

# **TESIS DOCTORAL**

## **Prospección de la colaboración utilizando herramientas de minería de datos en ambiente abiertos de aprendizaje colaborativo con el objetivo de mejorar la gestión del proceso de colaboración**

Antonio Rodríguez Anaya

Licenciado en Ciencias Físicas

por la Universidad Complutense de Madrid

Dpto. de Inteligencia Artificial

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA

UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA

UNED,

2009



Departamento de Inteligencia Artificial,

Escuela Técnica Superior de Ingeniería Informática.

Prospección de la colaboración utilizando herramientas de minería de datos en ambiente abiertos de aprendizaje colaborativo con el objetivo de mejorar la gestión del proceso de colaboración

Antonio Rodríguez Anaya

Licenciado en Ciencias Físicas

por la Universidad Complutense de Madrid

Jesús González Boticario



# Agradecimientos

Este trabajo no hubiera sido posible sin la ayuda de un conjunto de personas que de alguna forma me han ayudado.

En primer lugar agradezco a Jesús González Boticario por su paciencia e incidencia en las debilidades de este trabajo para un constante perfeccionamiento.

Agradezco a mi familia, padres y hermana, por estar ahí y por sus cuidados y críticas.

Agradezco a Mila por todos sus consejos, comentarios y advertencias. Eso es una amiga y lo demás son tonterías.

Agradezco a los compañeros del departamento de Inteligencia artificial por las conversaciones amigables que me han ayudado a saber qué es una tesis.

Agradezco a aquel poeta que me encontré un día camino a Soria por el cual advertí que lo bueno o malo que hay en la vida empieza por uno mismo y su propio esfuerzo.

Dedico este trabajo a mi familia, padres y hermana, y con especial cariño a mi abuelo de casi 93 años.



"HOMINES, ALTER ALTERIUS CAUSA NATI SUNT.  
EOS IGITUR DOCE AUT FER."

"LOS HOMBRES HAN NACIDO LOS UNOS PARA LOS OTROS.  
INSTRÚYELOS O SOPÓRTALOS."

"MEN EXIST FOR THE SAKE OF ONE ANOTHER.  
TEACH THEM THEN OR BEAR WITH THEM."

[Marcus Aurelius, Meditationes, Libri VIII, 59]





## Resumen

La motivación de esta tesis ha sido la mejora del aprendizaje en un entorno de educación a distancia. Debido a sus condicionamientos se aconseja el uso de la colaboración como estrategia educativa y la nuevas técnicas de información y comunicación hacen posible su uso en entornos de e-learning. Sin embargo, que haya servicios para poder colaborar no significa que suceda el aprendizaje colaborativo. El modelo de la experiencia de aprendizaje colaborativo debe cumplir una serie de condiciones como la de analizar la colaboración que se produce de manera regular y frecuente.

Debido a la diversidad de los estudiantes en la educación a distancia y a las recomendaciones de centrarse en el estudiante, el control sobre los procesos que intervienen en el aprendizaje es responsabilidad de los propios estudiantes. Un ejemplo de estudiantes de la educación a distancia son los alumnos de la UNED (Universidad Nacional de Educación a Distancia), los cuales controlan su aprendizaje por lo que también deben controlar sus procesos de colaboración en un entorno colaborativo.

Los condicionamientos de la educación a distancia, las condiciones que un entorno colaborativo debe cumplir y el control de los estudiantes sobre los procesos de colaboración, aconsejan realizar un análisis de lo que sucede en el entorno colaborativo para obtener valoraciones sobre la colaboración de forma regular y frecuente para asegurarse y mejorar el aprendizaje colaborativo con la menos intervención directa de los participantes en la experiencia.

Éste ha sido el trabajo descrito en esta tesis. Se ha propuesto una experiencia de aprendizaje colaborativo durante los cursos 2006/2007, 2007/2008 y 2008/2009. Durante esos cursos se ha analizado mediante las interacciones en los foros del entorno la colaboración realizada y se han propuesto dos métodos de inferencia basados en tecnologías de aprendizaje automático. Durante el curso 2008/2009 se ha modelado los estudiantes según la colaboración utilizando uno de los dos métodos de inferencia propuesto. Con el modelo creado se han construido cuatro herramientas, las cuales mostraban distintas partes del modelo para entender cual de ellas era más útil y

de dos estrategias de visualización distintas, una de ella mostraba la información, y la otra permitía la gestión de la misma desde un punto de vista escrutable.

Las aportaciones de esta tesis al campo del modelado del estudiante en entornos colaborativos han sido evaluadas. Se ha propuesto un método de comparación de los dos métodos de inferencia utilizando los datos de los tres cursos. La información almacenada en el modelo de colaboración del estudiante y la forma de mostrar dicha información han sido evaluados utilizando distintas fuentes centradas en el estudiante.

Se concluye que un método de análisis de la colaboración de forma regular y frecuente con la menor intervención directa de docentes o estudiantes es posible y que dicha información ayuda a mejorar el control del proceso de colaboración por lo que se alcanza mayor aprendizaje colaborativo. La generalidad del modelo educativo y sus componentes hace posible que se pueda utilizar en otros entornos colaborativos.

Hay puntos abiertos que no cierra esta tesis como son la mejora de la exactitud de las inferencias de los métodos propuestos y su comparación con otros, así como la mejora del uso de las herramientas que utilizan la información inferida sobre la colaboración para la mejora del control de la colaboración, sobre todo aquellas que utilizan la estrategia de escrutabilidad.

## **Abstract**

The motivation of this thesis has been the improvement of learning in a distance education environment. The use collaboration as an educational strategy is advised because of the conditioning factor of the distance education, and the new information and communication technologies make possible its use in e-learning environments. However, the use of collaborative services does not mean that collaborative learning happens. The model of collaborative learning experience must achieve certain conditions, as analyzing the collaboration that occurs regularly and frequently, to assure collaborative learning happens.

Due to the diversity of students in distance education and recommendations collaborative learning experiences are focused on the student, the control over the processes involved in learning is the responsibility of students themselves. An example of this kind of students are students of the UNED (National University for Distance Education), which control their learning so also they must control their collaboration processes in a collaborative environment.

The constraints of distance education, the conditioning factor which must achieve a collaborative environment and students' control on collaborative processes, recommend an analysis on interactions in the collaborative environment to obtain evaluations on collaboration regularly and frequently to ensure and improve the collaborative learning with less direct intervention of participants in the experience.

This is the work described in this thesis. It has proposed a collaborative learning experience during the academic year 2006/2007, 2007/2008 and 2008/2009. During these years has been analyzed the interactions on environment in collaboration forums and proposed methods of inference based on machine learning technologies. During the academic year 2008/2009 has been modeled the collaboration of students using one of the proposed methods of inference. With the model four tools have been built, which showed different parts of the model to understand what part of them was more useful, and the tools are displayed in two different strategies, one, showed the information, and the other, allowed the management of the information in scrutable terms.

The contributions of this thesis to the field of student modeling in collaborative environments have been evaluated. We have proposed a method of comparing the two methods of inference using data from the three courses. The information stored in the collaborative model student and how to display that information have been evaluated using various student-centered sources.

We conclude that a method of analysis of collaboration regularly and frequently with less direct intervention of teachers or students is possible and that such information helps to improve the management of the collaborative process so that collaborative learning increases. The generality of the educational model and its components makes possible to transfer to other collaborative environments.

There are open issues, which this thesis do not close, such as improving the accuracy of the inferring methods and their comparison with others methods, and improving the use of tools that use the inferred information about collaboration for better monitoring collaboration, especially those that use the scrutability strategy.

# Índice de contenidos

<b>PARTE I: INTRODUCCIÓN A LA INVESTIGACIÓN</b>	<b>22</b>
<b>CAPÍTULO I: INTRODUCCIÓN</b>	<b>23</b>
Motivación	24
Planteamiento	25
Aproximación al problema	27
Objetivos	30
Preguntas que esta tesis se ha planteado	32
Organización de la memoria	32
<b>CAPÍTULO II: TRABAJOS RELACIONADOS</b>	<b>34</b>
<b>Introducción</b>	<b>34</b>
Caracterización del entorno colaborativo	35
Discusión	37
Análisis de la colaboración	38
Análisis de la colaboración mediante minería de datos	40
Investigaciones	42
Análisis crítico	46
Modelado de la colaboración	49
Modelo de la colaboración	50
Identificadores de la colaboración	53
Análisis crítico	58
Estrategias de mejora del aprendizaje colaborativo	60
La autorregulación como estrategia de aprendizaje	61
Realizar seguimiento de la colaboración	63
Investigaciones	64
Escrutabilidad	66
Investigaciones	67
Privacidad	68
Análisis crítico	68
Conclusiones generales	70
<b>PARTE II: DESARROLLO</b>	<b>75</b>

<b>CAPÍTULO III: MARCO CONCEPTUAL, MODELOS Y DESARROLLO DEL PLANTEAMIENTO</b>	<b>76</b>
<b>Contexto educativo</b>	<b>77</b>
Introducción a la educación a distancia	77
Modelo educativo de la UNED	80
La asignatura	81
Objetivos	81
Destrezas y competencias	82
Metodología docente	83
Estructura y contenidos teóricos	83
Evaluación	84
Los alumnos	84
<b>La experiencia de aprendizaje colaborativo</b>	<b>86</b>
Modelo educativo de la experiencia colaborativa	88
Objetivos docentes	88
Requisitos	89
Destrezas y competencias	89
Contextualización	90
Metodología	91
Experiencia pedagógica	91
Análisis de las interacciones	101
Inferencia de la colaboración	106
Modelo de colaboración	112
Herramientas metacognitivas	119
Evaluación	128
<b>PARTE III: EXPERIMENTACIÓN</b>	<b>131</b>
<b>CAPÍTULO IV: DESARROLLO Y RESULTADOS</b>	<b>132</b>
<b>Análisis de las interacciones</b>	<b>133</b>
<b>Resultados de los métodos de inferencia de la colaboración</b>	<b>137</b>
Resultados del análisis mediante métricas de la colaboración	138
Métricas propuestas	141
Análisis crítico	144
Resultados de la clasificación mediante clustering	146
<b>Datos de las herramientas de autorregulación</b>	<b>156</b>

Experimentación	158
Aplicación web	159
Encuesta	160
Notas	163
Portlet indicadores estadísticos	165
Encuesta	165
Notas	167
Portlet niveles de colaboración	169
Encuesta	169
Notas	171
Portlet indicadores estadísticos y niveles de colaboración	172
Encuesta	172
Notas	175
<b>Resultados del entorno colaborativo</b>	<b>176</b>
<b>CAPÍTULO V: VALORACIÓN</b>	<b>181</b>
<b>Validez de la prospección de la colaboración</b>	<b>181</b>
Análisis de la validez del acercamiento de la métrica	183
Análisis de la validez del acercamiento de clustering	186
Métricas vs clustering	188
Procedimientos de automatización de los acercamientos de minería de datos	189
<b>Validez de las herramientas metacognitivas de autorregulación</b>	<b>191</b>
Encuestas	191
Notas	193
<b>Validez de la experiencia de aprendizaje colaborativo</b>	<b>195</b>
<b>PARTE IV: CONCLUSIONES Y TRABAJO FUTURO</b>	<b>198</b>
<b>CAPÍTULO VI: CONCLUSIONES</b>	<b>199</b>
Logros	200
Valoración de contribuciones	203
Análisis de las limitaciones	204
Respuesta a las preguntas planteadas por esta tesis	206
Respuesta a la primera pregunta	206
Respuesta a la segunda pregunta	207
Respuesta a la tercera pregunta	207
Respuesta a la cuarta pregunta	208

Respuesta a la quinta pregunta	208
<b>CAPÍTULO VII: TRABAJOS FUTUROS</b>	<b>209</b>
Análisis de la colaboración	209
Fuente de datos	209
Método de inferencia	210
Mejora de la métrica	211
Qué hacer con la información obtenida en el análisis	212
Modelo de la colaboración	212
Información sobre la colaboración	212
Herramienta escrutable	213
Esquema de los trabajos futuros	214
<b>BIBLIOGRAFÍA</b>	<b>215</b>
<b>APÉNDICES</b>	<b>227</b>
<b>Apéndice I: encuestas</b>	<b>227</b>
Encuesta inicial curso 2006/2007	227
Encuesta inicial curso 2007/2008	230
Encuesta inicial curso 2008/2009	233
Encuesta final curso 2006/2007	236
Encuesta final curso 2007/2008	240
Encuesta final curso 2008/2009	243
Encuesta final sobre la aplicación web del curso 2008/2009	246
Encuesta final sobre la herramienta de indicadores estadísticos del curso 2008/2009	247
Encuesta final sobre la herramienta que mostraba los niveles de colaboración en el curso 2008/2009	248
Encuesta final sobre la herramienta que mostraba los indicadores estadísticos y los niveles de colaboración en el curso 2008/2009	249
<b>Apéndice II: ontología de la colaboración</b>	<b>250</b>
<b>Apéndice III: Summary and Conclusions</b>	<b>257</b>
Summary	257
Framework	259
Critical Analysis	261
Research	263
Results	264
Conclusions	265





# Índice de tablas

Tabla 1: Caracterización de investigaciones en relación al análisis de la colaboración .....	72
Tabla 2: Características de los alumnos .....	85
Tabla 3: resumen de las evaluaciones realizadas. ....	130
Tabla 4: criterio de filtrado de equipos con poca actividad.....	134
Tabla 5: datos estadísticos de los datasets.....	135
Tabla 6: datos sobre las lista del nivel de colaboración de los alumnos.....	136
Tabla 7: N° de árboles de decisión que utilizan cada uno de los indicadores estadísticos. ....	139
Tabla 8: indicadores estadísticos más frecuentes según tres tipos de suma.....	140
Tabla 9: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-I-06-07. ....	142
Tabla 10: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-II-06-07. ....	142
Tabla 11: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-I-07-08.....	143
Tabla 12: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-II-07-08. ....	143
Tabla 13: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-I-08-09.....	143
Tabla 14: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-II-08-09. ....	144
Tabla 15: relación entre la lista suministrada por el experto sobre la colaboración de los alumnos y los cluster obtenidos de los datasets sin filtrar de equipos poco activos.....	155

Tabla 16: relación entre la lista suministrada por el experto sobre la colaboración de los alumnos y los cluster obtenidos de los datasets filtrado de equipos poco activos. ....	155
Tabla 17: datos sobre la experimentación con las herramientas metacognitivas. ....	158
Tabla 18: resultados de la evaluación del trabajo colaborativo de los alumnos y del resultado del examen del primer parcial. ....	164
Tabla 19: valoración del portlet de indicadores estadísticos. ....	166
Tabla 20: evaluación de los alumnos con el portlet de indicadores estadísticos. ....	168
Tabla 21: valoración del portlet de los niveles de colaboración inferidos. ....	169
Tabla 22: evaluación de los alumnos con el portlet de niveles de colaboración inferidos. ....	171
Tabla 23: valoración de los portlets de los indicadores estadísticos y de los niveles de colaboración inferidos. ....	173
Tabla 24: evaluación de los alumnos con los portlet de indicadores estadísticos y niveles de colaboración inferidos. ....	176
Tabla 25: valoración de los alumnos a la experiencia. ....	177
Tabla 26: comparación de resultados de los alumnos de la experiencia colaborativa respecto a los que no hicieron dicha experiencia. ....	178
Tabla 27: resultados de los alumnos de cursos anteriores. ....	178
Tabla 28: pendientes y errores de las rectas que representan las métricas frente al espacio de niveles de colaboración. ....	184
Tabla 29: pendiente y % de varianza respecto a los distintos acercamientos. ....	188
Tabla 30: comparativa de resultados de las evaluaciones según la herramienta metacognitiva de autorregulación. ....	194

# Índice de figuras

Figura 1: ejemplo de modelo de restricciones (Baghaei y Mitrovic, 2007).....	51
Figura 2: modelo de una acción (Martínez et al., 2003) .....	52
Figura 3: ejemplo de actividad (Barros et al., 2002).....	53
Figura 4: esquema de tareas de la experiencia de aprendizaje colaborativo .....	92
Figura 5: Espacio general del primer periodo de la experiencia.....	93
Figura 6: Espacio virtual específico de un equipo.....	96
Figura 7: árbol lógico.....	110
Figura 8: Atributos de la ontología en las clases de los datos estáticos.....	115
Figura 9: clases de la ontología que informan sobre las interacciones en los foros. .	116
Figura 10: clase de la ontología que revela la colaboración inferida de los alumnos. .....	117
Figura 11: atributos y propiedades de la clase Usuario.....	118
Figura 12: pantalla de identificación del usuario de la aplicación web. ....	122
Figura 13: página de entrada en le modelo de colaboración. ....	122
Figura 14: pantalla que muestra los datos de una muestra de la clase Usuario.....	123
Figura 15: pantalla de edición de los datos de la muestra Profesor.....	124
Figura 16: portlet de las estadísticas de interacción en los foros.....	125
Figura 17: portlet que muestra el nivel de colaboración inferido de los alumnos.....	126
Figura 18: rectas obtenidas por análisis de mínimos cuadrados de los valores medios de las métricas para el dataset D-II-06-07. ....	145
Figura 19: datasets sin filtrar de equipos poco activos clasificados en 2 cluster.....	148
Figura 20: datasets filtrados de equipos poco activos clasificados en 2 cluster.....	149
Figura 21: datasets sin filtrar de equipos poco activos clasificados en 3 cluster.....	150
Figura 22: datasets filtrados de equipos poco activos clasificados en 3 cluster.....	151
Figura 23: datasets sin filtrar de equipos poco activos clasificados en 4 cluster.....	152

Figura 24: datasets filtrado de equipos poco activos clasificados en 4 cluster.....	153
Figura 25: relación entre los indicadores estadísticos Num_msg y Num_reply_msg en el caso de clasificar los datasets sin filtrar de equipos poco activos en 3 cluster. ....	154
Figura 26: relación entre los indicadores estadísticos Num_msg y Num_reply_msg en el caso de clasificar los datasets filtrados de equipos poco activos en 3 cluster. ....	154
Figura 27: comparativa de resultados en los últimos seis cursos. ....	179
Figura 28: esquema del trabajo futuro.....	214

# PART E I: INTRODUCCIÓN A LA INVESTIGACIÓN

En esta primera parte de la tesis se introduce la tesis con dos capítulos. En el capítulo I se introduce la motivación hipótesis y objetivos de la investigación que ha dado como resultado esta tesis. En el capítulo II se enmarcar la investigación haciendo un análisis de los trabajos relacionados discutiendo los aspectos más relevantes para esta.

# Capítulo I: introducción

La investigación que da vida a esta tesis parafrasea la sentencia de Marco Aurelio utilizada como *lema* en esta memoria. Marco Aurelio, emperador del imperio romano a la vez que filósofo estoico, tenía claro que el ser humano es un animal social y necesita de otros. Él, como emperador y filósofo, sintió la obligación de instruir a sus semejantes, en el caso de que fuera posible, o dejarlos en paz, en el caso contrario.

El autor de la tesis ni es emperador ni filósofo, es físico y está obligado, contractualmente, a instruir a los alumnos que estudian Ingeniería informática en la UNED (Universidad Nacional de Educación a Distancia). Sin embargo, el lema de Marco Aurelio resume el trabajo de esta investigación, con algunos matices.

Admitimos y afirmamos que los seres humanos son animales sociales. Las relaciones sociales nos rodean y dan forma a procesos cognitivos que nos diferencian de otros seres vivos. La comunicación, el habla, es un rasgo principal en el ser humano e imprescindible históricamente en la evolución de la cultura humana, del pensamiento y de la sociedad.

Donde hay algún cambio es en el consejo de instruir o soportar a los demás. Como docente, la instrucción es fundamental y hay que soportar, con el significado de dar soporte, dicha instrucción con todos los medios y estrategias posibles. El matiz es cambiar el significado de soportar, dejar en paz, por el de sustentar, mantener, dar un soporte. Es verdad que estamos haciendo un poco de trampa al traducir el término *soportar* como más conviene a la investigación, pero eso es lo bueno de la ambigüedad del lenguaje natural.

La tesis une los conceptos del *lema* en una única investigación: *instruir* a los estudiantes dando *soporte* al trabajo en equipo que requiere de *interacciones sociales*. El aprendizaje colaborativo es una estrategia educativa muy apropiada y ventajosa en algunos contextos, siendo la educación a distancia (ED) o los entornos de e-learning uno de ellos (García Aretio, 2009). La colaboración se basa en las habilidades sociales de los seres humanos, los cuales las utilizan con gusto todos los días. A su vez, las

nuevas Tecnologías de Información y Comunicación (TIC) hacen la colaboración posible en entornos a distancia.

Esta tesis no habla de soportar (sustentar, dar soporte a) la colaboración para alcanzar la instrucción. Esta tesis habla de dar un soporte al aprendizaje colaborativo, es decir, que se instruya a los alumnos mientras estos utilizan habilidades sociales por medio de herramientas o servicios ofrecidos en un contexto de ED o e-learning. La diferencia no es tan sutil. Una cosa es dejar que los estudiantes hablen o se comuniquen, y otra es que aprendan realizando tareas que requieran del uso de habilidades sociales.

Por lo tanto, el *lema* ha sido el guía que ha motivado la tesis y el resumen de la misma. Profundizando se establece un planteamiento que lleva a identificar los problemas a los que hay que enfrentarse para alcanzar los objetivos de la tesis, que son los que indica la frase de Marco Aurelio con la nueva interpretación.

## **Motivación**

La ED es un modelo de enseñanza cada vez más frecuente que tiene una serie de características relevantes, como la ruptura de las barreras de lugar y tiempo. El docente puede instruir al estudiante sin tener que estar ambos en el mismo sitio y al mismo tiempo. Esta ventaja trae el inconveniente que la comunicación entre ambos tiene que utilizar otros medios menos ricos e intensos que el cara a cara. Lo mismo sucede con la comunicación entre los propios estudiantes. Es decir, la gran ventaja de la ED es paradójicamente también su principal condicionamiento (García Aretio, 2009). Otra ventaja es que los estudiantes en cualquier lugar y tiempo (ruptura de las barreras de lugar y tiempo) puedan participar en un curso, esto ocasiona que los estudiantes no formen grupos homogéneos y puedan ser caracterizados por varios criterios (edad, lugar de residencia, tipo de trabajo, historial académico, etc.) (Boticario y Gaudio, 2000b).

La ED es un modelo adecuado para introducir el aprendizaje colaborativo o cooperativo. Estos términos se utilizarán como sinónimos en esta tesis, aunque hay autores que plantean distintas definiciones. Barkley et al. (2004) dice: el aprendizaje colaborativo es un medio de trabajo de los estudiantes en parejas o grupos pequeños para alcanzar metas comunes. Johnson y Johnson (2004) comentan: el aprendizaje



cooperativo o colaborativo es una relación de un grupo de estudiantes que requiere una interdependencia positiva, la rendición de cuentas individuales, las habilidades interpersonales, la promoción la interacción y de transformación. Estos autores comentan las ventajas de utilizar el aprendizaje colaborativo en entornos web (web, Internet o red son términos que se utilizarán indistintamente en esta tesis). Las TIC ofrecen nuevos servicios de comunicación y gestión de esta comunicación que hacen factible la colaboración a través de la web. Sin embargo, el uso de servicios y herramientas que posibilitan la comunicación no es suficiente para que el aprendizaje colaborativo se produzca o éste sea superior al aprendizaje competitivo o individual (Jonhson y Jonhson, 1994).

La motivación de esta tesis ha sido la mejora del aprendizaje en un entorno de ED a través de Internet mediante la estrategia del aprendizaje colaborativo. Aunque parezcan evidentes las ventajas del uso de la colaboración en la ED, como ya se ha señalado, no es suficiente con ofrecer servicios de comunicación por lo que un diseño apropiado y un estudio de la colaboración es necesario para intentar asegurarse que el aprendizaje colaborativo se lleva a cabo.

## **Planteamiento**

El uso de las TIC en entornos educativos vía web se incrementa cada vez más. Internet ha hecho posible, por medio de sistemas de gestión del aprendizaje (LMS en sus siglas en inglés), reunir estudiantes en espacios de trabajo compartido que facilitan la comunicación y la colaboración. La colaboración, algo que era difícil aplicar en los antiguos contextos de la ED, es fácilmente aplicable en los actuales entornos de educación vía web. Existen plataformas web especializadas en éste área tales como Blackboard<sup>1</sup>, dotLRN<sup>2</sup>, Moodle<sup>3</sup>, Sakai<sup>4</sup>, las cuales ofrecen servicios apropiados para organizar cursos colaborativos. En concreto, esta investigación toma como base dotLRN, y se enmarca en los trabajos que hace el grupo de investigación

---

<sup>1</sup> [www.blackboard.com/](http://www.blackboard.com/)

<sup>2</sup> <http://dotlrn.org/>

<sup>3</sup> <http://docs.moodle.org/>

<sup>4</sup> <http://sakaiproject.org/>

aDeNu<sup>5</sup> (sistemas educativos online de adaptación dinámica basado en el modelado del usuario) (Boticario et al., 2007).

Sin embargo, la oferta de herramientas que favorecen la comunicación entre estudiantes no garantiza que la colaboración se realice. Según Johnson y Johnson (1994), el aprendizaje cooperativo es aquel que se consigue mientras se establecen comunicaciones con otros compañeros de aprendizaje. Con esta definición no es necesario colaborar para aprender cooperativamente. Los mismos autores establecen ciertas condiciones para que el esfuerzo cooperativo sea más productivo que el esfuerzo competitivo e individual, las cuales también se deben cumplir en entornos basados en TIC (Johnson y Johnson, 2004). Estas condiciones son:

- Percibirse claramente la interdependencia positiva, es decir, que uno pueda tener éxito sólo cuando todos lo tienen.
- Promocionar la interacción de unos con otros por distintos medios. Los autores aconsejan que los estudiantes utilicen varios medios para tener una comunicación más completa y flexible, e incluso la comunicación cara a cara por las ventajas que aporta, aunque no sea la más probable en un entorno e-learning (Johnson y Johnson, 1989).
- Percibirse claramente la rendición de cuentas individual y la responsabilidad personal para alcanzar los objetivos del grupo.
- Uso frecuente de las relaciones interpersonales y habilidades, las cuales son fomentadas al trabajar en pequeños grupos.
- Frecuentes y regulares procesamientos del funcionamiento del grupo para mejorar su eficacia en el futuro.

Estas condiciones requieren un trabajo de diseño del entorno, de las herramientas de comunicación, del propio curso colaborativo, e incluso un trabajo de análisis o modelado de los participantes. Cumpliendo las condiciones se garantiza que el aprendizaje colaborativo se lleva a cabo. Hacemos notar que la quinta condición se tiene que cumplir durante la realización de la experiencia de aprendizaje. El procesamiento puede ser realizado por tutores y estudiantes. Pero este procesamiento puede no ser posible en entornos de ED con servicios centrados en los estudiantes. Estos entornos se caracterizan por la gran cantidad de estudiantes y la poca

---

<sup>5</sup> <https://adenu.ia.uned.es/web/>

homogeneidad de los mismos (Boticario y Gaudioso, 2000b). Los tutores de este tipo de entornos pueden tener grandes dificultades si no se les ofrece algún tipo de ayuda, al igual que a los estudiantes. Las dificultades son evidentes al tener en cuenta la gran cantidad de interacciones posibles que pueden realizar una gran cantidad de estudiantes heterogéneos. El trabajo necesario de profesores, tutores o expertos se incrementa, por lo que es apropiado utilizar las ventajas del uso de ordenadores (gran capacidad de almacenamiento y de cálculos sencillos) (Romero y Ventura, 2007). Técnicas de modelado del usuario (User Modelling en inglés) pueden ser de gran ayuda (Brusilovsky y Millan, 2007).

Durante los últimos años se ha avanzado en el modelado de sistemas educativos, considerando igualmente la colaboración, pero se han investigado en gran medida sólo características de los alumnos enmarcadas en escenarios de aprendizaje caracterizados por el trabajo individual (Brusilovsky y Millan, 2007). Sin embargo, se detecta una carencia significativa de falta de análisis sobre cómo los estudiantes colaboran (Strijbos y Fisher, 2007).

Esta tesis se plantea, dentro del marco de trabajo del grupo de investigación aDeNu, el diseño de un entorno web colaborativo que, cumpliendo las anteriores condiciones, favorezca o ayude a los estudiantes en la mejora de la gestión del proceso de colaboración, con el objetivo principal de mejorar el aprendizaje. Por lo cual, el análisis de la colaboración de los estudiantes y las acciones a consecuencia del análisis son parte del planteamiento.

## **Aproximación al problema**

Tanto el diseño de un entorno colaborativo como el análisis de la colaboración ha caído históricamente en el trabajo de expertos. Los profesores o tutores realizaban un seguimiento de los alumnos con lo que se era capaz de cumplir con las condiciones de (Johnson y Johnson, 2004). ¿Sucede lo mismo en los entornos educativos vía web? Las TIC han hecho más fácil unir la colaboración con la ED vía web. En la actualidad, los entornos de ED vía web hacen posible que una gran variedad y número de estudiantes puedan utilizar gran cantidad de recursos para su educación. Esto provoca que los estudiantes de la ED no formen un grupo homogéneo y puedan diferenciarse por su edad, experiencia laboral, currículo académico, lugar de

residencia, preferencias, responsabilidades, etc. (Boticario y Gaudio, 2000b). La estrategia de aprendizaje colaborativo puede resolver los problemas de comunicación inherente a la ED, pero puede ser de difícil aplicación teniendo en cuenta la diversidad de los estudiantes. Así sucede con los alumnos de la UNED.

Los estudiantes de la UNED son principalmente adultos con otras responsabilidades además del aprendizaje. El modelo educativo utilizado ofrece una instrucción adaptada a este tipo de estudiantes por lo que permite el trabajo de un gran número de alumnos con distintas características (García Aretio, 2009). En este modelo son los estudiantes los que necesitan tener el control sobre sus procesos para realizar el aprendizaje (Gaudio et al., 2003).

Como la motivación de la tesis también incluye generalidad para que este estudio se pueda aplicar a otros entornos, se ha tenido en cuenta la diversidad de los tipos de estudiantes y la recomendación del control sobre el proceso de colaboración. Si tenemos un entorno de aprendizaje colaborativo vía web donde puede haber un gran número de estudiantes, los cuales tienen el control sobre el proceso de colaboración, ¿el tutor o profesor podría realizar análisis de la colaboración de forma regular y frecuente? No se puede dar una respuesta con absoluta certeza, pero sí se puede decir que esa tarea será de más difícil realización. Sin embargo, justo en entornos semejantes, y desde hace tiempo, el área del modelado del usuario se ha utilizado con el objetivo de modelar a los usuarios para distintas funciones (Brusilovsky y Millan, 2007). El modelado del usuario estudia los usuarios de un sistema y obtiene valoraciones o caracterizaciones sobre él. Esto se puede aplicar a entornos colaborativos (Bratitsis et al., 2008), e incluso en relación con la colaboración (Redondo et al., 2003).

Ha habido ciertas investigaciones que se han centrado en el análisis de la colaboración. Éstas se pueden caracterizar, según Romero y Ventura (2007), por la fuente de datos que utilizan, el método de inferencia, y el proceso que aplican a sus resultados. Los dos primeros puntos, dentro del área del modelado de usuarios es lo que se llama “modelado”, cuyo objetivo es obtener un modelo. El modelo es la base para establecer una serie de acciones para alcanzar los objetivos finales. Estos puntos forman pasos necesarios en un análisis de minería de datos.

La minería de datos consiste en la extracción no trivial de información que reside de manera implícita en los datos. Dicha información era previamente desconocida y

podrá resultar útil para algún proceso. En otras palabras, la minería de datos prepara, sondea y explora los datos para sacar la información latente en ellos. Los distintos análisis de minería de datos se pueden distinguir primero por los métodos de adquisición de datos y según de dónde se adquieran los datos (de respuestas de los participantes o expertos a preguntas, o de estadísticas de las interacciones). Una vez que se adquieren los datos, se aplican métodos de inferencia para deducir el valor de ciertas características, como la colaboración. Estos métodos de inferencia puede ser el análisis de un experto, lo cual hay que distinguirlo de las respuestas a unas preguntas dada por el experto, comparación entre modelos, análisis de las interacciones, análisis estadístico complejos de los datos, o ningún análisis.

Antes de proponer o establecer un comportamiento del sistema después del análisis, ha habido un grupo de investigaciones que han modelado de alguna forma la colaboración. Dos estrategias principales se han identificado. Ha habido investigaciones que han propuesto un modelo para la colaboración, mientras que otras han propuesto una serie de indicadores relacionados con la colaboración.

Por último, se pueden caracterizar los sistemas según el método de postprocesado de los datos, o de tratamiento de los resultados obtenidos del proceso de extracción de conocimiento. Según (Soler et al., 2005) los sistemas de aprendizaje colaborativo soportado mediante ordenador (en sus siglas en inglés CSCL), y en este caso los sistemas que aquí estamos tratando ya que analizan la colaboración en un entorno web, se pueden diferenciar según sea el tratamiento de los resultados del análisis. Unos sistemas muestran el seguimiento realizado al usuario, otros muestran, además, los resultados de los análisis debidos al seguimiento al usuario, y otros proponen guías o recomendaciones según los resultados de los análisis.

En resumen, en el problema de mejorar el aprendizaje colaborativo en entornos web generales de ED caracterizando al estudiante según su colaboración, el análisis de la colaboración utilizando tecnología de modelado y de minería de datos es recomendable.

El problema de esta tesis se centra en la caracterización de los alumnos según su colaboración. Para ello debe concretarse qué se va a analizar, cómo se va a analizar, y qué se hará con los resultados del análisis. No hay que olvidar la última pregunta ya que sin utilizar los resultados no habría una relación directa con la mejora del

aprendizaje colaborativo de los estudiantes y no se alcanzaría uno de los objetivos que sustentan el origen de esta tesis.

Una vez definido, hay que evaluar la solución al problema para descubrir la calidad de la misma. La evaluación de sistemas adaptativos es otro campo de investigación abierto. (Velsen et al., 2006) inciden claramente en que la evaluación debe centrarse en los usuarios del sistema y requiere utilizar más de un medio de evaluación para evitar sesgos innecesarios. La evaluación del sistema y del aprendizaje, para comprobar que la solución planteada al problema de mejora del aprendizaje colaborativo es correcta.

## Objetivos

La motivación y la definición del problema de esta tesis han dirigido los objetivos de la misma, además de las hipótesis que se utilizan para la resolución del problema definido. Para la realización de la investigación se han postulado un conjunto de hipótesis:

1. que es posible caracterizar la colaboración en entornos abiertos de educación utilizando tecnología en vez del análisis de un experto,
2. que el aumento de control sobre el proceso de colaboración incita a colaborar y mejora la gestión del propio proceso,
3. que el aumento del control para ser efectivo debe realizarse mientras el proceso de colaboración está teniendo lugar, y
4. que un análisis de la colaboración sin conocimiento del dominio es posible y permite el que se repeta en otros entornos educativos.

Las hipótesis requieren la evaluación de la investigación en un escenario real de aprendizaje colaborativo y establecen un conjunto de objetivos. Los objetivos generales han sido:

1. caracterizar a los estudiantes según su colaboración,
2. utilizar un método de análisis que obtenga resultados de manera regular y frecuente,
3. suministrar la caracterización a los estudiantes y evaluar su aprendizaje, y
4. ofrecer una solución general para que los trabajos de esta tesis sean aplicables en otros entornos de aprendizaje colaborativo.

Al enmarcar el problema, se ha dividido el análisis de la colaboración para la mejora del aprendizaje en cuatro partes: qué datos utilizar, qué método de análisis emplear, qué modelo construir, y qué tipo de herramienta utilizar para facilitar los resultados de los análisis a los estudiantes. Para cada parte hay un conjunto de objetivos que se pueden definir. Una herramienta educativa tiene que cerrar el círculo (Soler et al., 2005), es decir, una vez que el resultado final llega al estudiante, hay que evaluar los progresos de éste y utilizarlos como retroalimentación para mejorar cada uno de los pasos de la investigación. La evaluación de los estudiantes también tiene sus objetivos.

Los objetivos del método de adquisición de datos son: que los datos puedan obtenerse durante el proceso de colaboración, y que los datos sean independientes del dominio. Si los datos son independientes del dominio se aumenta la compatibilidad con otros entornos de colaboración, aunque el dominio académico, los contenidos que se enseñan, sean distintos.

Los objetivos del método de análisis son: que se obtengan inferencias relativas a la colaboración, que las inferencias se obtengan durante el proceso de colaboración, y que el método sea independiente del dominio.

Los objetivos del modelo de la colaboración son: que sea utilizado por los estudiantes y tutores, que sea fácil de entender por personas, y que contenga información relativa a la colaboración.

Los objetivos de las herramientas son: que suministren información sobre la colaboración, que sean de fácil uso y fácil de entender los contenidos que muestren, y que ayuden a los estudiantes a mejorar la gestión del proceso de colaboración.

En el caso de la evaluación, los objetivos han sido seguir las recomendaciones de (Velsen et a., 2006) haciendo hincapié en los estudiantes como principales responsables de la evaluación tanto por su opiniones, por sus acciones, como por sus resultados. Los objetivos de la evaluación de los estudiantes son: que se centre en los estudiantes, y que utilice distintas fuentes de información.

## **Preguntas que esta tesis se ha planteado**

En esta sección queremos plantear unas preguntas directamente que la tesis tiene que responder. Estas preguntas están implícitas en el planteamiento de la tesis y objetivos, pero se considera útil su clarificación. Las preguntas son:

- ¿Es posible desvelar la colaboración de los estudiantes mediante métodos automáticos de inferencia? ¿Cómo?
- ¿Qué datos son los necesarios para realizar la inferencia?
- ¿Es posible, aportando más información sobre la colaboración, ayudar a los estudiantes en la gestión del proceso de la colaboración?
- ¿La mejora del trabajo colaborativo mejora el aprendizaje?
- ¿Es útil la estrategia escrutable en la mejora del proceso de colaboración? ¿Qué características son apropiadas para conseguir dicha mejora?

Estas preguntas serán contestadas a lo largo del trabajo expuesto en esta memoria y finalmente en el Parte IV: conclusiones y trabajo futuro.

## **Organización de la memoria**

En esta memoria se explica la investigación realizada en la tesis. La división en capítulos ha querido seguir las partes en que se ha fragmentado la investigación.

En el Capítulo I se ha introducido las líneas generales que han motivado y planteado esta investigación. El problema del análisis de la colaboración se ha identificado y definido. A continuación se enmarca el problema y se introduce las hipótesis de trabajo, y los objetivos a los que las hipótesis y la motivación han guiado. Para aclarar el trabajo de esta tesis, se plantean una serie de preguntas que serán respondidas a lo largo de la memoria y, especialmente, en el capítulo de las conclusiones.

El Capítulo II es el encargado de enmarcar la investigación en detalle. El marco de trabajo se divide en las partes de la investigación, se describen trabajos relacionados en cada uno de los campos y se hace un análisis crítico. En este análisis crítico se desvelan puntos abiertos del campo y se argumenta nuestra propuesta en esos puntos.

Una vez que la investigación ha sido enmarcada, se describe el trabajo realizado en el Capítulo III. En la introducción se describe la investigación en las partes



importantes en que se ha dividido y esa misma división ha sido la empleada en el Capítulo III para detallar el trabajo realizado.

Toda investigación necesita una parte experimental con la que comprobar las hipótesis y encontrar nuevo conocimiento. Esto es lo que se describe en el Capítulo IV. La complejidad de una investigación en el entorno educativo ha provocado que se hayan propuesto varios experimentos, principalmente referidos al análisis de la colaboración y a las herramientas mostradas a los alumnos para la mejora de su aprendizaje. Todo ello se describirá en el Capítulo IV.

Los resultados de los experimentos y evaluaciones serán mostrados en el Capítulo V. Este capítulo tiene sentido para aclarar la diversidad de los datos obtenidos en la parte experimental. También se deducirán las conclusiones que se desprenden de los resultados de los experimentos y evaluaciones.

Esta memoria termina con el Capítulo VI donde se establecen las conclusiones de la tesis. Como conclusiones se responderán a las preguntas planteadas en el Capítulo I. En este capítulo se acabará proponiendo trabajos futuros para completar y mejorar puntos abiertos que se han detectado en esta investigación.

# Capítulo II: Trabajos relacionados

## Introducción

En el Capítulo I se ha introducido esta tesis y se ha definido el problema al que se enfrenta. La motivación de la tesis se apoya en el trabajo de (Johnson y Johnson, 2004), los cuales indican cinco condiciones para garantizar que el aprendizaje colaborativo se lleva a cabo. La estrategia de aprendizaje colaborativo se utiliza por la compatibilidad con la ED. La falta de interacción social entre alumnos y estos con el equipo docente se puede paliar utilizando el aprendizaje colaborativo.

La unión de la ED y la colaboración produce que una gran cantidad de estudiantes, con características muy diversas (Boticario y Gaudio, 2000b), puedan participar en experiencias de aprendizaje colaborativo. Los entornos colaborativos se generalizan y hacen más difícil cumplir las cinco condiciones, para garantizar el aprendizaje colaborativo, únicamente con el trabajo de los tutores de la experiencia.

Es en este punto donde nace la tesis ya que se propone, mediante métodos de análisis de la colaboración, mejorar el aprendizaje colaborativo. Por tanto, esta investigación abarca tres cuestiones fundamentales: 1º caracterización del entorno colaborativo, 2º análisis de la colaboración y 3º crear un modelo que proporcione la suficiente información para realizar las acciones correctivas apropiadas.

Este capítulo trata de enmarcar la tesis dentro de las investigaciones actuales sobre la colaboración. En primer lugar se presentará los entornos colaborativos existentes. En segundo lugar se tratará de las investigaciones que han abordado el análisis de la colaboración. Pasaremos a continuación a discutir sobre distintos acercamientos al modelado de la colaboración. Por último se describirán distintos enfoques sobre la utilización de la información relativa a la colaboración centrándonos en sus usos pedagógicos, pero no en exclusiva.

## **Caracterización del entorno colaborativo**

Los entornos colaborativos son aquellos en los que se realizan tareas de trabajo en equipo. Se puede definir la colaboración como un trabajo grupal para obtener metas compartidas (Johnson y Johnson, 1994). Otros autores plantean distintas definiciones matizando lo dicho (Barkley et al., 2004). En esta sección se caracterizarán los entornos educativos donde se trabaja en equipos para obtener metas comunes, es decir, entornos colaborativos.

Los entornos colaborativos se distinguen por suministrar un conjunto de servicios para la colaboración, lo que incluye medios de comunicación entre los participantes. Sin embargo, cuando hablamos de educación puede no ser suficiente con suministrar herramientas de comunicación para que ocurra el aprendizaje colaborativo (Johnson y Johnson, 1989; McGrath, 1984). Que los estudiantes puedan comunicarse entre sí, no significa que colaboren, al igual que un alumno al que se le proporciona un libro no significa que lo vaya a leer. Sobre todo en entornos de ED hay que asegurarse que el aprendizaje colaborativo ocurre con un trabajo de diseño y estudio del entorno antes y durante la experiencia de aprendizaje colaborativo.

Johnson y Johnson (2004) estudiaron entornos de aprendizaje colaborativo que utilizan las TIC. Gran parte de estos entornos utilizan las ventajas de la red. Estos autores establecen cinco condiciones para que el esfuerzo cooperativo sea más productivo que el esfuerzo competitivo e individual. Estas condiciones ya han sido descritas en el Capítulo I. Si el entorno y la experiencia de aprendizaje colaborativo cumplen las condiciones se puede garantizar, o más bien controlar, que el aprendizaje colaborativo se produce.

Los entornos colaborativos vía Web están basados en investigaciones en torno a los sistemas colaborativos soportados por ordenador (CSCL es sus siglas en inglés), que es un área de innovación pedagógica que desarrolla un modelo educativo basado en aumentar la interacción entre los miembros de un grupo cuyo interés común es el aprendizaje y que cuenta con el ordenador como el elemento para interactuar y el medio para comunicarse.

(Jermann et al., 2001) y (Soler et al., 2005) han estudiado en profundidad los sistemas CSCL y han propuesto una serie de fases con las que describir el proceso de análisis de la colaboración:

- Fase primera o recogida de datos. Esta fase cubre la observación y recogida de las interacciones.
- Fase segunda o la construcción de un patrón o plantilla de la interacción. Esta fase se refiere a la selección e inferencia de uno o más indicadores de alto nivel con los que representar el actual estado de interacción.
- Fase tercera o donde se compara el actual estado de interacción con uno deseado. El estado deseado se define como un conjunto de valores de los indicadores con los que se puede discriminar la productividad de las interacciones.
- Fase cuarta o de guía/consejo, donde se proponen una serie de actuaciones si hay diferencias entre el estado deseado y el actual de interacción.
- Fase quinta o de evaluación y diagnóstico. Después de la fase de evaluación/consejo, se evalúa dicha actuación.

E incluso (Soler et al., 2005) ha clasificado los sistemas CSCL según la fase cuarta en:

- Herramientas de monitorización o seguimiento que recogen automáticamente los datos de los estudiantes acerca de la interacción, y muestran esta información.
- Herramientas metacognitivas que muestran la información deducida o inferida en el proceso de minería, además de las interacciones.
- Sistemas de guía los cuales proponen medidas correctivas para ayudar al estudiante, una vez inferida la información adecuada.

Un sistema CSCL se puede dividir en dos espacios de trabajo: uno para cada usuario individual, que es privado, y otro común para los miembros del equipo. El tutor o tutores de la experiencia de aprendizaje colaborativo atienden a los alumnos tanto de forma individual como colectiva. El sistema o los tutores proponen a los alumnos una serie de tareas colaborativas. El sistema proporciona las herramientas y servicios necesarios para realizar las tareas de forma colaborativa. La comunicación entre los miembros del mismo equipo se puede realizar síncrona, al mismo tiempo, o asíncrona. En (Jermann et al., 2001) y (Soler et al., 2005) se pueden encontrar gran cantidad de sistemas que siguen lo anterior expuesto.

Conviene dejar claro que, para conseguir que la colaboración se realice, muchos sistemas imponen restricciones temporales a las tareas (Jermann et al., 2001; Fiaidhi y

Mohammed, 2006). Esto es obvio en el caso de la colaboración síncrona, cuando todos los participantes en la experiencia colaborativa tienen que estar al mismo tiempo conectados al sistema. Este planteamiento no aprovecha una de las ventajas de la ED que es la que no tienen porqué los estudiantes trabajar al mismo tiempo. Sin embargo, para controlar la colaboración cuando la comunicación es asíncrona (Dewiyanti et al., 2007), los sistemas CSCL suelen proponer restricciones temporales, obligando, por ejemplo, a que la comunicación se realice durante un periodo de tiempo determinado. Esto sólo puede ocurrir cuando el sistema, tutores o el equipo docente pueden actuar conforme a dichas exigencias, proponiendo actividades para que los alumnos las resuelvan en un plazo.

## **Discusión**

Los sistemas CSCL animan a los participantes a trabajar en objetivos comunes. En entornos colaborativos esto se traduce en la promoción del aprendizaje colaborativo. Sin embargo, promover o motivar el aprendizaje colaborativo, ofreciendo medios de comunicación y tareas de trabajo en equipo, no es garantía de que el aprendizaje colaborativo se produzca. Los sistemas CSCL plantean normalmente algún tipo de restricción, para facilitar el trabajo de los profesores o tutores a la hora de proponer y comprobar el aprendizaje colaborativo.

(Gaudioso y Boticario, 2003) señalan que hay un cierto perfil de alumnos que no permite estrictas restricciones a las tareas colaborativas en un sistema CSCL. Este es el caso de los alumnos participantes en la experiencia educativa que se ha investigado en esta tesis. Los alumnos de la UNED que pueden ser caracterizados por la diversidad, es decir, tienen distintos objetivos y motivaciones, y se distinguen por las diferencias de edad, lugar de residencia, tipo de trabajo, tipo de responsabilidades, etc. (Boticario y Gaudioso, 2000b), y han elegido estudiar en esta universidad por ser la que más se adapta a sus necesidades. Si se desea que estos estudiantes participen en una experiencia de aprendizaje colaborativo, hay que flexibilizar las restricciones en las tareas colaborativas para que los alumnos, en su diversidad, puedan participar en ellas. Por ejemplo, los estudiantes que han participado en esta investigación tienen una edad media superior a los 30 años, y alrededor de la mitad trabajan y tienen familia. Si se desea que una proporción alta de estudiantes pueda participar hay que considerar que los participantes tendrán que realizar al mismo tiempo otras tareas,

ajenas al aprendizaje, y distintas para los distintos miembros de los grupos que trabajen en equipo.

En esta investigación se ha ofrecido una experiencia de larga duración de aprendizaje colaborativo, con un modelo educativo que ofrecer a los participantes la posibilidad de gestionar su propio proceso de colaboración. De esta forma se han flexibilizado las restricciones de los sistemas CSCL. Las tareas colaborativas que se han propuesto animaban al trabajo en equipo y han permitido que cada grupo de participantes pudiera gestionarla para adaptarse a todas sus restricciones y necesidades. Esto produjo un aumento en la dificultad del análisis del funcionamiento de los grupos por parte del tutor de la experiencia, ya que éste no tenía el control sobre las tareas que los distintos grupos o equipos estaban haciendo.

Si el análisis de expertos no es un medio posible de análisis de los procedimientos colaborativos, hay que recurrir a otras técnicas de análisis en las que tutores o expertos tengan menor participación.

## **Análisis de la colaboración**

Para realizar análisis dentro del entorno web es necesario aplicar técnicas capaces de obtener, filtrar y analizar ingentes cantidades de datos provenientes de múltiples fuentes. Este campo es el objeto de las técnicas de minería de datos.

La minería de datos es un campo de Inteligencia Artificial (IA) que consiste en la extracción no trivial de información que reside de manera implícita en los datos. Dicha información era previamente desconocida y podrá resultar útil para algún proceso. Más precisamente, la minería de datos prepara, sondea y explora los datos para sacar la información oculta en ellos (Kantardzic, 2003). Esta técnica ha demostrado su utilidad en distintos dominios, como el comercio electrónico (Srinivasa Raghavan N. R., 2005), educación (Baker R., 2007), Astrofísica (Sarro et al., 2009), Geografía (Wang et al., 2006), etc. El objetivo de la minería de datos es descubrir conocimiento y éste se suele expresar mediante reglas (Clark y Niblett, 1998; Sánchez y Otero, 2007), una clasificación de los datos (Quinlan, 1993; Otero y Sánchez, 2005), un agrupamiento de los datos según cierta semejanza entre ellos (Minaei-Bidgoli y Punch, 2003; Rustagi, 1994), etc. Desde el inicio de la minería de datos ha ido surgiendo gran cantidad de software específico para ayudar en las labores de

prospección y minería de datos. Este trabajo no tiene como objeto la realización de un estudio detallado de cada sistema. Hay gran cantidad de libros de interés sobre este tema como (Han y Kamber, 2006; Berry y Linoff, 2004).

Aunque se ha utilizado desde hace tiempo la minería de datos en el dominio de la educación en los últimos años ha habido un aumento representativo del interés de utilizar técnicas de minería de datos en educación (Véase el ciclo de conferencias sobre EDM (2008, 2009)).

Dentro de este área la investigación se está dirigiendo hacia la mejora del software educativo gracias a los análisis de los datos de interacción (Bratitsis et al., 2008), la representación del dominio, la evaluación de las intervenciones de los profesores en entornos soportados por ordenador (Feng et al., 2009), caracterización del comportamiento y emociones de los estudiantes (Nugent et al., 2009), la integración de la minería en las teorías pedagógicas, la mejora del soporte al profesor (Bravo y Ortigosa, 2009), y la transferencia de técnicas de minería a otros dominios de aplicación (Ben-Naim et al., 2009). Estos son los tópicos de la conferencia EDM'09. Como puntos abiertos en esta área se pueden citar la falta de metodología y estándares en el análisis de las interacciones, así como la falta de análisis empíricos (Bratitsis y Dimitracopoulou, 2006), siendo un ejemplo de estas cuestiones el dominio de la colaboración (Strijbos y Fischer, 2007). Estos autores señalan que el análisis de la colaboración en entornos educativos vía ordenadores o web es un campo todavía poco explorado.

Romero y Ventura (2007) recopilan los trabajos en este ámbito concluyendo que la minería de datos o el descubrimiento de conocimiento desde bases de datos es la técnica que se centra en la extracción automática de patrones implícitos e interesantes desde grandes colecciones de datos. Este descubrimiento de conocimiento desde bases de datos puede utilizarse no sólo para aprender el modelo del proceso de aprendizaje o el modelado del estudiante, sino también para evaluar y mejorar los sistemas de e-learning descubriendo información útil para el aprendizaje. La minería de datos es un paso en la proceso general de descubrimiento de conocimiento desde bases de datos que consiste en preprocesamiento, minería de datos y postprocesado. Los autores afirman que la minería de conocimiento debe entrar en bucle con el sistema y así orientar, facilitar y mejorar el aprendizaje en su conjunto. No sólo convertir los datos en conocimientos, sino también el uso de los conocimientos extraídos para la toma de

decisiones.

## **Análisis de la colaboración mediante minería de datos**

Como establecieron Romero y Ventura (2007), el proceso de extracción de conocimiento se divide en tres fases: preproceso, minería de datos y postproceso. En este apartado nos centramos en las técnicas de minería de datos empleadas por las investigaciones que han estudiado la colaboración.

Aunque hay una falta de metodología y estándares en el análisis de la colaboración en entornos educativos, se han realizado distintos experimentos para medir o identificar la colaboración que se realizaba entre usuarios de un sistema. De estos experimentos se deduce, en primer lugar, que hay que tener en cuenta el método de adquisición de la información, lo que se corresponde con el preproceso de la técnica de minería de datos utilizado. Se pueden identificar tres métodos:

- Cualitativo (Meier et al., 2007; Kahrimanis et al., 2009): donde o se pregunta directamente a los individuos participantes en la investigación, o expertos evalúan las actividades de los participantes.
- Cuantitativo (Talavera & Gaudioso, 2004; Redondo et al., 2003; Hong, 2001; Bratitsis et al., 2008): donde se recoge información de estadísticas de las actividades de los participantes.
- Mixta (Collazos et al., 2002; Collazos et al., 2007; Daradoumis et al., 2006; Martínez et al., 2006; Perera et al., 2007; de Pedro, 2007): el uso de ambos métodos a la vez.

Lo anterior es una clasificación de distintos preprocesos. Cuando ya hablamos de minería de datos, estos sistemas se pueden caracterizar según el método de inferencia utilizado para deducir el valor de ciertas características, como la colaboración, que se ha producido o se está produciendo. Los métodos pueden ser:

- análisis de un experto (Meier et al., 2007; Kahrimanis et al., 2009): donde un experto analiza los datos y realiza un juicio,
- comparación con un modelo preexistente (Redondo et al., 2003),
- diferentes técnicas estadísticas o de aprendizaje automático, como árboles de decisión, cluster, minería de patrones, etc., (Talavera & Gaudioso, 2004; Redondo et al., 2003; Hong, 2001; Perera et al., 2007; de Pedro, 2007),



- técnicas de análisis de las interacciones, las cuales obtienen indicadores estadísticos de las interacciones de los usuarios del sistema (Daradoumis et al., 2006; Martínez et al., 2006; Bratitsis et al., 2008),
- incluso se pueden caracterizar los sistemas por no utilizar ningún sistema de inferencia (Collazos et al., 2002; Collazos et al. 2007).

Evidentemente, la extracción de conocimiento desde bases de datos, sin hacer una definición técnica de lo que es una base de datos, lo podrá hacer un experto, pero cuando tratamos con grandes cantidades de datos, como los que pueden ser recogidos en un entorno educativo vía web, lo que se entiende por una técnica de minería de datos es el uso de técnicas estadísticas o de aprendizaje automático.

Por último, se pueden caracterizar los sistemas según el método de postproceso de los datos, o lo que hacen con los resultados obtenidos del proceso de extracción de conocimiento. En este punto volvemos a comentar el trabajo de (Soler et al., 2005), donde se clasifican los sistemas CSCL según lo que hagan con los resultados del análisis como: (1) herramientas de monitorización y seguimiento, donde no se muestra ninguna valoración, inferencia o juicio sobre el usuario (Meier et al., 2007; Kahrimanis et al., 2009; Collazos et al., 2007; Daradoumis et al., 2006; Martínez et al., 2006; Bratitsis et al., 2008); (2) herramientas metacognitivas que muestran la información inferida en el proceso de minería, además de las interacciones (Redondo et al., 2003; Talavera y Gaudioso, 2004; Perera et al., 2007); (3) sistemas de guía los cuales proponen medidas correctivas para ayudar al estudiante, una vez inferida la información adecuada.

En pocas palabras, hay sistemas cuyo objetivo es mostrar los datos sobre los alumnos, por lo cual lo principal es saber cómo adquieren los datos, los filtran y los muestran. Para otros lo principal es inferir a partir de los datos recogidos información útil, y los últimos, realizan guías o hacen recomendaciones a los alumnos gracias a los resultados inferidos. Ha habido ejemplos de sistemas que guían en entornos colaborativos, pero no utilizan la colaboración como la característica a guiar (Goodman et al., 2005; Santos y Boticario, 2008).

No se ha encontrado en la literatura actual ningún sistema que analizando la colaboración proponga recomendaciones o guías a los estudiantes participantes, lo que no significa que al gran ritmo de avances en este campo sea previsible encontrar ejemplos de este planteamiento en breve.

## **Investigaciones**

En este apartado entramos en detallar los sistemas antes clasificados, aunque se dará más importancia al método de inferencia, es decir, a lo que Romero y Ventura (2007) llamaban minería de datos y con el que es posible inferir juicios gracias a los datos. Debido a la importancia que se da en esta tesis, comenzamos explicando los sistemas que han utilizado la tecnología de aprendizaje automático.

**DomoSim-TPC** (Redondo et al., 2003) es una herramienta metacognitiva, donde se adquieren datos cuantitativos de las interacciones de los usuarios en un curso. El curso se encuentra estructurado en actividades, por lo que las interacciones se etiquetan según la actividad. Un experto construyó un modelo de la colaboración según las interacciones. El modelo es comparado mediante un algoritmo de lógica borrosa con las interacciones del usuario, dando como resultado el grado de pertenencia en un nivel de colaboración que el experto ha establecido. Esta información es mostrada al tutor. Este sistema se caracteriza por un método de adquisición cuantitativo, donde el método de inferencia es la lógica borrosa en comparación con un modelo establecido, siendo una herramienta metacognitiva, al mostrar el valor inferido de la colaboración. Este sistema, al utilizar una comparación con un modelo preestablecido, es dependiente del dominio. Al cambiar el dominio, el modelo cambia y primero habría que realizar una investigación para encontrar un nuevo modelo para el nuevo dominio.

**Talavera y Gaudioso (2004)** explican un procedimiento de generación de agrupaciones utilizando tecnología de clustering con los datos de un curso. Definen un conjunto de atributos cuantitativos con los que construyen el dataset, o colección de datos en forma tabular donde cada columna representa un atributo y una fila una determinada instancia, que el algoritmo de clustering EM utiliza en su ejecución. Las columnas del dataset representan los atributos cuantitativos, mientras que las filas o instancias son los valores de dichos atributos por estudiante. El algoritmo agrupa las instancias del dataset, o estudiantes, según la semejanza entre ellos. Una vez que se tienen los grupos hechos, los comparan con la colaboración que conocen por parte de los alumnos, y que el tutor o tutores del curso han dado. De la comparación hay que los grupos encontrados identifican diferentes comportamientos de los alumnos respecto a la colaboración, por lo que encuentran que el método utilizado identifica en los alumnos la colaboración realizada. El método de obtención de datos es únicamente

cuantitativo y el método de inferencia un algoritmo de clustering que se utiliza para comparar los cluster o grupos con características de los alumnos deducidas por un experto.

**Hong (2001)** presenta un método de evaluación del trabajo colaborativo. Propone un conjunto de indicadores que son el resultado de un análisis estadístico de las interacciones realizadas por los alumnos. El objetivo es monitorizar a los alumnos y a los grupos, por lo que los indicadores son presentados al tutor. Una vez que el curso terminó y tubo información de cómo concluyo las tareas, comparó dichos resultados con las interacciones para dar un significado a los indicadores definidos previamente en relación a la colaboración. En este sentido uno de los objetivos es evaluar sus propios indicadores.

**de Pedro (2007)** realiza un estudio de la colaboración en un entorno de aprendizaje colaborativo basado en la técnica de Wiki<sup>6</sup>. En el método que propone, los estudiantes tienen que escribir un documento de forma colaborativa al estilo de las wikis. Solicita a los estudiantes que etiqueten sus contribuciones al documento con una serie de etiquetas establecidas o que propongan nuevas. De este modo, las contribuciones están etiquetadas e identificado el autor. A partir de ahí, realiza un análisis estadístico del número, tamaño, tipo y corrección de las contribuciones de un estudiante. El tutor también revisa las contribuciones aportando el análisis. Evalúa la metodología propuesta en tres casos de estudio y concluye advirtiéndolo de las ventajas del método al dar una herramienta que hace posible la autorregulación y de las desventajas de un análisis cuantitativo de las contribuciones al no considerar la información semántica y estructural de los documentos.

**Perera et al. (2007)** han investigado el trabajo colaborativo de una serie de estudiantes en un curso de informática, el cual consistía en programar en equipo. La comunicación se realizó mediante “tickers” (servicio que consiste en la asignación de tareas a un alumno) y la wiki, que permiten la creación y mantenimiento de un documento por varios autores. De los dos servicios obtienen datos cuantitativos para realizar un análisis estadístico y otro aplicando técnicas de clustering. Para validar ambas técnicas utilizan los resultados finales en el curso de los equipos y alumnos. Además de las dos técnicas de análisis, utilizan minería de secuencia de patrones, que

---

<sup>6</sup> <http://es.wikipedia.org/wiki/Wiki>

consiste en dividir secuencias de comunicación o trabajo en eventos ocurridos. Esta división la realizan de forma manual. A estos datos se les analiza comparándolos con los resultados de los equipos y los alumnos en el curso. Realizan tres tipos de análisis con los datos: análisis estadísticos sencillos (no obtienen nada en claro), análisis de cluster (sólo identifican al mejor grupo), y análisis de secuencias donde se tiene en cuenta la secuencia temporal (sí consiguen identificar los mejores grupos de los peores). Obtienen como conclusiones que el trabajo y análisis no es completo, pero que existe una relación entre el mejor liderazgo y la mejor interacción del grupo. Es decir, cuando el líder del grupo era mejor y cuando la interacción ha sido más equilibrada, el grupo ha sido mejor.

Además de las investigaciones anteriores, otras también han investigado sobre la colaboración pero sin utilizar técnicas de aprendizaje automático. Empezamos describiendo un sistema que utiliza el juicio de expertos como método de inferencia.

En **Meier et al. (2007)** y **Kahrimanis et al., 2009** se propone una metodología de análisis de la colaboración llamado “plan de evaluación” (rating scheme en inglés). El plan de evaluación es un método cualitativo de análisis de la colaboración y su objetivo es identificar la validez de los atributos cualitativos propuestos. Este método consiste en hacer estimaciones de la colaboración entre una pareja en una tarea colaborativa, que se realiza vía web y que se graba en vídeo para un análisis posterior. Proponen varios indicadores para identificar la colaboración y unos expertos evalúan las escenas gravadas en video según los indicadores propuestos. Estas investigaciones no tienen como finalidad desvelar la colaboración sino evaluar el propio método de análisis de la colaboración.

Unas cuantas investigaciones no han realizado inferencias o juicios respecto a la colaboración de los estudiantes. Los objetivos eran variados, desde la validación de indicadores a la monitorización y seguimiento del estudiante.

**Collazos et al. (2007)** proponen un método principalmente cuantitativo de detección de la colaboración. Utilizan ciertos atributos obtenidos de un análisis estadístico de las interacciones y un análisis de los contenidos. El primer objetivo de la investigación fue identificar la validez de los atributos utilizados (Collazos et al., 2002). Al continuar la investigación el objetivo que siguieron fue el de monitorizar al estudiante según los indicadores que habían propuesto. No proponen ningún juicio o inferencia.

Dentro de el contexto del que estamos hablando, el análisis de la colaboración, hay que mencionar un conjunto de investigaciones que realizan análisis de las interacciones. Estas investigaciones se agrupan dentro de la red de excelencia llamada Kaleidoscope (<http://www.noe-kaleidoscope.org/pub/>). Kaleidoscope es la Red Europea para la innovación científica en materia de Tecnologías para la Educación. A este fin, para mejorar la calidad de la experiencia del aprendizaje y la propia forma de aprender, Kaleidoscope se apoya en equipos cooperativos de investigadores, líderes en campos clave: ciencias de la educación, tecnologías de la información, ciencias sociales.

La técnica del análisis de las interacciones se centran en la obtención de indicadores de nivel medio obtenidos mediante procesos estadísticos principalmente del conjunto de interacciones, en bruto, que un usuario de un sistema realiza. Dentro de Kaleidoscope se ha preparado un marco de trabajo en el análisis de las interacciones de un estudiante en un entorno web educativo (Gómez-Sánchez et al., 2009). De esta manera, proponen un conjunto de indicadores basados en las interacciones de los estudiantes y un conjunto de métodos de análisis para modelar las interacciones. En ese trabajo se echa en falta propuestas de inferencia sobre la colaboración de los estudiantes. Dentro de Kaleidoscope, un grupo de investigaciones se han centrado en la colaboración. El objetivo de estas investigaciones ha sido monitorizar y hacer seguimiento de las interacciones de los estudiantes en un entorno de aprendizaje colaborativo (Daradoumis et al., 2006; Martínez et al., 2006; Bratitsis et al., 2008).

**(Daradoumis et al., 2006)** y **(Martínez et al., 2006)** proponen un marco teórico para el análisis de la colaboración dividido en tres capas. El modelo de capas se divide en indicadores de alto nivel, que son evaluaciones cualitativas de los tutores. En el nivel medio realizan un análisis de redes sociales, y utilizan datos de dicho análisis y del log de comunicaciones para establecer unos indicadores cuantitativos en el nivel inferior (densidad de la red, número de comunicaciones). No proponen un método de inferencia. Su objetivo es mostrar el análisis en capas que han especificado.

