelo obtenido por `Mod_Ajuste` mejora el comportamiento del modelo establecido por los expertos, reduciendo el error y aproximándose al error considerado admisible de 0.5 puntos.

5.6 Conclusiones

En este capítulo se ha descrito el proceso de definición y evaluación de un conjunto de rúbricas para la evaluación de los TFG acorde al proceso establecido por la metodología FPM.

El proceso de definición y evaluación de las rúbricas ha contado con la colaboración de un grupo de expertos en la supervisión y evaluación de TFG que han definido una rúbrica de tipo analítico y otra de tipo atomístico para cada uno de los elementos evaluables definidos en la metodología. A partir de la definición inicial de las rúbricas y siguiendo un proceso iterativo de evaluación, se han ido ajustando hasta obtener una versión cuyo contenido y comportamiento satisface al profesorado que las ha utilizado.

El contenido de las rúbricas ha sido sometido a la evaluación del alumnado, por ser éste también parte importante del proceso. Los resultados de la evaluación han sido muy buenos. Las únicas dimensiones que han generado rechazo entre el alumnado han sido las correspondientes a ciertos formalismos como las referencias bibliográficas, pero el profesorado opina que dichos formalismos son importantes, por lo que se han mantenido estos aspectos en las rúbricas de evaluación.

También se ha preguntado tanto a profesorado como a alumnado sobre su opinión respecto a las distintas versiones de las rúbricas. Ambos actores coinciden en indicar que la rúbrica atomística proporciona una retroalimentación más detallada y adecuada. Sin embargo, entre el profesorado no existe unanimidad sobre el tipo de rúbrica que prefieren utilizar para realizar la evaluación, lo que sugiere que puede resultar adecuado no obligar a usar un tipo de rúbrica concreto y dejarlo a elección de la persona que tenga que rellenarla.

Una vez instaurado el uso de las rúbricas para la evaluación de los TFG del área de la Ingeniería Informática, se han usado los datos recogidos para ajustar, a través del módulo de ajuste de AdESMuS, los modelos matemáticos implicados en el proceso de evaluación de los TFG.

Por un lado se ha obtenido el TAM, que indica qué influencia tiene cada uno de los elementos evaluables definidos por la metodología FPM en el cálculo de la nota final del TFG. El modelo obtenido varía respecto al modelo fijado por los expertos, pero mejora sensiblemente sus resultados. Esto es, si se calculara la calificación de los estudiantes usando los porcentajes indicados por el TAM para cada uno de los elementos evaluables, dichas calificaciones serían más parecidas a las otorgadas realmente, que si se usaran los porcentajes indicados por los propios expertos.

Como el sistema de evaluación establecido por la metodología utiliza de rúbricas de evaluación, también se ha obtenido el REM correspondiente. El REM obtenido indica que hay varias dimensiones en las rúbricas que no tienen influencia real en la calificación y que, por lo tanto, podrían eliminarse. Asimismo, el modelo obtenido modifica sustancialmente el porcentaje de influencia de otras dimensiones, respecto al otorgado inicialmente por los expertos. Sin embargo, al igual que ocurre con el TAM, el modelo obtenido presenta mejores resultados que el modelo de los expertos a la hora de calcular la calificación correspondiente a cada estudiante.

Por lo tanto, se ha demostrado la utilidad de *Mod_Ajuste*, ya que se han obtenido unos modelos matemáticos cuyo comportamiento está más próximo al objetivo de permitir asignar a los trabajos la calificación más similar posible a la otorgada usando la opinión personal del tribunal de evaluación.

Sin embargo, no hay que olvidar que los modelos dependen directamente de los datos que se han utilizado para su obtención y que el profesorado no tiene por qué estar de acuerdo con ellos. Los modelos obtenidos pueden ser la base de una reflexión sobre la evaluación que se está llevando a cabo.

Por ejemplo, el TAM obtenido en el experimento indica que el elemento evaluable correspondiente a la defensa oral (*Eva_Presentación*), tiene una influencia sólo del 2% en la nota final. Este valor tan bajo puede responder al hecho de que los componentes del tribunal, tras la lectura de la documentación escrita, tal vez ya tengan decidida en gran medida la calificación que van a otorgar al TFG. Lo que convierte la defensa oral del mismo en un mero trámite que, salvo imprevistos, no modifica de manera significativa la calificación previamente decidida por los componentes del tribunal. Si el profesorado considerara que la presentación debiera tener una influencia mayor en la calificación final, tendría que modificar su forma de calificar y valorarla más a la hora de evaluar el trabajo.

La misma situación se produce con los porcentajes de influencia establecidos por el REM. Según los resultados obtenidos, por ejemplo, en la documentación escrita (*Eva_Memoria*), ni el apartado de conclusiones, ni la calidad global del documento influyen en la calificación. Si el profesorado considerara que dichos aspectos debieran tenerse en cuenta en la nota final que se otorga al documento, tendría que

considerar los valores obtenidos por el modelo como un aviso, y modificar consecuentemente su forma de calificar. Si, por el contrario, el profesorado considera que efectivamente dichos aspectos no tienen que influir en la calificación, se podrían simplificar las rúbricas correspondientes eliminando los contenidos de dichas dimensiones.

6

Evaluación

En este capítulo se detalla la evaluación de las mejoras a los procesos de realización y evaluación de los TFG propuestas.

Por un lado, se ha propuesto la metodología FPM, que define qué elementos deben formar parte de la evaluación de un TFG y quiénes deben realizar la evaluación de cada uno de esos elementos. Para poder implantar una metodología de este tipo en un entorno real, es necesario evaluar los beneficios que puede proporcionar. En este contexto, en este trabajo se han evaluado el aspecto de evaluación continua y las rúbricas definidas para ello.

Por otro lado, en esta memoria se ha presentado TraMeA, la plataforma que da soporte informático a la metodología FPM. En este capítulo se presenta la evaluación de usabilidad de TraMeA, ya que el éxito de una plataforma informática está directamente relacionado con su grado de usabilidad (Hosseini-Khayat et al., 2010; Mayhew, 1999).

Dado que TraMeA requiere los servicios de AdESMuS para determinadas funcionalidades, la evaluación de la plataforma ha contemplado también la de AdESMuS.

A continuación se detallan los procesos de evaluación realizados.

6.1 Evaluación de los aspectos de evaluación de la metodología FPM

En este apartado se presenta la evaluación de los aspectos de evaluación propuestos en la metodología FPM. En esta metodología se establece la evaluación de los TFG como un proceso continuo donde se evalúan 6 elementos a través de rúbricas de evaluación (ver sección 2.4).

Para realizar esta evaluación, se han analizado la adecuación de los elementos evaluables establecidos en la metodología y el impacto del uso de las rúbricas.

6.1.1 Adecuación de los elementos evaluables

En este apartado se presenta la evaluación realizada para determinar si los elementos evaluables identificados son adecuados o no.

Por un lado, en la sección 5.3 se ha mostrado que los actores que intervienen en la realización de un TFG consideraban que las dimensiones definidas son adecuadas. Como los elementos evaluables definidos cubren todas estas dimensiones, se puede concluir que todos los aspectos que el profesorado considera necesario evaluar están incluidos.

Por otro lado, se han realizado varios experimentos para determinar si los elementos evaluables identificados permiten determinar la calificación final del TFG mejor que el tribunal calificando únicamente la documentación final, la defensa oral y la complejidad técnica del trabajo.

En un estudio preliminar con los resultados obtenidos en la evaluación de 32 TFG (Villamañe et al., 2015), se comparó el modelo presentado en este trabajo con diversos modelos en los que se tenían en cuenta sólo algunos de los elementos evaluables. En dicho estudio se obtuvo que los mejores resultados se conseguían al considerar los 6 elementos evaluables.

Estos resultados se corroboraron con el estudio realizado con 68 trabajos que se ha presentado en la sección 5.5.2. En este estudio se obtuvieron los resultados que se muestran a continuación (ver Tabla 6-1).

Tabla 6-1.- Resultados evaluación elementos evaluables

| Elemento evaluable | Influencia en la calificación | Correlación | RMSE |
|--------------------|-------------------------------|-------------|------|
| Eva_Requisitos | 19% | 0.96 | 0.35 |
| Eva_Diseño | 12% | | |
| Eva_Actitud | 22% | | |
| Eva_Memoria | 23% | | |
| Eva_Presentación | 2% | | |
| Eva_Complejidad | 22% | | |

Si se analizan estos resultados, se observa que en el mejor modelo obtenido se contemplan todos los elementos evaluables definidos en la metodología (no aparece ninguna influencia con valor 0). Esto implica que la calificación que se obtiene utilizando los 6 elementos es mejor que cualquier otra combinación, obteniendo además un 0.35 RMSE, lo que está dentro del rango admisible definido para la evaluación de los TFG (ver sección 5.5.2).

Estos datos además corroboran los datos que se obtuvieron en el estudio preliminar con los resultados obtenidos en la evaluación de 32 TFG ((Villamañe et al., 2015)) en el que se compararon diversos modelos en los que se tenían en cuenta sólo algunos de los evaluables.

6.1.2 Beneficios del uso de rúbricas de evaluación

Una de las principales razones para la elección de rúbricas como instrumento de evaluación en la metodología FPM es la objetividad que supuestamente proporcionan (ver sección 2.2.1).

Para analizar si realmente el uso de rúbricas contribuye a una evaluación más objetiva, se realizó un estudio (Villamañe, Álvarez, et al., 2016a) donde se calculó el coeficiente de correlación intraclase (CCI) (Hallgren, 2012). El CCI es uno de los enfoques más comunes para analizar en qué medida evaluadores diferentes otorgan puntuaciones similares, ya que permite determinar la correlación entre las puntuaciones de diferentes evaluadores para su posterior análisis (Tinsley & Weiss, 2000).

Para calcular el CCI se seleccionaron aquellos TFG en los que los evaluadores habían realizado al menos un número suficiente de evaluaciones. Para ello se seleccionaron los datos correspondientes a 24 de los 32 TFG. Los ocho TFG restantes fueron desechados para evitar desviaciones, ya que las personas que los evaluaron sólo habían participado en la evaluación de uno o dos de los TFG.

El estudio se realizó sobre los elementos evaluables *Eva_Memoria* y *Eva_Presentación*, ya que son los elementos que deben evaluar varias personas y permiten comprobar si se consigue un CCI aceptable. Los demás evaluables sólo son evaluados por una persona y por lo tanto no es posible calcular el CCI.

El CCI se calculó buscando el acuerdo absoluto (McGraw & Wong, 1996). Para *Eva_Memoria*, se obtuvo un CCI de 0.8, lo que implica una correlación excelente (Cicchetti, 1994). Por su parte, *Eva_Presentación* obtuvo una correlación de 0.75, valor considerado bueno y muy cercano a la categoría de excelente.

Estos resultados demuestran que las rúbricas utilizadas proporcionan objetividad en la evaluación, ya que hay un alto grado de acuerdo en las calificaciones.

Otra de las razones para la elección de rúbricas como instrumento de evaluación es su utilidad para clarificar y explicitar los criterios de evaluación. Para evaluar si las rúbricas definidas cumplen este objetivo, se suministraron las rúbricas al alumnado y se realizó una encuesta para recoger su opinión al respecto (ver Tabla 6-2).

Tabla 6-2.- Extracto de la encuesta sobre rúbricas al alumnado

| Pregunta | | Respuestas posibles |
|-----------|--|--|
| E1 | Un sistema de evaluación para los TFG basado en rúbricas te parece... | Similar al actual / Mejor que el actual / Peor que el actual |
| E2 | Usando rúbricas, te parece que los criterios en los que se basa la evaluación... | Quedan claros / Quedan más claros que en la actualidad pero siguen siendo confusos / Quedan más confusos que en la actualidad / No quedan claros |
| E3 | ¿En qué tipo de rúbrica consideras que quedan más claros los criterios de evaluación? | Analítica / Atomística / Indistinta |
| E4 | ¿Qué tipo de rúbrica consideras que proporciona una mejor retroalimentación a los estudiantes? | Analítica / Atomística / Indistinta |
| E5 | Otros comentarios y/o sugerencias | Libre |

La encuesta fue respondida por 37 estudiantes tras la finalización y evaluación de su TFG. Las respuestas indican que el 100% del alumnado encuestado considera que un sistema de evaluación mediante rúbricas es, al menos, igual de bueno que el sistema tradicional de evaluación y el 67% lo considera mejor. Además, el 92% opina que las rúbricas deberían ser públicas y el 100% de las respuestas indican que el uso de rúbricas clarifica los criterios de evaluación.

6.2 Evaluación de TraMeA

TraMeA se ha evaluado mediante un estudio de usabilidad ya que el éxito de cualquier plataforma informática está directamente relacionado con su grado de usabilidad (Hosseini-Khayat et al., 2010; Mayhew, 1999). Este tipo de evaluación permite evaluar la satisfacción de los usuarios (Wieggers & Beatty, 2013) y observar cómo de intuitivo es el uso de la plataforma, los problemas que se encuentran los usuarios durante su utilización así como la formación previa requerida para usarla correctamente. Para realizar esta evaluación de TraMeA se implementó un prototipo funcional de la plataforma.

Entre los distintos métodos para evaluar la usabilidad de un sistema (Nielsen & Mack, 1994), se optó por el método del paseo cognitivo, (*Cognitive Walkthrough*). Este método consiste en proponer al usuario la realización de una o varias tareas y dejarle navegar por el sistema intentado solucionarlas (Wharton et al., 1994). Se eligió este método de evaluación porque permite evaluar cómo de intuitivo es el uso de TraMeA y la cantidad de formación previa que requiere su correcta utilización. De esta manera además, los futuros usuarios aprenden a manejar el sistema mientras realizan la prueba de usabilidad, lo que reduce el tiempo de formación necesario a la hora de implantarlo (Carroll & Rosson, 1987).

El alumnado, normalmente, hará uso TraMeA una única vez a lo largo de su vida académica dado que sólo realiza un TFG. Sin embargo, el profesorado tendrá que usar el sistema con todos los TFG que supervise y cada vez que forme parte del tribunal o tenga que dar el visto bueno al alcance de una propuesta de TFG. Por lo tanto, el profesorado va a ser un usuario habitual de TraMeA y es imprescindible satisfacer sus necesidades si se desea que el sistema tenga éxito (Lin & Wang, 2012). Además, gran parte de la funcionalidad del alumnado coincide con la del profesorado. Por estas razones, el estudio de usabilidad llevado a cabo se centró en el profesorado.

El proceso de evaluación de TraMeA se realizó desde dos perspectivas diferentes en función del rol de usuario involucrado en cada una de ellas. Por un lado se analizó la perspectiva de administración del sistema (rol Persona administradora) y por otro, la perspectiva de supervisión/evaluación de los TFG (roles Persona supervisora y Componente tribunal).

El proceso de evaluación se dividió en tres fases (ver Fig. 6-1) que se repitieron para evaluar cada una de las perspectivas. Siguiendo las pautas del método del paseo cognitivo, en la primera fase («Fase1: preparación estudio») se definieron las tareas a realizar por los participantes en el estudio.

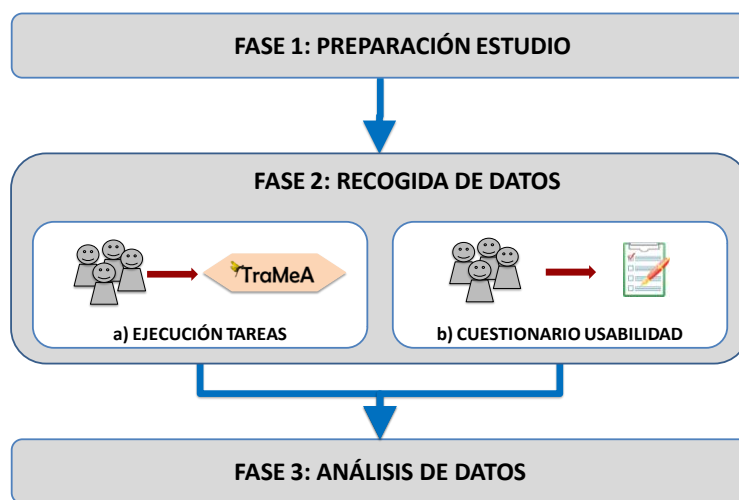


Fig. 6-1.- Proceso de evaluación

En la segunda fase («Fase2: recogida de datos») se solicitó a los participantes la ejecución de las tareas definidas en la fase anterior (Fig. 6-1→Fase 2.a) y se les pidió que respondieran a un cuestionario para recoger sus impresiones sobre el sistema (Fig. 6-1→Fase 2.b). Los datos recogidos durante la ejecución de las tareas y los resultados del cuestionario se sometieron a distintos análisis (ver Fig. 6-1→Fase 3: análisis de datos) para evaluar la usabilidad de TraMeA.

A continuación se presenta cada una de las fases del proceso de evaluación cuyos resultados y conclusiones se presentan en la sección 6.2.5.

6.2.1 Fase 1: Preparación estudio

Para aplicar el método del paseo cognitivo lo primero que se debe hacer es identificar las tareas o casos de prueba de usabilidad (CPU) que deben realizar los usuarios. Para este estudio, las tareas se definieron a partir de los UO de TraMeA directamente relacionados con las tareas que debe realizar el profesorado durante las distintas fases de un TFG.

En la Tabla 6-3 se muestran las tareas definidas para el estudio, agrupadas por la fase a la que están asociadas. Junto a la descripción de cada tarea se indica el rol de usuario encargado de su ejecución y el código del UO de TraMeA al que da respuesta la ejecución de la tarea.

Tabla 6-3.- Tareas definidas para el estudio de usabilidad y su correspondencia con los UO de TraMeA

| Fase | Descripción Tarea | Rol | Código UO |
|-------------|----------------------------------|--|------------------|
| Inicio | Añadir criterios | Persona administradora | UOF01 |
| | Asignar estudiante y supervisión | Persona administradora | U003-A1 |
| Cierre | Finalizar TFG | Persona administradora | U003-A2 |
| | Asignar tribunal | Persona administradora | U003-A3 |
| Seguimiento | Consultar planificación TFG | Persona supervisora Componente tribunal | UOF07 |
| | Consultar actividad estudiante | Persona supervisora Componente tribunal | UOF08 |
| | Consultar detalles tarea | Persona supervisora Componente tribunal | UOF09 |
| | Consultar documento | Persona supervisora Componente tribunal | UOF10 |
| | Dar retroalimentación | Persona supervisora | U002-T2 |
| | Consultar retroalimentación | Persona supervisora | UOF06 |
| | Solicitar cierre | Persona supervisora | U004-T |
| Evaluación | Evaluar un elemento | Persona supervisora Componente tribunal | UOF11 |
| | Consultar evaluación | Persona supervisora Componente tribunal | UOF12 |

La Tabla 6-4 recoge el código asignado a cada una de las tareas definidas para llevar a cabo el estudio de usabilidad (ver Tabla 6-3). Dicho código está formado por las letras CPU seguidas de un índice numérico y un sufijo que indica el rol de TraMeA al que se asocia la tarea: A para el rol Persona administradora, T para el rol Persona supervisora y Tr para el rol Componente tribunal. Las tareas asociadas a más de un rol, se nombran concatenando los sufijos correspondientes. También se indica a qué perspectiva del estudio, administración o supervisión/evaluación, corresponde cada tarea.

Tabla 6-4.- Códigos de las tareas del estudio de usabilidad

| Código Tarea | Descripción Tarea | Perspectiva del estudio |
|--------------|----------------------------------|-------------------------|
| CPU0_A | Añadir criterios | Administración |
| CPU1_A | Asignar estudiante y supervisión | |
| CPU2_A | Finalizar TFG | |
| CPU3_A | Asignar tribunal | |
| CPU4_T_Tr | Consultar planificación TFG | Supervisión/Evaluación |
| CPU5_T_Tr | Consultar actividad estudiantes | |
| CPU6_T_Tr | Consultar detalles tarea | |
| CPU7_T_Tr | Consultar documento | |
| CPU8_T | Dar retroalimentación | |
| CPU9_T | Consultar retroalimentación | |
| CPU10_T | Solicitar cierre | |
| CPU11_T_Tr | Evaluar un elemento | |
| CPU12_T_Tr | Consultar evaluación | |

6.2.2 Fase 2: Recogida de datos

En esta fase se recogieron los datos generados durante la ejecución de las tareas y posteriormente los obtenidos a partir de los cuestionarios de usabilidad (ver Fig. 6-1).

A continuación se describe cómo se realizó esta fase con cada grupo de datos.

a) Ejecución tareas

El objetivo principal de esta fase es la recogida de datos relativos a la ejecución por parte de los participantes de las tareas definidas en la primera fase.

Para recoger los datos relativos a la ejecución de las tareas y con el objetivo de conseguir la mayor cantidad posible de usuarios para probar el sistema, se optó por realizar un prototipo que soportara una evaluación remota controlada (*instrumented remote evaluation*), de modo que fuera el propio prototipo el encargado de almacenar los datos necesarios para conducir el estudio de usabilidad (Hartson et al., 1996). Este tipo de prototipo facilita la participación ya que cada usuario puede ejecutar la prueba de manera independiente, en el momento que prefiera y sin necesidad de contar con ninguna otra persona.

El prototipo fue desarrollado usando el software Justinmind³⁶, que permite obtener prototipos interactivos funcionales basados en HTML y JavaScript. El prototipo desarrollado se desplegó posteriormente en un servidor web para permitir el acceso remoto a los distintos usuarios.

El modelo del sistema de TraMeA (ver sección 3.2.2) establece que el sistema generado debe respetar el aspecto gráfico del sistema Redmine. En la Fig. 6-2 puede observarse cómo el prototipo realizado (Fig. 6-2→b) sigue fielmente el aspecto original proporcionado por Redmine (Fig. 6-2→a), limitándose a añadir los elementos necesarios para las funcionalidades que se desean evaluar. Por ejemplo, en la Fig. 6-2→b se indica mediante una flecha la nueva opción añadida a un menú de Redmine para poder proporcionar funcionalidades no presentes en el sistema original.

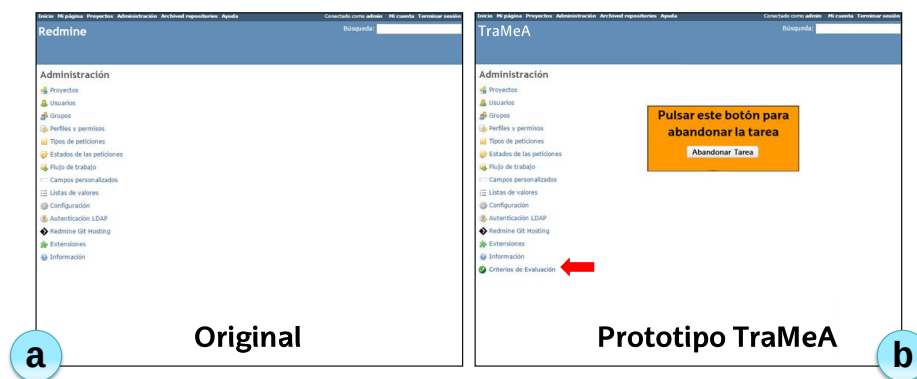


Fig. 6-2.- Comparación del sistema original (a) y del prototipo realizado (b)

El prototipo implementado muestra una interfaz con el listado de las posibles tareas a realizar en función del rol de acceso (ver Fig. 6-3→a). Para este estudio, se solicitó a cada participante que ejecutase las tareas en un orden determinado. Dicho orden se obtuvo de manera aleatoria para cada sujeto, intentando evitar que en los resultados del estudio repercutieran el orden de ejecución de las tareas (Losada et al., 2012; Tullis & Albert, 2013) y el efecto que el cansancio o el aburrimiento acumulado durante la prueba pudieran tener sobre las últimas tareas a realizar.

³⁶ <http://www.justinmind.com>

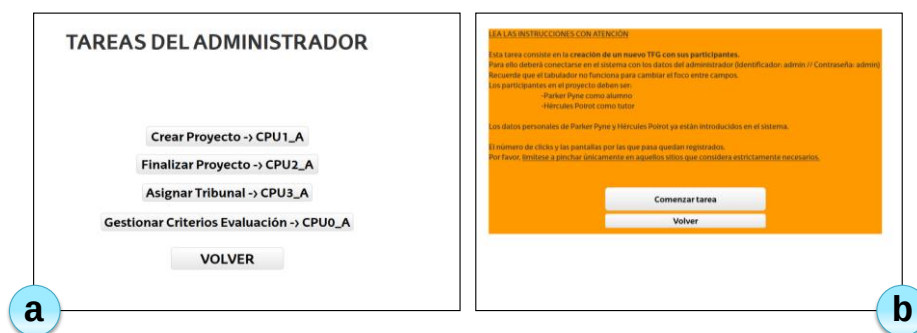


Fig. 6-3.- Interfaces gráficas de elección (a) y descripción (b) de tareas

Una vez que se selecciona una tarea, se muestran las instrucciones para llevarla a cabo (Fig. 6-3→b).

Cuando la persona que esté realizando la prueba, considere que ha comprendido en qué consiste la tarea, puede comenzar su ejecución. A partir de ese momento, dispone de acceso completo al prototipo pudiendo navegar libremente e interactuar con todos los elementos del sistema que desee para completar la tarea propuesta.

El prototipo almacena automáticamente información sobre la resolución o el abandono de la tarea presentada, lo que facilita el posterior análisis de la usabilidad del sistema. Automatizar la recogida de datos presenta ventajas como el ahorro temporal a la hora de registrar y analizar los resultados. Además, proporciona la posibilidad de analizar en detalle dónde se han producido las desviaciones del camino correcto en la actividad de los usuarios y qué elementos fueron los que les desviaron de su objetivo mientras intentaban completar la tarea seleccionada (Ivory & Hearst, 2001).

Para cada tarea realizada, el prototipo permite visualizar un resumen de la información almacenada. Se recogen, entre otros datos: el número de veces que se ha pulsado algún elemento en la interfaz gráfica (Fig. 6-4→a), el momento de inicio y de finalización de la prueba para el cálculo del tiempo empleado (Fig. 6-4→b) y el camino seguido durante la resolución de la tarea propuesta (Fig. 6-4→c). De este modo, es posible extraer información sobre qué partes del sistema han resultado más confusas para las personas usuarias y en qué zonas se han perdido con mayor facilidad.

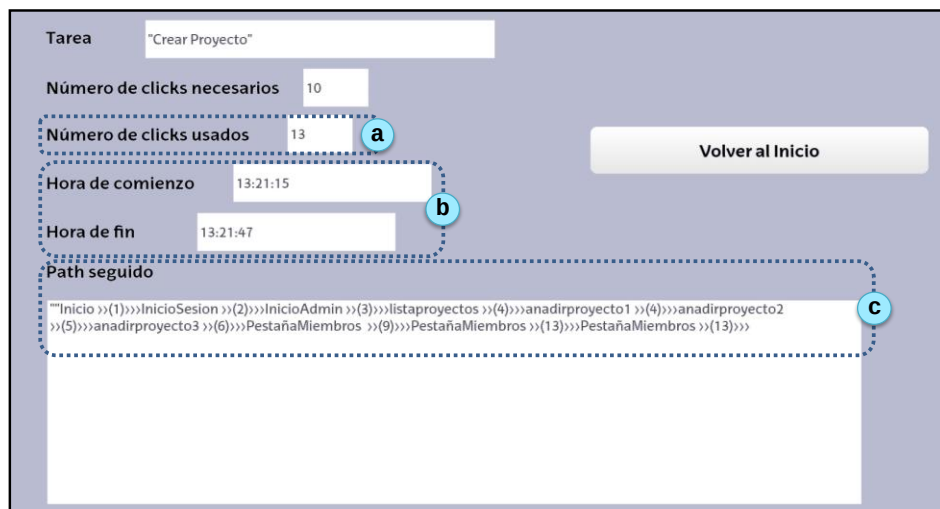


Fig. 6-4.- Resumen de una prueba de usabilidad

b) Cuestionario usabilidad

Para recoger las opiniones de los usuarios sobre el sistema se usó el cuestionario con la Escala de Usabilidad del Sistema (*System Usability Scale, SUS*). Esta escala es una herramienta eficaz y eficiente a la hora de evaluar la usabilidad de un sistema (J. Brooke, 1996; John Brooke, 2013).

En la Tabla 6-5 se muestra el cuestionario SUS utilizado para el estudio de usabilidad de TraMeA. Los participantes en el estudio rellenaron el cuestionario utilizando una escala de Likert, donde el 1 indica que se está completamente de desacuerdo y el 5 completamente de acuerdo, con cada una de las 10 afirmaciones.

Tabla 6-5.- Cuestionario de la Escala de Usabilidad del Sistema (SUS)

| | 1 | 2 | 3 | 4 | 5 |
|--|----------|----------|----------|----------|----------|
| 1.- Creo que me gustará usar el sistema con frecuencia | | | | | |
| 2.- Encontré el sistema innecesariamente complejo | | | | | |
| 3.- Me pareció que el sistema era fácil de usar | | | | | |
| 4.- Creo que necesitaría el apoyo de un experto para poder usar el sistema | | | | | |
| 5.- Encontré las distintas funcionalidades del sistema bastante bien integradas | | | | | |
| 6.- Me pareció que había demasiada inconsistencia en el sistema | | | | | |
| 7.- Imagino que la mayoría de las personas aprenderían a usar el sistema muy rápidamente | | | | | |
| 8.- Encontré el sistema muy engorroso de usar | | | | | |
| 9.- Me sentí cómodo usando el sistema | | | | | |
| 10.- Necesité aprender muchas cosas antes de manejarlo con el sistema | | | | | |

6.2.3 Fase 3: Análisis de datos

Una vez recogidos los datos sobre la realización de las tareas y los datos de los cuestionarios de usabilidad, se procedió a su análisis a través de diversas métricas de usabilidad. La norma ISO 9241-11 (ISO, 1998) define la usabilidad como «*El grado en el que un producto puede ser utilizado por usuarios para lograr objetivos concretos con efectividad, eficiencia y satisfacción*». Por ello, los datos de ejecución de las tareas se utilizaron para evaluar la efectividad y la eficiencia del sistema. Con los resultados de los cuestionarios se calculó el grado de satisfacción de los usuarios.

A continuación se describen las métricas usadas para evaluar la usabilidad (la efectividad, eficiencia y satisfacción) y las técnicas para su cálculo. También se presentan los grafos utilizados para representar de manera gráfica la navegación de los usuarios y facilitar su análisis.

6.2.3.1 Efectividad

La métrica utilizada para evaluar la efectividad del sistema fue la tasa de éxito de las tareas (Tullis & Albert, 2013). La tasa de éxito muestral (p) de una tarea se calcula mediante la ecuación

$$p = x/n \quad (6-1)$$

donde x representa el número de usuarios que completaron la tarea correctamente y n el número total de usuarios que participaron en el estudio.

Con la tasa de éxito calculada a partir de los datos recogidos de los participantes (la tasa de éxito muestral) se puede estimar cual será la tasa de éxito de la población en general. Multiplicando por 100 la tasa de éxito de la población general se obtiene el porcentaje de futuros usuarios que serán capaces de completar con éxito cada tarea. Dicho porcentaje se considera aceptable a partir del 70% para versiones preliminares del sistema (Rubin & Chisnell, 2008).

Para tasas de éxito muestral cuyo valor se encuentre entre 0.5 y 0.9, la estimación de la tasa de éxito poblacional se corresponderá con la tasa de éxito muestral calculada mediante la ecuación (6-1). Sin embargo, para otros valores, existen estimadores más precisos (Lewis & Sauro, 2006).