El caso de **(Bratitsis et al., 2008)** es distinto ya que utilizan dos capas obteniendo los datos de los ficheros log y de la base de datos, donde está incluido un análisis de

las redes sociales. Tampoco realizan ninguna inferencia con los datos obtenidos pero sí los preparan para su visualización en una herramienta de monitorización.

### **Análisis crítico**

Las investigaciones descritas anteriormente han dejado algunos puntos abiertos si se desea asegurar que el aprendizaje colaborativo se realiza. Estos puntos abiertos son: la dependencia o independencia del dominio, la fuente de datos utilizada para poder realizar análisis frecuentes y regulares, y el propio método de análisis de la colaboración. Aunque aquí nos centramos en el análisis y extracción de inferencias, evaluaciones o conclusiones sobre la colaboración, el ciclo del análisis de la minería de datos se cierra al utilizar el conocimiento extraído para la toma de decisiones.

La dependencia del dominio se hace patente en (Redondo et al., 2003), ya que el modelo de las interacciones, es decir, el modelo que señala qué interacción debería ocurrir, se realiza antes de que las interacciones se hagan. Esto indica que antes de que el curso comience se sabe, o se cree saber, cómo los alumnos se deben comportar en ciertas circunstancias. En unas acciones donde intervienen dos o más personas donde deben comunicarse, el conocimiento apriorístico es dudoso salvo que el entorno limite el número de acciones y la forma de desarrollarlas en un momento determinado. Esto requiere, como poco, un alto conocimiento del dominio donde las acciones se realizan. Por lo tanto, el modelo tiene intrínseco conocimiento o información del dominio que, al extrapolar este acercamiento el entorno de destino tendría que dar soporte a unas actividades colaborativas tan estructuradas como las que proponen estos autores.

La obtención de resultados de un análisis de forma regular y frecuente obliga a que el método de adquisición de los datos y el método de análisis se puedan realizar igualmente de forma regular y frecuente. Un método de adquisición de datos cuantitativo cumple la restricción anterior, pero un método de adquisición de datos que requiera la intervención regular y frecuente de los participantes en la experiencia colaborativa o de expertos, es mucho más complicado de realizarse por la gran carga de trabajo, tanto para los estudiantes como para los expertos. En estos contextos, comunes en la ED y e-learning, para asegurarse de que el análisis se puede realizar de forma regular y frecuente en cualquier circunstancia, el método de adquisición de datos bastante apropiado es el método cuantitativo.

En los entornos colaborativos las interacciones puede llegar a ser muy complejas. Una solución es reducir el problema centrándose en un solo medio de comunicación o un tipo de trabajo colaborativo. Esto es lo que ha hecho de Pedro (2007) al analizar la colaboración en un entorno de trabajo en wikis. Solamente pidiendo a los estudiantes que etiqueten sus contribuciones a un documento común consigue hacer análisis estadístico con el objetivo de ayudar a los estudiantes en la autorregulación de su trabajo.

Otro problema es el análisis en sí. La quinta condición de Johnson y Johnson (2004) aconseja un procesamiento del trabajo por lo que el análisis de la colaboración debería proveer con una valoración, deducción o inferencia sobre la colaboración de los estudiantes. Hong (2001) realiza un sencillo análisis estadístico de las interacciones sin aportar una valoración de la colaboración, al igual que (Collazos et al., 2002) y (Collazos et al., 2007), los cuales obtienen una compilación de las interacciones. (Daradoumis et al., 2006), (Martínez et al., 2006) y (Bratitsis et al., 2008) realizan un análisis de las interacciones definiendo un conjunto de indicadores de nivel medio para ayudar a los usuarios de los respectivos sistemas a que ellos realicen la propia valoración de la colaboración sin informar de ello al sistema. Por lo tanto, tampoco proponen un método de análisis de la colaboración que valore a los usuarios según la colaboración.

Los acercamientos anteriores no cubren las hipótesis de esta tesis por la falta de un método que infiera juicios sobre la colaboración. Para realizar dicho método de forma frecuente y regular, se puede recurrir a la tecnología del aprendizaje automático. Así, Talavera y Gaudioso (2004) utilizaron la técnica de aprendizaje automático llamada clustering, demostrando su utilidad, y se puede afirmar que el método por ellos propuesto se puede realizar de forma regular y frecuente.

En el aprendizaje automático hay diferentes técnicas de abordar los mismos problemas (Borrajo et al., 2006) y no hay un acuerdo o metodología en relación al análisis de la colaboración (Strijbos y Fischer, 2007). Es conveniente un estudio empírico para deducir que técnica es mejor que otra en un determinado problema. Este es otro punto abierto dentro del análisis de la colaboración.

Los puntos abiertos de este campo ya han sido comentados en relación a la necesidad de realizar análisis de la colaboración de forma regular y frecuente. Por un lado, tenemos la fuente de datos o el método de adquisición de información. En este

punto la ventaja es clara si se usa un método de adquisición cuantitativo. En esta tesis hemos utilizado como fuente de datos estadísticas de las interacciones en los foros de los equipos de la experiencia de aprendizaje colaborativo, es decir, un método de adquisición cuantitativo. Las ventajas han sido: la adquisición se podía realizar en cualquier momento; al realizarse las comunicaciones de los estudiantes principalmente en los foros, las interacciones estaban relacionadas con la colaboración entre los estudiantes; al ser los foros un medio común de comunicación en entornos de e-learning, el método podrá ser trasladado a otros entornos con facilidad; las estadísticas realizadas no ofrecen información semántica por lo que son independientes del dominio; la literatura informa de la utilidad de analizar los foros como medio de extraer información sobre los usuarios (Dringus y Ellis, 2005).

Otro punto abierto es el método de análisis. De los análisis que hemos visto en otros sistemas que han analizado la colaboración, aquellos que han utilizado técnicas de aprendizaje automático son los más idóneos según los objetivos de la tesis (Redondo et al., 2003; Talavera y Gaudioso, 2004; Perera et al., 2007). Las técnicas de aprendizaje automático son apropiadas en estos contextos ya que después de un estudio sobre las mismas, son capaces de aplicarse de forma regular y frecuente automatizando el proceso del análisis (Russel & Norvig, 1995). Los principales problemas de estas técnicas son que se necesita una fuente de datos alternativa para validar los resultados, y que no es conocido a priori cuál técnica es mejor que otra para un determinado problema.

En este trabajo se han probado dos acercamientos al análisis de la colaboración que utilizan tecnologías de aprendizaje automático distintas. La elección de la técnica ha sido guiada por el tipo de inferencia que se quería obtener. Por un lado se deseaba tener un modelo de los alumnos que nos dijera qué indicador estaba más relacionado con la colaboración (Prata et al., 2009). En esta caso la técnica utilizada ha sido los algoritmos de árbol de decisión. Por otro lado, se quería clasificar a los estudiantes según su colaboración dependiendo de los datos, algo que ya ha propuesto (Talavera y Gaudioso, 2004; Perera et al., 2007). La técnica utilizada para este objetivo ha sido la de clustering.



## **Modelado de la colaboración**

Con modelado queremos señalar los procesos que se realizan para la creación de un modelo. En este sentido, el modelado es dinámico y se basa en acciones, análisis, seguimientos, etc., con el objetivo de obtener un modelo, es decir, un objeto, lo cual podría decirse que es estático, en tanto que recoge la instantánea de una situación a lo largo del tiempo.

Aunque la colaboración es una estrategia ampliamente utilizada en entornos web no ha sido el foco de muchas investigaciones centradas en modelar al usuario. En entornos educativos el nivel de conocimiento del alumno ha sido la característica principal investigada en la mayoría de sistemas de modelado del alumno (Kobsa, 2001; Brusilovsky y Millan, 2007), incluso hay investigaciones que centran su estudio en otras características del alumno teóricamente más difíciles de determinar, como es la motivación (Cocea y Weibelzahl, 2007).

En este apartado se analizan las contribuciones de otras investigaciones al modelado de la colaboración en entornos educativos vía web u ordenador. Se pueden distinguir las investigaciones que han modelado la colaboración según los resultados obtenidos en:

- la construcción de un modelo de la colaboración,
- la determinación de indicadores de la colaboración.

Un modelo contiene indicadores, pero su finalidad es distinta a la de los indicadores que posee. Un modelo conlleva una descripción “del mundo” para un fin o variedad de fines. Los indicadores “asumen” un modelo implícito y se centran en la toma de decisiones.

Los modelos son estructuras utilizadas como contenedores de información sobre diferentes contenidos y que en la propia estructura del modelo relaciona la información contenida entre sí (Inaba et al., 2003). En resumen, un modelo es un conjunto de atributos relacionados entre sí.

Los indicadores son atributos con su valor. La diferencia respecto a los modelos que aquí se considera es que los indicadores no se relacionan entre sí, o esa no es la principal característica. No hay una jerarquía ni comunicación entre ellos.

## Modelo de la colaboración

Algunos trabajos de investigación han construido un modelo de colaboración, es decir, una serie de atributos estructurados entre sí e incluso con partes externas al propio modelo de la colaboración. El objetivo de estos modelos ha sido tener una referencia donde comparar el modelo extraído del alumno e identificar la colaboración del mismo.

Redondo et al. (2003) presentan un sistema, **DomoSim-TPC**, que modela la colaboración en un entorno de aprendizaje colaborativo vía web. El modelo describe la colaboración de los estudiantes en un curso estructurado. Las interacciones que realizan los alumnos en el curso, al estar éste claramente estructurado en distintas actividades, pueden ser etiquetadas. De esta forma, el modelo del alumno va siendo creado después de que el alumno vaya interactuando. Los atributos cuantitativos que utilizan en el modelo son los siguientes: *número de accesos al sistema que cada estudiante realiza al trabajar en la actividad; número y la media de las contribuciones hechas; la media del tamaño de las contribuciones; tipo de las contribuciones; número de contribuciones respondidas; la profundidad de la discusión para cada tarea del problema; número de accesos a las últimas noticias; número de mensajes enviados (clasificados por su tipo: planificación, coordinación o sistema); número de mensajes instantáneos*. Con los anteriores indicadores construyen variables de más alto nivel relativas al estudiante (*iniciativa, creatividad, elaboración, conformidad, disconformidad*), relativas al grupo (*trabajo, discusión, coordinación, cooperación y colaboración*), y relativas a la participación del estudiante (*modelado, revisión, crítica, justificación, participación*). Estas variables de alto nivel son calculadas mediante lógica difusa. Una vez obtenido el modelo, lo comparan mediante un algoritmo de lógica difusa con un modelo ideal de las interacciones realizadas preestablecido por expertos. Las interacciones del alumno comparadas con el modelo de la colaboración ideal indica la colaboración del alumno en cada momento.

Baghaei y Mitrovic (2007) presentan **COLLECT-UML** un sistema tutorial inteligente. La parte principal del sistema consiste en un modelo del dominio. El modelo del dominio representa, mediante restricciones, como un problema se debe solucionar, y aporta un feedback cuando una restricción no se cumple. El modelo del usuario consiste en los problemas que ha ido haciendo. Estos problemas se comparan

con el dominio y si alguna restricción no es seguida por el usuario, el feedback es enviado al usuario. Un ejemplo del modelo se muestra en la Figura 1, donde la restricción número 54 aparece el código que se asegura de que todos los métodos del problema hayan sido especificados.

```
(54
  "Check whether you have defined all the methods aspecified
  by the problem. You are missing some methods."
  (and (match IS METHODS (?* "@" ?tag ?name ?class_tag ?*))
        (match SS CLASSES (?* "@" ?class_tag ?*)))
  (match SS METHODS (?* "@" ?tag ?name2 ?class_tag ?*))
  "methods"
  (?class_tag))
```

Figura 1: ejemplo de modelo de restricciones (Baghaei y Mitrovic, 2007)

La parte de la colaboración sigue el mismo patrón. Hay un modelo de colaboración basado en restricciones, con un feedback cuando una restricción es violada. Este modelo de colaboración consistente en un conjunto de 25 meta-restricciones con las que representar la colaboración. Cada meta-restricción consiste en una condición de relevancia, otra condición de satisfacción y un mensaje de feedback, que se presenta cuando la restricción es violada.

**Martínez et al. (2003)** describen un modelo para representar acciones de colaboración, destinadas a apoyar el análisis de la interacción en CSCL. Sus principales contribuciones fueron la propuesta de un modelo del contexto, y de una taxonomía genérica basada en un enfoque abajo-arriba que se centra en los actores que intervienen en cada tipo de acción. No presentan atributos. El modelo es un esquema de la estructura que debe seguir un modelo formado por atributos con su valor determinado. En el modelo que presentan, una acción se representa como se puede ver en la Figura 2, donde se puede observar características que definen la acción como la fuente, el documento de que consta la acción, por ejemplo.

```
<!-- Example of an action in a document workspace -->
  <ACTION>
    <ACT.TIMESTAMP>21:04</ACT.TIMESTAMP>
    <ACT.SOURCE ref="laox11"/>
    <ACT.DESC>
      <ACT.IND>
        <ACT.IND.DOC type="Read">
          <DOC key="97577" type="File">
            <DOC.PATH>:laox11</DOC.PATH>
```

```

        <DOC.PATH>Curso 2002-2003</DOC.PATH>
        <DOC.PATH>GENERAL</DOC.PATH>
        <DOC.PATH>Curso01-02</DOC.PATH>
        <DOC.CONTENT>text/plain</DOC.CONTENT>
    </DOC>
</ACT.IND.DOC>
</ACT.IND>
</ACT.DESC>
</ACTION><SITUATION id="PUZZLE-COLAB">

```

**Figura 2: modelo de una acción (Martínez et al., 2003)**

**Vidou et al. (2006)** propusieron varios modelos útiles para describir Comunidades de Práctica (entorno virtual que agrupa usuarios con las mismas aficiones o conocimientos): la comunidad, el actor, el perfil de alumno, competencia, colaboración, proceso o actividad, y lecciones aprendidas. Los modelos que proponen son de alto nivel y se encuentran relacionados entre sí y estructurados en una ontología. Para el modelo de la colaboración consultaron la literatura específica sobre la colaboración y crearon un modelo que está compuesto de las actividades y definido por objetivos, que implica a actores y que necesita/produce recursos. El modelo que proponen establece las relaciones del mismo con otros modelos, los cuales se refieren a otras partes de los componentes de un curso, como las actividades del curso, los recursos que contiene, los estudiantes o actores. Sin embargo, no proponen ningún atributo, sólo la estructura en general.

**(Barros et al., 2001; Barros et al., 2002)** utilizan la Teoría de la actividad para proponer una ontología para el análisis de la colaboración. La Teoría de la actividad es una meta-teoría, paradigma, o marco de estudio psicológico, con raíces dadas por la psicología histórico-cultural del psicólogo soviético Lev Vygotsky. Sus fundadores buscaban entender las actividades humanas como complejos fenómenos socialmente situados, e ir más allá de los paradigmas del psicoanálisis y de la psicología conductista. (Barros et al., 2001, Barros et al., 2002) se basan en esta teoría por lo que la ontología que proponen no es de la colaboración del alumno como característica del mismo, sino una ontología de las actividades que puede realizar un alumno para colaborar. Un ejemplo de la estructura y propiedades, como el tipo de herramienta, el tipo de reglas, a quienes está dirigida o el objeto de la actividad, que proponen se puede ver en la Figura 3.

```

-Activity
  p/o@n"Tools":Artefact

```

```
p/o@n"Rules":Norms
p/n"Division of Labour":Division of Labour
p/o"Community":Group
p/o"Subject":Individual
p/o@n"Object":Motive
p/o"Outcome":Result
-Action
p/o@n"Action":Activity
p/o"Goal":Goal
```

Figura 3: ejemplo de actividad (Barros et al., 2002)

**Durán (2006)** a fin de lograr una asistencia personalizada que facilite una colaboración eficiente en un entorno CSCL diseña un Modelo de Estudiante Colaborativo que incluye, además de los rasgos personales y del grupo, las destrezas colaborativas que posee, las que serán identificadas a partir de experiencias de colaboración previas. El modelo o perfil de la colaboración contempla el contexto de la colaboración, habilidades para mediar o argumentar en conflictos, y habilidades para motivar, manejar información, solicitar trabajo de otros, delegar, reconocer a otros en una conversación, mantener la cohesión del grupo y para cambiar de tarea durante la conversación. Propone generar distintos perfiles y compararlos con los generados por los alumnos para encontrar el tipo de colaboración.

Las investigaciones anteriores proponen distintos acercamientos al modelo de la colaboración. Describen y estructuran los modelos de acuerdo a la relación con otras características dentro del modelado del usuario, o las características generales internas del modelo de colaboración. Aunque sus objetivos son semejantes, se observan diferencias importantes al entrar en detalle.

### **Identificadores de la colaboración**

En el apartado anterior se han presentado investigaciones que proponen un modelo para la colaboración, o la meta-estructura que debería tener. Otra forma de contemplar el análisis de la colaboración es proponiendo indicadores sin necesidad de estructurarlos o relacionarlos entre sí o con otros atributos ajenos a la colaboración. Estos indicadores pueden venir de un análisis de las teorías sobre la colaboración o dirigidos por los datos de interacción de los usuarios. El objetivo de estos es el mismo que del modelo, dar pistas de la colaboración para deducir la utilidad, eficiencia o

grado de colaboración, y así descubrir si se está colaborando o no, y si esa colaboración mejora, o no, el aprendizaje.

(Collazos et al., 2002; Collazos et al., 2007) proponen 5 indicadores basados en la teoría de Johnson y Johnson (1978) en un entorno de aprendizaje colaborativo vía web. Estos son: *estrategia aplicada, coordinación intragrupal, revisión de los criterios de éxito, monitorización y desarrollo*. Son indicadores de alto nivel pero también proponen atributos de más bajo nivel, alguno de los cuales pueden ser calculados de forma automática. Los indicadores propuestos de la colaboración son calculados mediante relaciones matemáticas de los atributos de bajo nivel. Realizan varios experimentos para concluir que los indicadores parecen apropiados y que la actividad de monitorizar las tareas con estos indicadores mejora el aprendizaje. En (Collazos et al., 2007) proponen, además, la inclusión de otros indicadores socioculturales y de la personalidad para completar los propuestos, pero siempre con el objetivo de monitorizar la colaboración del alumno.

Baldiris et al. (2007) describen una aproximación a la adaptación educativa en **ADAPTAPLAN**. La aproximación consiste en modelar a los alumnos según sus estilos de aprendizaje, nivel de conocimiento y grado de competencia en colaboración. Ellos consideraron seis niveles de competencia en colaboración definidos en (Santos y Boticario, 2004). Los niveles propuestos son: 1) *alumno útil*, 2) *alumno intuitivo*, 3) *alumno con iniciativa*, 4) *alumno comunicativo*, 5) *alumno colaborativo* y 6) *alumno participativo*. Decidieron separar este tipo de competencias ya que definen aspectos importantes en la colaboración y la cooperación del comportamiento del estudiante. Se interesaron por estas características en el modelado del usuario con el fin de establecer su relación con el éxito del proceso de aprendizaje.

**Soller (2001)** propuso el Modelo de Aprendizaje Colaborativo donde un conjunto de indicadores describen la colaboración eficiente en un entorno educativo. Los indicadores son: *participación, conocimiento elemental social, conversaciones de aprendizaje activo, análisis del desarrollo y procesado de grupos, y promover las interacciones*. El principal objetivo de estos indicadores es ayudar a los desarrolladores de un entorno de aprendizaje colaborativo mientras desarrollan el sistema o entorno.

**Park y Hyun (2006)** proponen nuevos elementos para la evaluación de los alumnos en un entorno de aprendizaje colaborativo vía web. Plantean que los alumnos

evalúen a sus colegas (de equipo) y a ellos mismos utilizando como indicadores la *interacción*, la *“colaborabilidad”*, y la *responsabilidad*. Alcanzan dichos indicadores después de un análisis de la literatura y de su propia experiencia. La colaborabilidad representa las actividades del grupo para conseguir los objetivos de aprendizaje del equipo, donde la responsabilidad representa un esfuerzo o contribución individual a su equipo.

**Baeza-Yates y Pino (2006)** hacen un análisis dirigido por los datos de las interacciones en un curso de aprendizaje colaborativo vía web. Ellos proponen un método de evaluación de la colaboración consistente en medir: *la calidad*, *el tiempo gastando*, y *la cantidad de trabajo* que se ha hecho en la experiencia. Realizan un análisis comparativo de los valores anteriores respecto al número de participantes y al tipo de experiencia (en grupos colaborativos o no). Los resultados que obtienen son los siguientes: La calidad aumenta con el número de personas pero menor que linealmente; el tiempo aumenta con el número de personas pero menor que linealmente; el trabajo realizado aumenta con el número de personas pero más que linealmente. Quiere decir que dependiendo de la tarea, habrá un número de personas que la realizarán de la forma más eficiente. Aumentando ese número, la eficiencia disminuye. Los autores buscan un método cuyo objetivo es identificar el número de miembros de un grupo para que este trabaje de forma más eficiente.

**Meier et al. (2007)** y **Kahrimanis et al. (2009)** proponen identificar la eficiencia de la colaboración en un entorno de colaboración mediante ordenador. Realizan dos tipos de análisis con el objetivo de identificar atributos o indicadores de la colaboración. Por un lado encuentran después de una análisis de la literatura cinco aspectos del proceso de la colaboración. Por el otro realizan un análisis de las conversaciones grabadas para encontrar los atributos que identifican la colaboración. En los indicadores que proponen es donde se diferencian las investigaciones. Meier et al. (2007) realiza la investigación en un entorno de aprendizaje médico mientras que Kahrimanis et al. (2009) el entorno de aprendizaje es el de la Ingeniería Informática. Los indicadores que propone Meier et al. (2007) son: *mantenimiento del entendimiento mutuo*; *gestión del dialogo*; *puesta en común de la información*; *alcanzar el consenso*; *división de tareas*; *gestión del tiempo*; *coordinación técnica*; *interacción recíproca*; *orientación individual a la tarea*. Mientras que Kahrimanis et al. (2009) reutilizan parte de los indicadores y proponen algunos nuevos más

adaptados al entorno: *flujo de la colaboración; mantenimiento del entendimiento mutuo; intercambio de conocimiento; discusión; estructuración del proceso de solución de problemas; orientación cooperativa; orientación a las tareas individuales*. El objetivo de los autores es encontrar en el dominio que tratan unos indicadores generales para identificar la calidad de la colaboración.

**Bayón et al. (2007)** presentan una metodología para introducir el modelo educativo del Marco lógico colaborativo en una plataforma orientada a la educación como dotLRN. Afirman que para que la colaboración se realice, no es suficiente con ofrecer servicios de comunicación y tareas comunes. Por ello proponen un módulo que analiza las interacciones para obtener indicadores del: comportamiento de los alumnos, del uso activo de las herramientas de comunicación, del trabajo en las tareas pedagógicas, de los comentarios y valoraciones a distintos objetos de aprendizaje, y uso pasivo de los servicios de comunicación. Los indicadores relativos al comportamiento del alumno intentan identificar el perfil (participativo, intuitivo, útil, no colaborativo, estudiante con iniciativa, estudiante no comunicativo). Sugieren el uso de los indicadores para realizara recomendaciones y ayudar a que la colaboración suceda.

La siguientes investigaciones (Martínez el al., 2006; Daradoumis et al., 2006; Bratitsis et al., 2008) comparten una metodología semejante, análisis de interacciones e indicadores. Esto se debe a que forman parte de la red de excelencia llamada Kaleidoscope. Una de las áreas de estudio de Kaleidoscope es el análisis de las interacciones. El análisis de las interacciones es algo que la gran mayoría de las investigaciones, que han modelado la colaboración, han hecho. Sin embargo, el análisis de las interacciones en Kaleidoscope (Gómez-Sánchez et al., 2009) no se centra exclusivamente en la colaboración, y propone un gran conjunto de indicadores posibles, tanto para la colaboración como para otras características y habilidades de los estudiantes. Por otro lado, las tres investigaciones siguientes tienen como objetivo la mejora de la colaboración por lo que utilizan un conjunto reducido de indicadores de los propuestos en el marco de trabajo de Kaleidoscope.

**Martínez el al. (2006)** y **Daradoumis et al. (2006)** proponen una herramienta de realiza seguimiento de las interacciones de los usuarios sin proponer ninguna valoración o inferencia sobre la colaboración. La propuesta consiste en una serie de indicadores que son presentados en la herramienta. Los indicadores se dividen en tres



capas. El nivel superior contiene evaluaciones del tutor y respuestas de los alumnos a encuestas y test. El nivel medio presenta el análisis de redes sociales. El análisis de las redes sociales es una técnica que muestra quien se ha comunicado con quien y de que forma. Se construye, por tanto, una red de nodos colocando a los actores o individuos más comunicativos en el centro de la red. Los atributos del análisis cuantitativo del nivel inferior están relacionados con el análisis de las redes sociales y estos atributos son: *densidad de la red* (como está tejida la red), *grado de centralidad de actor* (indica el prestigio) y *grado de centralización de la red* (es una medida a nivel de grupo basada en el grado de centralidad del actor, y da idea sobre la dependencia de la red de un pequeño grupo de actores). Los tres niveles son mostrados con el objetivo de monitorizar al alumno y para que, gracias a los atributos, se deduzca la colaboración de cada alumno. Las diferencias entre ambas investigaciones radica los objetivos. Martínez et al. (2006) se centra en la prueba del acercamiento en distintos contextos, mientras que Daradoumis et al. (2006) investigan la colaboración de los usuarios con el acercamiento.

La tercera investigación cambia de acercamiento al dividir a los indicadores utilizados en dos niveles: nivel medio de análisis de red sociales, y nivel inferior de análisis cuantitativo. **Bratitsis et al. (2008)** proponen una metodología de monitorización de la colaboración de los participantes en una red científica (entorno de colaboración vía web de usuarios que comparten una relación científica). El objetivo es analizar las interacciones para rebelar la colaboración, pero sin que el sistema realice ninguna valoración o inferencia sobre la misma. Los autores proponen 13 indicadores divididos en cuatro categorías con el objetivo de ayudar a los usuarios en la colaboración. Las categorías son:

- Indicadores de visitas del sitio. Esta categoría incluye tres indicadores: a) visitantes únicos por intervalo de tiempo, b) visitas por intervalo de tiempo, y c) tasas de tráfico.
- Indicadores de visita a los recursos. Estos indicadores son: a) los recursos y páginas más populares, b) los puntos de entrada más populares, y c) los puntos de salida más populares.
- Indicadores de la manipulación de los recursos. Son los siguientes: a) recursos publicados por un Miembro, b) recursos publicados por un rol, c) visitantes que

han accedido a la página o recurso, y d) los recursos relacionados a un determinado recurso o página.

- Indicadores de comportamiento del usuario: Se trata de: a) los usuarios más activos, b) usuarios relevantes, y c) análisis de las redes sociales.

### **Análisis crítico**

En este apartado hemos estudiado distintos acercamientos al modelado de la colaboración y hemos visto que hay dos estrategias significativas, la construcción de un modelo o el uso de distintos indicadores que desvelen la colaboración. Como muchos han advertido (Park y Hyun, 2006; Strijbos y Fischer, 2007), el análisis o el modelado de la colaboración es un campo aún abierto y que necesita un estándar. Como se habrá podido observar, cada investigación propone diferentes modelos basados en distintas teorías, experiencias, o indicadores de la colaboración. Las teorías que han utilizado van desde la inspirada en el conocido psicólogo Vygotsky (Nardi, 1966) a otras más específicas del comportamiento colaborativo (Johnson y Johnson, 1978), obteniendo de ellas modelos e indicadores distintos y de difícil relación.

Un esfuerzo reseñable es el que se está produciendo dentro de la red de excelencia Kaleidoscope. Distintos grupos han propuesto un marco de trabajo con el objetivo de establecer una metodología al análisis de las interacciones en un entorno educativo (Gómez-Sánchez, 2009). Como se puede deducir de las investigaciones sobre la colaboración que están en el ámbito de Kaleidoscope, en entornos colaborativos sólo se ha considerado la monitorización de los estudiantes y no se plantea ningún juicio o inferencia sobre la colaboración.

Aunque las investigaciones dentro del ámbito de Kaleidoscope proponen indicadores de la colaboración semejantes, se observa que no son compartidos por otras investigaciones. Es más, se observa una gran variedad de indicadores de la colaboración propuestos, basados o no en teorías, debidos al propio contexto de la investigación.

De todo ello se deduce que no hay un consenso en la forma de modelar la colaboración. Esto principalmente se debe a la falta de análisis comparativo entre los distintos acercamientos al respecto y a la falta de base experimental para la misma comparación (Bratitsis y Dimitracopoulou, 2006). Este trabajo debe ser

multidisciplinar, por lo que sobrepasa los límites de esta tesis y del autor de la misma. Pero se detecta la necesidad de un análisis de la colaboración en entornos de aprendizaje colaborativo cuya finalidad sea identificar indicadores que estén relacionados con la colaboración de los alumnos y un método de calcularlos durante el proceso colaborativo.

Al no haber un acuerdo sobre cómo modelar la colaboración la investigación de esta tesis se ha guiado por los objetivos que se querían obtener. El objetivo es mejorar el aprendizaje colaborativo y una de las hipótesis de trabajo de esta tesis es que aumentando el control sobre el proceso de colaboración se mejora el aprendizaje. En la siguiente sección se introducirán herramientas de monitoreo y escrutables que pensamos apropiadas para cumplir los objetivos, dar más control al alumno sobre el proceso y poder extrapolar el acercamiento de esta tesis a otros entornos y dominios. Por lo tanto un modelo de la colaboración es útil, recomendable y necesario.

Sin embargo, por los objetivos de la tesis, el usuario final del modelo de la colaboración serán los alumnos. Hay un conjunto de restricciones que el modelo debe cumplir ya que debe ser entendido por los estudiantes (Bull y Kay, 2008). Las ideas principales que el modelo ha seguido han sido: utilidad, simplicidad y claridad. Además se ha tenido en cuenta a (Muehlenbrock, 2005) el cual afirma que para la agrupación de estudiantes para trabajar en equipo hay que tener en cuenta el contexto de la colaboración. Nos ha parecido apropiado también para los estudiantes que conocieran su contexto y el de los compañeros. También nos ha parecido útil que hubiera información sobre el propio proceso de la colaboración, la cual pudiera indicar a los alumnos una forma de comparar el trabajo de cada uno. Para finalizar, hemos considerado que el modelo debía dar salida a toda la información obtenida por el análisis de la colaboración.

En las figuras 1, 2 y 3 se pueden ver tres modelos distintos orientados al uso por parte de un motor de inferencia. Sin embargo, en esta tesis se ha propuesto un modelo de la colaboración que se usó por los propios estudiantes y cuyas interacciones sean las que se muestran en el modelo. La meta es que los estudiantes lo puedan utilizar para mejorar la gestión del proceso de colaboración. Por estas razones, el modelo debe ser entendible y auto-explicativo, de lo que los modelos mostrados en las figuras 1, 2 y 3 no son un ejemplo. Además de los estudiantes, el tutor de la experiencia de aprendizaje colaborativo también ha sido un usuario final de este modelo, y lo ha

podido utilizar como una herramienta más para la mejora de la colaboración. El modelo que se ha propuesto tiene tres partes diferenciadas:

- Información del contexto de la colaboración. El objetivo de esta información ha sido suministrar a los estudiantes de un mismo grupo datos prácticos para realizar las comunicaciones y que ellos suelen ofrecer a los compañeros, tales como: horarios de conexión a Internet, lugar de residencia, horarios de trabajo, dirección de correo, etc.
- Información del proceso de colaboración. Como la colaboración se ha realizado mediante los foros, estadísticas de las interacciones de los estudiantes en los foros han sido el medio utilizado para informar sobre el proceso de colaboración. Estas estadísticas han sido las utilizadas para realizar el análisis de la colaboración. Las estadísticas tenían que indicar la iniciativa, la actividad, la regularidad y la estimación del estudiante.
- Información sobre la colaboración de los estudiantes. Esta información ha sido el resultado del análisis de la colaboración realizado mediante técnicas de aprendizaje automático. Esta información tenía que dar un juicio alternativo a la colaboración del estudiante de forma frecuente y regular.

Hemos pensado que lo anterior daba información suficiente y fácil de entender para que el alumno tuviera una ayuda a la hora de deducir la calidad del trabajo colaborativo que él u otros compañeros estaban realizando.

## **Estrategias de mejora del aprendizaje colaborativo**

En este apartado hablaremos de estrategias y herramientas para la mejora del aprendizaje en entornos educativos dirigidos a la colaboración. Siguiendo a Soller et al. (2005) hay tres tipos de sistemas o herramientas: los que hacen un seguimiento del usuario o monitorizan las interacciones, aquellos que ofrecen herramientas metacognitivas al alumno mostrando una inferencia o juicio según sus acciones, y de guiado, que ofrecen al alumno una guía o recomendación.

Si se quiere que el sistema cumpla la quinta condición de Johnson y Johnson (2004), la herramienta debe suministrar un juicio sobre la colaboración de forma regular y frecuente. Por definición, sólo las herramientas metacognitivas y de guiado pueden llegar a cumplir dicha condición.

No ha habido ningún estudio comparativo de qué tipo de herramienta ayuda más a los alumnos en la colaboración. Las investigaciones que han tenido un objetivo claro en este punto, es decir, las que tenían como objetivo la creación de una herramienta que mejorara la colaboración, han propuesto complejas herramientas de monitorización (Martínez et al., 2006; Daradoumis et al., 2006; Bratitsis et al., 2008). Otras investigaciones que han realizado juicios sobre la colaboración o han propuesto un acercamiento de difícil traslado a otros entornos o dominios (Redondo et al., 2003), o no han estudiado las reacciones de los estudiantes a la herramienta planteada (Perera et al., 2007). Otras investigaciones han planteado recomendaciones pero no con el objetivo de mejorar/controlar la colaboración, sino las conversaciones en un entorno donde se puede colaborar (Goodman et al., 2005), o algunos aspectos de la colaboración, como la lectura de un mensaje de un compañero o acudir a un compañero para resolver una duda, pero sin analizar la colaboración (Santos y Boticario, 2008).

Éste es un punto abierto y esta tesis se ha decantado por la herramienta metacognitiva debido a que una herramienta de ese tipo es también una herramienta de autorregulación.

### **La autorregulación como estrategia de aprendizaje**

La autorregulación se puede definir como el proceso de aprendizaje consistente en el desarrollo de un conjunto de conductas constructivas que afectan al aprendizaje. Estos procesos se han planificado y adaptado para dar soporte a la búsqueda de las metas personales en un entorno de aprendizaje cambiante. Se pueden manejar otras definiciones como medio de tener un meta-conocimiento del conocimiento, o medio de regular el propio comportamiento. No hay un acuerdo definitivo.

Steffens (2001) establece en su estudio que las herramientas de autorregulación en entornos por ordenador mejoran el aprendizaje, en comparación con el mismo entorno sin dichas herramientas. Además de esto, Chang (2005) ha estudiado la relación de la autorregulación en entornos web respecto a la motivación, dando como respuesta que en entornos educativos donde se promueve y potencia el uso de la autorregulación, la motivación del alumno se incrementa.

Se ha descubierto que ayudar al alumno en el uso de sus habilidades de autorregulación mejora su aprendizaje (Pintrich, 2000) e incluso se ha propuesto caracterizar la autorregulación con tres características (Zimmerman, 1990):

- auto-observación (seguimiento de las actividades propias),
- auto-juicio (auto-evaluación de su rendimiento) y
- auto-reacciones (reacciones a los resultados del rendimiento).

También Steffens (2001) establece que la autorregulación es más que la regulación de las propias actividades cognitivas (metacognición), ya que también involucra aspectos motivacionales y emocionales. Es evidente, según el libro de "libro de mano de la autorregulación" (Boekaerts et al., 2000), que la autorregulación no sólo se emplea en la monitorización de los procesos de aprendizaje de uno mismo. También juega un rol importante en la gestión de las actividades sociales.

Dentro del ámbito que nos interesa, la educación vía ordenador, una herramienta de autorregulación es aquella que ofrece: una monitorización sobre el usuario, o pide que el usuario la haga; ayudas para un juicio sobre lo que el usuario ha realizado o solicita dicho juicio; e informa sobre las reacciones del usuario.

Según lo anteriormente expuesto, cualquier tipo de herramienta según Soller et al. (2005) podría ser una herramienta de autorregulación, mientras esté bien diseñada. Está claro que la herramienta metacognitiva no necesita un diseño especial ya que muestran una monitorización del usuario y un juicio sobre él. Si, además, la información que muestra cambia de forma regular y frecuente, el sistema mostrará las reacciones del usuario también (Zimmerman, 1990). Se ha dicho que las habilidades sociales (Boekaerts et al., 2000) juegan un rol importante en la autorregulación. Esto justifica el uso de herramientas metacognitivas para la mejora de la colaboración en un entorno de aprendizaje colaborativo.

En este apartado también introduciremos otro tipo de herramientas, que están adquiriendo una creciente importancia, y cuyo objetivo es el fomento del aprendizaje dando soporte al escrutinio o indagación de datos por parte del usuario al que los datos se refieren. Esta idea se denomina escrutabilidad (Kay, 1999). Bajo este enfoque una herramienta escrutable muestra un modelo sobre el usuario y que el usuario tiene la capacidad de gestionarlo. Una herramienta escrutable puede ser perfectamente una herramienta de monitorización o metacognitiva, con la diferencia de que muestra la información y "debe permitir" que el usuario la gestione. Se deduce que una

herramienta escrutable no tiene ningún impedimento para no convertirse, además, en una herramienta de autorregulación.

### **Realizar seguimiento de la colaboración**

El seguimiento o la monitorización de las interacciones del usuario es algo inherente al uso de los ordenadores, ya que estos tienen la facilidad de poder guardar información de lo realizado por los usuario, aunque en formatos no pensados para la mejora del aprendizaje. En este apartado se tendrán en cuenta investigaciones que han hecho un esfuerzo en monitorizar las interacciones de un alumno en un curso web colaborativo. Aquí matizamos monitorizar, en el sentido de herramienta de monitorización que establece Soler et al. (2005). Monitorizar es mostrar al usuario un modelo sobre sí mismo, sus interacciones, etc. Hacer seguimiento es identificar las interacciones de un usuario, que es lo que se hace en un fichero log de un servidor web. A partir de un seguimiento de las interacciones de un usuario se puede extraer un modelo de las interacciones, el cual se mostrará en una herramienta de monitorización.

Santos et al. (2004) y Gaudioso et al. (2003) investigaron en un marco de colaboración llamado Marco lógico, implementado en algunos cursos virtuales de la UNED. El Marco Lógico es una metodología utilizada por las Agencias de Cooperación Internacional para gestión de proyectos, ya que ayuda a aclarar los objetivos de cualquier proyecto al facilitar la identificación de las relaciones de causalidad previstas y determinar los riesgos que podrían impedir el logro de los objetivos del mismo. Con él se consigue implicar a las partes interesadas para que intervengan en el proceso de planificación y seguimiento. De estas investigaciones se obtuvieron la conclusión de que la monitorización no es suficiente en colaboración, hace falta suministrar más información sobre el proceso de colaboración para que el alumno tenga un punto de apoyo en el juicio sobre su propio trabajo.

Podemos diferenciar sistemas que, después de hacer un seguimiento de las interacciones de los usuarios, construyen una serie de indicadores relacionados con dichas interacciones pero únicamente muestran datos relacionados con las interacciones de los alumnos (Collazos et al., 2007; Hong, 2001). Sin embargo, estas investigaciones no tienen como objetivo la mejora del aprendizaje mientras la colaboración se llevaba a cabo, sino la identificación de los indicadores como útiles

en la monitorización de la colaboración. Otras investigaciones construyen un modelo de las interacciones de los usuarios y lo muestran en una herramienta de monitorización (Daradoumis et al., 2006; Martínez et al., 2006; Bratitsis et al., 2008). Otras hacen seguimiento de las interacciones y una vez hecho el seguimiento, infieren juicios sobre la colaboración de los estudiantes (Redondo et al., 2003; Talavera & Gaudioso, 2004; Perera et al. 2007). Sin embargo, ninguna de las anteriores han investigado la utilidad de los juicios, es decir, no han presentado la información a los estudiantes durante la experiencia educativa. Por tanto, son herramientas metacognitivas pero que todavía no han sido utilizadas. En este apartado nos centramos en los sistemas que han propuesto herramientas para la mejora del aprendizaje colaborativo, las cuales sólo han sido las que han construido herramientas de monitorización (Daradoumis et al., 2006; Martínez et al., 2006; Bratitsis et al., 2008).

### **Investigaciones**

Ha habido investigaciones donde han propuesto una serie de indicadores cuyo objetivo ha sido desvelar la colaboración realizada tanto de los equipos como de los participantes o miembros de los equipos de colaboración. Como ejemplo de estas investigaciones citamos (Collazos et al. 2007).

**Collazos et al. (2007)** han presentado un conjunto de indicadores y herramientas de software que les permitió experimentar en la evaluación del trabajo de colaboración, y, en particular, para estudiar la colaboración en los procesos que ocurren durante el aprendizaje en colaboración. Concluyen que los indicadores proporcionan una dirección a seguir en la comprensión y evaluación del proceso de CSCL. Afirman que el sistema basado en indicadores puede utilizarse para detectar las debilidades del grupo en el proceso de aprendizaje colaborativo. El objetivo es proporcionar a los tutores dicha información pero no proponen ningún tipo de herramienta definitiva.

Otras investigaciones se centran más en la monitorización y por ello trabajan más los datos antes de enseñarlos. **Daradoumis et al. (2006)** y **Martínez et al. (2006)** muestran los datos recogidos de los alumnos y, principalmente, mediante un análisis de redes sociales realizado de forma automática y apoyado por un nivel de análisis cualitativo y otro de análisis cuantitativo, es decir, los indicadores comentados



anteriormente. El acercamiento en ambas investigaciones es el mismo, pero el contexto de aplicación distinto. Evalúan el acercamiento de análisis en distintos casos de estudio tales como entornos de trabajo colaborativo a distancia o no, o un entorno de comunicación síncrona. No evalúan la herramienta como medio de mejorar la colaboración sino como medio de análisis de la colaboración.

**Bratitsis et al. (2008)** presentan una aplicación de análisis de las interacciones basado en ordenador que monitoriza a los usuarios de una red científica. Utilizan un conjunto de indicadores obtenidos de las interacciones registradas en los ficheros log y base de datos, como fuente de la información. Una vez obtenidos estos indicadores, los presentan en la herramienta con el objetivo de dar mayor control a los usuarios en la colaboración. De esta forma la herramienta de monitorización puede actualizarse según las interacciones se realizan.

Aunque el sistema **DIAS** (Bratitsis, 2009) no se centra en el análisis de la colaboración, se ha incluido esta investigación ya que monitoriza las interacciones asíncronas en un entorno colaborativo. La herramienta que proponen muestra visualmente gráficos que son combinaciones de indicadores de las interacciones en los foros y de la relación entre usuarios, mediante un análisis de redes sociales. Aunque no monitorizan la colaboración, sí monitorizan los foros, que es un medio apropiado para colaborar. Además, la herramienta que presentan es un ejemplo de interés de seguimiento de las interacciones.

Para terminar este análisis, también hay que tener en cuenta las investigaciones cuyo objetivo ha sido inferir conclusiones sobre la colaboración (Redondo et al., 2003; Talavera y Gaudioso, 2004; Perera et al., 2007). Estas investigaciones pueden mostrar, por tanto, no sólo los datos de donde obtienen el análisis, sino también conclusiones sobre esos datos, aunque no establecen herramientas que realicen acciones según los juicios inferidos. Uno de los objetivos de (Redondo et al., 2003) y (Perera et al., 2007) sí ha sido presentar los resultados a los tutores, mientras que en (Talavera y Gaudioso, 2004) se ha quedado en propuesta de uso de las inferencias. Hay que observar que (Perera et al., 2007) proponen mostrar los resultados de los análisis a los alumnos, pero que no lo han hecho en el curso que han investigado. Es decir, en potencia las investigaciones anteriores pueden crear una herramienta metacognitiva para ayudar al proceso de colaboración de los estudiantes, pero todavía no lo han hecho ni han analizado su aplicación.

## **Escrutabilidad**

El diccionario de la Real Academia Española dice de escrutar “indagar, examinar cuidadosamente, explorar; reconocer y computar los votos que para elecciones u otros actos análogos se han dado secretamente por medio de bolas, papeletas o en otra forma”. Quedémonos con el primer significado que es el que coincide con las investigaciones en este campo. De esto, deducimos que un sistema escrutable es aquel en el que es posible explorar, examinándolo y así deducir o simplemente darse cuenta de algo que no se conocía. En esta sección se justificará que una herramienta escrutable tiene las características de una herramienta de autorregulación.

El término escrutabilidad (*scrutability* en inglés) se empezó a utilizar ya a finales de la última década del siglo XX por Judy Kay, quien ha seguido explorando la escrutabilidad en sistemas de modelado de usuarios desde entonces, cuyas bases se exponen en 1999 (Kay, 1999). En este planteamiento se propone construir modelos reusables y escrutables. Escrutabilidad significa que el modelo pueda ser “escudriñado” por el alumno para determinar qué sabe el modelo sobre él y cómo opera el proceso de modelado. La escrutabilidad es importante, en este sentido, al proveer acceso a la información personal a los propios usuarios, a los cuales se refiere la información. Esto hace posible que el usuario corrija y controle el modelo.

La escrutabilidad es una característica que también puede ser utilizada en el modelado de los alumnos (Bull y Kay, 2008). Estas autoras establecieron lo que es un modelo abierto del alumno. Esta clase de modelo tiene como objetivo ser un medio para que el alumno pueda conocer lo que el sistema sabe de él con la característica de escrutabilidad, por lo que el modelo puede ser escudriñado, explorado o indagado por el dueño de la información que está representada. El modelo abierto del alumno tiene que contener información que el usuario entienda, sea de su confianza, sea responsable de ella y pueda utilizar y gestionar de manera eficiente para su aprendizaje.

Un modelo abierto del alumno posee características de las herramientas de autorregulación como la auto-observación ya que el modelo se refiere al usuario, el auto-juicio ya que lo hace responsable al usuario del modelo, y al cambiar con las interacciones del usuario, también lo hace de sus propias reacciones. Teóricamente el modelo abierto de los alumnos refuerza las características de autorregulación por lo que es una técnica deseable en el entorno educativo.

## Investigaciones

Hasta la fecha pocas investigaciones se han centrado en el estudio de la escrutabilidad y ninguna ha tenido como principal objetivo la colaboración. Sin embargo, comentaremos algunos sistemas que han utilizado la escrutabilidad para facilitar la comprensión de esta técnica y su utilidad en el campo de la colaboración.

(Kay y Kummerfeld, 2006) han diseñado y construido **Personis**, el cual es una arquitectura de modelado y presentación con características de escrutabilidad. Personis garantiza servicios personalizados donde el usuario puede mantener el control en todos los niveles: lo que sucede en su modelo de usuario, lo que está disponible para los distintos servicios y la forma en que el modelo es administrado y mantenido.

(Assad et al., 2007) han desarrollado un servicio, **MyPlace Locator**, cuyo objetivo ha sido mostrar al usuario dónde él se encuentra y el resto de personas de la investigación. MyPlace Locator utiliza un dispositivo bluetooth para localizar a cada usuario de la investigación. Además, MyPlace Locator informa al usuario del lugar en que éste se encuentra y el resto de personas que poseen el localizador. Al usuario final se le muestra un modelo con lo que sabe el sistema de él, y se le da la capacidad de gestionar dicho modelo según su propio criterio. Por lo tanto, el usuario tiene la capacidad de permitir que algún otro miembro de la investigación sepa donde está él, que todos lo sepan, o ninguno. Este sistema es un ejemplo de sistema escrutable fuera del dominio del modelado en educación.

(Kay y Li, 2006) presentaron **KrAssess**, un sistema de educación para aprender programación, que tiene por objeto facilitar la autoevaluación de los estudiantes y promover la ponderación del alumno a través de modelos escrutables. KrAssess promueve el aprendizaje ayudando a los estudiantes a aprender a "ver" las soluciones de acuerdo con los criterios evaluados por su profesor.

(Tanimoto, 2007) propone un sistema llamado **INFACT** que facilita la creación y captura de las evidencias del aprendizaje del alumno mientras éste participa en la solución de problemas y en la construcción de actividades. El sistema enriquece los ficheros log mediante técnicas de minería de datos. Mostrando el fichero log enriquecido a los estudiantes, afirman que ellos pueden mejorar el entendimiento de

cómo han sido evaluados, por lo que especialmente se entiende que el sistema INFACT modelos abiertos en el sentido definido por (Bull y Kay, 2008).

## **Privacidad**

En la era Web 2.0, la privacidad cobra cada vez más importancia (Parsell, 2008), debido a la facilidad de perderla. La personalización, entendida como las acciones de un sistemas para adaptarse a los requerimientos de los usuarios, es un paradigma a seguir ya que el éxito o el fracaso de un entorno se puede deber a las recomendaciones personalizadas que suministre (Adomavicius y Tuzhilin, 2005).

La privacidad y la personalización se sienten como características opuestas. Un sistema con mayor privacidad no puede establecer tantas personalizaciones como otro en donde se ignore la privacidad (Mark y Kobsa, 2005). Sin embargo, la unión de ambos términos es un objetivo deseable debido a la falta de privacidad de algunos entornos Web 2.0 y a las futuras consecuencias que puede acarrear (Kobsa, 2007).

En el marco de trabajo que este capítulo establece, la privacidad no ha sido un característica importante en gran parte de las investigaciones comentadas. Se debe esto al hecho de que los resultados de los análisis y modelados han sido dirigidos al tutor de la experiencia de aprendizaje colaborativo (Hong, 2001; Redondo et al., 2003; Collazos et al., 2007). En los sistemas de monitorización descritos (Martínez et al., 2006; Daradoumis et al., 2006; Bratitsis et al., 2008; Bratitsis, 2009) mostraban un análisis de las interacciones de los estudiantes, y datos sobre ellos, los cuales eran solicitados mediante encuestas. Si al estudiante se le permitía responder o no según a que pregunta, él sería el responsable de su propia privacidad.

La privacidad es una característica importante en las herramientas escrutables. El usuario tiene el poder de gestionar el modelo que se ha hecho sobre él, por lo que el modelo será tan privado como el usuario lo desee.

## **Análisis crítico**

Como se ha ido diciendo a lo largo de este capítulo, en el análisis y modelado de la colaboración hay una falta de estándares y metodología provocada por, entre otros, un escaso número de investigaciones en este campo y la falta de estudios comparativos entre dichas investigaciones (Strijbos y Fischer, 2007; Perera et al., 2007; Bratitsis y

Dimitracopoulou, 2006). Tampoco hay un acuerdo a la hora de encontrar una estrategia pedagógica que mejore, después de realizado el análisis y el modelado, el proceso de colaboración (Barkley et al., 2004).

Si acudimos a Soler et al. (2005) se señala que las herramientas posibles en un entorno colaborativo son las de monitorización, metacognitivas y de guiado, pero si también tenemos en cuenta las condiciones de Johnson y Johnson (2004), después del análisis y modelado de la colaboración, el sistemas debería suplir con un procesamiento frecuente y regular del trabajo colaborativo de los equipos. Por la conveniencia de realizar un juicio o inferencia sobre la colaboración, el tipo de herramienta más apropiado sería la herramienta metacognitiva o de guiado.

Hemos indicado que los entornos educativos vía web u ordenador son un campo que acoge adecuadamente cualquier intento de promover la autorregulación, y hemos señalado que las habilidades sociales también constituyen un dominio en donde la autorregulación cobra importancia (Steffens, 2001). Además, se han descrito las ventajas y características de la autorregulación (Boekaerts et al., 2000; Zimmerman, 1990). Concluimos que una herramienta de autorregulación sobre la colaboración es posible y apropiada para la mejora del aprendizaje colaborativo.

De las características de la autorregulación se intuye la relación cercana con una herramienta metacognitiva, ya que ésta propone juicios sobre las acciones del usuario lo cual promueve el auto-juicio del usuario y el control sobre sus reacciones. En este sentido, una herramienta de guiado, la cual propone reacciones al juicio sobre el usuario, tiene mayores dificultades para cumplir las condiciones de una herramienta de autorregulación, salvo que también suministre la información que una herramienta metacognitiva mostraría.

De las investigaciones estudiadas se observan dos puntos abiertos, por un lado, el análisis de la colaboración para establecer valoraciones o juicios (Redondo et al., 2003; Talavera & Gaudio, 2004; Perera et al., 2007), y por el otro, el medio de mostrar el análisis para la mejora de los procesos de colaboración (Daradoumis et al., 2006; Martínez et al., 2006; Bratitsis et al., 2008). Las investigaciones que proponen una solución a uno de los puntos abiertos, no la proponen en el otro.

Llegado a este punto, nos hemos decantado por la creación de una herramienta metacognitiva de autorregulación para mejorar el aprendizaje colaborativo, después de analizada y modelada la colaboración. Consultando la literatura hemos visto que

una característica más podría ser añadida a las deseables para la herramienta metacognitiva de autorregulación. Esta característica es la escrutabilidad (Kay, 1999).

La técnica de escrutabilidad pensamos que es muy recomendable en entornos educativos. Además de monitorizar al alumno, éste tiene la responsabilidad de gestionarlo. La escrutabilidad, por tanto, aplica la teoría de intercambio social (Hummel et al., 2005), la cual se puede simplificar diciendo que a mayor información y responsabilidad mayor la motivación. Y un alumno motivado es un alumno con facilidad para el aprendizaje (Burlison, 2005; Boticario y Gaudioso, 2000a). Un modelo del alumno escrutable, como señalan (Bull y Kay, 2008), tiene que ser entendible por los usuarios del modelo, cierto, fácil de gestionar, etc.

También ha sido considerada la privacidad en entornos semejantes. Se ha comentado la conveniencia de su consideración en la actualidad y su unión con los sistemas de personalización (Kobsa, 2007). Se ha advertido, además, de la relación positiva de las herramientas escrutables con la privacidad.

Para cubrir los objetivos de la tesis se ha propuesto la creación de una herramienta metacognitiva y para evaluar la hipótesis sobre las ventajas de utilizar la estrategia escrutable, se han propuesto dos formas de mostrar la información. Una ha sido mostrar la información recogida, y la otra permite además gestionar parte de dicha información, desde un punto de vista escrutable. Para considerar la privacidad de los estudiantes que han utilizado ambos tipos de herramientas, la información obtenida del estudiante directamente se ha realizado mediante encuestas de respuestas voluntarias.

## **Conclusiones generales**

En este capítulo hemos descrito el marco de trabajo del ciclo de la mejora del aprendizaje colaborativo en un entorno web educativo. El ciclo podría dividirse en: el diseño del entorno de aprendizaje colaborativo, análisis de la colaboración, modelado de la colaboración y herramientas de mejora de la colaboración.

La ED ha dado un gran paso hacia delante en la era de Internet. Los contenidos que un alumno aprende pueden estar accesibles a través de la web y éste puede consultar dudas a los profesores en el momento que desee. La posible falta de interacciones sociales se puede resolver creando entornos virtuales de colaboración. Sin embargo, el

diseño del entorno de aprendizaje colaborativo tiene que cumplir una serie de condiciones (Johnson y Johnson, 2004). En esta tesis consideramos la quinta condición (procesamiento regular y frecuente del trabajo de los equipos al colaborar) como de mayor dificultad en un entorno web de ED. Los alumnos de la ED se caracterizan, entre otras cosas, por su diversidad (Boticario y Gaudioso, 2000b), cuya consecuencia es la necesidad de tener el control sobre el proceso de aprendizaje (Gaudioso et al., 2003). En un entorno colaborativo esto significa que es conveniente que el estudiante tenga el control sobre el proceso de colaboración. Si a esto le sumamos que en un entorno de ED basado en web la participación puede ser muy numerosa, resulta que el tutor o profesor de la experiencia de aprendizaje colaborativo tendrá grandes dificultades en hacer el procesamiento regular y frecuente de la colaboración. En la UNED el número de alumnos en una asignatura popular en carreras con un gran número de estudiantes supera los 1000 alumnos sin problemas. Por ejemplo, más de 400 alumnos han cursado la asignatura de IA-IC durante el periodo de esta investigación en los cursos 2006/2007, 2007/2008 y 2008/2009.

A la hora de establecer un marco de trabajo para la mejora del aprendizaje colaborativo, se ha estudiado la literatura que aborda el problema o alguna de sus partes. Con lo dicho anteriormente, en el diseño del entorno tiene que considerarse el análisis del comportamiento de los participantes en un equipo de manera regular y frecuente con la menor intervención de los tutores, debido al contexto de la ED. Se han establecido un conjunto de características para ir caracterizando las investigaciones. Estas caracterizaciones se puede dividir en: cómo se obtienen los datos (preproceso); qué método de análisis se utiliza (inferencia); qué tipo de modelado se aplica (modelado); qué tipo de herramienta se utiliza (herramienta). En la Tabla 1 se resumen las características de las investigaciones utilizadas para enmarcar el trabajo de esta tesis.

**Tabla 1: Caracterización de investigaciones en relación al análisis de la colaboración**

Investigaciones	Preproceso			Análisis					Modelado		Herramienta			
	Cualitativo	Cuantitativo	Mixto	Experto	Estadístico	Análisis de interacciones	Aprendizaje automático	Ninguno	Modelo	Indicadores	Monitorización	Metacognitiva	Guiado	Escurtable
Baeza-Yates y Pino (2006)										X				
Baghaei y Mitrovic (2007)									X					
Baldiris et al. (2007)										X				
Barros et al. (2002)									X					
Bayón et al. (2007)										X				
Bratitsis et al. (2008)		X				X				X	X			
Bratitsis (2009)		X				X				X	X			
Collazos et al. (2007)			X					X		X	X			
Daradoumis et al. (2006)			X			X				X	X			
Durán (2006)									X					
Goodman et al. (2005)													X	
Hong (2001)		X			X					X				
Kahrimanis et al. (2009)	X			X						X				
Martínez et al. (2003)									X					
Martínez et al. (2006)		X				X				X	X			
Meier et al. (2007)	X			X						X				
Park y Hyun (2006)										X				
Perera et al. (2007)			X				X			X		P		P
Redondo et al. (2003)		X					X		X			X		
Santos y Boticario (2004)													X	
Soller (2001)										X				
Talavera y Gaudioso (2004)		X					X			X		P		
Vidou et al. (2006)									X					

La “X” indica una de las características de la investigación, mientras que “P” indica que dicha característica se ha propuesto pero no se ha evaluado. Se puede observar que algunas investigaciones sólo están caracterizadas por una o dos características. Estas investigaciones no se proponen completar el ciclo del análisis de la colaboración pero sirven de ejemplo de la caracterización empleada. Un caso aparte es el de la investigación Bratitsis (2009), ya que su objetivo no ha sido el análisis de la colaboración sino el de la comunicación asincrónica, pero al estar tan relacionada la



comunicación asíncrona en un entorno e-learning con la colaboración, ha sido incluida en la tabla.

Debido a la falta de metodología, estándares y estudios empíricos (Strijbos y Fischer, 2007; Perera et al., 2007; Bratitsis y Dimitracopoulou, 2006), hay varios puntos abiertos en el análisis de la colaboración, los cuales también se observan perfectamente en la tabla anterior. El primero de ellos es la fuente de datos. Debido al contexto educativo y a los objetivos de la investigación, se ha pensado que lo conveniente era utilizar como fuente de datos información cuantitativa sobre las interacciones de los estudiantes relacionadas con la colaboración. De esta forma se favorece la obtención regular y frecuente de los datos, y la menor intervención posible de tutores y estudiantes.

Otro punto abierto en el análisis de la colaboración es el método de inferencia. Ya que el análisis se tiene que realizar de forma regular y frecuente con la menor intervención posible de tutores o expertos, se hace necesario utilizar técnicas de aprendizaje automático. Los objetivos sugieren dicha deducción ya que es un método de análisis que infiere información que estaba oculta en los datos, que se puede utilizar cuando los datos sean obtenidos, y puede aplicarse en gran cantidad de entornos (Rusell y Norvig, 1996).

El análisis de la colaboración ha sido la parte principal del modelado del alumno, pero no la única. Modelar al alumno es un reto fundamental para la personalización de entornos educativos (McCalla et al., 2000). Una estrategia muy interesante desde el punto de vista pedagógico es el modelo abierto (Bull y Kay, 2008), donde el resultado del modelado es mostrado al estudiante y éste lo puede gestionar. Por lo tanto el modelo obtenido del proceso de modelado debe ser entendible por el estudiante.

En entornos colaborativos tanto la información sobre el proceso como sobre el contexto de la colaboración son de interés (Muehlenbrock, 2005). Se concluye que la información útil para que el estudiante pueda colaborar es la debida al contexto de la colaboración, al proceso de la colaboración, y a la inferida por el análisis, ya que establece juicios sobre la colaboración del estudiante.

Para cerrar el ciclo de la mejora del aprendizaje colaborativo, en este capítulo se han comentado distintas herramientas de uso pedagógico. Nos hemos centrado en los beneficios de las herramientas de autorregulación en el contexto educativo y se han relacionado con la colaboración, ya que la autorregulación interviene en procesos

sociales (Boekaerts et al., 2000). Se han establecido las características de las herramientas de autorregulación y se han comparado con las herramientas metacognitivas que un entorno colaborativo puede suministrar. Nos hemos decantado en esta investigación por una herramienta metacognitiva de autorregulación.

Siguiendo el hilo de los modelos abiertos a los estudiantes, se ha estudiado la escrutabilidad (Kay, 1999). Al definir la escrutabilidad se ha detectado la afinidad de la escrutabilidad con las herramientas metacognitivas y con la autorregulación. Se ha señalado las ventajas teóricas de la escrutabilidad para aumentar la responsabilidad y, por tanto, la motivación y el aprendizaje activo (Hummel et al., 2005; Burtleson, 2005; Boticario y Gaudio, 2000a).

En resumen, se concluye que un entorno de aprendizaje colaborativo vía web dentro del paradigma de la ED necesita que el sistema realice análisis de la colaboración de forma regular y frecuente. Gracias a los análisis se puede modelar al estudiante. Una vez modelado el resultado, es decir, el modelo, se puede convertir en una herramienta de metacognitiva de autorregulación con la que proveer al estudiante de un medio de ayuda a la gestión del proceso de colaboración. Esta herramienta puede añadir las ventajas de la escrutabilidad para aumentar la responsabilidad del alumno en todo el proceso de colaboración.

## PARTE II: DESARROLLO

En esta parte se describe el desarrollo de la investigación. Dado que la mejora del aprendizaje en un entorno colaborativo es un objetivo esencial de esta tesis, en el desarrollo se describe toda la experiencia educativa sin olvidar el modelo educativo desarrollado. Forman parte del desarrollo los análisis, métodos de inferencia, modelos y estrategias aplicadas para alcanzar los objetivos de la investigación, estando estos elementos descritos en el capítulo III.

# **Capítulo III: Marco conceptual, modelos y desarrollo del planteamiento**

La investigación se ha centrado en la mejora del aprendizaje colaborativo y en este capítulo se describe el marco conceptual. Hay que considerar el contexto educativo del que la experiencia de aprendizaje colaborativo ha formado parte. Hay que describir, por tanto, el modelo de la UNED, la asignatura y el tipo de alumnos que han participado en la experiencia.

En el Capítulo II se estudió trabajos relacionados para dar un marco a la investigación que se ha desarrollado. Se comenzó hablando de los entornos colaborativos como experiencias de aprendizaje y de sus características comunes. En este capítulo se explicará el modelo seguido en la experiencia de aprendizaje colaborativo. También se enmarcó la investigación de la colaboración y aquí se presentarán los pasos seguidos para el análisis de las interacciones de los equipos colaborativos con el objetivo de inferir evaluaciones sobre la colaboración. El modelado del alumno en el entorno colaborativo también ha sido una materia de estudio del Capítulo II y ahora presentaremos una propuesta dirigida al uso posterior de la información por los propios estudiantes. Se comentaron las ventajas de esta perspectiva al compararla con la autorregulación y se describieron herramientas que en otras investigaciones han sido utilizadas para el beneficio de los estudiantes. Este acercamiento es el seguido en esta tesis y el que será explicado.

El objetivo del marco ha sido el cierre del ciclo de la investigación de la colaboración por lo que se desarrolla en este capítulo el modelo educativo de la experiencia educativa centrándose en la parte de la colaboración, su análisis, inferencia, modelado y uso de los resultados como estrategia pedagógica para el beneficio del estudiante y de su aprendizaje.

Para finalizar, hay que evaluar la experiencia de aprendizaje colaborativo. Se explicará el método de evaluación propuesto centrado en el estudiante.

## **Contexto educativo**

En el contexto de esta investigación se hará una breve descripción del contexto académico, es decir, de la educación a distancia, del modelo educativo de la UNED, de la asignatura de Inteligencia artificial e ingeniería basada en conocimiento (IA-IC), y de los alumnos que la cursan, dado que parte de ellos se han convertido en los sujetos y objetos de esta investigación. Se hará hincapié en las características de los estudiantes que los hacen diferentes a los estudiantes de las universidades presenciales y en la relación de estos condicionamientos con el planteamiento del modelo realizado en esta investigación.

Este apartado terminará con un análisis crítico del contexto justificando el planteamiento de la investigación, los objetivos docentes y las técnicas de minería de datos utilizadas para su obtención.

## **Introducción a la educación a distancia**

La UNED es una universidad que oferta estudios a distancia. Fue inaugurada en 1972 y posee su propio modelo educativo de educación a distancia (ED). En este apartado se definirá la ED dando un breve apunte histórico relacionado con las tecnologías usadas y se hablará del modelo educativo de la UNED.

La ED es el proceso de creación de experiencias educativas de valor que se adapten a las necesidades de los alumnos fuera de las aulas (Holmberg, 2005). Aquí aula tiene un sentido literal y se refiere al lugar físico donde se reúnen los alumnos y los docentes para que los primeros aprendan y los segundos enseñen. La definición es suficientemente flexible para englobar a todas las experiencias de ED a las que se les ha dado ese nombre a lo largo de la historia.

Aunque hay cierta controversia, la ED comenzó en la época post-industrial al aparecer los primeros cursos que se impartían por correo (Peters, 1993). Aquí aparece ya una tecnología y una forma de comunicación. Realmente, puede ser esto lo que caracterice más a la ED. En la ED la comunicación vertical (docente – alumno) como la horizontal (alumno – alumno) es realizada principalmente utilizando algún medio tecnológico (Perraton, 1988). La comunicación cara a cara no es el medio

fundamental de entablar relación entre los docentes y los alumnos, y estos entre sí. Esta causa tiene como efecto la ruptura de las barreras o condiciones para la realización de una charla o clase presencial. La ED asume dichas rupturas como características y nos fijaremos en dos: la ruptura espacial y temporal. La comunicación, por tanto, y por definición, la transmisión de conocimiento, el aprendizaje, no tiene que ocurrir cuando el docente y el alumno se encuentre en el mismo lugar al mismo tiempo.

Se puede abordar la historia de la ED enfatizando las tecnologías utilizadas de comunicación y los medios docentes utilizados (García Aretio, 2009; Sutherland et al., 2008):

- En un primer momento, la tecnología utilizada fue el correo ordinario entre docente y alumno sin adaptación pedagógica en un principio.
- Al aparecer nuevas tecnologías (nuevos medios de comunicación, teléfono, etc.) se fueron empleando ampliándose el uso de las nuevas tecnologías a los contenidos docentes.
- Instituciones universitarias de ED aparecieron y utilizaron los medios de comunicación para la transmisión del conocimiento y comunicación. El diseño, producción y generación de materiales didácticos son objetivos primordiales, dejando de lado la comunicación horizontal.
- La llegada del ordenador o enseñanza asistida por ordenador supuso una adaptación de la ED al uso frecuente de los ordenadores y los sistemas multimedia.
- El siguiente paso es Internet. Se establece el modelo de aprendizaje flexible y se utiliza los sistemas multimedia interactivos, la comunicación mediada por ordenador, y, en síntesis, la comunicación educativa a través de Internet.
- En la actualidad los sistemas multimedia, el ordenador e Internet pueden llevar la ED a todos los ciudadanos. La instrucción se ha democratizado con el paso del tiempo, y generalizado. A todos y en cualquier momento es capaz de llegar. El paradigma de la educación continuada es un hecho y los sistemas educativos buscan adaptarse al estudiante como medio de mejora del aprendizaje.

Debido a las características comentadas casi cualquier persona puede realizar estudios a distancia sin salir de su casa. En la actual sociedad de la información cualquier persona tiene la opción de instruirse ya que es posible manteniendo las

responsabilidades de trabajo y familia para ser un miembro útil de la sociedad, algo que es difícil de hacer compatible en la educación presencial tradicional (Finn, 1988). El paradigma de la educación continuada (o en inglés Life-Long Learning) es un objetivo plausible para una gran mayoría de los miembros de una sociedad moderna (Field, 2006). La ED es una estrategia educativa apta para los momentos actuales y realiza esfuerzos para adaptarse tanto a los nuevos tiempos como a los estudiantes. Estos alumnos se caracterizan por su diversidad ya que no forman grupos homogéneos de individuos al no tener en común la edad, un lugar de residencia, parecidos currículos académicos (Boticario y Gaudioso, 2000). Un ejemplo son los alumnos de la UNED, ya que forman un subconjunto significativo de la sociedad española. La motivación de estos alumnos es variada, por ejemplo, hay alumnos que estudian por enriquecimiento o reto personal, o por alcanzar una promoción profesional, o alumnos que buscan una formación permanente y la necesaria actualización de sus conocimientos (Fernández y Montalvo, 2001). En otras palabras, representan el tipo de alumno que vendrá al generalizarse el modelo de educación continuada (Comision of the European Communities, 2006). Modelo educativo que es uno de los objetivos de la Unión Europea. La ED se está adaptando a esta clase de estudiantes que buscan, cada vez en mayor número, la educación continuada o formación permanente.

Dejamos claro que unas singularidades de la ED son: el uso de tecnologías para establecer la comunicación vertical y horizontal, y el amplio espectro de tipos de personas que puede acceder. Estas singularidades han formado parte de la historia de la ED.

Es en el último punto donde se encuadra esta investigación. Desde las instituciones se solicita un esfuerzo para desarrollar entornos de e-learning centrados en el usuario y alumno (European Comision, 2008). La TIC abren la puerta pero para alcanzar la adaptación y el aprendizaje personalizado se necesita investigaciones multidisciplinares y estudios empíricos sobre distintas propuestas. Esta tesis aborda este problema y propone una metodología para alcanzar la adaptación intentando ser lo suficientemente general y poder ser extensible en otros entornos.

## **Modelo educativo de la UNED<sup>7</sup>**

El modelo educativo de la UNED está basado en la creencia de que la relación directa profesor/alumno no es completamente imprescindible, ya que es posible aprender en solitario determinados contenidos científicos, siempre que se use una metodología adecuada, que garantice la calidad y claridad de los mismos, y se disponga de los cauces tecnológicos adecuados para asegurar una comunicación fluida con el profesorado (tutores y profesores). Para lograr este objetivo se necesitan unos instrumentos pedagógicos de una calidad universitaria contrastada, que se materializan en el material didáctico.

El profesorado de la UNED cuenta para organizar sus enseñanzas con un sistema de comunicación de carácter multimedia que incluye materiales escritos, medios audiovisuales, apoyos informáticos y nuevos recursos tecnológicos.

De acuerdo con el principio anterior, cada equipo docente prepara el material didáctico de su asignatura y selecciona los medios a utilizar para su transmisión, ateniéndose a las normas generales que al respecto se fijan. La decisión adoptada queda reflejada en el plan docente anual del Departamento correspondiente.

La UNED se ha centrado en tres tecnologías para la transmisión de los contenidos educativos: la televisión, la radio y los cursos virtuales a través de internet. Con las tres tecnologías la UNED pretende cubrir las necesidades diversas de todos los alumnos.

Los Cursos Virtuales de la UNED, tanto de Enseñanzas Regladas (Primer, Segundo y Tercer Ciclo) como de Educación Permanente, tienen por objeto incluir recursos didácticos a través de las plataformas de teleformación de la Universidad, de gran utilidad para el seguimiento de las asignaturas. Estos cursos tienen muy presente el componente a distancia de nuestra universidad y, por ello, hacen especial hincapié en aportar al estudiante una ayuda para el seguimiento óptimo de sus cursos. Su consulta es gratuita y voluntaria y supone un valor añadido a la docencia a distancia tradicional y por tanto, no la sustituye.

Los Cursos Virtuales tienen una metodología reglada con el objetivo de adaptarse a las necesidades de los alumnos y donde se da un importante papel a la comunicación

---

<sup>7</sup> Ver <http://www.uned.es/webuned/areasgen/info/guia2006-07/regimeneducativo.htm>



y colaboración entre ellos. Se promueve el aprendizaje activo, colaborativo y en la resolución de problemas reales. Dentro de la sociedad actual, los cursos virtuales cobran gran importancia al ser la herramienta principal de la formación permanente. En estos entornos, la colaboración es implícita al ser un medio de comunicación entre todos los estudiantes.

La UNED presenta una metodología bien establecida y reglada para la realización de las pruebas de evaluación. Hay dos modalidades, pruebas de evaluación a distancias y presenciales. Las pruebas de evaluación a distancia se encuentra unidas a los cursos virtuales donde se ofrece a los estudiantes medios de autoevaluación.

El otro factor que diferencia sustancialmente la enseñanza libre de la enseñanza a distancia es la existencia de vías organizadas de comunicación profesorado/alumno y alumno/alumno. Las TIC abren un abanico de medios de comunicación reunidos en los Cursos Virtuales organizados por la UNED. De esta forma, la comunicación y colaboración entre profesorado y alumnado, y entre los propios alumnos se facilita para la mejora del aprendizaje.

## **La asignatura**

Los estudiantes con los que se ha realizado la investigación cursaban la asignatura de Inteligencia artificial e ingeniería basada en conocimiento (IA-IC). Es una asignatura de tipo troncal de 9 créditos del cuarto curso de Ingeniería de Informática en la UNED. La experiencia de aprendizaje colaborativo se ha centrado en una parte del temario de IA-IC por lo que se describe ésta en relación a dicha experiencia. A continuación se describirán los objetivos, competencias, metodología de la asignatura centrada en la relación con la experiencia.

### *Objetivos*

El objetivo principal de la materia de IA-IC es conseguir que los alumnos del primer curso del Segundo Ciclo de la Escuela Técnica Superior de Ingeniería Informática de la UNED sean capaces de adquirir los conocimientos teóricos y prácticos de temas avanzados relacionados con la Inteligencia Artificial así como aquellos relativos al desarrollo de los Sistemas Basados en el Conocimiento. En

esencia, se desea ofrecer una perspectiva global, científica, metodológica y práctica de la resolución de problemas de esta área de conocimiento.

### *Destrezas y competencias*

Respecto a la asignatura:

- Conocimiento de materias básicas y tecnologías en relación a la Inteligencia artificial, que le capacite para el aprendizaje de nuevos métodos y tecnologías, así como que le dote de una gran versatilidad para adaptarse a nuevas situaciones.
- Capacidad para diseñar, desarrollar, implementar, verificar y documentar soluciones software inteligentes, como es el trabajo que se solicita en la experiencia de aprendizaje colaborativo, centrada en el problema de planificación.

Generales:

- Capacidad de resolver problemas con iniciativa, toma de decisiones, creatividad, y de comunicar y transmitir conocimientos, habilidades y destrezas, comprendiendo la responsabilidad ética y profesional de la actividad del ingeniero técnico en informática.
- Competencias de gestión y planificación: Iniciativa y motivación. Planificación y organización (establecimiento de objetivos y prioridades, secuenciación y organización del tiempo de realización, etc.). Manejo adecuado del tiempo
- Aplicación de los conocimientos a la práctica. Resolución de problemas en entornos nuevos o poco conocidos. Pensamiento creativo. Razonamiento crítico. Toma de decisiones
- Competencias de expresión y comunicación (a través de distintos medios y con distinto tipo de interlocutores): Comunicación y expresión escrita.
- Trabajo en equipo desarrollando distinto tipo de funciones o roles.

Estas destrezas y competencias son las que la experiencia de aprendizaje colaborativo ha querido incentivar con un trabajo planificación en un entorno colaborativo.

## *Metodología docente*

Los contenidos teóricos se imparten a distancia o en la tutorías de los centros asociados, donde se imparte la asignatura, de acuerdo con las normas y estructuras soporte temático de la enseñanza en la UNED. Además, desde la página de la asignatura y, sobre todo, desde el espacio virtual los alumnos pueden encontrar toda la información necesaria y suficiente para estudiar la asignatura.

El material docente incluye información detallada sobre los contenidos de cada tema informando, además, de material docente complementario que el alumno podrá encontrar en las bibliotecas de la UNED o a través de Internet. El material docente se encuentra estructurado en el curso virtual de la asignatura.

El curso virtual es el principal medio de comunicación entre docentes y alumnos, y entre ellos mismos mediante unos foros estructurados. Un tutor del curso virtual atenderá, primero, las dudas y peticiones de los alumnos a través de los foros del curso virtual, dejando las preguntas de contenidos para el equipo docente. El equipo docente atiende, también, a los alumnos a través de llamadas telefónicas, correos electrónicos a la cuenta de la asignaturas o personales, pero promoviendo los foros del curso virtual como principal medio de contacto.

La asignatura es de carácter práctico y, para alcanzar los objetivos y competencias, se solicita de los alumnos trabajos prácticos. Uno de los trabajos prácticos es la base de la experiencia de aprendizaje colaborativo, la cual utiliza los medios virtuales para ofrecer un entorno de trabajo en equipo.

## *Estructura y contenidos teóricos*

Los contenidos de la asignatura se estructuran en seis temas: Introducción a la Inteligencia Artificial; Fundamentos de búsqueda y representación; Planificación clásica; Aspectos metodológicos del desarrollo de SBCs; Ingeniería ontológica; Sistemas multiagente. Estos contenidos introducen el área de la Inteligencia Artificial y distintas técnicas que aportan distintas aportaciones al desarrollo de sistemas informáticos.

Respecto a la experiencia educativa que se expone en esta tesis, la planificación tiene gran importancia ya que se ha propuesto un trabajo práctico sobre este tema de

forma colaborativa para alcanzar las competencias y destrezas solicitadas en la asignatura.

### *Evaluación*

La evaluación de los conocimientos enseñados en la asignatura se ha realizado con las pruebas de evaluación presenciales, obligatorias. En la primera prueba presencial se evalúan los contenidos de los tres primeros capítulos del temario y en la segunda prueba el resto de capítulos. Ambas pruebas parciales son eliminatorias de materia y de resultados compensables desde una nota mínima de 4 puntos.

Además de se han propuestos pruebas de evaluación a distancia. Una práctica obligatoria ha sido ofrecida a principios de curso. La práctica obligatoria ha consistido en la solución de un problema de planificación y ha sido la utilizada para probar un modelo pedagógico de aprendizaje colaborativo, que se explicará a continuación. La practica se evaluaba a los estudiantes que habían participado en la experiencia antes del examen del primer parcial.

El contenido docente que formaba parte de la práctica obligatoria ha sido la planificación. De este tema se preguntaba en el examen del primer parcial. Las evaluaciones que se han tenido en cuenta para evaluar el trabajo de los estudiantes que han participado en la experiencia de aprendizaje colaborativo han sido el resultado del trabajo práctico, y el resultado del examen del primer parcial.

## **Los alumnos**

Los alumnos son el centro de los procesos pedagógicos y para esta investigación es importante estudiar sus peculiaridades y características.

Los alumnos de la UNED no responden al estereotipo (Losada, 1996). Normalmente se forman en la universidad a la vez que trabajan, por lo que estudian lo que pueden y cuando pueden. El perfil actual del retrato robot podría ser: de mediana edad, con estudios universitarios medios, con ocupación laboral estable en cuadros intermedios y con inquietudes de progreso personal y/o profesional. Ellos son, y deben ser, responsables de su tiempo por lo que ellos tienen que gestionar su aprendizaje.

En la Tabla 2 se dan datos de ciertas características de los alumnos que han realizado la asignatura de IA-IC. Los datos han sido obtenidos de la encuesta inicial que se pasaba a los alumnos que querían realizar el trabajo práctico obligatorio de la forma colaborativa y de la que se hablará más adelante.

**Tabla 2: Características de los alumnos**

Curso	Nº de alumnos	Edad media	Varianza de la edad media	Porcentaje de hombre y mujeres	Porcentaje de alumnos con responsabilidades laborales	Porcentaje de alumnos con responsabilidades familiares
06/07	476	32	35,83	84,11/15,89		
07/08	428	33	39,64	80,21/19,79	92,79	31,96
08/09	412	34	46,30	78,64/21,36	94,18	45,63

En la Tabla 2 aparecen las características más llamativas de un estudiante universitario de la asignatura, pero que suelen ser comunes en un estudiante de ED<sup>8</sup>. Primero observamos el gran número de alumnos de la asignatura en cada curso. También se observa que la edad media del alumnado es alta en comparación a la de los alumnos de las universidades presenciales. Hay que advertir de la gran varianza de la edad media. Esto significa que la distribución de alumnos según la edad es muy elevada. Es posible encontrar alumnos desde los 20 años hasta los 56. Lo que es mayoritario son las responsabilidades laborales. Un altísimo porcentaje de alumnos trabajan a la vez que estudian. Además, un poco menos de la mitad tienen responsabilidades familiares. También se señala la población masculina y femenina. Los estudios de informática, como todas las ingenierías, atraen más a la población masculina (Sainz et al, 2009), pero la brecha mengua con el paso de los años.

Las características anteriores son muy difíciles de encontrar en el alumnado de una universidad presencial, pero relativamente fáciles en el alumnado de una universidad a distancia como la UNED (Boticario y Gaudioso, 2000a). Los alumnos de la ED forman un grupo heterogéneo en relación a la edad, objetivos, experiencia, currículo. Los alumnos de la asignatura de IA-IC son un ejemplo del alumnado de la ED.

El modelo educativo de la UNED está adaptado a esta clase de alumnos, al igual que la asignatura IA-IC. Estos alumnos conocen al principio del curso la información metodológica suficiente para saber lo que tienen que hacer para aprender los temarios de la asignatura. Esto les ayuda a programar su propio aprendizaje. Los plazos que se

<sup>8</sup> [http://portal.uned.es/portal/page?\\_pageid=93.656896.93\\_20543314&\\_dad=portal&\\_schema=PORTAL](http://portal.uned.es/portal/page?_pageid=93.656896.93_20543314&_dad=portal&_schema=PORTAL)

les ofrece para realizar algún trabajo práctico siempre son lo suficientemente amplios para que el conjunto de todos los alumnos pueda realizarlo. Todo esto significa que el control sobre el aprendizaje no lo tiene el equipo docente sino el alumno individual. Los estudios realizados sobre estos alumnos indican la necesidad que tienen de gestionar su propio aprendizaje (Gaudioso et al., 2003). La diversidad y las otras actividades y responsabilidades de estos alumnos, impiden diseñar experiencias educativas pensando en un tipo de alumno en especial. Cualquier experiencia tiene que ser lo suficientemente general para que el alumno tenga opción de adaptar su proceso de aprendizaje.

## **La experiencia de aprendizaje colaborativo**

La UNED es una universidad a distancia que desde hace unos años se ha convertido una institución de referencia en la ED. Ofrece servicios con los que ha creado un modelo educativo de cuarta generación (García Aretio, 2009). Las últimas generaciones de ED han dado un paso cualitativo importante por dos factores: la rapidez con la que se puede obtener el feedback, y la gran cantidad de vías de comunicación tanto verticales como horizontales. Ambos factores son caldo de cultivo de la colaboración como estrategia de aprendizaje.

Gran cantidad de estudios avalan la conveniencia de utilizar la colaboración en entornos de ED (Martínez R. et al., 2007; Barkley et al., 2004; Johnson y Johnson, 2004; Elvheim, 2002). Todos los autores informan de las ventajas generales que aporta el aprendizaje colaborativo. Algunos establecen metodologías que dirigen la colaboración (Martínez R. et al., 2007), otros establecen el contexto apropiado (Barkley et al., 2004), y los últimos matizan los contextos proponiendo condiciones que se deben cumplir (Johnson y Johnson, 2004). En general se puede decir que la colaboración incentiva el aprendizaje activo, uno de los objetivos del aprendizaje (Bonwell y Eison, 1991), que tienen su origen en que el ser humano es un ser social, o animal social “zoon politikon” (Aristóteles, 1985).

Sin embargo, la estrategia de colaboración no está ampliamente desarrollada en los estudios de la UNED. La UNED ofrece servicios de colaboración (foros, Chat, correo electrónico), pero esto no es suficiente para que el aprendizaje colaborativo suceda

(Johnson y Johnson, 1994). En este punto la UNED se enfrenta a problemas que requieren soluciones originales debido a la diversidad de alumnos (Boticario y Gaudioso, 2000a), a la gran cantidad de ellos (la asignatura de IA-IC ha tenido los últimos años más de 400), a la necesidad de estos alumnos de gestionar su aprendizaje (Gaudioso et al., 2003). Por ello, si el equipo docente de la asignatura quiere aplicar estrategias de aprendizaje colaborativo de forma efectiva, el trabajo a realizar puede llegar a ser inabordable. Esta tesis investiga sobre una solución alternativa al problema que ha nacido con el objetivo de ser tan general como se pueda para poder ser aplicable en otros entornos colaborativos.

Esta investigación ha propuesto a los alumnos de IA-IC una experiencia de aprendizaje colaborativo de larga duración. La experiencia se enmarca dentro de la práctica obligatoria de IA-IC. Al ser uno de los objetivos de la asignatura llevar a cabo aplicaciones de los conocimientos teóricos, se pensó que la estrategia colaborativa ayudaría a los alumnos en la realización de la práctica. Considerando especialmente la adquisición de competencias de colaboración requeridas para un profesional de informática.

La experiencia siguió la metodología de la UNED para asegurarse que la mayor parte de los alumnos de IA-IC pudiera realizarla. Por eso los materiales se ofrecieron al comienzo de la experiencia y se informó de los plazos de las tareas colaborativas, algunos de estos plazos eran orientativos mientras otros obligatorios. La experiencia propuesta ha tenido una duración ligeramente superior a los tres meses. El objetivo es que el control de la colaboración recayera en los alumnos.

Sin embargo, para intentar garantizar la efectividad del aprendizaje colaborativo, se ha planteado un diseño de un sistema en cierto sentido adaptativo centrado en el alumno, lo cual ha requerido un análisis del comportamiento de los alumnos en el trabajo en equipo por medio de las acciones que iban realizando (Dimitracopoulou, 2008). A partir de ahí se han podido aplicar técnicas de minería de datos (Romero y Ventura, 2007) con las cuales inferir valoraciones sobre la colaboración útiles para garantizar que el aprendizaje colaborativo sucede (Johnson y Johnson, 2004). Se ha modelado a los alumnos según un modelo de la colaboración, y se han propuesto una serie de herramientas de autorregulación que ayuden a los estudiantes a gestionar el proceso de colaboración.

## **Modelo educativo de la experiencia colaborativa**

La motivación de la experiencia de aprendizaje colaborativo ha sido la de suministrar un entorno educativo de colaboración donde los estudiantes pudieran realizar tareas eminentemente prácticas trabajando en equipo. Los estudiantes que podían participar eran todos los matriculados en la asignatura IA-IC. El autor de esta tesis fue el responsable del seguimiento de la experiencia. Debido a las condiciones de los alumnos de la UNED, la participación en la experiencia fue voluntaria, y quien la finalizara terminaba la práctica obligatoria de la asignatura. Como se pensó que la experiencia sería beneficiosa para los estudiantes, se les motivó ofreciendo sustituir la nota que obtuvieran en la experiencia por parte de las preguntas teóricas del examen del primer parcial de IA-IC.

### **Objetivos docentes**

Los estudiantes debían realizar un ejercicio práctico que consistía en resolver un problema de planificación. La planificación es una área dentro de la Inteligencia artificial que forma parte del temario de la asignatura de IA-IC. El objetivo era realizar una serie de tareas secuenciales de forma individual o en equipo que los participantes tendrían que planificar, desarrollar y documentar. Se permitía que el trabajo se realizara de forma colaborativa si el alumno así lo quería. Los objetivos se pueden dividir en individuales y grupales.

- **Objetivos individuales:**
  - que supiera utilizar un entorno educativo utilizando las TIC,
  - que cada estudiante supiera planificar el desarrollo de una solución a un problema de planificación,
  - que supiera discutir con otros compañeros problemas o soluciones de las tareas solicitadas,
  - que ofreciera soluciones creativas al problema,
  - que documentara su trabajo adecuadamente.
- **Objetivos grupales debido al trabajo en equipo:**
  - que el grupo supiera planificar el trabajo común adaptándose a los miembros del equipo,
  - que a través del diálogo y la negociación se alcanzaran objetivos comunes,



- que supiera utilizar distintos medios de comunicación para alcanzar los acuerdos.

## **Requisitos**

Los requisitos de conocimientos eran los mismos que los de la asignatura IA-IC. Sin embargo, el trabajo colaborativo requería unas condiciones que los estudiantes debían cumplir:

- Amplio conocimiento de servicios de comunicación a través de Internet.
- Capacidad de planificación y desarrollo de tareas secuenciales en un periodo de tiempo no muy amplio.
- Motivación por trabajar de forma colaborativa.

Los requisitos fueron puestos a prueba para excluir de la experiencia a los estudiantes que tuvieran grandes dificultades para trabajar en entornos colaborativos vía web, antes de empezar el periodo colaborativo.

## **Destrezas y competencias**

Las destrezas y competencias que los estudiantes debían desarrollar es un subconjunto de las debidas a la asignatura IA-IC. Nos centramos en la parte colaborativa ya que era la que introducía más novedades.

- Capacidad para diseñar, desarrollar, implementar, verificar y documentar soluciones software a un problema de planificación utilizando distintos lenguajes de implementación y distintos planificadores.
- Competencias en el uso de las herramientas y recursos de la Sociedad del Conocimiento: Manejo de las TIC. Competencia en la búsqueda de información relevante. Competencia en la gestión y organización de la información. Competencia en la recolección de datos.
- Trabajo en equipo desarrollando distinto tipo de funciones o roles. En la Sociedad del Conocimiento se presta especial atención a las potencialidades del trabajo en equipo y a la construcción conjunta de conocimiento, por lo que las competencias relacionadas con el trabajo colaborativo son particularmente relevantes: Habilidad para coordinarse con el trabajo de otros. Habilidad para negociar de forma eficaz. Habilidad para la mediación y resolución de

- conflictos. Habilidad para coordinar grupos de trabajo. Liderazgo (cuando se estime oportuno).
- Competencias cognitivas superiores: selección y manejo adecuado de conocimientos, recursos y estrategias cognitivas de nivel superior apropiados para el afrontamiento y resolución de diverso tipo de tareas/problemas con distinto nivel de complejidad y novedad: Análisis y Síntesis.
  - Competencias de expresión y comunicación (a través de distintos medios y con distinto tipo de interlocutores): Comunicación y expresión escrita. Comunicación y expresión oral. Comunicación y expresión matemática, científica y tecnológica (cuando sea requerido y estableciendo los niveles oportunos)

## **Contextualización**

El contexto general de la experiencia de aprendizaje colaborativo es el debido a la asignatura, la UNED y los alumnos. Al entrar en detalle el contexto es el de la práctica de planificación y el contexto tecnológico utilizado para proporcionar un entorno colaborativo.

El campo de la planificación en Inteligencia Artificial, también llamada planificación automática, busca construir algoritmos de control que permitan a un agente sintetizar una secuencia de acciones que le lleve a alcanzar sus objetivos. Un problema de planificación puede definirse como el problema de encontrar una secuencia de acciones que transformen el estado inicial dado en el estado final deseado.

Según Ghallap et al. (2004) una de las motivaciones de la planificación automática es muy práctica: diseñar herramientas de procesado de información que ofrezcan acceso a recursos de planificación eficientes y asequibles. Por eso mismo, en la experiencia de aprendizaje colaborativo se pedía solucionar un problema de planificación para que los alumnos aprendieran el uso y utilidad de la técnica.

Respecto a la tecnología utilizada para dar soporte a un entorno colaborativo, se ha utilizado la plataforma dotLRN. dotLRN es una plataforma de teleformación (e-learning) que facilita la colaboración y gestión de clases a través de Internet. Está basado en la plataforma de gestión de comunidades virtuales OpenACS. Es una

plataforma potente, escalable y flexible que puede soportar un uso fuerte por parte de los usuarios.

dotLRN permite gestionar los cursos y las clases del centro educativo. Pero sobre todo la relación de todos los usuarios del centro educativo: alumnos, profesores, coordinadores, administradores, personal externo, etc., a través de Internet. Esto significa una gestión en cualquier momento, desde cualquier lugar y desde cualquier dispositivo con conexión a Internet. Esta plataforma se utiliza desde hace tiempo en la UNED obteniendo buenos resultados (Santos et al., 2006).

## **Metodología**

En esta sección se describirá la experiencia de aprendizaje colaborativo explicando la metodología seguida. La experiencia no consiste sólo en lo que se les ofrece a los estudiantes. También incluye todo el proceso de análisis que desde las interacciones de los estudiantes llega a las herramientas propuestas para la mejora del control del proceso de colaboración. Por lo tanto, se describirán todos los elementos relacionados: la experiencia pedagógica, el análisis de las interacciones, el método de inferencia, el modelado de la colaboración, las herramientas de autorregulación ofrecidas, y la evaluación realizada para validar la experiencia de aprendizaje colaborativo.

### **Experiencia pedagógica**

Las necesidades de un estudiante actual en el aprendizaje de competencias necesarias para una mejor utilidad de sus conocimientos en un entorno laboral, por ejemplo, y los condicionamientos de la ED, hacen aconsejable y apropiado el uso de la colaboración como estrategia de aprendizaje (Barkley et al., 2004; García Aretio, 2009; European Comision, 2008). Las competencias y objetivos de la asignatura IA-IC hacen también apropiado el trabajo práctico de forma colaborativa.

Por estas razones se ha diseñado una experiencia colaborativa de larga duración en la que han participado los alumnos de la asignatura de IA-IC de cuarto curso de la UNED durante los cursos 2006/2007, 2007/2008 y 2008/2009. Se incentivo la realización del trabajo práctico por la forma colaborativa al valorar el trabajo colaborativo en 2 puntos máximo y sustituyendo dicha puntuación por las cuestiones

que se preguntaban en la parte de planificación del examen del primer parcial, cuya valoración ha sido de 2 puntos.

La experiencia colaborativa se ha diseñado de tal forma que el alumno tenga control sobre su propio aprendizaje y proceso de colaboración. De esta forma se busca un aprendizaje activo (Bonwell y Eison, 1991) y que el alumno se sienta responsable de su propio aprendizaje. La experiencia de aprendizaje colaborativo ha sido de larga duración, en donde tenían que realizar una serie de tareas de forma colaborativa pero sin la obligatoriedad de unas fechas determinadas de entrega, salvo el trabajo final. La experiencia de aprendizaje colaborativo ha consistido en dos fases, como se puede ver en la Figura 4.

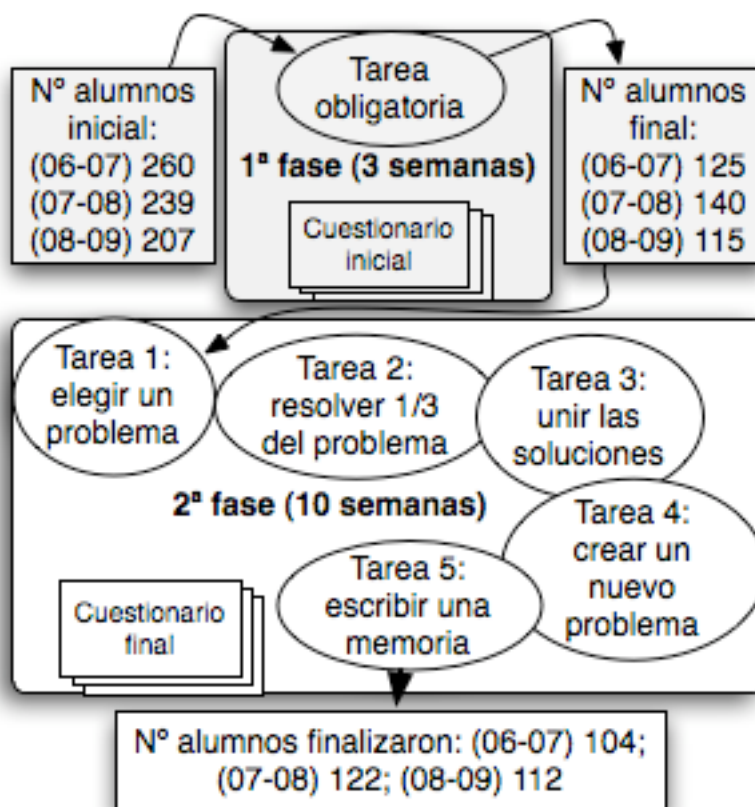


Figura 4: esquema de tareas de la experiencia de aprendizaje colaborativo

### *Primer periodo*

Lo primero que tenían que hacer los alumnos que desearan comenzar la experiencia colaborativa era registrarse en la plataforma dotLRN. El tutor daba acceso a las peticiones. Además de dar acceso a los alumnos, el tutor dio acceso a otros

miembros del equipo docente y tutores de la asignatura para que pudieran estar informados de lo que se realizaría en la experiencia.

Una vez que los alumnos ya estaban registrados, se les ofrecía los siguientes servicios:

- Un servicio de foros diferenciados según el tema para comunicarse con el tutor o el resto de alumnos.
- Un servicio de noticias gestionado por el tutor.
- Un área de documentación estructurada en directorios donde se encontraba toda la documentación necesaria para realizar la práctica.
- Una guía donde se explicaba lo que se debía realizar en la práctica y que estaba almacenada en el área de documentación.
- Una tarea sencilla que los alumnos tenían que realizar para pasar al segundo periodo y que se podía descargar desde el área de documentación.
- Un servicio de encuestas.
- Una encuesta inicial que preguntaba información personal relativa al alumno con el objetivo de conocer su contexto de colaboración para seleccionar los mejores candidatos a forma un equipo para el segundo periodo.

Estos servicios se ofrecían en un espacio de la plataforma que desde ahora llamaremos espacio general.

The screenshot shows the LIRN platform interface for the 'Práctica de planificación 08/09' community. The interface includes a navigation menu, a news section, a frequently asked questions section, a forums section with a list of topics, an assessment section with a table of open assessments, and a members section.

**Open Assessments Table:**

Title	Status	Actions	Review	Administration	Results
Encuesta de opinión sobre la práctica	Untaken	Take		Administration	Results

**Figura 5: Espacio general del primer periodo de la experiencia.**

Como tareas de este periodo se pospusieron:

- responder a la encuesta, y
- realizar una tarea sencilla obligatoria.

Aunque se propuso como tarea del primer periodo, la respuesta a la encuesta era voluntaria, pero así se consiguió que la mayor parte de los alumnos respondieran.

La tarea sencilla obligatoria fue distinta en cada curso. Esta tarea consistía en resolver un problema simple de planificación con dos planificadores distintos. Se aconsejaba el uso de los planificadores Prodigy 4.0 y SGP, y se ofrecía documentación necesaria para su instalación y primer contacto. El objetivo de esta tarea era instalar los dos planificadores. El problema que se pedía realizar era lo suficientemente sencillo, desde el punto de vista del tutor, para que los principales problemas con los que se tuviera que enfrentar el alumno fueran de instalación. El objetivo de la tarea también fue limitar el número de alumnos que empezaran el segundo periodo a aquellos que hubieran tenido el tiempo y la organización necesaria para terminarla. El tutor comprobó la totalidad de las soluciones de los alumnos para evitar engaños.

Como se ha dicho, se propuso la respuesta a una encuesta inicial como tarea en el primer periodo. Los alumnos respondieron masivamente a las encuestas y sus respuestas fueron utilizadas para crear los equipos de 3 miembros. Se preguntó:

- Preguntas sobre datos personales.
- Preguntas sobre datos académicos relacionados con la UNED.
- Preguntas sobre datos laborales.
- Preguntas sobre preferencias de estudio y en trabajo de equipo.
- Preguntas sobre conocimiento previo.

El primer curso las preguntas se redujeron a los anteriores apartados. En el curso 2007/2008 se añadió un apartado más, datos demográficos y socio-económicos. En el curso 2008/2009 se añadió el apartado anterior y se cambiaron algunas preguntas para descubrir con más facilidad datos necesarios para la colaboración. Las tres encuestas se pueden encontrar en su totalidad en el Apéndice I: encuestas.

### ***Segundo periodo***

El segundo periodo empezaba a mediados o finales de noviembre y terminaba a finales de enero del año siguiente. Los alumnos que terminaron la tarea sencilla del primer periodo, y lo solicitaran, fueron agrupados en equipos de 3 miembros, de 4 en

el caso de que el número de alumnos que pedían realizar el segundo periodo, y hubieran terminado la tarea sencilla obligatoria, fuera un número no múltiplo de 3.

El tutor, una vez que hubo evaluado las tareas sencillas, construyó equipos de 3 miembros. Cuando el tutor consultó la literatura sobre la construcción de grupos de colaboración (Cooper, 1990; Johnson et al., 1991; Smith, 1986) no encontró un criterio claro y sencillo que seguir, salvo la recomendación de crear grupos de colaboración de pocos miembros pero, si pudiera ser, de más de dos. Le pareció apropiado, sin embargo, seguir la recomendación de (Muehlenbrock, 2005) relativa a que es conveniente tener en cuenta, entre otros factores, el contexto de la colaboración a la hora de crear los grupos para que sus miembros colaboren. El contexto de la colaboración fue preguntado en la encuesta inicial que los alumnos respondieron al comenzar el primer periodo. Como el objetivo era que los equipos colaboraran, dio más importancia a los horarios de conexión a internet y la localización geográfica. Para cada equipo se abrió un subespacio donde los miembros podían comunicarse. Los servicios ofrecidos fueron:

- Un servicio de foros.
- Área de documentación privada para el equipo.
- Servicio de encuestas.
- Calendario.
- Material de clase.
- Información del grupo, donde se daba información de las tareas a realizar, de los miembros del equipo y de información en ayuda a la colaboración, pero esto último sólo en el año 2008/2009.
- Servicio de Chat, pero sólo en el año 2008/2009.

The screenshot displays the LMS interface for a team-specific space. At the top, there is a navigation bar with links for Home, Terms, Departments, Subjects, Classes, Communities, Users, Admin, and Show Xtra Info. Below this, there is a search bar and a user profile section. The main navigation area includes tabs for Home, Courses, Communities, Control Panel, Administration, and Equipo-2. Under the Equipo-2 tab, there are sub-tabs for Subgroup Home, Subgroup Info, Calendar, Files, Class Material, Sala, and Class Admin. The content area is divided into three main sections: News, Forums, and Assessment. The News section shows 'No News' and a link to 'Add a News Item'. The Forums section shows a link to 'Equipo-2 > Discusión'. The Assessment section shows a table of 'Open Assessments' with the following data:

Title	Status	Actions	Review	Administration	Results
Encuesta de opinión sobre la información de los indicadores estadísticos de los foros y el nivel de colaboración	Untaken	Take		Administration	Results

**Figura 6: Espacio virtual específico de un equipo.**

A cada equipo se le propuso una serie de tareas a realizar y cuya documentación se podía descargar desde Material de clase. El espacio de cada equipo ofrecía los servicios necesarios para que los miembros se comunicaran entre sí sin tener que recurrir a medios externos (correo electrónico, Skype, Messenger, etc.). Se les aconsejó utilizar los servicios del espacio del equipo para que el tutor de la experiencia pudiera realizar las tareas de seguimiento de la colaboración, aunque el seguimiento fue realizado por otros medios. Las tareas que se tenían que realizar fueron las siguientes:

- Tarea 1: a los alumnos se les ofrecían 3 problemas de planificación distintos, que tenían que ser solucionados con dos planificadores distintos, y se les pedía que seleccionaran uno de ellos dejando constancia en los foros del espacio del equipo. Además se les pedía que dieran una estimación del tiempo que tardaría en realizar la tarea 2 y por quién.
- Tarea 2: los problemas de planificación estaban divididos en 3 partes y cada parte la tenía que realizar un miembro del equipo de forma individual. La ayuda entre miembros estuvo siempre permitida. A los equipos de 4 miembros se les ofrecieron 4 partes del problema que en la tarea 1 habían elegido.



- Tarea 3: los miembros del equipo tenían que unir las 3 partes de la tarea 2 en una sola, lo que era el problema completo de planificación que se había elegido en la tarea 1.
- Tarea 4: una vez que el problema completo de planificación estaba resuelto, se pidió que los miembros realizaran variaciones para estudiar el comportamiento de los dos planificadores. En esta tarea se daba total libertad para cambiar del problema original lo que quisiera, advirtiéndole que el objetivo sería realizar experimentos siguiendo el método científico.
- Tarea 5: los equipos tenían que escribir una memoria del trabajo realizado.

Al final del segundo periodo se les pedía que se respondiera a una encuesta final donde se pedía opinión sobre la experiencia colaborativa, la plataforma, el tutor y los compañeros.

Una vez terminada la experiencia, o a punto de terminar, se les propuso a los estudiantes que respondieran una serie de encuestas finales. A los estudiantes se les preguntó sobre:

- Datos personales y preferencias educativas.
- Opinión sobre diferentes aspectos de la experiencia de aprendizaje colaborativo.
- Relación entre la colaboración y la motivación.
- Trabajo en equipo realizado.

La fecha en la que se habilitó la encuesta cambió de curso en curso para que aumentaran el número de estudiantes que respondieran. Siempre se advirtió que no existía ninguna influencia con la nota del examen. En el Apéndice I: encuestas, se han incluido las encuestas finales a cada curso.

Se comentó que durante el segundo periodo de la experiencia del curso 2008/2009 se fueron ofreciendo distintas herramientas a ciertos equipos para suministrar más información sobre la colaboración y las interacciones con el objetivo de ayudar a mejorar el proceso de colaboración. Para evaluar la información y las herramientas, se habilitaron encuestas para que los miembros de los equipos implicados dieran su opinión sobre las mismas. En las encuestas se preguntaba sobre: la usabilidad (solamente a los equipos a los que se les ofreció la herramienta escrutable, la aplicación web de la que se hablará más adelante), la valoración general sobre la

información presentada y la valoración general sobre la herramienta. Las encuestas se encuentran en el Apéndice I: encuestas.

### ***Discusión y justificación de la propuesta pedagógica***

Realmente la experiencia de aprendizaje colaborativo empezaba en el segundo periodo. La tarea obligatoria y los plazos para realizarla tenían como objetivo familiarizarse con el método de trabajo y comprobar que el estudiante podría realizar un trabajo colaborativo. Al no existir dificultad en la tarea obligatoria, los estudiantes que no pudieron fue por problemas de motivación o tiempo. La participación fue incentivada por la novedad de la experiencia de aprendizaje colaborativo, y por los beneficios en el examen del primer parcial de la asignatura de IA-IC, ya que la nota obtenida (máximo 2 puntos) sería sustituida por la nota del apartado de cuestiones (máximo 2 puntos) del examen del primer parcial.

El diseño que ha seguido la experiencia pedagógica ha sido guiado para que se cumplieran las cinco condiciones de Johnson y Johnson (1994) para asegurar que el aprendizaje colaborativo sucede, pero también se ha tenido en cuenta la instrucción de los estudiantes en el uso del entorno colaborativo y de sus habilidades de trabajo en equipo. En entornos colaborativos, además de proponer un conjunto de tareas de trabajo en equipo, los estudiantes tienen que saber interactuar y desarrollar sus habilidades o competencias de trabajo en equipo. Por lo tanto, el entorno debería instruir a los estudiantes, mientras ocurre la experiencia colaborativa, en el uso de la tecnología para interactuar y en las habilidades colaborativas (Macdonald, 2003)

Durante el primer periodo las tareas propuestas se caracterizaron por su simplicidad para que los principales problemas fueran el acercamiento a la plataforma y comenzaran a interactuar con ella y el resto de compañeros. Al pasar al segundo periodo, la primera tarea propuesta obligaba a los miembros de un mismo equipo a comunicarse entre sí mediante los servicios de comunicación de la plataforma o mediante otros, y proponía una elección para provocar una discusión que llevara a un acuerdo, y que ese acuerdo se concretara en un trabajo por escrito. La tarea es lo suficiente sencilla para servir de entrenamiento o instrucción de las habilidades de comunicación y trabajo colaborativo. Las siguientes tareas estaban diseñadas para aumentar el uso de las habilidades de trabajo en equipo, llegado a la tarea 4 donde la

discusión, el acuerdo, el trabajo en equipo la crítica ese trabajo en equipo son los objetivos.

Además de que las tareas se fueran haciendo más complejas desde el punto de vista del trabajo en equipo, también fueron diseñadas para cumplir las condiciones de Johnson y Johnson (2004). La primera condición es que se perciba con claridad la interdependencia positiva. Desde el primer momento se explicó a los alumnos que la valoración que se les iba a dar dependía del trabajo final y de la colaboración realizada, penalizando a los alumnos menos colaboradores. Para ello se recomendó que las interacciones se realizaran mediante los servicios suministrados por la plataforma web dotLRN. También se informó que los alumnos que no colaboraran serían expulsados sin penalización al resto de compañeros. Evidentemente, si la comunicación entre los alumnos se realizaba fuera de la plataforma, el tutor no podía tener conocimiento de la colaboración del equipo. Sin embargo, al ser la experiencia en un escenario con alumnos reales, no se podía restringir la comunicación a los servicios de la plataforma únicamente. El objetivo de la experiencia ha sido incrementar el aprendizaje pero dejando a los alumnos el control sobre el proceso. Ellos tenían que decidir los medios de comunicación apropiados.

La segunda condición era la de promocionar la interacción entre los miembros de los equipos, sin impedir la comunicación cara a cara. Justificando esta interacción se pidió a los alumnos que respondieran a la encuesta inicial. En ella se solicitaba datos relativos al contexto de la colaboración como en qué parte del día se podían conectar a internet o dónde residían. Los equipos de tres miembros para empezar el periodo colaborativo fueron hechos teniendo en cuenta la facilidad en horarios para colaborar, pensando que conectándose al mismo tiempo a internet los alumnos podrían comunicarse con más facilidad. El lugar de residencia de los estudiantes también fue un factor a considerar. Se hace notar que los alumnos de la asignatura provenían de todas las regiones de España y alguno del extranjero. Agrupando estudiantes residentes en la misma ciudad se daba la oportunidad de que pudieran comunicarse cara a cara y de que hubiera más elementos comunes que facilitarían la interacción social y la comunicación. Esto último era responsabilidad de los miembros del equipo. No hay que olvidar que la motivación original ha sido la mejora del aprendizaje. Aunque el objetivo de la investigación ha sido aumentar el control del proceso de colaboración mediante inferencias respecto al trabajo en equipo, por lo que

ha sido necesario un análisis de las interacciones. Evidentemente, no se dejaría constancia de las interacciones en el espacio del equipo si el equipo se comunicaba cara a cara. Sin embargo, en la experiencia no se ha impedido que los estudiantes hicieran algo que pudiera mejorar la colaboración, aunque no se dejará constancia en las interacciones.

La tercera condición fue que los participantes rindieran cuentas individuales y de responsabilidad para alcanzar los objetivos del equipo. Para facilitar esta clase de control, las tareas que se debían realizar siguieron un orden secuencial y se propusieron unas fechas de inicio y final. Se aconsejó a los alumnos que hicieran un calendario propio, aunque se podían basar en las fechas propuestas. El tutor se guió por el calendario de tareas para observar el trabajo de cada equipo. En casos excepcionales se puso en contacto con los equipos o miembros del equipo que no cumplían sus propios calendarios. A su vez, los alumnos también fueron conscientes de su trabajo y responsabilidad, así como del de sus compañeros. Se ha detectado la relativa importancia de las fechas propuestas aconsejadas de finalización de las tareas. Esto a muchos les ha ayudado a gestionar su trabajo.

La cuarta condición es que hubiera un uso frecuente de las relaciones interpersonales y habilidades del equipo. Por ello los equipos eran de 3 o 4 miembros para incentivar la comunicación entre y se han propuesto tareas, como la Tarea 1, donde la comunicación e intercambio de ideas era sencilla, lo que incentiva las relaciones interpersonales.

La quinta y última condición es el frecuente y regular procesamiento del funcionamiento del equipo. Este punto ha sido el más difícil de abordar en el diseño de la experiencia. Durante los dos primeros años el tutor fue el encargado de procesar el funcionamiento del equipo. Pero este trabajo fue absolutamente ineficiente debido al gran número de alumnos. Era necesario un medio que procesara el funcionamiento de los equipos y reportara los resultados tanto al tutor como a los alumnos. Debido a esta última condición, la investigación que se explica en esta tesis ha sido realizada.

#### *Cambios realizados durante los tres cursos*

Dentro de la propia experiencia pedagógica los principales ajustes que se han realizado de una experiencia a otra han sido:

- Las fechas de finalización de la tarea obligatoria de la primera fase y la fecha de inicio de la segunda fase. En el origen, se propuso una fecha de finalización de la tarea obligatoria y de inicio de la segunda fase. Sin embargo, una gran cantidad de estudiantes terminaron la tarea obligatoria de la primera fase con rapidez y solicitaron no tener que esperar. Se les hizo caso y se ofrecieron dos fechas de inicio de la segunda fase ya en la experiencia de 2006/2007. Esta misma metodología ha seguido en las experiencias de los cursos 2007/2008 y 2008/2009, donde todos los estudiantes conocían al empezar la fase primera que podían elegir entre dos fechas de inicio de la segunda fase si terminaba la tarea obligatoria en el plazo.
- En las encuestas finales también hubo cambios debidos a las respuestas. En este caso, los cambios no se produjeron por problemas de privacidad, sino por dudas en la veracidad de las respuestas. En las encuestas finales de los cursos 2006/2007 y 2007/2008 se les preguntaba a los estudiantes participantes en el segundo periodo de la experiencia de aprendizaje colaborativo que evaluaran la colaboración de sus compañeros y marcaran a alguno de ellos, si se daba el caso, como un prototipo de un estudiante muy colaborativo. Por alguna razón, los estudiantes valoraban a sus compañeros con altos grados de colaboración y no solían distinguir a uno de sus compañeros como el mejor. Esto indicaba que no hacían una evaluación objetiva y deseaban elogiar a sus compañeros de su propio equipo. Las respuestas que podían considerarse objetivas eran tan pocas que en la experiencia del curso 2008/2009 no se preguntó dicha evaluación de la colaboración de los compañeros.

### **Análisis de las interacciones**

El análisis de las interacciones es un área de investigación que pretende mejorar el aprendizaje analizando las interacciones de los alumnos en un entorno colaborativo. Se ha advertido de las ventajas del análisis y de la conveniencia del uso de sus resultados para mejorar la autorregulación de los alumnos (Dimitracopoulou, 2008).

El análisis de las interacciones obtiene indicadores relativos a las interacciones pero no establecen juicios sobre las mismas. El objetivo del análisis de las interacciones ha sido proveer de una base al método de inferencia sobre la colaboración con el cual obtener juicios de forma regular y frecuente. Dicho análisis

se debía referir a la colaboración, que proveyera de información mientras el proceso de colaboración ocurre de forma regular y frecuente, y para que sea genérico y pueda aplicarse en otros entornos, se propuso un método que fuera independiente del dominio.

Ya que el principal medio de comunicación entre los miembros de los equipos fueron los foros, el análisis de las interacciones se centraron en este servicio. Los foros es un medio de comunicación asíncrono de uso muy extendido en LMS. Por lo tanto, un acercamiento de análisis de las interacciones relacionadas con la colaboración puede ser llevado a cabo también en cualquier otro LMS donde la experiencia de aprendizaje colaborativo utilice fundamentalmente los foros como medio de comunicación entre los miembros de los equipos. Siendo un medio de comunicación muy empleado en distintos entornos educativos el análisis de los foros mediante técnicas de minería de datos es muy recomendable por los resultados que se obtienen (Dringus y Ellis, 2005).

Se han propuesto una serie de indicadores estadísticos de las interacciones de los estudiantes en los foros como resultados del análisis de las interacciones. Estos indicadores nos hablan del proceso de colaboración mediante valores cuantitativos y pueden ser obtenidos mientras el proceso de colaboración ocurre, con alta frecuencia y regularidad, y puede ser calculados en cualquier otro entorno donde se utilicen los foros.

Para los métodos de inferencia que se querían descubrir para obtener juicios en relación a la colaboración, además del análisis de las interacciones era necesario otra fuente de información que nos sirviera como un modelo donde comparar los resultados en relación con la colaboración. En este caso un experto clasificó a los alumnos según su nivel de colaboración después de leer los mensajes de todos los alumnos enviados a los foros. Esta fuente de datos es cualitativa y es necesario que la colaboración haya terminado para obtener resultados. La lista de alumnos según su colaboración ha sido utilizada para validar los métodos de inferencia propuestos para deducir juicios sobre la colaboración.

Por lo tanto, las fuentes de datos han sido dos: análisis de las interacciones a los foros que han sido obtenidas por medios automáticos, y una lista de los estudiantes participantes en las experiencias de aprendizaje colaborativo de cada curso etiquetados según su nivel de colaboración deducido de forma manual por un experto.

### ***Indicadores estadísticos***

La plataforma dotLRN ofrece un servicio de foros que, al igual que en cualquier otro foro de Internet, permite crear y enviar mensajes respondiendo a un mensaje individual dentro de un hilo dado. De esta forma, unos mensajes son réplicas a otros por lo que es posible identificar qué usuario respondió al mensaje de otro usuario. DotLRN se utiliza ampliamente en la UNED principalmente en la enseñanza abierta y cursos de postgrado. Aporta la ventaja del almacenamiento de todas las interacciones activas en una base de datos, por lo que se permite hacer minería de datos empotrada. Aporta otros servicios como los logs del servidor web con el que obtener todas las interacciones de los usuarios, tanto activas como pasivas. Debido a que la colaboración es una estrategia de aprendizaje activa, hemos analizado únicamente las interacciones activas relacionadas con la colaboración que se han hecho principalmente en los foros de los espacios virtuales de cada equipo.

Para obtener los indicadores estadísticos de las interacciones en los foros, se crearon un conjunto de funciones en lenguaje psql. Estas funciones calculan los indicadores estadísticos durante un determinado periodo de tiempo, pudiendo también discriminar entre todos los foros, un único foro individual, o los foros pertenecientes a un curso determinado, o a los foros pertenecientes a subcursos de un curso padre. Las funciones fueron guardando el valor de los indicadores en una tabla. Durante los cursos se fue llenando esa tabla con los indicadores estadísticos semana a semana o en la totalidad del tiempo transcurrido del curso. De esta forma se obtuvieron los indicadores estadísticos de cada semana del curso, de lunes a lunes, y de la totalidad del curso. Esta tabla ha sido la fuente de datos. El desarrollo es genérico para la plataforma dotLRN lo que significa que el mismo método de análisis de las interacciones se puede extender a cualquier otro curso, colaborativo o no, que utilice dotLRN.

Los indicadores estadísticos de las interacciones son los siguientes:

- Num\_thrd (N\_thrd): número de hilos o conversaciones iniciadas en los foros del equipo.
- Med\_thrd (M\_thrd): el promedio de hilos o conversaciones iniciadas por día. Es el num\_thrd dividido por la duración del periodo colaborativo.
- Var\_thrd (V\_thrd): cuadrado de la varianza del número de hilos o conversaciones iniciadas en los foros del equipo durante el periodo

colaborativo. La formula es:

$var\_thrd = varianza^2 = (1/n) \sum^n (x_i - med)^2$ ,  $n$  es el número de días del periodo colaborativo;  $x_i$  es el número de hilos creados en el día ( $i$ ) del periodo colaborativo;  $med$  es el promedio de hilos creados por día o  $med\_thrd$ .

- Level\_thrd (L\_thrd): el número de hilos o conversaciones iniciadas dividido por la varianza [ $level\_thrd = num\_thrd / (var\_thrd)^{1/2}$ ].
- Num\_msg (N\_msg): número de mensajes enviados por un alumno a los foros del equipo.
- Med\_msg (M\_msg): el promedio de mensajes enviados por día a los foros del equipo.
- Var\_msg (V\_msg): varianza del número de mensajes enviados. La formula es la misma que para  $var\_thrd$  sustituyendo  $x_i$  por el número de mensajes enviado el día ( $i$ ) y  $med$  por  $med\_msg$ .
- Level\_msg (L\_msg): el número de mensajes dividido por la raíz cuadrada de  $var\_msg$ .
- Num\_reply\_thrd (N\_reply\_thrd): el número de réplicas a los hilos creados por el alumno.
- Med\_reply\_thrd (M\_reply\_thrd): el promedio de mensajes en los hilos creados por el alumno. La fórmula es:  $med\_reply\_thrd = num\_reply\_thrd / num\_thrd$ .
- Num\_reply\_msg (N\_reply\_msg): número de réplicas a los mensajes enviados por el alumno.
- Med\_reply\_msg (M\_reply\_msg): promedio de réplicas a los mensajes enviados por el alumno. La fórmula es:  $med\_reply\_msg = num\_reply\_msg / num\_msg$ .

Las réplicas son mensajes enviados a los foros respondiendo a un mensaje determinado.

Siguiendo (Santos et al., 2003) se puede dar un significado semántico a los indicadores. El significado propuesto es el siguiente:

- Num\_thrd: iniciativa del alumno. A más hilos creados, más iniciativa.
- Med\_thrd: iniciativa media del alumno durante un tiempo.
- Var\_thrd: constancia o regularidad en la iniciativa. A valores cercanos a 0 significaría que ha mantenido su nivel de iniciativa constante con el tiempo.



- **Level\_thr**: este atributo es un indicador de la iniciativa y regularidad, a más alto el valor quiere decir que el alumno ha tenido más iniciativa y la ha mantenido constante durante tiempo.
- **Num\_msg**: actividad del alumno. A más mensajes creados, más activo el alumno.
- **Med\_msg**: actividad media del alumno.
- **Var\_msg**: constancia o regularidad de la actividad.
- **Level\_msg**: a mayor valor indica que una alta actividad se ha mantenido constante durante tiempo.
- **Num\_reply\_thr**: actividad causada por la iniciativa del alumno.
- **Med\_reply\_thr**: media de la actividad causada por la iniciativa del alumno.
- **Num\_reply\_msg**: actividad causada por la actividad del alumno.
- **Med\_reply\_msg**: media de la actividad causada por la actividad del alumno.

Con los indicadores se crean datasets por curso, es decir, se recolectan los datos en forma de tabla donde cada columna es un indicador estadísticos y donde cada fila se refiere al estudiante que ha realizado las interacciones.

#### *Otros indicadores estadísticos propuestos*

Dentro del análisis de las interacciones se propusieron más indicadores estadísticos de las interacciones en los foros que los explicados. Estos fueron:

- Indicadores estadísticos de interacción de un estudiante con el tutor y viceversa. Las interacciones con el tutor podrían indicar gran interés del estudiante o la identificación de algún problema por parte del tutor. La causa principal por la que no se utilizaron estos indicadores fue por la gran cantidad de estudiantes. También por esta razón nació el método de inferencia ya que un único tutor no podía atender convenientemente a todos los estudiantes en el entorno colaborativo. El número de alumnos que tuvieron alguna interacción con el tutor fue tan escaso que la introducción de indicadores que tuvieran en cuenta estas interacciones falsearían los análisis y métodos de inferencia.

#### ***Lista de alumnos según su colaboración***

Un experto, que fue el tutor de las experiencias de aprendizaje colaborativo, después de terminada cada una, leyó todo los mensajes enviados a los foros de los equipos durante el periodo colaborativo. Fue leyendo equipo por equipo y evaluó a

los alumnos según su colaboración. Utilizó una escala numérica de 9 valores: 1, 2 y 3 los utilizó con alumnos muy colaborativos; 4, 5 y 6 para alumnos de colaboración media; 7, 8 y 9 para alumnos muy poco colaborativos. La escala es inversa, 1 es el valor de colaboración más alta. No utilizó nunca el valor 9 ya que los tan poco colaborativos no enviarían ningún mensaje y utilizó sólo una vez el valor de 1.

### **Inferencia de la colaboración**

En este apartado se aclara el acercamiento realizado en el método de inferencia de la colaboración. Primero revisaremos la literatura existente sobre la conveniencia de establecer un procedimiento que analice la colaboración. Después describiremos las características generales del acercamiento según los objetivos de la tesis e investigaciones realizadas en el mismo sentido. Por último, entraremos en los detalles de nuestro planteamiento.

La quinta condición que propusieron (Johnson y Johnson, 2004) establece que es necesario un procesamiento frecuente y regular del funcionamiento del equipo. El objetivo de esta investigación ha sido el realizar dicho procesamiento de forma automática o con la menor participación posible de los tutores.

De (Johnson et al., 1991) hemos cogido la idea de que la evaluación de la colaboración de los alumnos mientras esta se está realizando, y mostrando esta evaluación, provoca la competencia en la colaboración de los miembros de un mismo equipo.

En definitiva, para cumplir con los objetivos de la investigación en un entorno de aprendizaje colaborativo es necesario que, la colaboración de los participantes cumplan las condiciones de Johnson y Johnson (2004) además que se suministre una información útil para los participantes. Esta información sobre la colaboración se convierte en el medio por el cual podremos verificar la hipótesis de que ayudará a los participantes en la mejora del control del proceso de colaboración. Esta ayuda es muy recomendable ya que el control sobre dicho proceso lo tienen los alumnos según el diseño de la propia experiencia de aprendizaje colaborativo. Los otros objetivos son que la información sobre la colaboración se obtenga mientras el proceso de la colaboración está siendo llevado a cabo o poco después, así como que el análisis pueda ser realizado en otros dominios y entornos.

Ha habido muchos estudios sobre las habilidades personales de los alumnos, como estilos de aprendizaje (Felder, 1993), estilos cognitivos (Liu y Ginther, 1999), de pensamiento (Sternberg, 1997). Sin embargo, todavía no se ha llegado a un acuerdo sobre las habilidades personales en relación con el trabajo en equipo. Dentro de la teoría de las múltiples inteligencias (Gardner, 1999), se considera una inteligencia interpersonal, pero no se han encontrado unos parámetros de cálculo de dicha inteligencia y su relación con el trabajo en equipo.

Pocas investigaciones se han centrado en analizar la colaboración para obtener juicios o inferencias sobre la colaboración de los estudiantes (Redondo et al., 2003; Talavera & Gaudioso, 2004; Perera et al., 2007), y no hay un acuerdo en la técnica más útil para el problema, aunque está claro que si se quiere automatizar la inferencia, técnicas de aprendizaje automático son adecuadas.

La hipótesis de aumentar el control sobre el proceso de colaboración requiere que el alumno tenga más información sobre la colaboración. De la literatura no se encuentra con claridad qué información sería la más indicada. En esta investigación se postula, como se explicará en breve, que la idea sobre la colaboración de un equipo puede ser alcanzada comparando la colaboración de unos alumnos con otros o clasificando a los alumnos según su colaboración.

Según (Torres et al., 1999) existen dos líneas fundamentales de pensamiento acerca de las categorías:

- Visión clásica de los conceptos (basada en Aristóteles): un concepto puede ser caracterizado por una serie de atributos definitorios. Los atributos definitorios son una lista de atributos "necesarios y suficientes" para que un miembro entre en una categoría. En esta visión si algo no tiene alguno de los atributos, no forma parte de la categoría, y todos los miembros de la categoría son igual de representativos que ella.
- Visión prototípica de los conceptos (Rosch, 1978): información o conocimientos ligados a un cierto concepto se puede describir mediante una lista de atributos. No existe un conjunto de atributos necesarios y suficientes para categorizar. La inclusión dependerá de su parecido global con otros miembros de la categoría y los que más se parecen a todos serán los más prototípicos. Los límites no son claros por lo que las categorías son conjuntos borrosos.

Si se desea clasificar estudiantes dependiendo de su colaboración, ¿cuál visión se empleará? Esta claro que no hay atributos necesarios y suficientes que definan la colaboración. Hay que decantarse por la visión prototípica. Por lo que habrá que comparar unos estudiantes con otros, o lo que cada uno ha realizado, para deducir la colaboración de cada uno. Esta clasificación mediante comparación es la que se quiere realizar y ofrecer a los alumnos como medio de obtener mayor control sobre el proceso de colaboración. Hemos seguido dos caminos. Por un lado, se podría medir la colaboración con un valor y que los alumnos se compararan con dicho valor y, por otro lado, se podría clasificar a los alumnos según su colaboración.

Para alcanzar el primer acercamiento, se debería crear una métrica que midiera la colaboración de los alumnos según las interacciones realizadas. Esta métrica debería discriminar entre todos los posibles indicadores de las interacciones aquellos más relacionados con la colaboración.

Para alcanzar el segundo acercamiento se debía plantear un procedimiento que clasificara a los alumnos según su colaboración tomando como fuente de datos las interacciones realizadas por éstos. La clasificación obtenida tendría que ser cotejada con la colaboración de los alumnos obtenida desde otra fuente que la utilizada para realizar la clasificación.

En esta tesis se han planteado ambos acercamientos como método de inferencia sobre la colaboración. Hemos tenido en cuenta los objetivos de la investigación en el diseño, por lo tanto, los enfoques planteados (métrica y clasificación) deberán dar información sobre la colaboración mientras el proceso de colaboración está siendo realizado o poco después, y deberán poder aplicarse en otros entornos o dominios.

Al utilizar datos cuantitativos de las interacciones de los alumnos, estos pueden ser obtenidos incluso en el mismo momento en el que las interacciones se realizan, además no es necesario que haya un trabajo por parte del tutor o de otro miembro del equipo docente. Los datos cuantitativos informan de la cantidad, lo cual un sistema informático puede calcular, mientras que informar de la cualidad o calidad requiere de, normalmente, conocimiento lo cual lo aportan expertos. El ofrecer los resultados poco tiempo después de las interacciones, se debe a que el objetivo de la investigación es ir dando información de forma regular y frecuente de la colaboración a los alumnos para que ellos mejoren el proceso de colaboración. Si esta información se ofrece

mucho más tarde de la interacción, la utilidad de la información disminuye, y puede llegar a desaparecer.

La metodología de trabajo llevada a cabo ha sido la siguiente:

1. Analizar las interacciones relacionadas con la colaboración. Se han propuesto una serie de indicadores estadísticos de las interacciones de los alumnos en los foros. Se han creado datasets con el valor de los indicadores. Cada fila han sido los indicadores estadísticos de un alumno.
2. Un experto, que ha sido el propio tutor de las experiencias colaborativas, ha etiquetado a los alumnos según su nivel de colaboración.
3. Se comparan ambas fuentes de datos utilizando tecnología de minería de datos:
  - 3.1. La obtención de la métrica se ha realizado discriminando qué indicadores estadísticos estaban más relacionados con la colaboración. Esta relación se ha obtenido entrenando algoritmos de árbol de decisión. La etiqueta es el valor de la colaboración dada por el experto. El objetivo es identificar qué indicadores son los más utilizados por los árboles de decisión para aprender la colaboración.
  - 3.2. Por otro lado, se ha ejecutado el algoritmo de agrupamiento o “clusterización” (proveniente del término inglés “clustering”, el cual será utilizado a menudo en esta tesis) EM con los datasets para obtener distintos grupos de alumnos. Estos grupos han sido comparados con los valores de la colaboración dada por el experto. De esta forma se ha obtenido una relación entre el cluster y la colaboración.
4. Una vez realizado el análisis de las interacciones y alcanzadas las inferencias, se ha rellenado un modelo de la colaboración, el cual ha sido presentado a los estudiantes mediante diversas herramientas construidas al efecto.

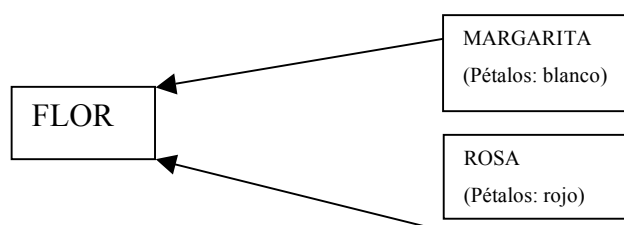
La metodología de la inferencia de la colaboración anterior resume todo el proceso, desde el análisis de las interacciones explicado anteriormente y resumido en los puntos 1 y 2 de la lista, hasta la creación del modelo de la colaboración y el uso del mismo en distintas herramientas, lo cual se ha resumido en el punto 4 y se explicará en profundidad en breve. En la metodología se comentan todos los pasos para dar un contexto al método de inferencia.