Si la tasa de éxito muestral es inferior o igual a 0.5, el estimador más adecuado para la tasa de éxito poblacional es el método Wilson (Wilson, 1927):

$$Wilson = (x + z_{\alpha/2}^2/2)/(n + z_{\alpha/2}^2) \quad (6-2)$$

siendo x el número de usuarios que completaron la tarea correctamente; n el número total de sujetos participantes en el estudio; α el nivel de significación; y $z_{\alpha/2}$ el valor crítico de la distribución normal estandarizada.

En cambio, para tasas de éxito muestral superiores a 0.9, el método Laplace (Laplace, 1812; Lewis & Sauro, 2006) es el más preciso:

$$Laplace = (x + 1)/(n + 2) \quad (6-3)$$

donde x es el número de usuarios que completaron la tarea correctamente y n el número total de sujetos participantes en el estudio.

Sin embargo, dado que la población de muestra del estudio era inferior a 150 sujetos, la forma más precisa de expresar las distintas tasas de éxito es a través de intervalos de confianza calculados mediante el método ajustado de Wald (Agresti & Coull, 1998; J. Sauro & Lewis, 2005):

$$\hat{p} \pm z_{\alpha/2} \times \sqrt{\frac{\hat{p} \times (1 - \hat{p})}{n + z_{\alpha/2}^2}} \quad (6-4)$$

donde \hat{p} representa el estimador de Wilson calculado a través de la ecuación (6-2).

Al calcular intervalos de confianza es práctica habitual hacerlo con un nivel de significación $\alpha = 0.05$, lo que proporciona intervalos con un nivel de confianza del 95%.

Al efectuar los cálculos anteriores teniendo en cuenta únicamente el éxito o el fracaso a la hora de realizar una tarea, se pueden obtener resultados demasiado optimistas ya que debería tenerse en cuenta que, a pesar del éxito final, durante la ejecución podrían haberse presentado distintos problemas. Con objeto de obtener resultados lo más precisos posible y conocer el porcentaje de usuarios que no tendrán dificultades para finalizar las tareas, se recomienda el granular de los niveles de finalización y el cálculo de las tasas de éxito en base a dicha granularidad (Tullis & Albert, 2013).

Para este estudio se definieron 3 niveles de finalización: *abandono* (la tarea no se finalizó con éxito); *con problemas* (la tarea finalizó correctamente, pero el proceso seguido no fue el óptimo); y *sin problemas* (la tarea se finalizó correctamente siguiendo el proceso óptimo).

Teniendo en cuenta los distintos niveles de éxito definidos, se volvieron a estimar las tasas de éxito y sus intervalos de confianza para calcular el porcentaje de futuros usuarios que serán capaces de finalizar cada tarea en el nivel *sin problemas*, esto es, que realizarán la tarea siguiendo el proceso óptimo de resolución.

6.2.3.2 Eficiencia

La eficiencia del sistema se calculó analizando los tiempos de ejecución de las tareas y el nivel de pérdida (*lostness*) junto con el recorrido que los usuarios realizaron a través del sistema para resolver las tareas.

Tiempo de ejecución

Al no existir datos previos sobre el tiempo más adecuado para la ejecución de cada tarea (ya que hasta ahora no se habían llevado a cabo), ni otro sistema similar a TraMeA con el que realizar comparativas, especificar qué máximo de tiempo se considera aceptable para la ejecución de cada tarea es complicado (Dumas & Redish, 1999).

Sin embargo, la definición de un máximo de tiempo admisible basado en datos es mejor que establecer un límite aleatorio o no establecer ningún límite. Por esta razón, se fija el máximo de tiempo admisible para la ejecución tomando como base el tiempo empleado por los participantes en el estudio. Dicho máximo será el resultado de calcular el percentil 95 de los tiempos de ejecución de los participantes que finalizaron la tarea y valoraron satisfactoriamente el sistema (más de 71 puntos en el cuestionario SUS, ver apartado *Satisfacción*). Se considera únicamente a los participantes que indican estar bastante o muy satisfechos con el sistema porque de su satisfacción se puede deducir que también consideran adecuado el tiempo que les ha llevado realizar las tareas. Los tiempos máximos calculados mediante este método son altamente variables, ya que dependen de los participantes de la muestra. No obstante, establecen un límite razonable basado en datos (Jeff Sauro & Kindlund, 2005).

Los tiempos de ejecución de las tareas también se analizaron para estimar el tiempo que tardaría en realizar cada tarea un nuevo usuario sin experiencia previa en el uso del sistema. En los estudios con menos de 25 participantes, el mejor estimador para calcular la mediana del tiempo de realización de las tareas es la media geométrica (G) (Jeff Sauro & Lewis, 2010).

Sin embargo, los tiempos recogidos durante el estudio no se pueden usar directamente, ya que no siguen una distribución normalizada. Es decir, los valores recogidos no se distribuyen proporcionalmente a lo largo del tiempo si no que presentan dos tendencias: o se agrupan alrededor del tiempo necesario para realizar la tarea de manera óptima; o presentan valores muy altos correspondientes a aquellos usuarios que se han encontrado con problemas y, por lo tanto, han dedicado más tiempo del estrictamente necesario en resolver la tarea. Por este motivo, la distribución de los tiempos de ejecución de las tareas suele presentar una distribución sesgada a la derecha (Jeff Sauro & Lewis, 2012).

Para normalizar la distribución de los tiempos, éstos se transforman a través de la función logarítmica. Posteriormente, la media geométrica (G) de los tiempos se calcula como la media aritmética de los logaritmos tal y como se muestra en la ecuación (6-5):

$$G = \frac{\sum_{i=1}^x \ln t_i}{x} \quad (6-5)$$

donde x es el número de participantes en el estudio que finalizaron la tarea y t_i el tiempo que el participante i tardó en hacerlo. El valor G obtenido se transforma exponencialmente para expresar el resultado en la misma escala en la que se han recogido los tiempos originalmente.

Una vez normalizada la distribución de los tiempos, y dado que el tamaño de la muestra es menor a 30, se puede usar la *t* de Student para calcular los intervalos de confianza a través de la ecuación

$$\bar{x} \pm t_{n-1, \alpha/2} \times \frac{s}{\sqrt{n}} \quad (6-6)$$

donde \bar{x} es la media aritmética de la muestra; n es el tamaño de la muestra; $t_{n-1, \alpha/2}$ es el valor crítico de la distribución para $n - 1$ grados de libertad y un nivel de significación α ; y s es la desviación estándar muestral calculada mediante la ecuación

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (6-7)$$

Siguiendo el uso habitual en estadística, los intervalos de confianza se calcularán con un nivel de significación $\alpha = 0.05$.

Nivel de pérdida

El nivel de pérdida (*lostness*) representa en una escala numérica de 0 a 1, cuánto se ha perdido el sujeto durante la ejecución de una tarea, teniendo en cuenta el recorrido realizado (Smith, 1996).

El nivel de pérdida (L) de un sujeto en una tarea se calcula mediante la ecuación (6-8):

$$L = \sqrt{\left(\frac{N}{S} - 1\right)^2 + \left(\frac{R}{N} - 1\right)^2} \quad (6-8)$$

donde N representa el número de páginas distintas que ha visitado para resolver la tarea; S el número total de páginas visitadas para resolver la tarea incluyendo cada vez que se ha revisitado alguna página; y R , el número mínimo de páginas que hay que visitar para resolver la tarea de manera óptima.

El valor ideal para el nivel de pérdida es 0, pero con valores inferiores a 0.4 no se considera que la persona realmente se haya perdido.

Para estimar el nivel de pérdida de la población, se calculó el nivel medio de pérdida de cada tarea y, debido al reducido tamaño de la muestra, mediante la ecuación (6-6) se construyó un intervalo de confianza al 95%.

6.2.3.3 Grafos de navegación

Para facilitar la identificación de las razones por las que los usuarios se pierden en el sistema, en este trabajo de tesis se ha definido un tipo de grafo que permite visualizar de manera agregada los datos de navegación de todos los usuarios durante la realización de una tarea específica.

Este tipo de grafo permite mostrar de manera gráfica (ver ejemplo en Fig. 6-5) las páginas visitadas (los nodos del grafo) y el recorrido realizado por los distintos usuarios (los arcos dirigidos entre los nodos).

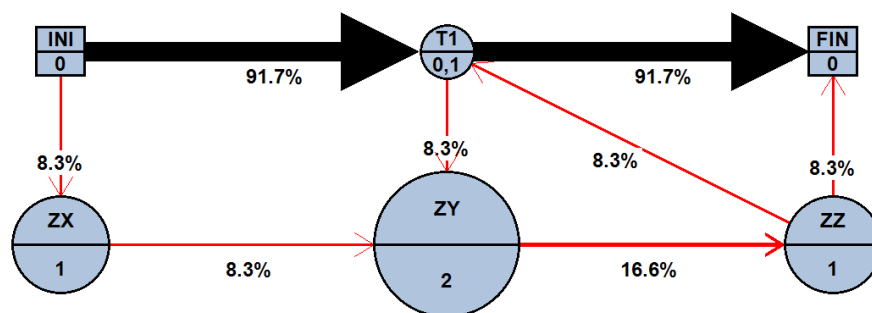


Fig. 6-5.- Ejemplo de grafo con el recorrido de ejecución

En este tipo de grafo, la forma de los nodos indica si se trata de nodos de comienzo y/o finalización de la tarea (nodos cuadrados) o de nodos intermedios (nodos redondos). Cada nodo está dividido en dos partes, en la parte superior se encuentra un identificador correspondiente a la página del sistema representada por dicho nodo; en la parte inferior se reseña el promedio de acciones extra que los usuarios llevaron a cabo en esa página. Se consideran acciones extra aquellas que no son necesarias para la resolución de la tarea y/o que tengan lugar en una página que no forme parte del camino correcto. Para el cálculo del promedio de acciones extra sólo se tiene en cuenta el número de usuarios que pasaron por dicha página. El tamaño de cada nodo viene determinado por el valor de su promedio de acciones extra.

Los arcos del grafo representan la navegación realizada por los usuarios entre las distintas páginas del sistema. El color y el tipo de flecha expresan si el paso de una página a otra está contenido en el recorrido óptimo para realizar la tarea (flechas negras con cabeza cerrada y rellena) o no (flechas rojas con cabeza abierta). A su vez, el grosor del arco y la etiqueta del mismo indican el porcentaje de usuarios que realizaron ese paso.

6.2.3.4 Satisfacción

Para estimar el grado de satisfacción de los usuarios, se calculó la puntuación obtenida en los cuestionarios SUS usando la ecuación(6-9):

$$SUS = \left(\sum_{i=1,3,5,7,9} (q_i - 1) + \sum_{j=2,4,6,8,10} (5 - q_j) \right) \times 2.5 \quad (6-9)$$

donde q_i y q_j representan la valoración dada por el usuario a las correspondientes preguntas i (preguntas con connotación positiva) y j (preguntas con connotación negativa) del cuestionario (J. Brooke, 1996).

Para categorizar los resultados obtenidos se utilizó la escala definida por (Bangor et al., 2009), según la cual a partir de una puntuación de 51, el grado de satisfacción con la usabilidad del sistema es aceptable; siendo deseable alcanzar al menos una puntuación de 71 que representa el límite inferior para una buena usabilidad. Debido al limitado tamaño de la muestra, con los valores obtenidos a través de los cuestionarios se computó la media aritmética de las puntuaciones obtenidas y se construyó un intervalo de confianza al 95% para estimar la puntuación que se obtendría al preguntar a todos los posibles usuarios. Para la construcción del intervalo de confianza se usó la ecuación (6-6).

6.2.4 Participantes

Todos los sujetos participantes en el estudio eran personal docente e investigador perteneciente a la UPV/EHU y relacionados con el campo de la Ingeniería Informática. Los participantes no tenían conocimientos previos del sistema, pero sí experiencia en la supervisión y evaluación de Trabajos Fin de Grado y en la realización de proyectos en general.

Para evaluar tanto la parte de administración de TraMeA como la parte de supervisión/evaluación se formó un grupo de usuarios a los que se les proporcionó la lista de tareas a ejecutar y el orden en el que debían hacerlo. El número de sujetos en cada grupo se fijó en 12 ya que una población de ese tamaño proporciona una fiabilidad de entre el 90% y el 100% a los resultados obtenidos a través del SUS (Tullis & Stetson, 2004).

Un total de 18 sujetos participó en la prueba y se distribuyeron aleatoriamente en dos grupos disjuntos de 12 sujetos cada uno. Cada grupo de 12 participantes evaluó una parte del sistema.

6.2.5 Resultados de la evaluación

A continuación se resumen los principales resultados obtenidos en el estudio de usabilidad para cada una de las perspectivas del estudio sobre TraMeA: administración y supervisión/evaluación y un resumen final que recoge las conclusiones generales de la evaluación.

6.2.5.1 Administración

En este apartado se desglosan los resultados obtenidos para la perspectiva de administración en el estudio de evaluación de usabilidad de TraMeA. El apartado está dividido siguiendo los tres aspectos de la usabilidad descritos anteriormente: efectividad, eficiencia y satisfacción.

Efectividad

El análisis de la efectividad se realizó en base a la tasa de éxito de las tareas. En la Tabla 6-6 se muestran los resultados recogidos sobre cada una de las tareas. Se representa para cada sujeto participante en el estudio, si finalizó la tarea correctamente (valor «SI»), o no (valor «NO»). A partir de esos datos, la tabla muestra también el porcentaje de tareas completadas con éxito por cada sujeto y la tasa de éxito muestral de cada tarea.

Tabla 6-6.- Porcentaje de tareas completadas con éxito y su tasa de éxito muestral

| | CPU0_A | CPU1_A | CPU2_A | CPU3_A | Porcentaje de tareas completadas con éxito |
|-------------------------------|---------------|---------------|---------------|---------------|---|
| Sujeto 1 | SI | SI | SI | SI | 100% |
| Sujeto 2 | SI | SI | SI | SI | 100% |
| Sujeto 3 | SI | SI | SI | SI | 100% |
| Sujeto 4 | SI | SI | SI | SI | 100% |
| Sujeto 5 | SI | SI | SI | SI | 100% |
| Sujeto 6 | SI | SI | SI | SI | 100% |
| Sujeto 7 | NO | SI | SI | SI | 75% |
| Sujeto 8 | SI | SI | SI | SI | 100% |
| Sujeto 9 | SI | SI | SI | SI | 100% |
| Sujeto 10 | SI | SI | SI | SI | 100% |
| Sujeto 11 | SI | SI | SI | SI | 100% |
| Sujeto 12 | SI | SI | SI | SI | 100% |
| Tasa de éxito muestral | 0.917 | 1 | 1 | 1 | |

A excepción de un sujeto (el número 7) que no finalizó correctamente la tarea CPU0_A, el resto de usuarios presenta un porcentaje de tareas completadas con éxito del 100% y todas las tareas, a excepción de la CPU0_A, presentan una tasa de éxito de 1.

Como todas las tasas de éxito muestrales son superiores a 0,9, el método Laplace (ecuación (6-3)) es el más preciso para estimar la tasa de éxito de la población general para cada una de las tareas (ver sección 6.2.3). En la Tabla 6-7 se muestran las tasas de éxito poblacionales estimadas para cada una de las tareas usando el método de Laplace.

Tabla 6-7.- Estimación de las tasas de éxito poblacionales

| | Estimación tasa de éxito poblacional |
|--------|---|
| CPU0_A | 0.857 |
| CPU1_A | 0.929 |
| CPU2_A | 0.929 |
| CPU3_A | 0.929 |

De acuerdo a las tasas de éxito poblacionales estimadas, el porcentaje de usuarios que se espera sea capaz de resolver satisfactoriamente las tareas es el 85.70% para la tarea CPU0_A y el 92.90% para el resto. Estos resultados reflejan que la gran mayoría de usuarios serán capaces de finalizar las tareas planteadas en la perspectiva de administración de TraMeA con éxito y superan con creces el estándar establecido del 70%.

En la Tabla 6-8 se muestran los intervalos de confianza ajustados de Wald al 95% para cada una de las tareas del estudio en función de si se terminaron con éxito o no.

Tabla 6-8.- Intervalos de confianza ajustados de Wald al 95% (éxito/fracaso)

| | Límite inferior | Límite superior |
|--------|------------------------|------------------------|
| CPU0_A | 0.625 | 0.999 |
| CPU1_A | 0.784 | 1 |
| CPU2_A | 0.784 | 1 |
| CPU3_A | 0.784 | 1 |

Todas las tareas excepto la CPU0_A, presentan un intervalo de confianza que comienza en 0.784, lo que significa que se puede asegurar con una confianza del 95% que como mínimo el 78.40% de los usuarios será capaz de finalizar dicha tarea con éxito. En el caso de la tarea CPU0_A, si se calcula el intervalo de confianza con un nivel de confianza del 85%, los valores para la tarea CPU0_A pasan a ser 0.720 y 0.990 respectivamente. Por lo tanto, se puede asegurar con un 85% de confianza, que se satisfará el 70% establecido como límite de aceptabilidad y que, en el peor de los casos, el 72% de los usuarios será capaz de finalizar correctamente la tarea.

Al analizar los datos en función de los distintos niveles de éxito posibles a la hora de finalizar una tarea (*sin problemas, con problemas, abandono*), los resultados varían. En la Tabla 6-9 se muestra la categorización de cada una de las ejecuciones en función de los niveles de éxito, junto con la tasa de éxito de las tareas y el porcentaje de tareas completadas por cada usuario teniendo en cuenta únicamente el nivel *sin problemas*, esto es, considerando exclusivamente aquellas ejecuciones que fueron completamente perfectas.

Tabla 6-9.- Resultados en función del nivel de finalización

| | CPU0_A | CPU1_A | CPU2_A | CPU3_A | Porcentaje de tareas completadas sin problemas |
|---|---------------|---------------|---------------|---------------|--|
| Sujeto 1 | Con problemas | Con problemas | Sin problemas | Sin problemas | 50% |
| Sujeto 2 | Con problemas | Sin problemas | Sin problemas | Sin problemas | 75% |
| Sujeto 3 | Sin problemas | Sin problemas | Sin problemas | Sin problemas | 100% |
| Sujeto 4 | Con problemas | Sin problemas | Sin problemas | Sin problemas | 75% |
| Sujeto 5 | Con problemas | Con problemas | Sin problemas | Sin problemas | 50% |
| Sujeto 6 | Con problemas | Con problemas | Sin problemas | Con problemas | 25% |
| Sujeto 7 | Abandono | Sin problemas | Sin problemas | Sin problemas | 75% |
| Sujeto 8 | Con problemas | Sin problemas | Sin problemas | Sin problemas | 75% |
| Sujeto 9 | Con problemas | Con problemas | Con problemas | Sin problemas | 25% |
| Sujeto 10 | Sin problemas | Con problemas | Sin problemas | Sin problemas | 75% |
| Sujeto 11 | Sin problemas | Sin problemas | Sin problemas | Con problemas | 75% |
| Sujeto 12 | Sin problemas | Sin problemas | Sin problemas | Sin problemas | 100% |
| Tasa de éxito muestral (sin problemas) | 0.333 | 0.583 | 0.917 | 0.833 | |

Independientemente de la gravedad de los problemas encontrados, que se analizará mediante el nivel de pérdida (ver sección *Eficiencia*), el 67% de los participantes terminó de manera óptima el 75% de las tareas y el 83% de los participantes fueron capaces de realizar de manera óptima al menos la mitad de las tareas.

Las tasas de éxito de las distintas tareas presentan valores muy diversos, por lo que el método más preciso para estimar la tasa de éxito poblacional varía para cada una de ellas. En la Tabla 6-10 se muestran las tasas de éxito estimadas para cada una de las tareas y el método utilizado para calcularla.

Tabla 6-10.- Estimación de las tasas de éxito poblacional y método de cálculo

| | Estimación tasa de éxito poblacional | Método de cálculo |
|--------|---|--------------------------|
| CPU0_A | 0.374 | Wilson |
| CPU1_A | 0.583 | Tasa de éxito muestral |
| CPU2_A | 0.852 | Laplace |
| CPU3_A | 0.833 | Tasa de éxito muestral |

Observando la estimación de la tasa de éxito poblacional para cada tarea, se puede concluir que todas las tareas excepto la CPU0_A, se ejecutarán de manera óptima como mínimo por más de la mitad de los usuarios. Destacan muy positivamente los resultados de las tareas CPU2_A y CPU3_A, cuyos datos indican que más del 80% de los usuarios las resolverán de manera óptima.

También se calcularon de nuevo los intervalos de confianza ajustados de Wald, pero en este caso, teniendo en cuenta solamente los usuarios que finalizaron la ejecución de las tareas en el nivel de éxito *sin problemas* (ver Tabla 6-11). De este modo se obtuvo la probabilidad al 95% que tienen los usuarios de resolver las distintas tareas de manera óptima.

Tabla 6-11.- Intervalos de confianza ajustados de Wald al 95% (*sin problemas*)

| | Límite inferior | Límite superior |
|--------|------------------------|------------------------|
| CPU0_A | 0.136 | 0.612 |
| CPU1_A | 0.319 | 0.807 |
| CPU2_A | 0.625 | 0.999 |
| CPU3_A | 0.540 | 0.965 |

Los resultados de la Tabla 6-11 muestran que es muy probable que un nuevo usuario realice las tareas CPU2_A y CPU3_A de forma óptima en la primera ocasión, sin ningún tipo de desvío. En el caso de la tarea CPU1_A, aunque el límite inferior del intervalo es menor, el superior es bastante alto, lo que indica que el porcentaje de usuarios que resolverán la tarea de forma óptima puede llegar a ser hasta del 80.70%.

Sin embargo, la tarea CPU0_A presenta unos límites inferior y superior bajos, lo que hace vislumbrar que puede existir algún problema con el diseño de dicha funcionalidad. Por lo tanto, debe prestarse especial atención a los resultados relativos a dicha tarea en el resto de apartados de la evaluación de usabilidad.

Eficiencia

La eficiencia del sistema se analizó a través de los tiempos de ejecución de las tareas y del camino recorrido por los participantes para resolver cada tarea.

La Tabla 6-12 presenta el tiempo, expresado en segundos, que cada uno de los participantes en la evaluación de usabilidad tardó en resolver cada una de las tareas. El sujeto 7 fue el único que abandonó una tarea sin finalizarla (la tarea CPU0_A) por lo que su tiempo para esta tarea, ni figura en la tabla, ni se ha tenido en cuenta en los distintos análisis llevados a cabo.

Para establecer el tiempo máximo admisible de ejecución de cada tarea se seleccionaron únicamente los tiempos correspondientes a los participantes que habían valorado el sistema satisfactoriamente (ver sección *Satisfacción*) y que aparecen sombreados en la Tabla 6-12.

Tabla 6-12.- Tiempos de ejecución de las tareas (segundos)

| | CPU0_A | CPU1_A | CPU2_A | CPU3_A |
|----------------|---------------|---------------|---------------|---------------|
| Sujeto 1 | 175 | 252 | 20 | 386 |
| Sujeto 2 | 159 | 90 | 15 | 37 |
| Sujeto 3 | 161 | 111 | 74 | 46 |
| Sujeto 4 | 124 | 70 | 53 | 36 |
| Sujeto 5 | 233 | 168 | 37 | 54 |
| Sujeto 6 | 184 | 98 | 62 | 981 |
| Sujeto 7 | -- | 88 | 45 | 51 |
| Sujeto 8 | 164 | 87 | 34 | 115 |
| Sujeto 9 | 132 | 224 | 64 | 34 |
| Sujeto 10 | 189 | 79 | 26 | 49 |
| Sujeto 11 | 65 | 69 | 20 | 149 |
| Sujeto 12 | 118 | 157 | 93 | 69 |
| Mediana | 148 | 113 | 39 | 84 |

Con dichos tiempos se calculó el percentil 95 para obtener el tiempo máximo admisible de ejecución de cada tarea (ver Tabla 6-13). Una vez obtenidos los tiempos máximos admisibles de ejecución para cada tarea, se estimó cuánto tardarán los futuros usuarios en ejecutarlas. Para ello, usando los tiempos de

ejecución de todos los participantes (Tabla 6-12), se calcularon los intervalos de confianza al 95% que se muestran en la Tabla 6-13.

Tabla 6-13.- Intervalos de confianza al 95% y máximo admisible para el tiempo de ejecución de las tareas

| | Tiempo máximo admisible de ejecución (Percentil 95) | Límite inferior | Límite superior |
|--------|--|------------------------|------------------------|
| CPU0_A | 187.25 | 118 | 186 |
| CPU1_A | 242.20 | 85 | 150 |
| CPU2_A | 63.30 | 27 | 56 |
| CPU3_A | 772.75 | 43 | 164 |

Todos los límites superiores de los intervalos de confianza son inferiores al tiempo máximo admisible de ejecución de las tareas por lo que se puede afirmar con una confianza del 95% que los usuarios serán capaces de completar las tareas en un tiempo razonable. La tarea CPU0_A es la que presenta unos resultados más ajustados, ya que el límite superior de su intervalo de confianza se aproxima mucho al tiempo máximo admisible.

La segunda métrica usada para medir la eficiencia fue el nivel de pérdida de los participantes. Mediante la ecuación (6-8) se calculó el nivel de pérdida de cada participante en cada una de las tareas. La Tabla 6-14 recoge los resultados obtenidos.

Al analizar los niveles de pérdida, sólo deben tenerse en cuenta los valores iguales o superiores a 0.4. Por este motivo, únicamente se considera que se perdieron durante la ejecución el sujeto 6 en la tarea CPU3_A y el sujeto 9 en la tarea CPU2_A (elementos sombreados en la Tabla 6-14). El resto de datos entra dentro de la normalidad, destacando la tarea CPU2_A donde, a excepción del sujeto 9, todos los participantes realizaron la tarea de forma óptima.

Tabla 6-14.- Nivel de pérdida de los participantes

| | CPU0_A | CPU1_A | CPU2_A | CPU3_A |
|--------------|---------------|---------------|---------------|---------------|
| Sujeto 1 | 0.15 | 0.37 | 0 | 0 |
| Sujeto 2 | 0.11 | 0 | 0 | 0 |
| Sujeto 3 | 0 | 0 | 0 | 0 |
| Sujeto 4 | 0.11 | 0 | 0 | 0 |
| Sujeto 5 | 0.38 | 0.29 | 0 | 0 |
| Sujeto 6 | 0.15 | 0.17 | 0 | 0.49 |
| Sujeto 7 | 0.29 | 0 | 0 | 0 |
| Sujeto 8 | 0.15 | 0 | 0 | 0 |
| Sujeto 9 | 0.27 | 0.29 | 0.73 | 0 |
| Sujeto 10 | 0 | 0.17 | 0 | 0 |
| Sujeto 11 | 0 | 0 | 0 | 0.35 |
| Sujeto 12 | 0 | 0 | 0 | 0 |
| Media | 0.13 | 0.11 | 0.06 | 0.07 |

Con los niveles de pérdida obtenidos para cada tarea (ver Tabla 6-14), se calculó el intervalo de confianza al 95% para poder estimar cuál será el nivel de pérdida de los usuarios futuros. La Tabla 6-15 muestra los límites inferior y superior de los intervalos de confianza correspondientes a cada una de las tareas del estudio.

Tabla 6-15.- Intervalos de confianza al 95% para el nivel de pérdida de las tareas

| | Límite inferior | Límite superior |
|--------|------------------------|------------------------|
| CPU0_A | 0.054 | 0.215 |
| CPU1_A | 0.017 | 0.198 |
| CPU2_A | 0* | 0.195 |
| CPU3_A | 0* | 0.176 |

*Valores negativos

Como el nivel de pérdida no puede tener valores negativos, los límites inferiores de los intervalos de confianza de CPU2_A y CPU3_A, se interpretan como 0. A la vista de que el límite superior del intervalo de confianza para todas las tareas es inferior a 0.4, se puede afirmar con un 95% de confianza que los nuevos usuarios no

se perderán para realizar ninguna de las tareas de administración contempladas en el estudio.

Aunque podamos afirmar con un nivel alto de confianza que los nuevos usuarios no se van a perder, el nivel de pérdida de los participantes para la tarea CPU0_A (añadir criterios de evaluación) es el que más valores distintos de 0 presenta (ver Tabla 6-14). Además, esa misma tarea es la que presentaba un límite inferior para el intervalo de confianza de la tasa de éxito *sin problemas* especialmente bajo (ver Tabla 6-11) y un límite superior para el tiempo de ejecución de la tarea muy cercano al tiempo máximo admisible (ver Tabla 6-13).

Teniendo en cuenta todos estos indicadores se deduce que CPU0_A es la tarea que más problemas generó a los participantes en el estudio. Por ello, esta tarea se analizó de manera más detallada mediante el grafo con los datos de navegación que se muestra en la Fig. 6-6.

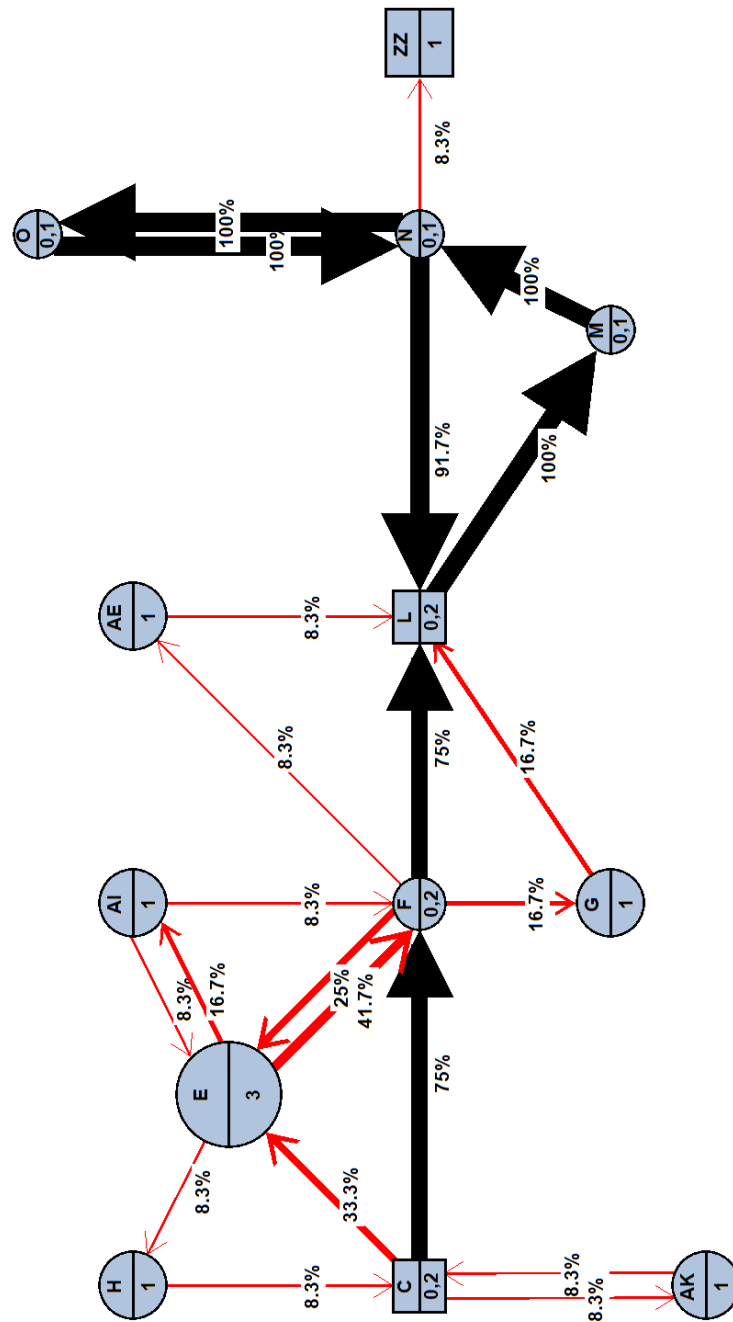


Fig. 6-6.- Grafo con la navegación de la tarea CPU0_A

En la figura se observa que los errores de navegación (flechas rojas de cabeza abierta) se producen básicamente en el camino desde el nodo inicial para esta tarea, la página de inicio del sistema para el rol Persona Administradora (nodo C), hasta la página que muestra los criterios de evaluación definidos (nodo L) y por la que había que pasar en dos ocasiones. Tras llegar al nodo L por primera vez, todos los participantes fueron capaces de continuar la ejecución correctamente, volviendo al nodo L que era a su vez el nodo final. El único error que se produjo en esa parte de la tarea fue el abandono por parte de uno de los participantes, reflejado en el grafo con el nodo ZZ.