### ***Métrica de la colaboración***

Los objetivos de la investigación estarían cumplidos si se encontrara una forma de medir de forma matemática la colaboración a través de indicadores cuantitativos de las interacciones. En esta investigación se han propuesto un conjunto de indicadores estadísticos de las interacciones en los foros. Si una función matemática que empleara algunos de estos indicadores identificara la colaboración, o diera una medida de la colaboración en relación de unos alumnos con otros, dichos objetivos estarían satisfechos.

Una investigación utilizando técnicas de aprendizaje automático ha sido llevada a cabo con los datos de los tres años durante los que se ha realizado la experiencia de aprendizaje colaborativo. El objetivo ha sido encontrar entre los indicadores estadísticos propuestos aquellos que estuvieran relacionados en mayor medida con la colaboración. Una vez identificados, se podrían proponer distintas fórmulas matemáticas o métricas que nos dieran una relación de alumnos, con valores más altos para los más colaborativos, y más bajos para los menos colaborativos.

Lo primero de todo tenía que ser encontrar los indicadores estadísticos, entre los propuestos, en relación con la colaboración. Para ello, se ha requerido la ayuda de algoritmos de árbol de decisión. Estos suministran en un árbol lógico la relación lógica ente los atributos y la clase inferida (Borrajo et al., 2006), por lo que se identifica el indicador empleado que utilizan para aprender la clasificación. Una vez que se tiene los indicadores, había que proponer unas métricas y medir de alguna forma su relación con la colaboración. Un ejemplo de árbol lógico es el que aparece en la Figura 7:



**Figura 7: árbol lógico.**

Todas las instancias son flores, pero las que tienen los pétalos blancos son margaritas y las que tienen los pétalos rojos son rosas. En el ejemplo el atributo discriminante es el color de los pétalos.

Los algoritmos de árbol de decisión clasifican las instancias de un dataset dado de acuerdo a un etiquetado de las instancias según algún criterio dado, y las clasifica. Estos algoritmos ofrecen un árbol lógico el cual muestra los atributos del dataset utilizado para realizar la clasificación. Como es conocida la dificultad lógica y teórica para establecer un patrón que indique la mejor forma de colaborar (Johnson y Johnson, 1994; Johnson et al., 1991), no buscamos ese patrón de la colaboración. En su lugar buscamos una estimación de los indicadores enlazados con la colaboración de los alumnos. Esto es lo que los algoritmos de árbol de decisión ofrecen en sus árboles lógicos, al enlazar los atributos del dataset con el etiquetado de las instancias. Otros métodos como el Bagging (Breiman, 1996), cuyo acercamiento es similar al método que se está describiendo, mejora el patrón aprendido. La técnica de Bagging genera un número finito de nuevos conjuntos de entrenamiento reduciendo el conjunto de datos original, y promedia los resultados obtenidos para crear un único árbol lógico. De esta forma, el patrón aprendido y mostrado en el árbol de decisión es más exacto al reducir los valores de la varianza de la muestra. Debido a que teóricamente no hay un patrón correcto de la colaboración, descartamos esta técnica, ya que podrían filtrar indicadores de interés.

Los árboles de decisión son muy sensibles a pequeños cambios en las instancias de entrada. Debido a esto, si se hacen pruebas con varios conjuntos de datos referidos al mismo problema de clasificación, se obtienen resultados sensiblemente diferentes, debido a los sesgos inherentes a los propios algoritmos de árbol de decisión. Para mitigar esos sesgos, un camino es utilizar más de un algoritmo con el conjunto de todos los datos. Los atributos más relacionados con la clasificación a aprender aparecerían en un mayor número de pruebas, y de esta forma los sesgos inherentes a los propios algoritmos de árbol de decisión se eliminarían. En resumen, los algoritmos de árbol de decisión suministran la relación entre los atributos en relación a la clasificación obtenida, algo que es deseable para saber que indicadores estadísticos están más relacionados con la colaboración, y al utilizar varios es posible eliminar los sesgos que cada uno tiene. Estas han sido las razones para utilizar algoritmos de árbol de decisión y el método estadístico empleado en este acercamiento.

### ***Agrupar alumnos según su colaboración***

Otra forma de alcanzar los objetivos sería encontrar una clasificación de los mismos según su colaboración. Esto nos daría una forma entendible con la que informar a los alumnos de su colaboración.

El procedimiento seguido en este caso ha sido utilizar los algoritmos de agrupamiento o clustering. Estos algoritmos agrupan instancias según la semejanza entre ellas. Para calcular la semejanza utilizan normalmente la métrica euclídea. Una vez calculadas las semejanzas, agrupan las instancias en grupos o cluster. Por esta razón, los algoritmos de clustering no necesitan ser entrenados por lo que los resultados de la clasificación por clustering puede ser obtenida una vez que se tienen las interacciones, sin necesidad de etiquetar instancias. El algoritmo seleccionado ha sido EM debido a sus buenos resultados en otras investigaciones que analizaban la colaboración (Talavera y Gaudioso, 2004; Perera et al., 2007).

La fuente de datos son las intervenciones en los foros, las cuales pueden obtenerse mientras las interacciones se realizan. A partir de las mismas se ejecuta, con los datos recogidos, un algoritmo de clustering para obtener la clasificación de los alumnos.

Finalmente, la clasificación obtenida tiene que compararse con otra fuente de datos para relacionar la clasificación con la colaboración. Esto es lo que se hizo con los datos de los cursos de 2006/2007 y 2007/2008. Avanzamos que se observó una relación entre la clasificación de los alumnos y su colaboración.

Durante el curso 2008/2009 ha sido utilizado este método para obtener información sobre la colaboración. Esta información ha sido mostrada a los alumnos creando así una herramienta metacognitiva. Pero esto es parte del postproceso.

### ***Modelo de colaboración***

La propuesta de esta investigación ha sido la facilitación de herramientas para que el alumno gestione mejor el proceso de colaboración, en vez de proponer una serie de recomendaciones de acuerdo a la colaboración inferida. En un contexto de ED tan abierto y flexible donde el alumno tiene libertad para gestionar su propio proceso, el control se pierde. En estas circunstancias dar una recomendación en el momento adecuado al alumno indicado es una tarea muy complicada. La dificultad de la recomendación radica no sólo en la recomendación como objeto, sino también en el momento y el sujeto al que recomendar (Soller, 2001). Algunas investigaciones que



han propuesto recomendaciones relativas sobre la colaboración plantean un modelo en el cual se especifica qué recomendación, cuándo, y a quién (Baghaei & Mitrovic, 2007; Santos & Boticario, 2008). Estas investigaciones proponen estudios relativos al objeto de la recomendación en vez de abordar el análisis en la colaboración de los estudiantes mientras participan en la experiencia educativa. En esta investigación se ha optado por dar al alumno la información suficiente para que él mismo valore la inferencia y actúe en consecuencia.

Tanto un sistema recomendador como el propuesto necesitan un modelo del alumno centrado en la colaboración. De ahí es donde se extrae la información necesaria. Una de las principales diferencias de los modelos en sistemas recomendadores con el que esta investigación propone es el sujeto que los va a utilizar. Aquí investigación los alumnos y el tutor tendrán acceso a ese modelo, mientras que en otros sistemas sólo los utilizan los motores de inferencia.

Se han demostrado las ventajas de utilizar ontologías en la especificación de modelos en el campo de los sistemas adaptativos en educación (Mizoguchi, 2005; Bull y Kay, 2008; Kay y Lum, 2005). Hay distintos significados y usos del término “ontología” según del dominio o materia que estemos tratando. En el dominio de la Informática<sup>9</sup>:

- El término **ontología** en informática hace referencia a la formulación de un exhaustivo y riguroso esquema conceptual dentro de uno o varios dominios dados; con la finalidad de facilitar la comunicación y la compartición de la información entre diferentes sistemas y entidades. Aunque toma su nombre por analogía, ésta es la diferencia con el punto de vista filosófico de la palabra ontología.

La definición es un poco vaga pero se ajusta concepto de modelar alumnos y sus características, y compartir dicho modelo. En este caso los diferentes sistemas o entidades son, por un lado, la plataforma web y, por otro, los alumnos y el tutor. En conclusión, la ontología guarda y estructura la información que la plataforma web recoge. El uso que se quiera hacer de la información guardada y estructurada depende de los objetivos de la investigación. En entornos educativos, las ontologías ofrecen la posibilidad de rehusar la información según el alumno cambia de entorno (Mizoguchi,

---

<sup>9</sup> [http://es.wikipedia.org/wiki/Ontolog%C3%ADa\\_\(Inform%C3%A1tica\)](http://es.wikipedia.org/wiki/Ontolog%C3%ADa_(Inform%C3%A1tica))

2005), tanto para dicha información sea utilizada por un motor de inferencia o por los propios estudiantes y profesores (Bull y Kay, 2008).

Recordemos que, en el caso de que sean las personas quienes utilicen la ontología, ésta tiene que tener ciertas características para que sea usable y entendible por dichas personas (Bull y Kay, 2008).

Se concluye que una ontología informática es recomendable para ser utilizada como contenedor de un modelo del alumno (Mizoguchi, 2005). Además, la estructuración, claridad y ordenación que aporta una ontología son adecuadas para un modelo abierto del alumno (Bull y Kay, 2008). También aporta la ventaja de compartir con facilidad el modelo estructurado en la ontología con otros sistemas (Heckmann, 2006). La principal decisión técnica ha sido la elección del lenguaje ontológico de la herramienta de autoría. El entorno Protégè<sup>10</sup> ofrece una herramienta ampliamente utilizada y que aporta facilidades para la creación de ontologías en distintos lenguajes. El lenguaje elegido ha sido OWL-LD (Ontology Web Language – Logical Description) por la orientación a la Web semántica y el uso extensivo (Antoniou y van Harmelen, 2004).

Partiendo de estos conceptos, se ha propuesto en la investigación una ontología que almacene y estructure información relativa a la colaboración. La información elegida ha sido: información sobre el contexto de la colaboración (Muehlenbrock, 2005), información del proceso de la colaboración (Dimitracopoulou, 2008) y la información sobre colaboración inferida de los alumnos (Johnson y Johnson, 2004).

- El contexto de la colaboración se entiende como la información que hay alrededor de la colaboración y que es necesario conocer, al menos recomendable, para realizar una colaboración eficiente (por ejemplo, información personal, horarios posibles de comunicación, etc.).
- El proceso de la colaboración se entiende como la información que explica cómo se colabora o qué se ha hecho para colaborar (por ejemplo, información relativa al medio de comunicación/colaboración utilizado).
- En la sección anterior de Inferencia de la colaboración se han explicado dos acercamientos que utilizan algoritmos de aprendizaje automático con el objetivo de inferir, a partir de datos de interacción, la colaboración de los

---

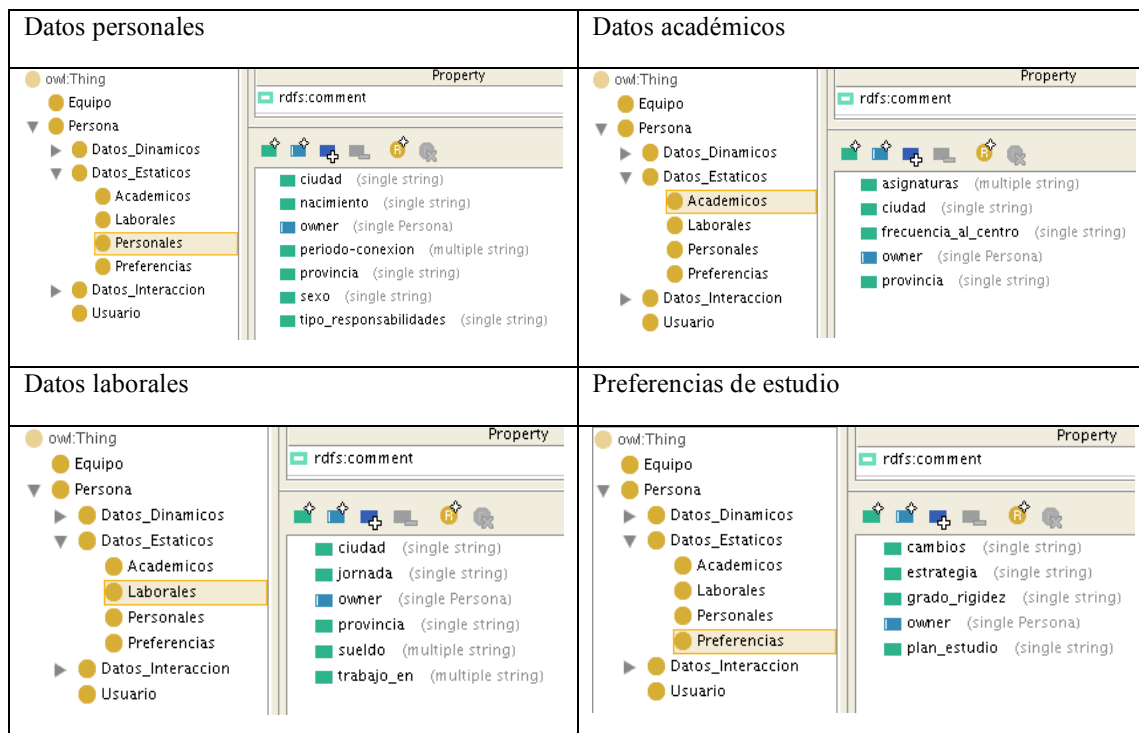
<sup>10</sup> Protégè: herramienta de creación de ontologías (<http://protege.stanford.edu/>)

alumnos. Los resultados de aplicar los acercamientos son lo que aquí llamamos “colaboración inferida” de los alumnos.

Ya que se considera útil la información sobre el contexto de la colaboración (Muehlenbrock, 2005), se tuvo en cuenta la información que los estudiantes solían intercambiar en las primeras conversaciones por los foros. En estas conversaciones solían presentarse. Por esto, en la encuesta inicial se fue preguntando información a los estudiantes y fue la utilizada para extraer la información sobre el contexto de la colaboración. Esta información se dividió en:

- Datos personales: nombre, año de nacimiento, e-mail, lugar de residencia, tipo de responsabilidades.
- Datos académicos: asignaturas que cursa, lugar del centro asociado, frecuencia de visitas al centro asociado.
- Datos laborales: tipo de trabajo, lugar de trabajo, tipo de jornada, sueldo.
- Preferencias de estudio: tipo de plan de estudio, rigidez del plan, estrategia de estudio, causas de cambios en el plan.

En Figura 8 aparecen el esquema general de la ontología y los datos de cada división.



**Figura 8: Atributos de la ontología en las clases de los datos estáticos.**

Como se puede ver en la Figura 8, en la ontología del modelo de la colaboración se almacenaba información bastante completa sobre el contexto de la colaboración. Por

ejemplo, se almacenaba información sobre la empresa donde trabajaba, dónde residía dicha empresa, tipo de jornada laboral, etc.

Como información del proceso de colaboración se han utilizado los indicadores estadísticos de la colaboración de los alumnos. Los indicadores estadísticos ya se explicaron anteriormente.

Los indicadores estadísticos utilizados han sido: num\_thrd, num\_msg, num\_reply\_thrd y num\_reply\_msg. Se pensó que 12 indicadores estadísticos eran demasiados y confundirían al alumno, por lo que se decidió presentar sólo los básicos, de donde los demás se calculan. Estos indicadores se han guardado en la ontología de dos formas, una que indica el valor de los indicadores desde el inicio hasta el final, y otra que señala el valor de los indicadores por semana.

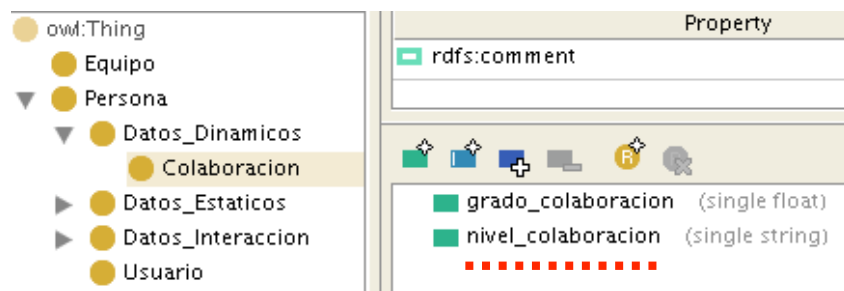
En la Figura 9 se muestran los indicadores estadísticos en la ontología.



**Figura 9: clases de la ontología que informan sobre las interacciones en los foros.**

Como se puede ver, los nombre de los atributos eran autoexplicativos. Por ejemplo, el indicador estadístico “num\_reply\_msg” era representado en la ontología por el atributo “num\_replicas\_a\_mensajes”.

El método de inferencia utilizado en esta investigación ha sido el de clustering. Ha sido elegido por los buenos resultados obtenido con los datos de los cursos 2006/2007 y 2007/2008. En la sección de resultados se explicará en profundidad estos resultados. Se ha elegido un número de 3 cluster porque facilitan su identificación y entendimiento del significado de los mismos. En este caso, los clusters podían significar un nivel de colaboración bajo, medio o alto. En la Figura 10 se indica cómo aparecen en la ontología.



**Figura 10: clase de la ontología que revela la colaboración inferida de los alumnos.**

El nivel de colaboración era almacenado en el atributo “nivel\_colaboracion”. El otro atributo que aparece, “grado\_colaboracion”, no fue utilizado ya que el método de inferencia seleccionado para la experiencia del curso 2008/2009 fue el de clustering que sólo ofrecía el nivel de colaboración, pero la ontología del modelo de la colaboración se preparó para algún otro método de inferencia.

Como se ha podido observar en las Figuras 8, 8 y 10, las clases de la ontología se organizan en forma jerárquica. Bajo la clase raíz *Thing*, obligatoria en todas las ontologías creadas en Protégè, cuelgan dos subclases, la clase equipo, la cual contiene a los alumnos agrupados por *Equipo*, y la subclase *Persona*. Bajo la clase *Persona* cuelgan las clases *Datos\_Dinamicos*, *Datos\_Estaticos* y *Datos\_Interaccion*. Bajo la clase *Datos\_Dinamicos* cuelga la clase *Colaboracion* que contiene los datos sobre la colaboración de los alumnos. Bajo la clase *Datos\_Estaticos* cuelgan las clases *Personales*, *Academicos*, *Laborales* y *Preferencias*. Bajo la clase *Datos\_Interaccion* cuelga la información relativa a los indicadores estadísticos de los foros. En las figuras sobre los indicadores estadísticos se ve que también hay unas clases para las interacciones en los Chat, pero esto no ha sido utilizado durante el curso.

La clase *Datos\_Estaticos* contiene datos del alumno que no cambian con el tiempo. La clase *Datos\_Interaccion* contiene datos de las interacciones del alumno, en este caso sólo en los foros. La clase *Datos\_Dinamicos* contiene datos que sí cambian o pueden cambiar con el tiempo, es decir, con las interacciones o acciones del alumno.

Hay una clase *Usuario* que cuelga de la clase *Persona* que enlaza o centraliza mediante propiedades la distinta información que contienen las otras clases. Como el objetivo ha sido mostrar esta ontología que representa el modelo de colaboración del alumno, se pensó que centralizar toda la información desde una clase favorecía la navegación y la usabilidad (Nielsen, 2000). En la Figura 11 se puede ver toda la estructura jerárquica de la ontología y los contenidos de la clase *Usuario*.

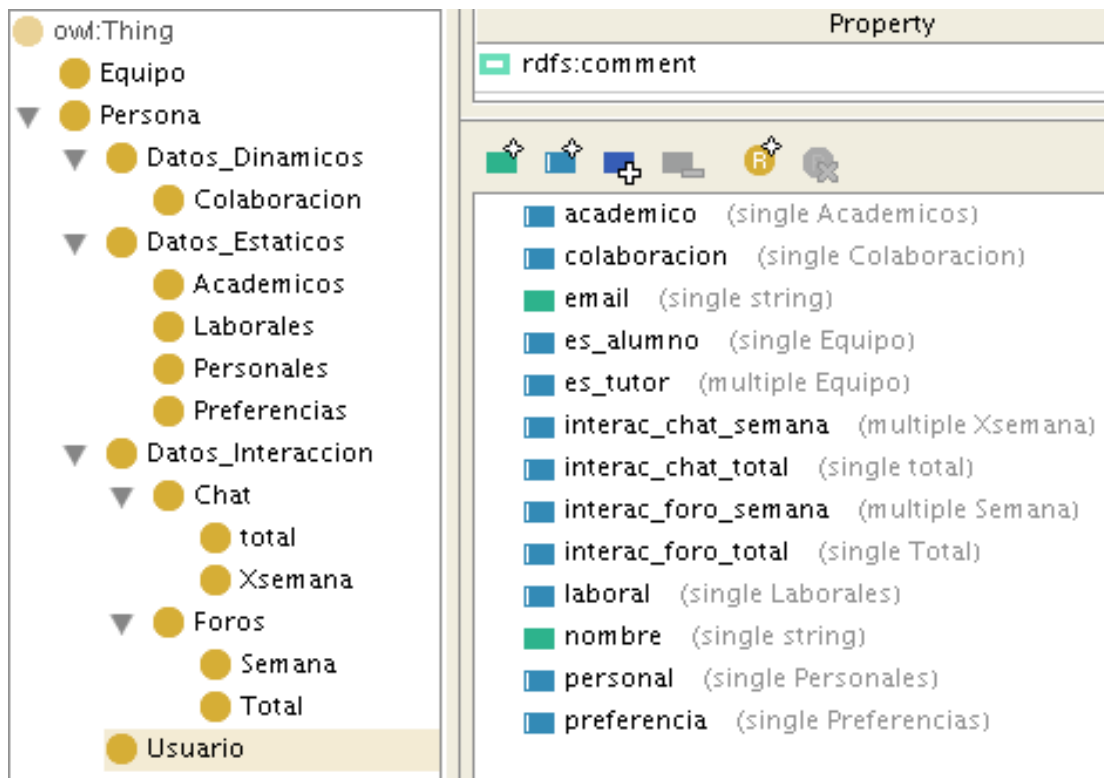


Figura 11: atributos y propiedades de la clase Usuario.

Las propiedades en esta ontología funcionan como enlaces unidireccionales. Como se ve la Figura 11, desde la clase *Usuario* se enlazan el resto de clases. Para favorecer la navegación, hay propiedades que enlazan desde las clases que contienen datos a la clase *Usuario*, pero esta clases tienen una característica especial. Las propiedades de la ontología que enlazan las clases contenidas en *Datos\_Estaticos* con la clase *Usuario* se han nombrado **owner**. De esta forma se puede saber quién es el propietario de los datos que contienen la clase de origen. Cuando se describa la aplicación web escrutable, se explicará la importancia de esta propiedad.

En el

Apéndice II: ontología de la colaboración se encuentra el código de la ontología del modelo de la colaboración de los alumnos.

### *Análisis crítico*

Lo explicado en esta sección contiene un modelo de la colaboración de un estudiante dividido en tres partes: contexto, proceso e inferencia sobre la colaboración. El modelo ha nacido con la vocación de ser abierto, es decir, usable y entendible por los propios alumnos (Bull y Kay, 2008). Además, el modelo ha sido

unido a la ingeniería ontológica por los beneficios que esto conlleva respecto a claridad, estructura y reutilización (Mizoguchi, 2005).

Hay que aclarar que el uso de ontologías en educación ha sido empleado en su mayoría como forma de expresar los objetos que forman parte del contexto educativo (Barros et al., 2002; Villasclaras-Fernandez et al., 2009), antes que en expresar las características de los estudiantes. Esto último forma parte del modelado del usuario o alumno y su uso se está empezando a generalizar recientemente (Kyriacou et al., 2009; Heckmann, 2006).

Como se ha visto al enmarcar el modelado de la colaboración, no hay acuerdo ni estudios comparativos en este asunto. El modelado de la colaboración es un punto abierto del campo y nosotros lo hemos abordado aportando un modelo con ciertos atributos para abarcar aquello que es importante para colaborar según (Muehlenbrock, 2005), (Dimitracopoulou, 2008) y (Johnson y Johnson, 2004).

Creemos que la solución de este punto requiere un trabajo multidisciplinar orientado al estudiante donde se estudie qué es apropiado para modelarlo en un entorno de aprendizaje colaborativo y cómo obtener dicha información.

### **Herramientas metacognitivas**

En el apartado de trabajos relacionados con la escrutabilidad y autorregulación se han descrito distintas teorías que hablan de las ventajas en el aprendizaje de utilizar la autorregulación y, por tanto, la adecuación de los sistemas que ofrecen herramientas de autorregulación. Ésta ha sido una de las ideas principales de esta investigación, ofrecer una herramienta que ayude a la autorregulación del alumno en un entorno de enseñanza colaborativo, considerando que para esto lo más conveniente es una herramienta que ayude a la mejor comprensión del proceso de colaboración (Boekaerts et al., 2000; Dimitracopoulou, 2008).

Muchas investigaciones han planteado gran variedad de herramientas para la mejora del aprendizaje (Bratitsis, 2009; Santos et al., 2003; Daradoumis et al., 2006, Barkley et al., 2004). Estudiando la literatura se descubren distintos tipos de herramientas que pueden ser útiles: herramientas metacognitivas, de guiado (Soller et al., 2005), y escrutables (Kay, 1999). Las investigaciones que han propuesto herramientas de guiado o de recomendación en entornos de aprendizaje educativo, muestran que el análisis realizado se centra en la recomendación a ofrecer (Baghaei y

Mitrovic, 2007; Santos y Boticario, 2008). Estas investigaciones proponen un modelo de las recomendaciones. Baghaei y Mitrovic (2007) ofrecen un modelo basado en restricciones mientras que Santos y Boticario (2008) sugieren utilizar distintos tipos de filtrado colaborativo para encontrar distintos ítems a recomendar. Otra investigación (Goodman et al., 2005) analiza las conversaciones en el Chat mediante un modelo preestablecido y propone recomendaciones para mantener la conversación dentro de los límites de los contenidos de un curso. Esta investigación analiza un medio de comunicación pero no la colaboración. Ninguna de las investigaciones se centran en analizar la colaboración en el entorno, algo que esta tesis considera necesario para asegurar que el aprendizaje colaborativo sucede (Johnson y Johnson, 2004).

La opción más significativa es la herramienta metacognitiva que ya se ha relacionado con las herramientas de autorregulación. Se ha discutido en la sección anterior sobre el tipo de información más adecuada para alcanzar los objetivos de la investigación. Sin embargo, también hay que tener en cuenta la intrusividad y la privacidad (Schreck, 2003). Una forma intrusiva de adquisición de información son las encuestas, formularios, interrogatorios, entrevistas, etc. En esta investigación se han pasado varias encuestas. La encuesta inicial tenía más de 20 preguntas y no se buscaba más que información general sobre el contexto de la colaboración. Ya se ha investigado la conveniencia de no “cansar” al alumno con un enorme número de preguntas ya que sus respuestas pierden veracidad (McFarland, 1981). También hay que tener en cuenta la privacidad. En Internet todo puede ser público o publicable. Por eso mismo, la mayor parte de las preguntas sobre de la encuesta inicial eran de respuesta voluntaria. Se ha discutido sobre la privacidad, necesaria en los entornos adaptativos pero sin menguar la capacidad de adaptación (Kobsa, 2007).

Una estrategia que da la posibilidad de gestionar al usuario su propia privacidad sin perder adaptación ni capacidad de autorregulación es la escrutabilidad. Sin embargo, es una estrategia más intrusiva ya que exige del usuario una gestión, es decir, la herramienta pide más trabajo del usuario que otra que sólo mostrase la información.

Nos encontramos con dos alternativas: una herramienta metacognitiva, o escrutable. Por esta razón no hemos ofrecido en esta investigación una herramienta a



los estudiantes, sino varias para comprobar la utilidad de los distintos acercamientos. Por un lado, hemos ofrecido distintas herramientas metacognitivas de autorregulación con diferencia en la información que mostraban. De esta forma se pretende identificar el tipo de información más adecuado para el objetivo de mejorar el aprendizaje colaborativo. Además hemos ofrecido también una herramienta escrutable de autorregulación que mostraba toda la información del modelo de colaboración de un modo escrutable. Al comparar los resultados de los estudiantes a los que se les ofreció la herramienta escrutable con el resto, se podrá establecer una cierta relación entre la escrutabilidad y el aprendizaje colaborativo.

En resumen, se ha ofrecido a unos estudiantes una aplicación web que es una herramienta escrutable que muestra todo el modelo de colaboración, a otros tres grupos de estudiantes se les ha ofrecido una, y sólo una a cada grupo, de las herramientas metacognitivas creadas mediante portlets o pequeñas ventanas enmarcada en una de las secciones de la plataforma dotLRN. Estos portlets mostraban: *información sobre las interacciones*, *información sobre el proceso de colaboración*, o *ambos tipos de información*.

### ***Aplicación web***

La aplicación web es un desarrollo basado en el plugin de Protège Web Browser<sup>11</sup>. El plugin se pensó como un medio de crear ontologías a través de la web, por lo que replica la funcionalidad de la versión 3.0 de Protège. Sin embargo, se ha utilizado dicho plugin por la facilidad de presentación y navegación. La aplicación web ha consistido en una personalización y adaptación del plugin a los objetivos de esta investigación.

Lo que no se ha utilizado y se ha eliminado de la aplicación web han sido las facilidades del plugin para crear y gestionar instancias de la ontología, y las búsquedas. Se ha añadido una gestión de los contenidos para identificar al usuario al que se referían los datos. Además, se ha cambiado la apariencia para adaptarla al contexto educativo de la UNED.

En detalle, la aplicación web consiste en una página de inicio donde se pide el nombre y la contraseña del usuario que va a entrar en la aplicación. Se pensó que un

---

<sup>11</sup> [Protège Web Browser](#)

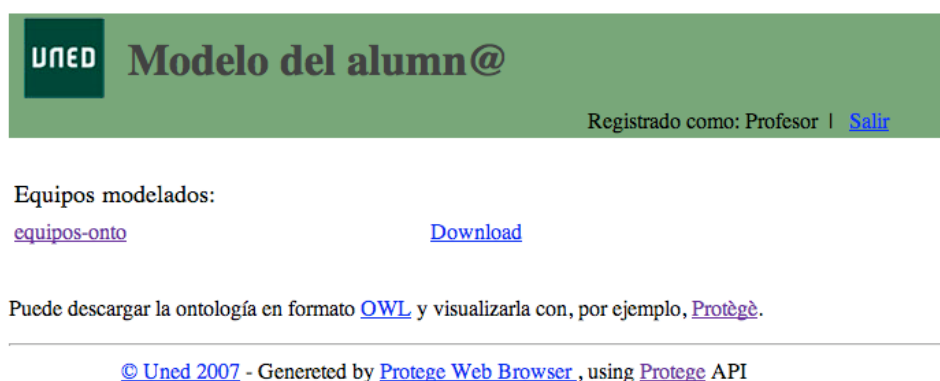
control de los usuarios era necesario para permitir a los dueños de la información editar dicha información. En la Figura 12 se muestra la pagina inicial.



The screenshot shows a green header bar with the UNED logo on the left and the text 'Modelo del alumn@' in the center. Below the header, there is a login form with two input fields: 'Nombre:' and 'Contraseña:'. A 'Login' button is positioned below the password field.

**Figura 12: pantalla de identificación del usuario de la aplicación web.**

Una vez identificado el usuario, se le daba entrada a una página donde podía descargarse la ontología del modelo o entrar a navegar en ella.



The screenshot shows the user dashboard for 'Modelo del alumn@'. The header bar is green and contains the UNED logo, the text 'Modelo del alumn@', and a user status indicator 'Registrado como: Profesor | [Salir](#)'. Below the header, there is a section titled 'Equipos modelados:' with a link to 'equipos-onto' and a 'Download' button. A note below states: 'Puede descargar la ontología en formato [OWL](#) y visualizarla con, por ejemplo, [Protégè](#).' At the bottom, there is a footer: '© Uned 2007 - Genereted by [Protege Web Browser](#) , using [Protege](#) API'.

**Figura 13: página de entrada en le modelo de colaboración.**

Como se observa en la parte superior derecha de la Figura 13, el usuario ya ha sido registrado, en este caso el usuario *Profesor*. Pulsando sobre en enlace del centro-izquierda se entra en la aplicación en sí.

The screenshot shows the UNED EQUIPOS-ONTO application interface. At the top, the logo 'UNED' is on the left, 'EQUIPOS-ONTO' is in the center, and 'Tu usuario es : Profesor | Salir' is on the right. The main content is divided into three vertical panels:

- Left Panel:** A hierarchical tree of classes. The 'Usuario' class is highlighted in green, indicating it is the current view. Other classes include Owl:Thing (2,0), Equipo (0,9), Persona (4,0), Datos\_Dinamicos (1,0), Datos\_Estaticos (4,0), Datos\_Interaccion (2,0), and Usuario (0,28).
- Center Panel:** A list of instances under the heading 'Muestras'. The instance 'alejandro-donoso-86' is highlighted in green. Other instances include Profesor, alfonso-sancho-93, antonio-garcia-72, antonio-perez-64, cristina-fletes-32, daniel-navarro-84, david-oter-29, francisca-haro-71, francisco-jose-99, francisco-miguel-08, ivan-sanchez-86, javier-banon-66, javier-jimenez-98, jon-gonzalez-74, jose-ignacio-12, jose-luis-28, lorena-prida-85, miguel-pastor-02, nuria-fuentes-11, and oliver-reves-13.
- Right Panel:** A detailed view of the 'alejandro-donoso-86' instance. It shows the following data:
  - Instance of:** Usuario
  - email:** alejandro.donos@gmail.com
  - nombre:** Alejandro Donoso Franco
  - academico:** Academicos\_alejandro-donos-86
  - colaboracion:** Bajo
  - es\_alumno:** Equipo\_32
  - interac\_foro\_semana:**
    - Semana\_2008-12-22-alejandro-donos-86
    - Semana\_2008-12-01-alejandro-donos-86
    - Semana\_2008-12-29-alejandro-donos-86
    - Semana\_2008-11-03-alejandro-donos-86
    - Semana\_2008-11-24-alejandro-donos-86
    - Semana\_2008-11-17-alejandro-donos-86
    - Semana\_2009-01-05-alejandro-donos-86
    - Semana\_2009-01-12-alejandro-donos-86

At the bottom of the screenshot, there is a footer: '© Uned 2007 - Generated by Protege Web Browser, using Protege API'.

**Figura 14:** pantalla que muestra los datos de una muestra de la clase Usuario.

Como se puede ver en el marco izquierdo de la Figura 14, la aplicación muestra una estructura jerárquica de las clases que contienen la información tanto sobre el contexto de la colaboración, como del proceso de la colaboración y los niveles de colaboración inferidos del usuario. La clase por la que está navegando se marca en verde, en este caso la clase *Usuario*. En el marco central se ven las instancias que la clase *Usuario* contiene, en este caso los usuarios a los que está permitido la navegación por la aplicación. La instancia por la que el usuario está navegando en esos momentos se encuentra marcada en verde (alejandro-donos-86). El marco de la derecha muestra los datos que contiene la instancia, en este caso, los referidos al contexto como el e-mail, el nombre (Alejandro Donoso Franco), el equipo del alumno (Equipo\_32), un enlace a los datos académicos, los referidos al proceso como los enlaces a los datos estadísticos de interacción por semana del alumno, y los referidos a la información inferida, es decir, el nivel de colaboración (Bajo). Un usuario que sea el dueño de la información, es decir, a quien se refiere la información, tiene la posibilidad de editar la información. En la Figura 15 se muestra esta pantalla.

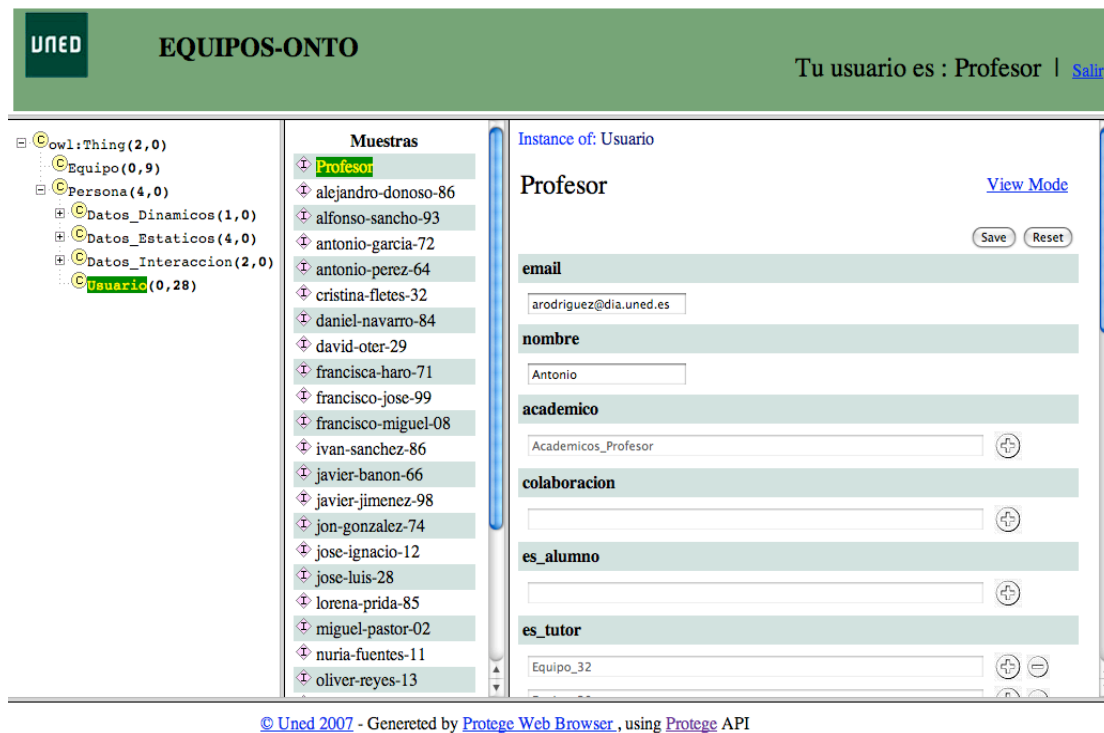


Figura 15: pantalla de edición de los datos de la muestra Profesor.

Se observa en la Figura 15 que el usuario **Profesor**, marcado en verde en el marco central, está navegando por sus propios datos, como se ve en el nombre del marco de la derecha. El usuario pulsó sobre el enlace de editar información por lo que en el marco de la derecha aparece un formulario con los datos y con la posibilidad de cambiarlos. Por ejemplo, su correo electrónico ([arodriguez@dia.uned.es](mailto:arodriguez@dia.uned.es)) o su nombre (Antonio). El dueño de los datos es identificado por el valor de la propiedad de la ontología **owner**, de la que ya se habló antes. Este valor es comparado con el usuario. A aquel usuario que es el mismo que el propietario, indicado por el valor de la propiedad **owner** le aparece un enlace en el marco de la derecha que poder editar los datos.

Se advierte que la aplicación web se hizo independiente de la plataforma dotLRN por lo que el usuario debía seleccionar pasar por dos pantallas antes de ver sus datos, como se ha visto en las figuras de esta sección.

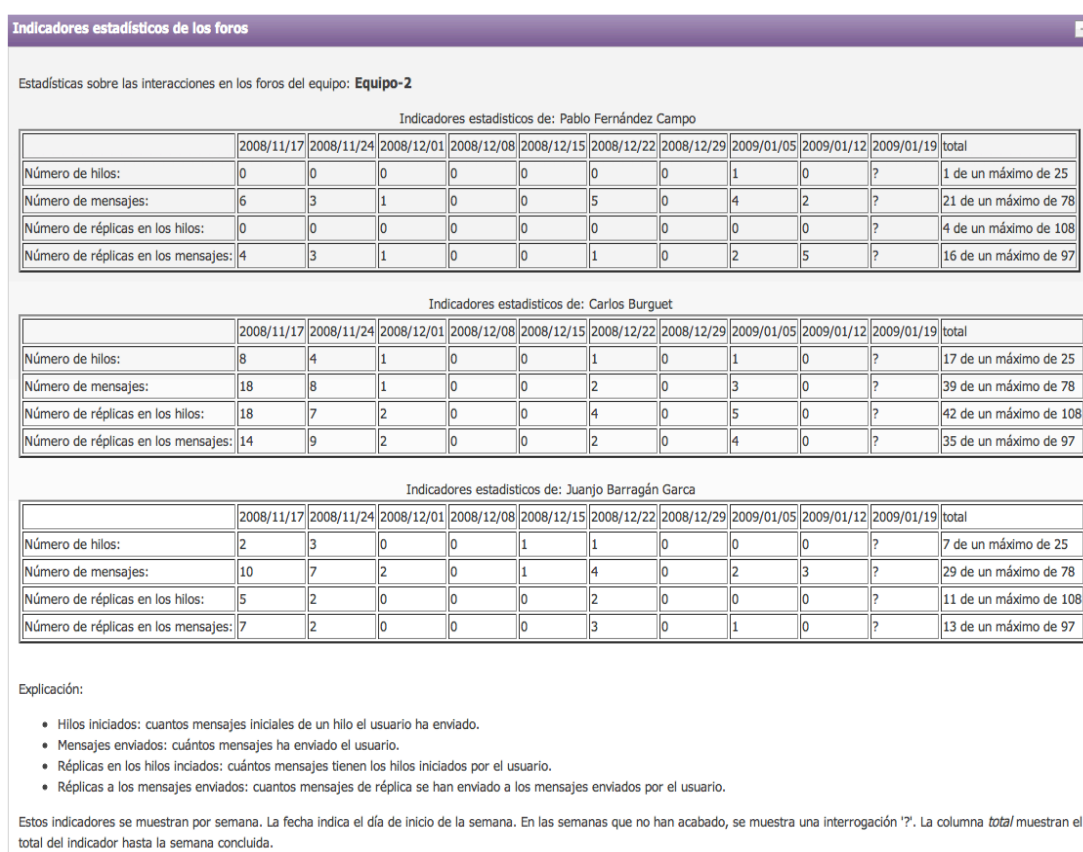
### ***Portlet sobre interacciones y colaboración***

Siguiendo la perspectiva de mostrar información escueta y esquemática de la manera menos intrusiva, se pensó en los indicadores estadísticos sobre las interacciones de los alumnos en los foros y el nivel de colaboración inferido. En

realidad la información que se mostró era la misma que la mostrada en la aplicación web salvo la del contexto de la colaboración.

Esta información se colocó en un portlet de dotLRN accesible desde una de las secciones del espacio privado del equipo. La información de este portlet ha sido actualizada una vez por semana todos los lunes durante la experiencia colaborativa.

Por un lado, la información presentada era los indicadores estadísticos de las interacciones a los foros. En la Figura 16 aparece el portlet.



**Figura 16: portlet de las estadísticas de interacción en los foros.**

En la Figura 16 aparecen tres tablas. Cada tabla contiene los datos de interacción de un miembro del equipo. Las filas representan un indicador estadístico distinto y las columnas una semana de la 2ª fase de la experiencia de aprendizaje colaborativo. Cada celda es el valor del indicador estadístico (según la fila) en una semana (según la columna). En la parte inferior se aclara la información mostrada en las tablas, explicando lo que significan los indicadores.

En lo relativo al nivel de colaboración, es decir, la información inferida por el método de inferencia, se creó otro portlet donde plasmar la información. El portlet que se presentó se muestra en la Figura 17.

**Información sobre el trabajo en equipo**

**Nivel de colaboración inferido por las interacciones**

Se ha inferido el nivel de colaboración de los miembros del equipo según las interacciones realizadas en los foros de esta plataforma. Se ha utilizado herramientas de Inteligencia Artificial para agrupar a todos los alumnos según su colaboración. El objetivo es dar más información sobre el trabajo en equipo para que el alumno pueda actuar en consecuencia, por lo que no es una evaluación. Se advierte que los resultados son aproximados. El nivel de colaboración inferido puede ser:

- Alto, lo que significa que el/la alumn@ en cuestión trabaja aproximadamente mejor en equipo que el resto de alumnos de esta experiencia.
- Medio, lo que significa que el/la alumn@ en cuestión trabaja aproximadamente igual en equipo que el resto de alumnos de esta experiencia.
- Bajo, lo que significa que el/la alumn@ en cuestión trabaja aproximadamente peor en equipo que el resto de alumnos de esta experiencia.

Los resultados son:

Alumn@	Nivel de colaboración
Cecilia Gómez Jáuregui	Medio
Javier Palacio Val	Medio
Jon Joseba Zurutuza Salsamendi	Alto

**Figura 17: portlet que muestra el nivel de colaboración inferido de los alumnos.**

Como se puede ver en la Figura 17, a los miembros del equipo se les informaba del nivel de colaboración inferido. Por ejemplo, el alumno Javier Palacio Val se le asignó un nivel de colaboración medio. Todos los miembros del equipo sabían el nivel de colaboración propio y el de los compañeros. Además del nivel de colaboración, a los estudiantes se les informaba escuetamente (ver parte superior de la Figura 17) de la forma de cálculo del nivel de colaboración y sus posibles valores (alto, medio o bajo).

### ***Análisis crítico***

Al estudiar otras investigaciones que ofrecieron herramientas para la mejora del aprendizaje y de la colaboración, no se encontró una guía muy clara sobre la herramienta con la que alcanzar los objetivos de la investigación después de analizar las interacciones e inferir juicios sobre la colaboración de los estudiantes. Han quedado claras las ventajas de la autorregulación en la instrucción e incluso en las actividades que implican relaciones sociales (Steffens, 2001; Boekaerts et al., 2000). Las herramientas de autorregulación imponen ciertas condiciones (Zimmerman, 1990), las cuales son compatibles con una herramienta metacognitiva (Soller et al., 2005). La información metacognitiva se encuentra después de inferir juicios sobre la colaboración (Talavera y Gaudioso, 2004).

La estrategia de las recomendaciones implica un análisis de la colaboración y un análisis de la propia recomendación (qué, cuándo y cómo). Esto entraña grandes dificultades en un entorno dinámico como son los colaborativos. Los sistemas que han propuesto recomendaciones han modelado la propia recomendación en vez de la colaboración del alumno (Baghaei y Mitrovic, 2007; Santos y Boticario, 2008), lo que reduce la complejidad del proceso pero no realizan ningún análisis metacognitivo.

Se ha elegido la herramienta metacognitiva pero al tener distinta información sobre la colaboración de los estudiantes (del contexto, del proceso, inferencias), se han propuesto distintas herramientas para identificar el tipo de información más útil.

También se ha tenido en cuenta la estrategia escrutable (Kay, 1999) como herramienta de autorregulación. Se ha discutido las ventajas de esta estrategia tanto como herramienta de autorregulación, para aumentar la responsabilidad y motivación debido a la posibilidad de gestión de sus contenidos, como de privacidad por la propia posibilidad de gestión. Sin embargo, esta estrategia aumenta la intrusividad de la herramienta ya que el usuario tiene que realizar el esfuerzo de gestionarla. Por estas razones se ha propuesto, además, una herramienta escrutable.

En resumen, las cuatro herramientas han sido una aplicación web y tres ventanas de contenidos que se podían añadir en cualquier sección del espacio virtual, y que llamaremos Portlet:

- *Aplicación web* mostraba información sobre el contexto de la colaboración, el proceso de la colaboración y la inferencia de la colaboración desde un punto de vista escrutable (i.e., monitorización y metacognitiva).
- Protlet I, ventana pequeña enmarcada en una página que mostraba escuetamente sólo la información del proceso de la colaboración (i.e., monitorización).
- Protlet II, ventana pequeña enmarcada en una página que mostraba escuetamente sólo la información inferida de la colaboración de los alumnos (i.e., metacognitiva).
- Portlet III, ventana pequeña enmarcada en una página que mostraba escuetamente información del proceso de la colaboración y la inferida de la colaboración de los alumnos (i.e., monitorización y metacognitiva).

Durante la experiencia de aprendizaje colaborativo del curso 2008/2009 se han propuesto a distintos equipos las herramientas anteriores. Se ha experimentado para

descubrir el tipo de información más útil y la estrategia más apropiada. Para ello se ha organizado la experimentación realizada de forma que a algunos de los equipos se les ofreció una sola de las herramientas, pero dejando un conjunto de equipos sin ninguna para así poder comparar las posibles diferencias en el aprendizaje colaborativo alcanzado.

### **Evaluación**

En este apartado no sólo planteamos la evaluación del estudiante de la experiencia de aprendizaje colaborativo, sino de la propia experiencia y de las adaptaciones realizadas para mejorar el aprendizaje colaborativo o cerciorarse de que ocurre.

Para evaluar el conocimiento y el trabajo de los estudiantes hay gran cantidad de distintos tipos de pruebas o exámenes, examen de desarrollo, de tipo test, trabajos prácticos, cuestionarios, etc. Hay mucha literatura sobre el uso de las pruebas de evaluación (Viñao, 2004), y de lo conveniente o no para medir el conocimiento (Merchán, 2005). En este sentido nos tenemos que dejar guiar por la metodología de la UNED en el caso de las pruebas de evaluación.

Sin embargo, para evaluar la adaptación ofrecida en la experiencia hemos de recurrir a ejemplos de otras investigaciones y a la literatura especializada en la evaluación de sistemas adaptativos. Velsen et al. (2006) inciden claramente en que la evaluación debe centrarse en los usuarios del sistema, y requiere utilizar más de un medio de evaluación para evitar sesgos innecesarios. Ha habido investigaciones que han utilizado este marco de trabajo (Santos y Boticario, 2008), pero el principal problema al que se enfrenta este planteamiento es el número de medios de evaluación. No se puede obligar a los estudiantes de un entorno de aprendizaje real que respondan a un número ilógico de pruebas de evaluación (Merchán, 2005).

En la experiencia de aprendizaje colaborativo se ha pedido una memoria del trabajo realizado durante el periodo colaborativo como medio de evaluación. También se ha solicitado el código desarrollado para verificar que los problemas de planificación hayan sido resueltos.

La memoria del trabajo realizado, la cual se pedía en la tarea 5, recogía las labores del equipo completo. Cada una de las tareas del periodo colaborativo se veían reflejadas en la memoria y era posible evaluar cada una de las tareas de la experiencia. Sin embargo, es difícil percatarse del esfuerzo colaborativo de cada uno



de los integrantes del equipo. Para ello el tutor tuvo en cuenta los mensajes enviados a los foros de los equipos para identificar el esfuerzo colaborador de cada uno de los miembros.

El trabajo de la experiencia de aprendizaje colaborativo fue evaluado entre un 2 (máximo) y un 0 (mínimo), y representaba el trabajo colaborativo de cada uno de los componentes de los equipos, por lo que la nota no fue siempre la misma para los miembros de un mismo equipo.

La evaluación del conocimiento aprendido se realizó con las pruebas de evaluación presencial que establece la UNED para todas las asignaturas de los estudios reglados. Entre finales de enero y principios de febrero de cada año, la UNED organiza las pruebas de evaluación presenciales del primer parcial. El examen que se propone a los estudiantes de las asignaturas tiene una duración máxima de 2 horas y el equipo docente de cada asignatura se encargar de pedir los conocimientos apropiados para el contenido enseñado. En el caso de la asignatura de IA-IC, el examen se evaluaba en un máximo de 10 puntos y, durante los cursos donde se ha ofrecido la experiencia de aprendizaje colaborativo, los contenidos pedidos se dividían principalmente entre planificación y un problema práctico en lenguaje CLISP. La parte de planificación se dividía entre 2 puntos para una o dos preguntas teóricos y el resto para un ejercicio práctico. La fuente principal de evaluación del conocimiento de los estudiantes ha sido en examen del primer parcial.

Con la memoria y el examen del primer parcial se ha evaluado el conocimiento de los estudiantes. Sin embargo, en esta experiencia hay más cosas que evaluar, la propia experiencia y las herramientas que proponen una adaptación.

Para evaluar la propia experiencia de aprendizaje colaborativo se han utilizado dos medios: el examen del primer parcial comparando a estudiantes que participaron en la experiencia con los que no lo hicieron, y las respuestas a las encuestas finales donde se pedían opiniones sobre la experiencia.

Para evaluar la adaptación ofrecida se han utilizado tres medios: el examen del primer parcial comparando a estudiantes a los que se les ofreció una determinada herramienta de autorregulación con los que no se les ofreció ninguna, la evaluación de la memoria del trabajo colaborativo de los estudiantes a los que se les ofreció alguna de las herramientas de autorregulación con los que no se les ofreció ninguna, y las

opiniones a las encuestas finales ofrecidas sólo a aquellos equipos que pudieron utilizar alguna de las herramientas de autorregulación.

### ***Análisis crítico***

Para realizar la evaluación se ha propuesto una metodología centrada en el estudiante donde se han utilizado distintos medios de evaluación. En esta investigación ha sido necesario evaluar la propia experiencia de aprendizaje colaborativo en los cursos 2006/2007, 2007/2008 y 2008/2009, y las adaptaciones ofrecidas en el curso 2008/2009. En la Tabla 3 se resume los tipos de evaluación y los medios.

**Tabla 3: resumen de las evaluaciones realizadas.**

	Cursos			Examen 1er parcial	Trabajo colaborativo	Encuesta final de la experiencia	Encuestas finales sobre herramientas
	06/07	07/08	08/09				
Experiencia colaborativa	X	X	X	X		X	
Herramientas de autorregulación			X	X	X		X

# P A R T E I I I : E X P E R I M E N T A C I Ó N

La parte tercera de la tesis está dividida en dos partes o capítulos. Por un lado el capítulo IV describe los datos obtenidos en los experimentos realizados dando un pequeño análisis de los resultados, y el capítulo V profundiza en el análisis con la intención de validar las propuestas examinadas en la experimentación mediante los datos obtenidos.

## Capítulo IV: desarrollo y resultados

Partiendo del marco conceptual, los modelos y herramientas descritos en el capítulo anterior, en este capítulo se mostrará el desarrollo y los resultados de la experimentación realizada para validar las hipótesis establecidas en esta tesis. La mejora del aprendizaje colaborativo se ha basado en una serie de hipótesis y se han definido una serie de objetivos. Se quiere validar que una análisis automático o con la mínima intervención de expertos es posible y que aporta información sobre la colaboración mientras el proceso de colaboración ocurre, que se puede repetir regularmente, que puede ser genérico e independiente del dominio para tener más posibilidades de ser llevado a otros entornos de aprendizaje colaborativo, y que dicho análisis aporta información que el estudiante y tutores de la experiencia de aprendizaje colaborativo pueden utilizar para mejorar la gestión del proceso de colaboración y el aprendizaje colaborativo.

La experimentación ha abarcado el análisis de las interacciones, el desarrollo del método de inferencia y del modelo de colaboración, como partes del modelado o análisis de la colaboración, y una serie de herramientas de autorregulación que han suministrado la información debida a los análisis y el modelo. Estas herramientas también han sido evaluadas para validar los objetivos y las hipótesis de la tesis.

Se hace notar que en el proceso de modelado o análisis se mostrarán los datos obtenidos de los cursos 2006/2007, 2007/2008 y 2008/2009. Con estos datos se discutirá la validez de la propuesta. La evaluación de las herramientas de autorregulación se ha realizado sólo durante el curso 2008/2009. En ese curso ya se tenía un método de inferencia y un modelo validado y con la información suministrada se han creado las herramientas de autorregulación.

La duración de la parte experimental ha provocado que se fueran obteniendo datos antes de finalizar. Los datos obtenidos sugirieron que un método de inferencia de la colaboración correcto era el método que clasificaba mediante técnicas de clustering a los estudiantes (Anaya y Boticario, 2009a). Esta técnica ha sido utilizada para las valoraciones sobre la colaboración que fueron mostradas durante el curso 2008/2009.

Los datos del último curso respecto a la clasificación mediante clustering fueron comparados con los anteriores obteniendo unos resultados semejantes, lo que daba soporte a la decisión realizada (Anaya y Boticario, 2009c).

Durante los tres cursos se fue preparando distintos tipos de información para la construcción del modelo de la colaboración. Este trabajo ha soportado una investigación con la cual se ha obtenido el modelo de la colaboración presentado en esta tesis (Anaya y Boticario, 2009b).

Este capítulo describe el análisis de las interacciones realizado, de los resultados obtenidos de los dos métodos de inferencia propuestos, se muestran los resultados de la evaluación de las herramientas de autorregulación, y termina mostrando unos resultados generales de la propia experiencia de aprendizaje colaborativo durante los cursos que se ha aplicado comparándolos con resultados de cursos pasados, cuando no se ofreció la experiencia de aprendizaje colaborativo.

## **Análisis de las interacciones**

Con el análisis de las interacciones propuesto se hace posible realizar el seguimiento del usuario en el sistema. Esta estrategia no es intrusiva al recoger las interacciones sin pedir otro trabajo al usuario. Esto se puede hacer mientras el usuario interactúa con el sistema por lo que los resultados son obtenidos mientras el proceso de interacción ocurre. Además, se puede aplicar en gran cantidad de entornos mientras se cumplan ciertas condiciones tecnológicas (acceso a base de datos, ficheros log del servidor web, etc.), además de ser independiente del dominio (asignatura, tutores y sus contenidos) (Dimitracopoulou, 2008). Por lo tanto, el análisis de las interacciones cumple los objetivos de la tesis ya que obtiene información de las interacciones mientras estas ocurren y puede ser transferido a otros entornos.

Se ha utilizado el análisis de las interacciones para obtener indicadores estadísticos de las interacciones de los estudiantes en los foros de los espacios privados de los equipos, que era donde se producía la colaboración en gran medida. En el capítulo anterior, que introdujo el modelo educativo de la experiencia de aprendizaje colaborativo, se presentaron 12 indicadores estadísticos. Con los indicadores se

crearon datasets, o datos en formato de tabla, por curso, donde cada fila son los indicadores estadísticos de las interacciones de un alumno. Por el tamaño de estos datasets no aparecen en esta tesis pero pueden ser solicitados al autor de la misma por correo electrónico (arodriguez@dia.uned.es).

Teniendo en cuenta que el objetivo de la investigación ha sido mejorar el control sobre el proceso de colaboración hallando un método de inferencia con el que obtener valoraciones sobre la colaboración de los estudiantes, el análisis de las interacciones ha sido usado como fuente de datos para el método de inferencia, el cual utiliza tecnología de aprendizaje automático, ya que esta tecnología cumple los objetivos y obtiene inferencias que estaban ocultas en los datos en bruto (Talavera y Gaudioso, 2004; Romero y Ventura, 2007).

Como hubo equipos que se comunicaron muy poco mediante los foros, o por problemas de colaboración o porque utilizaban otros medios de comunicación, se pensó en la conveniencia de hacer el análisis sólo con algunos equipos, sobre los que no se tenía dudas que la colaboración la habían realizado por los medios suministrados en la experiencia de aprendizaje colaborativo. Por esta razón se crearon otros datasets con los mismos datos pero sin los alumnos de equipos con poca interacción. Es decir, del dataset original de cada curso se eliminaron los alumnos de equipos con poca actividad.

El criterio para filtrar ha sido: eliminar a los alumnos de un equipo, cuya interacción sea inferior a la mitad de la interacción promedio de todos los equipos. En la Tabla 4 se dan los resultados.

**Tabla 4: criterio de filtrado de equipos con poca actividad.**

	2006/2007	2007/2008	2008/2009
Nº de equipos	42	45	38
Promedio de mensajes enviados	65,45	73,04	87,11
Nivel de filtrado	32	36	44
Nº equipo con poca interacción	12	11	8

La tabla señala que un equipo durante todo el periodo de colaboración enviaron sus miembros una media de 65 mensajes en el curso 2006/2007, 73 en el curso 2007/2008 y 87 en 2008/2009. Por supuesto los miembros de algunos equipos enviaron más mensajes y otros menos. Algunos equipos tuvieron muy poca actividad y para medir cuantitativamente la poca actividad, se estableció un límite o nivel de filtrado. Si todos los mensajes enviados por los miembros de un mismo equipo superaban el

límite, se consideró al equipo insuficientemente activo. En caso contrario se consideró al equipo poco activo.

Como fuente de datos del método de inferencia han sido utilizados los dos tipos de datasets, los filtrados y los sin filtrar. Los datasets sin filtrar se han llamado: **D-I-06-07**, **D-I-07-08** y **D-I-08-09**. Los datasets filtrados se han llamado: **D-II-06-07**, **D-II-07-08** y **D-II-08-09**. Antes de describir los resultados se desea presentar datos estadísticos de los datasets filtrados y sin filtrar para distinguir sus diferencias.

**Tabla 5: datos estadísticos de los datasets.**

	Población	N_thrd		N_msg		N_reply_thrd		N_reply_msg	
		M	V	M	V	M	V	M	V
D-I-06-07	117	4,97	15,37	22,79	270,05	17,90	431,65	17,17	240,69
D-II-06-07	82	6,23	15,28	30,41	184,88	24,46	461,74	24,21	200,63
D-I-07-08	131	4,35	9,31	24,54	254,87	20,26	452,30	20,21	244,73
D-II-07-08	103	4,59	9,70	25,47	271,53	20,94	435,69	20,91	257,94
D-I-08-09	112	4,64	17,64	28,73	249,14	23,96	608,56	23,89	266,26
D-II-08-09	88	5,30	19,94	33,91	185,15	28,59	663,74	28,49	235,20

Leyenda:

- M – media.
- V – varianza.

Se han considerado sólo cuatro indicadores (N\_thrd, N\_msg, N\_reply\_thrd, N\_reply\_msg), pero al ser el resto de indicadores combinaciones matemáticas de los cuatro, los resultados de media y varianza no muestran cambios importantes de los mostrados en la tabla. La población indica el número de instancias de cada dataset. Por ejemplo, el dataset D-I-06-07 tenía 117 instancias, que son los estudiantes, y éstos crearon en promedio 4,97 hilos en los foros o conversaciones durante la experiencia.

El comportamiento de los indicadores según se trate de los datasets sin filtrar o de los filtrados no muestran un comportamiento común que sirva para obtener alguna conclusión sobre la validez o no del filtrado.

Además de los indicadores estadísticos de las interacciones de los estudiantes en los foros, un experto analizó todos los mensajes en los foros de los espacios privados de los equipos en las experiencias de los cursos 2006/2007, 2007/2008 y 2008/2009. Con ello se trataba de etiquetar a los estudiantes según un número que representaba el nivel de colaboración. Utilizó una escala entre 1 y 9. El valor de 9 representaba la menor colaboración posible y nunca fue utilizado, mientras 1 representaba la mayor

colaboración posible y sólo fue utilizado una vez. Se creó una lista que ha sido utilizada para validar los métodos de inferencia de la colaboración.

Como el experto etiquetó a los estudiantes empezando por los miembros del primer equipo y terminando por los miembros del último, se pensó que los primeros estudiantes podrían haber sido erróneamente etiquetados debido a que el experto todavía no habría tenido un esquema mental del modelo de etiquetas, ya que no tenía la experiencia. Es decir, el experto al evaluar la colaboración de los estudiantes tenía que comparar lo que éstos habían enviado a los mensajes con un modelo mental de la colaboración en dicho entorno. Este modelo se obtiene con la experiencia por lo que los primeros casos podrían no haber sido etiquetados correctamente. Lo mismo se pensó de los últimos estudiantes etiquetados, ya que el cansancio de la tarea podría hacer al experto cometer errores. Por estas razones, la lista de estudiantes etiquetados según su colaboración por el experto también fue filtrada. El criterio fue: eliminar a los estudiantes de los primeros y últimos 5 equipos.

Ya que este filtro no se debió a ninguna característica cuantitativa, se comenzó el análisis de la colaboración con la lista sin filtrar y con la lista filtrada observando si los primeros datos obtenidos indicaban la utilidad del filtrado.

Los datos del etiquetado se encuentran en la Tabla 6.

**Tabla 6: datos sobre las lista del nivel de colaboración de los alumnos.**

	2006/2007	2007/2008	2008/2009
Nº de alumnos	125	140	115
Nº alumnos etiquetados	104	122	108
Promedio de la colaboración	4,18	4,43	4,33
Nº alumnos etiquetados al eliminar los filtrados	83	90	77
Promedio de la colaboración en la lista filtrada	4,20	4,43	4,47

Se observa al comparar los promedios de la lista original y filtrada por cada curso que no hay grandes diferencias. Los experimentos realizados con los datasets se realizaron también con la lista filtrada y sin filtrar en el caso de los datasets de los años 2006/2007 y 2007/2008. Adelantamos que obtenidos los primeros resultados se comprobó que no había diferencias reseñables entre los resultados de los datasets filtrados y sin filtrar de la lista de niveles de colaboración de los alumnos. Por eso, no se continuó el filtrado y se prefirió realizar los experimentos con una única lista por año.



## **Resultados de los métodos de inferencia de la colaboración**

Los objetivos de los experimentos realizados han sido: 1) encontrar una forma de medir la colaboración de los alumnos mediante indicadores estadísticos de las interacciones en los foros del entorno colaborativo, 2) que los resultados fueran obtenidos durante el proceso de colaboración o poco después, y 3) que los análisis pudieran ser trasladados a otros entornos o dominios sin perder utilidad.

Ya se han explicado los indicadores propuestos, aquí sólo se insiste en que estos indicadores son cuantitativos y fueron calculados una vez por semana, casi siempre los lunes. En realidad, gracias al método aplicado, los indicadores podían ser obtenidos cuando se quisiera, para ello se crearon unas funciones en lenguaje psql y unas tablas donde almacenarlos.

De la base de datos de la plataforma dotLRN se obtuvieron los datasets, los cuales fueron etiquetados según la colaboración de los alumnos. Se hace notar que se etiquetó a los alumnos según su colaboración siguiendo el criterio del experto. Este método fue el aplicado en segundo lugar.

Antes de utilizar al experto, la idea original fue identificar a los estudiantes más colaboradores por las respuestas de los estudiantes a la encuesta final. En las encuestas finales de los cursos 2006/2007 y 2007/2008 se pidió a los estudiantes que evaluaran el trabajo colaborativo de sus compañeros de equipo. Las respuestas a las encuestas no fueron muy numerosas y menos los que evaluaron a sus compañeros. Además, surgieron dudas sobre la veracidad de las respuestas debido a que evaluaban a sus compañeros con la máxima valoración para todos. Se realizaron algunos análisis por medio de árboles de decisión etiquetando a los estudiantes según las respuestas de sus compañeros. Los primeros resultados aportaban conclusiones contradictorias. Por esta razón se filtraron las respuestas para quedarse con las evaluaciones de la colaboración de los compañeros más objetivas. La lista de estudiantes etiquetados se redujo bastante haciendo que los análisis mediante árboles de decisión obtuvieran resultados contradictorios y con errores muy altos. Esta fuente de datos fue lamentablemente desechada y se planteó entonces el etiquetado del experto. La lista del experto suministraba el nivel de colaboración para la mayoría de los estudiantes

utilizando un criterio común, todo lo cual hace a la lista suministrada por el experto un medio más útil y efectivo para el método de inferencia propuesto.

Se ha explicado que los dataset fueron filtrados debido a la poca actividad de algunos equipos por lo que se eliminaron los alumnos de equipos donde el número total de mensajes enviados a los foros del equipo era menor a la mitad de la media de mensajes enviados a un equipo. De esta forma, por cada curso, se presentarán los resultados de dos datasets, los no filtrados de equipos poco activos y los filtrados de equipos poco activos.

En el Capítulo III se han descrito dos acercamientos al análisis de la colaboración dependiendo de los resultados que se desean obtener y teniendo en cuenta los objetivos, tanto del análisis como de lo realizado con los resultados. Por un lado, se ha descrito un procedimiento de minería de datos cuya meta ha sido construir un métrica que represente el valor de la colaboración de los estudiantes, y por otro, se ha descrito un procedimiento de clasificación de los alumnos según su colaboración por medio de sus datos de interacción.

En cada uno de los acercamientos de minería de datos se han utilizado distintas técnicas de aprendizaje automático: algoritmos de árbol de decisión para encontrar la relación entre la colaboración y los indicadores estadísticos, y un algoritmo de clustering para clasificar las interacciones de los alumnos según la colaboración. A continuación se mostrarán los resultados obtenidos con ambos planteamientos.

## **Resultados del análisis mediante métricas de la colaboración**

Se ha propuesto un métrica que dé un valor a cada estudiante según su colaboración. Este acercamiento cumple los objetivos y ofrece una valor cuantitativo con los que los estudiantes puedan compararse entre sí y deducir el grado de colaboración ya que la colaboración es una característica prototípica (Rosch, 1978).

En este acercamiento hay dos puntos sensibles: qué variables utilizar en la métrica y qué forma tiene la métrica. Como posibles variables se han propuesto los indicadores estadísticos extraídos del análisis de las interacciones. Para identificar los indicadores más relevantes en relación a la colaboración, se buscó un método de

inferencia que cumpliera los objetivos y cuyos resultados informaran de la relación entre los indicadores y la colaboración. Se acudió a las técnicas de aprendizaje automático y, entre éstas, a los árboles de decisión. Éstos informan en un árbol lógico de los atributos utilizados para clasificar una serie de instancias. En el capítulo anterior se ha justificado esta decisión.

Para el caso de encontrar la relación entre la colaboración y los indicadores, por un lado, se tenía una clasificación de los estudiantes según su colaboración dada por el análisis del experto, y por el otro, los doce indicadores estadísticos para cada uno de los estudiantes etiquetados. Se propuso el siguiente método de trabajo:

- Entrenar con el dataset etiquetado todos los algoritmos de árbol de decisión que suministraran un árbol lógico. Los algoritmos han sido: *Best first decision tree*, *DecisionStump*, *Functional trees*, *J48*, *Logistic model trees*, *Naïve Bayes tree*, *Random tree*, *REPTree*, *Simple Cart*.
- De cada árbol obtener los indicadores estadísticos utilizados.
- Contar cuantos algoritmos de árbol de decisión utilizan un determinado indicador estadístico.
- El indicador o indicadores más utilizados se les supone más relacionados con la colaboración de los alumnos.
- Construir una métrica con los indicadores más utilizados.
- Relacionar dicha métrica con las medidas de la colaboración dadas por el tutor.

En la Tabla 7 se muestra la frecuencia de uso de los indicadores estadísticos y el número de algoritmos de árbol de decisión utilizados para cada dataset.

**Tabla 7: N° de árboles de decisión que utilizan cada uno de los indicadores estadísticos.**

	D-I-06-07	D-II-06-07	D-I-07-08	D-II-07-08	D-I-08-09	D-II-08-09
Num_thr	4	3	3	2	7	8
Med_thr	2	2	3	2	3	4
Var_thr	4	3	4	3	4	5
Level_thr	5	4	4	4	5	4
Num_msg	5	4	5	4	2	5
Med_msg	2	2	2	1	3	4
Var_msg	4	4	4	3	5	5
Level_msg	4	5	4	5	4	8
Num_reply_thr	3	3	3	2	4	4
Med_reply_thr	2	3	4	4	2	5
Num_reply_msg	4	3	5	5	3	8
Med_reply_msg	3	3	5	5	5	5
N° algoritmos utilizados	6	6	7	7	7	9

El objetivo de este análisis ha sido encontrar los indicadores estadísticos que mejor discriminan la colaboración. Se pueden utilizar varios criterios para identificarlos (por ejemplo, que el número de árboles de decisión que los han utilizado supere cierto límite, que sean el más utilizado en más de un dataset, etc.). El criterio utilizado ha sido: los más utilizados contando con todos los datasets. Por lo que, sumando las cantidades de cada fila sería suficiente. Sin embargo, como se ve en la Tabla 7, ha habido datasets que han funcionado en más algoritmos que otros. Para minimizar esto se han propuesto distintas ponderaciones de la suma:

- Suma sin ponderar de las filas de la tabla superior.
- Suma ponderada I: donde cada valor de la tabla superior es dividido por el valor máximo en la columna y después se suma los valores de cada fila.
- Suma ponderada II: donde cada valor de la tabla superior es dividido por el número de algoritmos utilizados (el valor de la última fila) y después se suman los valores de cada fila.

**Tabla 8: indicadores estadísticos más frecuentes según tres tipos de suma.**

Indicadores estadísticos	Suma de los usos sin ponderar	Suma ponderada I	Suma ponderada II
Num_thr	<b>27</b>	4,40	3,77
Med_thr	16	2,73	2,25
Var_thr	23	4,00	3,29
Level_thr	26	<b>4,61</b>	<b>3,80</b>
Num_msg	25	4,51	3,63
Med_msg	14	2,33	1,97
Var_msg	25	4,34	3,60
Level_msg	<b>30</b>	<b>5,17</b>	<b>4,25</b>
Num_reply_thr	19	3,27	2,73
Med_reply_thr	20	3,51	2,82
Num_reply_msg	<b>28</b>	<b>4,83</b>	<b>3,91</b>
Med_reply_msg	26	4,54	3,70

Como se observa en la Tabla 8, los tres indicadores estadísticos más frecuentes están marcados en cada columna con negrita. Es interesante observar que los dos indicadores más utilizados no varíen y son “level\_msg” y “num\_reply\_msg”. El tercer indicador estadístico más utilizado es, según casos, o “level\_thr” o “num\_thr”.

Con estos resultados se han propuesto tres métricas que utilizan algunos de los indicadores como variables y se ha comparado los resultados con la lista suministrada por el experto sobre el nivel de colaboración de los alumnos. En el siguiente apartado se habla de ellas.

## Métricas propuestas

Las métricas que se proponen incluyen los indicadores más utilizados en los tres cursos. Así se pretende conseguir una métrica general que se pueda utilizar en un futuro, en lugar de una métrica por curso, la cual tendría otra vez que inferirse para cada nuevo curso, lo cual no es lo buscado en los objetivos de la investigación. Los indicadores más utilizados en los tres años han sido: *level\_msg* (número de mensajes dividido por la varianza), *num\_reply\_msg* (número de réplicas a los mensajes enviados) y *level\_thr* (número de conversaciones iniciadas dividido por la varianza), aunque en la suma sin ponderar (ver columna segunda de la Tabla 8) el tercer indicador más utilizado es *num\_thr* (número de conversaciones iniciadas). Se han propuesto tres métricas donde aparecen los indicadores más utilizados según la tabla anterior. A los valores de cada indicador se les ha normalizado dado un valor de entre 1 (máximo) y 0 (mínimo) para que ningún indicador estadístico sobresaliera por encima del resto. Las métricas propuestas son:

- *Métrica I* =  $(level\_msg/\max(level\_msg)) + (num\_reply\_msg/\max(num\_reply\_msg)) + (level\_thr/\max(level\_thr))$
- *Métrica II* =  $(level\_msg/\max(level\_msg)) + (num\_reply\_msg/\max(num\_reply\_msg)) + (level\_thr/\max(level\_thr)) + (med\_reply\_msg/\max(med\_reply\_msg))$
- *Métrica III* =  $(level\_msg/\max(level\_msg)) + (num\_reply\_msg/\max(num\_reply\_msg)) + (level\_thr/\max(level\_thr)) + (num\_thr/\max(num\_thr))$

Para que cada indicador tuviera un valor entre 0 y 1, se ha dividido por el valor máximo del indicador en cada dataset. Las métricas así definidas cumplen las condiciones de espacio métrico (Arregui, 2002) que son:

- La distancia entre dos puntos siempre es mayor o igual a 0 ( $d(x,y) \geq 0$ ).
- La distancia entre dos puntos que son el mismo es cero ( $d(x,x) = 0$ ).
- La distancia entre un punto X y otro Y es igual a la distancia entre un punto Y y otro X ( $d(x,y) = d(y,x)$ ).
- Cumple la desigualdad triangular ( $d(x,y) \leq d(x,z) + d(y,z)$ ).

Lógicamente no se conocen qué métrica será la mejor. Para conocer cuál de ellas funcionaría mejor se han agrupado las instancias que representan los alumnos según su nivel de colaboración dado por el experto. Recordamos que el experto asignó a los

alumnos un valor del nivel de colaboración entre 1 (realmente muy colaborativo) a 9 (muy poco colaborativo), aunque el valor de 9 nunca fue utilizado. Se han agrupado los alumnos por su valor en el nivel de colaboración y se han medido las métricas para cada uno. Una vez que se tenía esto, se ha calculado la media para cada conjunto de alumnos y la varianza. A continuación se muestran los resultados con los 6 datasets.

Las tablas que vienen a continuación (tablas 9, 10, 11, 12, 13, 14) muestran los niveles de colaboración (Nivel) que hay en el dataset junto con el número de instancias de cada nivel (Nº instancias). De todas las instancias del dataset se calculan el valor de las métricas propuestas. En las tablas aparecen la media o promedio (Media) y la varianza (Varianza) de los valores de las métricas en los grupos de instancias que tienen el mismo nivel de colaboración.

**Tabla 9: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-I-06-07.**

Nivel de colaboración		Métrica I		Métrica II		Métrica III	
Nivel	Nº instancias	Media	Varianza	Media	Varianza	Media	Varianza
2	10	1,89	0,09	1,92	0,11	2,10	0,20
3	24	1,61	0,13	1,81	0,15	2,04	0,27
4	34	1,37	0,11	1,59	0,18	1,69	0,23
5	19	1,08	0,18	1,27	0,22	1,34	0,32
6	10	0,78	0,11	0,91	0,13	0,94	0,17
7	4	0,69	0,03	0,92	0,04	0,89	0,10
8	3	0,60	0,08	0,68	0,12	0,72	0,14

**Tabla 10: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-II-06-07.**

Nivel de colaboración		Métrica I		Métrica II		Métrica IV	
Nivel	Nº instancias	Media	Varianza	Media	Varianza	Media	Varianza
2	9	2,10	0,07	2,10	0,09	2,29	0,20
3	24	1,68	0,15	1,88	0,18	2,10	0,31
4	29	1,54	0,09	1,77	0,16	1,89	0,21
5	12	1,30	0,20	1,52	0,20	1,58	0,39
6	4	1,21	0,16	1,40	0,22	1,46	0,17
7	1	1,07		1,23		1,54	

Como se muestra en la Tabla 10, al eliminar los equipos poco activos en este dataset, se eliminaron también los alumnos con un nivel de colaboración de 8, y de nivel de colaboración 7 sólo hubo 1, por lo que no se calculó la varianza y la media es el valor de la métrica en ese único caso.

**Tabla 11: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-I-07-08.**

Nivel de colaboración		Métrica I		Métrica II		Métrica IV	
Nivel	Nº instancias	Media	Varianza	Media	Varianza	Media	Varianza
1	1	1,03		1,69		1,31	
2	5	1,89	0,14	2,52	0,18	2,21	0,31
3	27	1,53	0,21	2,09	0,26	1,88	0,39
4	32	1,50	0,15	2,05	0,21	1,83	0,29
5	36	1,34	0,23	1,83	0,32	1,69	0,46
6	16	0,95	0,13	1,44	0,11	1,16	0,30
7	1	0,86		1,44		1,14	
8	6	1,22	0,21	1,72	0,30	1,58	0,46

Muestra la Tabla 11 que sólo hubo un alumno con el nivel de colaboración igual a 1, y otro con el nivel de colaboración igual a 7.

**Tabla 12: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-II-07-08.**

Nivel de colaboración		Métrica I		Métrica II		Métrica IV	
Nivel	Nº instancias	Media	Varianza	Media	Varianza	Media	Varianza
2	3	1,64	0,02	2,24	0,02	1,83	0,04
3	24	1,58	0,22	2,15	0,26	1,97	0,40
4	25	1,49	0,19	2,02	0,26	1,83	0,35
5	26	1,43	0,23	1,94	0,31	1,81	0,45
6	14	0,97	0,13	1,46	0,22	1,20	0,31
7	1	0,86		1,44		1,14	
8	4	1,27	0,11	1,79	0,16	1,67	0,31

Se ve en la Tabla 12 que al eliminar los equipos poco activos, no aparece el alumno cuyo nivel de colaboración es 1.

**Tabla 13: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-I-08-09.**

Nivel de colaboración		Métrica I		Métrica II		Métrica IV	
Nivel	Nº instancias	Media	Varianza	Media	Varianza	Media	Varianza
2	3	1,83	0,18	2,36	0,18	2,29	0,14
3	19	1,50	0,13	2,02	0,22	1,78	0,23
4	47	1,27	0,08	1,80	0,10	1,43	0,13
5	25	0,98	0,06	1,46	0,07	1,08	0,10
6	11	0,93	0,09	1,44	0,13	1,04	0,12
7	1	0,54		1,03		0,58	
8	1	0,95		1,63		1,05	

Se observa en la Tabla 13 que sólo un alumno ha sido etiquetado con el valor del nivel de colaboración de valor igual a 7 y lo mismo ha pasado con el valor igual a 8.

**Tabla 14: valores promedios y varianzas de las métricas según el nivel de colaboración en el dataset D-II-08-09.**

Nivel de colaboración		Métrica I		Métrica II		Métrica III	
Nivel	Nª instancias	Media	Varianza	Media	Varianza	Media	Varianza
2	3	1,83	0,18	2,36	0,18	2,29	0,14
3	18	1,50	0,13	2,02	0,22	1,78	0,23
4	39	1,33	0,07	1,87	0,09	1,49	0,13
5	20	1,02	0,06	1,49	0,07	1,13	0,11
6	7	1,12	0,02	1,66	0,02	1,26	0,04
7	0						
8	1	0,95		1,63		1,05	

La Tabla 14 muestra que ha desaparecido el alumno etiquetado con un valor en el nivel de colaboración de 7.

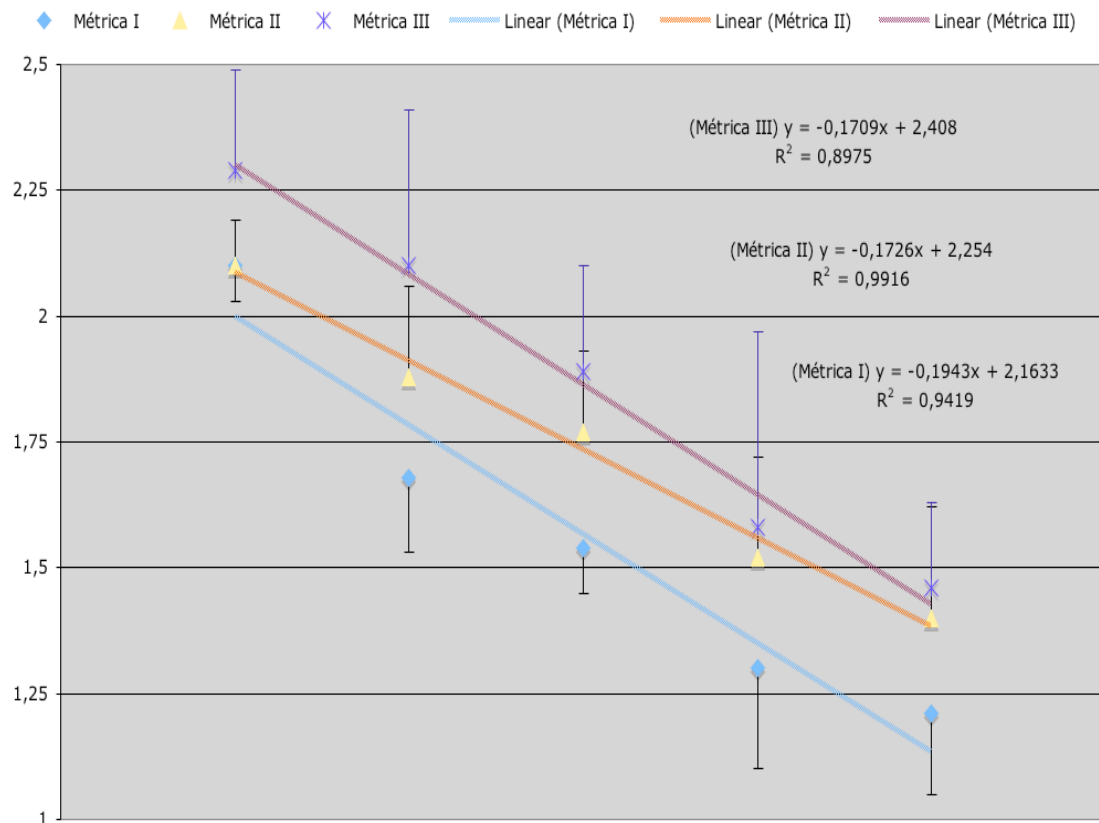
### **Análisis crítico**

De los resultados expuestos en las tablas anteriores hay que hacer notar que los valores que muestra la varianza indican que los grupos, entendiendo como grupo al conjunto de alumnos que han sido etiquetados con el mismo valor del nivel de colaboración, se solapan en muchos casos, acercándose la varianza al 10% del valor medio. Esto hace que muchos alumnos tengan valores en las métricas que podrían entrar en dos grupos según el nivel de colaboración. Es decir, un alumno con valores en las métricas tales que podrían entrar en el grupo del nivel de colaboración igual a 4, también lo podría hacer en el grupo del 3, pero si se da ese caso ese mismo alumno no podría entrar en el grupo de nivel de colaboración 5 o 2. El resultado no es muy exacto, por lo que un alumno podría tener un nivel según alguna de las métricas de 3, o nivel alto, pero también podría tener un nivel de 4, nivel medio, pero nunca podría llegar a un nivel 6, nivel bajo. La poca exactitud de la asignación de nivel de colaboración según la métrica indica que los conjuntos del mismo nivel de colaboración tienen los límites borrosos.

Para obtener una métrica idónea hay que atender, entre otros, a dos factores: la varianza que indica el grado de borrosidad o poca proximidad de la predicción del nivel de colaboración, y la pendiente de la recta que forma la métrica. Si se crea una gráfica donde cada punto es el valor de la métrica para una instancia, y los puntos se ordenan de mayor a menor, se obtendrá una gráfica que puede ser representada por una recta, utilizando el cálculo de mínimos cuadrados por ejemplo. La pendiente de la



recta, la derivada en cada uno de sus puntos, indicará también el grado predictivo de la métrica. Teniendo en cuenta que la variable “x” son los indicadores estadísticos que componen la métrica e “y” el nivel de colaboración, a mayores diferencias entre los distintos niveles en el valor “y”, indicará mayor poder predictivo. En la Figura 14 se muestran las curvas, obtenidas después de un análisis de mínimos cuadrados, de los valores para las métricas con el dataset D-II-06-07.



**Figura 18: rectas obtenidas por análisis de mínimos cuadrados de los valores medios de las métricas para el dataset D-II-06-07.**

La Figura 18 muestra una gráfica donde los valores del eje Y son los valores para las métricas y los puntos en el eje X se ordenan según el nivel de colaboración. Los puntos más situados a la izquierda representan los valores medios en cada una de las métricas para el conjunto de estudiantes agrupados según el nivel de colaboración “2”. Los puntos más a la derecha representan el valor medio de las métricas (eje Y) para los estudiantes agrupados por el nivel de colaboración “7” (eje X). Las barras de error que aparecen muestran la varianza, y se ve con claridad que la barra de error de la métrica III es la mayor. La gráfica es una representación gráfica de los datos de la Tabla 10.

El análisis de mínimos cuadrados nos dice que la métrica y los niveles de colaboración se relacionan entre sí con una recta. Esto indica que un valor más alto de la métrica indica un valor más alto de la colaboración y viceversa. El análisis de mínimos cuadrados también nos muestra la pendiente de la curva. En la esquina superior derecha de la Figura 14 aparecen las fórmulas de las rectas que han sido calculadas por medio de un análisis de mínimos cuadrados de los puntos para cada una de las métricas. La pendiente de las curvas es el valor numérico que aparece al realizar la deriva en “x” de la fórmula. La pendiente para la métrica I = -0,1943; para la métrica II = -0,1726; para la métrica III = -0,1709. El valor absoluto de la pendiente indica la distancia entre dos nivel de colaboración contiguos, es decir, la diferencia en el valor de las métricas entre dos niveles de colaboración cualesquiera. A valores más altos, indica que la métrica es capaz de distinguir mejor entre dos niveles.

En geometría la pendiente de una recta es el indicativo del cambio de la variable Y respecto a la variable X. Hay distintas formas de medirla<sup>12</sup> y nosotros nos hemos decantado por el cálculo de la derivada de la recta. La pendiente nos dice que dejando constante un intervalo de la variable X, es decir, en nuestro caso tomado un número fijo de un intervalo de instancias o estudiantes, a mayor pendiente indica mayor diferencia entre los valores de la variable Y, en este caso la métrica en cuestión. Mayor diferencia entre los valores en la variable Y, es decir, los valores de la métrica, indica más facilidad en discriminar entre la colaboración de los estudiantes que forma el intervalo y menos probabilidad de que haya solapamientos.

Por lo tanto se han definido dos variables para indicarnos de forma cuantitativa qué métrica es mejor: la pendiente de la curva y el error representado por la varianza en las Tablas 9, 10, 11, 12, 13 y 14. En el Capítulo V, el análisis realizado con el dataset D-II-06-07 será repetido con todos los datasets para encontrar la mejor métrica de las tres propuestas.

## **Resultados de la clasificación mediante clustering**

La tecnología de clustering agrupa instancias según cierta semejanza entre ellas por medio de una métrica. No se necesita un entrenamiento ni un etiquetado de las

---

<sup>12</sup> [http://es.wikipedia.org/wiki/Pendiente\\_de\\_la\\_recta](http://es.wikipedia.org/wiki/Pendiente_de_la_recta)

instancias para guiar la clasificación. Lo que sí es necesario es otra fuente de información que informe del significado de la clasificación alcanzada. El camino es primero clasificar y después comparar la clasificación con otra fuente de datos para obtener un significado. Esto se puede realizar mediante la técnica de clustering. El método de clustering cubre los objetivos de la investigación y una vez validado, es decir, comparada la clasificación que ofrece con una fuente de datos distinta que informe de la característica que se quiere observar en la clasificación, puede ser utilizada de forma regular y frecuente. Hemos elegido un solo algoritmo, EM (Gama y Gaber, 2007). Los algoritmos de clustering agrupan instancias según una medida de la distancia que suele ser la distancia euclídea. Los buenos resultados del algoritmo EM en otras experiencias nos ha hecho decantarnos por él (Talavera y Gaudioso, 2004; Perera et al., 2007; Meilâ y Heckerman, 2001).

Al algoritmo EM se le puede dar el número de cluster o grupos que se quiere utilizar para la agrupación de instancias. Ya que el objetivo es relacionar los grupos obtenidos con la colaboración, el número de clusters o grupos tiene que ser tal que la clasificación obtenida sea entendible. Un planteamiento extremo poco operativo sería la obtención de un gran número de clusters o grupos de instancias y que cada grupo significara un nivel de colaboración distinto. Al dar una etiqueta del lenguaje natural a cada grupo, podría ser difícil de entender o de establecer la diferencia entre distintos grupos. En esta investigación, el resultado de la clasificación se ofrece a los propios alumnos que han sido clasificados, por lo cual es aconsejable utilizar unas etiquetas de clasificación entendibles por cualquiera.

Al no conocer las etiquetas previamente, se ha experimentado para que el algoritmo EM suministrara 2, 3 y 4 cluster. Las etiquetas propuestas:

- Caso de 2 clusters: colaboración alta, colaboración baja.
- Caso de 3 clusters: colaboración alta, colaboración media, colaboración baja.
- Caso de 4 clusters: colaboración muy alta, colaboración alta, colaboración baja, colaboración muy baja.

En las Figuras 19, 20, 21, 22, 23 y 24 se muestran los indicadores estadísticos básicos frente a los cluster, es decir, una gráfica de las instancias según un indicador estadístico (en el eje X) frente a la clasificación obtenida al aplicar el algoritmo de clustering (en el eje Y). El objetivo es mostrar la importancia del indicador en la clasificación y, de paso, deducir cual es el mejor número de clusters para los datasets.

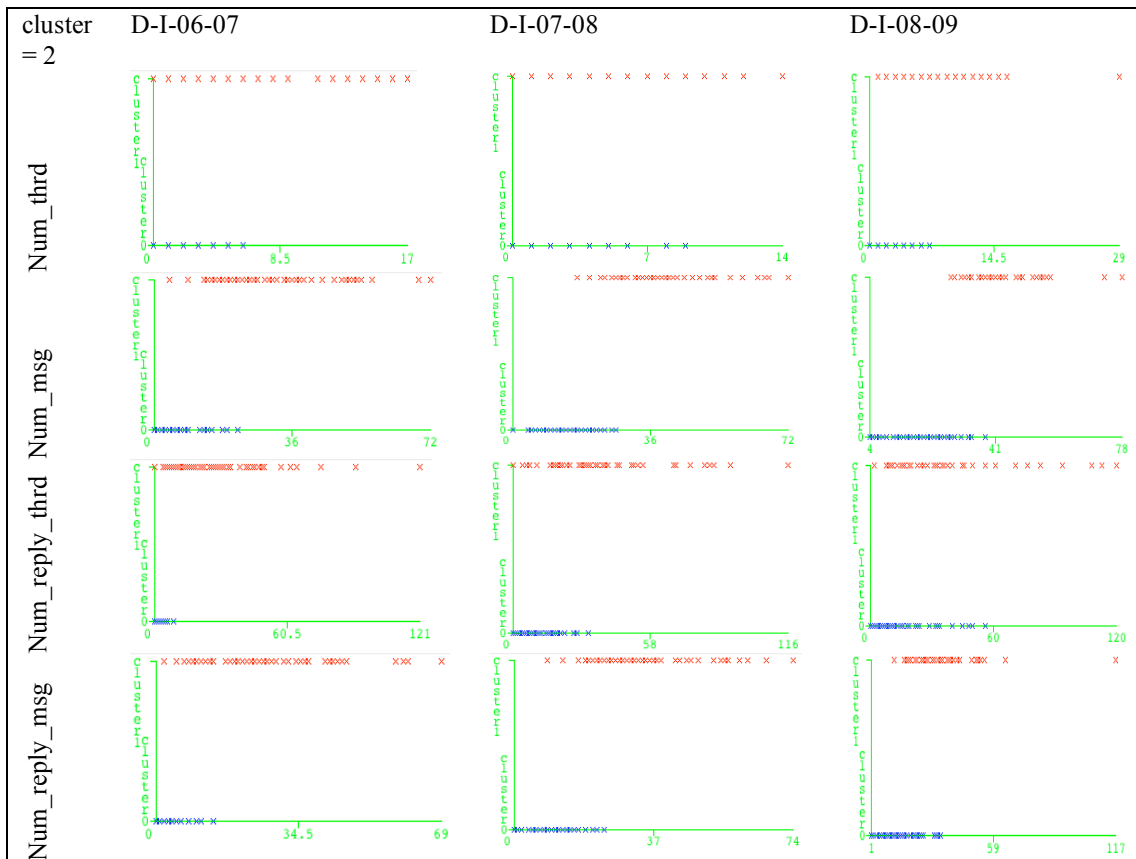


Figura 19: datasets sin filtrar de equipos poco activos clasificados en 2 cluster.

La Figura 19 muestra la clasificación con dos cluster o grupos con los datasets D-I-06-07, D-I-07-08 y D-I-08-09. En las gráficas aparecen las instancias clasificadas según el indicador estadístico que se muestra a la izquierda. Una de las clasificaciones aparece en la parte superior de cada gráfica (en rojo) y la otra en la parte inferior (en azul). Se puede decir que un ejemplo de clasificación es la que muestra la gráfica del indicador Num\_msg frente a los dos cluster en el dataset D-I-08-09, ya que el solapamiento es el mínimo, es decir, hay muy pocas instancias de un cluster que tengan valores en el indicador estadístico Num\_msg iguales a instancias del otros cluster. Se observa que con algunos indicadores los cluster se diferencia con claridad aunque con solapamiento (Num\_msg), pero que en otros casos, no se pueden identificar con tanta claridad los cluster (Num\_thr). Los cluster caracterizados por el indicador estadístico Num\_thr es un ejemplo de poca claridad ya que los dos cluster definidos tienen instancias con valores semejantes, es más, según los valores en el indicador estadístico, el cluster inferior (de color azul) es un subconjunto del cluster superior (de color rojo) ya que todas las instancias del cluster inferior comparten valores con alguna instancia del cluster superior.

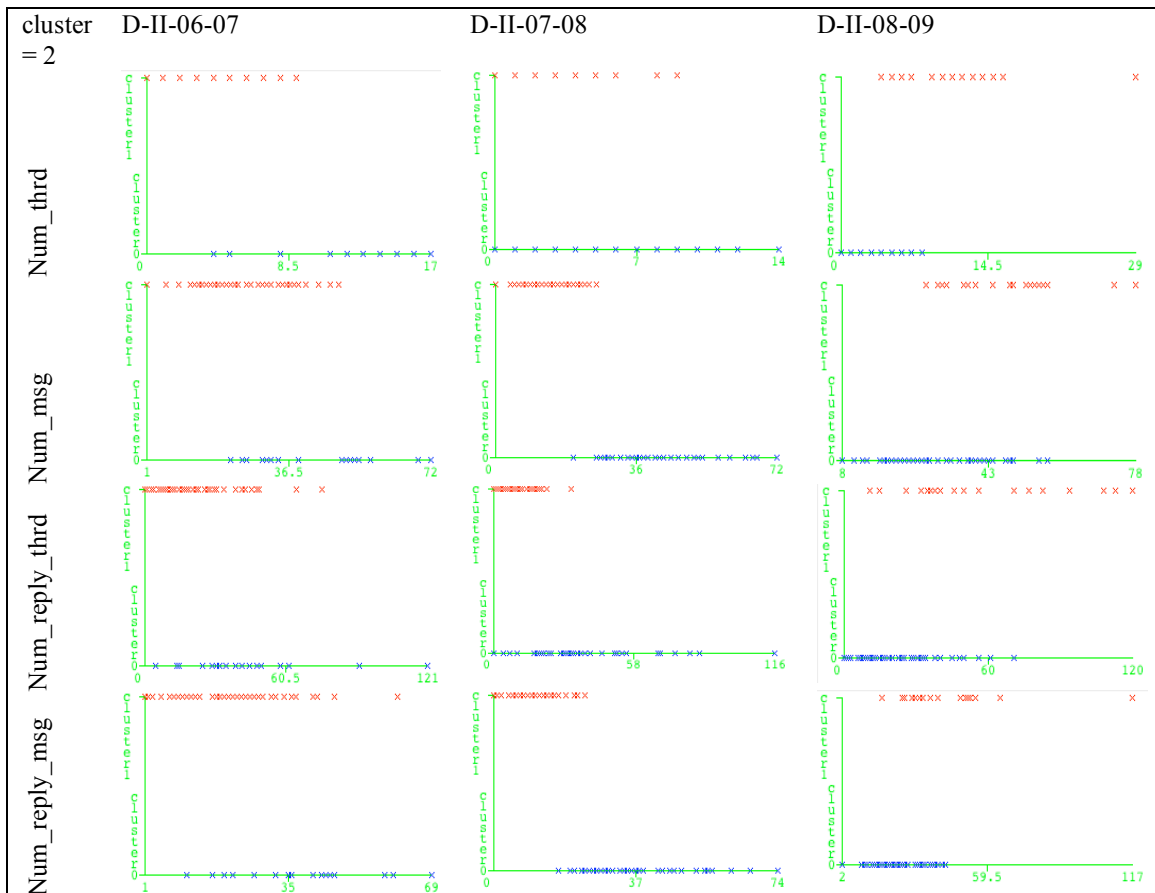
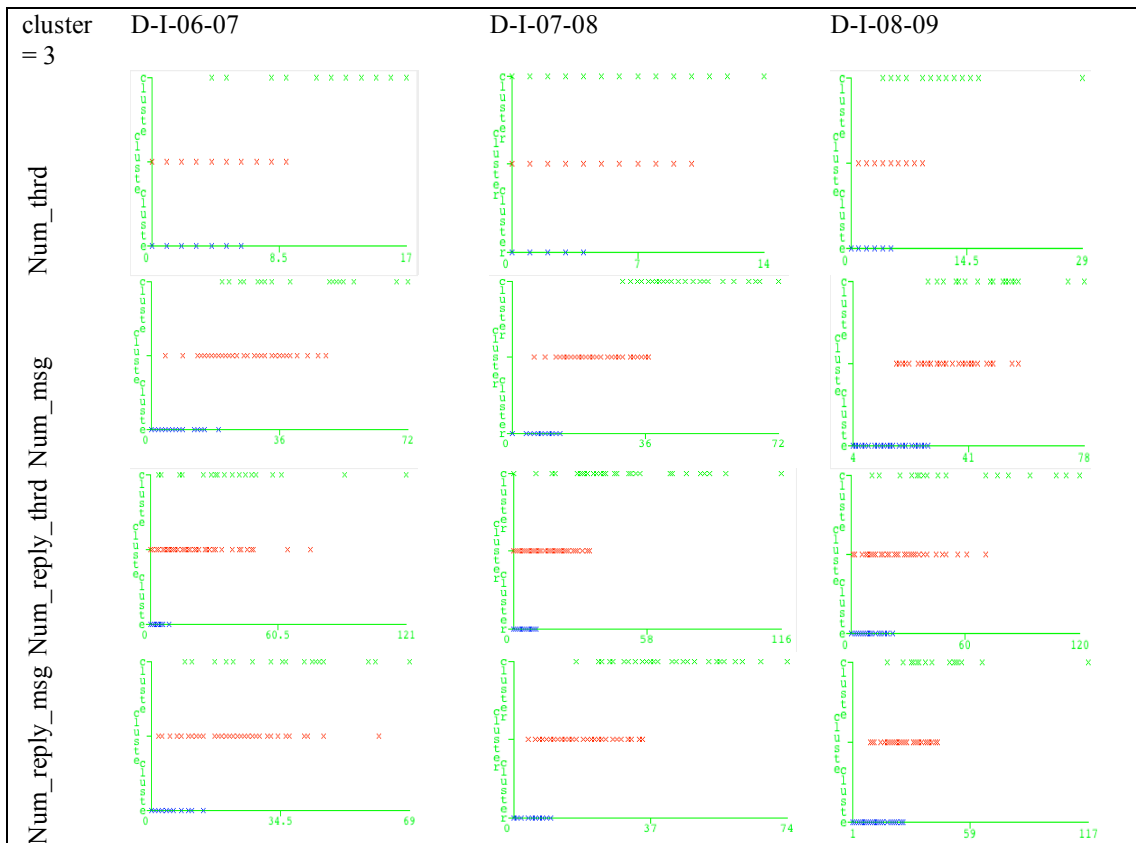


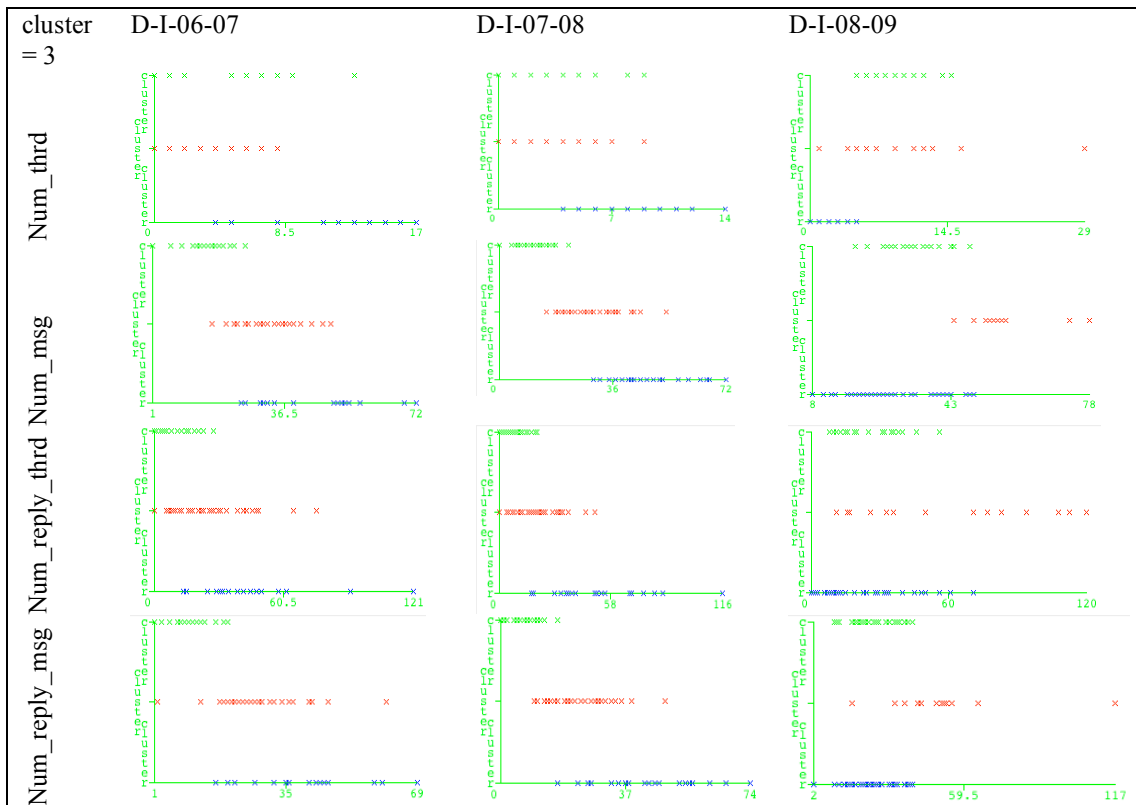
Figura 20: datasets filtrados de equipos poco activos clasificados en 2 cluster.

La Figura 20 muestra la clasificación con dos cluster o grupos con los datasets D-II-06-07, D-II-07-08 y D-II-08-09. En las gráficas aparecen las instancias clasificadas según el indicador estadístico que se muestra a la izquierda. Una de las clasificaciones aparece en la parte superior de cada gráfica (en rojo) y la otra en la parte inferior (en azul). Se puede decir que un ejemplo de clasificación correcta es la que muestra la gráfica del indicador Num\_msg frente a los dos cluster en el dataset D-I-07-08. En comparación con las gráficas mostradas en la Figura 19, el filtrado de los datasets no aumenta la claridad. Examinando las gráficas no se puede establecer la utilidad o no del filtrado.



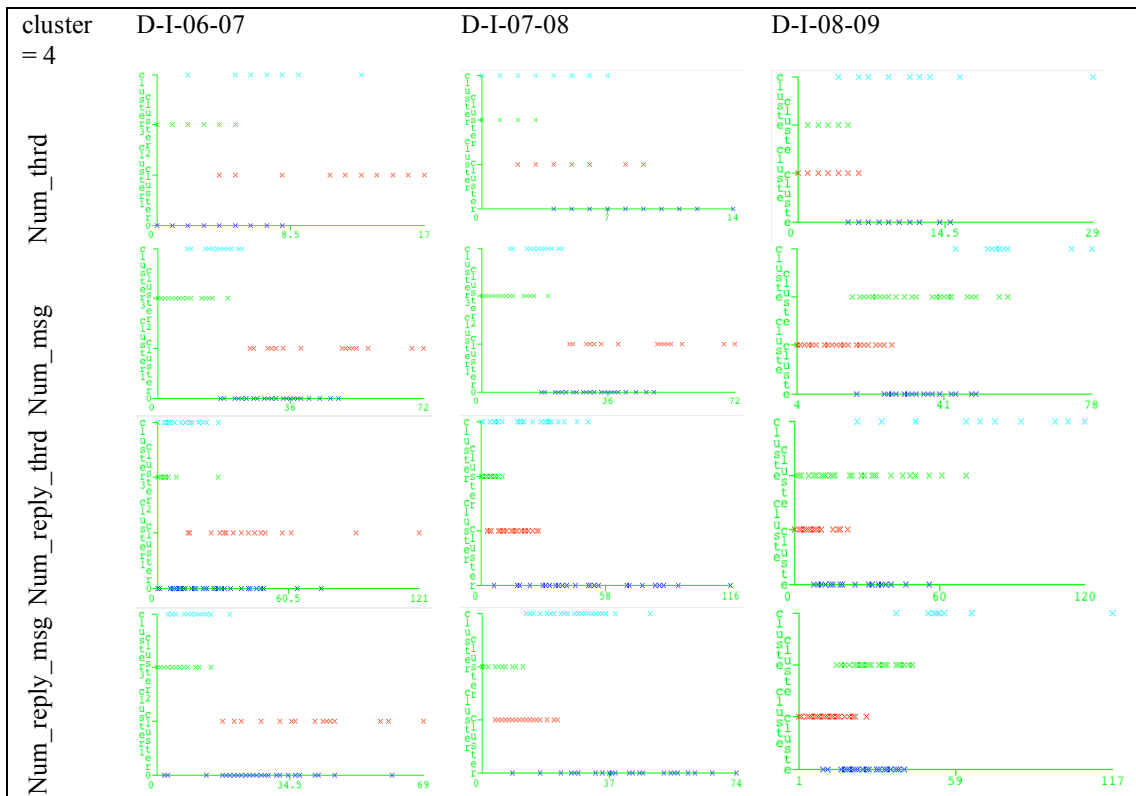
**Figura 21: datasets sin filtrar de equipos poco activos clasificados en 3 cluster.**

Ahora, en vez de clasificar en 2 cluster, en la Figura 21 se muestra la clasificación con tres cluster o grupos con los datasets D-I-06-07, D-I-07-08 y D-I-08-09. En las gráficas aparecen las instancias clasificadas según el indicador estadístico que se muestra a la izquierda. Una de las agrupaciones o clusters aparece en la parte superior de cada gráfica (en verde), otra en la mitad (en rojo) y la tercera en la parte inferior (en azul). Se observa que los cluster se solapan los unos con los otros. En el caso de algunos indicadores (Num\_msg y Num\_reply\_msg) el cluster situado en la mitad de las gráficas (en rojo) solapa con los otro dos, pero los otros dos no se solapan entre sí.



**Figura 22: datasets filtrados de equipos poco activos clasificados en 3 cluster.**

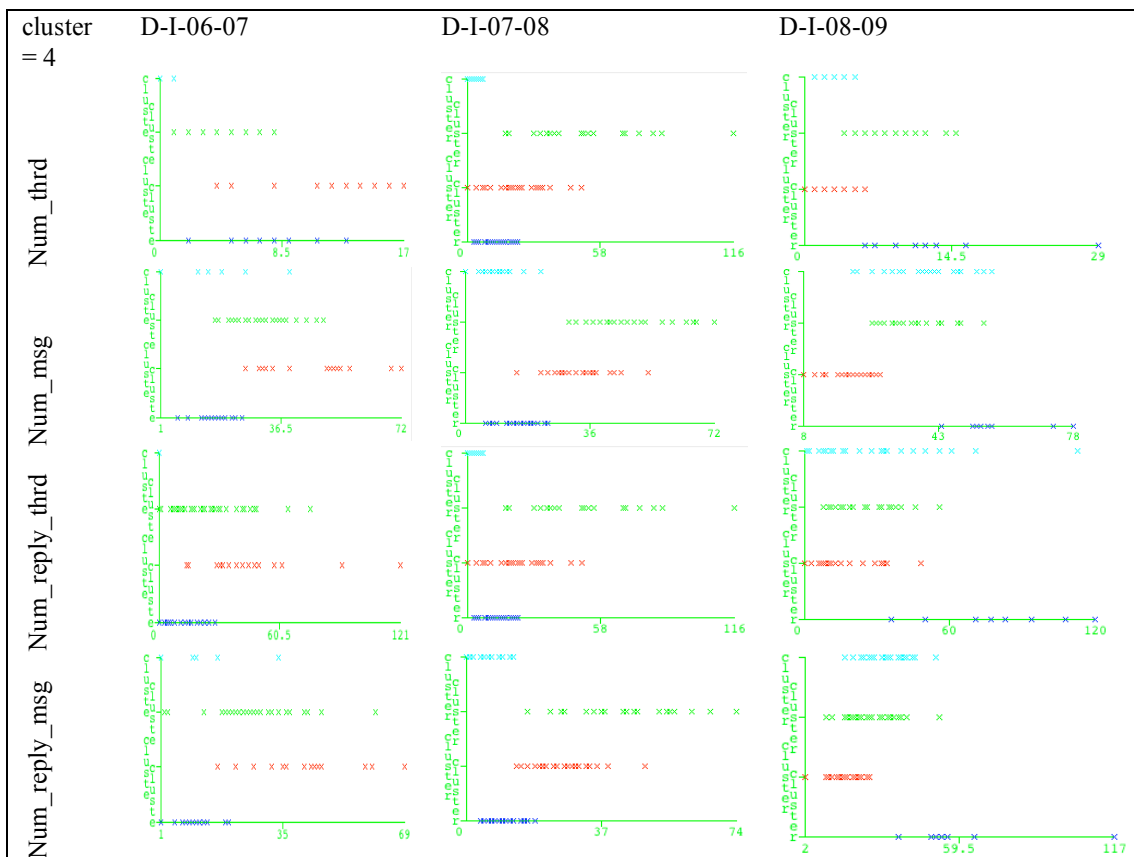
Los datasets filtrados D-II-06-07, D-II-07-08 y D-II-08-09 se han clasificado en 3 cluster. Es la única diferencias entre las gráficas de la Figura 21 y esta, la Figura 22. De un primer análisis se puede afirmar que lo dicho para las gráficas de la Figura 21, se puede decir para las gráficas de la Figura 22. No hay diferencias significativas entre la clasificación con los datasets sin filtrar de los filtrados.



**Figura 23: datasets sin filtrar de equipos poco activos clasificados en 4 cluster.**

La Figura 23 muestra la clasificación con cuatro cluster o grupos con los datasets D-I-06-07, D-I-07-08 y D-I-08-09. En las gráficas aparecen las instancias clasificadas según el indicador estadístico que se muestra a la izquierda. Una de las clasificaciones aparece en la parte superior de cada gráfica (en azul claro), otra en la mitad superior (en verde), una tercera en la parte media inferior (en rojo), y la cuarta en la parte inferior (en azul). Se observa que los clusters se solapan y es difícil distinguirlos, sobre todo a aquellos que no colectan instancias con valores extremos en los indicadores estadísticos, es decir, valores o muy altos o muy bajos. Por ejemplo, el cluster de la parte superior central (color verde) y el cluster inferior (color azul) colectan instancias con valores muy parecidos en algunos indicadores.





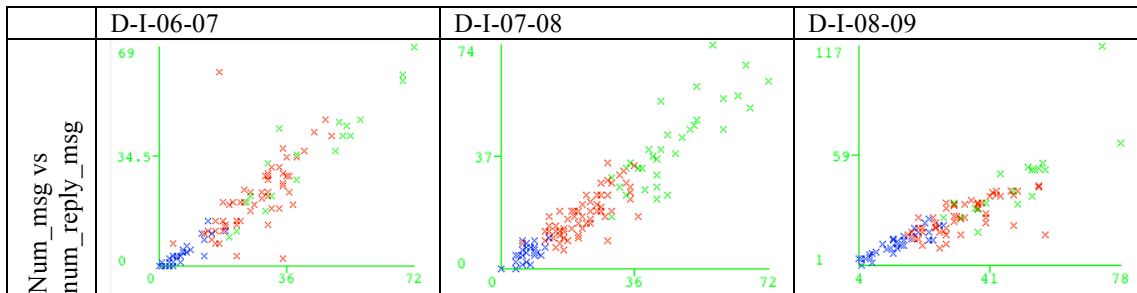
**Figura 24: datasets filtrado de equipos poco activos clasificados en 4 cluster.**

En la Figura 24 se muestran el mismo caso que en la Figura 23 pero con los datasets filtrados D-II-06-07, D-II-07-08 y D-II-08-09. Los dicho para el análisis en el caso de los dataset sin filtrar con la clasificación con cuatro cluster (Figura 23) también es válido aquí. No hay grandes diferencias entre los datasets filtrados de los sin filtrar.

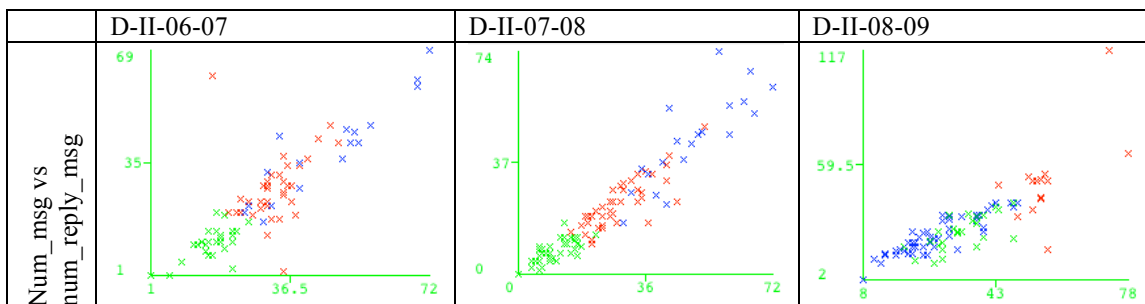
Se observa en todas las figuras anteriores (19, 20, 21, 22, 23 y 24) que los gráficos donde aparecen los indicadores “num\_msg” y “num\_reply\_msg” frente a los cluster, muestran estos con más claridad y menos solapamientos. Esto quiere decir que estos indicadores están más relacionados que otros con la clasificación realizada.

Se ha comentado que en el caso de la clasificación con 4 agrupamientos, no hay una diferencia clara entre los cluster que no tienen valores extremos en los indicadores. En el caso de la clasificación con 2 clusters, se observa que se pueden identificar con claridad y sin solapamientos los grupos en algunos casos. Sin embargo, esto ocurre con gran claridad en el caso de la clasificación con 3 clusters entre los grupos con valores extremos en los indicadores estadísticos. Pensando en el uso pedagógico de la clasificación, se considera conveniente la clasificación de las instancias, que representan las interacciones de los estudiantes, en 3 clusters o grupos.

Esto facilita el entendimiento de dicha clasificación por parte de los estudiantes. En las figuras 25 y 26 se muestran la relación de los dos indicadores que han sido identificados como los más relacionados a la clasificación (Num\_msg y Num\_reply\_msg) para dejar constancia de la claridad de la clasificación de los estudiantes en tres cluster o grupos.



**Figura 25: relación entre los indicadores estadísticos Num\_msg y Num\_reply\_msg en el caso de clasificar los datasets sin filtrar de equipos poco activos en 3 cluster.**



**Figura 26: relación entre los indicadores estadísticos Num\_msg y Num\_reply\_msg en el caso de clasificar los datasets filtrados de equipos poco activos en 3 cluster.**

En las gráficas que se muestran en las figuras 25 y 26 se representan las instancias donde el eje X es el indicador estadístico Num\_msg y el eje Y el indicador estadístico Num\_reply\_msg. Las instancias están coloreadas según el cluster en donde están recogidas. En todos los casos se observa que los cluster están, más o menos, ordenados y diferenciados según se incrementa el valor en los indicadores. Primero se encuentran las instancias recogidas en el cluster de valores en los indicadores más bajos, después la recogidas en el cluster de valores medios, y, por último, están las instancias recogidas por el cluster de valores extremos. Esto se verá más claro en las tablas 15 y 16. La ordenación no es perfecta y los cluster se solapan, pero se identifican bien por lo que se puede decir que los indicadores Num\_msg y Num\_reply\_msg son los más significativos en relación a la clasificación. No se muestra las figuras 25 y 26 en relación con otros indicadores porque la ordenación de

las instancias según los cluster no era tan significativa. No se observa diferentes significativas entre los datasets sin filtrar (Figura 25) de los filtrados (Figura 26).

Falta mostrar la relación de las medidas de la colaboración realizadas por el tutor con los grupos o clusters. El mecanismo ha sido medir la media de la valoración de la colaboración en cada grupo o cluster y observar si la relación se repetía en dataset de distintos cursos. En las tablas 15 y 16 se muestran los promedios de la valoración de la colaboración por cluster. Para aumentar la claridad, los clusters han sido etiquetados según el número de mensajes, por lo que tendremos un cluster muy interactivo o de interacción alta (C-I-A), donde las instancias agrupadas son las que han enviado más mensajes a los foros, otro cluster de interacción media (o C-I-M), donde las instancias agrupadas son las que han enviado a los foros un número de mensajes dentro de promedio, y un cluster de interacción baja (o C-I-B), donde las instancias agrupadas son las que han enviado menos mensajes a los foros.

**Tabla 15: relación entre la lista suministrada por el experto sobre la colaboración de los alumnos y los cluster obtenidos de los datasets sin filtrar de equipos poco activos.**

	D-I-06-07			D-I-07-08			D-I-08-09		
	Nº	Media	Varianza	Nº	Media	Varianza	Nº	Media	Varianza
C-I-A	19	3,26	0,61	36	3,89	2,37	19	3,26	0,51
C-I-M	60	3,85	1,32	56	4,39	1,02	51	4,25	0,78
C-I-B	38	5,62	1,47	29	5,21	2,16	42	4,86	0,95

**Tabla 16: relación entre la lista suministrada por el experto sobre la colaboración de los alumnos y los cluster obtenidos de los datasets filtrado de equipos poco activos.**

	D-II-06-07			D-II-07-08			D-II-08-09		
	Nº	Media	Varianza	Nº	Media	Varianza	Nº	Media	Varianza
C-I-A	17	3,29	0,68	31	3,81	2,03	14	3,57	0,53
C-I-M	37	3,58	0,93	45	4,49	1,01	25	3,80	0,96
C-I-B	26	4,27	1,35	27	5,40	1,54	49	4,51	0,94

En las tablas 15 y 16 se puede apreciar que cada cluster se diferencia de los otros según su nivel de colaboración promedio. La columna “Nº” representa el número de instancias recogidas por el cluster en cuestión. La columna “Media” representa la media del valor del nivel de colaboración según el análisis del experto de todas las instancias del cluster, y la columna “Varianza” la varianza del nivel de colaboración. Por ejemplo, para el dataset D-II-08-09 el valor medio del nivel de colaboración etiquetado por el experto es para el cluster C-I-A de 3,57, para el cluster C-I-M de 3,80, y de 4,51 para el cluster C-I-B. Independientemente del dataset del que se trate,

el cluster de mayor interacción, que agrupa instancias con valores más altos en los indicadores estadísticos, está señalado por el promedio del nivel de colaboración más bajo. Se advierte que debido a la escala dada por el experto, un nivel de colaboración con un número bajo, significa una alta colaboración. También, independientemente del dataset, todos los clusters de poca interacción, los cuales agrupan instancias con valores bajos en los indicadores estadísticos, contienen instancias etiquetadas con menor colaboración.

De esto se deduce que los agrupamientos representan distintos niveles de colaboración. Esta clasificación se repite en todos los clusters filtrados o no filtrados de los cursos 2006/2007, 2007/2008 y 2008/2009. Los agrupamientos con instancias, cuyos indicadores estadísticos son más altos, agrupan a los alumnos más colaborativos. Los clusters con instancias, cuyos indicadores toman valores medios, agrupan alumnos de colaboración media. Los clusters con instancias, cuyos indicadores estadísticos tienen los valores más bajos, agrupan a estudiantes de muy poca colaboración. Esto indica que la clasificación por clustering de los indicadores estadísticos de las interacciones en los foros está relacionada con la colaboración. Al repetirse el mismo tipo de clasificación en los tres cursos, se puede decir que el método de inferencia de clasificación por clustering está validado como medio de clasificar a los estudiantes según su colaboración.

Sin embargo, la varianza que se muestran en las tablas 15 y 16 indica el grado de aproximación del acercamiento. A mayor varianza menor es la exactitud de la predicción del nivel de colaboración. Según la varianza, cualquier instancia del cluster con grado de interacción media podría contener a alumnos con un nivel de colaboración medio, pero también alto o bajo. Por lo tanto, el factor de incertidumbre de las clasificaciones sigue siendo significativo.

De lo anterior, se puede concluir que el método de inferencia de la colaboración por clustering está validado pero que sus inferencias son aproximadas.

## **Datos de las herramientas de autorregulación**

Para alcanzar el objetivo de mejorar el aprendizaje colaborativo se ha propuesto un análisis con el que obtener valoraciones o inferencias sobre la colaboración. Los

procedimientos descritos previamente para alcanzar dicho objetivo mediante inferencias no son suficientes. Hay que establecer una estrategia pedagógica que, utilizando las inferencias, ayude a los estudiantes y haga mejorar el aprendizaje colaborativo.

Hemos propuesto la estrategia de mostrar al estudiante información sobre la colaboración en herramientas de autorregulación con el objetivo de mejorar la gestión del proceso de colaboración. Se han discutido ya las ventajas de la autorregulación en el aprendizaje (Boekaerts et al., 2000), así como el beneficio de mostrar información metacognitiva en los entornos de aprendizaje para mejorar el aprendizaje (Papadopoulos et al., 2009).

Tomando como base dichos fundamentos un conjunto de herramientas de autorregulación fueron ofrecidas a los alumnos del curso 2008/2009. Durante el periodo colaborativo se fueron analizando las interacciones en los foros de los distintos equipos colaborativos y se aplicó el acercamiento de clustering para inferir el nivel de colaboración de los alumnos.

Se utilizó el acercamiento de clasificar a los estudiantes según su colaboración mediante el método de clustering porque ofreció buenos resultados en los cursos 2006/2007 y 2007/2008, aunque aproximados. El acercamiento de las métricas, según se ha diseñado, necesitaba un número suficientes de pruebas con distintos datasets y se decidió probar su validez con los datasets de los tres cursos antes de empezar a mostrar sus resultados a los estudiantes.

Para probar las distintas herramientas se ofrecieron alternativamente a distintos equipos de la experiencia colaborativa del curso 2008/2009. Los equipos fueron seleccionados de forma aleatoria para que entre seis a nueve equipos disfrutaran de cada una de las herramientas, dejando a unos pocos equipos sin ninguna herramienta para tomarlos como grupo de control (Kirk, 1995). El uso de un grupo de control nos indicará la utilidad de las herramientas en sí, sin necesidad de comparar las herramientas, ya que esto nos establece la utilidad entre unas u otras.

Las herramientas ofrecidas fueron una aplicación web, un portlet mostrando los indicadores estadísticos, otro portlet mostrando el nivel inferido de colaboración de los alumnos, y los dos portlets anteriores ofrecidos a los mismos equipos.

La evaluación de los sistemas adaptativos es todavía un campo abierto (van Velsen et al., 2008) y no hay un método que dé mejores resultados en vez de otro. En concreto, en esta investigación para la evaluación de las herramientas se ha tenido en cuenta las respuestas a las encuestas de opinión que se suministraron a los alumnos a los que se les ofreció cada una de las herramientas, y las notas finales tanto de la experiencia de aprendizaje colaborativo como la del examen del primer parcial de la asignatura IA-IC.

A continuación se describe la experimentación realizada durante el curso 2008/2009 señalando el número de estudiantes a los que se les ofreció cada herramienta, cuándo, y cómo se han evaluado las herramientas. Se hace un análisis detallado de la evaluación según la herramienta ofrecida.

Antes hay que advertir que en la experiencia de aprendizaje colaborativo de 2007/2008 ya se ofreció a los estudiantes una versión preliminar de la aplicación web. Sin embargo, problemas de potencia de máquina hizo imposible su utilización por más de cinco estudiantes. En la experiencia del curso 2008/2009 se utilizaron máquinas de mayor potencia, además de ampliar la oferta de herramientas de autorregulación.

## Experimentación

Para probar las distintas herramientas de autorregulación se ofrecieron a algunos equipos de la experiencia colaborativa del curso 2008/2009. A unos equipos se ofreció la aplicación web, a otros el portlet con la información de los indicadores, a distintos equipos el portlet con la información de los niveles de colaboración, y, por último, se les dio a otros equipos el portlet con los datos de los indicadores y con los niveles de colaboración. En la Tabla 17 se muestra información sobre los alumnos que disfrutaron de las distintas herramientas.

**Tabla 17: datos sobre la experimentación con las herramientas metacognitivas.**

	Nº alumnos	Inicio	Final
Aplicación web sin niveles de colaboración	12*	24/11/2009	25/01/2009
Aplicación web con niveles de colaboración	15	15/12/2008	25/01/2009
Portlet indicadores estadísticos	18	24/11/2009	25/01/2009
Portlet niveles de colaboración	21	15/12/2008	25/01/2009
Portlet indicadores estadísticos y niveles de colaboración	18**	24/11/2009	25/01/2009

\* A estos alumnos no se les ofreció la información de sus niveles de colaboración pero podían ver el nivel de colaboración de otros en la aplicación web, ya que todos los alumnos a los que se les ofreció la aplicación web, podían navegar por los datos de todos los compañeros, del mismo equipo o de otros.

\*\* A estos alumnos se les ofreció el 24 de noviembre de 2008 el portlet con los indicadores estadísticos, y el 15/12/2008 el portlet se completó con la información de los niveles de colaboración.

Al principio, la información que podían encontrar en la aplicación web era la introducida en la encuesta inicial. Una semana después se actualizaron los datos añadiendo las interacciones totales y las de la semana anterior de cada uno de los alumnos. La información en la aplicación web y en los portlets se actualizó cada nueva semana, siempre los lunes. A partir del 15 de diciembre, se añadió información del nivel de colaboración en la aplicación web y se ofrecieron portlets con esta información.

Estas herramientas se han evaluado por tres medios:

- Respuesta de los estudiantes a las encuestas finales. En las encuestas finales se les preguntaba por el uso, la opinión general sobre la herramienta y sobre la información que se mostraba.
- Evaluación del trabajo de los estudiantes en la experiencia de aprendizaje colaborativo. Los equipos debieron hacer una memoria final que el tutor de la experiencia evaluó entre 2 puntos y 0. En esta evaluación se tuvo en cuenta, además del trabajo en relación a la materia que se pedía en la experiencia, el trabajo colaborativo de los estudiantes.
- Evaluación en el examen del primer parcial. Éste tuvo que ser hecho por los estudiantes poco después de terminar la experiencia de aprendizaje colaborativo y tuvo una evaluación entre 10 y 0 puntos.

En los siguientes apartados se describen los resultados de la evaluación de las distintas herramientas que fueron ofrecidas.

## **Aplicación web**

Se ha explicado en profundidad la aplicación web que se ofreció a los alumnos en el curso 2008/2009. El total de alumnos a los que se ofreció la aplicación web fueron

27, a 15 de estos alumnos se les fue completando la información con los niveles de colaboración inferidos. El resto de alumnos también podían ver el nivel de colaboración de esos alumnos, aunque no el de ellos mismos.

La aplicación web se ofreció al comienzo de los equipos en el periodo colaborativo y cada lunes se fueron actualizando los datos de los indicadores estadísticos de las interacciones en los foros de los equipos. A partir del 15 de diciembre de 2008 se añadieron los niveles de colaboración inferidos, por lo que se completaba toda la información que contenía el modelo de colaboración (sobre el contexto de colaboración, sobre el proceso de colaboración y sobre la colaboración de los alumnos).

Para evaluar se preparó una encuesta de opinión y un análisis de las notas de los alumnos a los que se les ofreció la aplicación web.

### **Encuesta**

En la encuesta se hicieron varias preguntas para tener en cuenta el uso hecho por los alumnos, la usabilidad de la aplicación, la utilidad de la misma respecto a la colaboración. A continuación los resultados:

- La aplicación se ofreció a 27 alumnos.
- La encuesta se ofreció a 26 de los 27 alumnos. Uno de los alumnos abandonó la experiencia colaborativa durante el curso.
- Sólo 12 alumnos respondieron a la encuesta, un 46%.
- Para conocer el uso de la aplicación se les preguntó a los alumnos:
  - a. Con qué frecuencia entraban en la aplicación. Las posibles respuestas eran: varias veces por semana, apenas una vez por semana, cada dos semanas, apenas nunca y nunca. Las respuestas obtenidas fueron:
    - i. Varias veces por semana: 4.
    - ii. Apenas una vez por semana: 0.
    - iii. Cada dos semanas: 2.
    - iv. Apenas nunca: 5.
    - v. Nunca: 1.
  - b. Cuándo solían entrar en la aplicación. Las posibles respuestas fueron: al principio del periodo colaborativo, a mediados del



periodo colaborativo, hacia el final del periodo colaborativo, o no sé. Las respuestas fueron:

- vi. Al principio del periodo colaborativo: 7.
- vii. A mediados del periodo colaborativo: 1.
- viii. A finales del periodo colaborativo: 2.
- ix. No sé: 2.