La tarea CPU0_A consiste en añadir los criterios de evaluación para un elemento del TFG, lo que implica la definición de una rúbrica de evaluación con todos sus elementos: dimensiones, escala, etc. La falta de experiencia en el trabajo con rúbricas de evaluación por parte de los participantes, podría justificar el peor resultado en la ejecución de esta tarea en comparación con el resto.

Satisfacción

La satisfacción de los participantes con el sistema se analizó a partir de los cuestionarios SUS que rellenaron los usuarios tras ejecutar las tareas. Las respuestas de los cuestionarios se analizaron y la puntuación correspondiente a cada participante se calculó usando la ecuación (6-9).

La Tabla 6-16 muestra las puntuaciones obtenidas para cada participante y la media aritmética de las mismas.

El intervalo de confianza al 95% para el SUS correspondiente a la perspectiva de administración de TraMeA calculado a partir de las puntuaciones de la Tabla 6-16, es [70.096, 89.453]. Esto implica que, como mínimo, la usabilidad del sistema podrá categorizarse como muy aceptable o casi buena (la buena usabilidad se considera a partir de los 71 puntos). Por su parte, el límite superior del intervalo está por encima de la puntuación necesaria para considerar que el sistema tiene una excelente usabilidad (85 puntos). Por lo tanto, se puede concluir que con un 95% de probabilidad, el nivel de satisfacción de los usuarios respecto a la usabilidad de la perspectiva de administración de TraMeA será bueno o excelente.

Tabla 6-16.- Puntuaciones del SUS

| | Puntuaciones |
|--------------|---------------------|
| Sujeto 1 | 92.5 |
| Sujeto 2 | 82.5 |
| Sujeto 3 | 70 |
| Sujeto 4 | 90 |
| Sujeto 5 | 65 |
| Sujeto 6 | 77.5 |
| Sujeto 7 | 80 |
| Sujeto 8 | 45 |
| Sujeto 9 | 82.5 |
| Sujeto 10 | 95 |
| Sujeto 11 | 87.5 |
| Sujeto 12 | 95 |
| Media | 80.21 |

6.2.5.2 Supervisión/Evaluación

La perspectiva de supervisión/evaluación de TraMeA se evaluó siguiendo los mismos tres aspectos de usabilidad descritos anteriormente: efectividad, eficiencia y satisfacción. A continuación se enumeran los resultados obtenidos en cada una de las métricas empleadas.

Efectividad

El análisis de la efectividad se realizó en base a la tasa de éxito de las tareas. En la Tabla 6-17 se muestran los resultados recogidos sobre cada una de las tareas. Se representa para cada sujeto participante en el estudio, si finalizó la tarea correctamente (valor «SI»), o no (valor «NO»). A partir de esos datos, la tabla muestra también el porcentaje de tareas completadas con éxito por cada sujeto y la tasa de éxito de cada tarea.

Tabla 6-17.- Porcentaje de tareas completadas con éxito y su tasa de éxito muestral

| | CPU4_I_Tr | CPU5_I_Tr | CPU6_I_Tr | CPU7_I_Tr | CPU8_I | CPU9_T | CPU10_I | CPU11_I_Tr | CPU12_I_Tr | Porcentaje de tareas completadas con éxito |
|-------------------------------|-----------|-----------|-----------|-----------|--------|--------|---------|------------|------------|--|
| Sujeto 1 | SI | SI | SI | SI | SI | SI | SI | SI | SI | 100% |
| Sujeto 2 | SI | SI | SI | SI | SI | SI | SI | SI | SI | 100% |
| Sujeto 3 | SI | SI | SI | SI | SI | SI | SI | SI | SI | 100% |
| Sujeto 4 | SI | SI | SI | SI | SI | SI | SI | SI | SI | 100% |
| Sujeto 5 | SI | SI | SI | SI | SI | SI | SI | SI | SI | 100% |
| Sujeto 6 | SI | SI | SI | SI | SI | SI | SI | SI | SI | 100% |
| Sujeto 7 | SI | SI | SI | SI | SI | SI | SI | SI | SI | 100% |
| Sujeto 8 | SI | SI | SI | SI | SI | SI | NO | SI | SI | 88.89% |
| Sujeto 9 | SI | SI | SI | SI | SI | NO | NO | SI | SI | 77.78% |
| Sujeto 10 | SI | SI | SI | SI | SI | SI | SI | SI | SI | 100% |
| Sujeto 11 | SI | SI | SI | SI | SI | SI | SI | SI | SI | 100% |
| Sujeto 12 | SI | SI | SI | SI | SI | SI | NO | SI | SI | 88.89% |
| Tasa de éxito muestral | 1 | 1 | 1 | 1 | 1 | 0.917 | 0.750 | 1 | 1 | |

El 75% de los participantes fueron capaces de finalizar correctamente el 100% de las tareas y el 25% restante fue capaz de realizar correctamente al menos el 77% de las tareas.

A excepción de las tareas CPU9_T y CPU10_T, con una tasa de éxito de 0.92 y 0.75 respectivamente, el resto de tareas poseen una tasa de éxito de 1.

Para aquellas tasas de éxito muestrales superiores a 0.9 (todas menos CPU10_T), el método Laplace (ver ecuación (6-3)) es el más preciso para estimar la tasa de éxito poblacional. Sin embargo, para la tarea CPU10_T (cuya tasa de éxito muestral es 0.75) el estimador más preciso es la propia tasa de éxito muestral. En la Tabla 6-18 se muestran las tasas de éxito poblacionales estimadas para cada una de las tareas de la perspectiva de supervisión/evaluación.

Tabla 6-18.- Estimación de las tasas de éxito poblacional

| | Estimación tasa de éxito poblacional |
|------------|---|
| CPU4_T_Tr | 0.929 |
| CPU5_T_Tr | 0.929 |
| CPU6_T_Tr | 0.929 |
| CPU7_T_Tr | 0.929 |
| CPU8_T | 0.929 |
| CPU9_T | 0.857 |
| CPU10_T | 0.750 |
| CPU11_T_Tr | 0.929 |
| CPU12_T_Tr | 0.929 |

Todas las tasas de éxito poblacionales estimadas superan el estándar establecido del 70%, lo que significa que la mayoría de usuarios serán capaces de realizar las tareas propuestas con éxito.

Para profundizar en el análisis de las tasas de éxito se calcularon los intervalos de confianza ajustados de Wald al 95% para cada una de las tareas. Los resultados obtenidos se muestran en la Tabla 6-19.

Tabla 6-19.- Intervalos de confianza ajustados de Wald al 95% (éxito/fracaso)

| | Límite inferior | Límite superior |
|------------|-----------------|-----------------|
| CPU4_T_Tr | 0.784 | 1 |
| CPU5_T_Tr | 0.784 | 1 |
| CPU6_T_Tr | 0.784 | 1 |
| CPU7_T_Tr | 0.784 | 1 |
| CPU8_T | 0.784 | 1 |
| CPU9_T | 0.625 | 0.999 |
| CPU10_T | 0.462 | 0.917 |
| CPU11_T_Tr | 0.784 | 1 |
| CPU12_T_Tr | 0.784 | 1 |

El intervalo de confianza para la tasa de éxito de todas las tareas, excepto los de CPU9_T y CPU10_T, presentan un límite inferior de 0.784, lo que implica que aún en el peor de los casos, el 78.40% de los usuarios resolverá dichas tareas con éxito. En el caso de la tarea CPU9_T, reduciendo la confianza del intervalo al 85%, se obtiene un límite inferior de 0.720, por lo que se puede afirmar con una confianza del 85% que dicha tarea también superará el 70% de éxito establecido como mínimo aceptable.

El intervalo de confianza de la tarea CPU10_T, a pesar de tener un límite superior bastante alto, su límite inferior es bajo, por lo que habrá que prestar especial atención a sus resultados en el resto de métricas.

En la Tabla 6-20 se muestran los resultados de la ejecución de las tareas teniendo en cuenta los niveles de éxito en la finalización (*sin problemas, con problemas, abandono*). La categoría *sin problemas* hace referencia exclusivamente a las ejecuciones óptimas, en las que no se cometió ni un solo error. La importancia de los errores cometidos dentro de la categoría *con problemas* se tratará en la sección de *Eficiencia* a través del nivel de pérdida de los usuarios.

En la tabla también se muestran la tasa de éxito muestral de las tareas y el porcentaje de tareas completadas por cada usuario para la categoría *sin problemas* (columna %), que supera el 50% para todos los participantes en el estudio.

Tabla 6-20.- Resultados en función del nivel de finalización

| | CPU4_I_Tr | CPU5_I_Tr | CPU6_I_Tr | CPU7_I_Tr | CPU8_I | CPU9_I | CPU10_I | CPU11_I_Tr | CPU12_I_Tr | % |
|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|--------|
| Sujeto 1 | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | 100% |
| Sujeto 2 | Sin problemas | Con problemas | Sin problemas | Sin problemas | Con problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | 77.78% |
| Sujeto 3 | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Con problemas | Sin problemas | Sin problemas | 88.89% |
| Sujeto 4 | Sin problemas | Sin problemas | Con problemas | Sin problemas | Sin problemas | Sin problemas | Con problemas | Con problemas | Sin problemas | 66.67% |
| Sujeto 5 | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Con problemas | Sin problemas | Con problemas | Sin problemas | Sin problemas | 77.78% |
| Sujeto 6 | Sin problemas | Sin problemas | Con problemas | Sin problemas | Con problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | 77.78% |
| Sujeto 7 | Sin problemas | Sin problemas | Con problemas | Sin problemas | Con problemas | Con problemas | Con problemas | Sin problemas | Sin problemas | 55.56% |
| Sujeto 8 | Sin problemas | Sin problemas | Con problemas | Sin problemas | Con problemas | Sin problemas | Abandono | Sin problemas | Sin problemas | 66.67% |
| Sujeto 9 | Sin problemas | Sin problemas | Sin problemas | Con problemas | Sin problemas | Abandono | Abandono | Sin problemas | Sin problemas | 66.67% |
| Sujeto 10 | Sin problemas | Sin problemas | Con problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | 88.89% |
| Sujeto 11 | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Con problemas | Sin problemas | Sin problemas | Sin problemas | Sin problemas | 88.89% |
| Sujeto 12 | Sin problemas | Sin problemas | Sin problemas | Sin problemas | Con problemas | Sin problemas | Abandono | Sin problemas | Sin problemas | 77.78% |
| Tasa de éxito muestral (sin problemas) | 1 | 0.917 | 0.580 | 0.917 | 0.417 | 0.830 | 0.417 | 0.917 | 1 | |

El 83% de las tareas presentan una tasa de éxito superior a 0.5. Debido a la variabilidad de las tasas de éxito entre las distintas tareas, el método más adecuado para estimar la tasa de éxito de la población general es distinto para cada una de ellas (ver sección 6.2.3). En la Tabla 6-21 se muestran las tasas de éxito estimadas para cada una de las tareas y el método empleado en su cálculo.

Tabla 6-21.- Estimación de las tasas de éxito poblacionales y método de cálculo

| | Estimación tasa de éxito poblacional | Método de cálculo |
|------------|---|--------------------------|
| CPU4_T_Tr | 0.929 | Laplace |
| CPU5_T_Tr | 0.852 | Laplace |
| CPU6_T_Tr | 0.583 | Tasa de éxito muestral |
| CPU7_T_Tr | 0.852 | Laplace |
| CPU8_T | 0.437 | Wilson |
| CPU9_T | 0.833 | Tasa de éxito muestral |
| CPU10_T | 0.437 | Wilson |
| CPU11_T_Tr | 0.852 | Laplace |
| CPU12_T_Tr | 0.929 | Laplace |

Seis de las nueve tareas planteadas presentan una tasa estimada de éxito superior al 80%, lo que significa que más del 80% de los usuarios las resolverán de manera óptima la primera vez que se enfrenten a ellas. Las tareas CPU8_T y CPU10_T presentan una estimación de la tasa de éxito por debajo del 50% por lo que habrá que analizarlas más en profundidad para comprender las razones de estos datos.

En la Tabla 6-22 se recogen los intervalos de confianza ajustados de Wald al 95%, esto es, los límites entre los que se encontrará la tasa de éxito para la categoría *sin problemas* de las distintas tareas.

Tabla 6-22.- Intervalos de confianza ajustados de Wald al 95% (sin problemas)

| | Límite inferior | Límite superior |
|------------|------------------------|------------------------|
| CPU4_T_Tr | 0.784 | 1 |
| CPU5_T_Tr | 0.625 | 0.999 |
| CPU6_T_Tr | 0.319 | 0.807 |
| CPU7_T_Tr | 0.625 | 0.999 |
| CPU8_T | 0.193 | 0.681 |
| CPU9_T | 0.540 | 0.965 |
| CPU10_T | 0.193 | 0.681 |
| CPU11_T_Tr | 0.625 | 0.999 |
| CPU12_T_Tr | 0.784 | 1 |

A la vista de los límites inferiores de los intervalos, se puede asegurar con una confianza del 95% que más de la mitad de los usuarios resolverán sin ningún tipo de error cualquiera de las tareas planteadas, a excepción de CPU8_T y CPU10_T que son las que presentan unos valores más ajustados y por lo tanto requerirán una mayor atención a lo largo de la evaluación.

Eficiencia

Para analizar la eficiencia de la perspectiva de supervisión/evaluación de TraMeA, se usaron los tiempos de ejecución de las tareas y el estudio del camino seguido por los participantes durante la resolución de las mismas.

En la Tabla 6-23 se recoge el tiempo, en segundos, que empleó cada participante en resolver cada una de las tareas planteadas y para cada tarea la mediana de los tiempos empleados redondeada al segundo superior. En aquellos casos en los que se terminó abandonando la tarea, no se ha tenido en cuenta el tiempo transcurrido.

Tabla 6-23.- Tiempos de ejecución de las tareas (segundos)

| | CPU4_I_Tr | CPU5_I_Tr | CPU6_I_Tr | CPU7_I_Tr | CPU8_I | CPU9_I | CPU10_I | CPU11_I_Tr | CPU12_I_Tr |
|----------------|-----------|-----------|-----------|-----------|--------|--------|---------|------------|------------|
| Sujeto 1 | 30 | 13 | 13 | 21 | 60 | 26 | 19 | 45 | 23 |
| Sujeto 2 | 45 | 90 | 211 | 31 | 50 | 20 | 25 | 60 | 30 |
| Sujeto 3 | 9 | 14 | 29 | 113 | 47 | 25 | 57 | 50 | 26 |
| Sujeto 4 | 4 | 6 | 40 | 17 | 23 | 29 | 203 | 60 | 85 |
| Sujeto 5 | 68 | 12 | 15 | 27 | 85 | 17 | 66 | 40 | 15 |
| Sujeto 6 | 8 | 17 | 82 | 23 | 176 | 31 | 21 | 104 | 19 |
| Sujeto 7 | 7 | 11 | 206 | 55 | 351 | 95 | 358 | 49 | 35 |
| Sujeto 8 | 4 | 16 | 41 | 208 | 128 | 39 | -- | 48 | 22 |
| Sujeto 9 | 7 | 7 | 24 | 36 | 39 | -- | -- | 29 | 43 |
| Sujeto 10 | 5 | 16 | 101 | 16 | 30 | 69 | 24 | 44 | 19 |
| Sujeto 11 | 5 | 5 | 40 | 15 | 209 | 14 | 15 | 86 | 85 |
| Sujeto 12 | 8 | 10 | 28 | 25 | 243 | 26 | -- | 33 | 23 |
| Mediana | 11 | 13 | 46 | 34 | 85 | 31 | 47 | 51 | 31 |

Los tiempos que están sombreados en la Tabla 6-23 son los correspondientes a los participantes que puntuaron el sistema con más de 70 puntos (ver sección *Satisfacción*). Calculando el percentil 95 de dichos tiempos se obtuvo el tiempo máximo admisible de ejecución de cada tarea.

La Tabla 6-24 recoge los tiempos máximos admisibles de ejecución de cada tarea junto con los intervalos de confianza en los que al 95% de probabilidad se encontrará el tiempo que tardará un nuevo usuario en resolver cada tarea. Los intervalos de confianza se calcularon considerando todos los tiempos recogidos en la Tabla 6-23.

Tabla 6-24.- Intervalos de confianza al 95% y máximo admisible para el tiempo de ejecución de las tareas

| | Tiempo máximo admisible de ejecución (Percentil 95) | Límite inferior | Límite superior |
|------------|--|------------------------|------------------------|
| CPU4_T_Tr | 58.8 | 6 | 19 |
| CPU5_T_Tr | 60.8 | 8 | 21 |
| CPU6_T_Tr | 167 | 26 | 83 |
| CPU7_T_Tr | 170 | 20 | 56 |
| CPU8_T | 216.2 | 48 | 149 |
| CPU9_T | 57 | 21 | 45 |
| CPU10_T | 161.9 | 20 | 110 |
| CPU11_T_Tr | 86.4 | 41 | 64 |
| CPU12_T_Tr | 63 | 21 | 43 |

El tiempo máximo admisible de ejecución de todas las tareas es considerablemente superior al respectivo límite superior del intervalo de confianza. Esto permite asegurar con una confianza del 95% que un usuario sin experiencia en el sistema, será capaz de terminar el 100% de las tareas en un tiempo razonable.

La eficiencia del sistema también se midió a través del nivel de pérdida de los participantes y del camino seguido durante la ejecución de cada tarea. La Tabla 6-25 recoge el nivel de pérdida de cada participante en cada tarea calculado mediante la ecuación (6-8) (ver sección 6.2.3).

Tabla 6-25.- Nivel de pérdida de los participantes

| | CPU4_I_Tr | CPU5_I_Tr | CPU6_I_Tr | CPU7_I_Tr | CPU8_T | CPU9_T | CPU10_T | CPU11_I_Tr | CPU12_I_Tr |
|--------------|-----------|-----------|-----------|-----------|--------|--------|---------|------------|------------|
| Sujeto 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Sujeto 2 | 0 | 0.4 | 0 | 0 | 0.35 | 0 | 0 | 0 | 0 |
| Sujeto 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0.32 | 0 | 0 |
| Sujeto 4 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0.6 | 0.3 | 0 |
| Sujeto 5 | 0 | 0 | 0 | 0 | 0.25 | 0 | 0.54 | 0 | 0 |
| Sujeto 6 | 0 | 0 | 0.26 | 0 | 0.33 | 0 | 0 | 0 | 0 |
| Sujeto 7 | 0 | 0 | 0.46 | 0 | 0.19 | 0.48 | 0.84 | 0 | 0 |
| Sujeto 8 | 0 | 0 | 0.26 | 0 | 0.14 | 0 | 0.91 | 0 | 0 |
| Sujeto 9 | 0 | 0 | 0 | 0.26 | 0 | 0.72 | 0.85 | 0 | 0 |
| Sujeto 10 | 0 | 0 | 0.26 | 0 | 0 | 0 | 0 | 0 | 0 |
| Sujeto 11 | 0 | 0 | 0 | 0 | 0.58 | 0 | 0 | 0 | 0 |
| Sujeto 12 | 0 | 0 | 0 | 0 | 0.14 | 0 | 0.73 | 0 | 0 |
| Media | 0 | 0.03 | 0.12 | 0.02 | 0.17 | 0.1 | 0.4 | 0.03 | 0 |

Dado que sólo los valores iguales o superiores a 0.4 (sombreados en la Tabla 6-25) se consideran relevantes, todas las tareas, a excepción de CPU10_T, presentan un nivel medio de pérdida excelente. Destacan especialmente las tareas CPU4_T_Tr y CPU12_T_Tr cuyo nivel medio de pérdida es 0, lo que significa que todos los participantes fueron capaces de seguir el camino óptimo de resolución de la tarea sin desviarse del mismo.

A partir de los datos recogidos en la Tabla 6-25, se calcularon los intervalos de confianza al 95% de las distintas tareas. La Tabla 6-26 muestra los límites inferior y superior entre los que se encontrarán los niveles de pérdida de los futuros usuarios de la perspectiva de supervisión/evaluación de TraMeA para cada tarea.

Tabla 6-26.- Intervalos de confianza al 95% para el nivel de pérdida de las tareas

| | Límite inferior | Límite superior |
|------------|------------------------|------------------------|
| CPU4_T_Tr | 0 | 0 |
| CPU5_T_Tr | 0* | 0.107 |
| CPU6_T_Tr | 0.018 | 0.222 |
| CPU7_T_Tr | 0.169 | 0.264 |
| CPU8_T | 0.047 | 0.283 |
| CPU9_T | 0* | 0.252 |
| CPU10_T | 0.155 | 0.644 |
| CPU11_T_Tr | 0* | 0.080 |
| CPU12_T_Tr | 0 | 0 |

*Valores negativos

Los valores negativos en los límites inferiores se interpretan como 0, ya que el nivel de pérdida sólo puede tomar valores entre 0 y 1. Los límites superiores son, a excepción de para la tarea CPU10_T, sensiblemente inferiores al límite establecido de 0.4. Por lo tanto, se puede asegurar, con una confianza del 95%, que durante la ejecución de esas tareas los nuevos usuarios no se perderán.

A pesar de estos resultados positivos, se decidió analizar en detalle, mediante el uso de grafos con los datos de navegación, los niveles de pérdida de los participantes en aquellas tareas que habían dado algún resultado bajo en las métricas anteriores. Las tareas CPU8_T y CPU10_T habían presentado una tasa de éxito poblacional inferior al 50% (ver Tabla 6-21) y unos intervalos de confianza ajustados para la categoría *sin problemas* (ver Tabla 6-22), por lo que el análisis mediante grafos se centró en ellas.

En la Tabla 6-25 se puede observar cómo en la tarea CPU8_T (dar retroalimentación), sólo hubo un participante que se perdió (el sujeto 11), pero varios participantes también se equivocaron en algún momento ya que su nivel de pérdida es mayor que 0. La tarea CPU8_T consiste en proporcionar retroalimentación a un estudiante a través del foro de comunicación. El grafo que representa el camino recorrido por los participantes durante la ejecución (ver Fig. 6-7), muestra que los problemas en esta tarea (flechas de cabeza abierta) surgen en el paso desde la página de información del TFG (nodo Q), a la página con los foros de comunicación (nodo Y), donde más del 66% de los participantes se desviaron hacia la página con el historial de acciones de la persona supervisora (nodo T) y al repositorio de documentos de la persona supervisora (nodo W). Una vez llegado al nodo Y, el 100% de los participantes continuaron de manera correcta el camino de ejecución.

Aunque la tarea CPU8_T presentó problemas en varias métricas, analizando su grafo de navegación, se observa que se debieron a que llegados a un punto (la página de información del TFG, nodo Q), los participantes no sabían dónde encontrar el elemento sobre el cual debían proporcionar la retroalimentación. Este hecho queda justificado por el diseño de la comunicación entre profesorado y alumnado de TraMeA. En este diseño, todas las comunicaciones se realizan a través de foros de comunicación, por lo que para dar retroalimentación sobre un elemento concreto, hay que acceder a los foros de comunicación del sistema. La mayoría de los participantes buscaron el documento del estudiante por el sistema antes de percatarse de que debían acceder a los foros. Cuando se dieron cuenta que tenían que acceder a los foros de comunicación (nodo Y) fueron capaces de finalizar la tarea sin problemas.

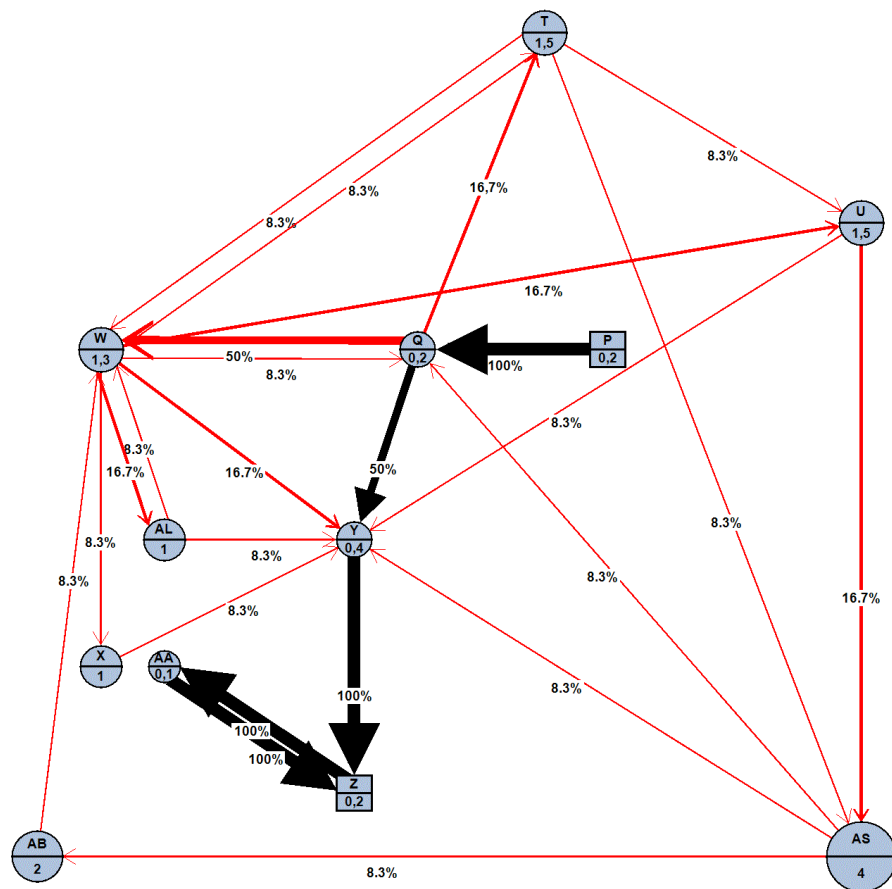


Fig. 6-7.- Grafo con la navegación de la tarea CPU8_T

Por su parte, la tarea CPU10_T (solicitar cierre) presenta un nivel medio de pérdida de 0.4, el valor considerado como límite aceptable. Esta tarea, cuyo grafo de navegación se muestra en la Fig. 6-8, consiste en, ejerciendo el rol de persona supervisora, solicitar el cierre del mismo. Para ello, a partir de la página de inicio de TraMeA para personas supervisoras (nodo P), hay que llegar a la página de información del TFG (nodo Q), y una vez allí, seleccionar la opción «cierre de TFG».

navegar por el sistema buscando cómo cerrar el TFG. Tres de los participantes desistieron en su búsqueda y abandonaron la tarea (nodo ZZ).

Satisfacción

A través de los cuestionarios SUS que los participantes en el estudio rellenaron tras finalizar la prueba, se calculó la satisfacción de los participantes con el sistema.

La Tabla 6-27 recoge las puntuaciones computadas aplicando la ecuación (6-9) a las respuestas de cada participante y la media aritmética de las mismas.

Tabla 6-27.- Puntuaciones del SUS

| | Puntuaciones |
|--------------|---------------------|
| Sujeto 1 | 95 |
| Sujeto 2 | 82.5 |
| Sujeto 3 | 95 |
| Sujeto 4 | 77.5 |
| Sujeto 5 | 92.5 |
| Sujeto 6 | 92.5 |
| Sujeto 7 | 62.5 |
| Sujeto 8 | 75 |
| Sujeto 9 | 47.5 |
| Sujeto 10 | 87.5 |
| Sujeto 11 | 52.5 |
| Sujeto 12 | 90 |
| Media | 79.17 |

A partir de las puntuaciones de la Tabla 6-27 se calculó el intervalo de confianza al 95% para el SUS que resultó ser [66.657, 91.676]. El intervalo calculado estima que la usabilidad del sistema podrá categorizarse como mínimo de manera muy aceptable. Por su parte, el límite superior del intervalo está por encima de la puntuación necesaria para considerar que el sistema tiene la mejor usabilidad posible (90 puntos). Por lo tanto, se puede concluir que con un 95% de probabilidad, el nivel de satisfacción con la usabilidad de la perspectiva de supervisión/evaluación de TraMeA será bueno o excelente.

6.2.5.3 Resumen de la evaluación de usabilidad

Los resultados obtenidos en las distintas métricas tanto para la perspectiva de administración como para la de supervisión/evaluación de TraMeA son excelentes, cumpliendo holgadamente con los mínimos establecidos en cada una de las métricas.

Destacan las tareas CPU2_A y CPU3_A de la perspectiva de administración y CPU4_T_Tr y CPU12_T_Tr de la de supervisión/evaluación por los resultados especialmente positivos que han obtenido en todas las métricas a lo largo de la evaluación.

Las métricas relativas a la efectividad en ambas perspectivas de TraMeA presentan muy buenos resultados incluso cuando se tienen en cuenta únicamente a los participantes que realizaron las tareas de manera óptima, sin el más mínimo error.

Respecto a la eficiencia del sistema, los resultados en ambas perspectivas son muy positivos ya que todas las tareas presentan unos tiempos de ejecución razonables y unos niveles de pérdida muy bajos que indican que los futuros usuarios del sistema no se perderán durante su uso.

En la perspectiva de administración los resultados más ajustados se han obtenido para la tarea CPU0_A. Esto puede estar justificado por la falta de experiencia previa de los usuarios en la definición de rúbricas de evaluación. Para mejorar los resultados para esta tarea y a la vista del grafo de navegación, parece recomendable rediseñar la interfaz de TraMeA correspondiente a la definición de rúbricas de evaluación o proporcionar a los usuarios algún tipo de guía que les facilite la ejecución de dicha tarea.

En cuanto a la perspectiva de supervisión/evaluación, las tareas con los resultados más moderados fueron las tareas CPU8_T y CPU10_T. En la tarea CPU8_T, los usuarios se perdieron por no recordar que para comunicarse con los estudiantes había que acceder a los foros de comunicación. Mientras que en la tarea CPU10_T, los participantes tuvieron problemas para encontrar la opción de cierre del TFG a pesar de haber llegado correctamente a la página donde ésta se encuentra.

En ambos casos, para solucionar los problemas parece aconsejable realizar algún tipo de formación previa a los futuros usuarios que les facilite los conocimientos básicos sobre el funcionamiento general de TraMeA. Asimismo, en el caso de la tarea CPU10_T, la modificación de la interfaz gráfica para ubicar la opción de cierre de un TFG en un lugar que facilite su localización y la haga más visible también se tendrá en cuenta.

La valoración de la satisfacción de los usuarios con la usabilidad de TraMeA, en cualquiera de las dos perspectivas, también presenta resultados muy positivos, ya que se sitúa entre las categorías muy buena y excelente.