- Para conocer la usabilidad de la aplicación se les preguntó a los alumnos:
  - c. Si habían entendido con facilidad la información que se mostraba. Las posibles respuestas eran: sí, mucho; sí, pero sólo algo; apenas. Las respuestas han sido:
    - x. Sí, mucho: 4.
    - xi. Sí, pero sólo algo: 7.
    - xii. Apenas: 1.
    - xiii. No, nada: 0.
  - d. Si les había resultado sencillo navegar por la aplicación. Las posibles respuestas eran: sí, mucho; sí, pero sólo algo; ha resultado un poco difícil; ha resultado muy difícil. Las respuestas han sido:
    - xiv. Sí, mucho: 3.
    - xv. Sí, pero sólo algo: 9.
    - xvi. Ha resultado un poco difícil: 0.
    - xvii. Ha resultado muy difícil: 0.
- Para saber sobre si la aplicación les ha sido útil a la hora de colaborar se les hicieron varias preguntas:
  - e. Sí la información mostrada ha ayudado a colaborar mejor. Las posibles respuestas eran: sí, mucho; sí, algo; apenas; no, nada. Las respuestas han sido:
    - xviii. Sí, mucho: 2.
    - xix. Sí, algo: 3.
    - xx. Apenas: 4.
    - xxi. No, nada: 3.
  - f. Si la información sobre los compañeros les ha ayudado a saber como colaboraban los compañeros. Las posibles respuestas eran: sí, mucho; sí, algo; apenas; no, nada. Las respuestas han sido:
    - xxii. Sí, mucho: 0.

- xxiii. Sí, algo: 3.
- xxiv. Apenas: 7.
- xxv. No, nada: 2.
- g. Si la herramienta ha facilitado el aprendizaje colaborativo y ha fomentado la interacción. Las posibles respuestas eran: sí, mucho; sí, un poco; ha sido un poco el efecto contrario; totalmente opuesto. Las respuestas dadas por los alumnos han sido:
  - xxvi. Sí, mucho: 2.
  - xxvii. Sí, un poco: 6.
  - xxviii. Ha sido un poco el efecto contrario: 4.
  - xxix. Totalmente opuesto: 0.
- También se ha preguntado sobre la utilidad de los indicadores estadísticos que se mostraban y se actualizan una vez por semana. Las posibles respuestas fueron: sí, muy útil; sí, ha sido útil; no, no ha sido útil; no ha sido para nada útil. Las respuestas de los alumnos han sido:
  - h. Sí, muy útil: 1.
  - i. Sí ha sido útil: 2.
  - j. No, no ha sido útil: 8.
  - k. No ha sido para nada útil: 0.
- Para terminar se preguntó sobre la utilidad de los niveles de colaboración que se mostraron y se actualizaban una vez por semana. Las posibles respuestas fueron: sí, muy útil; sí, ha sido útil; no, no ha sido útil; no ha sido para nada útil. Las respuestas de los alumnos han sido:
  - l. Sí, muy útil: 1.
  - m. Sí ha sido útil: 2.
  - n. No, no ha sido útil: 7.
  - o. No ha sido para nada útil: 2.

El 46% de los alumnos consultados respondieron. Aunque un 46% de respuestas es un número adecuado en muchos casos, nos encontramos que al final sólo ha habido 12 alumnos que respondieran. Por esa razón hemos realizado un análisis de la representatividad estadística de los alumnos que han respondido respecto a los que no lo han hecho. Hemos cogido dos fuentes de datos, el conjunto de indicadores estadísticos y los niveles de colaboración suministrados por el experto. Para calcular la representatividad estadística hemos utilizado la función T-test (Ipiña, 2008). Dos

conjuntos de datos son estadísticamente representativos el uno del otro si el resultado de aplicar t-test es superior a 0,05. Los resultados son:

- T-test del conjunto de datos de indicadores estadísticos: 0,0439.
- T-test del conjunto de datos de niveles de colaboración: 0,0004.
- Promedio de nivel de colaboración de los alumnos que han respondido a la encuesta: 4,25.
- Promedio de nivel de colaboración de los alumnos que no han respondido a la encuesta: 5,36.

Claramente se observa que los alumnos etiquetados como mejores colaboradores han respondido a la encuesta, mientras que los otros no. El conjunto de los alumnos que han respondido no es una muestra del conjunto.

De las respuestas dadas por los estudiantes a la encuesta final se puede deducir que:

- La aplicación web fue escasamente utilizada y principalmente a principios del periodo colaborativo.
- La facilidad de uso es apropiada pero mejorable.
- La información suministrada no mejora sensiblemente la colaboración.
- Los indicadores estadísticos no fueron muy útiles.
- Los niveles de colaboración inferidos no fueron muy útiles.

En general los alumnos suspenden la aplicación y sobre todo la información que muestra.

### **Notas**

El objetivo de la aplicación web era ayudar a los alumnos a mejorar la gestión de proceso colaborativo. La hipótesis principal de la estrategia de la colaboración es que el trabajo en equipo mejora el aprendizaje. La evaluación del aprendizaje es, por tanto, una evaluación también del trabajo colaborativo. Por eso mismo se ha realizado un análisis de las notas tanto del periodo colaborativo como del examen del primer parcial que se muestra en la siguiente tabla.

En la Tabla 18 aparecen los datos de población de los grupos y sus resultados en la evaluación de la experiencia colaborativa y del examen del primer parcial. Las filas representan el grupo de alumnos a los que se ha ofrecido la aplicación web, al resto de alumnos que participaron en la experiencia colaborativa, un conjunto de 28 alumnos,

9 equipos donde en uno había 4 miembros, los cuales formaron el grupo de control y a los que no se les ofreció ni la aplicación web ni los portlet, y el resto de alumnos que no participaron en la experiencia colaborativa. Según se observa en la Tabla 18 el conjunto de alumnos de control obtuvieron una valoración de la experiencia educativa cercana al 1,59 lo que indica que hicieron un trabajo colaborativo un poco por debajo de la media de todos los alumnos que participaron en la experiencia colaborativa, que fue de 1,62. Sin embargo la nota promedio obtenida en el examen 6,83 sí está bastante por encima de la nota promedio de todos los alumnos que participaron en el entorno colaborativo (6,38). Mientras que la diferencia entre las notas de la experiencia colaborativa es posible explicarlo, es normal que el grupo de control tenga un comportamiento cercano a la media, la diferencia en las notas del examen del primer parcial es complicado de explicar salvo pensando que al azar se eligieron unos alumnos que se prepararon mejor para el examen. Se informa que el examen constaba de distintas preguntas. En una parte del examen había que resolver un problema de planificación y responder a unas cuestiones de planificación. Esta parte era la relacionada directamente con la experiencia de aprendizaje colaborativo y tenía un valor que ha ido cambiando de curso en curso de 5 puntos a 6 puntos. En la evaluación del examen del primer parcial entran también variables dependientes de las que no se tiene control, como por ejemplo, el tiempo de estudio que el estudiante consiguió o la motivación.

**Tabla 18: resultados de la evaluación del trabajo colaborativo de los alumnos y del resultado del examen del primer parcial.**

	N	N-E	M-C	V-C	M-E	V-E
Alumnos a los que se les ofreció la aplicación web	27	19	1,53	0,09	5,72	1,66
Resto de alumnos del entorno colaborativo	88	76	1,65	0,06	6,54	2,31
Alumnos de los grupos de control del entorno colaborativo	28	26	1,59	0,06	6,83	1,77
Alumnos que no participaron en el entorno colaborativo		79			4,69	5,09

Leyenda:

- N: número de alumnos.
- N-E: número de alumnos presentados a examen.
- M-C: promedio de la nota parte colaborativa [0,2].
- V-C: Varianza del promedio de la nota de la parte colaborativa.

- M-E: promedio de la nota del examen [0,10].
- V-E: varianza del promedio de la nota del examen.

Lo que se observa llamativamente es que los alumnos a los que se les ofreció la aplicación web han obtenidos peores resultados, en promedio, tanto en la evaluación del periodo colaborativo como del examen del primer parcial que los otros alumnos que también participaron en el periodo colaborativo, tanto comparados con todo el conjunto restante como sólo con los alumnos de control (ver columnas 3 y 5 de las filas 2, 3 y 4). Sin embargo, estos alumnos mejoran los resultados obtenidos, en promedio, de los alumnos que no participaron en el periodo colaborativo (ver columna 5 de las filas 2 y 5). Se observa de los resultados de la varianza de la evaluación del trabajo colaborativo (0,09) que es la mayor, lo que indica que hubo más diferencias en la evaluación con los alumnos a los que se les ofreció la aplicación web. Esto indica que la aplicación web no sirvió de la misma forma a todo los estudiantes a los que se les ofreció. Sin embargo, la varianza en el caso de la evaluación del examen del primer parcial (1,66) es la menor, lo que indica que muchos de los estudiantes tuvieron una nota parecida por lo que su aprendizaje en general fue el mismo.

### **Portlet indicadores estadísticos**

El portlet de indicadores estadísticos fue ofrecido a 18 alumnos, 6 equipos. Este portlet mostraba los indicadores estadísticos por cada una de las semanas del periodo colaborativo además del total. Cada lunes se actualizaba el contenido añadiendo los indicadores estadísticos de la semana anterior y se actualizaba el valor de los indicadores estadísticos totales. Para evaluar la utilidad del portlet se ha preguntado a los alumnos a los que se les ofreció. Las respuestas son analizadas en el siguiente apartado. Además se ha analizado los resultados de la evaluación del periodo colaborativo y del examen del primer parcial.

### **Encuesta**

Se pidió a los alumnos de los equipos a los que se había ofrecido el portlet mostrando los indicadores estadísticos que respondieran una encuesta de opinión. La encuesta pedía, entre otras cosas, que valorara el portlet entre 0 (muy malo) y 5 (muy bueno). En la Tabla 19 se muestran los resultados de la valoración del portlet.

**Tabla 19: valoración del portlet de indicadores estadísticos.**

Nº alumnos preguntados	Nº alumnos que han respondido	Valoración media [0,5]	Varianza
18	9	3,3	0,22

La encuesta sólo fue respondida por la mitad de alumnos posibles, pero la valoración fue buena. Para saber si las respuestas de los 9 alumnos es representativa del total de 18 alumnos, hemos realizado el análisis T-test al igual que hicimos con la encuesta sobre la aplicación web. Los resultados son:

- T-test del conjunto de datos de indicadores estadísticos: 0,23.
- T-test del conjunto de datos de la lista de niveles de colaboración: 0,97.
- Promedio de nivel de colaboración de los alumnos que han respondido a la encuesta: 4,60.
- Promedio de nivel de colaboración de los alumnos que no han respondido a la encuesta: 4,57.

Al contrario que sucedía con las respuestas a la encuesta sobre la aplicación web, las respuestas a esta encuesta sí son representativas del total. Se observa, además, que los alumnos que han respondido están etiquetados como ligeramente menos colaboradores.

Además de preguntar por la valoración se les preguntó por la colaboración y la utilidad de la información presentada:

- Para saber sobre si la aplicación les ha sido útil a la hora de colaborar se les hicieron varias preguntas:
  - a. Sí la información mostrada ha ayudado a colaborar mejor. Las posibles respuestas eran: sí, mucho; sí, algo; apenas; no, nada. Las respuestas han sido:
    - i. Sí, mucho: 0.
    - ii. Sí, algo: 4.
    - iii. Apenas: 4.
    - iv. No, nada: 1.
  - b. Si la información sobre los compañeros les ha ayudado a saber como colaboraban los compañeros. Las posibles respuestas eran: sí, mucho; sí, algo; apenas; no, nada. Las respuestas han sido:
    - v. Sí, mucho: 1.
    - vi. Sí, algo: 5.

- vii. Apenas: 2.
- viii. No, nada: 1.
- c. Si la herramienta ha facilitado el aprendizaje colaborativo y ha fomentado la interacción. Las posibles respuestas eran: sí, mucho; sí, un poco; ha sido un poco el efecto contrario; totalmente opuesto. Las respuestas dadas por los alumnos han sido:
  - ix. Sí, mucho: 1.
  - x. Sí, un poco: 6.
  - xi. Ha sido un poco el efecto contrario: 1.
  - xii. Totalmente opuesto: 1.
- También se ha preguntado sobre la utilidad de los indicadores estadísticos que se mostraban y se actualizan una vez por semana. Las posibles respuestas fueron: sí, muy útil; sí, ha sido útil; no, no ha sido útil; no ha sido para nada útil. Las respuestas de los alumnos han sido:
  - d. Sí, muy útil: 0.
  - e. Sí ha sido útil: 5.
  - f. No, no ha sido útil: 4.
  - g. No ha sido para nada útil: 1.

No se preguntó sobre el uso y la facilidad de uso ya que el portlet estaba colocado en una de las secciones del espacio virtual de cada equipo y se podía acceder directamente al navegar por esa sección. En resumen, los resultados indican:

- La información suministrada ni mejora ni entorpece el conocimiento sobre la colaboración personal, y sí ayuda un poco en el caso de la información sobre los compañeros.
- Los indicadores estadísticos fueron útiles pero sólo sensiblemente, es decir, sí, pero un poco.

Aunque la valoración dada por la puntuación de los alumnos es de un bien alto, según el resto de respuestas, el portlet obtendría un aprobado raspado.

### **Notas**

También se ha hecho un análisis de las notas de estos alumnos para comprobar si las posibles ventajas colaborativas del portlet se ven representadas en un mejor conocimiento.

**Tabla 20: evaluación de los alumnos con el portlet de indicadores estadísticos.**

	N	N-E	M-C	V-C	M-E	V-E
Alumnos a los que se les ofreció el portlet	18	15	1,56	0,10	6,24	2,76
Resto de alumnos del entorno colaborativo	97	80	1,63	0,06	6,41	2,20
Alumnos de los grupos de control del entorno colaborativo	29	26	1,59	0,06	6,83	1,77
Alumnos que no participaron en el entorno colaborativo		79			4,69	5,09

Leyenda:

- N: número de alumnos.
- N-E: número de alumnos presentados a examen.
- M-C: promedio de la nota parte colaborativa [0,2].
- V-C: Varianza del promedio de la nota de la parte colaborativa.
- M-E: promedio de la nota del examen [0,10].
- V-E: varianza del promedio de la nota del examen.

Se observa que al igual, que los alumnos a los que se ofreció la aplicación web las notas tanto del periodo colaborativo como del examen del primer parcial son peores de los alumnos a los que se ha ofrecido el portlet de indicadores estadísticos que al resto de alumnos de la experiencia colaborativa (ver columnas 3 y 5 de las filas 2, 3 y 4). Sin embargo, la diferencia entre alumnos ha disminuido, respecto al caso de los alumnos con la aplicación web. Eso sí, las notas del examen del primer parcial siguen siendo mejores que la de los alumnos que no han participado en la experiencia colaborativa (ver columna 5 de las filas 2 y 5). La alta varianza en la evaluación del trabajo colaborativo (0,10) señala que el portlet con los indicadores estadísticos no ayudo a todos los estudiantes por igual, aunque al ser sólo 18, el juicio se puede alcanzar al comparar con otro grupo de población semejante, como es el caso del grupo de control. El grupo de control tuvo una media en la evaluación del trabajo colaborativo más alta (1,59) y una varianza más baja (0,06). Se deduce que el portlet que mostraba los indicadores estadísticos no ayudo mucho a los estudiantes a los que se les ofreció, y lo hizo de forma desigual.



## **Portlet niveles de colaboración**

El portlet que mostraba los niveles de colaboración inferidos con el acercamiento de clustering se ofreció a 21 alumnos, 7 equipos. A mediados de diciembre de 2008 se empezó a ofrecer dicha información. Era necesario una cierta cantidad de datos para que se produjeran unos clusters relacionados con la colaboración. Cada lunes se fue actualizado la información hasta el final del periodo colaborativo. El portlet aparecía en una sección del espacio virtual del equipo y ver dicha información era tan fácil como ver los mensajes del foro o quienes eran los compañeros del equipo. Al igual que el portlet de los indicadores estadísticos, este portlet se ha evaluado analizando las respuestas a una encuesta donde se pedía la opinión sobre el portlet, y analizando la evaluación de la experiencia colaborativa y del examen del primer parcial.

### **Encuesta**

Se pidió a los alumnos de los equipos a los que se había ofrecido el portlet mostrando los niveles de colaboración inferidos que respondieran una encuesta de opinión. La encuesta pedía, entre otras cosas, que valorara el portlet entre 0 (muy malo) y 5 (muy bueno). En la Tabla 21 se muestran los resultados de la valoración del portlet.

**Tabla 21: valoración del portlet de los niveles de colaboración inferidos.**

Nº alumnos preguntados	Nº alumnos que han respondido	Valoración media	Varianza
21	13	3	1,45

La encuesta fue respondida por más de la mitad de alumnos posibles, y la valoración fue buena, aunque la alta varianza indica que no hubo unanimidad. Para saber si las respuestas de los 13 alumnos es representativa del total de 23 alumnos, hemos realizado el análisis T-test al igual que hicimos con la encuesta sobre la aplicación web. Los resultados son:

- T-test del conjunto de datos de indicadores estadísticos: 0,0699.
- T-test del conjunto de datos de niveles de colaboración: 0,0480.
- Promedio de nivel de colaboración de los alumnos que han respondido a la encuesta: 3,83.
- Promedio de nivel de colaboración de los alumnos que no han respondido a la encuesta: 4,73.

No se llega a un acuerdo de si las respuestas de los 13 alumnos son representativas de todos los alumnos a los que se les ofreció el portlet, ya que en la función T-test ofrece al aplicarse a los datos de los indicadores estadísticos un valor que indica que hay representatividad estadística (0,0699), pero en el caso de los datos de los niveles de colaboración (0,0480) no. Por tanto, los alumnos que sí respondieron tienen un grado de colaboración más alto de los que no respondieron.

Además de preguntar por la valoración se les preguntó por la colaboración y la utilidad de la información presentada:

- Para saber sobre si la aplicación les ha sido útil a la hora de colaborar se les hicieron varias preguntas:
  - a. Sí la información mostrada ha ayudado a colaborar mejor. Las posibles respuestas eran: sí, mucho; sí, algo; apenas; no, nada. Las respuestas han sido:
    - i. Sí, mucho: 1.
    - ii. Sí, algo: 7.
    - iii. Apenas: 2.
    - iv. No, nada: 2.
  - b. Si la información sobre los compañeros les ha ayudado a saber como colaboraban los compañeros. Las posibles respuestas eran: sí, mucho; sí, algo; apenas; no, nada. Las respuestas han sido:
    - v. Sí, mucho: 1.
    - vi. Sí, algo: 7.
    - vii. Apenas: 2.
    - viii. No, nada: 2.
  - c. Si la herramienta ha facilitado el aprendizaje colaborativo y ha fomentado la interacción. Las posibles respuestas eran: sí, mucho; sí, un poco; ha sido un poco el efecto contrario; totalmente opuesto. Las respuestas dadas por los alumnos han sido:
    - ix. Sí, mucho: 1.
    - x. Sí, un poco: 8.
    - xi. Ha sido un poco el efecto contrario: 2.
    - xii. Totalmente opuesto: 0.

- Para terminar se preguntó sobre si los niveles de colaboración que se mostraron y se actualizaban una vez por semana les parecían correctos o aproximados. Las posibles respuestas fueron: sí, muy correctos; sí, pero aproximados; no, eran poco aproximados; no, eran incorrectos. Las respuestas de los alumnos han sido:
  - d. Sí, muy correctos: 2.
  - e. Sí, pero aproximados: 6.
  - f. No, eran poco aproximados: 3.
  - g. No, eran incorrectos: 1.

A continuación se da un análisis de las respuestas a la encuesta. No se preguntó sobre el uso y la facilidad de uso ya que el portlet estaba colocado en una de las secciones del espacio virtual de cada equipo y se podía consultar con navegar por esa sección. En resumen, los resultados indican que:

- Al contrario de las dos herramientas anteriores, la información que se suministra es positivamente valorada, aunque mejorable.
- Los niveles de colaboración inferidos también son positivamente valorados.

Tanto la valoración dada por la puntuación de los alumnos como las respuestas indican que el portlet obtendría un bien. Por ahora este portlet mejora los resultados de las otras dos herramientas.

### Notas

También se ha hecho un análisis de las notas de estos alumnos para comprobar si las posibles ventajas colaborativas del portlet se ven representadas en un mejor conocimiento.

**Tabla 22: evaluación de los alumnos con el portlet de niveles de colaboración inferidos.**

	N	N-E	M-C	V-C	M-E	V-E
Alumnos a los que se les ofreció el portlet	21	21	1,71	0,02	6,21	2,46
Resto de alumnos del entorno colaborativo	91	74	1,60	0,08	6,43	2,24
Alumnos de los grupos de control del entorno colaborativo	29	26	1,59	0,06	6,83	1,77
Alumnos que no participaron en el entorno colaborativo		79			4,69	5,09

Leyenda:

- N: número de alumnos.

- N-E: número de alumnos presentados a examen.
- M-C: promedio de la nota parte colaborativa [0,2].
- V-C: Varianza del promedio de la nota de la parte colaborativa.
- M-E: promedio de la nota del examen [0,10].
- V-E: varianza del promedio de la nota del examen.

Los alumnos con este portlet obtuvieron mejores notas en la experiencia colaborativa que el resto de alumnos de la experiencia colaborativa, pero peor que estos otros en el examen del primer parcial (ver columnas 3 y 5 de las filas 2, 3 y 4). Por el escaso valor de la varianza de las notas de la experiencia colaborativa (0,02), se observa con claridad que el trabajo colaborativo fue evaluado como mejor que el del resto de alumno de la experiencia sin este portlet. Se puede deducir, entonces, que los alumnos con este portlet colaboraron bastante bien en comparación con el resto. En las notas del examen del primer parcial, sin embargo, no se puede verificar nada en concreto ya que la gran varianza impide tal juicio (2,46). Los alumnos con el portlet de niveles de colaboración inferidos obtuvieron en general peores notas pero la diferencia es muy pequeña.

### **Portlet indicadores estadísticos y niveles de colaboración**

A un conjunto de 18 alumnos se les ofreció los portlets que muestran los indicadores estadísticos y los niveles de colaboración inferidos, es decir, los portlets de las dos secciones anteriores a la vez en la misma sección del espacio virtual del equipo. Como en el caso de las secciones anteriores, se actualizaron los datos cada lunes durante el periodo colaborativo. A estos alumnos, se les empezó a mostrar primero el portlet de los indicadores estadísticos y después, a partir del 15 de diciembre, se ofreció también el portlet de los niveles de colaboración inferidos. Para evaluar los portlets también se preparó una encuesta de opinión y se ha hecho un análisis de las respuestas y de las evaluaciones de la experiencia colaborativa y del examen del primer parcial.

#### **Encuesta**

Se pidió a los alumnos de los equipos a los que se había ofrecido los portlets mostrando los indicadores estadísticos y los niveles de colaboración inferidos que respondieran una encuesta de opinión. La encuesta pedía, entre otras cosas, que

valorara los portlets entre 0 (muy malo) y 5 (muy bueno). En la Tabla 23 se muestran los resultados de la valoración de los portlets.

**Tabla 23: valoración de los portlets de los indicadores estadísticos y de los niveles de colaboración inferidos.**

Nº alumnos preguntados	Nº alumnos que han respondido	Valoración media	Varianza
18	12	3,08	1,08

La encuesta fue respondida por más de la mitad de alumnos posibles, y la valoración fue buena, aunque la alta varianza indica que no hubo unanimidad. Para saber si las respuestas de los 12 alumnos es representativa del total de 18 alumnos, hemos realizado el análisis T-test al igual que hicimos con la encuesta sobre la aplicación web. Los resultados son:

- T-test del conjunto de datos de indicadores estadísticos: 0,67.
- T-test del conjunto de datos de niveles de colaboración: 0,07.
- Promedio de nivel de colaboración de los alumnos que han respondido a la encuesta: 3,58.
- Promedio de nivel de colaboración de los alumnos que no han respondido a la encuesta: 4,50.

Se ve que la respuestas de los 12 alumnos si es representativa de la de todos los alumnos preguntados. Se observa, además, que los alumnos que han respondido están etiquetados como ligeramente más colaboradores.

Además de preguntar por la valoración se les preguntó por la colaboración y la utilidad de la información presentada:

- Para saber sobre si la aplicación les ha sido útil a la hora de colaborar se les hicieron varias preguntas:
  - a. Sí la información mostrada ha ayudado a colaborar mejor. Las posibles respuestas eran: sí, mucho; sí, algo; apenas; no, nada. Las respuestas han sido:
    - i. Sí, mucho: 2.
    - ii. Sí, algo: 5.
    - iii. Apenas: 2.
    - iv. No, nada: 3.

- b. Si la información sobre los compañeros les ha ayudado a saber como colaboraban los compañeros. Las posibles respuestas eran: sí, mucho; sí, algo; apenas; no, nada. Las respuestas han sido:
- v. Sí, mucho: 1.
  - vi. Sí, algo: 6.
  - vii. Apenas: 1.
  - viii. No, nada: 4.
- c. Si la herramienta ha facilitado el aprendizaje colaborativo y ha fomentado la interacción. Las posibles respuestas eran: sí, mucho; sí, un poco; ha sido un poco el efecto contrario; totalmente opuesto. Las respuestas dadas por los alumnos han sido:
- ix. Sí, mucho: 3.
  - x. Sí, un poco: 8.
  - xi. Ha sido un poco el efecto contrario: 0.
  - xii. Totalmente opuesto: 1.
- También se ha preguntado sobre la utilidad de los indicadores estadísticos que se mostraban y se actualizan una vez por semana. Las posibles respuestas fueron: sí, muy útil; sí, ha sido útil; no, no ha sido útil; no ha sido para nada útil. Las respuestas de los alumnos han sido:
    - d. Sí, muy útil: 0.
    - e. Sí ha sido útil: 5.
    - f. No, no ha sido útil: 5.
    - g. No ha sido para nada útil: 2.
  - Para terminar se preguntó sobre la utilidad de los niveles de colaboración que se mostraron y se actualizaban una vez por semana. Las posibles respuestas fueron: sí, muy útil; sí, ha sido útil; no, no ha sido útil; no ha sido para nada útil. Las respuestas de los alumnos han sido:
    - h. Sí, muy útil: 2.
    - i. Sí ha sido útil: 4.
    - j. No, no ha sido útil: 3.
    - k. No ha sido para nada útil: 3.
  - Para terminar se preguntó sobre si los niveles de colaboración que se mostraron y se actualizaban una vez por semana les parecían correctos o aproximados. Las posibles respuestas fueron: sí, muy correctos; sí, pero aproximados; no,

eran poco aproximados; no, eran incorrectos. Las respuestas de los alumnos han sido:

- l. Sí, muy correctos: 0.
- m. Sí, pero aproximados: 7.
- n. No, eran poco aproximados: 1.
- o. No, eran incorrectos: 4.

A continuación se da un análisis de las respuestas a la encuesta. No se preguntó sobre el uso y la facilidad de uso ya que el portlet estaba colocado en una de las secciones del espacio virtual de cada equipo y se podía navegar directamente por esa sección.

- Al igual que pasó con las opiniones sobre la información suministrada en el portlet de niveles de colaboración inferidos, cuando se muestran los dos portlet a la vez, la valoración de la información y la ayuda de esta a la colaboración es positiva, aunque mejorable.
- La opinión sobre los indicadores estadísticos es muy parecida en este caso que al de sólo mostrar el portlet de los indicadores estadísticos. Las opiniones se dividen entre algo buenos y algo malos.
- Los niveles de colaboración inferidos son considerados acertados pero de utilidad no muy clara.

Las respuestas a esta encuesta es una mezcla de las respuestas de los alumnos que sólo se les ofreció un portlet, o el de indicadores estadísticos o el de niveles de colaboración inferidos. De la puntuación de los alumnos como las respuestas indican que el portlet obtiene una buena evaluación .

### **Notas**

También se ha hecho un análisis de las notas de estos alumnos para comprobar si las posibles ventajas colaborativas del portlet se ven representadas en un mejor conocimiento.

**Tabla 24: evaluación de los alumnos con los portlet de indicadores estadísticos y niveles de colaboración inferidos.**

	N	N-E	M-C	V-C	M-E	V-E
Alumnos a los que se les ofreció el portlet	18	16	1,74	0,02	6,84	2,10
Resto de alumnos del entorno colaborativo	97	79	1,60	0,07	6,29	2,28
Alumnos de los grupos de control del entorno colaborativo	29	26	1,59	0,06	6,83	1,77
Alumnos que no participaron en el entorno colaborativo		79			4,69	5,09

Leyenda:

- N: número de alumnos.
- N-E: número de alumnos presentados a examen.
- M-C: promedio de la nota parte colaborativa [0,2].
- V-C: Varianza del promedio de la nota de la parte colaborativa.
- M-E: promedio de la nota del examen [0,10].
- V-E: varianza del promedio de la nota del examen.

Los alumnos a los que se les ofreció a la vez los portlets de los indicadores estadísticos y de los niveles de colaboración inferidos han sido evaluados como los que hicieron el mejor trabajo colaborativo y obtuvieron mejores notas en el examen del primer parcial. La varianza en el caso de la evaluación del trabajo colaborativo indica que estos alumnos sí hicieron los mejores trabajos por lo que se deduce que colaboraron mejor. Sin embargo, debido a la varianza en el caso de las notas del examen del primer parcial no se puede asegurar que de forma absoluta que estos alumnos aprendieron mejor que los otros, aunque hayan obtenido en promedio mejores notas. Hay que advertir que en el examen no se pudieron controlar otras variables dependientes por lo que los resultados tienen siempre un grado de arbitrariedad que en estos análisis que se desea menguar usando la evaluación de un conjunto grande de estudiantes.

## **Resultados del entorno colaborativo**

Para descubrir si el entorno colaborativo ha mejorado el aprendizaje utilizaremos dos variables: el resultado del examen del primer parcial de la asignatura IA-IC y el



número de abandonos. El abandono es una característica de los alumnos de la UNED, y de todos los alumnos en general, que consiste en no terminar los estudios empezado. Aquí consideraremos a los alumnos que no se han presentado a examen. De los alumnos matriculados en la UNED cerca del 50% (Callejo, 2001) no se presentan a los exámenes.

Además de evaluar la validez de la experiencia de aprendizaje colaborativo con las evaluaciones realizadas a los alumnos, tendremos en cuenta su opinión como variable de la experiencia. Por esta razón se preguntó a los alumnos en las encuestas finales, una por curso, que valoraran la experiencia en general, el trabajo del equipo docente y de la ayuda suministrada por la principal fuente de comunicación, los foros. Los resultados se detallan a continuación.

**Tabla 25: valoración de los alumnos a la experiencia.**

	2006/2007	2007/2008	2008/2009
Nº de respuestas	49	45	72
Valoración media de la experiencia	3,98	3,62	3,65
Varianza de la valoración de la experiencia	0,51	0,77	0,71
Valoración media del equipo docente	4,06	3,42	3,57
Varianza de la valoración del equipo docente	0,67	0,96	0,80
Valoración de los foros como herramienta de comunicación	4,06	3,76	3,5
Varianza de la valoración de los foros	0,83	1,25	1,08

La valoración máxima que se podía dar era de 5 y la mínima de 0. Se observa que la valoración, año tras año, ha sido positiva. La media de la valoración general de la experiencia de aprendizaje colaborativo, el equipo docente que gestionaba la experiencia, y los foros como instrumento de ayuda, siempre ha obtenido un resultado muy aceptable. Sin embargo, las respuestas de los alumnos siempre pueden estar sesgadas, por ejemplo, por los intentos de no disgustar al equipo docente antes de corregir los exámenes. Por eso mismo también se analizan los resultados obtenidos por los alumnos en la experiencia colaborativa y en el examen del primer parcial.

En la Tabla 26 se comparan los resultados en los exámenes del primer parcial entre alumnos que realizaron la práctica por el entorno colaborativo con los que no la realizaron.

**Tabla 26: comparación de resultados de los alumnos de la experiencia colaborativa respecto a los que no hicieron dicha experiencia.**

	2006/2007	2007/2008	2008/2009
Nº alumnos asignatura	476	428	412
Nº alumnos entorno colaborativo	125	140	115
Nº alumnos presentados a examen	179	189	178
Nº alumnos del entorno colaborativo presentados a examen	91	93	95
% de alumnos presentados a examen	37,61%	44,16%	43,20%
% de alumnos del entorno colaborativo que se han presentado a examen	72,8%	66,43%	82,61%
Media de la nota del trabajo colaborativo	1,60	1,60	1,62
Media notas	4,74	4,69	5,59
Media de las notas de alumnos fuera del entorno colaborativo	3,65	3,73	4,69
Media alumnos del entorno colaborativo	5,80	5,58	6,38

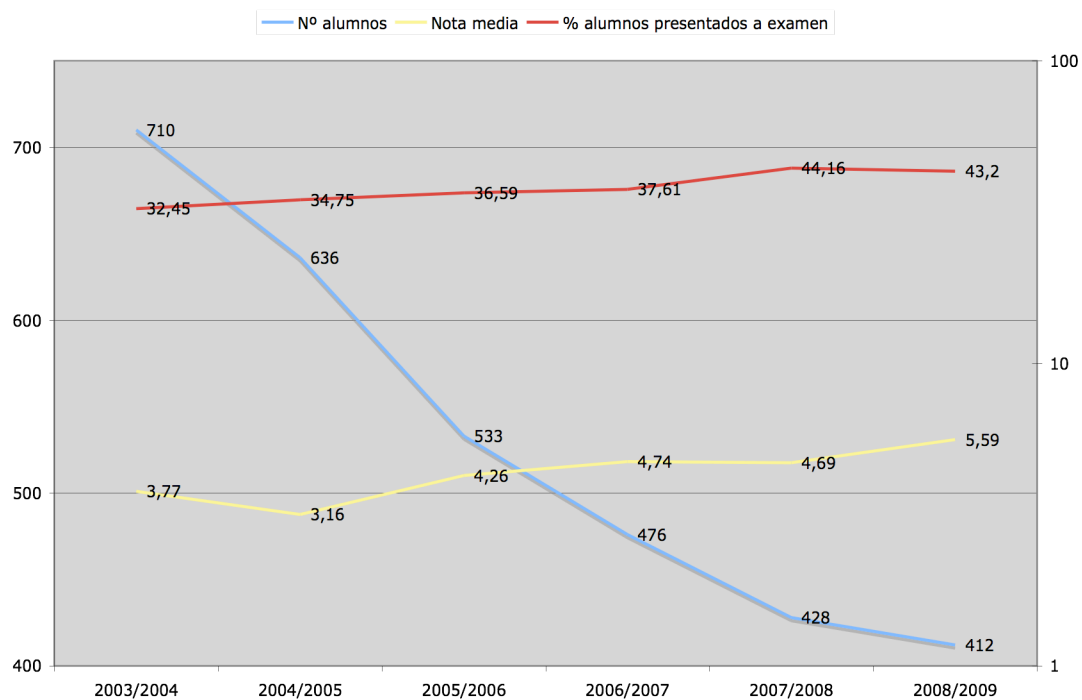
Como se puede ver, en la Tabla 26 se informa del número de alumnos matriculados en los cursos 2006/2007, 2007/2008 y 2008/2009, los que comenzaron el periodo colaborativo en la experiencia de aprendizaje colaborativo, el número y porcentaje de estudiantes presentados a examen, el porcentaje de presentados a examen de los que participaron en la segunda parte de la experiencia de aprendizaje colaborativo, la media de las notas en la evaluación del trabajo colaborativo, y la media en el examen del primer parcial de todos los estudiantes, de los que no participaron en la experiencia de aprendizaje colaborativo y de los que sí.

Para completar la descripción de los resultado en la Tabla 27 se presentan los datos de los tres cursos anteriores a la oferta del entorno colaborativo, es decir, los cursos 2003/2004, 2004/2005 y 2005/2006. Al no tener más información sólo se muestra el número de alumnos y presentados a examen del primer parcial, el porcentaje de presentados y la media.

**Tabla 27: resultados de los alumnos de cursos anteriores.**

	2003/2004	2004/2005	2005/2006
Nº de alumnos de la asignatura	710	636	533
Nº de presentados al examen	231	221	195
Nota media	3,77	3,16	4,26
% de alumnos presentados a examen	32,45%	34,75	36,59

En la Figura 27 se comparan el número de alumnos en los últimos 6 cursos, la nota medio de examen del primer parcial, y la tasa de o porcentaje de alumnos presentados.



**Figura 27: comparativa de resultados en los últimos seis cursos.**

Se observa que el número de alumnos que se matriculan en la asignatura de IA-IC ha disminuido en los últimos cursos al contrario que sucede con el porcentaje de presentados a examen y la nota media en el examen del primer parcial.

Pero si comparamos las notas medias en el examen del primer parcial de los estudiantes de los cursos 2003/2004, 2004/2005 y 2005/2006 (3,77, 3,16 y 4,26 respectivamente) con la nota media en el mismo examen de los estudiantes que no participaron en la experiencia de aprendizaje colaborativo de los cursos 2006/2007, 2007/2008 y 2008/2009 (3,65, 3,73 y 4,69 respectivamente), el incremento de la nota media deja de poder relacionarse con el descenso del número de matriculados. Se podría concluir que el incremento de la nota media en el examen del primer parcial se ha podido deber a la experiencia de aprendizaje colaborativo, por lo que se podría decir que la experiencia de aprendizaje colaborativo ha incrementado el aprendizaje.

Sin embargo, este análisis es muy superficial ya que no tiene en cuenta otras variables dependientes relacionadas con el aprendizaje. Es muy difícil conseguir resultados que muestren inequívocamente la correlación entre aprendizaje y el acceso a la información suministrada, si bien los indicios obtenidos se consideran esperanzadores. Hay que considerar que existen otras muchas variables dependientes, como por ejemplo el tiempo de dedicación, que pueden depender de situaciones

personales que no han sido controladas por la dificultad inherente y por no perturbar más a los alumnos. Hay que tener en cuenta que son estudiantes siguiendo un modelo de educación a distancia a los que se les suministra una docencia pero no se puede controlar cómo la utilizar o cuánto tiempo trabaja con ella (Barkley et al., 2004; García Aretio, 2004).

# Capítulo V: valoración

En este capítulo se recogen y resumen las consecuencias de la investigación deducidas a partir de los resultados obtenidos. En esta investigación se han probado diferentes acercamientos tanto, en el análisis de la colaboración como en herramientas que apoyen a los estudiantes en la gestión del proceso de colaboración. Una vez que los resultados se han descrito, la evaluación de esos resultados de forma individual es necesaria. En este capítulo se hará un análisis crítico de los resultados obtenidos para validar los acercamientos propuestos.

## Validez de la prospección de la colaboración

El análisis de la colaboración en entornos de aprendizaje colaborativo es todavía un campo de la investigación falto de estándares y metodología (Strijbos y Fisher, 2007). En esta investigación se ha experimentado con dos acercamientos para analizar, mediante técnicas de minería de datos, las interacciones de los alumnos en la experiencia de aprendizaje colaborativo.

Los objetivos de la prospección han sido:

1. obtener resultados que indiquen la diferente colaboración de los alumnos,
2. que los resultados sean obtenidos mientras la colaboración se está realizando o poco después, y
3. que el acercamiento de minería de datos pueda ser realizado en otros entornos o dominios.

La fuente de datos utilizada siempre ha sido las interacciones en los foros. Nos hemos centrado en los foros por ser este servicio de uso generalizado en entornos de colaborativos vía web donde se puede realizar comunicación asíncrona, y porque el análisis de los foros mediante técnicas de minería ha sido demostrado como muy útil (Dringus y Ellis, 2004). Además, el análisis, al utilizar las interacciones como fuentes de datos, puede ser iniciado mientras las interacciones están siendo realizadas.

Como fuente de datos de la investigación hemos utilizado indicadores estadísticos de las interacciones en los foros. Estas estadísticas podían realizarse en cualquier momento, aunque en la investigación las interacciones han sido calculadas por semana y así tener un control de las interacciones de los alumnos de semana en semana. Los indicadores estadísticos están relacionados con el grado de iniciativa y su constancia, el grado de actividad y su constancia, y la actividad generada en otros. Hemos considerado estas características apropiadas para el aprendizaje educativo (Santos et al., 2003).

Con esta fuente de datos hemos utilizado dos acercamientos de prospección:

- crear una métrica relacionada con la colaboración y que de resultados según se interaccione en los foros, y
- clasificar a los alumnos mediante algoritmo de clustering y relacionar dicha clasificación con la colaboración.

Tanto para la creación de la métrica como para relacionar la clasificación por clustering con la colaboración, se pidió a un experto que valorara la colaboración de los alumnos desde su punto de vista de experto. Esta fuente de datos no es cuantitativa pero ha sido utilizada como comprobación de la validez de los acercamientos. Una vez que los acercamientos fuesen validados, podrá no utilizarse el trabajo del experto en experiencias futuras.

Tanto por la fuente de datos usadas como por las técnicas de aprendizaje automático utilizadas, el segundo objetivo ha sido alcanzado. Una vez automatizado los acercamientos, los resultados de los análisis se alcanzarán mientras las interacciones se están llevando a cabo. Matiza que el tercer objetivo es alcanzable ya que estos acercamientos pueden ser aplicados en cualquier otro sistema que utilice los foros como medio de comunicación. Claro que este objetivo no ha sido probado en esta investigación ya que no han sido trasladados los acercamientos a otros entornos. Nos faltaría cubrir el primer objetivo, es decir, si los resultados de los acercamientos revelan la colaboración de los alumnos participantes en la experiencia de aprendizaje colaborativo.

En las secciones siguientes se deducirá la validez de los dos acercamientos propuestos gracias a los resultados mostrados en el capítulo anterior, se analizará la calidad de los resultados y se describirá un procedimiento de automatización.

## Análisis de la validez del acercamiento de la métrica

En pocas palabras la propuesta de métricas ha consistido en entrenar a un conjunto de algoritmos de árbol de decisión para que nos dijeran los indicadores estadísticos utilizados en los árboles lógicos para clasificar a los alumnos según su colaboración. Una vez que se tenían los árboles lógicos se contó el número de algoritmos que utilizaba un determinado indicador estadístico. De esta forma se ha obtenido una idea de qué indicadores están más relacionados con la colaboración. En la tabla de indicadores más frecuentes según distintos tipos de suma aparecen los indicadores más utilizados.

Se propusieron tres métricas:

- *Métrica I* =  $(level\_msg/\max(level\_msg)) + (num\_reply\_msg/\max(num\_reply\_msg)) + (level\_thrd/\max(level\_thrd))$
- *Métrica II* =  $(level\_msg/\max(level\_msg)) + (num\_reply\_msg/\max(num\_reply\_msg)) + (level\_thrd/\max(level\_thrd)) + (med\_reply\_msg/\max(med\_reply\_msg))$
- *Métrica III* =  $(level\_msg/\max(level\_msg)) + (num\_reply\_msg/\max(num\_reply\_msg)) + (level\_thrd/\max(level\_thrd)) + (num\_thrd/\max(num\_thrd))$

Los valores de esas métricas se calcularon para todos los datasets de los tres cursos. Una vez calculadas, se agruparon las instancias según el nivel de colaboración suministrado por el experto y se calculó el promedio y la varianza para los valores de las métricas en cada grupo. De los resultados obtenidos se observa que los valores de la métrica crecen según crece la colaboración. Es decir, en el espacio de los niveles de colaboración, donde un valor de 1 indica el máximo colaboración y un valor de 8 indica el mínimo, las métricas son monótonamente decrecientes, que es lo deseado, y pueden ser representadas por una recta.

El que las métricas sean monótonamente decrecientes, al igual que los niveles de colaboración, certifica la relación de las métricas con la colaboración, es decir, valida el acercamiento planteado como medio de medir la colaboración de los alumnos. Ahora nos encontramos en el punto de distinguir entre las métricas para saber cuál es la mejor.

Ya se apunto al final del Capítulo IV en la sección Análisis crítico del apartado sobre la prospección utilizando métricas que los valores de las métricas para cada dataset pueden ser representados por una recta aplicando mínimos cuadrados. La pendiente de la recta resultante indica la diferencia numérica entre dos niveles y a mayor diferencia, mayor claridad en la inferencia de la colaboración. Lo que indica que la pendiente de la recta es un indicador de la capacidad predictiva de la métrica.

El error inherente nos dirá la aproximación de la inferencia. El error (**E%**) como el promedio del porcentaje de la varianza respecto al valor de la métrica en cada caso. Matemáticamente se puede representar como  $E\% = (\sum^n_i (\text{varianza}_i/\text{media}_i)/n) \cdot 100$ ; donde  $n$  es el número de niveles dados por el experto,  $i$  son los niveles y toma valores desde 2 hasta  $n$ . Al haber menor error, en este caso menor varianza, se puede decir que la inferencia de la métrica es menos aproximada.

A mayor error menos acertada la inferencia. Por lo tanto, la pendiente de la curva y el error inherente son indicadores de la calidad de la métrica en relación a describir la colaboración de los alumnos.

Si las métricas en relación a los niveles de colaboración pueden ser representadas por una recta del tipo:  $y = \alpha x + \beta$ , donde  $x$  son las entradas e  $y$  las salidas, entonces la pendiente de la curva es la derivada de la función anterior, es decir,  $\alpha$ . El error viene dado por la varianza. Se calculará el error con el valor medio de la varianza en porcentaje respecto a los valores de la métrica. En la Tabla 28 se muestran las pendientes de las curvas para cada métrica en cada dataset junto al promedio del la varianza en porcentaje según los valores de la métrica.

**Tabla 28: pendientes y errores de las rectas que representan las métricas frente al espacio de niveles de colaboración.**

	Métrica I			Métrica II			Métrica II		
	Pendiente	R <sup>2</sup>	E (%)	Pendiente	R <sup>2</sup>	E (%)	Pendiente	R <sup>2</sup>	E (%)
D-I-06-07	0,225	0,97	<b>9,90</b>	0,2207	0,97	11,28	<b>0,2568</b>	0,96	15,57
D-II-06-07	<b>0,1943</b>	0,94	<b>9,34</b>	0,1726	0,99	10,35	0,1709	0,90	14,19
D-I-07-08	0,2014	0,95	13,20	<b>0,2186</b>	0,93	<b>12,07</b>	0,2163	0,94	22,14
D-II-07-08	<i>0,1096</i>	<i>0,61</i>	<i>11,01</i>	<i>0,0989</i>	<i>0,42</i>	<i>10,97</i>	<i>0,1189</i>	<i>0,65</i>	<i>18,48</i>
D-I-08-09	0,232	0,96	8,12	0,24	0,95	<b>7,58</b>	<b>0,32</b>	0,94	9,78
D-II-08-09	0,19	0,87	6,29	0,193	0,83	<b>5,85</b>	<b>0,271</b>	0,86	8,13

En la Tabla 28 se muestra la pendiente de la recta que es el valor absoluto de la derivada de la recta calculada con el método de mínimos cuadrados. R<sup>2</sup> es el coeficiente de correlación de la recta. Cuando R<sup>2</sup> = 1, la correlación es perfecta, al



contrario que cuando  $R^2 = 0$ . La columna E (%) es el promedio del porcentaje de la varianza respecto al valor de la métrica en cada caso. En cada fila se han señalado en negrita los valores más altos de la pendiente y los más pequeños del porcentaje de error. La fila señalada en cursiva indica un valor mínimo de la correlación  $R^2$ .

Se observa de Tabla 28 que el ajuste por mínimos cuadrados a una recta es acertado en todos los casos menos en uno. La excepción sucede con el dataset D-II-07-08 donde al mostrar un  $R^2$  demasiado bajo indica que el ajuste de los puntos a una recta no es apropiado. Por lo tanto, los datos de la fila del dataset D-II-07-08 no se tomarán en consideración.

La mejor métrica será aquella que tenga un valor más alto en la casilla de la Pendiente y menor en E (%). Lo que la Tabla 28 nos dice es que el error siempre es superior en la métrica III, aunque a veces es donde la pendiente es superior. La métrica III es la que valores más altos en la pendiente tiene en promedio, mientras que la métrica II es la menor porcentaje tiene también en promedio.

Las conclusiones que obtenemos son:

- Todas las métricas están relacionadas con la colaboración y son capaces de medir la misma siguiendo el criterio de a mayores valores, más colaboración. El acercamiento de calcular una métrica de la colaboración mediante los indicadores estadísticos de las interacciones a los foros es un procedimiento válido para medir la colaboración de los alumnos.
- No se puede deducir con toda seguridad cuál de las tres métricas propuestas funciona mejor en todos los casos.
- Si se busca que el resultado de aplicar una métrica sea el menor posible, lo aconsejable es no utilizar la métrica III.
- Si se busca que la diferencia de resultados sea la mayor posible, es decir, que la recta resultante tenga la mayor pendiente, lo aconsejable es utilizar la métrica III.
- Las principales características de los alumnos que indican su nivel de colaboración son: grado de constancia en la actividad (`level_msg`), grado de actividad provocada en otros (`num_reply_msg`) y el grado de constancia de la iniciativa (`level_thr`).

## **Análisis de la validez del acercamiento de clustering**

Este acercamiento al análisis de la colaboración ha sido distinto al explicado anteriormente. Mientras que algoritmos de aprendizaje supervisado han sido utilizados en el acercamiento de encontrar métricas relacionadas con la colaboración, ahora se utiliza un algoritmo de clustering que es de aprendizaje no supervisado.

El algoritmo de clustering EM se ha utilizado para clasificar instancias que son los indicadores estadísticos de las interacciones de los alumnos en los foros. Una vez que se obtuvo la clasificación, se utilizó la lista suministrada por el experto de los niveles de colaboración de los alumnos para establecer o no la relación de los cluster con la colaboración.

Aunque durante el curso 2008/2009 se ha desarrollado este experimento obteniendo la clasificación de los alumnos según tres niveles de colaboración, las conclusiones serán extraídas de los cursos 2006/2007, 2007/2008 y 2008/2009, durante los cuales se recopiló los indicadores estadísticos.

Se han utilizado datasets de varios cursos para validar el acercamiento. Si el algoritmo de clustering EM clasifica los alumnos siguiendo la misma relación con la colaboración aunque los datasets sean de cursos distintos, se podrá decir que el acercamiento propuesto es correcto y válido.

De las tablas que aparecen en el Capítulo IV sección de Resultados de la clasificación mediante clustering, se deduce que el acercamiento es válido. El algoritmo de clustering EM clasificó a las instancias en tres grupos y se observa con claridad que cada grupo tiene unos valores de los indicadores “num\_msg” y “num\_reply\_msg” en promedio distintos. Los grupos con los valores promedio de los indicadores estadísticos anteriores más altos fueron nombrados como de alta interacción. Y así se nombraron los grupos de media interacción y baja interacción. En todos los casos, los 6 datasets con los que se ha trabajado, se muestra que los grupos de alta interacción recogen a los alumnos más colaborativos, los grupos de interacción media recogen a los alumnos de colaboración media, y los grupos de interacción baja recogen a los alumnos de colaboración baja.

De lo anterior deducimos que los resultados validan el acercamiento de utilizar un algoritmo de clustering para clasificar las interacciones de los alumnos según la colaboración de los mismos. Sin embargo, en las mismas tablas, de las que en esta

sección se han hablado, aparece el valor medio del nivel de colaboración suministrado por un experto, que es lo que nos ha ayudado a validar el acercamiento, y también aparecen las varianzas o desviaciones medias de los niveles de colaboración en cada grupo respecto a la media. Se observa con claridad que la diferencia entre el valor promedio del nivel de colaboración en los grupos es inferior a las varianzas de los respectivos grupos. Por ejemplo la diferencia entre los grupos de interacción baja y media del dataset D-I-08-09 es de 0,71 ( $4,51 - 3,80$ ) menor que la varianza del grupo con baja interacción (0,94) y que la varianza del grupo con interacción media (0,94). Lo que indica esto es que un alumno en el grupo de baja interacción, y clasificado como de poca colaboración, podría perfectamente clasificarse como de colaboración media. Lo que no sucede es la posibilidad de clasificar a un alumno como de colaboración alta encontrándose dicho alumno en el grupo de baja interacción, clasificado de poca colaboración.

Las conclusiones sobre el acercamiento basado en la clasificación de la interacción de los alumnos por medio de un algoritmo de clustering son:

- El acercamiento de clustering es válido para agrupar a los alumnos según su colaboración.
- Los grupos o cluster con mayor interacción recogen a los alumnos más colaborativos.
- Los grupos o cluster con interacción media recogen a los alumnos con colaboración media.
- Los grupos o cluster con menor interacción recogen a los alumnos menos colaborativos.
- La clasificación anterior es válida en todas las pruebas realizadas.
- Las características del alumno que marcan su nivel de colaboración según este acercamiento son: el grado de actividad (`num_msg`) y el grado de actividad provocado en otros (`num_reply_msg`).
- Debido a la gran varianza o desviación media la clasificación es aproximada, sólo teniendo certeza de:
  - a. Si un alumno ha sido clasificado como de alta colaboración, ya que esta clasificación es aproximada, sólo se está seguro de que no tiene una colaboración baja, pudiendo ser esta o alta o media.

- b. Si un alumno ha sido clasificado como de baja colaboración, ya que esta clasificación es aproximada, sólo se está seguro de que no tiene una colaboración alta, pudiendo ser esta o baja o media.

## Métricas vs clustering

En esta sección se debería responder a la pregunta: ¿cuál de los dos acercamientos es mejor? Primero habrá que dar unos patrones para poder medir la calidad. Estos patrones son los mismos que los elegidos para deducir cuál de las métricas propuestas es la mejor. Para ello ya se ha concretado el uso de la pendiente de la recta que se obtiene con el procedimiento de mínimos cuadrados y el porcentaje de varianza respecto al valor medio por grupo. Para el caso de los datos suministrados por el análisis de clustering, no es necesario calcular la recta, la diferencia entre los valores medios indica la pendiente de esa recta. La media de la varianza indicará el error del acercamiento.

En la Tabla 29 se muestran las pendientes medias según la métrica y el dataset, además del error medio, y la diferencia media entre los grupos o cluster en el mismo dataset así como el porcentaje de varianza media.

**Tabla 29: pendiente y % de varianza respecto a los distintos acercamientos.**

	Métrica I	Métrica II	Métrica III	Clustering
Pendiente media	0,19	0,19	0,22	0,24*
E (%)	9,6	9,4	14,7	28,60

\*Se observa que al aplicar el acercamiento que utiliza técnicas de clustering, los resultados se dividen en tres niveles, y que los resultados de las métricas se dividen en siete niveles. En la Tabla 29 se ha hecho un cambio de escala en el caso del clustering para poder comparar este acercamiento con los de las métricas. También se observa que las medidas se han calculado con los resultados de los tres cursos y con los datasets sin filtrar y filtrados.

La pendiente para los resultados del clustering ha sido calculada como la diferencia en los niveles de colaboración entre dos grupos o cluster próximos dividido por el número de veces que se ha calculado la diferencia, es decir, 12, ya que estamos hablando de todos los datasets. El porcentaje de varianza respecto al valor medio se

ha calculado con una sencilla regla de tres:  $(\text{varianza} / \text{media}) * 100$ . Después se calcula el valor medio de los resultados.

El criterio para encontrar el mejor acercamiento será el que tenga una pendiente mayor y un error menor (E (%)). El acercamiento de clustering tiene el mayor error, aunque su pendiente sea alta. Mejor que el acercamiento de clustering es la métrica III. Las métricas I y II tienen un comportamiento muy semejante.

Las sencillas conclusiones que se obtienen de la comparación entre acercamientos son:

- El acercamiento de clustering, aunque diferencia bien los niveles de colaboración, tiene un error demasiado grande por lo que es conveniente utilizar el acercamiento de utilizar una métrica.
- Las características del alumno grado de constancia en la actividad (level\_msg) e iniciativa (level\_thr) identifican mejor la colaboración del alumno que el grado de actividad (num\_msg).
- La característica del alumno grado de actividad causado (num\_reply\_msg) es también un indicador apropiado de la colaboración del alumno.

## **Procedimientos de automatización de los acercamientos de minería de datos**

Los acercamientos de minería de datos planteados, el de análisis de clustering y el de creación de métricas de colaboración, son semiautomáticos ya que era necesario un trabajo “manual” por parte del experto que evaluó la colaboración de los alumnos en todos los cursos, y por parte del autor de la tesis. Sin embargo, no resultará difícil hacer cada uno de los procedimientos automáticos. En esta tesis se han verificado los acercamientos planteados y, partiendo de los resultados obtenidos, el análisis del experto sobre la colaboración de los alumnos ha dejado de ser imprescindible.

En el caso del acercamiento del cálculo de métricas de la colaboración, el procedimiento consistiría en:

1. Automatización del cálculo de los indicadores estadísticos. Las funciones psql que calculan estos indicadores ya están creadas y almacenadas en la base de datos, lo que se llama minería de datos empotrada. Sólo es necesario llamarlas

con los parámetros apropiados. La llamada automática puede ser realizada como un trigger de la base de datos al enviar, por ejemplo, un mensaje a los foros, o creando un procedimiento que cada determinado tiempo haga la llamada.

2. Una vez calculados los indicadores estadísticos, y almacenados estos en una tabla de la base de datos de la plataforma, una función psql podrá calcular la métrica y guardar el valor en una tabla junto al indicador del alumno en cuestión. Las métricas planteadas son matemáticamente fáciles de calcular porque sólo requieren de operaciones tan simples como divisiones y sumas.
3. Una vez guardadas en una tabla de la base de datos la relación de alumnos con sus valores según la métrica de la colaboración, habría que desarrollar el portlet para mostrar dicha información. Ya que la plataforma utilizada en la experiencia de aprendizaje colaborativo ha sido dotLRN<sup>13</sup>, habría que seguir las recomendaciones que los desarrolladores sugieren para realizar el desarrollo de forma inmediata.

La automatización del acercamiento de agrupación utilizando un algoritmo de clustering de los alumnos según su colaboración es idéntica al procedimiento explicado antes salvo en el punto 2. En vez de crear una función que realizara las operaciones matemáticas necesarias para el cálculo, lo conveniente será utilizar el algoritmo EM de alguna distribución de software especializado en minería de datos como WEKA<sup>14</sup> o R<sup>15</sup>. Estos dos paquetes de software tienen la posibilidad de obtener los datos directamente de una base de datos. Una vez que se obtuvieran los resultados, un pequeño plugin será necesario para guardar los resultados en la base de datos, añadiendo el indicador del alumno. De esta forma nos encontraríamos en el tercer punto de la lista anterior y el acercamiento habría sido automatizado totalmente.

Con los pasos anteriores se podría crear una aplicación capaz de funcionar en un LMS como es dotLRN. Se cumpliría el objetivo de poder trasladar los acercamientos planteados en esta tesis a otros dominios. Si se desea trasladar los acercamientos a

---

<sup>13</sup> [Desarrollo de dotLRN](#).

<sup>14</sup> Sitio web de [WEKA](#).

<sup>15</sup> Sitio web del proyecto [R](#).

otras plataformas, habría que tener en cuenta las especificaciones de la plataforma en cuestión, pero en general el procedimiento sería el mismo.

## **Validez de las herramientas metacognitivas de autorregulación**

En esta sección evaluaremos las herramientas metacognitivas de autorregulación que se han ofrecido a los alumnos y volvemos a recordar que han sido cuatro: una aplicación web que mostraba información sobre el contexto de la colaboración, del proceso y de la colaboración inferida de los alumnos; un portlet que mostraba información sobre el proceso de colaboración; un segundo portlet que mostraba información sobre la colaboración inferida de los alumnos; y para terminar se ofrecieron los dos portlet a la vez.

Estas herramientas fueron ofrecidas a principios del periodo de colaboración y la información se fue actualizando cada lunes. La información relativa a la colaboración inferida de los alumnos se empezó a mostrar a partir del 15 de diciembre de 2008.

La evaluación de las herramientas se ha realizado mediante encuestas de opinión, un tipo de encuesta por cada tipo de herramienta, y analizando los resultados de la evaluación de la experiencia de aprendizaje colaborativo y del examen del primer parcial.

El objetivo de la evaluación es comprobar las hipótesis planteadas en este ámbito: que el aumento de control sobre el proceso de colaboración incita a colaborar y mejora la gestión del propio proceso. Por eso mismo hay que comparar la colaboración de los equipos con las herramientas con un grupo de control sin ellas, teniendo también en cuenta los resultados de las notas para comprobar si las herramientas han influido algo en el conocimiento.

### **Encuestas**

Las encuestas nos han dicho:

1. Sobre la aplicación web:

- 1.1. El uso que se le dio fue muy bajo y se consultó la herramienta a principio del periodo colaborativo.
- 1.2. La facilidad de uso es apropiada pero mejorable.
- 1.3. La información suministrada no mejora sensiblemente la colaboración.
- 1.4. Los indicadores estadísticos no fueron muy útiles.
- 1.5. Los niveles de colaboración inferidos no fueron muy útiles.
- 1.6. En general los alumnos suspenden la aplicación y sobre todo la información que muestra.
2. Sobre el portlet de indicadores estadísticos:
  - 2.1. Valoración: 3,3 (5 máximo).
  - 2.2. La información suministrada no mejora ni empeora el conocimiento sobre la colaboración personal, aunque mejora la información sobre resto de compañeros, por lo que la colaboración mejorar pero sólo un poco.
  - 2.3. Los indicadores estadísticos fueron útiles pero sólo un poco.
  - 2.4. Aunque la valoración dada por la puntuación de los alumnos es de un bien alto, según el resto de respuestas, el portlet obtendría un aprobado raspado.
3. Sobre el portlet de los niveles de colaboración:
  - 3.1. Valoración: 3 (5 máximo).
  - 3.2. Al contrario de las dos herramientas anteriores, la información que se suministra es positivamente valorada, aunque mejorable.
  - 3.3. Los niveles de colaboración inferidos también son positivamente valorados.
  - 3.4. Tanto la valoración dada por la puntuación de los alumnos como las respuestas indican que el portlet obtendría un bien.
4. Sobre los dos portlet a la vez:
  - 4.1. Valoración: 3,08 (5 máximo).
  - 4.2. Al igual que pasó con las opiniones sobre la información suministrada en el portlet de niveles de colaboración inferidos, cuando se muestran los dos portlet a la vez, la valoración de la información y la ayuda de esta a la colaboración es positiva, aunque mejorable.
  - 4.3. La opinión sobre los indicadores estadísticos es muy parecida en este caso que al de sólo mostrar el portlet de los indicadores estadísticos. Las opiniones se dividen entre algo buenos y algo malos.
  - 4.4. Los niveles de colaboración inferidos son considerados acertados pero de utilidad no muy clara.



4.5. Las respuestas a esta encuesta es una mezcla de las respuestas de los alumnos que sólo se les ofreció un portlet, o el de indicadores estadísticos o el de niveles de colaboración inferidos. De la puntuación de los alumnos como las respuestas indican que esta herramienta, los dos portlet a la vez, obtendría un bien.

Aunque la información mostrada en la aplicación web es más completa que la mostrada en los portlet, la aplicación web es la herramienta peor valorada y, con toda seguridad, la menos utilizada. Esto último se puede deber a que eran necesarios pinchar sobre 4 enlaces desde la página inicial del equipo para llegar a la página que mostraba información sobre el propio usuario y que permitía navegar por la información propia y la de los compañeros. El número de enlaces que había que pinchar para llegar a ver la información del portlet desde la página de inicio del equipo era de uno solo. Esto puede ser una causa, pero la escasa valoración de la información que se mostraba indica que el problema no sólo se debe a la usabilidad.

Como era lógico y normal, la herramienta que mostraba los dos portlet a la vez ha sido mejor valorada que los portlet de forma individual. De entre los dos, el portlet que mostraba la colaboración o niveles de colaboración inferida mediante el acercamiento de clustering ha sido mejor valorado que el portlet que mostraba información sobre el proceso de colaboración, es decir, los indicadores estadísticos.

## **Notas**

Las evaluaciones al trabajo del periodo colaborativo y del examen del primer parcial de la asignatura de IA-IC también se han tenido en cuenta. La hipótesis es sencilla, si los equipos con las herramientas metacognitivas de autorregulación han colaborado mejor que otros, esto se reflejará en las notas de alguna manera. Hay más variables en juego que la de promover la colaboración y su mejora por parte de una herramienta. En el aprendizaje influyen muchos factores como la inteligencia, la motivación, el tiempo de estudio, etc. Todos estos factores son de muy difícil control en un experimento que dura, además, varias semanas. Esto se comenta para advertir de que es imposible estar seguro al 100% de que la mejora o empeoramiento del aprendizaje, medido por las evaluaciones comentadas antes, se debe sólo y exclusivamente a la herramienta. Sin embargo, podrá apreciarse su influencia.

En la Tabla 30 se muestran el promedio de las notas de la parte colaborativa y del examen del primer parcial de un grupo de alumnos a los que no se les ofreció ninguna herramienta, a los que se les ofreció la aplicación web, el portlet de indicadores estadísticos, el portlet de niveles de colaboración inferidos, y los dos portlets juntos. Además se muestra el porcentaje de abandono de cada conjunto de alumnos, entendido el abandono como aquellos que no se presentaron al examen del primer parcial de IA-IC.

**Tabla 30: comparativa de resultados de las evaluaciones según la herramienta metacognitiva de autorregulación.**

	Población final	% de no abandono	Evaluación trabajo colaborativo		Evaluación examen	
			Media [0, 2]	Varianza	Media [0, 10]	Varianza
Todos los alumnos de la experiencia	112	82,61	1,62	0,07	6,38	2,29
Alumnos de control	30	89,66	1,60	0,05	6,69	1,77
Alumnos con la aplicación web	27	70,37	1,53	0,09	5,72	1,66
Alumnos con el Portlet I	18	83,33	1,56	0,10	6,24	2,76
Alumnos con el Portlet II	20	87,5	1,72	0,02	6,31	2,71
Alumnos con el Portlet III	18	88,89	1,74	0,03	6,84	2,10

La Tabla 30 muestra el número de alumnos que terminaron la experiencia de aprendizaje colaborativo (columna 2). Tres alumnos abandonaron la experiencia mientras ésta se realizaba. También se observa la nota media y la varianza de las evaluaciones del trabajo colaborativo y del examen de IA-IC. Por ejemplo, el grupo de control, formado por 30 alumnos, obtuvo de promedio una evaluación del trabajo colaborativo peor que la media de todos los alumnos participantes en la experiencia (columna 3, filas 2 y 3), pero la evaluación del examen fue mejor a la media (columna 5, filas 2 y 3).