A la vista de los resultados en todos los aspectos de usabilidad evaluados se puede concluir que, salvo los pequeños ajustes de la interfaz gráfica mencionados, el diseño de TraMeA es bueno y los usuarios podrán utilizar el sistema podrá para lograr objetivos concretos con efectividad, eficiencia y satisfacción (ISO, 1998).

6.3 Conclusiones

En este capítulo se han presentado las evaluaciones de las propuestas planteadas en este trabajo de tesis que se han realizado.

Se ha evaluado la metodología FPM y se ha comprobado que los elementos evaluables definidos en la misma son los más adecuados para calcular la calificación final del TFG de forma admisible, esto es, con un error inferior a 0.5 puntos. Además, los resultados también indican que el hecho de incorporar la opinión de la persona supervisora a la evaluación del TFG mejora el cálculo de la calificación, lo que avala la metodología definida.

Asimismo, los resultados de la evaluación respaldan la decisión de usar rúbricas como instrumento de evaluación, ya que ha quedado demostrado que aportan objetividad a la evaluación y el alumnado ha valorado muy positivamente su uso. El 100% del alumnado encuestado opina que el uso de rúbricas clarifica los criterios de evaluación.

La plataforma TraMeA se ha evaluado a través de un estudio de usabilidad cuyos resultados permiten concluir que el diseño de la plataforma es adecuado y que presenta un alto grado de usabilidad.

Al realizar la evaluación de TraMeA mediante un prototipo funcional, la funcionalidad aportada por AdESMuS se ha evaluado simultáneamente. AdESMuS ha proporcionado sin inconvenientes todos los servicios de evaluación requeridos por TraMeA, por lo que su evaluación es plenamente satisfactoria.

7

Conclusiones y líneas futuras

Este capítulo presenta las principales conclusiones de este trabajo de tesis. Tras un resumen del trabajo realizado, se destacan sus principales resultados y contribuciones. Finalmente, se describen las líneas futuras de investigación derivadas del presente trabajo.

7.1 Resumen

Desde la implantación del Espacio Europeo de Educación Superior (EEES), el Trabajo Fin de Grado (TFG) es una asignatura obligatoria en todos los programas de grado en el sistema universitario español. La realización y evaluación de un TFG es un reto para todos los actores involucrados: estudiantes, supervisores y evaluadores.

En este contexto, el objetivo general establecido para este trabajo de tesis se ha definido como (ver sección 1.1.2):

«Analizar los problemas que pueda plantear tanto para el profesorado como para el alumnado la realización y evaluación de un TFG y establecer una metodología que evite o disminuya la aparición de dichos problemas y minimice su impacto. Se abordará la definición de los procesos, documentación y herramientas informáticas que den soporte a dicha metodología y la evaluación tanto de la metodología, como de los elementos definidos y los resultados de su aplicación».

A partir de este objetivo principal, se ha llevado a cabo un estudio con los actores implicados en el proceso de realización y evaluación del TFG para identificar, analizar y dar respuesta a los principales problemas que surgen en estos

procesos (Villamañe, Ferrero, Álvarez, et al., 2014). Los principales problemas que se han identificado en este estudio son: la falta de criterios para evaluar el alcance de los TFG, la escasez de canales de comunicación entre actores, la falta de un proceso continuo de evaluación y la ausencia de criterios de evaluación objetivos y públicos.

Una vez identificada la problemática en la realización y evaluación de los TFG (ver sección 1.2), se establecieron los objetivos específicos a cubrir para alcanzar el objetivo general (ver sección 1.3):

- OE—1: Asegurar una calidad mínima en todos los TFG.
- OE—2: Aumentar y mejorar la retroalimentación que recibe el alumnado.
- OE—3: Incorporar nuevos canales y posibilidades de comunicación.
- OE—4: Establecer un proceso de evaluación continuo.
- OE—5: Establecer un método de evaluación equitativo.
- OE—6: Desarrollar las herramientas informáticas necesarias.

Para cubrir estos objetivos, se ha propuesto la metodología FPM (*Final-year Project Mentoring*) para la realización y evaluación del TFG (ver sección 2.4.1) y se ha definido la plataforma TraMeA (*Tracking Mentoring and Assessment*, capítulo 3) para dar soporte tecnológico a las características de la metodología FPM.

La metodología FPM establece un proceso de evaluación continua basada en elementos evaluables según unos criterios públicos y objetivos, promoviendo los comentarios de retroalimentación por parte de la persona supervisora en los momentos clave de la realización de un TFG.

Con el ánimo de especificar unos criterios precisos y públicos, y así conseguir una evaluación objetiva, en la metodología se establece el uso de rúbricas como instrumento de evaluación. En base a la experiencia de un grupo de expertos y a la bibliografía relacionada, se ha definido un conjunto inicial de rúbricas para evaluar los elementos establecidos en la metodología (capítulo 5). Este conjunto inicial se ha ido refinando a través de un proceso iterativo, en el que han participado tanto docentes como estudiantes, hasta conseguir un conjunto de rúbricas que satisface tanto al alumnado como al profesorado.

Para garantizar el éxito de la metodología FPM, es necesario proporcionar el soporte tecnológico que facilite su aplicación. Para ello, se han determinado diez requisitos funcionales que una plataforma informática debe cumplir para dar soporte a la metodología FPM (ver sección 2.3.1). Estos requisitos hacen referencia, entre otros, a la necesidad de poder combinar distintas rúbricas de evaluación para

obtener una calificación y a la demanda de que la plataforma cuente con canales que faciliten la comunicación entre el alumnado.

Para cubrir los requisitos funcionales determinados en este trabajo, se ha definido la plataforma TraMeA (*Tracking Mentoring and Assessment*, capítulo 3). TraMeA se ha desarrollado tomando como base el sistema Redmine y ampliando sus funcionalidades en los aspectos relacionados con la evaluación.

La ampliación realizada en TraMeA en relación a los aspectos de evaluación, se ha generalizado para dar respuesta también a las necesidades de mejora de procesos de evaluación en entornos de aprendizaje combinados o *blended learning* (Graham, 2005).

En los entornos de aprendizaje combinados, el proceso de evaluación debe tener en cuenta tanto las tareas que los estudiantes realizan a través de los sistemas de aprendizaje on-line, como las que realizan en el aula de manera tradicional. Además, el proceso de evaluación puede involucrar distintos instrumentos y técnicas de evaluación, tales como el uso de rúbricas o la evaluación por pares para cada uno de los elementos evaluables. La gestión de un proceso de evaluación con esas características puede ser complicada. Para dar respuesta a las necesidades de evaluación en entornos *blended learning* se ha diseñado e implementado del sistema AdESMuS (*Adaptable Evaluation System using Multiple Sources*, capítulo 4).

AdESMuS permite definir la evaluación de una asignatura como una combinación de elementos evaluables, cada uno de los cuales puede ser calificado por una o varias personas usando distintas fuentes (instrumentos de evaluación). Para ello, entre otras funcionalidades, AdESMuS permite la definición de rúbricas de evaluación y de los distintos elementos que ello conlleva. La versatilidad que ofrece AdESMuS para definir la evaluación implica la toma de muchas decisiones que no siempre son sencillas, por ejemplo, decidir qué porcentaje de influencia tiene en la calificación final de la asignatura cada uno de los elementos evaluables o el peso que cada dimensión tiene en cada rúbrica.

Para facilitar esta toma de decisiones al profesorado, AdESMuS se ha enriquecido con la posibilidad de ajustar las distintas evaluaciones e instrumentos que se definan. Para ello, se ha incluido en AdESMuS un módulo de ajuste y optimización que, mediante técnicas de análisis del aprendizaje (*Learning Analytics*) y aprendizaje automático, permite automatizar el ajuste de los distintos elementos que forman parte de la evaluación.

Finalmente, se han evaluado las distintas propuestas realizadas obteniendo resultados muy positivos.

7.2 Resultados y contribuciones

El objetivo principal de este trabajo de tesis requiere establecer una metodología para evitar y disminuir la aparición de problemas durante la realización y la evaluación de los TFG. Con este fin, se ha propuesto la metodología FPM y se ha evaluado satisfactoriamente su utilidad para solucionar los problemas detectados.

El desarrollo de las herramientas informáticas necesarias para dar soporte a la metodología definida también está contemplado en el objetivo general, por lo que se ha diseñado y desarrollado la plataforma TraMeA. Esta plataforma da soporte completo a la metodología FPM y su usabilidad ha sido evaluada satisfactoriamente.

Durante el desarrollo de TraMeA se decidió generalizar uno de sus módulos para desarrollar un sistema autónomo que proporcionara servicios de evaluación (AdESMuS). AdESMuS ha sido evaluado mediante su uso en un entorno real en combinación con TraMeA.

A continuación se presentan los principales resultados de este trabajo, de acuerdo a cada uno de los objetivos específicos establecidos en la sección 1.3.

7.2.1 OE –1: Asegurar una calidad mínima en todos los TFG.

La metodología FPM da respuesta al objetivo específico OE –1 definiendo una primera etapa en la que se evalúa la propuesta de alcance del TFG a realizar usando rúbricas de evaluación.

Al hacer públicas las rúbricas de evaluación utilizadas para valorar las propuestas, se establecen los criterios público que determinan qué aspectos se deben contemplar en el TFG. Teniendo en cuenta estos criterios públicos, el alumnado puede estimar si su propuesta de TFG cumple o no con los requisitos, fomentando que presenten sus propias propuestas y no se limiten sólo a realizar alguna de las propuestas definidas por el profesorado.

De este modo, el alumnado puede tener la seguridad de que el trabajo que va a realizar cumple los requisitos mínimos establecidos para la titulación y que no van a surgir problemas en la posterior evaluación si cumple con lo establecido en la propuesta aprobada.

7.2.2 OE – 2: Aumentar y mejorar la retroalimentación que recibe el alumnado.

Para dar respuesta al objetivo OE-2, la metodología FPM establece un proceso de evaluación continua y una serie de elementos evaluables que aseguran la retroalimentación en los momentos clave de la realización del TFG.

Además, el uso de rúbricas como instrumento de evaluación permite establecer de forma explícita las expectativas sobre los distintos elementos a evaluar. El uso de rúbricas permite también que se pueda mejorar la retroalimentación sin incrementar de manera notable el tiempo dedicado a la evaluación por parte del profesorado, lo que es esencial para que se adopte cualquier cambio en el proceso de evaluación (Reimann et al., 2016).

Las rúbricas permiten además asegurar la calidad de la retroalimentación independientemente de la persona que realice la evaluación.

7.2.3 OE – 3: Incorporar nuevos canales y posibilidades de comunicación.

Al establecer las rúbricas de evaluación como instrumento de evaluación en la metodología FPM y debido a sus características y potencial como herramienta de retroalimentación, se está incrementando la comunicación entre el profesorado y el alumnado, ayudando a satisfacer el objetivo OE–3.

Además, TraMeA incorpora nuevos canales de comunicación que permiten la comunicación entre estudiantes que compartan intereses y amplían las posibilidades de comunicación habituales.

Para ello, TraMeA hace uso de los foros de comunicación y del complemento de Slack para Redmine. Este último permite la creación de canales temáticos y centralizar la comunicación de todos los actores implicados en la realización de los TFG. Cuenta con versión para dispositivos móviles (lo que ofrece la posibilidad de comunicarse en cualquier situación) y además, Slack también se puede integrar con otras herramientas como Google Calendar, Gmail, etc. ampliando aún más sus posibilidades como canal de comunicación.

7.2.4 OE – 4: Establecer un proceso de evaluación continuo.

Para satisfacer el objetivo específico OE–4 la metodología FPM divide la elaboración del TFG en etapas y establece la generación de distintos elementos que

son evaluados por la persona supervisora del TFG o por las personas que integran el tribunal de evaluación.

De este modo, el TFG ya no se evalúa sólo tras su finalización, si no en varias ocasiones a lo largo de su realización, permitiendo que el alumnado pueda reconducir su trabajo en caso de no estar obteniendo los resultados esperados.

Además, el uso de rúbricas de evaluación permite mejorar la retroalimentación indicándole al alumnado cuáles son los objetivos no alcanzados y propuestas para mejorarlo (Stevens et al., 2012). Esta retroalimentación enriquecida permitirá que el alumnado reconduzca su trabajo de una manera adecuada.

7.2.5 OE – 5: Establecer un método de evaluación equitativo.

La metodología FPM establece el uso de rúbricas como instrumento de evaluación. De este modo, los criterios para evaluar los distintos elementos definidos en la metodología son públicos e iguales para todo el alumnado, independientemente de la persona que esté realizando la evaluación.

Además de explicitar los criterios de evaluación, el uso de rúbricas aporta objetividad (ver sección 6.1.2), mejorando la opinión que el alumnado tiene sobre la claridad y objetividad del proceso. Por lo tanto, el uso de rúbricas como instrumento de evaluación en la metodología FPM permite cubrir el objetivo específico OE-5.

7.2.6 OE – 6: Desarrollar las herramientas informáticas necesarias.

Para satisfacer el objetivo OE-6, se ha definido el conjunto de requisitos funcionales que debe satisfacer una plataforma que de soporte a la metodología FPM (ver sección 2.2.3).

Para cumplir con los requisitos funcionales establecidos, se ha definido TraMeA (capítulo 2), un sistema que permite el seguimiento y la evaluación de los TFG. En su desarrollo TraMeA reutiliza y amplía el software de gestión de proyectos Redmine. Las capacidades de Redmine y sus posibilidades de expansión a través de la incorporación de complementos permiten satisfacer los requisitos referentes a la gestión de proyectos.

Por un lado se ofrece un apartado de administración que permite definir los TFG y todas las personas involucradas en el mismo, realizando asimismo la asignación de permisos correspondiente (REQ.10-Administración).

Gracias a la definición de permisos para las personas involucradas en cada TFG se establece quién es la persona encargada de actualizar la información del proyecto y quién tiene permiso para consultarla. De este modo se gestiona la información referente a la planificación temporal del TFG (REQ.4-Planificación temporal compartida) y a la documentación del mismo (REQ.5-Gestión de la documentación y REQ.6-Acceso a la documentación).

A través de los foros de comunicación de Redmine y de sus opciones de ampliación mediante complementos, TraMeA satisface los requisitos relacionados con la ampliación de los canales de comunicación y retroalimentación: REQ.8-Retroalimentación y REQ.9-Realizar consultas en base a intereses comunes.

Para satisfacer el requisito referente al control de versiones de todos los documentos y elementos generados (REQ.7-Gestión de versiones), se ofrece la posibilidad de usar distintos sistemas de gestión de versiones como Git, Subversion y otros.

El cumplimiento de todos estos requisitos por parte de TraMeA se evaluó mediante un estudio de usabilidad (ver sección 6.2) que arrojó resultados muy positivos sobre la plataforma y su facilidad de uso.

Durante el diseño de TraMeA se decidió ampliar toda la parte relacionada con la evaluación y generalizarla de manera que no se limitara a la evaluación de los TFG. Por ello, los requisitos relacionados con la evaluación (REQ.1-Definición y gestión de rúbricas, REQ.2-Consulta de criterios y REQ.3-Uso de rúbricas) se trasladaron al diseño del sistema AdESMuS, que posteriormente fue integrado en TraMeA para dar soporte completo a la metodología FPM.

AdESMuS es un sistema que surge a partir de los requisitos funcionales relacionados con la evaluación establecidos para el sistema TraMeA. Sin embargo, se decidió extender su uso más allá de la evaluación de los TFG y se realizó un diseño y una implementación que permitiera su uso autónomo en cualquier proceso de evaluación (ver capítulo 4).

Para ello se definieron tres ontologías teniendo en cuenta las distintas formas de evaluación que puedan tener lugar en una asignatura o curso. Gracias a este diseño, AdESMuS permite utilizar distintas fuentes (instrumentos de evaluación) para calificar un mismo elemento evaluable y definir cómo deben ser combinadas dichas fuentes para obtener la nota final. También soporta desde la definición hasta el uso de rúbricas de evaluación como instrumento de calificación.

La arquitectura de AdESMuS además contempla la posibilidad de que la evaluación de ciertos elementos sea realizada por más de una persona, por ejemplo, en los tribunales de evaluación de los TFG, tal y como establece la metodología FPM, o en situaciones de evaluación por pares.

El diseño de AdESMuS permite respetar la perspectiva de género a la hora de definir los procesos de evaluación teniendo en cuenta el género tanto de la persona que va a realizar la evaluación, como el de la persona que va a ser evaluada.

Además, AdESMuS incorpora un módulo de ajuste que a través de técnicas de aprendizaje automático y analítica del aprendizaje, y utilizando los datos de las evaluaciones que se han realizado, sugiere ajustes al proceso de evaluación definido. De este modo, el sistema sirve también de ayuda a la reflexión del profesorado sobre lo adecuado del proceso de evaluación definido y permite realizar mejoras al mismo.

AdESMuS se ha implementado en forma de API REST para permitir su reusabilidad y facilitar su comunicación con plataformas de aprendizaje como Moodle o MAgAdI. También se ha desarrollado un sistema web que consume los recursos de AdESMuS y que se ha utilizado en la evaluación de los TFG en varias convocatorias.

De este modo se ha comprobado que AdESMuS satisface los requisitos que se habían establecido en TraMeA referentes a la evaluación con rúbricas (REQ.1-Definición y gestión de rúbricas, REQ.2-Consulta de criterios y REQ.3-Uso de rúbricas). También se ha evaluado su utilidad en un entorno real y se ha comprobado que los modelos sugeridos por el módulo de ajuste del sistema mejoran los definidos por los expertos (ver sección 5.5).

7.3 Líneas futuras

En este apartado se describen algunas de las líneas de investigación que se derivan directamente de los resultados obtenidos en este trabajo de tesis. En primer lugar se presentan las líneas que se encuentran actualmente en desarrollo y han empezado a generar resultados. A continuación se mencionan otras líneas de trabajo que están planificadas y que se llevarán a cabo en el marco de futuros proyectos de investigación.

7.3.1 Uso de AdESMuS en nuevos ámbitos

AdESMuS está diseñado para ser empleado en el proceso de evaluación de cualquier tipo de asignatura o curso, independientemente de la temática del mismo y del tipo de evaluación que se lleve a cabo.

Hasta ahora AdESMuS se ha utilizado únicamente en la evaluación de los TFG en los distintos grados relacionados con la Ingeniería Informática que se imparten

en la UPV/EHU. Próximamente se ampliará su uso a distintas asignaturas y en distintos ámbitos.

Se prevé usar AdESMuS en asignaturas del ámbito de la programación que se imparten tanto en la Escuela Universitaria de Ingeniería Técnica Industrial de Bilbao como en la Escuela Universitaria de Ingeniería de Vitoria. Cada una de estas asignaturas tiene su propio proceso de evaluación: distintos elementos evaluables tales como exámenes, prácticas e informes de laboratorio; evaluaciones realizadas por parte del profesorado y en ocasiones evaluaciones por pares entre el alumnado y distintos tipos de herramientas de evaluación como rúbricas o calificación en base a la opinión personal.

También se utilizará en la evaluación de los TFG en aquellas titulaciones, donde se va a implantar la metodología FPM. Se han mantenido reuniones con las personas responsables de los grados relacionados con la Ingeniería Industrial que se imparten en la EUI y se están definiendo las rúbricas para evaluar los TFG que se realicen en el grado en Criminología de la UPV/EHU.

7.3.2 Incorporación de aspectos de recomendación

Cada vez existe una mayor demanda de contenido personalizado y de retroalimentación inteligente en los sistemas educativos on-line (Santos & Boticario, 2015). Los aspectos relacionados con la personalización, son una de las funcionalidades principales para la nueva generación de entornos digitales de aprendizaje (Brown et al., 2015).

Los sistemas de recomendación permiten guiar a los usuarios de una manera personalizada y recomendarles los objetos adecuados (Burke, 2002; Mikic Fonte et al., 2012). De esta manera, los sistemas de recomendación pueden dar una respuesta adecuada a los problemas de personalización y proporcionar retroalimentación inteligente apoyando al alumnado y reduciendo la carga de trabajo del profesorado (Drachsler et al., 2015; Shaw et al., 2014).

El grupo de investigación GaLan cuenta con experiencia en el desarrollo de sistemas de recomendación en entornos de aprendizaje combinado (A. Álvarez, 2010; A. Álvarez et al., 2014). Dicha experiencia servirá como base para la incorporación en AdESMuS de distintos aspectos de recomendación.

Los datos relativos a las evaluaciones contenidos en AdESMuS pueden servir como fuente de información para enriquecer procesos de recomendación que permitan, por ejemplo, proponer al alumnado el material docente o las actividades más adecuadas en función de las carencias de conocimiento detectadas y de las competencias que aún no hayan sido alcanzadas.

Actualmente el grupo de investigación está trabajando en la recomendación de recursos docentes en base a los resultados de la evaluación de actividades en plataformas como Moodle. Los resultados obtenidos en esta línea de investigación se trasladarán a AdESMuS para enriquecer los procesos de evaluación y recomendación.

También se incorporarán en AdESMuS recomendaciones y alertas orientadas a la mejora del proceso de evaluación por parte del profesorado. Para esta ampliación se tomarán como base los trabajos previos realizados en esta área por el grupo GaLan (A. Álvarez et al., 2014; Martín, 2014).

Al definir el proceso de evaluación, los docentes podrán configurar el sistema para recibir alertas ante ciertas situaciones como por ejemplo, si se detecta que alguno de los elementos evaluables definidos no es significativo a la hora de calcular la calificación del alumnado. También podrán indicar cuál es el máximo error que se considera admisible en el cálculo de las calificaciones, de modo que si la diferencia entre las calificaciones otorgadas por el profesorado y las calculadas por el sistema exceden dicho error, se recomiende un nuevo modelo para el cálculo de la calificación.

7.3.3 Incorporación de aspectos de visualización

A la hora de trasladar los sistemas de recomendación a entornos educativos se debe tener en cuenta que sus necesidades en cuanto a objetivos, características de los usuarios y condiciones de recomendación son muy diferentes a los de otras áreas (Buder & Schwind, 2012). También se debe tener en cuenta que el aprendizaje y los procesos de aprendizaje son difíciles de entender en detalle.

Para dar respuesta a esta necesidad aparecen las áreas de la analítica del aprendizaje y la minería de datos educativos (Siemens, 2012; Romero & Ventura, 2013) cuyo objetivo es ampliar el conocimiento sobre el proceso de aprendizaje.

Una parte importante de estas áreas es la visualización de los datos (Kay & Bull, 2015) y los análisis preliminares de los datos o informes integrados que permiten mejorar las intervenciones pedagógicas por parte del profesorado (Tervakari et al., 2014; Pardo & Dawson, 2016). Además, una adecuada visualización de los datos puede ayudar al alumnado a reflexionar sobre su propio proceso de aprendizaje (Klerkx et al., 2014).

Por ello, se está trabajando en el desarrollo de un módulo de visualización (Villamañe, Alvarez, et al., 2016; Villamañe, Álvarez, et al., 2016c) que permita al profesorado y alumnado consultar los datos relativos al proceso de evaluación utilizando distintas técnicas de visualización. De este modo se enriquecerá la

información disponible sobre el proceso de aprendizaje y evaluación mejorando las intervenciones del profesorado y fomentando la auto-reflexión por parte del alumnado.

7.3.4 Integración de AdESMuS con otras herramientas

La implementación de AdESMuS como una API REST facilita su integración con otras herramientas como Moodle o MAgAdI por ejemplo. Para ello es necesario desarrollar para cada plataforma, un complemento que haga uso de los recursos ofrecidos por AdESMuS.

Inicialmente se desarrollará un complemento para Moodle, por ser la plataforma de aprendizaje usada como apoyo a la docencia presencial en la UPV/EHU. De este modo, se trasladarán a AdESMuS los resultados de aquellas actividades que se realicen o califiquen en Moodle y se mostrará en Moodle toda la información relativa al proceso de evaluación que esté contenida en AdESMuS. El complemento de Moodle también permitirá consumir los recursos de la API para la definición, uso y ajuste de rúbricas de evaluación.

Las interacciones presenciales entre profesorado y alumnado proporcionan una gran cantidad de información útil para estimar el éxito del aprendizaje (VanLehn, 2008). En esta línea, en el grupo de investigación GaLan se está trabajando en la adquisición de este tipo de información mediante los sistemas PresenceClick (Ruiz et al., 2015) y WebClick (Ruiz et al., 2013, 2016). AdESMuS se integrará también con estas herramientas para enriquecer su información sobre la asistencia del alumnado las clases presenciales y las interacciones entre alumnado y profesorado. Esto permitirá enriquecer los procesos de evaluación llevados a cabo en AdESMuS.

7.3.5 Enriquecimiento del proceso de ajuste

Actualmente el módulo de ajuste de AdESMuS permite calcular el TAM y el REM más adecuado para un proceso de evaluación.

`Mod_Ajuste` proporciona el modelo matemático que indica la influencia que cada una de las tareas debe tener en el proceso de evaluación. Para obtener dicho modelo tiene en cuenta las evaluaciones almacenadas en AdESMuS, por lo que a partir de distintas evaluaciones, el modelo obtenido puede variar.

El módulo de ajuste se puede ampliar para que incorpore un catálogo de los distintos modelos que haya calculado y permita realizar comparaciones entre ellos. También se permitirá al profesorado la introducción de distintos modelos para compararlos con los obtenidos automáticamente por el módulo.

En el cálculo actual del REM existen una serie de restricciones. Cuando hay más de una rúbrica implicada, se ha determinado que todas las rúbricas tengan el mismo número de niveles de valoración. Además, el REM seleccionado es el más adecuado teniendo en cuenta todo el conjunto de rúbricas, su influencia en la nota final (el TAM) y un único modelo de niveles de la escala (SM) para todas ellas.

Este proceso de ajuste se enriquecerá para permitir calcular el REM de manera que cada rúbrica pueda tener distintos niveles de valoración. Además, se dará la opción de que el cálculo del REM se realice permitiendo la elección de un SM distinto para cada rúbrica.

También se ampliará el cálculo del REM permitiendo su cálculo en base a valores fijados previamente en algunos de sus elementos. Por ejemplo, se podrá calcular el peso más adecuado de las dimensiones para un SM concreto fijado por el docente o calcular el mejor SM para unos pesos concretos de las dimensiones.

Aunque en la evaluación de los TFG sea habitual redondear la calificación en intervalos de 0.5 puntos, esto no es generalizable a la mayoría de asignaturas. Además, no es equiparable una diferencia de 0.5 puntos entre las calificaciones de 9 y 9.5, con la misma diferencia entre una calificación de 8.5 ó 9. En este último caso la categoría de la calificación varía, pasando de un Notable a un Sobresaliente. Por lo tanto, a la hora de calcular el error debería tenerse en cuenta si supone un cambio de categoría de la calificación. Por ello, se adaptará el proceso de ajuste para que tenga en cuenta los cambios de categoría de la calificación a la hora de optimizar el modelo.

7.3.6 Otras extensiones

A continuación se exponen de manera más resumida otras líneas de ampliación del trabajo presentado.

Mapeado automático de rúbricas

AdESMuS está diseñado para que la persona que diseña la evaluación pueda ofrecer la posibilidad a las personas que van a realizar la evaluación de elegir entre distintas formas de evaluar un mismo elemento.

Por ejemplo, en el caso de las evaluaciones de TFG llevadas a cabo en este trabajo de tesis, cada uno de los elementos evaluables de la metodología FPM se puede evaluar usando una rúbrica analítica o una atomística. Sin embargo, el hecho de que cada tipo de rúbrica tenga un número distinto de dimensiones y de niveles de valoración, dificulta la asignación de pesos y valores numéricos para que la

calificación final sea la misma independientemente del tipo de rúbrica que se haya usado para su cálculo.

Actualmente, `Mod_Ajuste` calcula los pesos de las dimensiones y los valores numéricos para la escala de valoración de cada tipo de rúbrica de manera independiente. Esto es, el modelo devuelto por `Mod_Ajuste` no tiene en cuenta que todos los tipos de rúbrica disponibles para un mismo elemento deben tener un comportamiento uniforme.

Para asegurar la equidad en la evaluación independientemente de qué tipo de rúbrica seleccione la persona evaluadora, se realizará un proceso de mapeado automático entre las distintas rúbricas que se ofrezcan para evaluar un elemento concreto. De este modo se asegurará que el hecho de que la persona evaluadora elija un tipo u otro de rúbrica no afectará a la calificación obtenida.

Asistente para el proceso de evaluación

Haciendo uso de técnicas de analítica del aprendizaje y minería de datos educativos, se enriquecerá el módulo de ajuste de AdESMuS para que proporcione información útil de cara a la organización del proceso de evaluación.

Por ejemplo, a partir de los datos recopilados sobre las distintas evaluaciones y categorizando los trabajos a evaluar, se puede determinar quiénes son las personas cuya opinión es más discriminante para realizar una evaluación en una categoría. Esto permitirá la realización de propuestas automáticas de personas evaluadoras o personas integrantes de tribunales para el caso de los TFG en función de la categoría a la que pertenezca el TFG.

Así mismo, también se podrá proponer cuál es el tipo de evaluación más adecuado para cada categoría de trabajo o el tipo de rúbrica, en el caso de que existan varios tipos de rúbrica disponibles para realizar la evaluación.

Además, a partir de todos los datos también se podrá conocer otro tipo de información como si existen personas cuyas valoraciones siempre distan mucho de la del resto de personas evaluadoras o si hay personas que valoran positivamente todos los trabajos independientemente de su calidad. Este tipo de información permitirá al profesorado realizar una auto-reflexión sobre su forma de evaluar para mejorarla si lo considera necesario.

Generación y rellenado automático de rúbricas

Algunos de los criterios que se tienen en cuenta en las evaluaciones realizadas mediante rúbricas de evaluación son criterios difícilmente cuantificables, como la claridad de los textos o la idoneidad del lenguaje empleado. Sin embargo, otros criterios son cuantificables como el número de faltas de ortografía, el número de

veces que se ha actualizado la planificación temporal de un TFG o el índice de asistencia a clase.

Se incorporará en AdESMuS un módulo que permita completar rúbricas de manera semi-automática teniendo en cuenta los datos cuantificables que se tengan disponibles. Por ejemplo, gracias a la unión de AdESMuS con PresenceClick y WebClick, se podrán incorporar criterios de asistencia o interacción presencial a las rúbricas de evaluación. De este modo, si el índice de asistencia o el número de interacciones personales son criterios tenidos en cuenta en la rúbrica, el sistema automáticamente indicará cuál es la valoración correspondiente en función de los datos recogidos.

Asimismo, este nuevo módulo de AdESMuS también permitirá la generación automática de rúbricas en base a evidencias cuantificables. Esto es, teniendo en cuenta la calificación obtenida por los estudiantes y todos los datos cuantificables que se tengan disponibles, analizar cuáles de esos datos son estadísticamente significativos para el cálculo de la calificación e incorporarlos en la rúbrica de evaluación correspondiente.

BIBLIOGRAFÍA

- Agresti, A., & Coull, B. A. (1998). Approximate is Better than «Exact» for Interval Estimation of Binomial Proportions. *The American Statistician*, 52(2), 119-126.
- Aguirre, A., Lozano, A., Villamañe, M., Ferrero, B., & Matey, L. (2012). Olympus: an intelligent interactive learning platform for procedural tasks. En P. Richard, M. Kraus, R. S. Laramee, & J. Braz (Eds.), *Actas de International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization Theory and Applications* (pp. 543-550). SciTePress - Science and Technology Publications.
- Aguirre, A., Lozano-Rodero, A., Matey, L., Villamane, M., & Ferrero, B. (2014). A novel approach to diagnosing motor skills. *IEEE Transactions on Learning Technologies*, 7(4), 304-318.
- Allen, S., & Knight, J. (2009). A Method for Collaboratively Developing and Validating a Rubric. *International Journal for the Scholarship of Teaching and Learning*, 3(2), Art. 10.
- Álvarez, A. (2010). *MAgAdI, una propuesta de sistema adaptativo multi-agente de apoyo al aprendizaje en un marco combinado* (Tesis doctoral). Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU).
- Álvarez, A., Martín, M., Fernández-Castro, I., & Urretavizcaya, M. (2013). Blending traditional teaching methods with learning environments: Experience, cyclical evaluation process and impact with MAgAdI. *Computers & Education*, 68, 129-140.
- Álvarez, A., Martín, M., Fernández-Castro, I., & Urretavizcaya, M. (2014). Supporting Blended-Learning: tool requirements and solutions with OWLish. *Interactive Learning Environments*, 24(6), 1176-1197.
- Álvarez, M., & Pascual, M. M. (2012). Propuesta de evaluación del Trabajo Fin de Grado en Derecho. *Aula Abierta*, 40(1), 85-102.
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies*, 4(3), 114-123.
- Barney, S., Khurum, M., Petersen, K., Unterkalmsteiner, M., & Jabangwe, R. (2012). Improving Students With Rubric-Based Self-Assessment and Oral Feedback. *IEEE Transactions on Education*, 55(3), 319-325.

BIBLIOGRAFÍA

- Bonilla, M. I., & Martín, C. (2012). Evaluación de competencias en el Trabajo Fin de Grado en Administración y Dirección de Empresas: una propuesta de la Facultad de Ciencias Sociales de Talavera de la Reina. UCLM. *Revista de Formación e Innovación Educativa Universitaria*, 5(4), 241-253.
- Bonilla, M. J., Fuentes, L., Vacas, C., & Vacas, T. (2012). Análisis del proceso de evaluación del Trabajo Fin de Grado en las nuevas titulaciones. *Educade: revista de educación en contabilidad, finanzas y administración de empresas*, (3), 5–21.
- Bosque, I. (2012). *Sexismo lingüístico y visibilidad de la mujer*. RAE. Recuperado a partir de http://www.rae.es/sites/default/files/Sexismo_linguistico_y_visibilidad_de_la_mujer_0.pdf
- Brooke, J. (1996). SUS: A quick and dirty usability scale. En P. W. Jordan, B. Weerdmeester, A. Thomas, & I. L. Mclelland (Eds.), *Usability evaluation in industry*. London: Taylor and Francis.
- Brooke, J. (2013). SUS: A Restrospective. *Journal of Usability Studies*, 8(2), 29-40.
- Brookhart, S. M. (1999). *The Art and Science of Classroom Assessment. The Missing Part of Pedagogy* (Vol. 27). Washington, DC: ERIC Clearinghouse on Higher Education.
- Brown, M., Dehoney, J., & Millichap, N. (2015). The Next Generation Digital Learning Environment. *A Report on Research. ELI Paper*.
- Buder, J., & Schwind, C. (2012). Learning with Personalized Recommender Systems: A Psychological View. *Computers in Human Behaviour*, 28(1), 207–216.
- Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, 12(4), 331–370.
- Carroll, J. M., & Rosson, M. B. (1987). Paradox of the active user. En J. M. Carroll (Ed.), *Interfacing Thought: Cognitive Aspects of Human-computer Interaction* (pp. 80–111). Cambridge, MA, USA: MIT Press.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247-1250.
- Chan, K. L. (2001). Statistical analysis of final year project marks in the computer engineering undergraduate program. *IEEE Transactions on Education*, 44(3), 258-261.
- Chisholm, W., Vanderheiden, G., & Jacobs, I. (2001). Web content accessibility guidelines 1.0. *Interactions*, 8(4), 35-54.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290. doi:10.1037/1040-3590.6.4.284
- Dandis, M. A. (2014). Benefits and criticism. Should we support or neglect using rubrics? Evidences from a literature review. *EDMETIC*, 3(2), 91-113.

- Dirección para la igualdad. (2010). *Guía para el uso no sexista de la lengua castellana y de imágenes en la UPV/EHU*. UPV/EHU. Recuperado a partir de <https://www.ehu.eus/documents/2007376/0/Guía>
- Drachler, H., Verbert, K., Santos, O. C., & Manouselis, N. (2015). Panorama of Recommender Systems to Support Learning. En F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender Systems Handbook* (pp. 421-451). Springer US.
- Dumas, J. S., & Redish, J. (1999). *A practical guide to usability testing* (Rev. ed.). Exeter, England ; Portland, Or: Intellect Books.
- Dutson, A. J., Todd, R. H., Magleby, S. P., & Sorensen, C. D. (1997). A Review of Literature on Teaching Engineering Design Through Project-Oriented Capstone Courses. *Journal of Engineering Education*, 86(1), 17-28.
- Escudero, D., & Hernández, D. (2012). Aplicación de buenas prácticas para la mejora de la calidad de los trabajos fin de grado en Ingeniería en Diseño Industrial. Presentado en Simposio Internacional sobre Innovación y Calidad en la Formación de Ingenieros, Valladolid.
- European Ministers of Education. (1999, junio 19). The Bologna Declaration of 19 June 1999: Joint declaration of the European Ministers of Education. Recuperado a partir de http://www.eurashe.eu/library/modernising-phe/Bologna_1999_Bologna-Declaration.pdf
- Fernández-Castro, I., Urretavizcaya, M., Arruarte, A., Elorriaga, J. A., García-Alonso, A., Ferrero, B., ... López-Gil, J.-M. (2012). GaLan (Entornos de Aprendizaje Adaptativos, Lenguajes y Sistemas Informáticos), UPV/EHU. *Revista Iberoamericana de Informática Educativa*, 15, 47-56.
- Ferrer, V., Carmona, M., & Soria, V. (2012). *El trabajo de fin de grado: guía para estudiantes, docentes y agentes colaboradores*. Madrid: McGrawHill.
- Ferrero, B. (2004). *DETECTive: un entorno genérico e integrable para diagnóstico de actividades de aprendizaje. Fundamentos, diseño y evaluación* (Tesis doctoral). Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU).
- Fielding, R. T. (2000). *Architectural Styles and the Design of Network-based Software Architectures* (Tesis doctoral). University of California.
- Fraile, R., Argüelles, I., Sande, J. C. G., Arriola, J. M. G., Llorente, J. I. G., Encinas, L. A., ... Peces, C. B. (2009). Definition of the educational outcomes of final year projects. En *Actas de INTED2009* (pp. 4683-4691).
- Graham, C. R. (2005). Blended learning systems: Definition, current trends, and future directions. En C. J. Bonk, C. R. Graham, & J. Cross (Eds.), *The Handbook of Blended Learning: Global Perspectives, Local Designs* (pp. 3-21). San Francisco: John Wiley & Sons.
- Guerrero, S. (2013). *Guía para un uso adecuado y no sexista del lenguaje y de la imagen en la Universidad de Jaén* (Servicio de Igualdad. Universidad de Jaén.). Jaén: Servicio de publicaciones. Universidad de Jaén.

BIBLIOGRAFÍA

- Hafner, J., & Hafner, P. (2003). Quantitative analysis of the rubric as an assessment tool: an empirical study of student peer-group rating. *International Journal of Science Education*, 25(12), 1509-1528.
- Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23-34.
- Hartson, H. R., Castillo, J. C., Kelso, J., & Neale, W. C. (1996). Remote evaluation: the network as an extension of the usability laboratory. En *Actas de SIGCHI Conference on Human Factors in Computing Systems* (pp. 228-235). ACM Press.
- Heinze, A., & Heinze, B. (2009). Blended e-learning skeleton of conversation: Improving formative assessment in undergraduate dissertation supervision. *British Journal of Educational Technology*, 40(2), 294-305.
- Hosseini-Khayat, A., Hellmann, T. D., & Maurer, F. (2010). Distributed and Automated Usability Testing of Low-Fidelity Prototypes (pp. 59-66). IEEE.
- Huba, M. E., & Freed, J. E. (2000). *Learner-centered assessment on college campuses: shifting the focus from teaching to learning*. Boston: Allyn and Bacon.
- Hunter, D. M., Jones, R. M., & Randhawa, B. S. (1996). The use of holistic versus analytic scoring for large-scale assessment of writing. *The Canadian Journal of Program Evaluation*, 11(2), 61-85.
- ISO. (1998). *ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability*. Geneva, Switzerland: International Organization for Standardization.
- ISO. (2002). *ISO 639-1:2002, Codes for the representation of names of languages — Part 1: Alpha-2 code*. Geneva, Switzerland: International Organization for Standardization.
- Ivory, M. Y., & Hearst, M. A. (2001). The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys*, 33(4), 470-516.
- Jackson, C. W., & Larkin, M. J. (2002). Rubric: Teaching Students To Use Grading Rubrics. *TEACHING Exceptional Children*, 35(1), 40-45.
- Jawitz, J., Shay, S., & Moore, R. (2002). Management and Assessment of Final Year Projects in Engineering. *International Journal of Engineering Education*, 18(4), 472-478.
- Jiménez, M. L., Román, M., & Traverso, J. (2011). Lenguaje no sexista y barreras a su utilización. Un estudio en el ámbito universitario. *Revista de Investigación en Educación*, 9(2), 174-183.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130-144.
- Kay, J., & Bull, S. (2015). New Opportunities with Open Learner Models and Visual Learning Analytics. En C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo

- (Eds.), *Actas de Artificial Intelligence in Education* (Vol. 9112, pp. 666-669). Cham: Springer International Publishing.
- Kleppe, A. G., Warmer, J. B., & Bast, W. (2003). *MDA explained: the model driven architecture: practice and promise*. Boston: Addison-Wesley.
- Klerkx, J., Verbert, K., & Duval, E. (2014). Enhancing Learning with Visualization Techniques. En J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 791-807). Springer New York.
- Ku, H., & Goh, S. (2010). Final year engineering projects in Australia and Europe. *European Journal of Engineering Education*, 35(2), 161-173. doi:10.1080/03043790903497336
- Laplace, P. S. (1812). *Theorie analytique des probabilités*. Paris, France: Courcier.
- Lawson, C., & Hanson, R. (1995). *Solving Least Squares Problems* (Vol. 15). Philadelphia: Society for Industrial and Applied Mathematics.
- Lewis, J. R., & Sauro, J. (2006). When 100% Really Isn't 100%: Improving the Accuracy of Small-Sample Estimates of Completion Rates. *Journal of Usability Studies*, 1(3), 136-150.
- Lin, W.-S., & Wang, C.-H. (2012). Antecedences to continued intentions of adopting e-learning system in blended learning instruction: A contingency framework based on models of information system success and task-technology fit. *Computers & Education*, 58(1), 88-99.
- Littlefair, G., & Gossman, P. (2008). BE (Hons) final year project assessment - leaving out the subjectiveness. En *Actas de Australasian Association for Engineering Education conference* (pp. 568-573). Rockhampton, Queensland: Australasian Association for Engineering Education.
- López-Gil, J. M., Urretavizcaya, M., Losada, B., & Fernández-Castro, I. (2016). Field vs. Laboratory Usability Evaluations: a Study on a Context Dependent Mobile Application Developed with an Agile Methodology. *IEEE Latin America Transactions*, 14(1), 339-348.
- Losada, B. (2014). *InterMod: Un enfoque ágil, dirigido por modelos y centrado en el usuario, para desarrollar aplicaciones interactivas* (Tesis doctoral). Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU).
- Losada, B., Urretavizcaya, M., & Fernández-Castro, I. (2013). A guide to agile development of interactive software with a «User Objectives»-driven methodology. *Science of Computer Programming*, 78(11), 2268-2281.
- Losada, B., Urretavizcaya, M., López-Gil, J.-M., & Fernández-Castro, I. (2012). Combining InterMod agile methodology with usability engineering in a mobile application development. En *Actas de International Conference on Interacción Persona-Ordenador* (p. 39:1-39:8). Elche: ACM Press.

BIBLIOGRAFÍA

- Malik, M., Khusainov, R., Zhou, S., & Adamos, V. (2009). A two year case study: Technology Assisted Project Supervision (TAPaS). *Engineering Education*, 4(2), 76-83.
- Martín, M. (2014). *Análisis de interacciones de aprendizaje y generación de avisos multi-usuario en entornos combinados* (Tesis doctoral). Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU).
- Martin, R. C. (2003). *Agile software development: principles, patterns, and practices*. Upper Saddle River, N.J: Prentice Hall.
- Mateo, J. (2009). *Guía para la evaluación de competencias en el trabajo fin de grado en el ámbito de las Ciencias Sociales y Jurídicas*. Agència per a la Qualitat del Sistema Universitari de Catalunya. Recuperado a partir de http://www.aqu.cat/doc/doc_30156918_1.pdf
- Mayhew, D. J. (1999). *The usability engineering lifecycle: a practitioner's handbook for user interface design*. San Francisco, Calif: Morgan Kaufmann Publishers.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30-46.
- McMillan, J. H., & Hearn, J. (2008). Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement. *Educational Horizons*, 87(1), 40-49.
- Meyer, B. (1997). *Object-oriented software construction* (2nd ed.). Upper Saddle River, N.J: Prentice Hall PTR.
- Mikic Fonte, F. A., Burguillo, J. C., & Nistal, M. L. (2012). An intelligent tutoring module controlled by BDI agents for an e-learning platform. *Expert Systems with Applications*, 39(8), 7546-7554.
- Mishra, D., & Mishra, A. (2011). Complex software project development: agile methods adoption. *Journal of Software Maintenance and Evolution: Research and Practice*, 23(8), 549-564.
- Moskal, B. M. (2000). Scoring rubrics: what, when and how? *Practical Assessment, Research & Evaluation*, 7(3).
- Nielsen, J. (2000). *Designing Web usability*. Indianapolis, Ind: New Riders.
- Nielsen, J., & Mack, R. L. (Eds.). (1994). *Usability inspection methods*. New York: Wiley.
- Nilson, L. B. (2010). *Teaching at its best: a research-based resource for college instructors* (3rd ed.). San Francisco, CA: Jossey-Bass.
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129-144.
- Pardo, A., & Dawson, S. (2016). Learning Analytics: How can Data be used to Improve Learning Practice. En P. Reimann, S. Bull, M. Kickmeier-Rust, R. K. Vatrappu & B. Wasson (Eds.), *Measuring and visualizing learning in the information-rich classroom*, (pp. 41-55). Routledge.

- Peat, B. (2006). Integrating Writing and Research Skills: Development and Testing of a Rubric to Measure Student Outcomes. *Journal of Public Affairs Education*, 12(3), 295-311.
- Peeters, M. J., Sahloff, E. G., & Stone, G. E. (2010). A Standardized Rubric to Evaluate Student Presentations. *American Journal of Pharmaceutical Education*, 74(9), 171.
- Pressman, R. S. (2005). *Software engineering: a practitioner's approach* (6th ed.). Boston, Mass: McGraw-Hill.
- Real Decreto 1393/2007, de 29 de Octubre, por el que se establece la ordenación de las enseñanzas universitarias oficiales.* (2007). (Boletín Oficial del Estado No. 260) (pp. 44037-44048).
- Reimann, P., Bull, S., Kickmeier-Rust, M. D., Vatrappu, R., & Wasson, B. (Eds.). (2016). *Measuring and visualizing learning in the information-rich classroom*. New York, NY: Routledge.
- Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 12-27.
- Rosenberry, L. (2010). Using Quality Rubrics to Improve Student Performance, Enhance Grading Objectivity and Consistency, and Enhance Instructor Feedback. En J. Sanchez & K. Zhang (Eds.), *Actas de World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2010* (pp. 1451-1451). Orlando, Florida, USA: Association for the Advancement of Computing in Education (AACE).
- Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: how to plan, design, and conduct effective tests* (2nd ed.). Indianapolis, IN: Wiley Pub.
- Ruiz, S., Charleer, S., Urretavizcaya, M., Klerkx, J., Fernández-Castro, I., & Duval, E. (2016). Supporting Learning by Considering Emotions: Tracking and Visualization a Case Study. En D. Gašević & G. Lynch (Eds.), *Actas de International Conference on Learning Analytics & Knowledge* (pp. 254-263). New York, NY, USA: ACM.
- Ruiz, S., Urretavizcaya, M., & Fernandez-Castro, I. (2013). Monitoring F2F interactions through attendance control. En R. Shehab, J. Sluss, & D. Trytten (Eds.), *Actas de IEEE Frontiers in Education Conference* (pp. 226-232). IEEE Computer Society.
- Ruiz, S., Urretavizcaya, M., Fernández-Castro, I., & López-Gil, J.-M. (2015). Visualizing Students' Performance in the Classroom: Towards Effective F2F Interaction Modelling. En G. Conole, T. Klobučar, C. Rensing, J. Konert, & É. Lavoué (Eds.), *Actas de European Conference on Technology Enhanced Learning* (pp. 630-633). Cham: Springer International Publishing.
- Rullan, M., Estapé-Dubreuil, G., Fernández, M., & Márquez, M. D. (2010). La evaluación de competencias transversales en la materia trabajos fin de grado: Un estudio preliminar sobre la necesidad y oportunidad de establecer medios e

BIBLIOGRAFÍA

- instrumentos por ramas de conocimiento. *REDU: Revista de Docencia Universitaria*, 8(1), 74-100.
- Sabah, N. M. (2013). Initial Suggestions for Supervising and Mentoring Undergraduate Student Projects. Presentado en Evaluation & Assessment of the Undergraduate Project Workshop. Recuperado a partir de http://www.academia.edu/3655229/Initial_Suggestions_for_Supervising_and_Mentoring_Undergraduate_Student_Projects
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144. doi:10.1007/BF00117714
- Salinas, J. J., & Erochko, J. (2015). Using Weighted Scoring Rubrics in Engineering Assessment. En *Actas de Canadian Engineering Education Association Annual Conference*. Hamilton, Ontario: Canadian Engineering Education Association.
- Sanchez, F., Climent, J., Corbalan, J., Fonseca, P., Garcia, J., Herrero, J. R., ... Lopez, D. (2014). Evaluation and assessment of professional skills in the Final Year Project. En *Actas de IEEE Frontiers in Education Conference* (pp. 2352-2359). IEEE Computer Society.
- Sánchez, F., Moral, J., & López, D. (2015). Todos mis estudiantes acaban el TFG, y en el tiempo previsto. *ReVisión*, 8(3), 31-46.
- Sanchez, J. L., Gonzalez, C. S., & Alayon, S. (2011). Evaluation of transversal competences in the final year project in engineering. En *Actas de EAEEIE Annual Conference* (pp. 1-5). Maribor, Slovenia: Institute of Electrical and Electronics Engineers.
- Sánchez-Apellániz, M. J. (2009). Lenguaje y comunicación no sexista. En *Manual de agentes de igualdad* (pp. 255-268). Sevilla: Diputación de Sevilla.
- Santos, O. C., & Boticario, J. G. (2015). Practical guidelines for designing and evaluating educationally oriented recommendations. *Computers & Education*, 81, 354-374.
- Sauro, J., & Kindlund, E. (2005). How Long Should a Task Take? Identifying Specification Limits for Task Times in Usability Tests. En *Actas de Human Computer Interaction International Conference*. Las Vegas, Nevada, USA: Lawrence Erlbaum Associates.
- Sauro, J., & Lewis, J. R. (2005). Estimating Completion Rates from Small Samples Using Binomial Confidence Intervals: Comparisons and Recommendations. *Human Factors and Ergonomics Society Annual Meeting*, 49(24), 2100-2103.
- Sauro, J., & Lewis, J. R. (2010). Average task times in usability tests: what to report? En E. Mynatt (Ed.), *Actas de SIGCHI Conference on Human Factors in Computing Systems* (pp. 2347-2350). ACM Press.
- Sauro, J., & Lewis, J. R. (2012). *Quantifying the user experience: practical statistics for user research*. Amsterdam ; Waltham, MA: Elsevier/Morgan Kaufmann.

- Shaw, C., Larson, R., & Sibdari, S. (2014). An Asynchronous, Personalized Learning Platform—Guided Learning Pathways (GLP). *Creative Education*, 5(13), 1189-1204.
- Siemens, G. (2012). Learning analytics: envisioning a research discipline and a domain of practice. En *Actas de International Conference on Learning Analytics and Knowledge* (pp. 4–8). ACM.
- Smith, P. A. (1996). Towards a practical measure of hypertext usability. *Interacting with Computers*, 8(4), 365-381.
- Stevens, D. D., Levi, A. J., & Walvoord, B. E. (2012). *Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback, and Promote Student Learning* (2 edition.). DULLES: Stylus Publishing.
- Stewart, B. G., & Nesbitt, A. (2005). Quality assessment of BSc engineering honours projects at Glasgow Caledonian University. *International Journal of Electrical Engineering Education*, 42(1), 52-62.
- Stojanovic, Z., Dahanayake, A., & Sol, H. (2003). Modeling and Architectural Design in Agile Development Methodologies. Presentado en Exploring Modelling Methods for Systems Analysis and Design Workshop. Recuperado a partir de <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.108.3616&rep=rep1&type=pdf>
- Swartz, C. W., Hooper, S. R., Montgomery, J. W., Wakely, M. B., de Kruif, R. E. L., Reed, M., ... White, K. P. (1999). Using Generalizability Theory to Estimate the Reliability of Writing Scores Derived from Holistic and Analytical Scoring Methods. *Educational and Psychological Measurement*, 59(3), 492-506.
- Taras, M. (2005). Assessment – Summative And Formative – Some Theoretical Reflections. *British Journal of Educational Studies*, 53(4), 466-478.
- Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, 6(1), 35–39.
- Teo, Y., & Ho, D. J. (1998). A systematic approach to the implementation of final year project in an electrical engineering undergraduate course. *IEEE Transactions on Education*, 41(1), 25-30.
- Tervakari, A. M., Silius, K., Koro, J., Paukkeri, J., & Pirttilä, O. (2014). Usefulness of information visualizations based on educational data. En *Actas de IEEE Global Engineering Education Conference* (pp. 142-151). IEEE Computer Society.
- Tinsley, H. E. A., & Weiss, D. J. (2000). 4 - Interrater Reliability and Agreement. En H. E. A. T. D. Brown (Ed.), *Handbook of Applied Multivariate Statistics and Mathematical Modeling* (pp. 95-124). San Diego: Academic Press.
- Trice, A. D. (2000). *A handbook of classroom assessment*. New York: Addison Wesley Longman.

BIBLIOGRAFÍA

- Tullis, T., & Albert, B. (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics* (Second edition.). Amsterdam; Boston: Elsevier/Morgan Kaufmann.
- Tullis, T., & Stetson, J. N. (2004). A comparison of questionnaires for assessing website usability. En *Proceedings of the 13th UPA Conference*. Minneapolis.
- Valderrama, E. (2009). *Guía para la evaluación de competencias en los trabajos fin de grado y de máster en las Ingenierías*. Agència per a la Qualitat del Sistema Universitari de Catalunya. Recuperado a partir de http://www.aqu.cat/doc/doc_19718727_1.pdf
- Valderrama, E., Rullan, M., Sánchez, F., Pons, J., Mans, C., Giné, F., ... Bisbal, J. (2010). La Evaluación de Competencias en los Trabajos Fin de Estudios. *Revista Iberoamericana de Tecnologías del/da Aprendizaje/Aprendizagem*, 5(3), 107-114.
- Valderrama, E., Rullan, M., Sánchez, F., Pons, J., Mans, C., Giné, F., ... Peig, E. (2009). Guidelines for the final year project assessment in engineering. En *Actas de IEEE Frontiers in Education Conference* (pp. 1-5). San Antonio, Texas, EE.UU.: IEEE Computer Society.
- VanLehn, K. (2008). The Interaction Plateau: Answer-Based Tutoring < Step-Based Tutoring = Natural Tutoring. En B. P. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent Tutoring Systems* (Vol. 5091, pp. 7-7). Berlin, Heidelberg: Springer Berlin Heidelberg. Recuperado a partir de <http://www.springerlink.com/content/8x614214xw34n52q/>
- Vijayalakshmi, M., Desai, P. D., & Joshi, G. H. (2013). Outcome based education performance evaluation of capstone project using assessment rubrics and matrix (pp. 6-10). IEEE.
- Vilardell Riera, I. (2010). Experiencia en el trabajo de fin de Grado en Administración y Dirección de Empresas. *Revista de Educación en Contabilidad, Finanzas y Administración de Empresas*, 1(1), 101-122.
- Villamañe, M., Álvarez, A., Larrañaga, M., & Ferrero, B. (2016a). Desarrollo y validación de un conjunto de rúbricas para la evaluación de Trabajos Fin de Grado. *ReVisión, In Press*.
- Villamañe, M., Álvarez, A., Larrañaga, M., & Ferrero, B. (2016b). Recommending Dimension Weights and Scale Values in Multi-rubric Evaluations. En K. Verbert, M. Sharples, & T. Klobučar (Eds.), *Actas de European Conference on Technology Enhanced Learning* (Vol. 9891, pp. 678-681). Lyon: Springer International Publishing.
- Villamañe, M., Álvarez, A., Larrañaga, M., & Ferrero, B. (2016c). RubricVis: enriching rubric-based formative assessment with visual learning analytics. Presentado en International Conference Technological Ecosystems for Enhancing Multiculturality, Salamanca: In Press.
- Villamañe, M., Alvarez, A., Larrañaga, M., & Ferrero, B. (2016). Visual feedback in rubric based evaluation. Presentado en Learning Analytics Summer Institute (LASI

-
- 2016), Bilbao. Recuperado a partir de http://repositorio.grial.eu/bitstream/grial/629/2/Villamane_LASI2016.pdf
- Villamañe, M., Ferrero, B., & Álvarez, A. (2014). Uso de las tic en el seguimiento y evaluación del trabajo fin de grado en el campo de las ingenierías. En *Actas del I Congreso Interuniversitario sobre el Trabajo de Fin de Grado. Retos y oportunidades del TFG en la sociedad del conocimiento* (pp. 222-233). Servicio Editorial UPV/EHU.
- Villamañe, M., Ferrero, B., Álvarez, A., Larrañaga, M., Arruarte, A., & Elorriaga, J. A. (2014). Dealing with common problems in engineering degrees' Final Year Projects. En *Actas de IEEE Frontiers in Education Conference* (pp. 2663-2670). Madrid: IEEE Computer Society.
- Villamañe, M., Larrañaga, M., Álvarez, A., & Ferrero, B. (2015). Adjusting the weights of assessment elements in the evaluation of Final Year Projects. En *Actas de International Conference on Educational Data Mining* (pp. 596-597). Madrid (Spain).
- Wehlburg, C. M. (2006). *Meaningful Course Revision: Enhancing Academic Engagement Using Student Learning Data* (1 edition.). Bolton, Mass.: Anker.
- Wharton, C., Rieman, J., Lewis, C., & Polson, P. (1994). The cognitive walkthrough method: a practitioner's guide. En J. Nielsen & R. L. Mack (Eds.), *Usability inspection methods* (pp. 105-140). New York, NY, USA: John Wiley & Sons, Inc.
- Wieggers, K. E., & Beatty, J. (2013). *Software requirements* (Third edition.). Redmond, Washington: Microsoft Press, s division of Microsoft Corporation.
- Wilson, E. B. (1927). Probable Inference, the Law of Succession, and Statistical Inference. *Journal of the American Statistical Association*, 22(158), 209-212.
- Zipitria, I. (2011). *From Human to Automatic Summary Grading* (Tesis doctoral). Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU).