Se observa con claridad que los estudiantes, a los que se les ofreció la aplicación web, fueron los que peor colaboraron (media del trabajo colaborativo 1,53) en comparación a todos los demás estudiantes. La media en el examen también fue baja (5,72), pero, debido a la alta varianza en la medida, no se puede asegurar que el

aprendizaje también fuera el peor. Debido al poco uso de la aplicación web sólo se puede afirmar que la aplicación no ayudo a estos estudiantes.

Los estudiantes, a los que se les ofreció el Portlet I, obtuvieron resultados un poco mejores a los que se les ofreció la aplicación web. Por la igualdad de resultados no se puede sacar ninguna conclusión.

Los estudiantes, a los que se les ofreció el Portlet II, donde sólo se mostraban los niveles de colaboración, colaboraron muy bien (promedio del trabajo colaborativo de 1,72). Por tanto, el Portlet II cumplió con los objetivos.

Los estudiantes a los que se les ofreció el Portlet III, fueron los mejores en el trabajo colaborativo (promedio del trabajo colaborativo de 1,74) y obtuvieron buenas calificaciones en el examen (promedio del examen 6,84). Debido a la alta varianza en el promedio (2,10) no se puede asegurar que estos estudiantes fueran los que mejor aprendieron.

De los buenos resultados de los alumnos a los que se les ofreció el Portlet II o Portlet III se puede afirmar que la información inferida sobre la colaboración de los estudiantes, es decir, los niveles de colaboración, cumple los objetivos que se impuso esta investigación y ayudan a los alumnos a mejorar la gestión del proceso de colaboración. Por lo tanto, las hipótesis planteadas, que el análisis de la colaboración es posible y que el aumento de control sobre el proceso de colaboración incita a colaborar y mejora la gestión del propio proceso, han sido validadas.

## **Validez de la experiencia de aprendizaje colaborativo**

En este apartado se discute la validez de la experiencia de aprendizaje colaborativo en conjunto, sin tener en cuenta ni las inferencias realizadas ni las herramientas mecognitivas ofrecidas. En este punto se compararán los resultados entre los alumnos que realizaron la experiencia de aprendizaje colaborativo y aquellos otros que no la hicieron por el motivo que fuese. Debido a esto, se utilizará para realizar los juicios las notas obtenidas en el examen del primer parcial de la asignatura de IA-IC durante

los cursos de 2006/2007, 2007/2008 y 2008/2009. Además, se tendrá en cuenta la opinión de los alumnos en las encuestas finales.

Las encuestas finales podían ser respondidas por los dos tipos de alumnos, tanto los que realizaron la experiencia de aprendizaje colaborativa como los que no, pero la gran mayoría de los alumnos que respondieron fueron los que realizaron la experiencia colaborativa. Por ejemplo, en la encuesta final del curso 2006/2007 hubo 6 alumnos, de los 49 que respondieron, que no hicieron la experiencia de aprendizaje colaborativo. En la encuesta final del curso 2007/2008 todas las respuestas fueron de alumnos que habían realizado la experiencia de aprendizaje colaborativo. En la encuesta final del curso 2008/2009 hubo 5 alumnos, de los 72 que respondieron, que no habían realizado la experiencia de aprendizaje colaborativo. Por el escaso e irregular número de respuestas de los alumnos que no realizaron la experiencia de aprendizaje colaborativo, se tendrá en cuenta sólo la opinión de los alumnos que sí la realizaron. Estas opiniones aparecen en la Tabla 25.

La Tabla 26 muestra una comparativa entre las notas de los alumnos que realizaron la experiencia colaborativa y aquellos que no la realizaron. Ésta será la principal fuente para la realización de los juicios sobre la validez de la experiencia de aprendizaje colaborativo. Según esta tabla, hay una diferencia en las notas del examen del primer parcial entre los alumnos que no la realizaron y los que sí, que se resume en lista:

- La diferencia entre las notas del examen del primer parcial entre los alumnos que no realizaron la experiencia de aprendizaje colaborativa y los que sí en el curso 2006/2007, fue: 2,15.
- La diferencia entre las notas del examen del primer parcial entre los alumnos que no realizaron la experiencia de aprendizaje colaborativa y los que sí en el curso 2007/2008, fue: 1,75.
- La diferencia entre las notas del examen del primer parcial entre los alumnos que no realizaron la experiencia de aprendizaje colaborativa y los que sí en el curso 2008/2009, fue: 1,79.

Hay que tener en cuenta que a los alumnos que realizaron la experiencia de aprendizaje colaborativo se les dio la oportunidad de cambiar la nota de la experiencia por la de las cuestiones, que también tenían un valor máximo de 2 puntos. En la Tabla 26 se muestran estos resultados y se observa que la media de las notas del trabajo

colaborativo siempre es inferior a la diferencia mostrada, por curso, en la lista anterior.

La Tabla 25 muestra la valoración de los alumnos de la experiencia, que es positiva en los tres cursos. La valoración no es exageradamente alta pero siempre por encima de un aprobado.

De los datos que hemos tratado se deduce que la experiencia de aprendizaje colaborativo ha resultado positiva tanto en la forma de trabajo, deducido por la opinión de los alumnos, como por la mejora en el aprendizaje, lo que se observa por la diferencia positiva de las notas de los alumnos que sí hicieron la experiencia de aprendizaje colaborativo. La experiencia de aprendizaje colaborativo ha sido validada como una herramienta útil y positiva de aprendizaje, tanto colaborativo como individual.

# PARTE IV: CONCLUSIONES Y TRABAJO FUTURO

Esta parte concluye esta memoria y se expondrán las conclusiones generales que los resultados de la investigación sugieren, además de proponer trabajos futuros que completen el trabajo realizado hasta la fecha.

## Capítulo VI: conclusiones

En el Capítulo V se evaluaron los componentes de esta tesis. A continuación se evaluará la investigación en general discutiendo el grado de cumplimiento de los objetivos y los resultados con respecto a las hipótesis iniciales de este trabajo, además se responderán a las preguntas planteadas durante esta tesis.

La motivación principal de la investigación ha sido la mejora del aprendizaje de los alumnos a través de la estrategia de trabajo en equipo. Esta tesis se ha centrado en las acciones que hay que realizar para mejorar el aprendizaje colaborativo, mediante la comprobación de que dicho aprendizaje se lleva a cabo y se han propuesto las siguientes hipótesis:

1. que es posible caracterizar la colaboración en condiciones reales,
2. que el aumento de control sobre el proceso de colaboración incita a colaborar y mejora la gestión del propio proceso,
3. que el aumento del control para ser efectivo debe realizarse mientras el proceso de colaboración está teniendo lugar o poco después, y
4. que un análisis de la colaboración sin conocimiento del dominio es posible y permite el que se repitido en otros entornos educativos.

Para comprobar las hipótesis se ha propuesto analizar la colaboración de los alumnos e informar de las inferencias a los propios alumnos para darles mayor control sobre el proceso de colaboración. Los objetivos del análisis han sido:

1. caracterizar a los estudiantes según su colaboración,
2. utilizar un método de análisis que obtenga resultados de manera regular y frecuente,
3. suministrar la caracterización a los estudiantes y evaluar su aprendizaje, y
4. ofrecer una solución general para que los trabajos de esta tesis sean aplicables en otros entornos de aprendizaje colaborativo.

Estos objetivos han guiado el desarrollo de dos acercamientos de análisis y el desarrollo de cuatro herramientas de autorregulación que muestran, entre otras cosas, las inferencias de los análisis de la colaboración.

## Logros

El análisis de la literatura en relación a la mejora del aprendizaje en entornos colaborativos o el aumento del aprendizaje colaborativo nos ha llevado a crear un marco del problema. En entornos de ED se ha comentado la utilidad de aplicar la estrategia de aprendizaje colaborativo (Barkley et al., 2004; García Aretio, 2004), pero debido a la diversidad de estudiantes y a los propios condicionantes de la ED, se sugiere desarrollar un modelo educativo centrado en el estudiante, por lo que se le da el control sobre sus procesos de aprendizaje y colaboración (Boticario y Gaudioso, 2000b; European Comision, 2008). Aunque el aprendizaje colaborativo tenga grandes ventajas en entornos de ED, la oferta de la estrategia y medios de colaboración o comunicación no significa que se alcance el aprendizaje colaborativo (Johnson y Johnson, 1989). Se hace necesario que el modelo educativo del entorno colaborativo cumpla una serie de condiciones, las cuales han sido enumeradas en el Capítulo I (Johnson y Johnson, 2004).

En los nuevos entornos de ED donde es aconsejado que el control del proceso lo tenga el propio estudiante, una de las condiciones cobra gran importancia. La quinta condición que sugiere un procesamiento regular y frecuente del trabajo en equipo o colaborativo de un grupo de estudiantes. Teniendo en cuenta los estudiantes de la ED, su diversidad, cantidad, y control sobre los procesos de aprendizaje, y colaboración, los procesamientos o análisis sobre la colaboración tienen que acudir a metodologías donde la intervención de estudiantes, tutores o expertos sea mínima. La intervención de estos grupos de personas implica un aumento del trabajo fuera de su propia actividad (aprender, instruir, enseñar), lo que repercutiría en tanto el aprendizaje como en el análisis.

Desde aquí se inicia la tesis. En los entornos de ED vía web, plataformas educativas como dotLRN, tienen la gran ventaja del almacenamiento de información sobre las actividades de los alumnos lo cual puede ser utilizado para realizar seguimiento del alumno. En estos entornos técnicas de minería de datos son perfectamente aplicable (Romero y Ventura, 2007).

En el Capítulo II se explica el marco de trabajo del análisis de la colaboración y las acciones posteriores para alcanzar el objetivo de mejorar el aprendizaje colaborativo. Los logros de esta tesis han sido el desarrollo de un modelo educativo que cumple los



objetivos basado en un análisis de la colaboración en un entorno colaborativo y las consiguientes acciones para aumentar el control de los estudiantes en sus procesos de colaboración.

Los logros de la investigación se dividen: en el proceso de análisis de la colaboración, el modelo propuesto de colaboración y las herramientas que se han propuesto para la mejora del proceso de colaboración.

1. El análisis de la colaboración se divide en la adquisición de los datos y el método de inferencia. Para alcanzar los objetivos definidos para esta parte y que se reducen a que se obtenga información sobre la colaboración, mientras el proceso de colaboración ocurre y sin información semántica, se ha argumentado que: la intervención de expertos, tutores o estudiantes tiene que ser mínimas; hay que utilizar una fuente de datos en relación a la colaboración; el método de adquisición de datos a de ser cuantitativo para obtener los datos de forma regular y frecuente; evaluaciones sobre la colaboración son necesarias para asegurarse que el aprendizaje colaborativo ocurre; técnicas de aprendizaje automático pueden ser utilizadas para inferir evaluaciones sobre la colaboración. Un método de análisis que siga dichos argumentos puede ser automatizado y tiene grandes posibilidades de ser transferido a otros entornos. En la tesis se han descrito dos acercamientos los cuales han sido validados en el Capítulo V. Por lo tanto, un logro de esta tesis ha sido: encontrar en vez de uno, dos métodos de análisis de la colaboración con los cuales se obtienen evaluaciones sobre la colaboración de forma regular y frecuente, capaces de ser automatizados y transferidos a otros entornos.
2. En el Capítulo II se han discutido varias investigaciones las cuales, en algún sentido, modelaban la colaboración. Lo que se dejó claro es la falta de acuerdo en este sentido y la falta de estudios comparativos (Bratitsis y Dimitracopoulou, 2006). En primer lugar hay que tener en cuenta los objetivos del modelo y el uso que se hará de él. Hay sistemas cuyo objetivo es ofrecer recomendaciones para mejorar la colaboración (Baghaei y Mitrovic, 2007) y que el modelo sea usado por un motor de inferencia o por humanos (Daradoumis et al, 2006). Nosotros hemos seguido una estrategia que cuyo objetivo es ampliar el control del usuario sobre sus procesos, y el uso es dado directamente al usuario. Esta estrategia es denominada modelo abierto (Bull y Kay, 2008). Nosotros hemos seguido esta

perspectiva porque estructura la información en un modelo, el cual puede ser entendido como una ontología (Mizoguchi, 2005), el modelo lo usa el usuario y la información, por tanto, tiene que ser usada y entendida por el usuario. Estas ideas aportan características pedagógicas, ya que su objetivo es la mejora del aprendizaje, y de reutilidad en otros momentos y contextos. Ambas características son apropiadas a los objetivos de la investigación. Nosotros hemos propuesto un modelo de colaboración dividido en tres partes para acoger información sobre el contexto, sobre el proceso y sobre evaluaciones de la colaboración. El logro ha sido, crear un modelo de la colaboración que informe de los aspectos relevantes de la colaboración de forma entendible por estudiantes y transferible con facilidad a otros entornos.

3. Para cerrar el ciclo de un investigación en entornos educativos, hemos estudiado las posibles acciones correctivas que hay que realizar para asegurarse que el aprendizaje colaborativo se realiza y, además, proporcionando mayor control sobre el proceso de la colaboración. En el Capítulo II se enmarca este problema pero no se plantea una solución al problema. Estudiando el problema nos hemos dado cuenta de las ventajas que una herramienta de autorregulación aporta a un entorno educativo y colaborativo (Steffens, 2001). Esta idea está muy relacionada con el modelo abierto (Bull y Kay, 2008) y con las herramientas metacognitivas (Soller et al., 2005). Hemos propuesto a los estudiantes cuatro herramientas de autorregulación con el objetivo de comprobar el tipo de información más útil y de la forma o tipo de herramienta. Por estas razones se ha preparado una herramienta escrutable (Kay, 1999) que mostraba todo el modelo de colaboración, y otras tres herramientas de autorregulación, una de ellas de monitorización ya que informaba del proceso de colaboración, y otras dos que han sido metacognitivas ya que informaban de la evaluación de la colaboración. Un proceso de evaluación ha sido realizado y los resultados son discutidos en el Capítulo V. Los logros de esta parte han sido: comprobar las ventajas que aporta en la colaboración de una herramienta metacognitiva como de autorregulación en entornos de aprendizaje colaborativo.

Esta tesis tienen como logros generales que ha propuesto una metodología con la motivación de mejorar el aprendizaje colaborativo aumentando el control sobre el proceso de colaboración. Para ello ha analizado y modelado la colaboración, ha propuesto un modelo el cual ha sido utilizado como base para distintas herramientas de autorregulación que han sido evaluadas para comprobar si se alcanzan los

objetivos. Se ha demostrado que existe una relación entre el trabajo realizado (análisis, modelo) con el aumento del aprendizaje colaborativo (ver Capítulo V). Además, se ha intentado que el tutor, estudiante o experto tuvieran la menor intervención posible para alcanzar los objetivos de la investigación. El método seguido se ha intentado hacer lo suficientemente general para poder ser aplicado a otros entornos. Hay que decir que se ha conseguido el logro de la mejora del aprendizaje colaborativo, que estudiantes, tutores o expertos tengan la menor intervención posible, y se ha propuesto una automatización que hace posible la transferencia del método a otros entornos, algo que todavía no ha sido probado.

## **Valoración de contribuciones**

Los logros de esta tesis por los resultados antes mostrados se puede resumir en:

1. un método de análisis de la colaboración,
2. una herramienta que mejora el aprendizaje colaborativo, y
3. un modelo educativo que motiva a los estudiantes a no abandonar la asignatura.

Ambos métodos de análisis de la colaboración cubren los objetivos de la tesis para el análisis. La calidad de las inferencias es buena, aunque habría que establecer un método objetivo de cálculo de dicha calidad. Nunca fue un objetivo que las inferencias sobre la colaboración tuvieran que ser totalmente certeras. Nos estamos moviendo por un dominio donde la característica principal, la colaboración, no tiene ni una teoría, ni una metodología, ni unos estándares detrás para poder realizar un análisis de exactitud de una inferencia sobre la colaboración. Sin embargo, el encontrar un procedimiento cuantitativo automatizable que infiera información sobre la colaboración de los alumnos participantes en una experiencia de aprendizaje colaborativo, es un buen avance.

La mejora del aprendizaje colaborativo mostrado por los alumnos a los que se les ofreció las herramientas metacognitiva de autorregulación consistentes en mostrar los niveles de colaboración por un lado, y los niveles de colaboración más las estadísticas de las interacciones a los foros, por otro, sugiere que suministrar mayor control sobre el proceso de colaboración mejora la colaboración. Se ha intentado cuantificar la mejor del aprendizaje colaborativo al comparar los resultados de un conjunto de estudiantes, los que tenían dichas herramientas, con el otros grupos. Se observa esta

mejora la ver los datos de la Tabla 30. Esto indica que el método seguido cumple la motivación de la investigación. El avance no lo da estos resultados, ya que muchas otras herramientas en entornos parecidos podrían aportar los mismos resultados. El principal avance es la evaluación de todo el modelo educativo y método planteado como válidos ya que han alcanzado los objetivos y la mejora del aprendizaje.

No queremos dejar de notar un pequeño análisis comentado en el Capítulo V y que compara el número de abandonos de los estudiantes de la asignatura IA-IC teniendo en cuenta si han participado en la experiencia de aprendizaje colaborativo o no, y el tipo de herramienta ofrecida (ver Tabla 30). No era un objetivo de la investigación, pero se observa un porcentaje de estudiantes a los que se les ofreció la información metacognitiva (Portlet II y III) que se presentaron a examen (87,5 y 88,89 respectivamente). No tenemos más datos ahora para deducir todas causas y la exacta relación pero es un avance para la investigación que aquellos estudiantes que han validado el método, además se hayan presentado más a examen, lo que puede indicar una mayor motivación.

## **Análisis de las limitaciones**

Las principales limitaciones apreciadas en la tesis son:

- incertidumbre en la inferencia del análisis de la colaboración,
- métodos de aprendizaje automático utilizados,
- dificultad en el uso de la herramienta escrutable, y
- evaluación de los resultados sobre el aprendizaje.

Ya se ha dicho que nunca fue un objetivo de esta tesis el obtener inferencias exactas sobre la colaboración. Debido a la ambigüedad teórica, metodológica y de estándares, hay grandes dificultades en establecer que es la mejor colaboración. Sin embargo, los porcentajes de error calculados para ambos acercamientos de análisis de la colaboración son elevados. Los alumnos informaron de sus dudas sobre la exactitud de los valores de las inferencias. Según las encuestas de opinión los alumnos piensan que los resultados de las inferencias son acerados pero aproximados, lo que corresponde con el objetivo de esta tesis.

En esta investigación se ha analizado la colaboración utilizando dos acercamientos. Se ha probado que estos acercamientos analizan la colaboración aunque sea de forma

aproximada. Los resultados que ofrecen son aproximados pero tienen el valor de rapidez y flexibilidad para ser utilizados en otros entornos. Ambos acercamientos fueron elegidos por sus características teóricas o por experiencias realizadas con ellos en otras investigaciones. Los dos enfoques han sido probados de forma empírica y se han comparado entre sí. En esta investigación se ha establecido un criterio de comparación entre los dos acercamientos observando que el enfoque que calcula métricas de colaboración muestra menor error en las inferencias y una capacidad predictiva semejante al acercamiento de clustering. Sin embargo, las características teóricas de ambos acercamientos no nos proporcionaban la base para dicho comportamiento. Por esto mismo, no se puede establecer sin un análisis empírico como el que en esta investigación se ha utilizado cuál de los posibles enfoques que utilicen aprendizaje automático es el mejor. Esta es una de las limitaciones ya que no se puede determinar a priori que método de aprendizaje automático resolvería el problema del análisis de la colaboración según sea el caso.

El punto más abierto de esta tesis ha sido el uso de la herramienta escrutable. Requirió un gran trabajo de diseño e implementación que no se ha visto reflejado por el uso que de ella hicieron los alumnos. Al contrario que los portlets, cuyo diseño e implementación fueron triviales, han obtenido mucho mejores resultados. Las características teóricas de la escrutabilidad sugieren que los usuarios de una herramienta escrutable podría aumentar su motivación, ya que se le da más información y responsabilidad. Sin embargo, los estudiantes que pudieron utilizar la aplicación, la herramienta metacognitiva de autorregulación escrutable, no mejoraron su aprendizaje y se produjo un mayor número de abandonos en este conjunto de estudiantes (ver Tabla 30). Todo hace indicar que los estudiantes estaban menos motivados que otros, a los que se les ofreció otras herramientas metacognitivas de autorregulación. Por lo que se ve, la aplicación ha obtenido unos resultados opuestos a los esperados.

Hemos comentado que no hay metodología totalmente establecida de evaluación de sistemas adaptativos (van Velsen et al., 2008). Hemos seguido una evaluación centrada en el estudiante y, principalmente, en los resultados de sus trabajos, es decir, del trabajo colaborativo y del aprendizaje en general con el examen del primer parcial de la asignatura. Mientras que en la realización del trabajo colaborativo intervenían variables dependientes, como la colaboración o el tiempo dedicado, que el tutor ha

podido controlar en alguna medida gracias al modelo educativo de la experiencia de aprendizaje colaborativo, las variables dependientes en la realización del examen del primer parcial no podían ser controladas, como la motivación o el tiempo de estudio. Por eso mismo, a las conclusiones obtenidas de analizar los resultados del examen del primer parcial se les da un carácter de aproximación. El aprendizaje de los estudiantes a los que se les ofreció el Portlet III parece ser mejor que el del resto de estudiantes, pero la influencia de la herramienta (Portlet III) no es fácil de cuantificar por el descontrol o desconocimiento de las variables dependientes que han intervenido en la realización del examen del primer parcial (Chi, 2001).

## **Respuesta a las preguntas planteadas por esta tesis**

Las preguntas que esta tesis se ha planteado han sido:

1. ¿Es posible desvelar la colaboración de los estudiantes mediante métodos automáticos de inferencia? ¿Cómo? ¿Con qué método?
2. ¿Qué datos son necesarios para realizar la inferencia?
3. ¿Es posible, aportando más información sobre la colaboración, ayudar a los estudiantes en la gestión del proceso de la colaboración?
4. ¿La mejora del trabajo colaborativo mejora el aprendizaje?
5. ¿Es útil la estrategia escrutable en la mejora del proceso de colaboración? ¿Qué características son apropiadas para conseguir dicha mejora?

En las secciones anteriores se han dado las claves para responder a estas preguntas. Ahora las haremos explícitamente.

### **Respuesta a la primera pregunta**

Sí es posible desvelar la colaboración de los estudiantes analizando sus interacciones y hacerlo de forma automática. En esta tesis esto se ha realizado con los dos métodos propuestos: la creación de una métrica de la colaboración y la clasificación de los estudiantes de acuerdo a la colaboración mediante clustering. Una vez que ambos métodos se han verificado, la automatización se ha descrito en tres pasos. El método, de los dos propuestos, que mejores resultados a obtenido en la inferencia de la colaboración es el de la métrica ya que el error calculado es el menor. Incluso este error podría reducirse ponderando cada una de las partes que forma la

métrica con el objetivo de disminuir el error. También hay que discutir sobre otros métodos de inferencia, por ejemplo Bagging, por lo que para encontrar el método de inferencia para el problema en cuestión es necesario más estudios empíricos. Esto es uno de los trabajos futuros.

### **Respuesta a la segunda pregunta**

Los datos con los que se ha conseguido inferir la colaboración de los estudiantes han sido indicadores estadísticos de las interacciones en los foros. (Drigus y Ellis, 2004) aconsejan el análisis de las interacciones en los foros como medio de obtener juicios sobre la colaboración de los estudiantes. Los indicadores se han referido a interacciones activas ya que la colaboración es un aprendizaje activo. Como las comunicaciones entre los alumnos se realizaron principalmente por los foros, el análisis de los foros es necesario para inferir conclusiones sobre la colaboración de los estudiantes.

Los indicadores estadísticos han sido una respuesta lógica y simple de responder a la pregunta de qué atributos utilizar para inferir la colaboración. A partir de aquí se podrían añadir otras variables relacionadas también con la colaboración o con la comunicación. No obstante, el servicio de los foros puede suministrar más información sobre la colaboración. Se puede tener en cuenta las interacciones pasivas (Burr y Spennemann, 2004), es decir, cuantos mensajes ve o lee un estudiante, también información del dominio al realizar un análisis semántico de los mismos (Ravi et al., 2007), o una codificación de los mensajes mediante géneros (Lin et al., 2009). Estos últimos indicadores ya no serían independientes del dominio y habría que probar la posibilidad de transferir el acercamiento con estos nuevos indicadores a otros entornos y dominios.

### **Respuesta a la tercera pregunta**

Las notas obtenidas en la evaluación del trabajo colaborativo (ver **Error! Reference source not found.**<sup>25</sup>) demuestran que los estudiantes, a los que se les ofreció los Portlets II y III trabajaron colaborativamente mejor que el resto de los estudiantes que participaron en la experiencia de aprendizaje colaborativo del curso 2008/2009. Ya que las dos herramientas (Portlet II y Portlet III) compartían parte de

la información, exactamente la información sobre los niveles de colaboración, se puede afirmar que esta información, inferida gracias a un análisis de la colaboración, ha ayudado a los alumnos a gestionar su proceso de colaboración. El análisis de la colaboración ha ayudado a la mejora del proceso de colaboración.

### **Respuesta a la cuarta pregunta**

En la **Error! Reference source not found.**<sup>30</sup> se observa que los estudiantes, a los que se les suministró la información sobre los niveles de colaboración y los indicadores estadísticos (Portlet III), obtuvieron buenas calificaciones en el examen del primer parcial de IA-IC. Pero también se observa la alta varianza. No se puede asegurar, por tanto, que la mejora de la colaboración acarrea un mejor aprendizaje, aunque se sugiere. Existen variables dependientes que en un examen no pueden ser controladas y para minimizar dicho sesgo se haría necesario aumentar la población del análisis, es decir, el número de estudiantes a los que se les ofrece la herramienta y después se evalúe su examen. Más resultados son necesarios.

### **Respuesta a la quinta pregunta**

De los datos obtenidos podemos afirmar con claridad que la herramienta escrutable no ha alcanzado los objetivos. La valoración de la información presentada en esta herramienta fue evaluada como de poca calidad, en contraste a la valoración sobre la misma información, aunque mostrada de otra forma, que dieron los estudiantes a los que se les ofreció las otras herramientas. Si la información es la misma pero la valoración es distinta, eso indica que la forma de presentar dicha información ha sido lo evaluado negativamente.

Los resultados sugieren que la estrategia de escrutabilidad no es apropiada, al menos, en el entorno descrito, pero también hay que tener en cuenta el uso de la herramienta. Se ha dicho que el uso ha sido escaso, es más, demasiado escaso para poder realizar juicios de valor. Este es un punto abierto que la investigación no ha resuelto y que necesita más trabajo en el futuro.



# Capítulo VII: trabajos futuros

En este apartado se comentarán los trabajos a realizar debido a los resultados de la investigación de esta tesis. Esta investigación ha dejado algunos puntos abiertos que no se han respondido en su totalidad. A continuación se dividirá el trabajo realizado comentando en cada punto de interés el trabajo futuro a realizar.

## **Análisis de la colaboración**

Esta parte la vamos a dividir en la fuente de datos, los métodos de inferencia y la propia inferencia a alcanzar.

### **Fuente de datos**

Para realizar un análisis con el que obtener inferencias utilizando aprendizaje automático es necesario ricas fuentes de datos que informe sobre todas las partes del problema que se considera (Lehmann y Hitzler, 2009). La fuente de datos en esta investigación han sido indicadores estadísticos de las interacciones en los foros. Ya se ha justificado por qué se han tenido en cuenta en las interacciones en los foros en vez de otras interacciones en otros servicios. Sin embargo, al no ser los foros el único servicio de comunicación utilizado de la plataforma, ya que en el curso 2008/2009 también se suministro un servicio de Chat, habría que analizar qué indicadores estadísticos en esos servicios son los más apropiados para desvelar la colaboración. Lo malo es que los alumnos también utilizaron servicios que la experiencia de aprendizaje colaborativo no ofrecía como el correo electrónico ordinario, Messenger, Skype, etc. Estos servicios sólo se pueden monitorizar desde la propia máquina del usuario del servicio. Para evitar que se utilizaran servicios de Chat con Messenger o Skype, se ofreció el servicio de Chat en la experiencia de aprendizaje colaborativo. Lamentablemente la calidad del Chat ofrecido no se puede comparar con las anteriores marcas comerciales, por lo que el Chat fue poco utilizado.

El camino es lograr integrar herramientas de comunicación comerciales en entornos educativos, y poder acceder a la información que estas herramientas almacenan en relación a los usuario del entorno. En la actualidad todavía no se ha conseguido.

Debido a esto, es importante investigar la comunicación en los foros para extraer más información sobre la colaboración. Las interacciones pasivas también se pueden tener en cuenta (Burr y Spennemann, 2004), aunque la colaboración al ser una estrategia de aprendizaje activo requiera principalmente interacciones activas para llevarse a cabo. En la tesis no se ha realizado análisis semántico de los mensajes con el objetivo de facilitar todo lo posible la transferencia del método de análisis a otros entornos o dominios, pero en una futura investigación habría que medir si es un grave impedimento y si mejora el análisis de la colaboración. Por ejemplo, (Ravi et al., 2007) han hecho un análisis semántico de los mensajes a los foros para etiquetar dichos mensajes como preguntas, respuestas o ambas. El etiquetado no aporta información sobre el dominio pero sí por la clase de comunicación. Por lo tanto, utilizar este etiquetado como un atributo más no perjudicaría la transferencia del análisis. También (Lin et al., 2009) codifican los mensajes. Estos los hacen por el género del que trata el mensaje, lo cual tampoco aporta conocimiento del dominio. Habría que investigar que relación hay entre los códigos o etiquetas de los mensajes con la colaboración, lo que requeriría unos métodos de inferencia distintos a los expuestos en esta tesis.

## **Método de inferencia**

En esta tesis se ha justifica el uso de dos métodos de inferencia aplicando técnicas de aprendizaje automático. Por un lado se han aplicado árboles de decisión y por otro un algoritmo de clustering.

Debido a las características de la colaboración y a la aproximación de todos los juicios que sobre ella se realizan, no hay un criterio teórico que indique el método de inferencia más apropiado. En esta tesis hemos utilizado dos y los hemos comparado cuantitativamente.

Para la identificación de los indicadores estadísticos como más relacionados a la colaboración hemos utilizado un método de inferencia sensible a pequeños cambios

en los datos, es decir, los árboles de decisión, pero con un sencillo análisis estadístico hemos utilizado dicha característica para eliminar los sesgos que los algoritmos poseen. La técnica de Bagging (Breiman, 1996) mejora el modelo aprendido combinando múltiples árboles de decisión. Se consigue mejorar la medida de la varianza, por ejemplo, y se obtiene un único modelo aprendido o árbol lógico, pero el problema es la optimización. En principio, al ser la clasificación según la colaboración no clásica (ver Rosch, 1978), un modelo que clasifica según la colaboración no tiene base teórica. Por eso mismo la técnica de Bagging no fue utilizada pero se debería realizar un estudio empírico para compararlo con los realizados de esta tesis.

En el caso de la clasificación dirigida por los datos, técnica utilizada en esta tesis utilizando un algoritmo de clustering, en los trabajos futuros se podría utilizar otros algoritmos de clustering y mapas autoorganizativos de Kohonen (Kohonen, 1989). Después se podría realizar un análisis empírico que comparara los resultados de cada algoritmo para encontrar el mejor.

En el caso de ampliar los atributos de la fuente de datos con indicadores que mostrasen la codificación de los mensajes (ver apartado anterior) sería conveniente aplicar el método de minería de patrones secuenciales (Sequential Pattern Mining en inglés) como lo han hecho (Perera et al., 07).

### **Mejora de la métrica**

En el apartado Análisis de la validez del acercamiento de la métrica se evaluaron tres métricas distintas desde dos criterios distintos: la capacidad predictiva dada por la pendiente de la curva asociada a la métrica, y el error indicado por la varianza media. Se dijo que la mejor métrica sería aquella con mayor capacidad predictiva y menor error.

Las métricas propuestas incluían valores normalizados de algunos indicadores estadísticos (level\_msg, level\_thrd, num\_reply\_msg, num\_thrd, med\_reply\_msg) y estos indicadores eran sumados sin ponderar. Ya que se han encontrado los indicadores estadísticos más relacionados con la colaboración, el siguiente paso es ponderar la suma para dar más importancia a unos indicadores frente a otros. Para conocer los términos de la ponderación se pueden utilizar el número de veces que se han utilizado de media en cada dataset, los valores ponderados que aparecían en la

**Error! Reference source not found.**8, o realizar una regresión lineal multivariable para obtener una fórmula como  $Y = \sum \beta_i X_i + \varepsilon$ ; donde “i” toma el valor del número de indicadores estadísticos utilizados en la métrica. Los parámetros  $\beta_i$  pueden ser tomados como factores para ponderar la métrica.

## **Qué hacer con la información obtenida en el análisis**

En este apartado se discutirán trabajos futuros que se pueden realizar para mejorar el uso que se haga de las inferencias obtenidas de los análisis de la colaboración. En primer lugar, hay que tener en cuenta el modelo de la colaboración, después qué información mostrar, y por último, se discutirá sobre las herramientas metacognitivas de autorregulación y principalmente de la herramienta escrutable.

### **Modelo de la colaboración**

El modelo de la colaboración que aquí se ha planteado se dividía en: información del contexto de la colaboración, información del proceso de colaboración, información sobre la colaboración de los estudiantes. Solamente los estudiantes a los que se ofreció la herramienta escrutable se les preguntó su opinión sobre todos los datos del modelo y que se mostraban en la herramienta. Debido al poco uso de la herramienta escrutable, la validación del modelo no se ha podido hacer completamente. Un trabajo futuro de especial interés es validar cuantitativamente el modelo. Debido a la preferencia de los estudiantes participantes en la experiencia de aprendizaje colaborativo por la información simple y esquemática. En el siguiente curso se les mostrará a los participantes la totalidad del modelo de esa forma.

### **Información sobre la colaboración**

En el curso 2008/2009 se ha mostrado a algunos estudiantes el nivel de colaboración inferidos gracias a la técnica de clustering. Del estudio empírico de comparación de los acercamientos de clasificación por clustering y el cálculo de métricas se ha observado que la métrica no tiene un error asociado tan alto como el caso de la clasificación por clustering. En la experiencia de aprendizaje colaborativo

del curso 2009/2010 se mostrarán ambas inferencias y se preguntará a los estudiantes por su opinión.

Para ello se crearán dos portlets, uno con los valores de la métrica y otro con los niveles de colaboración inferidos por el clustering. Los portlets se irán actualizando de forma automática cada semana y se mostrarán los valores de las semanas pasadas.

## **Herramienta escrutable**

Ya que el fracaso de la herramienta escrutable en el curso 2008/2009 se le puede achacar principalmente al poco uso, en futuros cursos se realizarán actualizaciones sobre la herramienta.

Para evitar la complejidad e incentivar el uso se podría integrar la aplicación web, que daba soporte a la herramienta escrutable, dentro del LMS, es decir, dotLRN, y utilizar el usuario del estudiante en dotLRN como identificador en la herramienta escrutable.

Además de integrar la aplicación web, al usuario que utilizará la herramienta habrá que motivarle y explicar el uso. La escrutabilidad es una estrategia muy poco común y a la que los estudiantes no están acostumbrados. Por estos motivos, la herramienta debería integrarse de tal forma que el usuario la considere como una parte de la propia plataforma educativa y que poco a poco vaya aprendiendo a utilizarla. En el ejemplo de la experiencia de aprendizaje colaborativo esto se podría conseguir pidiendo que el usuario rellene la información como si fuese una encuesta.

El problema de la integración de la herramienta escrutable en dotLRN puede ser muy complejo. La plataforma dotLRN separa una página web clásica en tres partes: los datos que se muestran, el trabajo de obtención de esos datos, y la presentación. Como los datos en dotLRN se encuentran todos en la base de datos de la plataforma, el desarrollo de una aplicación de dotLRN se divide en: el esquema de tablas o modelo de datos, las sentencias que acceden al esquema o modelo de datos, y la presentación de los datos. Este método de trabajo consigue rapidez, modularidad y robustez, características loables en un LMS. Sin embargo, se eligió para la herramienta escrutable una aplicación web basada en java por la sencilla razón que aportaba flexibilidad a la herramienta para así añadir o quitar atributos del modelo. Es decir, era deseado que el modelo pudiera ser actualizado mientras se estuviera presentado en

la herramienta escrutable. El método de trabajo de dotLRN hace muy complicado esa flexibilidad al utilizar tablas de una base de datos. Una base de datos no está pensada para que sus tablas puedan cambiar constantemente. Si se desea esto último, lo mejor es utilizar otro software en vez de una base de datos.

La integración de la herramienta escrutable en dotLRN debe de cambiar el método de trabajo y esto se puede conseguir cambiando el módulo intermedio entre base de datos y presentación de los mismos. Por lo tanto, como trabajo futuro se desarrollará un módulo de acceso a datos que en vez obtenerlos de unas tablas lo haga de una ontología escrita en OWL. Con este pequeño cambio la herramienta escrutable se integrará en dotLRN sin perder la flexibilidad deseada.

## Esquema de los trabajos futuros

Lo comentado anteriormente se puede resumir en el siguiente gráfico de trabajos futuros a realizar.

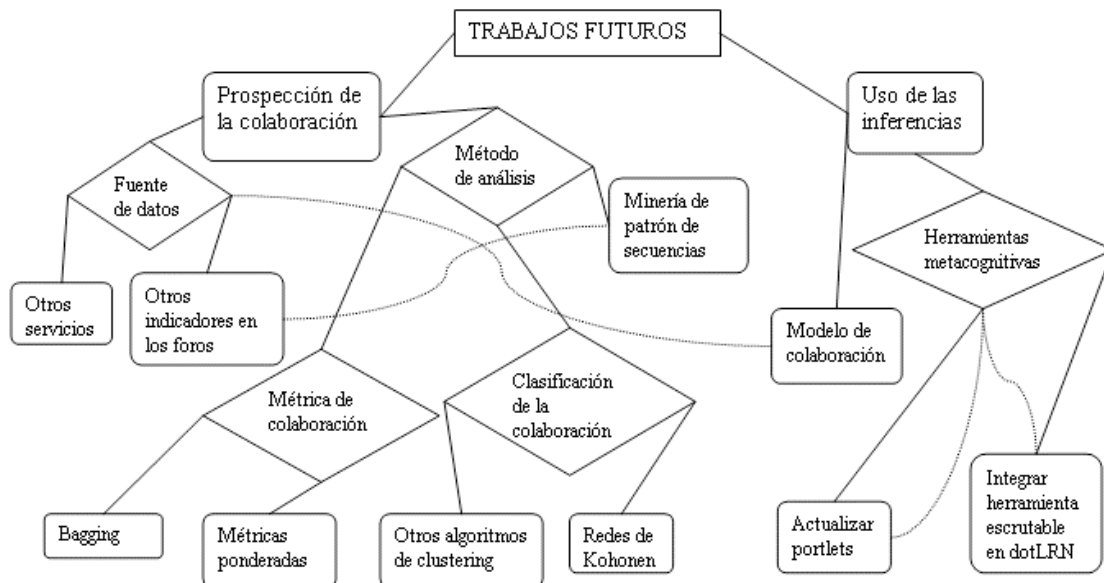


Figura 28: esquema del trabajo futuro.

# Bibliografía

- [1] Adams, D. and Hamm, M. (1996). Cooperative Learning, Critical Thinking and Collaboration Across The Curriculum. Second Edition, Charles Thomas Publisher, 1996
- [2] Adomavicius, G. and A. Tuzhilin (2005). "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions." Knowledge and Data Engineering, IEEE Transactions on Knowledge and Data Engineering. 17(6): 749
- [3] Anaya, A. R., Boticario, F. G. (2009). Clustering learners according to their collaboration. Proceedings of the 13th international conference on computer supported cooperative work in design (CSCWD 2009): IEEE Computer Society Press, 2009.
- [4] Anaya, A. R., Boticario, F. G. (2009). Reveal the Collaboration in a Open Learning Environment. Methods And Models In Artificial And Natural Computation, PT I Vol: 5601. 464-475. 2009.
- [5] Anaya, A. R., Boticario, F. G. (2009). A data mining approach to reveal representative collaboration indicators in open collaboration frameworks. Proceedings of the second international conference on educational data mining (EDM09), 2009.
- [6] Antoniou G., & Harmenlen F. "Web Ontology Language: OWL". Staab S. & Studer R. (Eds.), Handbook on Ontologies, ISBN 3-540-40834-7, Springer-Verlag Berlin Heidelberg Nes York, 2004
- [7] Aristóteles. Ética Nicomáquea · Ética Eudemia. Biblioteca clásica Gredos, 89. Ed.: Editorial Gredos. 1985.
- [8] Arregui Fernández, J. (2002). Topología. Eds: UNED. 440 páginas. ISBN: 8436216741
- [9] Assad M., Carmichael D. J., Kay J., Kummerfeld B. (2007). MyPlace Locator: Flexible Sharing of Location Information. C. Conati, K. McCoy, and G. Paliouras (Eds.): UM 2007, LNAI 4511, pp. 410–414, 2007
- [10] Baeza-Yates, R. & Pino, J.A. (2006) Towards formal evaluation of collaborative work. Information Research, 11(4) paper 271
- [11] Baghaei N., Mitrovic A. (2007). From Modelling Domain Knowledge to Metacognitive Skills: Extending a Constraint-Based Tutoring System to Support Collaboration. C. Conati, K. McCoy, and G. Paliouras (Eds.): UM 2007, LNAI 4511, pp. 217–227, 2007.
- [12] Baker R. (2007) "Is Gaming the System State-or-Trait? Educational Data Mining Through the Multi-Contextual Application of a Validated Behavioral Model". Workshop on Data Mining for User Modeling 2007.
- [13] Baldiris S., Santos, O. C., Barrera, C., Boticario, J. G., Velez, J., Fabregat, R.: "Linking educational specifications and standards for dynamic modeling in ADAPTAPlan". International

Workshop on Representation models and Techniques for Improving e-Learning: Bringing Context into Web-based Education (ReTLeL'07). Denmark (2007)

- [14] Barkley, E., Cross, K.P, and Major, C.H., Collaborative Learning Techniques: A Practical Guide to Promoting Learning in Groups, San Francisco, Cal.: Jossey Bass, 2004.
- [15] Barros, B. Mizoguchi, R. & Verdejo, M.F.: A platform for collaboration analysis in CSCL. An ontological approach. Proc. of Artificial Intelligence in Education AIED-2001, San Antonio, USA (2001)
- [16] Barros, B., Verdejo, M.F., Read, T. & Mizoguchi, R.: APPLICATIONS OF A COLLABORATIVE LEARNING ONTOLOGY. MICAI'2002 Mexican International Conference on Artificial Intelligence). LNAI, Springer-Verlag (2002)
- [17] Bates, A. W., Poole, G. 2003. Effective Teaching with Technology in Higher Education: Foundations for Success. Ed. Jossey-Bass
- [18] Ben-Naim, D., Bain, M., Marcus, N. (2009). A User-Driven and Data-Driven Approach for Supporting Teachers in Reflection and Adaptation of Adaptive Tutorials. Second International Conference on Educational Data Mining (EDM'09). Córdoba, Julio de 2009.
- [19] Berkhin P. (2006). A Survey of Clustering Data Mining Techniques. Grouping Multidimensional Data (2006), pp. 25-71.
- [20] Berry, M., Linoff, G. (2004). Data Mining Techniques: For Marketing, Sales, And Customer Relationship Management. John Wiley & Sons Inc. 2004/04/09
- [21] Boekaerts, M., Pintrich, P. & Zeidner, M. (2000) (Eds.) Handbook of self-regulation. New York: Academic Press.
- [22] Bonwell, C. & Eison, J. (1991). Active Learning: Creating Excitement in the Classroom AEHE-ERIC Higher Education Report No.1. Washington, D.C.: Jossey-Bass.
- [23] Borrajo, D., Boticario, J.G., Isasi, P. (2006). Aprendizaje Automático. Ed. Sanz y Torres, Madrid, Spain.
- [24] Boticario, J.G., Gaudioso, E. (2000a). Towards a Personalized Web-based educational system. Mexican international conference on artificial intelligence 2000, Acapulco, Mexico, Springer Verlag, pp. 729-740
- [25] Boticario, J.G., Gaudioso E. (2000b). Adaptive web-site for distance learning. Campus-Wide Information Systems: The Journal of technology on campus, Vol. 17, Num. 4, 120-128, 2000
- [26] Boticario, J.G., Santos, O.C., Barrera, C., Rodríguez, A. (2007). Usage of Educational Standards supported by dotLRN in aDeNu projects. International Conference and Workshops on Community Based Environments: OpenACS and .LRN Spring Conference, 04/2007, Vienna, Austria.
- [27] Bratitsis, T., Dimitracopoulou, A. (2006). Indicators for Measuring Quality in Asynchronous Discussion Forae. The 12th International Workshop on Groupware, CRIWG 2006, Spain: Springer Verlag, 54-61



- [28] Bratitsis, T., Dimitracopoulou, A., Martínez-Monés, A., Marcos-García, J.A., Dimitriadis, Y. (2008). Supporting members of a learning community using interaction analysis tools: the example of the Kaleidoscope NoE scientific network Proceedings of the IEEE International Conference on Advanced Learning Technologies, ICALT 2008, 809-813, Santander, Spain, July 2008.
- [29] Bratitsis, T. (2009). Visualized Interaction Analysis indicators in Asynchronous Discussion learning activities. Research results and trends. Workshop Interaction Analysis and Visualization of International Conference CSCL'09. Rhodes, Greece, June, 2009
- [30] Bravo, J., Ortigosa, O. (2009). Detecting Symptoms of Low Performance Using Production Rules. Second International Conference on Educational Data Mining (EDM'09). Córdoba, Julio de 2009.
- [31] Breiman, Leo (1996). "Bagging predictors". *Machine Learning* 24 (2): 123-140. doi:10.1007/BF00058655
- [32] Breiman, Leo (2001). "Random Forests". *Machine Learning* 45 (1), 5-32.
- [33] Bricall, J.M., 2000. Informe Universidad 2000, CRUE, Madrid.
- [34] Brusilovsky P. & Millan E. (2007). User Models for Adaptive Hypermedia and Adaptive Educational Systems. P. Brusilovsky, A. Kobsa, and W. Nejdl (Eds.): *The Adaptive Web*, LNCS 4321, pp. 3 – 53.
- [35] Bull, S. & Kay, J. (2008). Metacognition and Open Learner Models, in I. Roll & V. Alevén (eds), *Proceedings of Workshop on Metacognition and Self-Regulated Learning in Educational Technologies*, International Conference on Intelligent Tutoring Systems, 7-20.
- [36] Burleson, W. (2005). Developing creativity, motivation, and self-actualization with learning systems. *Int. J. Human-Computer Studies*, 62 (2005), 664-685.
- [37] Burr, L., Spennemann, D. HR. (2004). Patterns of User Behavior in University Online Forums. *International Journal of Instructional Technology and Distance Learning*. 2004.
- [38] Caballé S., Xhafa F., Daradoumis T., Fernández R. "Enabling Efficient Real Time User Modeling in On-Line Campus". C. Conati, K. McCoy, and G. Paliouras (Eds.): *UM 2007*, LNAI 4511, pp. 365–369, 2007.
- [39] Callejo Gallego, Manuel Javier. Estudio de cohorte de estudiantes de la UNED: una aproximación al análisis del abandono. *RIED: revista iberoamericana de educación a distancia*, ISSN 1138-2783, Vol. 4, Nº 2, 2001, pags. 33-70.
- [40] Chang, M-M. (2005). Applying self-Regulated Learning Strategies in a Web-Based Instruction-an Investigation of Motivation Perception. *Computer Assited Language Learning*, Vol. 18(3), pp. 217-230
- [41] Chin, D. (2001). Empirical Evaluation of User Models and User-Adapted Systems. *User Modeling and User-Adapted Interaction* 11: 181^194, 2001
- [42] Clark, P., Niblett, T. (1998). The CN2 Induction Algorithm. *Machine Learning*, 3(4), pp.261-283.

- [43] Cocea M., & Weibelzahl S., "Eliciting Motivation Knowledge from Log Files Towards Motivation Diagnosis for Adaptive Systems". C. Conati, K. McCoy, and G. Paliouras (Eds.): UM 2007, LNAI 4511, pp. 197–206, 2007.
- [44] Collazos C. A., Guerrero L. A., Pino J. A., Ochoa S. F. (2002). Evaluating Collaborative Learning Processes. M. Haake and J.A. Pino (Eds.): CRIWG 2002, LNCS 2440, pp. 203–221, 2002. © Springer-Verlag Berlin Heidelberg 2002
- [45] Collazos, C. A., Guerrero, L. A., Pino, J. A., Renzi, S., Klobas, J., Ortega, M., Redondo, M. A., & Bravo, C. (2007). Evaluating Collaborative Learning Processes using System-based Measurement. *Educational Technology & Society*, 10 (3), 257-274.
- [46] Commission of the European Communities: "Adult learning: It is never too late to learn". COM(2006) 614 final. Brussels, 23.10.2006.
- [47] Cooper, J. "Cooperative Learning and College Teaching: Tips from the Trenches." *Teaching Professor*, 1990, 4(5), 1-2.
- [48] Daradoumis, T., Martínez-Mónes, A., & Xhafa, F. (2006). A Layered Framework for Evaluating OnLine Collaborative Learning Interactions". *International Journal of Human-Computer Studies*, Volume 64 , Issue 7 (July 2006), Pages 622-635
- [49] de Pedro, X. (2007). New Method Using Wikis and Forums to Evaluate Individual Contributions in Cooperative Work while Promoting Experiential Learning: Results from Preliminary Experience. Proc. of the 2007 Int. Symposium On Wikis (WikiSym), October 2007, Montreal, Canada.
- [50] Dewiyanti, S., Brand-Gruwel, S., Jochems, W., Broers N., J. (2007). Students' experiences with collaborative learning in asynchronous Computer-Supported Collaborative Learning environments. *Computers in Human Behavior* 23 (2007) 496–514
- [51] DFEE-UK, (1998). Higher Education in the Learning Society - The Report of the National Committee of Inquiry into Higher Education. <http://www.lifelonglearning.co.uk/dearing/index.htm>
- [52] Dimitracopoulou A. (2008). Computer based Interaction Analysis Supporting Self-regulation: Achievements and Prospects of an Emerging Research Direction, In Kinshuk, M.Spector, D.Sampson, P. Isaias (Guest editors). *Technology, Instruction, Cognition and Learning (TICL)*.
- [53] Dringus, L.P & Ellis, E. (2005). Using data mining as a strategy for assessing asynchronous discussion forums. *Computers & Education*, 45(2005), 140-160.
- [54] Durán E. B. "MODELO DEL ALUMNO EN SISTEMAS DE APRENDIZAJE COLABORATIVO". WAIFE -2006
- [55] Elvheim, M. (2002). Supporting Group Work in Distance Education. Conference Proceedings XP 2002, 3rd International Conference on Extreme Programming and Flexible Processes in Software Engineering. Alghero: Università di Cagliari, Italia & Free University of Bolzano-Bolzen.
- [56] European Comision. The use of ICT to support innovation and lifelong learning for all - A report on progress. Commission Staff Working Document (October 2008).

- [57] Fayyad, U., Piatetsky-shapiro, G., Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17, 37-54.
- [58] Felder, R. M. (1993). Reaching the Second Tier: Learning and Teaching Styles in College Science Education. *J. College Science Teaching*, 23(5), 286-290 (1993).
- [59] Feng, M., Beck, J., Heffernan, N. (2009). Using Learning Decomposition and Bootstrapping with Randomization to Compare the Impact of Different Educational Interventions on Learning. Second International Conference on Educational Data Mining (EDM'09). Córdoba, Julio de 2009.
- [60] Fernández, F., Montalvo, J. (2001). La enseñanza universitaria a distancia: una reflexión desde la uned. Universidad Nacional de Educación a Distancia, UNED. ISBN: 84-362-4248-3
- [61] Fiaidhi J. A. W. & Mohammed M. A. "Collaborative Virtual Learning Model for Web Intelligence". W. Shen et al. (Eds.): CSCWD 2005, LNCS 3865, pp. 563 – 572, 2006.
- [62] Field, J. (2006). *Lifelong Learning and the New Educational Order*. Trentham Books. ISBN. 1858563461
- [63] Fikes, R., Nilsson, N. (1971). STRIPS: a new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2:189-208.
- [64] Finn, C. E. (1988, Jul.-Aug.). Judgment time for higher education: In the court of public opinion. *Change*, 20(4), 34-39.
- [65] Gama, J., and Gaber, M. M. (Eds), *Learning from Data Streams: Processing Techniques in Sensor Networks*, a book published by Springer Verlag, (2007)
- [66] García Aretio, L. (2009). ¿Por qué va ganando la educación a distancia? Madrid: UNED, pp. 419, ISBN: 978-84-362-5879-0.
- [67] Gardner, Howard. (1999) "Intelligence Reframed: Multiple Intelligences for the 21st Century." Basic Books.
- [68] Gaudioso, E., Boticario, J. G. (2003) Towards web-based adaptive learning communities. *Artificial Intelligence in Education: Shaping the future of Learning through Intelligent*. 237-244. 2003
- [69] Gaudioso, E., Santos, O. C., Rodriguez, A., Boticario, J. G. (2003). A Proposal for Modelling a Collaborative Task in a Web-Based Learning Environment. Papers for the UM'03 Workshop 'User and Group models for web-based adaptive collaborative environments' in conjunction with User Modelling 2003. 22 June University of Pittsburg.
- [70] Georgsen, M., Nyvang, T. (2007). Collaborative e-learning design method. Aalborg : Institut for kommunikation, Aalborg Universitet, 2007. 24 p. (E-Learning Lab Publication Series; 12).
- [71] Ghallab, M., Nau, D., Traverso, P. (2004). *Automated Planning - Theory and Practice*, Morgan Kaufmann Publishers, 2004
- [72] Gómez-Sánchez, E., Bote-Lorenzo, M.L., Jorrín-Abellán, I.M., Vega-Gorgojo, G., Asensio-Pérez, J.I., Dimitriadis, Y. (2009). Conceptual framework for design, technological support and

- evaluation of collaborative learning. *International Journal of Engineering Education*. 25(3):557-568, May 2009.
- [73] Goodman, B. A., Linton, F. N., Gaimari, R. D., Hitzeman, J. M., Ross, H. J., Zarrella, G. (2005). Using Dialogue Features to Predict Trouble During Collaborative Learning. *User Modeling and User-Adapted Interaction* (2005) 15: 85–134
- [74] Han, J., Kamber, M. (2006). *Data Mining : Concepts and Techniques*, 2nd edition, Morgan Kaufmann, ISBN 1558609016.
- [75] Heckmann, D. (2006). Situation Modeling and Smart Context Retrieval with Semantic Web Technology and Conflict Resolution. T.R. Roth-Berghofer, S. Schulz, and D.B. Leake (Eds.): *MRC 2005, LNAI 3946*, pp. 34–47, 2006.
- [76] Holmberg, B. (2005). The evolution, principles and practices of distance education. *Bibliotheks- und Informationssystem der Universität Oldenburg*. p. 13.
- [77] Hong, Wu. "Spinning Your Course Into A Web Classroom - Advantages And Challenges". *International Conference on Engineering Education August 6 – 10, 2001 Oslo, Norway*
- [78] Hoppe, Ulrich. (1995) The use of Multiple Student Modelling to Parameterize Group Learning. In J. Greer (Ed.), *Proceedings of the 7th World Conference on Artificial Intelligence in Education*, (pp. 234-241), Washington, DC.
- [79] Hummel, H.G.K., Burgos, D., Tattersall, C., Brouns, F., Kurvers, H., Koper, R. (2005). Encouraging contributions in learning networks using incentive mechanisms. *Journal of Computer Assisted Learning*, Volume 21, Number 5, October 2005 , pp. 355-365(11)
- [80] Inaba, A., Ohkubo, R., Ikeda, M., & Mizoguchi, R. (2003). Models and Vocabulary to Represent Learner-to-Learner Interaction Process in Collaborative Learning. *Proceedings of the International Conference on Computers in Education (ICCE2003)*, pp.1088-1096, Wanchai, HongKong, December 2-5, 2003.
- [81] Ipiña, S. L. (2008). *Inferencia estadística y análisis de datos*. Pearson Prentice Hall, XII, 483 p. : 9788483224045
- [82] Jameson, A. (2003). Adaptive interfaces and agents. In Jacko, J., Sears, A., eds.: *Human-Computer interaction Handbook*. 305–330
- [83] Jermann, P., Soller, A., Muehlenbrock, M.: From Mirroring to Guiding: A Review of State of the Art Technology for Supporting Collaborative Learning. *Proceedings of the First European Conference on Computer-Supported Collaborative Learning*, Maastricht, The Netherlands, 324-331. (2001).
- [84] Johnson, D., Johnson, R. Cooperative, competitive, and individualistic learning. *Journal of Research and Development in Education*, Vol. 12,p.p. 8 –15, (1978).
- [85] Johnson, D. W., Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Edina, MN: Interaction Book Company.

- [86] Johnson, D. W., Johnson, R. T., Smith, K. A. Cooperative Learning: Increasing College Faculty Instructional Productivity. ASHE-FRIC Higher Education Report No.4. Washington, D.C.: School of Education and Human Development, George Washington University, 1991.
- [87] Johnson, R., Johnson, D. (1994). Creativity and Collaborative Learning. J. Thousand, A. Villa and A. Nevin (Eds), Brookes Press, Baltimore, 1994.
- [88] Johnson, D. W., Johnson, R. (2004) Cooperation and the use of technology. In D. Jonassen (Ed.). Handbook of research on educational communications and technology. 785-812, 2004.
- [89] Kahrimanis, G., Meier, A., Chounta, I-A., Voyiatzaki, E., Spada, H., Rummel, N., Avounis, N. (2009). Assessing Collaboration Quality in Synchronous CSCL Problem-Solving Activities: Adaptation and Empirical Evaluation of a Rating Scheme. Learning in the Synergy of Multiple Disciplines. 4th European Conference on Technology Enhanced Learning, EC-TEL 2009. Nice, France, September 29–October 2, 2009
- [90] Kantardzic, M. (2003). Data Mining: Concepts, Models, Methods, and Algorithms. John Wiley & Sons.
- [91] Kay, J. (1999) Ontologies for reusable and scrutable student models, Mizoguchi, R, (ed) AIED Workshop W2: Workshop on Ontologies for Intelligent Educational Systems, 72-77
- [92] Kay, J., Lum, A. (2005). Ontology-based user modelling for semantic web. 10th International Conference on User Modeling (UM'05) Workshop 8, Lora Aroyo, Vania Dimitrova, Judy Kay, pp 11-19.
- [93] Kay, J., Kummerfeld, B. (2006). Scrutability, User Control and Privacy for Distributed Personalization. CHI 2006 Workshop on Privacy-Enhanced Personalization
- [94] Kay, J., Li, L.: Scrutable Learner Modelling and Learner Reflection in Student Self-assessment. M. Ikeda, K. Ashley, and T.-W. Chan (Eds.): ITS 2006, LNCS 4053, pp. 763 – 765, (2006).
- [95] Kirk, R. E. (1995). Experimental design: Procedures for the behavioral sciences. Pacific Grove, CA: Brooks/Cole.
- [96] Kobsa, A. (2001). Generic User Modeling Systems. In P. Brusilovsky, A. Kobsa, W. Nejdl, eds.: The Adaptive Web: Methods and Strategies of Web Personalization. Heidelberg, Germany: Springer Verlag.
- [97] Kobsa, A. (2007). Privacy-enhanced personalization. Communications of the ACM. Volume 50 , Issue 8 (August 2007)
- [98] Kohonen, T. (1989). Self-organization and associative memory. Springer Verlag 1989
- [99] Kyriacou, D., Davis, H. C., Tiropanis, T. A. (2009). (multi-domain'sional) Scrutable User Modelling Infrastructure for Enriching Lifelong User Modelling. Lifelong User Modelling. UMAP, 2009
- [100] Lehmann, J., Hitzler, P. (2009). Concept learning in description logics using refinement operators. Machine Learning (in press).

- [101] Lin, F-R., Hsieh, L-S., Chuango, F-T. (2009). Discovering genres of online discussion threads via text mining. *Computers & Education*, 52. 481–495
- [102] Liu, Y. Ginther, D. "Cognitive Styles and Distance Education". *Online Journal of Distance Learning Administration*, Volume II, Number III, Fall1999
- [103] Losada, N. (1996). *Estudio Del Perfil Psicopedagógico Del Alumno Del Curso de Acceso Directo para Mayores de 25 Años de la UNED*. ISBN 10: 8436233468. Editorial: Universidad Nacional de Educación a Distancia
- [104] Macdonald, J. (2003) Assessing online collaborative learning: process and product. *Computers & Education*, 40,377-391
- [105] Mardia, K. V., Kent, J. T., Bibby, J. M. (1979). *Multivariate Analysis*. Academic Press,
- [106] Mark, G. and A. Kobsa (2005). The Effects of Collaboration and System Transparency on CIVE Usage: An Empirical Study and Model. *Presence* 14(1), 60-80, MIT Press.
- [107] Martínez, A., Fuente, P. De La, Dimitriadis. Y.: An Xml-Based Representation Of Collaborative Interaction. *Proceedings of CSCL 2003*. 379-383. (2003)
- [108] Martínez, A., Dimitriadis, Y., Gómez, E., Jorrín, I., Rubia, B., Marcos, J. A. (2006). Studying participation networks in collaboration using mixed methods. *International Journal of Computer-Supported Collaborative Learning*. Volume 1, Number 3 / septiembre de 2006. 383-408
- [109] Martínez-Monés, A., Dimitriadis, Y., de la Fuente Redondo, P. Contributions to analysis of interactions for formative evaluation in CSCL in *Computers and education. Towards a lifelong learning society*, Llamas Nistal, M. (Editor), Fernández Iglesias, M.J. (Editor), Anido Rifon, L.E. (Editor), Kluwer Academic Publishers, Dordrecht/Boston/London, 2007.
- [110] Martínez, R., Bosch M., Herrero, M. M., Nuño, A. S.: "Psychopedagogical components and processes in e-learning. Lessons from an unsuccessful on-line course". *Computers in Human Behavior* 23 (2007) 146–161
- [111] McCalla, G., Vassileva, J., Greer, J, Bull, S. Active Learner Modelling. In Gauthier, G.,Frasson, C., and VanLehn, K., eds, *Intelligent Tutoring Systems: 5th International Conference, ITS 2000*, Montreal, Canada, June 19-23, 2000, Heidelberg, Germany: Springer-Verlag, pages 53-62.
- [112] McFarland, S. G. 1981 (Efectos del orden de las preguntas de respuestas en encuestas) "Effects of question order on survey responses." *Public Opinion Quarterly*. Verano. Vol. 45 # 2). Págs 208-205.
- [113] McGrath, J. E. (1984). *Groups: Interaction and Performance*. Inglewood, N. J.: Prentice Hall, Inc.
- [114] Meilã, M., Heckerman, D. "An Experimental Comparison of Model-Based Clustering Methods".*Machine Learning*, 42, 9–29, 2001
- [115] Meier, A., Spada, H., Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *Computer-Supported Collaborative Learning* 2:63–86

- [116] Merchán, F. J. (2005). Enseñanza, examen y control. Profesores y alumnos en la clase de historia. Barcelona: Octaedro.
- [117] Minaei-Bidgoli, B., Punch, W. (2003). Using Genetic Algorithms for Data Mining Optimization in an Educational Web-based System. *Genetic and Evolutionary Computation, Part II*. 2003. pp.2252–2263.
- [118] Mira, J, Delgado, AE., Boticario, J.G., Díez, F.J. Aspectos básicos de la Inteligencia Artificial. Sanz y Torres, edición del 1995 o edición del 2003.
- [119] Mizoguchi, R. (2005). The Role of Ontological Engineering for AIED Research. *ComSIS Vol.2, No.1, June 2005*.
- [120] Muehlenbrock, M. (2005). Formation of Learning Groups by using Learner Profiles and Context Information. In Looi, C.-K. and McCalla, G., editors. *Proceedings of the 12th International Conference on Artificial Intelligence in Education AIED-2005*. Amsterdam, The Netherlands, 2005
- [121] Nardi, B. A. (1966): MIT Press. pp (1-20) ISBN 0-262-14058-6
- [122] Nielsen, J. (2000). Usabilidad: Diseños de sitios Web. Pearson Educación. S. A., Madrid, 2000.
- [123] Nugent, R., Ayers, E., Dean, N. (2009). Conditional Subspace Clustering of Skill Mastery: Identifying Skills that Separate Students. *Second International Conference on Educational Data Mining (EDM'09)*. Córdoba, Julio de 2009.
- [124] Otero, J., Sánchez, L. (2005). Induction of Descriptive Fuzzy Classifiers with the Logitboost Algorithm. *Soft Computing* 2005, 10(9), pp. 825-835.
- [125] Papadopoulos, P. M., Demetriadis, S. N., Stamelos, I. G. (2009). The Impact of Prompting in Technology-Enhanced Learning as Moderated by Students' Motivation and Metacognitive Skills. *Learning in the Synergy of Multiple Disciplines. 4th European Conference on Technology Enhanced Learning, EC-TEL 2009*. Nice, France, September 29–October 2, 2009
- [126] Park, C. J., Hyun, J. S. (2006). Comparison of Two Learning Models for Collaborative e-Learning". *Z. Pan et al. (Eds.): Edutainment 2006, LNCS 3942*, pp. 50–59, 2006.
- [127] Parsell, M. (2008). Pernicious virtual communities: Identity, polarisation and the Web 2.0. *Ethics and Information Technology. Volume 10, Number 1 / marzo de 2008*
- [128] Perera, D., Kay, J., Yacef, K., Koprinska, I. (2007). Mining learners' traces from an online collaboration tool. *Workshop Educational Data Mining, Proceedings of the 13th International Conference of Artificial Intelligence in Education*. Marina del Rey, CA. USA. July 2007
- [129] Perraton, H. (1988). A theory for distance education. In D. Sewart, D. Keegan, & B. Holmberg (Eds.), *Distance education: International perspectives* (pp. 34-45). New York: Routledge.
- [130] Peters, O. (1993). Distance education in a post-industrial society. En Keegan (ed) *Theoretical principles of distance education*. London & New York: Routledge

- [131] Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology*, 92, 544-555.
- [132] Prata, D., Baker, R., Costa, E., Rose, C., Cui, Y. Detecting and Understanding the Impact of Cognitive and Interpersonal Conflict in Computer Supported Collaborative Learning Environments. EDM'09 (2009)
- [133] Quinlan, J.R. (1993). C4.5: Programs for Machine Learning. Morgan Kaufman.
- [134] Ravi, S., Kim, J., Shaw, E. (2007). Mining On-line Discussions: Assessing Technical Quality for Student Scaffolding and Classifying Messages for Participation Profiling. 13th International Conference of Artificial Intelligence in Education. 2007.
- [135] Redondo, M.A., Bravo, C., Bravo, J., Ortega, M. (2003) Applying Fuzzy Logic to Analyze Collaborative Learning Experiences in an e-Learning Environment. *USDLA Journal*. (United States Distance Learning Association).17.2, 19-28.
- [136] Romero, C., Ventura, S.: Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications* 33 (2007) 135–146
- [137] Rosch, E. (1978). Principles of Categorization. *Cognition and Categorization* ed. Roach E & Lloyd B. 27-48. 1978
- [138] Russell, S., Norvig, P. *Inteligencia Artificial: Un enfoque moderno*. Prentice Hall, 1995, (traducido en Prentice Hall Hispanoamericana, 1996).
- [139] Rustagi, J.S. (1994). *Optimization Techniques in Statistics*. Academic Press.
- [140] Sáinz, M., Meneses, J., Fatsini, E., Prat, P. (2009). " Las estudiantes de secundaria ante los estudios de Informática y Telecomunicaciones. ". A: XI Congreso Nacional de Psicología Social. Universidad Rovira i Virgili. TARRAGONA, 1 octubre.
- [141] Sánchez, L., Otero, J. (2007). Boosting Fuzzy Rules in Classification Problems under Single-winner Inference (in press). *International Journal of Intelligent Systems*.
- [142] Santos, O.C., Rodríguez, A., Gaudioso, E., Boticario, J.G. Helping the tutor to manage a collaborative task in a web-based learning environment. *AIED 2003: Supplementary Proceedings*. 153-162. (2003)
- [143] Santos, O.C., Boticario, J.G. (2004). Supporting a collaborative task in a web-based learning environment with Artificial Intelligence and User Modelling techniques. *Proceedings of the VI International Symposium on Educative Informatics (SIIE'04)*.
- [144] Santos, O.C., Rodríguez, A., Gaudioso, E., Boticario, J.G. (2004). Cómo gestionar la colaboración en la tarea del Marco Lógico Colaborativo en un entorno de aprendizaje adaptativo basado en web. *Revista Iberoamericana de Inteligencia Artificial, Asociación española de inteligencia artificial (AEPIA)*, Volume 8, Number 24, Spain, p.121-129 (2004)
- [145] Santos, O.C., Boticario, J.G., Raffenne, E., Pastor, R. (2006). Why using dotLRN? UNED use cases. *1stst International Conference on FLOSS: Free/Libre/Open Source Systems*. Jerez de la Frontera, Spain, Universidad de Cadiz, p. 195-212, 10/2006.