ÍNDICE DE FIGURAS

| | |
|---|-----|
| Fig. 1-1.- Resultados de las preguntas 35 (a) y 36 (b) al profesorado | 5 |
| Fig. 1-2.- Resultados de la pregunta 38 al profesorado | 5 |
| Fig. 1-3.- Resultados de las preguntas 11 (a) y 13 (b) al alumnado..... | 6 |
| Fig. 1-4.- Resultados a las preguntas 31 (a) y 32 (b) al profesorado..... | 7 |
| Fig. 1-5.- Resultados a las preguntas 36 (a) al alumnado y 52 (b) al profesorado..... | 8 |
| Fig. 1-6.- Resultados a las preguntas 30 (a) y 31 (b) al alumnado | 9 |
| Fig. 1-7.- Resultados a la pregunta 34 al alumnado | 9 |
| Fig. 1-8.- Resultados a las preguntas 33 (a) al alumnado y 42 (b) al profesorado..... | 10 |
| Fig. 2-1.- Resumen evaluación proyectos (Jawitz et al., 2002) | 26 |
| Fig. 2-2.- Evaluación de los proyectos final de carrera de Ingeniería Eléctrica en la Universidad de Victoria (Ku & Goh, 2010) | 28 |
| Fig. 2-3.- Clasificación de métodos de evaluación. Adaptada de (Hunter et al., 1996) | 30 |
| Fig. 2-4.- Metodología FPM..... | 58 |
| Fig. 3-1.- Proceso de InterMod (Losada et al., 2012) | 65 |
| Fig. 3-2.- Interfaz gráfica de Redmine..... | 67 |
| Fig. 3-3.- Arquitectura inicial de TraMeA | 81 |
| Fig. 3-4.- Arquitectura final de TraMeA con AdESMuS | 82 |
| Fig. 4-1.- Arquitectura modular de AdESMuS | 91 |
| Fig. 4-2.- Recolección de datos necesarios para los procesos de ajuste | 95 |
| Fig. 4-3.- Proceso de obtención del TAM | 96 |
| Fig. 4-4.- Proceso de obtención del REM..... | 98 |
| Fig. 4-5.- Representación UML de la ontología Ont_Universidad..... | 105 |
| Fig. 4-6.- Definición de las clases Centro, Titulación y Grupo..... | 105 |
| Fig. 4-7.- Definición de las clases Asignatura y GrupoAsignatura | 106 |

ÍNDICE DE FIGURAS

| | |
|---|-----|
| Fig. 4-8.- Ejemplo de objetos de las clases Centro , Titulación y Grupo | 107 |
| Fig. 4-9.- Ejemplo de objetos de las clases Asignatura y GrupoAsignatura | 107 |
| Fig. 4-10.- Representación UML de la ontología Ont_Definición | 108 |
| Fig. 4-11.- Definición de las clases FormaEvaluación , PesoEvaluable , ElementoEvaluable y Competencia | 109 |
| Fig. 4-12.- Definición de las clases PesoFuente y Fuente | 110 |
| Fig. 4-13.- Ejemplo de objetos de las clases FormaEvaluación , PesoEvaluable y ElementoEvaluable , PesoFuente y Fuente | 111 |
| Fig. 4-14.- Ejemplo de objetos de las clases FormaEvaluación , ElementoEvaluable y PesoEvaluable | 112 |
| Fig. 4-15.- Ejemplo de objetos de las clases ElementoEvaluable , PesoFuente y Fuente | 113 |
| Fig. 4-16.- Definición de las clases Rúbrica , Nivel , Dimensión y DescripciónNivelDimensión | 114 |
| Fig. 4-17.- Ejemplo de objetos de las clases Rúbrica , Nivel , Dimensión y DescripciónNivelDimensión | 116 |
| Fig. 4-18.- Ejemplo de objetos de las clases Rúbrica , Nivel , Dimensión y DescripciónNivelDimensión | 117 |
| Fig. 4-19.- Definición de la clase Perspectiva | 119 |
| Fig. 4-20.- Ejemplo de objetos de la clase DescripciónNivelDimensión incluyendo la perspectiva de género | 120 |
| Fig. 4-21.- Ejemplo de objetos de la clase Perspectiva | 120 |
| Fig. 4-22.- Representación UML de la ontología Ont_Evaluación (en verde) | 121 |
| Fig. 4-23.- Definición de la clase TareaEvaluación | 121 |
| Fig. 4-24.- Definición de las clases EvaluaciónElementoEvaluable y EvaluaciónCompleta | 122 |
| Fig. 4-25.- Definición de la clase DetalleEvaluaciónRúbricaEvaluador | 123 |
| Fig. 4-26.- Ejemplo de objetos de la clase TareaEvaluación (I) | 123 |
| Fig. 4-27.- Ejemplo de objetos de la clase TareaEvaluación (II)..... | 124 |
| Fig. 4-28.- Ejemplo de objetos de la clase TareaEvaluación (III)..... | 124 |
| Fig. 4-29.- Ejemplo de objetos de las clases EvaluaciónElementoEvaluable y EvaluaciónCompleta | 125 |

| | |
|--|-----|
| Fig. 4-30.- Ejemplo de objeto de la clase DetalleEvaluaciónRúbricaEvaluador | 125 |
| Fig. 4-31.- Definición de la clase Internacionalización | 126 |
| Fig. 4-32.- Ejemplo de internacionalización | 127 |
| Fig. 4-33.- Interfaces del sistema web | 129 |
| Fig. 5-1.- Proceso de definición de las rúbricas | 135 |
| Fig. 5-2.- Iteraciones del proceso de definición de rúbricas..... | 136 |
| Fig. 5-3.- Resultados de las preguntas P1 (a) y P2 (b) de la encuesta sobre rúbricas (profesorado)..... | 140 |
| Fig. 5-4.- Resultados de las preguntas E1 (a) y E2 (b) de la encuesta sobre rúbricas (alumnado)..... | 142 |
| Fig. 6-1.- Proceso de evaluación | 160 |
| Fig. 6-2.- Comparación del sistema original (a) y del prototipo realizado (b) | 163 |
| Fig. 6-3.- Interfaces gráficas de elección (a) y descripción (b) de tareas..... | 164 |
| Fig. 6-4.- Resumen de una prueba de usabilidad..... | 165 |
| Fig. 6-5.- Ejemplo de grafo con el recorrido de ejecución | 171 |
| Fig. 6-6.- Grafo con la navegación de la tarea CPU0_A | 182 |
| Fig. 6-7.- Grafo con la navegación de la tarea CPU8_T; Error! Marcador no definido. | |
| Fig. 6-8.- Grafo con la navegación de la tarea CPU10_T | 197 |
| Fig. B - 1.- Resultados pregunta 3..... | 239 |
| Fig. B - 2.- Resultados pregunta 11 (a) y pregunta 13 (b) | 240 |
| Fig. B - 3.- Resultados pregunta 22 (a) y pregunta 23 (b) | 242 |
| Fig. B - 4.- Resultados pregunta 30 (a) y pregunta 31 (b) | 243 |
| Fig. B - 5.- Resultados pregunta 33 en función de la pregunta 35..... | 245 |
| Fig. C - 1.- Resultados pregunta 7 (a) y pregunta 8 (b)..... | 248 |
| Fig. C - 2.- Resultados pregunta 10 y 11 | 251 |
| Fig. C - 3.- Resultados pregunta 15 | 252 |
| Fig. C - 4.- Resultados preguntas 31 (a) y 32 (b) | 255 |
| Fig. C - 5.- Resultados preguntas 35 (a) y 36 (b) | 256 |
| Fig. C - 6.- Respuestas a la pregunta 38..... | 257 |
| Fig. C - 7.- Resultados preguntas 39 (a) y 40 (b) | 257 |
| Fig. C - 8.- Resultados preguntas 41 (a) y 42 (b) | 258 |
| Fig. C - 9.- Resultados pregunta 43 | 258 |
| Fig. C - 10.- Resultados preguntas 45 (a) y 46 (b)..... | 260 |

ÍNDICE DE FIGURAS

Fig. C - 11.- Resultados pregunta 52260

ÍNDICE DE TABLAS

| | |
|--|----|
| Tabla 2-1.- Evaluación propuesta por (Valderrama et al., 2009) | 17 |
| Tabla 2-2.- Evaluación de los TFG del área de Ciencias Sociales y Jurídicas (Mateo, 2009) | 18 |
| Tabla 2-3.- Ejemplo de indicadores para un elemento evaluable. Adaptado de (Valderrama et al., 2010)..... | 19 |
| Tabla 2-4.- Evaluación propuesta por (Vilardell Riera, 2010)..... | 21 |
| Tabla 2-5.- Evaluación propuesta por (M. I. Bonilla & Martín, 2012)..... | 22 |
| Tabla 2-6.- Evaluación propuesta por (M. Álvarez & Pascual, 2012) | 23 |
| Tabla 2-7.- Ejemplo de rúbrica para la evaluación de una competencia. Adaptada de (M. Álvarez & Pascual, 2012)..... | 24 |
| Tabla 2-8.- Evaluación propuesta por (Teo & Ho, 1998) | 25 |
| Tabla 2-9.- Evaluación de los proyectos en la Escuela de Ingeniería Mecánica de la Universidad de Western Australia | 29 |
| Tabla 2-10.- Extracto de una rúbrica holística | 32 |
| Tabla 2-11.- Extracto de una rúbrica analítica | 33 |
| Tabla 2-12.- Extracto rúbrica atomística | 34 |
| Tabla 2-13.- Extracto de una rúbrica holística con sugerencia de valores numéricos para los niveles | 36 |
| Tabla 2-14.- Extracto de rúbrica con niveles de valoración equitativos | 37 |
| Tabla 2-15.- Extracto de una rúbrica con niveles de valoración no equidistantes..... | 37 |
| Tabla 2-16.- Comparativa herramientas usadas actualmente en la UPV/EHU.... | 44 |
| Tabla 2-17.- Comparativa herramientas de gestión de documentación..... | 48 |
| Tabla 2-18.- Comparativa herramientas de comunicación | 50 |
| Tabla 2-19.- Comparativa herramientas de gestión de proyectos..... | 53 |
| Tabla 2-20.- Comparativa herramientas de trabajo con rúbricas | 56 |
| Tabla 3-1.- Correspondencia de los UO extraídos de las encuestas con los requisitos..... | 70 |

ÍNDICE DE TABLAS

| | |
|---|-----|
| Tabla 3-2.- Correspondencia de los UO de administración y gestión con los requisitos..... | 71 |
| Tabla 3-3.- Lista completa de Objetivos de Usuario iniciales | 72 |
| Tabla 3-4.- Objetivos de Usuario indirectos obtenidos en las iteraciones para los roles Persona usuaria y Persona administradora | 73 |
| Tabla 3-5.- Objetivos de Usuario indirectos obtenidos en las iteraciones para el rol Estudiante | 74 |
| Tabla 3-6.- Objetivos de Usuario indirectos obtenidos en las iteraciones para los roles Docente, Persona supervisora y Componente tribunal..... | 75 |
| Tabla 3-7.- Lista de Objetivos de Usuario indirectos fusionados (Parte I)..... | 76 |
| Tabla 3-8.- Lista de Objetivos de Usuario indirectos fusionados (Parte II) | 77 |
| Tabla 3-9.- Objetivos de Usuario finales de TraMeA..... | 78 |
| Tabla 3-10.- Objetivos de Usuario de TraMeA en relación a las funcionalidades proporcionadas por Redmine | 80 |
| Tabla 4-1.- Objetivos de Usuario heredados de TraMeA | 87 |
| Tabla 4-2.- Objetivos de Usuario de AdESMuS..... | 88 |
| Tabla 4-3.- División en Objetivos de Usuario Indirectos de los UO de Persona usuaria..... | 88 |
| Tabla 4-4.- División en Objetivos de Usuario Indirectos de los UO de Docente | 89 |
| Tabla 4-5.- División en Objetivos de Usuario Indirectos de los UO de Persona evaluadora | 89 |
| Tabla 4-6.- División de Ad_UO01-Est en Objetivos de Usuario Indirectos..... | 90 |
| Tabla 4-7.- División de Ad_UO01-A en Objetivos de Usuario Indirectos | 90 |
| Tabla 4-8.- Posible asignación de valores numéricos | 99 |
| Tabla 4-9.- Ejemplo de distintos SM para el rango [0, 10]..... | 100 |
| Tabla 4-10.- Posibles formas de evaluación | 104 |
| Tabla 4-11.- Estructura general de una rúbrica de evaluación | 113 |
| Tabla 4-12.- Extracto de la rúbrica analítica del «Planteamiento Inicial» | 115 |
| Tabla 4-13.- Extracto de la rúbrica atomística del «Planteamiento Inicial» | 116 |
| Tabla 4-14.- Ejemplo de redacción incluyendo la perspectiva de género | 119 |
| Tabla 4-15.- Ejemplos de mensajes para consumir recursos de AdESMuS | 128 |
| Tabla 5-1.- Valores para el conjunto preliminar de rúbricas | 137 |
| Tabla 5-2.- Dimensiones peor valoradas por el alumnado | 139 |
| Tabla 5-3.- Extracto de la encuesta sobre versiones de rúbricas (profesorado) .. | 140 |

| | |
|---|-----|
| Tabla 5-4.- Extracto de la encuesta sobre rúbricas (alumnado)..... | 141 |
| Tabla 5-5.- Comparación del TAM calculado por Mod_Ajuste y el TAM establecido por los expertos | 144 |
| Tabla 5-6.- Resultados validación modelos TAM | 146 |
| Tabla 5-7.- Posible asignación de valores numéricos | 147 |
| Tabla 5-8.- Valores numéricos del SM seleccionado | 147 |
| Tabla 5-9.- Comparación de los pesos de las dimensiones obtenidos por Mod_Ajuste y los establecidos por los expertos | 148 |
| Tabla 5-10.- Resultados validación MSEM..... | 150 |
| Tabla 6-1.- Resultados evaluación elementos evaluables | 157 |
| Tabla 6-2.- Extracto de la encuesta sobre rúbricas al alumnado | 158 |
| Tabla 6-3.- Tareas definidas para el estudio de usabilidad y su correspondencia con los UO de TraMeA..... | 161 |
| Tabla 6-4.- Códigos de las tareas del estudio de usabilidad | 162 |
| Tabla 6-5.- Cuestionario de la Escala de Usabilidad del Sistema (SUS)..... | 166 |
| Tabla 6-6.- Porcentaje de tareas completadas con éxito y su tasa de éxito muestral | 174 |
| Tabla 6-7.- Estimación de las tasas de éxito poblacionales | 174 |
| Tabla 6-8.- Intervalos de confianza ajustados de Wald al 95% (éxito/fracaso).... | 175 |
| Tabla 6-9.- Resultados en función del nivel de finalización | 176 |
| Tabla 6-10.- Estimación de las tasas de éxito poblacional y método de cálculo | 177 |
| Tabla 6-11.- Intervalos de confianza ajustados de Wald al 95% (<i>sin problemas</i>) .. | 177 |
| Tabla 6-12.- Tiempos de ejecución de las tareas (segundos)..... | 178 |
| Tabla 6-13.- Intervalos de confianza al 95% y máximo admisible para el tiempo de ejecución de las tareas | 179 |
| Tabla 6-14.- Nivel de pérdida de los participantes | 180 |
| Tabla 6-15.- Intervalos de confianza al 95% para el nivel de pérdida de las tareas | 180 |
| Tabla 6-16.- Puntuaciones del SUS..... | 184 |
| Tabla 6-17.- Porcentaje de tareas completadas con éxito y su tasa de éxito muestral | 185 |
| Tabla 6-18.- Estimación de las tasas de éxito poblacional..... | 186 |
| Tabla 6-19.- Intervalos de confianza ajustados de Wald al 95% (éxito/fracaso)..... | 187 |

ÍNDICE DE TABLAS

| | |
|--|-----|
| Tabla 6-20.- Resultados en función del nivel de finalización | 188 |
| Tabla 6-21.- Estimación de las tasas de éxito poblacionales y método de cálculo | 189 |
| Tabla 6-22.- Intervalos de confianza ajustados de Wald al 95% (<i>sin problemas</i>) .. | 190 |
| Tabla 6-23.- Tiempos de ejecución de las tareas (segundos)..... | 191 |
| Tabla 6-24.- Intervalos de confianza al 95% y máximo admisible para el tiempo de ejecución de las tareas | 192 |
| Tabla 6-25.- Nivel de pérdida de los participantes | 193 |
| Tabla 6-26.- Intervalos de confianza al 95% para el nivel de pérdida de las tareas | 194 |
| Tabla 6-27.- Puntuaciones del SUS..... | 198 |
| Tabla B - 1.- Preguntas categoría Contextualización | 238 |
| Tabla B - 2.- Preguntas categoría Definición del proyecto | 240 |
| Tabla B - 3.- Preguntas categoría Durante la realización (I) | 241 |
| Tabla B - 4.- Preguntas categoría Durante la realización (II) | 242 |
| Tabla B - 5.- Preguntas categoría Evaluación | 244 |
| Tabla C - 1.- Preguntas categoría Contextualización..... | 248 |
| Tabla C - 2.- Preguntas categoría Experiencia | 249 |
| Tabla C - 3.- Preguntas categoría Supervisión (I)..... | 250 |
| Tabla C - 4.- Preguntas categoría Supervisión (II) | 253 |
| Tabla C - 5.- Preguntas categoría Supervisión (III) | 254 |
| Tabla C - 6.- Preguntas categoría Evaluación (I) | 255 |
| Tabla C - 7.- Preguntas categoría Evaluación (II)..... | 256 |
| Tabla C - 8.- Preguntas categoría Evaluación (III) | 259 |

ANEXO A

ACRÓNIMOS

| | |
|-----------------|---|
| API | Application Programming Interfaces (interfaces de programación de aplicaciones) |
| BC | Base de Conocimiento |
| CCI | Coefficiente de Correlación Intraclase |
| CSS | Cascade Style Sheet (hoja de estilo en cascada) |
| CPU | Caso de Prueba de Usabilidad |
| EEES | Espacio Europeo de Educación Superior |
| EUI | Escuela Universitaria de Ingeniería de Vitoria |
| EUITI | Escuela Universitaria de Ingeniería Técnica Industrial de Bilbao |
| G.A.U.R. | Sistema de Gestión Académica Universitaria Renovada |
| GIIGySI | Grado en Ingeniería Informática de Gestión y Sistemas de Información |
| HTML | HyperText Markup Language (lenguaje de marcas de hipertexto) |
| JSON | JavaScript Object Notation (notación de objetos JavaScript) |
| LMS | Learning Management System (sistema de gestión de aprendizaje) |
| MDD | Model Driven Development (desarrollo dirigido por modelos) |
| MSEM | Multi-Source Evaluation Model (modelo de evaluación multi-fuente) |
| MVC | Modelo Vista Controlador |
| NNLS | Non-negative least squares (mínimos cuadrados con coeficientes positivos) |
| PAS | Personal de Administración y Servicios |
| PFC | Proyecto Fin de Carrera |
| PHP | PHP Hypertext Preprocessor (preprocesador de hipertexto PHP) |
| PIE | Proyecto de Innovación Educativa |
| RAE | Real Academia Española |
| REM | Rubric Element Model (modelo de elementos de las rúbricas) |
| REST | REpresentational State Transfer (transferencia de estado representacional) |
| RMSE | Root-Mean Squared Error (raíz del error cuadrático medio) |

ANEXO A

| | |
|----------------|--|
| SM | Scale-level Model (modelo de niveles de la escala) |
| TAM | Task Adjustmen Model (modelo ajustado de tareas) |
| TIC | Tecnologías de la Información y las Comunicaciones |
| TFG | Trabajo Fin de Grado |
| UML | Unified Modeling Language (lenguaje unificado de modelado) |
| UO | User Objective (objetivo de usuario) |
| UPV/EHU | Universidad del País Vasco / Euskal Herriko Unibertsitatea |

ANEXO B

ENCUESTA ALUMNADO

Para analizar los problemas y dificultades que el alumnado se encuentra durante la realización y evaluación de los Trabajos Fin de Grado se realizó una encuesta entre antiguos estudiantes que hubieran defendido su proyecto en el área de la Ingeniería Informática. Para ello, se diseñó una encuesta (utilizando Google Docs) y se distribuyó a través de las listas de distribución de la UPV/EHU y de las redes sociales para recabar la mayor cantidad de opiniones posibles.

La encuesta para estudiantes egresados se compone de 38 preguntas divididas en cuatro categorías: *Contextualización*, *Definición del proyecto*, *Durante la realización* y *Evaluación*. La encuesta contiene preguntas de tipo Sí/No, respuesta múltiple y abierta.

A continuación se presentan las preguntas que forman la encuesta (las marcadas con un asterisco son de respuesta obligatoria.) y un análisis de los principales datos y respuestas recogidas para cada una de las categorías.

Contextualización

En la categoría de *Contextualización* (ver Tabla B - 1) se encuentran las preguntas referentes a información demográfica y sobre los estudios cursados para poder contextualizar los resultados obtenidos.

Se recogieron las respuestas de 60 egresados, 46 hombres y 14 mujeres. Aunque la gran mayoría de estudiantes han cursado sus estudios en la UPV/EHU (pregunta 4), hay 5 personas que han estudiado en otras universidades.

El 83.3% de los estudiantes finaliza el primer proyecto que comienza (pregunta 7), mientras que el 16.7% restante trabajaron en una media de dos proyectos diferentes (pregunta 8) antes de finalizar uno de ellos.

Tabla B - 1.- Preguntas categoría Contextualización

| |
|--|
| 1. Sexo* |
| <input type="checkbox"/> Hombre |
| <input type="checkbox"/> Mujer |
| 2. Edad cuando comenzaste el proyecto* |
| <input type="checkbox"/> Menor de 21 años |
| <input type="checkbox"/> Entre 21 y 25 años |
| <input type="checkbox"/> Mayor de 25 años |
| 3. Número de asignaturas que tenías pendiente cuando comenzaste el proyecto* |
| <input type="checkbox"/> Ninguna |
| <input type="checkbox"/> Menos de 3 |
| <input type="checkbox"/> Entre 3 y 7 |
| <input type="checkbox"/> Más de 7 |
| 4. ¿En qué centro universitario realizaste tu proyecto?* |
| <input type="checkbox"/> Facultad de Informática de San Sebastián (UPV/EHU) |
| <input type="checkbox"/> Escuela Universitaria de Ingeniería Técnica de Bilbao (UPV/EHU) |
| <input type="checkbox"/> Escuela Universitaria de Ingeniería de Vitoria-Gasteiz (UPV/EHU) |
| <input type="checkbox"/> Facultad de Ciencias, Estudios Agroalimentarios e Informática (Universidad de La Rioja) |
| <input type="checkbox"/> Otro: |
| 5. ¿En qué titulación realizaste el proyecto?* |
| <input type="checkbox"/> Ingeniería Superior en Informática |
| <input type="checkbox"/> Ingeniería Técnica en Informática |
| <input type="checkbox"/> Ingeniería Técnica Industrial, especialidad Electricidad |
| <input type="checkbox"/> Ingeniería Técnica Industrial, especialidad Mecánica |
| <input type="checkbox"/> Ingeniería Técnica Industrial, especialidad Química Industrial |
| <input type="checkbox"/> Ingeniería Técnica Industrial, especialidad Electrónica Industrial |
| <input type="checkbox"/> Grado en Ingeniería Informática |
| <input type="checkbox"/> Grado en Ingeniería Eléctrica |
| <input type="checkbox"/> Grado en Ingeniería Electrónica Industrial y Automática |
| <input type="checkbox"/> Grado en Ingeniería Mecánica |
| <input type="checkbox"/> Grado en Ingeniería Química Industrial |
| <input type="checkbox"/> Otro: |
| 6. ¿En qué año defendiste el proyecto?* |
| 7. ¿El proyecto que finalmente defendiste, se correspondía con la primera idea de proyecto en la que trabajaste?* Entiéndase por trabajar en un proyecto el comenzar a buscar información, reunirse con la persona que lo va a supervisar, etc. |
| <input type="checkbox"/> Sí |
| <input type="checkbox"/> No |
| 8. Si el proyecto final no se correspondió con la primera idea en la que se trabajó, ¿en cuántas ideas distintas trabajaste? |
| 9. Si trabajaste en más de una idea, ¿bajo la supervisión de cuántas personas distintas comenzaste dichos trabajos? |
| 10. ¿Cuánto tiempo (en meses) tardaste en terminar el proyecto? (incluye los periodos de inactividad)* Considera el inicio del proyecto el momento en el que comenzaste a buscar información y/o a reunirte con la persona encargada de su supervisión. |

Cabe destacar que la gran mayoría de estudiantes comienza su proyecto cuando le quedan por cursar pocas asignaturas (pregunta 3, Fig. B - 1). Concretamente el 35% del alumnado lo comienza con tres o menos asignaturas pendientes y el 36.7% lo hace cuando ya no tiene ninguna asignatura pendiente. Esta forma de trabajar del alumnado no se corresponde con la intencionalidad de los planes de estudio de las titulaciones que han cursado los participantes en la encuesta, ya que todos ellos enmarcan la realización del proyecto en el último curso como una asignatura más del mismo.

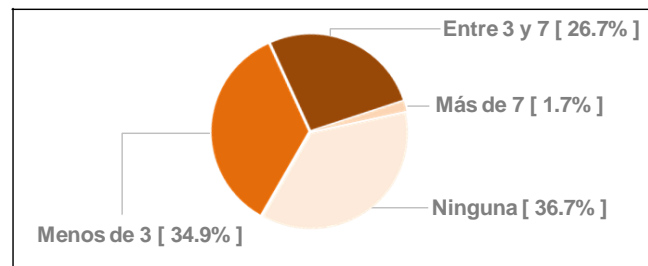


Fig. B - 1.- Resultados pregunta 3

En cuanto a la duración del proyecto expresada en meses (pregunta 10), las respuestas obtenidas son muy dispares, siendo la menor duración de 2 meses y la mayor de 60. La disparidad de valores hace que los datos se distribuyan de manera muy dispersa presentando una desviación de 10.22. La duración media se sitúa en 12.36 meses mientras que la mediana tiene un valor de 9 meses.

Definición del proyecto

La categoría *Definición del proyecto* contiene 6 preguntas referentes a aspectos de la definición del proyecto y sobre el conocimiento de la normativa (ver Tabla B - 2).

Algo más de la mitad del alumnado, el 55%, no sabe qué requisitos tiene que cumplir una idea para ser considerada un proyecto válido (pregunta 11, Fig. B - 2→a). Lo cual podría explicar porqué sólo el 25% de los proyectos que se llevan a cabo surgen de ideas del propio alumnado (pregunta 13, Fig. B - 2→b). Como se puede observar, la mayoría realiza proyectos definidos por el profesorado (40%) o por empresas (25%) a través del programa de prácticas en empresa.

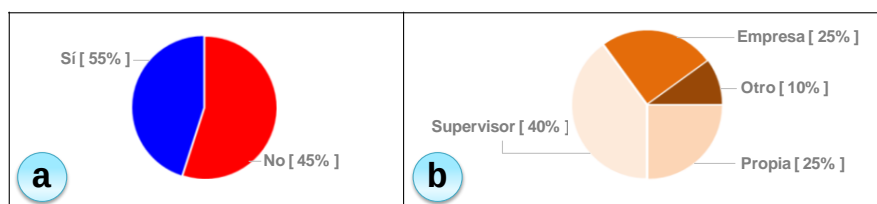


Fig. B - 2.- Resultados pregunta 11 (a) y pregunta 13 (b)

Es remarcable que únicamente el 71.7% del alumnado tiene definido el alcance del proyecto antes de comenzar (pregunta 14) y que sólo el 30.23% del alumnado realiza una planificación de cuánto tiempo y trabajo le va a costar llevar a cabo su proyecto (pregunta 16).

Tabla B - 2.- Preguntas categoría Definición del proyecto

| |
|---|
| <p>11. ¿Tenías claro qué requisitos tenía que cumplir una idea para poder ser considerada un proyecto válido en tu titulación?*</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>12. ¿Cómo obtuviste la información?*</p> <p><input type="checkbox"/> En la normativa</p> <p><input type="checkbox"/> A través del resto del alumnado</p> <p><input type="checkbox"/> A través de la persona encargada de la supervisión</p> <p><input type="checkbox"/> Consultando proyectos defendidos en convocatorias anteriores</p> <p><input type="checkbox"/> Otro:</p> |
| <p>13. ¿Cómo surgió la idea que finalmente llevaste a cabo como proyecto?*</p> <p><input type="checkbox"/> Idea propia</p> <p><input type="checkbox"/> Idea de la persona que supervisó el proyecto</p> <p><input type="checkbox"/> Idea de una empresa</p> <p><input type="checkbox"/> Otro:</p> |
| <p>14. ¿Estaba definido el alcance del proyecto desde el principio?*</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>15. ¿Se respetó la definición de dicho alcance?</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>16. ¿Realizaste una planificación de tu trabajo antes de comenzar?*</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |

Durante la realización

En la categoría denominada *Durante la realización* se encuentran 13 preguntas referentes a distintos aspectos relacionados con el trabajo realizado y con la relación con la persona encargada de la supervisión (ver Tabla B - 3 y Tabla B - 4).

Tabla B - 3.- Preguntas categoría Durante la realización (I)

| |
|---|
| 17. Aproximadamente, ¿cuántas veces te reuniste con la persona encargada de la supervisión de tu proyecto?* Entiéndase por reunión tanto reuniones presenciales como comunicaciones electrónicas |
| 18. ¿Consideras el número de reuniones adecuado?* Entiéndase por reuniones tanto reuniones presenciales como comunicaciones electrónicas |
| <input type="checkbox"/> Sí <input type="checkbox"/> No, fueron pocas reuniones <input type="checkbox"/> No, fueron demasiadas reuniones |
| 19. ¿Quién llevaba la iniciativa de convocar una reunión?* |
| <input type="checkbox"/> Tú <input type="checkbox"/> La persona encargada de la supervisión del proyecto <input type="checkbox"/> Estaban planificadas a priori |
| 20. ¿Alguna vez dejaste sin contestar algún requerimiento de información por parte de la persona que supervisaba tu proyecto?* |
| <input type="checkbox"/> No, nunca <input type="checkbox"/> Una o dos veces <input type="checkbox"/> Más de dos veces |
| 21. Si lo hiciste, ¿cuál fue la razón? |
| 22. ¿Existe algún proceso que haya que seguir para elaborar proyectos del tipo del que tú desarrollaste?* Entiéndase como proceso una serie de pasos más o menos ordenados indicando en qué momento hay que hacer cada parte del proyecto, su documentación, etc. |
| <input type="checkbox"/> Sí <input type="checkbox"/> No <input type="checkbox"/> No lo sé |
| 23. En caso de que exista y conozcas dicho proceso, ¿lo seguiste? |
| <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 24. Si existe el proceso y no lo seguiste, por favor, indica las razones |

De las respuestas recogidas a estas preguntas se extrae que el número medio de reuniones entre alumnado y profesorado supervisor es de 18.7 reuniones (pregunta 17, Tabla B - 3). Sin embargo, los valores recogidos son muy dispares presentando un valor mínimo de 3 reuniones y un máximo de 200. La mediana de la distribución es 10 y la desviación estándar tiene un valor de 26.32. El 70% del alumnado considera que el número de reuniones que tuvo con la persona

encargada de la supervisión fue el adecuado y sólo un 1.7% considera que fue excesivo (pregunta 18, Tabla B - 3).

Tabla B - 4.- Preguntas categoría Durante la realización (II)

| |
|---|
| <p>25. Durante la realización de tu proyecto, ¿echaste en falta en algún momento un mayor apoyo por parte de la persona que te supervisaba?*</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>26. Si tenías una planificación inicial realizada, ¿se cumplió?</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No, se retrasó</p> <p><input type="checkbox"/> No, se adelantó</p> |
| <p>27. Durante el desarrollo de tu proyecto, ¿fuiste anotando/recogiendo en algún sitio el trabajo que realizabas y cuánto tiempo te llevaba?*</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>28. Si fuiste anotando el trabajo que realizabas, ¿qué herramientas usaste para recogerlo? Marca todas las herramientas que usaste</p> <p><input type="checkbox"/> Hoja de cálculo</p> <p><input type="checkbox"/> Software de imputación de horas</p> <p><input type="checkbox"/> Blog o cualquier otro tipo de bitácora electrónica</p> <p><input type="checkbox"/> Anotaciones en papel</p> <p><input type="checkbox"/> Otro:</p> |
| <p>29. ¿Te parece que establecer una serie de entregas intermedias facilitaría el desarrollo del proyecto?*</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |

Aproximadamente la mitad del alumnado (el 53.33%) conoce la existencia de un proceso de desarrollo indicado para proyectos similar al realizado (pregunta 22 en la Tabla B - 3, Fig. B - 3→a). Sin embargo, a pesar de conocer su existencia, un 37.84% no lo aplica (pregunta 23 en la Tabla B - 3, Fig. B - 3→b), por centrarse en «conseguir algo que funcione rápidamente» (pregunta 24, Tabla B - 3).

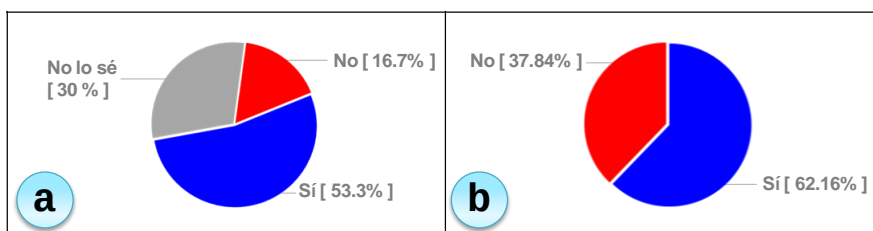


Fig. B - 3.- Resultados pregunta 22 (a) y pregunta 23 (b)

Un 30% del alumnado echa en falta un mayor apoyo por parte de la persona supervisora (pregunta 25, Tabla B - 4) y el 96.7% opina que establecer entregas intermedias facilitaría el desarrollo del proyecto (pregunta 29, Tabla B - 4).

De las planificaciones iniciales realizadas, únicamente el 19.64% se cumplen (pregunta 26, Tabla B - 4). En el 78.57% de los casos el proyecto finaliza retrasado y sólo un 1.79% de los proyectos finalizan antes de lo planificado. Algo más de la mitad del alumnado (exactamente el 53.3%) no toma anotaciones sobre el trabajo realizado y el tiempo dedicado a lo largo del desarrollo del proyecto (pregunta 27, Tabla B - 4). El 46.7% del alumnado que sí toma notas, usa para ello de manera independiente o complementaria: el 32.26% hojas de cálculo, el 29.03% anotaciones en papel, el 22.58% algún tipo de software de imputación de horas de trabajo y un 16.13% otro tipo de herramientas (pregunta 28, Tabla B - 4).

Evaluación

En la categoría de *Evaluación* se encuentran 9 preguntas relativas al proceso de evaluación del proyecto y a la opinión del alumnado sobre dicho proceso (ver Tabla B - 5).

La mayoría del alumnado (el 53.33%) desconoce antes de ser evaluado qué aspectos van a ser tenidos en cuenta a la hora de calificar su trabajo (pregunta 30, Fig. B - 4→a). Además, un 40% sigue sin saber qué se ha valorado tras haber sido evaluado (pregunta 31, Fig. B - 4→b).

Al ser preguntados sobre qué aspectos del trabajo consideran que influyen en la evaluación (pregunta 34), sólo el 15% del alumnado opina que el único factor influyente es el trabajo realizado. Un 85% opina que la composición del tribunal de evaluación influye en la calificación obtenida, y un 55% cree que quién sea la persona encargada de la supervisión del proyecto también influye.

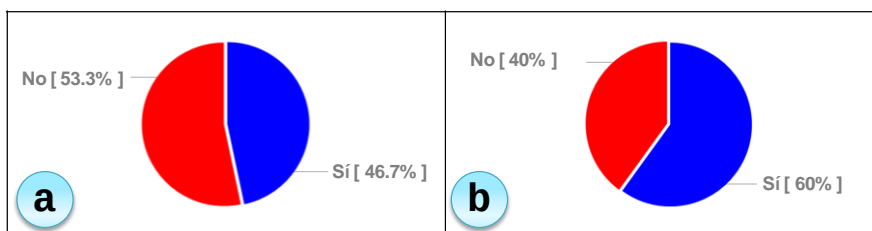


Fig. B - 4.- Resultados pregunta 30 (a) y pregunta 31 (b)

Tabla B - 5.- Preguntas categoría Evaluación

| | |
|--|---|
| 30. ¿Conocías qué aspectos del proyecto se iban a valorar antes de que te evaluaran?* | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 31. Una vez evaluado tu proyecto, ¿sabes qué aspectos del mismo se valoraron?* | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 32. ¿Consideras que la evaluación de tu proyecto fue justa?* | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 33. ¿Consideras que la evaluación de todos los proyectos es equitativa?* | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 34. ¿Qué factores consideras que influyen en la evaluación de un proyecto?*. Marca todas las opciones que consideres | <input type="checkbox"/> Sólo el trabajo realizado <input type="checkbox"/> La persona supervisora <input type="checkbox"/> El tribunal de evaluación <input type="checkbox"/> El expediente académico <input type="checkbox"/> Otro: |
| 35. ¿Asististe a la evaluación de algún proyecto anterior a la tuya?* | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 36. ¿Considerarías útil que se fuera evaluando tu trabajo según lo vas realizando?* | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 37. ¿Qué nota obtuviste en tu trabajo?* | <input type="checkbox"/> Suficiente <input type="checkbox"/> Notable <input type="checkbox"/> Sobresaliente |
| 38. Cualquier otro comentario/sugerencia que quieras realizar sobre el proceso de evaluación de un proyecto: | |

A pesar de que el 85% del alumnado considera que la evaluación de su proyecto fue justa (pregunta 32), el 70% considera que no es equitativa para todos los proyectos (pregunta 33). Analizando los datos en profundidad y tal y como se muestra en la Fig. B - 5, el porcentaje del alumnado que considera que la evaluación no es equitativa, es muy superior entre aquellos que asistieron a otras evaluaciones aparte de la suya propia (pregunta 35).

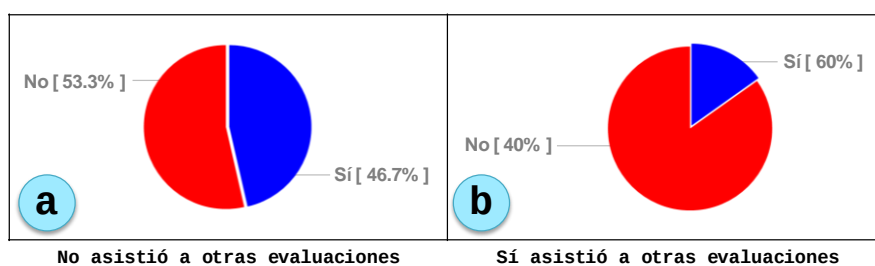


Fig. B - 5.- Resultados pregunta 33 en función de la pregunta 35

Un 80% del alumnado opina que sería útil establecer algún modo de evaluación continua que permitiera ir evaluando el trabajo conforme se va realizando (pregunta 36).

ANEXO C

ENCUESTA PROFESORADO

Para analizar los problemas y dificultades que el profesorado se encuentra durante la supervisión y la evaluación de los Trabajos Fin de Grado se realizó una encuesta entre el profesorado. Para ello, se diseñó una encuesta (utilizando Google Docs) y se distribuyó a través de las listas de distribución de la UPV/EHU.

La encuesta para el profesorado se compone de 52 preguntas divididas en cuatro categorías: *Contextualización*, *Experiencia*, *Supervisión* y *Evaluación*. La encuesta contiene preguntas de tipo Sí/No, respuesta múltiple y abierta. Además, incorpora también una pregunta de respuesta abierta para permitir al profesorado expresar opiniones o sugerencias sobre cualquier aspecto relacionado con los TFG no tratado en las categorías anteriores.

A continuación se presentan las preguntas que forman la encuesta (las marcadas con un asterisco son de respuesta obligatoria.) y un análisis de las respuestas recogidas para cada una de las categorías.

Contextualización

La categoría *Contextualización* se compone de 5 preguntas referentes a la información demográfica de los participantes (ver Tabla C - 1).

Se recogieron las respuestas de 56 docentes, 34 hombres y 22 mujeres, pertenecientes todos ellos a centros de la UPV/EHU (preguntas 1 y 3). El 78.6% del profesorado participante ha finalizado el doctorado, mientras que el 21.4% restante ha finalizado una ingeniería superior o similar (pregunta 5).

Tabla C - 1.- Preguntas categoría Contextualización

| |
|--|
| 1. Sexo* |
| <input type="checkbox"/> Hombre |
| <input type="checkbox"/> Mujer |
| 2. Edad* |
| <input type="checkbox"/> Menor de 35 años |
| <input type="checkbox"/> 36 - 45 años |
| <input type="checkbox"/> 46 -55 años |
| <input type="checkbox"/> Más de 55 años |
| 3. Centro de adscripción* |
| <input type="checkbox"/> Facultad de Informática (UPV/EHU) |
| <input type="checkbox"/> Escuela Universitaria de Ingeniería Técnica Industrial de Bilbao (UPV/EHU) |
| <input type="checkbox"/> Escuela Universitaria de Ingeniería de Vitoria-Gasteiz (UPV/EHU) |
| <input type="checkbox"/> Facultad de Ciencias, Estudios Agroalimentarios e Informática (Universidad de La Rioja) |
| <input type="checkbox"/> Otro: |
| 4. Departamento de adscripción* |
| 5. Formación académica* |
| <input type="checkbox"/> Doctorado |
| <input type="checkbox"/> Ingeniería Superior |
| <input type="checkbox"/> Ingeniería Técnica |
| <input type="checkbox"/> Otro: |

Experiencia

La categoría de Experiencia contiene 3 preguntas referentes a la experiencia del profesorado participante en la supervisión de proyectos (ver Tabla C - 2).

El 85.8% del profesorado encuestado tiene más de cinco años de experiencia en la supervisión de proyectos y más de la mitad del total del profesorado (el 55.4%) superan los diez años (pregunta 7, Fig. C - 1→a).

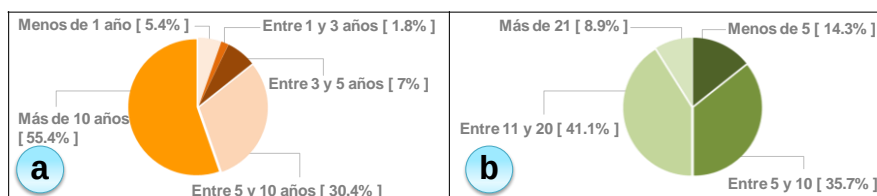


Fig. C - 1.- Resultados pregunta 7 (a) y pregunta 8 (b)

Durante los últimos cinco años, el 85.7% del profesorado encuestado ha supervisado más de cinco proyectos. Concretamente, el 8.9% ha supervisado más de 21 proyectos, el 41.1% entre 11 y 20 proyectos y el 35.7% entre 5 y 10 (pregunta 8, Fig. C - 1→b)

Tabla C - 2.- Preguntas categoría Experiencia

| |
|---|
| <p>6. Indique las titulaciones en las que ha supervisado al menos un proyecto* Aunque sólo figuran los Grados, marque la opción si supervisó proyectos en las titulaciones correspondientes en el plan anterior a Bolonia</p> <p><input type="checkbox"/> Grado en Ingeniería Informática</p> <p><input type="checkbox"/> Grado en Ingeniería Informática de Gestión y Sistemas de Información</p> <p><input type="checkbox"/> Grado en Ingeniería Eléctrica</p> <p><input type="checkbox"/> Grado en Ingeniería Electrónica Industrial y Automática</p> <p><input type="checkbox"/> Grado en Ingeniería Mecánica</p> <p><input type="checkbox"/> Grado en Ingeniería Química Industrial</p> <p><input type="checkbox"/> Otro:</p> |
| <p>7. ¿Cuánto tiempo hace que comenzó a supervisar proyectos?*</p> <p><input type="checkbox"/> Menos de un año</p> <p><input type="checkbox"/> Entre uno y tres años</p> <p><input type="checkbox"/> Entre tres y cinco años</p> <p><input type="checkbox"/> Entre cinco y diez años</p> <p><input type="checkbox"/> Más de diez años</p> |
| <p>8. Número de proyectos supervisados en los últimos 5 años*</p> <p><input type="checkbox"/> Menos de 5</p> <p><input type="checkbox"/> Entre 5 y 10</p> <p><input type="checkbox"/> Entre 11 y 20</p> <p><input type="checkbox"/> Más de 21</p> |

Supervisión

La categoría *Supervisión* contiene 25 preguntas relativas al proceso de supervisión de los proyectos por parte del profesorado (Tabla C - 3, Tabla C - 4 y Tabla C - 5).

Como el profesorado puede supervisar proyectos en distintas titulaciones, se solicitó a los participantes que respondieran a las preguntas ateniéndose exclusivamente a la titulación en la que tuvieran mayor experiencia (pregunta 9, Tabla C - 3).

De acuerdo a las respuestas del profesorado, para realizar un proyecto de calidad media (calificación de 6-7 puntos) el alumnado debería trabajar una media de 29.4 horas por cada crédito (preguntas 10 y 11, Tabla C - 3).

Tabla C - 3.- Preguntas categoría Supervisión (I)

| |
|---|
| <p>9. Por favor, indique respecto a qué titulación va a responder a las preguntas*</p> <p><input type="checkbox"/> Ingeniería Superior en Informática</p> <p><input type="checkbox"/> Ingeniería Técnica en Informática</p> <p><input type="checkbox"/> Ingeniería Técnica Industrial, especialidad electricidad</p> <p><input type="checkbox"/> Ingeniería Técnica Industrial, especialidad Mecánica</p> <p><input type="checkbox"/> Ingeniería Técnica Industrial, especialidad Electrónica Industrial</p> <p><input type="checkbox"/> Ingeniería Técnica Industrial, especialidad Química Industrial</p> <p><input type="checkbox"/> Grado en Ingeniería Informática</p> <p><input type="checkbox"/> Grado en Ingeniería Informática de Gestión y Sistemas de Información</p> <p><input type="checkbox"/> Grado en Ingeniería Eléctrica</p> <p><input type="checkbox"/> Grado en Ingeniería Electrónica Industrial y Automática</p> <p><input type="checkbox"/> Grado en Ingeniería Mecánica</p> <p><input type="checkbox"/> Grado en Ingeniería Química Industrial</p> <p><input type="checkbox"/> Otro:</p> |
| <p>10. ¿Cuántos créditos tiene el proyecto en la titulación?*</p> |
| <p>11. Aproximadamente, ¿cuántas horas considera que debería trabajar el alumnado para realizar un proyecto de calidad media?* Entiéndase por calidad media un proyecto que en su valoración obtendría una nota de 6-7</p> |
| <p>12. De manera general, ¿considera que el alumnado trabaja más horas de las que debiera para obtener un proyecto de calidad media?* Entiéndase por calidad media un proyecto que en su valoración obtendría una nota de 6-7</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>13. Si ha respondido positivamente a la pregunta anterior, ¿cuáles cree que son las razones por las que el alumnado invierte más horas de las necesarias? Puede marcar tantas opciones como desee</p> <p><input type="checkbox"/> Mala planificación</p> <p><input type="checkbox"/> Poca supervisión por parte de la persona encargada</p> <p><input type="checkbox"/> Mala formación de base</p> <p><input type="checkbox"/> Proyecto demasiado ambicioso</p> <p><input type="checkbox"/> Otro:</p> |
| <p>14. ¿Qué porcentaje de los proyectos que ha supervisado no se ha terminado debido al abandono del alumnado?*</p> |
| <p>15. ¿Qué porcentaje de los proyectos que ha supervisado se han retrasado sustancialmente con respecto a la planificación inicial?*</p> |
| <p>16. ¿En alguna ocasión ha rechazado continuar con la supervisión de algún proyecto con el que se había comprometido?*</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>17. En caso afirmativo, por favor indique las razones que le llevaron a rechazar la continuación de la supervisión</p> |

Sin embargo, los resultados recogidos son muy variados (ver Fig. C - 2) y van desde las tres personas que opinan que el número de horas a invertir por cada crédito debería ser inferior a 10, hasta las dos personas que creen que el alumnado debería trabajar más de 80 horas por cada crédito del proyecto. La

mediana de la distribución que indica el número de horas a invertir por cada crédito en opinión del profesorado es 25.

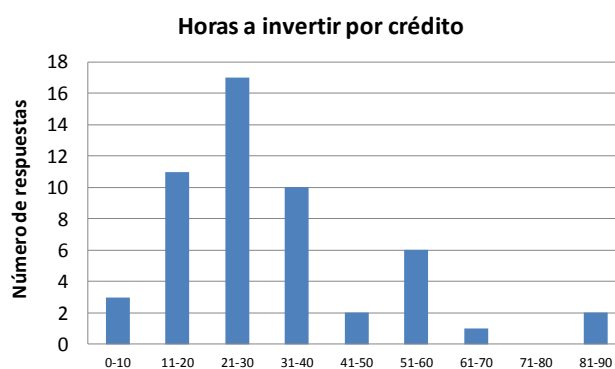


Fig. C - 2.- Resultados pregunta 10 y 11

La mayoría del profesorado (concretamente el 66.1%) opina que el alumnado trabaja más horas de las que debiera para realizar su proyecto (pregunta 12, Tabla C - 3). En opinión del profesorado, las principales razones para este exceso son (pregunta 13, Tabla C - 3): la mala planificación realizada (65.8%), la mala formación del alumnado (44.7%), la complejidad excesiva del proyecto (44.7%) y la falta de supervisión por parte de la persona encargada (21.1%).

De acuerdo a las respuestas recogidas, el porcentaje de proyectos abandonados por el alumnado antes de su finalización presenta una media del 7.25% y una mediana del 5% (pregunta 14, Tabla C - 3). De esas respuestas también se desprende que el porcentaje de proyectos retrasados presenta una media del 43.10% y una mediana del 40% (pregunta 15, Tabla C - 3 y Fig. C - 3)

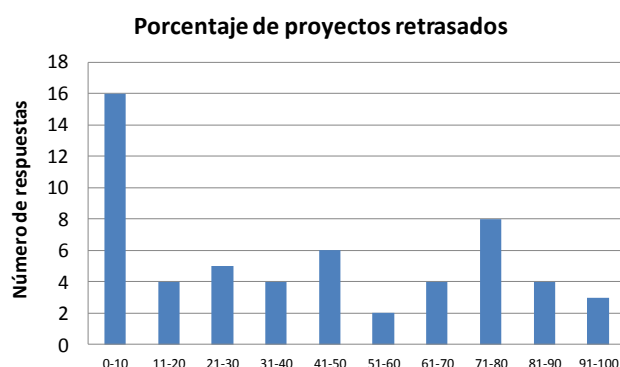


Fig. C - 3.- Resultados pregunta 15

En la mayoría de los casos (el 57.1%) el profesorado responde que no existe ningún proceso que establezca los pasos que el alumnado debe dar durante la realización de su proyecto (pregunta 18, Tabla C - 4). Pero en los casos en los que sí existe un proceso definido, sólo una media del 68.40% del alumnado lo ha seguido correctamente (pregunta 19, Tabla C - 4).

Para mejorar la supervisión sobre el proceso de realización de un proyecto (pregunta 21, Tabla C - 4), el profesorado sugiere: «...marcar hitos que los alumnos tuviesen que cumplir obligatoriamente...», «...realizar evaluaciones intermedias...», «...establecer entregas parciales y penalizar su retraso...» e «...incrementar la comunicación con el alumno» entre otras medidas. Estas sugerencias concuerdan con las respuestas a la pregunta 20 (Tabla C - 4), donde el 91.1% del profesorado a respondido que sí cree que establecer hitos intermedios durante el proceso pueda ser beneficioso.

Tabla C - 4.- Preguntas categoría Supervisión (II)

| |
|--|
| <p>18. ¿Existe en su titulación algún proceso definido que el alumnado debería seguir para elaborar su proyecto?* Entiéndase como proceso una serie de pasos con un orden más o menos riguroso que sería conveniente que el alumnado siguiera para finalizar con éxito su proyecto</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>19. En caso de que exista dicho proceso... ¿qué porcentaje de los proyectos que ha supervisado han seguido correctamente dicho proceso?</p> |
| <p>20. ¿Considera que establecer hitos en la realización del proyecto podría ayudar a controlar el proceso?* Entiéndase por hitos una serie de entregas obligatorias con partes concretas del proyecto</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>21. Por favor, indique cualquier tipo de acción que considera que podría mejorar la supervisión del proceso de realización de un proyecto</p> |
| <p>22. En general, ¿Cómo organiza las reuniones con el alumnado que supervisa?*Obligatorio</p> <p><input type="checkbox"/> Se establece un calendario a priori (ya sea propuesto por el alumnado o por la persona encargada de la supervisión)</p> <p><input type="checkbox"/> Se organiza una reunión cuando lo solicita el alumnado</p> <p><input type="checkbox"/> A instancia suya se organiza una reunión de vez en cuando sin seguir un calendario establecido</p> <p><input type="checkbox"/> No se reúne con el alumnado</p> <p><input type="checkbox"/> Otro:</p> |
| <p>23. Si establece un calendario de reuniones a priori, ¿suele cumplirse dicho calendario?</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>24. ¿Cuáles son los principales temas de conversación de las reuniones?* Marque tantas casillas como desee</p> <p><input type="checkbox"/> Estado del proyecto</p> <p><input type="checkbox"/> Dudas técnicas</p> <p><input type="checkbox"/> Dudas administrativas (requisitos para la presentación, criterios de evaluación, etc.)</p> <p><input type="checkbox"/> Otro:</p> |

El 50% del profesorado se reúne con el alumnado que supervisa en función de cuándo lo solicita el alumnado (pregunta 22, Tabla C - 4). Sin embargo, un 32.1% del profesorado considera que debería reunirse con el alumnado más a menudo (pregunta 25, Tabla C - 5).

El 73.2% del profesorado no utiliza ninguna herramienta para realizar el seguimiento del proyecto (pregunta 29, Tabla C - 5). Entre el profesorado que usa alguna herramienta para supervisar los proyectos, la usada mayoritariamente son los servicios de almacenamiento en la nube como Dropbox y similares (pregunta 30, Tabla C - 5).

Tabla C - 5.- Preguntas categoría Supervisión (III)

| | |
|---|--|
| 25. En general, ¿considera que debería reunirse más a menudo con el alumnado bajo su supervisión?* | Obligatorio |
| <input type="checkbox"/> Sí | |
| <input type="checkbox"/> No | |
| 26. ¿Cómo se comunica con el alumnado al que supervisa el proyecto?* | Obligatorio |
| <input type="checkbox"/> Sobre todo de manera presencial | |
| <input type="checkbox"/> Sobre todo de manera electrónica | |
| <input type="checkbox"/> Presencial o electrónicamente de manera indistinta | |
| 27. ¿Ha supervisado algún proyecto de manera virtual?* | Entiéndase por supervisar virtualmente el hecho de hacerlo sin establecer contactos presenciales con el alumnado |
| <input type="checkbox"/> Sí | |
| <input type="checkbox"/> No | |
| 28. Si ha supervisado algún proyecto de manera virtual, por favor explique brevemente su experiencia indicando aspectos positivos y negativos de la misma | |
| 29. ¿Utiliza alguna herramienta para realizar el seguimiento del proyecto?* | Obligatorio |
| <input type="checkbox"/> Sí | |
| <input type="checkbox"/> No | |
| 30. ¿Cuáles? | |
| 31. ¿Tiene alguna forma de comprobar el estado de un proyecto si el alumnado que lo está realizando no responde a sus comunicaciones?* | Obligatorio |
| <input type="checkbox"/> Sí | |
| <input type="checkbox"/> No | |
| 32. Si no tiene forma de comprobar el estado del proyecto sin la colaboración del alumnado, ¿le parecería útil poder hacerlo? | |
| <input type="checkbox"/> Sí | |
| <input type="checkbox"/> No | |
| 33. ¿Cuáles son los principales problemas a los que se ha enfrentado durante la realización de un proyecto bajo su supervisión? | |

La mayoría del profesorado (el 87.5%) reconoce que no dispone de ninguna forma para comprobar el estado del proyecto sin precisar la colaboración del alumnado (pregunta 31, Tabla C - 5 y Fig. C - 4→a). Sin embargo, al 76.5% le parecería útil disponer de algún medio para poder hacerlo (pregunta 32, Tabla C - 5 y Fig. C - 4→b).

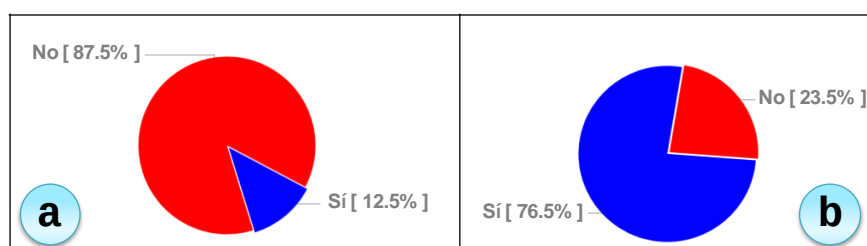


Fig. C - 4.- Resultados preguntas 31 (a) y 32 (b)

Al preguntar sobre los principales problemas que se ha encontrado el profesorado durante la realización de los proyectos (pregunta 33, Tabla C - 5), la respuesta más frecuente (34.2% de las respuestas) hace referencia a la poca autonomía del alumnado. También destaca (15.8% de las respuestas) que el ritmo de realización del proyecto no se corresponde con el planificado inicialmente debido a razones tales como causas laborales, problemas personales o necesidad de dedicar más tiempo a otras asignaturas.

Evaluación

La categoría de *Evaluación* está formada por 19 preguntas relativas a la opinión del profesorado sobre diferentes aspectos de la evaluación de los proyectos (Tabla C - 6, Tabla C - 7 y Tabla C - 8).

Tabla C - 6.- Preguntas categoría Evaluación (I)

| | |
|--|--|
| 34. ¿Considera que en su titulación se han iniciado proyectos cuya idea inicial no tenía la calidad mínima exigible?*Obligatorio | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 35. ¿Existe en su titulación algún criterio que establezca los requisitos mínimos que debe cumplir un proyecto?*Obligatorio | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 36. En caso de no existir dicho criterio, ¿considera que debería haberlo? | <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 37. En caso de existir dicho criterio, por favor, explíquelo | |

En cuanto al alcance de los proyectos, el 51.8% del profesorado cree que en su titulación se han iniciado proyectos cuya idea inicial no alcanzaba la calidad mínima exigible (pregunta 34, Tabla C - 6). Además, el 80.4% del profesorado reconoce que en su titulación no existe ningún criterio para establecer dicha

calidad mínima (pregunta 35, Tabla C - 6 y Fig. C - 5→a). Cabe destacar que el 84.4% opina que debería existir dicho criterio (pregunta 36, Tabla C - 6 y Fig. C - 5→b).

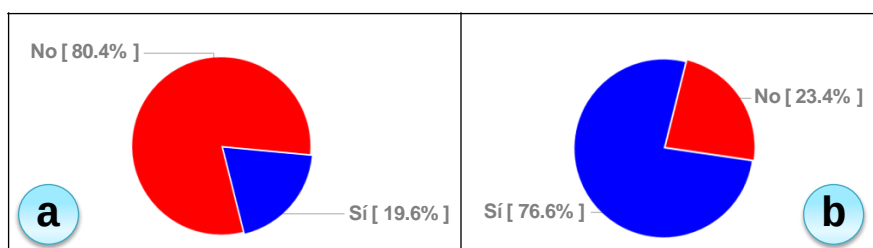


Fig. C - 5.- Resultados preguntas 35 (a) y 36 (b)

En los casos en los que según el profesorado sí existen criterios establecen la calidad mínima de los proyectos (19.6%, ver Fig. C - 5→a), estos criterios dependen de la opinión subjetiva de los componentes de una comisión de profesorado (pregunta 37, Tabla C - 6).

Esto influye en que el 73.2% del profesorado considere que se han defendido proyectos que no cumplían con la calidad mínima esperable (pregunta 38, Tabla C - 7 y Fig. C - 6).

Tabla C - 7.- Preguntas categoría Evaluación (II)

| |
|---|
| 38. ¿Considera que se han defendido proyectos que no llegaban a un mínimo de calidad?*Obligatorio <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 39. ¿Existe en su titulación alguna normativa o protocolo que asegure que todos los proyectos son evaluados de manera equitativa e imparcial?*Obligatorio <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 40. Si no existe dicha normativa, ¿considera que debería existir? <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 41. ¿Considera que han existido proyectos que han obtenido un aprobado y no deberían haberlo hecho debido a su calidad?*Obligatorio <input type="checkbox"/> Sí <input type="checkbox"/> No |
| 42. ¿Considera que han existido proyectos muy similares que han obtenido calificaciones significativamente diferentes?*Obligatorio <input type="checkbox"/> Sí <input type="checkbox"/> No |

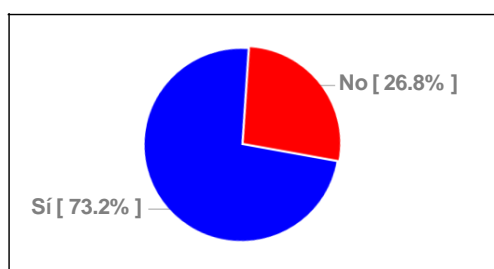


Fig. C - 6.- Respuestas a la pregunta 38

La mayoría del profesorado (el 76.8%) reconoce que en su titulación no existe ninguna normativa que asegure la evaluación equitativa e imparcial de todos los proyectos (pregunta 39, Tabla C - 7 y Fig. C - 7→a). Sin embargo, el 84.4% del profesorado opina que dicha normativa debería existir (pregunta 40, Tabla C - 7 y Fig. C - 7→b).

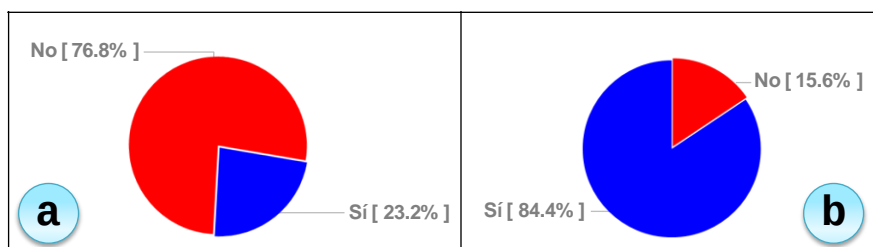


Fig. C - 7.- Resultados preguntas 39 (a) y 40 (b)

Al ser preguntados sobre los resultados de las evaluaciones, el 62.5% del profesorado opina que han existido proyectos que obtuvieron un aprobado sin merecerlo (pregunta 41, Tabla C - 7 y Fig. C - 8→a). Además, el 71.4% considera que se han evaluado proyectos muy similares con calificaciones significativamente diferentes (pregunta 42, Tabla C - 7 y Fig. C - 8→b).

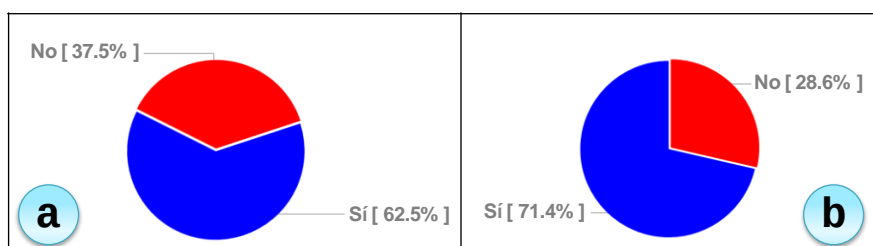


Fig. C - 8.- Resultados preguntas 41 (a) y 42 (b)

La Fig. C - 9 muestra los resultados respecto a los aspectos del proyecto que el profesorado tiene en cuenta a la hora de realizar la evaluación de un proyecto (pregunta 43, Tabla C - 8). Casi todo el profesorado valora la calidad de la documentación del proyecto (el 92.9%), la dificultad técnica del mismo (el 89.3%) y la defensa ante el tribunal (el 85.7%). Los aspectos relativos al proceso de realización y a la actitud del alumno también son valorados por más de la mitad del profesorado (el 64.3% en ambos casos).

En el apartado «Otros» de esta pregunta se han recogido aspectos como: el grado de innovación del proyecto (5 respuestas), su originalidad (3 respuestas) y su utilidad (1 respuesta).

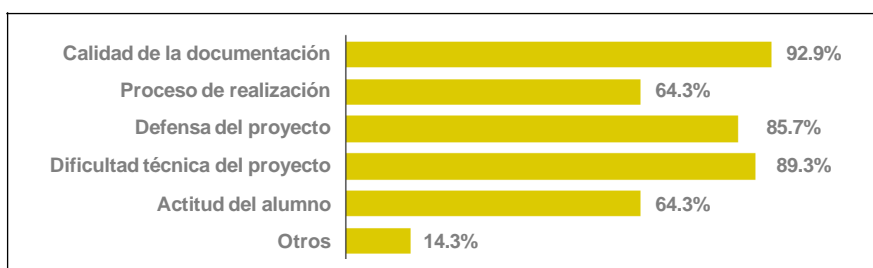


Fig. C - 9.- Resultados pregunta 43

En cuanto a las herramientas utilizadas, el 32.1% del profesorado encuestado no utiliza ninguna a la hora de evaluar los proyectos, mientras que el 62.5% utiliza simplemente anotaciones en papel (pregunta 44, Tabla C - 8).

Tabla C - 8.- Preguntas categoría Evaluación (III)

| |
|--|
| <p>43. ¿Qué aspectos del proyecto tiene en cuenta a la hora de valorarlo?* Por favor, indique todos los que use</p> <p><input type="checkbox"/> Calidad de la documentación</p> <p><input type="checkbox"/> Proceso de realización</p> <p><input type="checkbox"/> Defensa del proyecto</p> <p><input type="checkbox"/> Dificultad técnica del proyecto</p> <p><input type="checkbox"/> Actitud del alumno durante la realización</p> <p><input type="checkbox"/> Otro:</p> |
| <p>44. ¿Qué herramientas usa para realizar la evaluación de un proyecto?* Marque todas las que use</p> <p><input type="checkbox"/> Ninguna</p> <p><input type="checkbox"/> Hojas de cálculo</p> <p><input type="checkbox"/> Anotaciones en papel</p> <p><input type="checkbox"/> Moodle u otros sistemas similares</p> <p><input type="checkbox"/> Otro:</p> |
| <p>45. ¿Usa rúbricas o listas de requisitos mínimos a la hora de evaluar?*Obligatorio</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>46. En caso de que use rúbricas o listas de requisitos mínimos, ¿cómo son?</p> <p><input type="checkbox"/> Son de elaboración propia</p> <p><input type="checkbox"/> Están consensuados entre varios docentes, pero no son oficiales</p> <p><input type="checkbox"/> Son oficiales de la titulación</p> <p><input type="checkbox"/> Otro:</p> |
| <p>47. ¿Se permite la realización de proyectos en empresa?*Obligatorio</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>48. Si se permite la realización de proyectos en empresa, ¿participa alguien de la empresa en el proceso de evaluación del proyecto?</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>49. ¿Considera que debería hacerlo?</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |
| <p>50. ¿Cuáles son los principales problemas a los que se ha enfrentado durante la EVALUACIÓN de un proyecto? Puede ser tanto en un proyecto elaborado bajo su supervisión, como por su participación en tribunales de evaluación</p> |
| <p>51. ¿Cuándo se realiza la evaluación de los proyectos?*Obligatorio</p> <p><input type="checkbox"/> De manera continua durante la realización del mismo</p> <p><input type="checkbox"/> Cuando el proyecto ya está finalizado</p> <p><input type="checkbox"/> De manera continua a lo largo de la realización y a su finalización</p> |
| <p>52. ¿Considera que realizar una evaluación continua a lo largo del proyecto podría ser beneficioso para el alumnado?*Obligatorio</p> <p><input type="checkbox"/> Sí</p> <p><input type="checkbox"/> No</p> |

El 69.6% del profesorado no utiliza rúbricas ni ningún otro instrumento de evaluación similar (pregunta 45, Tabla C - 8 y Fig. C - 10→a). Entre el profesorado que sí utiliza este tipo de instrumentos de evaluación, en ningún caso son oficiales. En el 76.5% de los casos son instrumentos de elaboración propia y en el 23.5% restante están elaborados y consensuados entre varios docentes, pero sin llegar a ser oficiales (pregunta 46, Tabla C - 8 y Fig. C - 10→b).

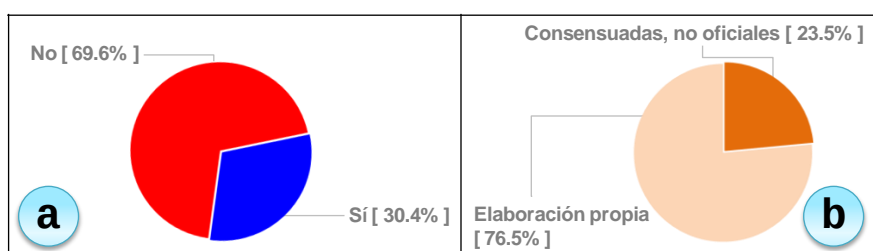


Fig. C - 10.- Resultados preguntas 45 (a) y 46 (b)

Respecto a los principales problemas encontrados por el profesorado durante la evaluación de los proyectos, destaca la diferencia de criterios entre los componentes del tribunal (el 39.39% de las respuestas). El 15.15% de las respuestas recogidas hace referencia a la mala calidad de la documentación entregada y el 12.12% a la falta de conocimientos de los componentes del tribunal sobre la temática del proyecto.

En el 78.6% de los casos el proyecto se evalúa una vez que está finalizado (pregunta 51, Tabla C - 8). Sin embargo, el 66.1% del profesorado considera que establecer una evaluación continua a lo largo de la realización del proyecto sería beneficioso para el alumnado (pregunta 52, Tabla C - 8).

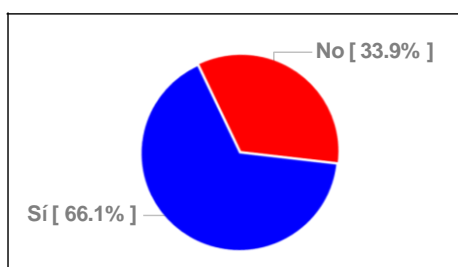


Fig. C - 11.- Resultados pregunta 52

ANEXO D

RÚBRICAS DE EVALUACIÓN

En este anexo se presentan las versiones definitivas de las rúbricas de evaluación desarrolladas en este trabajo de tesis para cada uno de los elementos evaluables establecidos por la metodología FPM.

En primer lugar se presentan las rúbricas analíticas y a continuación las rúbricas atomísticas.

Rúbricas analíticas

Eva_Requisitos

| | Excelente | Muy adecuado | Adecuado | Básico | Inadecuado |
|---------------------------------|--|---|---|---|--|
| Objetivos | Explica de forma clara, concisa y concreta cada uno de los objetivos propuestos, lo que permite comprobar fácilmente si se han alcanzado. Prioriza la secuencia de su realización. | Explica de forma clara los objetivos propuestos, expresándolos en términos que admiten una sola interpretación. | Incluye una relación de objetivos, redactados de forma apropiada, aunque la claridad y/o concreción de por lo menos algunos de ellos no permite comprobar fácilmente si se han alcanzado. | Expresa los objetivos propuestos de manera excesivamente general y/o no concretados adecuadamente. | Explica de forma difusa y muy poco concreta, con lo que se desprende una relación errática de objetivos concretos a alcanzar en el TFG. |
| Antecedentes | Indica claramente el punto de partida y realiza un estudio exhaustivo de soluciones similares y de las fuentes más relevantes. | Indica claramente el punto de partida del TFG. Realiza comparativas con soluciones similares. Menciona fuentes relevantes relacionadas. | Indica de manera confusa el punto de partida del TFG. Realiza comparativas básicas con soluciones similares. Menciona algunas de las fuentes relevantes. | Realiza comparativas con soluciones similares pero se queda en un nivel superficial. Las fuentes que menciona no son relevantes. | No realiza comparativas con soluciones similares. No menciona las fuentes bibliográficas relacionadas. |
| Planificación | La planificación está detallada y para cada tarea se justifican tanto el tiempo como los recursos necesarios. La planificación es la más indicada para el TFG. | Define correctamente las actividades necesarias para realizar el trabajo y su progresión temporal. La planificación no es la más adecuada. | Realiza una planificación donde las tareas y el tiempo están identificados de manera concreta para el TFG. No realiza ninguna justificación y no da detalles. | Presenta una planificación del tiempo y de los recursos para el TFG, pero no da detalles. Las tareas que definen son tan generales que podrían aplicarse a cualquier otro TFG. | No facilita ni una justificación de la programación de una secuencia. No estima ni gestiona los recursos necesarios para cumplir con plazos y objetivos. |
| Gestión de Riesgos | El plan de riesgos es detallado y ajustado al TFG. Los planes de contingencia y prevención están detallados para cada riesgo. El plan de seguimiento está detallado y justificado. | El plan de riesgos es concreto para el TFG. Los planes de contingencia y prevención para cada riesgo son adecuados. Existe plan de seguimiento, pero no está justificado. | La definición y cuantificación de los riesgos es apropiada para el TFG. Los planes de contingencia y prevención son ambiguos en algunos casos. Falta plan de seguimiento. | El plan de riesgos no es adecuado para el TFG, es demasiado general. | El plan de riesgos es insuficiente. Incluye planes de contingencia o prevención inadecuados. |
| Redacción y presentación | La redacción es correcta en todos sus aspectos, el estilo de escritura es apropiado y la lectura del documento es amena. | Redacta de forma clara, construyendo frases sintácticamente correctas, sin cometer faltas de ortografía y utilizando adecuadamente los signos de puntuación. | En general redacta de forma correcta aunque, en ocasiones, comete alguna falta de ortografía leve, utiliza incorrectamente algunos de los signos de puntuación y alguna frase sintácticamente incorrecta. | Redacta de forma poco clara, algunas de las frases son sintácticamente incorrectas, comete algunas faltas de ortografía y a menudo no utiliza adecuadamente los signos de puntuación. | Redacta de forma confusa y presenta una redacción que contiene gran cantidad de errores ortográficos, sintácticos, fonológicos y de puntuación. |

| | Excelente | Muy adecuado | Adecuado | Básico | Inadecuado |
|---------------------------------|--|--|--|---|---|
| Redacción y presentación | La redacción es correcta en todos sus aspectos, el estilo de escritura es apropiado y amena y el formato del documento es adecuado. | Redacta de forma clara, construyendo frases sintácticamente correctas, sin cometer faltas de ortografía y utilizando adecuadamente los signos de puntuación. | En general redacta de forma correcta aunque, en ocasiones, comete alguna falta de ortografía leve, utiliza incorrectamente signos de puntuación o escribe alguna frase sintácticamente incorrecta. | Redacta de forma poco clara, algunas de las frases son sintácticamente incorrectas, comete algunas faltas de ortografía y a menudo no utiliza adecuadamente los signos de puntuación. | Redacta de forma confusa y presenta una redacción que contiene gran cantidad de errores ortográficos y de puntuación. |
| Captura de requisitos | Tiene en cuenta todos los requisitos necesarios para alcanzar los objetivos del documento y los justifica. Justifica todas las decisiones. | Tiene en cuenta todos los requisitos necesarios para alcanzar los objetivos del TFG y los documenta adecuadamente. | Tiene en cuenta la gran mayoría de los requisitos necesarios para alcanzar los objetivos del TFG, pero su documentación es mejorable. | No profundiza lo bastante en los requisitos. Se queda en un nivel demasiado general. Falta algún requisito secundario, pero están los indispensables. | No se corresponden con los objetivos del TFG, falta gran parte de los requisitos. |
| Análisis y diseño | La documentación es detallada y justifica tanto el análisis como el diseño realizados. Justifica todos los requisitos a cumplir. | Justifica todas las decisiones tomadas y éstas son las más adecuadas al TFG. Falta algunos detalles. | Justifica todas o casi todas las decisiones tomadas, aunque no son las más adecuadas. La documentación es lo suficientemente completa. | La documentación cubre únicamente los aspectos básicos. Justifica las principales decisiones tomadas. | La documentación es incompleta, no se realiza un análisis y diseño suficiente para cubrir todos los objetivos del TFG. No se justifican las principales decisiones tomadas. |

| | Excelente | Muy adecuado | Adecuado | Básico | Inadecuado |
|---|---|--|--|--|--|
| Autonomía del estudiante | El alumno ha trabajado por su cuenta sin tener ningún tipo de problema y usando las reuniones sólo para informar de los avances logrados. | Ante un problema intenta buscar soluciones por su cuenta. Suele acudir a reuniones con distintas alternativas a sus problemas para recabar la opinión de la persona supervisora. El alumno tiene criterios propios para justificar sus decisiones. | Ante un problema intenta buscar soluciones por su cuenta. Suele acudir a reuniones con distintas alternativas a sus problemas para recabar la opinión de la persona supervisora. El alumno tiene criterios propios para justificar sus decisiones. | Ante un problema intenta buscar soluciones por su cuenta. Suele acudir a reuniones con distintas alternativas a sus problemas para recabar la opinión de la persona supervisora. El alumno tiene criterios propios para justificar sus decisiones. | Ante un problema acude donde la persona supervisora a que se lo solución sin haber buscado posibles soluciones. Acude a reuniones ante cualquier problema sin haber buscado posibles soluciones. Acude a reuniones ante cualquier problema sin haber buscado posibles soluciones. Acude a reuniones ante cualquier problema sin haber buscado posibles soluciones. |
| Receptividad ante críticas y cambios | El alumno acepta de buen grado las críticas que recibe, no vuelve a cometer el mismo error y cumple con las indicaciones de una manera excelente. | El alumno acepta de buen grado las críticas que recibe, es capaz de rebatir aquellas con las que no está de acuerdo y realiza las modificaciones que se le indican. | El alumno acepta correctamente todas las críticas que recibe y realiza las modificaciones que se le indican. | El alumno no asume las críticas pero realiza las modificaciones indicadas por la persona supervisora. | El alumno se cierra a recibir críticas sobre su trabajo y no realiza las modificaciones indicadas por la persona supervisora. |
| Interés por mejorar | El alumno realiza autocríticas sobre el trabajo realizado y es capaz de realizar mejoras de motu proprio. | El alumno incorpora las mejoras indicadas por la persona supervisora y busca posibles maneras de mejorar su trabajo. | El alumno incorpora las mejoras indicadas por la persona supervisora sin mayores problemas, pero no aporta mejoras propias. | El alumno se resiste a incorporar mejoras en su trabajo y cuando las acaba incorporando lo hace con dificultad. | No aporta ninguna mejora de motu proprio. Se limita a hacer lo que la persona supervisora le indica, incluso en ocasiones intentando rebajar las exigencias. |
| Constancia | El alumno ha informado y de manera periódica del trabajo que estaba realizando, y los avances conseguidos. | El alumno ha sido constante durante la elaboración del TFG. Pero sólo ha informado de manera periódica cuando quería finalizarlo. | El alumno ha trabajado por su cuenta, de manera constante durante la elaboración del TFG, pero no ha informado de que estaba haciendo en cada momento. | Ha desaparecido durante largos periodos de tiempo sin dar ningún tipo de explicación, lo que ha afectado al TFG, pero el único perjudicado ha sido el propio alumno. | Ha desaparecido durante largos periodos de tiempo sin dar ningún tipo de explicación, lo que ha afectado al TFG y a terceras personas. |

Eva_Memoria

| | Excelente | Muy adecuado | Adecuado | Básico | Inadecuado |
|--------------------------------------|---|---|--|---|---|
| Redacción y Presentación | La redacción es correcta en todos aspectos. El estilo es atractivo y adecuado. La lectura del documento es amena. | Redacta de forma clara, construyendo frases sintácticamente correctas, sin incorporear errores ortográficos y utilizando adecuadamente los signos de puntuación. | En general, redacta de forma correcta aunque, en ocasiones, comete alguna falta de ortografía o incorpore algunos errores de puntuación o escribe alguna frase sintácticamente incorrecta. | Redacta de forma poco clara, algunas de las frases son sintácticamente incorrectas, comete algunas faltas de ortografía y a menudo no utiliza adecuadamente los signos de puntuación. | Redacta de forma confusa y presenta una redacción que contiene gran cantidad de errores sintácticos, ortográficos y de puntuación. |
| Calidad global y coherencia | Presenta claramente y documenta de forma detallada y justificada la línea de trabajo que se va a seguir a lo largo del TFG. La línea de trabajo se ajusta completamente al TFG. | Presenta claramente y documenta la línea de trabajo que se va a seguir a lo largo del TFG. | Presenta adecuadamente los objetivos del TFG, pero la línea de trabajo no está claramente argumentada o no se ajusta a ellos. | Describe una línea de trabajo suficiente para cumplir con los objetivos principales del TFG. | No existe coherencia entre los objetivos planteados y la línea de trabajo a seguir. |
| Desarrollo | La documentación de cada parte es exhaustiva y justifica detalladamente cualquier cambio introducido. | El TFG sigue una línea de trabajo de manera competente según los objetivos y procesos descritos en el TFG. Justifica convenientemente cualquier cambio introducido. | Se cita parcialmente a los propósitos y la documentación de cada parte es suficientemente detallada y justifica los principales cambios introducidos. | Se cita medianamente a los objetivos y procesos necesarios para el desarrollo de su trabajo sin justificar los cambios producidos. | Ignora parte de los objetivos propuestos inicialmente sin justificar las razones. Falta parte de la documentación o las distintas partes no están cohesionadas. No justifica las decisiones tomadas a lo largo del TFG. |
| Evaluación | El proceso de evaluación del TFG realizado. Exponen los resultados obtenidos y su interpretación. | El proceso de evaluación realizado. Exponen los resultados obtenidos y los interpreta, aunque de forma superficial. | Expone de forma clara los resultados obtenidos y los realiza interpretaciones de los resultados. | El proceso de evaluación seguido es adecuado al TFG, aunque la muestra completa. Expone los resultados obtenidos de manera escueta. No incluye una discusión sobre los resultados obtenidos. | El proceso de evaluación no es adecuado al TFG. No presenta resultados o bien son incorrectos total o parcialmente. Expone los resultados de forma poco eficaz. |
| Conclusiones y trabajo futuro | Resume de manera clara y detallada el trabajo realizado y los principales hitos alcanzados. Realiza una comparativa entre los objetivos alcanzados. Realiza una valoración crítica sobre la valoración inicial y los realmente alcanzados. Realiza una valoración crítica sobre el TFG. Presenta de razonada forma posibles extensiones y mejoras relevantes del trabajo. | Resume de manera clara el trabajo realizado y los resultados alcanzados. Realiza una valoración superficial sobre la gestión llevada a cabo en el TFG. Presenta elementos no tratados y posibles extensiones o mejoras del trabajo realizado, pero no razona sobre ellas o no son relevantes. | El trabajo realizado y los resultados obtenidos no están explicados pero no de manera clara o concisa. Realiza una valoración superficial de la gestión realizada. Las líneas de posible mejora no son justificadas. | El trabajo realizado y los resultados obtenidos no están explicados pero no de manera clara o concisa. No presenta valoración sobre la gestión realizada. Las líneas de posibles mejoras no están bien razonadas o no son relevantes. | No realiza un análisis crítico sobre el trabajo realizado y los hitos logrados. No presenta conclusiones en los aspectos relativos a la gestión. No plantea posibles extensiones al trabajo realizado. |
| Bibliografía | Proporciona en la memoria las referencias bibliográficas completas, de las fuentes citadas y sólo de esas. Las fuentes citadas son relevantes y contrastables. Presenta las referencias de manera que permiten identificar unívocamente la fuente. Incluye en las tablas del documento, y | Proporciona en la memoria las referencias bibliográficas completas, pero estas no son las más relevantes. Presenta las referencias de manera que permiten identificar unívocamente la fuente. | Referencia de manera correcta todas las citas que realiza y las más relevantes. Presenta las referencias de manera que permiten identificar unívocamente la fuente. | Tiene alguna afirmación o cita que no ha referenciado. Tiene referencias que no cita en el trabajo. Presenta las referencias de manera que permiten identificar unívocamente la fuente. | No incluye referencias bibliográficas en las citas. Presenta referencias de Wikipedia u otras páginas donde la información no es contrastable. |

ANEXOD

Eva_Presentación

| | Excelente | Muy adecuado | Adecuado | Básico | Inadecuado |
|----------------------------------|---|--|---|---|--|
| Contenido y organización | La presentación está muy bien organizada. El tiempo dedicado a cada apartado ha sido el necesario. Esta división es clara y atractiva. El texto es muy atractivo. | La presentación está estructurada y sigue un guión lógico y fácil de seguir por la audiencia. Las partes están claramente identificadas y existe una relación coherente entre ellas. La presentación es clara y dedicada a cada uno de los apartados. La presentación no contiene errores gramaticales ni de ortografía y es fácil de leer con un diseño homogéneo, un tamaño de letra adecuado y una combinación de colores adecuada. | La presentación está estructurada y tiene un guión lógico. Existe una relación coherente entre las distintas partes de la exposición aunque no están claramente diferenciadas. Los apartados son los correctos aunque el tiempo que ha dedicado a cada uno de ellos no es el adecuado. La presentación es correcta aunque hay algunos apartados donde la lectura se hace difícil debido a los recursos empleados. | La presentación sigue un guión estructurado aunque no es el más adecuado. Existe relación entre las distintas partes, pero no es muy claro. La presentación tiene los apartados mínimos que debería tener. La presentación tiene una o dos faltas de ortografía. La lectura de la presentación es difícil debido al uso de párrafos muy largos, o fondos oscuros. | La presentación no sigue un guión estructurado. No existe coherencia entre las distintas partes de la exposición. Falta un apartado necesario. La presentación tiene varios errores gramaticales o faltas de ortografía. El texto es extremadamente difícil de leer con largos bloques de texto y tamaños de letra muy pequeños o colores. |
| Habilidades Comunicativas | Se nota que la presentación se ha preparado y ensayado a conciencia. Habla claramente y con un volumen de voz lo suficientemente alto. Su actitud durante la presentación es excelente y muestra un interés y una concentración en lo que cuenta. | Se nota que la presentación se ha preparado y ensayado. Habla claramente y con un volumen de voz lo suficientemente alto. Su actitud durante la presentación es adecuada. | La presentación ha sido ensayada y se nota que el/la ponente sabe qué decir en cada momento. La actitud durante la presentación no es la más adecuada en una exposición oral. | Se nota que la presentación se ha ensayado poco. Hay veces que se queda sin saber qué decir. A veces no se le entiende. | Se nota que la presentación no se ha ensayado previamente. Titubea constantemente sin saber qué decir en cada apartado. A menudo habla de manera que no se le puede entender. |
| Preguntas | La presentación ha sido tan clara que no ha sido necesario plantear preguntas. Responde con precisión y detalles a todas las preguntas que se le plantean. | Responde adecuadamente a todas las preguntas que se le plantean. | Responde adecuadamente a casi todas las preguntas que se le plantean. | Responde adecuadamente sólo a parte de las preguntas que se le plantean. El resto las responde contestar a lo que se le ha preguntado. | No es capaz de responder a la mayoría de las preguntas que se le plantean. |
| Uso del tiempo | El/la ponente se ajusta perfectamente al tiempo asignado para la presentación. Discute de manera detallada todos los aspectos de su trabajo. | El/la ponente se ajusta al tiempo asignado para la presentación, pero ha obviado algunos aspectos para conseguirlo. | Usa el tiempo adecuadamente, pero algunos apartados son prisa. | No se ajusta al tiempo indicado para terminar la presentación. Logra terminar la presentación. | No se ajusta al tiempo indicado por mucho tiempo. Termina muy pronto o se alarga en exceso sobre el tiempo. |

Eva_Complejidad

| Complejidad del TFG | Complejo | Adecuado | Básico | Inadecuado |
|--|---|--|---|------------|
| <p>Aporta utilidades de gran interés y novedosas. Hace uso de tecnologías/herramientas/planteamientos innovadores. La complejidad del TFG es notablemente mayor a los contenidos de las otras asignaturas del grado.</p> | <p>Las utilidades que aporta no son novedosas en el ámbito pero sí son de interés. Las tecnologías/herramientas/planteamientos empleados son bastante novedosos. El trabajo realizado no se limita a aplicar los conocimientos adquiridos en las otras asignaturas del grado.</p> | <p>No aporta ninguna utilidad no existente hasta el momento. Las tecnologías/herramientas/planteamientos empleados son actuales, pero no innovadoras. El trabajo reúne numerosos conocimientos adquiridos en el grado y los aplica de manera correcta.</p> | <p>No cumple con lo recogido en la propuesta de TFG aprobada. El TFG no aporta ninguna utilidad no existente hasta el momento. Las tecnologías/herramientas/planteamientos usados están ampliamente documentados o están obsoletos. La complejidad del trabajo no es mayor a los contenidos que se encuentran en las otras asignaturas del grado.</p> | |

Rúbricas atomísticas*Eva_Requisitos*

| Objetivos | | | |
|--|------------------------------|---------------------------|------------------------|
| Exposición de los objetivos del TFG | confusa | normal | muy clara |
| Posibilidad de analizar la consecución de los objetivos propuestos | inexistente | baja | alta |
| Diferenciación entre objetivos imprescindibles y secundarios | inexistente | confusa | clara o innecesaria |
| Antecedentes | | | |
| Explicación del punto de partida del trabajo a realizar | confusa | normal | muy clara |
| Mención a las fuentes bibliográficas o soluciones similares más relevantes para la contextualización del TFG | inexistente | normal | exhaustiva |
| Planificación | | | |
| Análisis de las tareas necesarias para llevar a cabo el TFG | inexistente | normal | exhaustivo |
| Tareas definidas para el TFG | genéricas | concretas poco detalladas | concretas y detalladas |
| Análisis de los recursos necesarios para llevar a cabo el TFG | inexistente | superficial | detallado |
| Planificación temporal de las tareas y recursos necesarios | superficial | normal | detallada |
| Gestión de riesgos | | | |
| Plan de riesgos | inexistente, malo o genérico | bueno | excelente |
| Cantidad de riesgos que incorporan una definición, un plan de prevención y un plan de contingencia | ninguno | algunos | todos |
| Planes de prevención y contingencia para cada riesgo | inexistentes o inadecuados | genéricos | particulares al TFG |
| Priorización de los riesgos | inexistente | sin justificar | justificada |
| Plan de seguimiento de los riesgos | inexistente | superficial | detallado |
| Redacción y presentación | | | |
| Redacción | mala | bueno | excelente |
| Uso de los signos de puntuación | malo | bueno | excelente |
| Nivel ortográfico del documento | malo | bueno | excelente |
| Nivel sintáctico y gramático del documento | malo | bueno | excelente |

Eva_Diseño

| Redacción y presentación | | | |
|---|-------------|----------------|---------------|
| Redacción | mala | buena | excelente |
| Uso de los signos de puntuación | malo | bueno | excelente |
| Nivel ortográfico del documento | malo | bueno | excelente |
| Nivel sintáctico y gramático del documento | malo | bueno | excelente |
| Captura de requisitos | | | |
| Relación entre los requisitos documentados y los objetivos definidos | nula | escasa | total |
| Requisitos imprescindibles para llevar a cabo el TFG que están reflejados en la documentación | ninguno | pocos | la mayoría |
| Documentación de los requisitos | escasa | buena | excelente |
| Análisis y diseño | | | |
| Explicaciones y justificaciones dadas | escasas | buenas | excelentes |
| Decisiones tomadas | incorrectas | poco adecuadas | muy adecuadas |
| Adecuación del análisis y diseño presentados para la consecución de los objetivos propuestos | escasa | buena | excelente |
| Documentación | escasa | buena | excelente |

ANEXO D

Eva_Actitud

| Autonomía | | | |
|--|--------------------------------|--|--|
| Persona que aportaba las soluciones a los problemas | la persona supervisora | entre la persona supervisora y el estudiante | el estudiante o no ha habido |
| Trabajo previo a las reuniones realizado por el estudiante | nulo | escaso | suficiente |
| Temas tratados en las reuniones | irrelevantes | poco relevantes | muy relevantes |
| Grado de aprovechamiento de las reuniones | bajo | escaso | alto |
| Receptividad ante críticas y cambios | | | |
| Aceptación de las críticas a su trabajo por parte del estudiante | sin aceptar | protestando | con buena actitud |
| Sugerencias de cambio indicadas por la persona supervisora | ignoradas | incorporadas parcialmente | No ha habido o están incorporadas totalmente |
| Capacidad del estudiante para rebatir razonadamente las críticas recibidas | nula | normal | alta |
| Interés por mejorar | | | |
| Persona que sugería las mejoras al TFG | la persona supervisora | entre la persona supervisora y el estudiante | el estudiante, el cliente o no es aplicable |
| Exigencias de la persona supervisora sobre el trabajo | protestadas | cumplidas al mínimo | aceptadas de buen grado |
| Actitud del estudiante a la hora de incorporar mejoras a su trabajo | mala | normal | buena |
| Mejoras sobre los enfoques y soluciones aportados por el estudiante | ninguno | sin interés | relevantes |
| Constancia | | | |
| Constancia del estudiante a lo largo de la elaboración del TFG | nada constante | constante a temporadas | constante |
| Personas afectadas negativamente por el ritmo de trabajo del estudiante | terceras personas | sólo el propio estudiante | Ninguna, el ritmo ha sido constante |
| Momento en el que se ha recibido información del estado del TFG | una vez estaba casi finalizado | de manera ocasional | de manera periódica |

Eva_Memoria

| Redacción y presentación | | | |
|---|---------------------------|-----------------|----------------|
| Redacción de la memoria | mala | buena | excelente |
| Uso de los signos de puntuación | malo | bueno | excelente |
| Nivel ortográfico de la memoria | malo | bueno | excelente |
| Uso de lenguaje formal en la memoria | malo | bueno | excelente |
| Coherencia de formatos y estilos a lo largo de la memoria | mala | bueno | excelente |
| Nivel sintáctico y gramatical de la memoria | malo | bueno | excelente |
| Calidad global y coherencia | | | |
| Descripción de los objetivos | confusa | normal | muy clara |
| Descripción de la línea de trabajo a seguir | inexistente | superficial | detallada |
| Relación entre la línea de trabajo propuesta y los objetivos planteados | inexistente o incoherente | algo confusa | coherente |
| Desarrollo | | | |
| Coherencia entre el trabajo desarrollado y los objetivos propuestos | escasa | normal | alta |
| Justificación de los cambios y decisiones tomadas a lo largo del TFG | escasa | bueno | excelente |
| Documentación sobre la realización del TFG | escasa | bueno | excelente |
| Evaluación | | | |
| Evaluación desarrollada en el TFG | poco adecuada | bueno | excelente |
| Proceso de evaluación llevado a cabo | escaso | bueno | excelente |
| Exposición de los resultados | confusa | poco detallada | detallada |
| Interpretación de los resultados obtenidos | escasa | bueno | excelente |
| Conclusiones y trabajo futuro | | | |
| Análisis del trabajo realizado | escaso | bueno | excelente |
| Explicaciones sobre el grado de cumplimiento de los objetivos | escasas | buenas | excelentes |
| Explicaciones sobre el trabajo de gestión realizado | escasas | buenas | excelentes |
| Relevancia de las líneas futuras propuestas | escasa | normal | muy alta |
| Análisis y justificación sobre las líneas futuras propuestas | escaso | bueno | excelente |
| Bibliografía | | | |
| Referencias bibliográficas incluidas | ninguna | alguna | bastantes |
| Número de referencias no contrastables (ej. Wikipedia) | más de dos | una o dos | ninguna |
| Relevancia de las referencias incluidas | no relevantes | poco relevantes | muy relevantes |
| Referencias incluidas que se citan a lo largo de la documentación | ninguna | algunas | todas |
| Tablas y figuras con pie descriptivo | ninguna | algunas | todas |
| Tablas y figuras referenciadas a lo largo de la documentación | ninguna | algunas | todas |

ANEXO D

Eva_Presentación

| Contenido y organización | | | |
|--|---------------|------------|--------------|
| Estructura de la presentación | mala | buna | excelente |
| Tiempo dedicado a cada apartado | inadecuado | aceptable | muy adecuado |
| Nivel sintáctico y gramático de la presentación | malo | bueno | excelente |
| Nivel ortográfico de la presentación | malo | bueno | excelente |
| Aspecto gráfico de la presentación | malo | bueno | excelente |
| Habilidades Comunicativas | | | |
| Preparación de la presentación | insuficiente | aceptable | buna |
| Momentos en los que el estudiante no ha sabido qué decir | muchos | pocos | ninguno |
| Tono de voz usado por el estudiante | ininteligible | bueno | excelente |
| Lenguaje y actitud del estudiante | malos | buenos | excelentes |
| Preguntas | | | |
| Respuestas del estudiante a las preguntas | malas | buenas | excelentes |
| Actitud del estudiante durante las preguntas | mala | buna | excelente |
| Uso del tiempo | | | |
| Uso del tiempo por parte del estudiante | breve | excesivo | ajustado |
| Aspectos imprescindibles cubiertos en el tiempo asignado | unos pocos | la mayoría | todos |

Eva_Complejidad

| Complejidad | | | |
|---|--------------------------|------------|-----------------------|
| Grado de cumplimiento de los objetivos mínimos aprobados en la propuesta | bajo | medio | alto |
| Grado de novedad de las utilidades que aporta | bajo | medio | alto |
| Interés de las utilidades que aporta | bajo | medio | alto |
| Herramientas/tecnologías/planteamientos usados | no adecuados u obsoletos | aceptables | adecuados o novedosos |
| Número de asignaturas involucradas en los contenidos trabajados en el TFG | una o dos | tres | más de tres |
| Autoformación en contenidos no vistos en el plan de estudios | nula | baja | alta |