- [146] Santos, O.C., Boticario, J.G. (2008). Users' experience with a recommender system in an open source standard-based learning management system. 4th Symposium of the WG HCI&UE of the Austrian Computer Society - Usability & HCI for Education and Work (USAB 2008), vol. 5298, Graz (Austria), Springer-Verlag, pp. 185-204, 20/11/2008.
- [147] Sarro, L. M., Debosscher, J., López, M., Aerts, C. (2009). Automated supervised classification of variable stars. II. Application to the OGLE database. *Astronomy and Astrophysics*, Volume 494, Issue 2, 2009, pp.739-768
- [148] Schreck, J. (2003). *Security and privacy in user modeling*. Publicado por Springer, 2003.
- [149] Smith, K. A. (1986). *Cooperative Learning Groups*. In S. F. Schmoberg (ed.), *Strategies for Active Teaching and Learning in University Classrooms*. Minneapolis: Office of Educational Development Programs, University of Minnesota, 1986.
- [150] Smith, M., K., Welty, C., McGuinness, D., L. Editors. (2004). *OWL Web Ontology Language Guide*. W3C Recommendation, 10 February 2004, <http://www.w3.org/TR/2004/REC-owl-guide-20040210/> . Latest version available at <http://www.w3.org/TR/owl-guide/>
- [151] Soller, A. (2001). Supporting Social Interaction in an Intelligent Collaborative Learning System. *International Journal of Artificial Intelligence in Education*, 12(1), 40-62. (2001)
- [152] Soller, A., Martínez-Monés, A., Jermann, P., Muehlenbrock, M. (2005). From Mirroring to Guiding: A Review of State of the Art Technology for Supporting Collaborative Learning. *International Journal of Artificial Intelligence in Education*, 15 (4), 261-290.
- [153] Soller, A. (2007). Adaptive support for distributed collaboration. In P. Brusilovsky, A. Kobsa, & W. Nejil (Eds.) *The Adaptive Web: Methods and Strategies of Web Personalization* (pp. 573-595). Berlin Heidelberg: Springer-Verlag.
- [154] Srinivasa, N. R. (2005). Data mining in e-commerce: A survey. *Sādhanā* Vol. 30, Parts 2 & 3, April/June 2005, pp. 275–289
- [155] Steffens, K. (2001). Self-regulation and computer based learning. *Anuarion de Psicología*, 2001, vol. 32(2), 77-94
- [156] Sternberg, R.J. (1997) *Thinking Styles*. Cambridge University Press, New York.
- [157] Strijbos, J-W., Fischer, F. (2007). Methodological challenges for collaborative learning research. *Learning and Instruction* 17.
- [158] Sutherland, R., Robertson, S., John. P. (2008). *Improving Classroom Learning with ICT*. Routledge. ISBN: 9780415461740
- [159] Talavera, L., Gaudioso, E. (2004). Mining Student Data To Characterize Similar Behavior Groups In Unstructured Collaboration Spaces. In: *Proceedings of the Workshop on Artificial Intelligence in CSCL*. 16th European Conference on Artificial Intelligence, (ECAI 2004), Valencia, Spain (2004) 17–23
- [160] Tanimoto, S. L. (2007). Improving the Prospects for Educational Data Mining. 11th International Conference on User Modeling (UM'07). (2007)

- [161] Torres, J. S., Mejías, F. T., Milán, E. G. (1999). *Procesos psicológicos básicos*. Ed. McGraw Hill, 1999.
- [162] van Velsen, L., van der Geest, T., Klaassen, R. Steehouder, M. (2008). User-centered evaluation of adaptive and adaptable systems: a literature review. *The Knowledge Engineering Review* Vol. 23:3, 261-281 (2008). Cambridge University Press.
- [163] Vidou, G., Dieng-Kuntz, R., Ghadi, A. E., Evangelou, C., Giboin, A., Tifous, A., Jacquemart, S. (2006). Towards an Ontology for Knowledge Management in Communities of Practice. U. Reimer and D. Karagiannis (Eds.): *PAKM 2006, LNAI 4333*, pp. 303 – 314, 2006.
- [164] Villasclaras-Fernandez, E. D., Isotani, S., Hayashi, Y., Mizoguchi, R. (2009). Looking into collaborative learning: design from macro- and micro-script perspectives. In *Proc. of 14th International Conference on Artificial Intelligence in Education (AIED2009)*, pp. 231-238, Brighton, UK, July 6-10, 2009.
- [165] VIÑAO, A. (2004). *Escuela para todos. Educación Modernidad en la España del siglo' XX*. Madrid: Marcial Pons.
- [166] Wang, M., Leung, Y., Zhou, C., Pei, T., Luo, J. (2006). A Mathematical Morphology Based Scale Space Method for the Mining of Linear Features in Geographic Data. *Data Mining and Knowledge Discovery*. Volume 12, Number 1 / enero de 2006
- [167] Winter, M., McCalla, G. (2003). An analysis of group performance in terms of the functional knowledge and teamwork skills of group members. *Workshop on User and Group Models for Web-based Collaborative Environments, Ninth International Conference on User Modeling (UM 2003)*, pp.35–45.
- [168] Witten, I. H., Frank, E. (2005). *Data Mining*. Morgan Kaufmann, June, 2005.
- [169] Zimmerman, B.J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25, 3-17.

# Apéndices

## Apéndice I: encuestas

### Encuesta inicial curso 2006/2007

#### *Encuesta inicial del curso virtual*

¿Trabajas o haces alguna otra actividad que te quite tiempo de estudio?

- Sí, jornada completa
- Sí, media jornada
- No, tengo tiempo para estudiar todo el día

¿Cuándo sueles tener más tiempo para estudiar?

- Por las mañanas.
- Por las tardes
- Por las mañanas y por las tardes
- Ni por las mañanas ni por las tardes

Cuando todavía no es época de exámenes, ¿qué momento dedicas más al estudio?

- Durante los días laborables
- Durante los días festivos y fines de semana
- Sólo estudio en época de exámenes
- Ninguna de las anteriores respuestas

¿Podrás trabajar en la práctica en Navidades?

- Sí, podré algunas mañanas o tardes
- Sí, algunas mañanas
- Sí, algunas tardes
- Sí podré, pero ahora mismo no sé cuando
- Es casi imposible que encuentre tiempo
- No podré

Cuando yo era estudiante de la UNED planificaba el estudio de la siguiente forma: cada día de la semana lo dejaba para una asignatura y el fin de semana para repasar, si tenía ganas. Otros compañeros lo hacían de otras formas. ¿Planificas de alguna manera tu también el estudio?

- Sí, estudio un poco de algunas asignatura todos los días
- Sí, estudio una asignatura por día de la semana
- Sí, estudio un poco de algunas asignaturas cada fin de semana ya que no tengo tiempo en días laborables
- Sí, estudio una asignatura o dos cada fin de semana ya que no tengo tiempo entre diario
- Sí, cada semana me centro en una asignatura
- Sí, estudio casi al completo una asignatura y cuando termino paso a otra
- Sí, pero sólo poco antes de los exámenes y sigo el orden en el que voy a hacer los exámenes
- No hago ningún plan. Estudio cuando tengo tiempo y lo que creo más importante
- Planifico pero de otra forma

¿Cómo sigues el programa o plan que haces para estudiar un curso de la UNED?

- A raja tabla (100%)
- Soy flexible pero no mucho (75%)
- Soy flexible (50%)
- Soy muy flexible (25%)
- El programa de estudio y lo que al final estudio, apenas tienen nada que ver (0%)

¿Cuál de las siguientes frases se ajusta mejor a tu forma de estudiar en la UNED?

- Estudio de forma continua desde el principio del cuatrimestre. Al final, como no, hago un sprint
- Al principio del cuatrimestre apenas estudio, y poco a poco lo hago con más intensidad. Al final, como no, hago un sprint
- Cuando creo que queda poco para los exámenes empiezo a estudiar con mucha intensidad
- Ninguna de las frases anteriores

¿Cambias o has cambiado alguna vez tu forma de estudiar?

- Sí, hago cambios dependiendo de la asignatura
- Sí, hago cambios dependiendo del tiempo disponible
- Sí, hago cambios dependiendo de lo motivado que esté
- Sí, hago cambios dependiendo de la asignatura y del tiempo disponible
- Sí, hago cambios dependiendo de la asignatura y de lo motivado que esté
- Sí, hago cambios dependiendo del tiempo disponible y de lo motivado que esté
- Sí, hago cambios dependiendo de la asignatura, del tiempo disponible y de lo motivado que esté
- No, siempre he estudiado con la misma forma y me va bien

¿Tienes algún comentario a las anteriores preguntas?

#### *Conocimientos*

¿Has estudiado anteriormente algo sobre Inteligencia Artificial?

- Sí, una asignatura optativa en el primer ciclo
- Sí, ya que me he matriculado antes de esta asignatura, aunque no me presenté a los exámenes
- Sí, ya que me he matriculado antes de esta asignatura, aunque no aprobé
- Sí, porque me he matriculado de asignaturas del segundo ciclo relacionadas con Inteligencia Artificial
- No, es la primera vez que estudiaré algo relacionado a Inteligencia Artificial
- Si has estudiado la asignatura optativa "Introducción a la IA" del primer ciclo, ¿podrías decirme la nota que obtuviste?

¿Si vieras el código fuente de un programa informático, podrías identificar en que lenguaje ha sido escrito? ¿Qué lenguajes identificarías?

- C
- Java
- XML
- LISP
- CLISP
- JESS
- Basic
- OWL

Como alumno de la UNED has tenido la oportunidad de utilizar los cursos virtuales. ¿Qué grado de uso les has dado?

- Los he utilizado con mucha frecuencia
- Los he utilizado con frecuencia
- Los he utilizado con poca frecuencia
- No los he utilizado o apenas los he utilizado

¿Has utilizado antes las plataformas de la UNED aLF o Innova, o una plataforma de educación llamada DotLRN?

- Sí
- Sí, pero muy poco
- No

¿Tienes algún comentario que hacer a las anteriores preguntas?

*Personal*

Sexo:

- Hombre
- Mujer

Edad:

Localidad de residencia:

Provincia de residencia:

¿Tienes algún comentario sobre las anteriores preguntas?

# Encuesta inicial curso 2007/2008

## *Instrucciones*

El objetivo de este cuestionario es obtener información sobre ti que pueda ser utilizada para mejorar la gestión de un equipo tanto por el tutor, como por los compañeros. Esta información, por tanto, será visible por tus compañeros, además de por el tutor.

Piensa detenidamente cada una de las preguntas y responde con sinceridad utilizando todo el tiempo que consideres oportuno. Una vez terminada la encuesta, podrás cambiar los datos cuando estimes oportuno.

## *Antecedentes laborales y académicos*

¿Qué has estudiado? ...

¿Qué título has obtenido? ...

¿En qué área has trabajado? ...

¿Cuántos años de experiencia? ...

¿Permites que estás respuestas puedan ser vistas por tus compañeros?

- Sí
- No

Comentarios: ...

## *Preferencias en el estudio y en el trabajo en equipo*

Cuando yo era estudiante de la UNED planificaba el estudio de la siguiente forma: cada día de la semana lo dejaba para una asignatura y el fin de semana para repasar, si tenía ganas. Otros compañeros lo hacían de otras formas. ¿Planificas de alguna manera tú también el estudio?

- Sí, una asignatura al día
- Sí, una asignatura a la semana
- Sí, una asignatura el fin de semana
- Sí, varias asignaturas al día
- Sí, varias asignaturas a la semana
- Sí, varias asignaturas el fin de semana
- Sí, en secuencia
- Sin plan
- Otros

¿Eres estricto con el plan de estudios?

- Muy estricto
- Estricto
- Poco estricto
- Caótico

¿Qué estrategia utilizas al estudiar?

- Constante desde el inicio al final
- Progresivo, a principio con más tranquilidad
- Explosivo, de vez en cuando estudio mucho
- Otras

¿Cambias la estrategia por el tiempo que tienes de estudio?

- Sí
- No

¿Cambias la estrategia por el tipo de asignatura?

- Sí
- No

¿Cambias la estrategia por la motivación que tienes?

- Sí
- No

¿Qué tipo de días podrás colaborar por Internet?

- Días laborables
- Fines de semana y festivos
- Antes de los exámenes
- Nunca tengo tiempo

¿Durante el día cuándo puedes colaborar?

- Por las mañanas
- Por las tardes
- Por las noches
- Nunca

¿Podrá colaborar en Navidad?

- Todas las mañanas
- Todas las tardes
- Algunas mañanas
- Algunas tardes
- Casi ninguna mañana
- Casi ninguna tarde
- Nunca

¿Permites que estás respuestas puedan ser vistas por tus compañeros?

- Sí
- No

Comentarios: ...

#### *Conocimiento previo*

¿Si vieras el código fuente de un programa informático, podrías identificar en que lenguaje ha sido escrito? ¿Qué lenguajes identificarías?

- C
- Java
- XML
- LISP
- CLISP
- JESS
- HTML
- Basic
- OWL
- Javascript

Como alumno de la UNED has tenido la oportunidad de utilizar los cursos virtuales. ¿Qué grado de uso les has dado?

- Los he utilizado con mucha frecuencia
- Los he utilizado con frecuencia
- Los he utilizado con poca frecuencia
- No los he utilizado o apenas los he utilizado

¿Has utilizado antes las plataformas de la UNED aLF o Innova, o una plataforma de educación llamada DotLRN

- Sí
- Sí, pero muy poco
- No

Comentarios: ...

#### *Datos demográficos personales y sobre el estudio y el trabajo*

Sexo:

- Varón
- Mujer

Año de nacimiento (Ejem.: 1974) ...

Localidad de residencia: ...

¿Tienes responsabilidades que te impidan estudiar durante todo el día, si quisieras?

- Responsabilidades laborales

- Responsabilidades familiares
- Ambas
- Ninguna

¿Qué asignaturas cursas este año? ...

¿Eres repetidor?

- Sí
- No

¿En qué localidad se encuentra el centro asociado de la UNED donde puedes ir a las tutorías? ...

¿Cuántas veces vas el centro asociado?

- Nunca
- Una vez por semana
- Varias por semana
- Algunas al mes

¿Qué jornada de trabajo, u otras tareas, tienes?

- Completa
- Mañana
- Tarde
- Noche
- Flexible

¿Cuál es el campo en el que trabajas? ...

¿En qué localidad trabajas? ...

¿A qué cantidad de euros se aproxima más tu sueldo al mes? (Suma la cantidad al año y divide por 12)

- 500
- 1000
- 1500
- 2000
- 2500
- 3000

¿Permites que estás respuestas puedan ser vistas por tus compañeros?

- Sí
- No

Comentarios: ...



# Encuesta inicial curso 2008/2009

## *Instrucciones*

El cuestionario que a continuación presentamos tiene el objetivo de mejorar la gestión de equipos de estudio, tanto por parte del tutor como por los compañeros, a partir de la información que nos proporciones sobre ti.

Piensa detenidamente cada una de las preguntas y responde con sinceridad utilizando todo el tiempo que consideres oportuno. Ten en cuenta que la información que des se podrá hacer pública. Salvo ciertos datos personales, como el sexo o la fecha de nacimiento, no será obligatorio responder a ninguna otra pregunta. Si te molesta responder a alguna pregunta, no lo hagas.

## *Datos necesarios de colaboración*

¿Dónde resides?

Ciudad:

Provincia:

Sexo:

- Varón
- Mujer

Año de nacimiento (Ejem.: 1974) ...

¿Tienes responsabilidades que te impidan estudiar durante todo el día, si quisieras? ¿De qué tipo?

- Responsabilidades laborales
- Responsabilidades familiares
- Ambas
- Ninguna

¿Cuándo podrás comunicarte por Internet?

- Todos los días
- Todos los días por las mañanas
- Todos los días por las tardes
- Días laborables
- Días laborables por las mañanas
- Días laborables por las tardes
- Fines de semana y festivos

## *Preferencias en el estudio*

¿De qué forma planificas tu estudio?

- Estudio una asignatura al día
- Estudio una asignatura a la semana
- Estudio una asignatura el fin de semana
- Estudio varias asignaturas al día
- Estudio varias asignaturas a la semana
- Estudio varias asignaturas el fin de semana
- Estudio en secuencia
- Estudio pero sin plan
- Otros

¿En qué grado eres rígido con tu plan de estudios?

- Muy rígido
- Rígido
- Poco rígido
- Caótico

¿Qué estrategia utilizas cuando estudias?

- Constante desde el inicio al final
- Progresivo, a principio con más tranquilidad
- Explosivo, de vez en cuando estudio mucho

- Otras

¿Cambias de estrategia según el tiempo de estudio de que dispones?

¿Por qué cambias de estrategia de estudio?

- Sí, hago cambios dependiendo de la asignatura
- Sí, hago cambios dependiendo del tiempo disponible
- Sí, hago cambios dependiendo de lo motivado que esté
- Sí, hago cambios dependiendo de la asignatura y del tiempo disponible
- Sí, hago cambios dependiendo de la asignatura y de lo motivado que esté
- Sí, hago cambios dependiendo del tiempo disponible y de lo motivado que este
- Sí, hago cambios dependiendo de la asignatura, del tiempo disponible y de lo motivado que esté
- No, siempre he estudiado con la misma forma y me va bien

#### *Conocimiento previo*

¿Si vieras el código fuente de un programa informático, podrías identificar en que lenguaje ha sido escrito? ¿Qué lenguajes identificarías?

- C
- Java
- XML
- LISP
- CLISP
- JESS
- HTML
- Basic
- OWL
- Javascript

Como alumno de la UNED, has tenido la oportunidad de utilizar los cursos virtuales. ¿Qué grado de uso les has dado?

- Los he utilizado con mucha frecuencia
- Los he utilizado con frecuencia
- Los he utilizado con poca frecuencia
- No los he utilizado o apenas los he utilizado

¿Has utilizado antes las plataformas de la UNED aLF o Innova, o una plataforma de educación llamada DotLRN

- Sí
- Sí, pero muy poco
- No

#### *Datos demográficos personales y sobre el estudio y el trabajo*

¿Qué asignaturas cursas este año? (Indícalas en una lista separada por comas)...

¿Eres repetidor?

- Sí
- No

¿En qué localidad se encuentra el centro asociado de la UNED donde puedes ir a las tutorías?

Ciudad: ...

Provincia: ...

¿Cuántas veces vas al centro asociado?

- Nunca
- Una vez por semana
- Varias por semana
- Algunas al mes

¿Cuál es tu profesión? ...

¿Qué jornada de trabajo, u otras tareas, tienes?

- Completa
- Mañana
- Tarde
- Noche
- Flexible

¿En qué localidad trabajas?

Ciudad:...

Provincia:...

¿A qué cantidad de euros se aproxima más tu sueldo al mes? Esta pregunta es un poco molesta en la actual sociedad, no respondas si no lo crees oportuno.

- 0
- 500
- 1000
- 1500
- 2000
- 2500
- 3000

Comentarios a la encuesta en general: ...

# Encuesta final curso 2006/2007

## *Instrucciones*

Ya que el método utilizado este año para hacer la práctica de planificación ha sido novedoso tanto para los alumnos como para el equipo docente, es necesario conocer tu opinión. Además espero sacar de esta experiencia un estudio que lleve por un lado a la conclusión de una tesis doctoral, la mía como profesor en esta experiencia, y por otro a la mejora de la enseñanza en la UNED gracias a las nuevas tecnologías unidas a la colaboración entre compañeros.

La principal hipótesis en esta experiencia ha sido que la colaboración directa, la que ha habido por la vía virtual, como la indirecta, la que ha habido por la vía normal, ha mejorado el aprendizaje. Por supuesto que esta hipótesis se puede descomponer en varias sólo con tener en cuenta distintos matices ya que no todos los alumnos se han comportado igual ni han utilizado la plataforma de la misma forma. Esto es lo que al estudio puede dar una gran riqueza y lo que espero conocer gracias a las interacciones que los alumnos han hecho en la plataforma y de las respuestas a las encuestas. Por eso te pido que pierdas unos minutos respondiendo a esta encuesta con sinceridad.

Para evitar malos entendimientos ya que estamos en época de exámenes y de corrección de los mismo, informo que responder a la encuesta es un acto voluntario y que no analizaré los resultados hasta después de haber corregido tanto las prácticas como los exámenes. Esto es muy fácil de justificar ya que analizar esta encuesta llevará mucho tiempo y el tiempo que tengo ahora lo tengo que dedicar casi por completo a corregir. Con esto último quiero advertir que no tendré en cuenta para nada las respuestas a la hora de evaluar ni los exámenes ni las prácticas. Tienes total libertad de responder lo que quieras.

Esta encuesta es la misma para todos los alumnos sin importar la vía elegida. Sin embargo, el último apartado es para evaluar la práctica por la vía virtual. Si elegiste hacer la práctica por la vía virtual, espero que respondas, pero si no elegiste la vía virtual, puedes pasar de largo sin problemas.

Lee con detenimiento las preguntas y las respuestas, y responde con sinceridad.

## *Datos personales*

1. Nombre
2. Apellidos
3. Tipo de vía:
  - a. Virtual
  - b. Normal
4. ¿Por qué elegiste un tipo de vía en vez de otro?
  - a. Elegí la vía normal porque mis circunstancias (trabajo, familia, etc.) me iban a absorber demasiado tiempo durante la fase de colaboración.
  - b. Elegí la vía normal para tener más libertad a la hora de planificar mi estudio.
  - c. Elegí la vía normal porque no me enteré de que podría elegir entre dos vías hasta que fue demasiado tarde.
  - d. Elegí la vía normal porque sabía que el trabajo en equipo iba a causar demasiados problemas.
  - e. Elegí la vía normal por otras causas.
  - f. Elegí la vía virtual sólo por las ventajas a la hora de hacer el examen.
  - g. Elegí la vía virtual sólo porque sabía que con el trabajo en equipo iba a aprender más.
  - h. Elegí la vía virtual por las ventajas en el examen y por trabajar en equipo.
  - i. Elegí la vía virtual porque gente que conocía la había elegido.
  - j. Elegí la vía virtual por otras razones
5. ¿Por qué estudias en la UNED?
  - a. No me llegó la nota de selectividad o del curso de acceso para ir a otra universidad.
  - b. Es la más cercana a donde yo vivo.
  - c. Es la que mejor se adapta a mis circunstancias de tiempo y esfuerzo (trabajo, familia, etc.).
  - d. Es una de las mejores universidades que conozco.
  - e. Es muy barata.
  - f. Otras razones.
6. Comentarios:

## *Evaluación de la práctica*

7. En general, evalúa la práctica entre 1 y 5:
  - a. 5 (Muy bien).
  - b. 4 (Bien).
  - c. 3 (Regular).
  - d. 2 (Mal).
  - e. 1 (Muy mal).
8. ¿Cuál es el principal problema que has encontrado?
  - a. Problemas de instalación.
  - b. Problemas por la escasa documentación.
  - c. Problemas de desarrollo con los planificadores.

- d. No entendía bien los enunciados y no sabía que había que hacer.
  - e. Los problemas causados por los compañeros.
  - f. El poco caso del profesor.
  - g. La plataforma.
  - h. Por mis circunstancias no tenía mucho tiempo para hacerla.
  - i. El tiempo que he gastado haciendo la práctica.
  - j. Otros problemas.
  - k. No he tenido ningún problema de consideración.
9. ¿Cuál ha sido la principal ventaja que te has encontrado? (Ídem a la anterior pregunta).
- a. La instalación.
  - b. La documentación.
  - c. El desarrollo con los planificadores.
  - d. La claridad de los enunciados de los problemas.
  - e. La ayuda de los compañeros.
  - f. La ayuda del profesor.
  - g. La plataforma.
  - h. La facilidad para organizar mi trabajo sin problemas.
  - i. La cantidad de cosas aprendidas respecto al tiempo gastado.
  - j. Otras ventajas.
  - k. No ha habido ninguna ventaja reseñable.
10. Selecciona de las siguientes frases la que más se ajuste a tu forma de utilizar la plataforma:
- a. Apenas he entrado en la plataforma.
  - b. No he utilizado mucho la plataforma, apenas para buscar en los foros si algún compañero había solucionado dudas parecidas a las mías.
  - c. He utilizado la plataforma para buscar respuestas a casi todas las dudas que han ido surgiendo
  - d. He utilizado la plataforma para buscar respuestas y para preguntar a los compañeros.
  - e. He utilizado la plataforma para principalmente preguntar las dudas que surgían en el momento.
  - f. He utilizado la plataforma para buscar respuestas y preguntar por dudas que no conseguía resolver después de mucho tiempo.
  - g. He utilizado la plataforma para buscar respuestas, preguntar dudas y resolver las dudas de otros compañeros.
  - h. No hay ninguna frase que se ajusta a la forma en la que he utilizado la plataforma.
11. Evalúa el trabajo del profesor:
- a. 5 (Muy bien).
  - b. 4 (Bien).
  - c. 3 (Regular).
  - d. 2 (Mal).
  - e. 1 (Muy mal).
12. Evalúa la ayuda que los foros del espacio virtual principal te han dado:
- a. 5 (Muy bien).
  - b. 4 (Bien).
  - c. 3 (Regular).
  - d. 2 (Mal).
  - e. 1 (Muy mal).
13. Evalúa la ayuda obtenida de los compañeros en el espacio virtual principal:
- a. 5 (Muy bien).
  - b. 4 (Bien).
  - c. 3 (Regular).
  - d. 2 (Mal).
  - e. 1 (Muy mal).
14. Comentarios:

*Evalúa la plataforma*

15. Evalúa la plataforma en general:
- a. 5 (Muy bien).
  - b. 4 (Bien).
  - c. 3 (Regular).
  - d. 2 (Mal).
  - e. 1 (Muy mal).
16. ¿Cuál es el principal problema que has encontrado?
- a. La dificultad de aprender a utilizarla.
  - b. Los problemas con el registro.
  - c. Los mensajes de error que aparecían.
  - d. La lentitud en la conexión.
  - e. La mala estructuración de los contenidos.
  - f. El poco contacto con los compañeros que permitía.
  - g. Otros problemas.
  - h. No ha habido ningún problema reseñable.
17. ¿Cuál es la principal ventaja que has encontrado en la plataforma?
- a. La facilidad de aprender a utilizarla.
  - b. El rápido registro en la plataforma.
  - c. La rapidez en la conexión.
  - d. La buena estructuración de los contenidos.

- e. El contacto con los compañeros que permitía.
  - f. Otras ventajas.
  - g. No ha habido ninguna ventaja reseñable.
18. Comentarios.

*Evalúa la colaboración y la motivación*

19. Evalúate a ti mismo según la colaboración que has prestado:
- a. 5 (Muy bien).
  - b. 4 (Bien).
  - c. 3 (Regular).
  - d. 2 (Mal).
  - e. 1 (Muy mal).
20. ¿Ha cambiado el grado de motivación por hacer la práctica en un ambiente colaborativo?
- a. El contacto con mis compañeros ha aumentado mi grado de motivación.
  - b. Por el trabajo del equipo docente ha aumentado mi motivación.
  - c. Por el contacto con mis compañeros y el trabajo del equipo docente ha aumentado mi motivación.
  - d. Mi motivación no ha cambiado por trabajar en un ambiente colaborativo.
  - e. El contacto con mis compañeros ha disminuido mi grado de motivación.
  - f. Por el trabajo del equipo docente ha disminuido mi motivación.
  - g. Por el contacto con mis compañeros y el trabajo del equipo docente ha disminuido mi motivación.
  - h. No sé que contestar.
21. Aunque esta pregunta ya se hizo en la encuesta que se publicó al abrir el espacio virtual, se vuelve a preguntar por si ha cambiado tu forma de estudio por hacer la práctica en un ambiente colaborativo. ¿Qué forma de estudio has utilizado para trabajar en la práctica?
- a. Forma continua (empezar a estudiar o trabajar a principio del curso y mantener el ritmo).
  - b. Forma en progreso (empezar a estudiar o trabajar con un ritmo muy suave pero aumentando el ritmo progresivamente).
  - c. Forma explosiva (estudiar o trabajar casi exclusivamente y con ritmo alto en época de exámenes).
  - d. Otra forma.
  - e. Mi forma de estudio ha cambiado pero no se ha debido al trabajo en un ambiente colaborativo.
22. Comentarios:

*Preguntas sólo para los alumnos de la vía virtual*

23. Define la estructura de tu grupo de trabajo en el área virtual colaborativa:
- a. Uno de los miembros tomaba la iniciativa al principio de todas las tareas.
  - b. La iniciativa no era siempre tomada por el mismo compañero. Ha cambiado según la carga de trabajo, la tarea, etc.
  - c. A todos los compañeros nos costaba tomar la iniciativa al principio. Se esperaba que otro la tomara.
  - d. Ni nadie tomaba la iniciativa de forma significativa ni nadie dejaba de tomarla.
  - e. Uno de los compañeros no tomaba nunca la iniciativa y esperaba a lo que los otros hacían.
  - f. No hay ninguna definición que explique adecuadamente mi grupo.
24. ¿Qué has echado en falta para mejorar la coordinación y la gestión de tu grupo en el subespacio virtual?
- a. Algún servicio de comunicación síncrona como un Chat o el Messenger.
  - b. Algún servicio de edición de documentos compartida como una wiki (mirar <http://www.wikipedia.org/> para saber que es una wiki).
  - c. Algún servicio de control sobre los compañeros para saber, por ejemplo, cuanto tiempo llevan sin entrar en el grupo, cuales son las horas más comunes para que se conecten, etc.
  - d. Más control por parte del profesor.
  - e. No he echado nada en falta.
  - f. Otras respuestas.
25. ¿Has tenido problemas de coordinación con los compañeros del grupo?
- a. Sí, y han producido un gran perjuicio a la coordinación del grupo.
  - b. Sí, pero apenas ha habido perjuicio.
  - c. No ha habido ningún problema reseñable.
26. Si ha habido problemas de coordinación, di cual de entre las siguientes opciones:
- a. Al principio el grupo no funcionó muy bien porque no nos conocíamos.
  - b. No nos conectábamos a la vez.
  - c. La comunicación por los foros era muy lenta.
  - d. Algún o algunos compañeros respondían muy tarde a los foros.
  - e. Nuestros horarios de trabajo en el grupo eran distintos (unos trabajaban más entre diario y otros los fines de semana, por ejemplo).
  - f. Durante las fiestas algún compañero o yo mismo no pudimos trabajar apenas nada.
  - g. Algún compañero iba a lo suyo y no colaboraba apenas con el resto.
  - h. Algún compañero abandono el grupo en mitad de la práctica.
  - i. Algún compañero se creyó el líder del grupo y no dejaba apenas que se hiciera nada que el no quisiera.
  - j. El profesor intervenía muy poco y no solucionaba los problemas con los otros compañeros.
  - k. Otros.
27. Valora el grado de colaboración de los compañeros. Esta información no será utilizada para evaluar a tus compañeros, sino para tener una información más detallada de la colaboración a la que yo no he podido acceder.

En la caja de texto indica el nombre del compañero y valórale entre 0 y 10. El 5 es un grado de colaboración normal

- a. Caja de texto para indicar el nombre
- b. Menú desplegable con valores de 0 a 10.

28. ¿De lo que has aprendido con esta experiencia, qué es lo que ha superado más tus expectativas? (Elección múltiple)

- a. He aprendido más de lo esperado en planificación como área de conocimiento.
- b. He aprendido más de lo esperado en el uso y programación de planificadores.
- c. He aprendido más de lo esperado en la gestión del trabajo en equipo a distancia.
- d. He aprendido más de lo esperado en la realización de experimentos.
- e. He aprendido más de lo esperado en la redacción de un documento explicativo del trabajo realizado.
- f. He aprendido más de lo esperado en la discusión sobre un tema desde un punto de vista científico.
- g. No he aprendido más de lo que esperaba.

Comentarios

# Encuesta final curso 2007/2008

## *Instrucciones*

Con esta encuesta, queremos conocer tu opinión sobre el espacio virtual, el método seguido en la práctica, el equipo docente y la colaboración de los compañeros. El cuestionario tiene dos objetivos principales. Por un lado, encontrar los puntos susceptibles de mejora en el espacio virtual y el método de realización. Por otro lado, encontrar una forma de calcular la colaboración de los alumnos, para que el equipo docente pueda ir anticipándose ante los posibles problemas. Ambos objetivos son muy importantes, por lo que solicitamos que respondas con objetividad e independencia. A pesar de que la realización de esta encuesta es totalmente voluntaria y de que **NO VA A SER TENIDA EN CUENTA PARA LA NOTA FINAL**, te solicitamos tu participación. Por ello, te pedimos que respondas a ella con la mayor sinceridad posible. Te aconsejamos que leas con detenimiento las preguntas y sus respectivas opciones de respuesta antes de responder.

1. ¿En qué tipo de modalidad has hecho la práctica?
  - a. Virtual
  - b. Normal
2. Si has elegido la modalidad normal, elige el motivo o motivos que te han conducido a elegirla
  - c. Porque mis circunstancias (trabajo, familia, etc.) me iban a absorber demasiado tiempo durante la fase de colaboración.
  - d. Para tener más libertad a la hora de planificar mi estudio.
  - e. Porque no me enteré de que podría elegir entre las dos modalidades hasta que fue demasiado tarde.
  - f. Porque sabía que el trabajo en equipo iba a causar demasiados problemas.
  - g. Por otros motivos
3. Si has elegido la modalidad virtual, elige el motivo o motivos que te han conducido a elegirla
  - h. Sólo por las ventajas a la hora de hacer el examen.
  - i. Sólo porque sabía que con el trabajo en equipo iba a aprender más.
  - j. Por las ventajas en el examen y por trabajar en equipo.
  - k. Porque gente que conocía la había elegido.
  - l. Por otros motivos
4. Si has elegido la modalidad virtual, ¿cuál es el número asignado a tu equipo?
5. ¿Por qué estudias en la UNED?
  - m. No me llegó la nota de selectividad o del curso de acceso para ir a otra universidad.
  - n. Es la universidad más cercana a donde yo vivo.
  - o. Es la que mejor se adapta a mis circunstancias de tiempo y esfuerzo (trabajo, familia, etc.).
  - p. Es una de las mejores universidades que conozco.
  - q. Es muy barata.
  - r. Otras razones
6. Si has elegido otras razones en la pregunta anterior, especifica cuáles:
7. Comentarios:

## *Evaluación de la práctica*

8. En general, evalúa la práctica entre 0 –puntuación más baja- y 5-puntuación más alta:
  - a. 5
  - b. 4
  - c. 3
  - d. 2
  - e. 1
  - f. 0
9. ¿Cuál o cuáles son los principales problemas que has encontrado? (Elige la opción u opciones que mejor reflejen tu opinión)
  - g. La instalación.
  - h. Problemas por la escasa documentación.
  - i. Problemas de desarrollo con los planificadores.
  - j. No entendía bien los enunciados y no sabía qué había que hacer.
  - k. Los problemas causados por los compañeros.
  - l. El poco caso del profesor.
  - m. La plataforma.
  - n. Por mis circunstancias no tenía mucho tiempo para hacerla.
  - o. El tiempo que he pasado haciendo la práctica.
  - p. Otros problemas.
  - q. No he tenido ningún problema de consideración.
10. ¿Cuál o cuáles han sido las principales ventajas que has encontrado? (Elige la opción u opciones que mejor reflejen tu opinión)
  - r. La instalación.
  - s. La documentación.
  - t. El desarrollo con los planificadores.
  - u. La claridad de los enunciados de los problemas.
  - v. La ayuda de los compañeros.



- w. La ayuda del profesor.
  - x. La plataforma.
  - y. La facilidad para organizar mi trabajo sin problemas.
  - z. La cantidad de cosas aprendidas respecto al tiempo empleado.
  - aa. Otras ventajas.
  - bb. No ha habido ninguna ventaja reseñable.
11. Evalúa el trabajo del equipo docente a nivel general (0 –puntuación más baja- y 5-puntuación más alta):
    - cc. 5 (Muy bien).
    - dd. 4 (Bien).
    - ee. 3 (Regular).
    - ff. 2 (Mal).
    - gg. 1 (Muy mal).
    - hh. 0
  12. Evalúa la ayuda que te han proporcionado los foros del espacio virtual principal (0 –puntuación más baja- y 5-puntuación más alta):
    - ii. 5 (Muy bien).
    - jj. 4 (Bien).
    - kk. 3 (Regular).
    - ll. 2 (Mal).
    - mm. 1 (Muy mal).
    - nn. 0
  13. Evalúa la ayuda obtenida de los compañeros en el espacio virtual principal (0 –puntuación más baja- y 5-puntuación más alta):
    - oo. 5 (Muy bien).
    - pp. 4 (Bien).
    - qq. 3 (Regular).
    - rr. 2 (Mal).
    - ss. 1 (Muy mal).
    - tt. 0
  14. Comentarios:

*Evalúa la colaboración y la motivación*

15. ¿Ha cambiado tu grado de motivación por hacer la práctica en un ambiente colaborativo?
  - a. Sí
  - b. No
  - c. No sé
16. En caso de contestar que haber contestado que SÍ a la pregunta anterior, elige cuál de los siguientes aspectos ha contribuido a incrementar tu motivación por hacer la práctica:
  - d. El contacto con mis compañeros
  - e. El trabajo del equipo docente.
  - f. Lo interesante de los problemas propuestos.
  - g. El estudio de la asignatura.
  - h. La cercanía de las fechas de los exámenes.
  - i. Otros aspectos.
17. En caso de que hayas contestado que NO a la pregunta 15, elige cuál de los siguientes aspectos ha contribuido a disminuir tu motivación por hacer la práctica:
  - j. El contacto con mis compañeros.
  - k. El trabajo del equipo docente.
  - l. Lo poco interesante de los problemas propuestos.
  - m. El estudio de la asignatura.
  - n. La cercanía de las fechas de los exámenes.
  - o. Otros aspectos.
18. A continuación identifica a algunos compañeros del espacio virtual que te hayan ayudado de forma importante por lo que podrías definirlos como prototipos de la colaboración (identificalos con nombre y algún apellido y separa la lista por comas):
19. Comentarios:

*Preguntas para que SÓLO las contesten los alumnos de la modalidad virtual*

20. Define la estructura de tu grupo de trabajo en el área virtual colaborativa:
  - a. En mi grupo ha habido un líder claro.
  - b. En mi grupo ha habido un líder claro, pero ha ido cambiando de persona según las circunstancias.
  - c. En mi grupo cada uno adoptó un rol fijo (de comportamiento o trabajo) al principio del curso.
  - d. En mi grupo cada miembro adoptó distintos roles, según las circunstancias.
  - e. En mi grupo todas las decisiones han sido tomadas en grupo después de un periodo de diálogo.
  - f. Mi grupo ha sido bastante caótico.
  - g. No hay ninguna definición que explique adecuadamente mi grupo.
21. ¿Has echado en falta un servicio de Chat o mensajería instantánea para mejorar la coordinación y la gestión de tu grupo?
  - h. Sí y la falta de dichos servicios ha influido muy negativamente.
  - i. Sí, pero la falta de dichos servicios no ha influido mucho.

- j. Sí, pero apenas ha habido influencia.
  - k. No, no he echado en falta estos servicios.
22. ¿En tu equipo, se ha utilizado alguna herramienta de comunicación además de los foros que se suministraban en el espacio virtual? Por favor, en caso afirmativo señala qué tipo de herramienta o herramientas y cual ha sido la frecuencia de uso.
23. ¿Crees que para mejorar la gestión y la coordinación de tu grupo hubiera sido necesario que los compañeros hubieran compartido entre sí alguna clase de información más personal (como por ejemplo, aquella información personal que pusiste en la encuesta primera)? Señala lo que crees que hubiera mejorado la gestión y la coordinación de tu grupo:
24. ¿Has tenido problemas de coordinación con los compañeros del grupo?
- l. Sí, y han producido un gran perjuicio a la coordinación del grupo.
  - m. Sí, pero apenas ha habido perjuicio.
  - n. No ha habido ningún problema reseñable.
25. Si ha habido algún problema de coordinación, elige qué opciones son las que mejor se ajustan a lo que has percibido:
- o. Al principio, el grupo no funcionó muy bien porque no nos conocíamos.
  - p. No nos conectábamos a la vez.
  - q. La comunicación por los foros era muy lenta.
  - r. Alguno de los compañeros respondían muy tarde a los foros.
  - s. Los horarios del grupo trabajo eran distintos (unos trabajaban más a diario y otros los fines de semana, por ejemplo).
  - t. Durante las fiestas algún compañero o yo mismo no pudimos trabajar apenas.
  - u. Algún compañero iba a lo suyo y no colaboraba apenas con el resto.
  - v. Algún compañero abandonó el grupo en mitad de la práctica.
  - w. Algún compañero se creyó el líder del grupo y no dejaba que se hiciera nada que el no quisiera.
  - x. El profesor intervenía muy poco y no solucionaba los problemas con los otros compañeros.
  - y. Otros.
26. En el documento “Presentación de las estadísticas de la colaboración”, que se encuentra en la sección de documentos, aparece una lista donde se te habrá asignado un nivel de interacción. ¿Cuál crees que debería ser tu nivel de interacción?
- z. Muy Alto
  - aa. Alto
  - bb. Medio
  - cc. Bajo
  - dd. Muy bajo
27. En el documento “Presentación de las estadísticas de la colaboración”, que se encuentra en la sección de documentos, aparece una lista donde se te habrá asignado un nivel de colaboración. ¿Cuál crees que debería ser tu nivel de colaboración?
- ee. Muy Alto
  - ff. Alto
  - gg. Medio
  - hh. Bajo
  - ii. Muy bajo
28. A continuación evalúa a tus compañeros del equipo según su nivel de interacción:  
 Compañero 1: Nombre: \_\_\_\_\_ Grado: (Muy alto, Alto, Medio, Bajo, Muy bajo)  
 Compañero 2: Nombre: \_\_\_\_\_ Grado: (Muy alto, Alto, Medio, Bajo, Muy bajo)  
 Compañero 3 (si se da el caso): Nombre: \_\_\_\_\_ Grado: (Muy alto, Alto, Medio, Bajo, Muy bajo)
29. A continuación evalúa a tus compañeros del equipo según su nivel de colaboración:  
 Compañero 1: Nombre: \_\_\_\_\_ Nivel: (Muy alto, Alto, Medio, Bajo, Muy bajo)  
 Compañero 2: Nombre: \_\_\_\_\_ Nivel: (Muy alto, Alto, Medio, Bajo, Muy bajo)  
 Compañero 3 (si se da el caso): Nombre: \_\_\_\_\_ Nivel: (Muy alto, Alto, Medio, Bajo, Muy bajo)
30. De lo que has aprendido con esta experiencia, ¿qué es lo que ha superado más tus expectativas?
- jj. He aprendido más de lo esperado en planificación como área de conocimiento.
  - kk. He aprendido más de lo esperado en el uso y programación de planificadores.
  - ll. He aprendido más de lo esperado en la gestión del trabajo en equipo a distancia.
  - mm. He aprendido más de lo esperado en la realización de experimentos.
  - nn. He aprendido más de lo esperado en la redacción de un documento explicativo del trabajo realizado.
  - oo. He aprendido más de lo esperado en la discusión sobre un tema desde un punto de vista científico.
  - pp. No he aprendido más de lo que esperaba.
31. Comentarios:

# Encuesta final curso 2008/2009

## *Instrucciones*

Con esta encuesta el equipo que ha creado y mantenido este ambiente de trabajo colaborativo quiere conocer tu opinión sobre el espacio virtual, el método seguido en la práctica, el equipo docente y la colaboración de los compañeros. El cuestionario tiene dos objetivos principales. Por un lado, encontrar los puntos susceptibles de mejora en el espacio virtual y el método de realización. Por otro lado, encontrar una forma de calcular la colaboración de los alumnos, para que el equipo docente pueda ir anticipándose ante los posibles problemas. Ambos objetivos son muy importantes, por lo que solicitamos que respondas con objetividad e independencia. Por ello, te pedimos que respondas a ella con la mayor sinceridad posible. Te aconsejamos que leas con detenimiento las preguntas y sus respectivas opciones de respuesta antes de responder.

## *Datos personales*

1. ¿En qué tipo de modalidad has hecho la práctica?
  - a. Virtual
  - b. Normal
2. Si has elegido la modalidad normal, elige el motivo o motivos que te han conducido a elegirla.
  - c. Porque mis circunstancias (trabajo, familia, etc.) me iban a absorber demasiado tiempo durante la fase de colaboración.
  - d. Para tener más libertad a la hora de planificar mi estudio.
  - e. Porque no me enteré de que podría elegir entre las dos modalidades hasta que fue demasiado tarde.
  - f. Porque sabía que el trabajo en equipo iba a causar demasiados problemas.
  - g. Por otros motivos
3. Si has elegido la modalidad virtual, elige el motivo o motivos que te han conducido a elegirla.
  - h. Sólo por las ventajas a la hora de hacer el examen.
  - i. Sólo porque sabía que con el trabajo en equipo iba a aprender más.
  - j. Por las ventajas en el examen y por trabajar en equipo.
  - k. Porque gente que conocía la había elegido.
  - l. Por otros motivos
4. Si has elegido la modalidad virtual, ¿cuál es el número asignado a tu equipo?
5. ¿Por qué estudias en la UNED?
  - m. No me llegó la nota de selectividad o del curso de acceso para ir a otra universidad.
  - n. Es la universidad más cercana a donde yo vivo.
  - o. Es la que mejor se adapta a mis circunstancias de tiempo y esfuerzo (trabajo, familia, etc.).
  - p. Es una de las mejores universidades que conozco.
  - q. Es muy barata.
  - r. Otras razones
6. Si has elegido otras razones en la pregunta anterior, especifica cuáles:

## *Evaluación de la práctica*

7. En general, evalúa la práctica entre 0 –puntuación más baja- y 5-puntuación más alta:
  - a. 5
  - b. 4
  - c. 3
  - d. 2
  - e. 1
  - f. 0
8. ¿Cuál o cuáles son los principales problemas que has encontrado? (Múltiple)
  - g. La instalación.
  - h. Problemas por la escasa documentación.
  - i. Problemas de desarrollo con los planificadores.
  - j. No entendía bien los enunciados y no sabía que había que hacer.
  - k. Los problemas causados por los compañeros.
  - l. El poco caso del profesor.
  - m. La plataforma.
  - n. Por mis circunstancias no tenía mucho tiempo para hacerla.
  - o. El tiempo que he pasado haciendo la práctica.
  - p. Otros problemas.
  - q. No he tenido ningún problema de consideración.
9. ¿Cuál o cuáles han sido las principales ventajas que has encontrado? (Múltiple)
  - r. La instalación.
  - s. La documentación.
  - t. El desarrollo con los planificadores.
  - u. La claridad de los enunciados de los problemas.
  - v. La ayuda de los compañeros.

- w. La ayuda del profesor.
  - x. La plataforma.
  - y. La facilidad para organizar mi trabajo sin problemas.
  - z. La cantidad de cosas aprendidas respecto al tiempo empleado.
  - aa. Otras ventajas.
  - bb. No ha habido ninguna ventaja reseñable.
10. Evalúa el trabajo del equipo docente a nivel general (0 –puntuación más baja- y 5-puntuación más alta):
- cc. 5
  - dd. 4
  - ee. 3
  - ff. 2
  - gg. 1
  - hh. 0
11. Evalúa la ayuda que te han proporcionado los foros del espacio virtual principal, no en el espacio del equipo (0 –puntuación más baja- y 5-puntuación más alta):
- ii. 5
  - jj. 4
  - kk. 3
  - ll. 2
  - mm. 1
  - nn. 0
12. Evalúa la ayuda obtenida de los compañeros en el espacio virtual principal , no en el espacio del equipo (0 –puntuación más baja- y 5-puntuación más alta):
- oo. 5
  - pp. 4
  - qq. 3
  - rr. 2
  - ss. 1
  - tt. 0

*Evalúa la colaboración y la motivación*

13. ¿Ha cambiado tu grado de motivación por hacer la práctica en un ambiente colaborativo?
- a. Sí
  - b. No
  - c. No sé
14. En caso de contestar que SÍ a la pregunta anterior, elige cuál de los siguientes aspectos ha contribuido a incrementar tu motivación por hacer la práctica: (Múltiple)
- d. El contacto con mis compañeros
  - e. El trabajo del equipo docente.
  - f. Lo interesante de los problemas propuestos.
  - g. El estudio de la asignatura.
  - h. La cercanía de las fechas de los exámenes.
  - i. Otros aspectos.
15. En caso de que hayas contestado que NO a la pregunta 13, elige cuál de los siguientes aspectos ha contribuido a disminuir tu motivación por hacer la práctica: (Múltiple)
- j. El contacto con mis compañeros.
  - k. El trabajo del equipo docente.
  - l. Lo poco interesante de los problemas propuestos.
  - m. El estudio de la asignatura.
  - n. La cercanía de las fechas de los exámenes.
  - o. Otros aspectos.
16. A continuación identifica a algunos compañeros del espacio virtual que te hayan ayudado de forma importante por lo que podrías definirlos como prototipos de la colaboración (identificalos con nombre y algún apellido y separa la lista por comas):

*Preguntas para que SÓLO las contesten los alumnos de la modalidad virtual*

17. Define la estructura de tu equipo de trabajo en el área virtual colaborativa:
- a. En mi grupo ha habido un líder claro.
  - b. En mi grupo ha habido un líder claro, pero ha ido cambiando de persona según las circunstancias.
  - c. En mi grupo cada uno adoptó un rol fijo (de comportamiento o trabajo) al principio del curso.
  - d. En mi grupo cada miembro adoptó distintos roles, según las circunstancias.
  - e. En mi grupo todas las decisiones han sido tomadas en grupo después de un periodo de diálogo.
  - f. Mi grupo ha sido bastante caótico.
  - g. No hay ninguna definición que explique adecuadamente mi grupo.
18. ¿Has quedado con tus compañeros de equipo alguna vez para tratar los problemas de la práctica?
- h. Sí, la comunicación cara a cara ha sido el principal medio de comunicación
  - i. Sí, la comunicación cara a cara ha sido un medio de comunicación del equipo pero no el principal
  - j. Sí, pero hemos quedado muy pocas veces
  - k. No, nunca hemos quedado
19. ¿Ha utilizado tu equipo otros medios de comunicación además de los ofrecidos en este espacio virtual (foros,

Chat)?

- l. Sí, y los hemos utilizado mucho más que los del espacio virtual
  - m. Sí, y los hemos utilizado algo más que los del espacio virtual
  - n. Sí, pero los hemos utilizado algo menos que los del espacio virtual
  - o. Sí, pero los hemos utilizado mucho menos que los del espacio virtual
20. ¿Qué otros medio de comunicación habéis utilizado además de los foros y el Chat del espacio virtual (teléfono, correo, mensajería instantánea, cara a cara, etc.)? Por favor escribe cuales en orden de importancia. El primero el más utilizado y así sucesivamente:
21. ¿Has tenido problemas de coordinación con los compañeros del grupo?
- p. Sí, y han producido un gran perjuicio a la coordinación del grupo.
  - q. Sí, pero apenas ha habido perjuicio.
  - r. No ha habido ningún problema reseñable.
22. Si ha habido algún problema de coordinación, comenta según tu criterio los que hubo. Si puedes ordénalos según la tarea que estabais realizando:
23. Durante esta experiencia educativa nosotros hemos investigado una forma de evaluar la colaboración de vosotros, los alumnos. Para terminar la investigación necesitamos que digas el nivel de colaboración de tus compañeros. Por favor, contesta con sinceridad. A continuación escribe el nombre de tus compañeros y a continuación el nivel de colaboración que crees que ha demostrado (nivel alto, medio o bajo):
24. De lo que has aprendido con esta experiencia, ¿qué es lo que ha superado más tus expectativas? (Múltiple)
- s. He aprendido más de lo esperado en planificación como área de conocimiento.
  - t. He aprendido más de lo esperado en el uso y programación de planificadores.
  - u. He aprendido más de lo esperado en la gestión del trabajo en equipo a distancia.
  - v. He aprendido más de lo esperado en la realización de experimentos.
  - w. He aprendido más de lo esperado en la redacción de un documento explicativo del trabajo realizado.
  - x. He aprendido más de lo esperado en la discusión sobre un tema desde un punto de vista científico.
  - y. No he aprendido más de lo que esperaba.

#### *Comentarios*

1. Comentarios a la encuesta:

# Encuesta final sobre la aplicación web del curso

2008/2009

## *Instrucciones*

A este equipo se le habilitó una aplicación web, enlazada en el portlet Información relativa a la colaboración, que consistía en datos de algunos alumnos por la que se podía navegar. La información se actualizaba cada lunes dando indicadores estadísticos de las interacciones en los foros de los equipos. Además, se ha suministrado indicaciones del nivel de la colaboración de algunos alumnos. Con esta encuesta queremos conocer tu opinión sobre la aplicación. Por ello, te pedimos que respondas con la mayor sinceridad posible. Asimismo, te aconsejamos que, antes de responder, leas con detenimiento las preguntas y sus respectivas opciones de respuesta.

1. ¿Has entrado alguna vez en la aplicación web?
  - a. Varias veces por semana
  - b. Al menos una vez por semana
  - c. Cada dos semanas, más o menos
  - d. Apenas nunca
  - e. Nunca
2. ¿Cuándo has entrado más en la aplicación?
  - f. Al principio de la experiencia colaborativa
  - g. Hacia la mitad de la experiencia colaborativa
  - h. Hacia el final de la experiencia colaborativa
  - i. No sé
3. ¿Has entendido fácilmente la información que se muestra en la aplicación?
  - j. Sí, mucho
  - k. Sí, pero sólo algo
  - l. Apenas
  - m. No, no la he entendido
4. ¿Te ha sido sencillo navegar por la aplicación?
  - n. Sí, mucho
  - o. Sí, pero sólo algo
  - p. Me ha resultado difícil
  - q. Me ha resultado muy difícil
5. ¿La información que se mostraba te ha ayudado a colaborar mejor con tus compañeros?
  - r. Sí, mucho
  - s. Sí, algo
  - t. Apenas
  - u. No, nada
6. ¿La información que se mostraba te ha ayudado a saber como colaboraban tus compañeros?
  - v. Sí, mucho
  - w. Sí, algo
  - x. Apenas
  - y. No, nada
7. ¿Crees que ha sido positivo que la información presentada sobre ti mismo fuera conocida por tus compañeros?
  - z. Sí, mucho
  - aa. Sí, algo
  - bb. Apenas
  - cc. No, nada
8. Cada lunes se actualizaban las estadísticas de interacción en los foros del equipo. ¿Crees que ha sido útil el hecho de se te suministrara esa información?
  - dd. Sí, muy útil.
  - ee. Sí, ha sido útil.
  - ff. No, no ha sido útil.
  - gg. No ha sido para nada útil.
9. Además de las estadísticas, se ha proporcionado información sobre el nivel de colaboración (nivel alto, medio o bajo). ¿Te ha sido útil conocer esta información?
  - hh. Sí, muy útil.
  - ii. Sí, ha sido útil.
  - jj. No, no ha sido útil.
  - kk. No ha sido para nada útil.
10. Selecciona lo que has podido echar en falta en la aplicación web:
  - ll. Que se hubiera habilitado antes, al comienzo del espacio virtual a finales de octubre.
  - mm. Que aparecieran sólo los compañeros de un mismo equipo
  - nn. Algún otro indicador estadístico de mis interacciones.
  - oo. Algún otro indicador de la colaboración.
  - pp. No he echado en falta nada.
11. ¿Crees que la aplicación web funcionaría para todos los miembros del espacio virtual?
  - qq. Sí, debería ser una herramienta que mostrara la información de todos los miembros del espacio virtual.
  - rr. Sí, pero sólo con grupos de personas reducidos como el caso de mi equipo.

- ss. No, no funcionaría en ningún caso.
12. ¿Añadirías alguna información más a la que ya se? En caso afirmativo, señala en el espacio que se te proporciona lo que añadirías:
13. ¿Te ha facilitado la herramienta el aprendizaje colaborativo y debido a ella has incrementado tu interacción?
- tt. Sí, mucho.  
uu. Sí, un poco.  
vv. Ha sido un poco el efecto contrario  
ww. Totalmente opuesto.
14. Comentarios:

## Encuesta final sobre la herramienta de indicadores estadísticos del curso 2008/2009

### *Instrucciones*

A este equipo se le habilitó un portlet en la sección Información del subgrupo que muestra indicadores estadísticos de las interacciones en los foros. La información se actualizaba cada lunes. Con esta encuesta queremos conocer tu opinión. Por ello, te pedimos que respondas con la mayor sinceridad posible. Asimismo, te aconsejamos que, antes de responder, leas con detenimiento las preguntas y sus respectivas opciones de respuesta.

1. En general, evalúa la información presentada práctica entre 0 –puntuación más baja- y 5-puntuación más alta:
- a. 5  
b. 4  
c. 3  
d. 2  
e. 1  
f. 0
2. ¿La información que se mostraba te ha ayudado a colaborar mejor con tus compañeros?
- g. Sí, mucho  
h. Sí, algo  
i. Apenas  
j. No, nada
3. ¿La información que se mostraba te ha ayudado a saber como colaboraban tus compañeros?
- k. Sí, mucho  
l. Sí, algo  
m. Apenas  
n. No, nada
4. Cada lunes se actualizaban las estadísticas de interacción en los foros del equipo. ¿Crees que ha sido útil el hecho de se te suministrara esa información?
- o. Sí, muy útil.  
p. Sí, ha sido útil.  
q. No, no ha sido útil.  
r. No ha sido para nada útil.
5. ¿Añadirías alguna información más a la que ya se? En caso afirmativo, señala en el espacio que se te proporciona lo que añadirías:
6. ¿Te ha facilitado la herramienta el aprendizaje colaborativo y debido a ella has incrementado tu interacción?
- s. Sí, mucho.  
t. Sí, un poco.  
u. Ha sido un poco el efecto contrario  
v. Totalmente opuesto.
7. Comentarios:

# Encuesta final sobre la herramienta que mostraba los niveles de colaboración en el curso 2008/2009

## *Instrucciones*

A este equipo se le habilitó un portlet en la sección Información del subgrupo que muestra los niveles de colaboración de los miembros. La información se actualizaba cada lunes. Con esta encuesta queremos conocer tu opinión. Por ello, te pedimos que respondas con la mayor sinceridad posible. Asimismo, te aconsejamos que, antes de responder, leas con detenimiento las preguntas y sus respectivas opciones de respuesta.

1. En general, evalúa la información presentada práctica entre 0 –puntuación más baja- y 5-puntuación más alta:
  - a. 5
  - b. 4
  - c. 3
  - d. 2
  - e. 1
  - f. 0
2. ¿La información que se mostraba te ha ayudado a colaborar mejor con tus compañeros?
  - g. Sí, mucho
  - h. Sí, algo
  - i. Apenas
  - j. No, nada
3. ¿La información que se mostraba te ha ayudado a saber como colaboraban tus compañeros?
  - k. Sí, mucho
  - l. Sí, algo
  - m. Apenas
  - n. No, nada
4. ¿Te ha sido útil conocer el nivel de colaboración?
  - o. Sí, muy útil.
  - p. Sí, ha sido útil.
  - q. No, no ha sido útil.
  - r. No ha sido para nada útil.
5. ¿Crees que los niveles de colaboración que se mostraban eran correctos?
  - s. Sí, muy correctos
  - t. Sí, pero aproximados
  - u. No, eran poco aproximados
  - v. No, eran incorrectos
6. ¿Añadirías alguna información más a la que ya se? En caso afirmativo, señala en el espacio que se te proporciona lo que añadirías:
7. ¿Te ha facilitado la herramienta el aprendizaje colaborativo y debido a ella has incremento tu interacción?
  - w. Sí, mucho.
  - x. Sí, un poco.
  - y. Ha sido un poco el efecto contrario
  - z. Totalmente opuesto.
8. Comentarios:



# Encuesta final sobre la herramienta que mostraba los indicadores estadísticos y los niveles de colaboración en el curso 2008/2009

## *Instrucciones*

A este equipo se le habilitó un portlet en la sección Información del subgrupo que muestra indicadores estadísticos de las interacciones en los foros. La información se actualizaba cada lunes. Además, se ha suministrado indicaciones del nivel de la colaboración de los miembros del equipo. Con esta encuesta queremos conocer tu opinión. Por ello, te pedimos que respondas con la mayor sinceridad posible. Asimismo, te aconsejamos que, antes de responder, leas con detenimiento las preguntas y sus respectivas opciones de respuesta.

1. En general, evalúa la información presentada práctica entre 0 –puntuación más baja- y 5-puntuación más alta:
  - a. 5
  - b. 4
  - c. 3
  - d. 2
  - e. 1
  - f. 0
2. ¿La información que se mostraba te ha ayudado a colaborar mejor con tus compañeros?
  - g. Sí, mucho
  - h. Sí, algo
  - i. Apenas
  - j. No, nada
3. ¿La información que se mostraba te ha ayudado a saber como colaboraban tus compañeros?
  - k. Sí, mucho
  - l. Sí, algo
  - m. Apenas
  - n. No, nada
4. Cada lunes se actualizaban las estadísticas de interacción en los foros del equipo. ¿Crees que ha sido útil el hecho de se te suministrara esa información?
  - o. Sí, muy útil.
  - p. Sí, ha sido útil.
  - q. No, no ha sido útil.
  - r. No ha sido para nada útil.
5. Además de las estadísticas, se ha proporcionado información sobre el nivel de colaboración (nivel alto, medio o bajo). ¿Te ha sido útil conocer esta información?
  - s. Sí, muy útil.
  - t. Sí, ha sido útil.
  - u. No, no ha sido útil.
  - v. No ha sido para nada útil.
6. ¿Crees que los niveles de colaboración que se mostraban eran correctos?
  - w. Sí, muy correctos
  - x. Sí, pero aproximados
  - y. No, eran poco aproximados
  - z. No, eran incorrectos
7. ¿Añadirías alguna información más a la que ya se? En caso afirmativo, señala en el espacio que se te proporciona lo que añadirías:
8. ¿Te ha facilitado la herramienta el aprendizaje colaborativo y debido a ella has incrementado tu interacción?
  - aa. Sí, mucho.
  - bb. Sí, un poco.
  - cc. Ha sido un poco el efecto contrario
  - dd. Totalmente opuesto.
9. Comentarios:

## Apéndice II: ontología de la colaboración

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns="http://www.semanticweb.org/ontologies/2008/8/Ontology1220535799039.owl#"
  xmlns:owl2xml="http://www.w3.org/2006/12/owl2-xml#"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"

xml:base="http://www.semanticweb.org/ontologies/2008/8/Ontology1220535799039.owl">
  <owl:Ontology rdf:about="" />
  <owl:Class rdf:ID="Preferencias">
    <rdfs:subClassOf>
      <owl:Class rdf:ID="Datos_Estaticos" />
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:ID="Foros">
    <rdfs:subClassOf>
      <owl:Class rdf:ID="Datos_Interaccion" />
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:ID="Colaboracion">
    <rdfs:subClassOf>
      <owl:Class rdf:ID="Datos_Dinamicos" />
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:ID="Persona" />
  <owl:Class rdf:about="#Datos_Estaticos">
    <rdfs:subClassOf rdf:resource="#Persona" />
  </owl:Class>
  <owl:Class rdf:ID="Equipo" />
  <owl:Class rdf:ID="Chat">
    <rdfs:subClassOf>
      <owl:Class rdf:about="#Datos_Interaccion" />
    </rdfs:subClassOf>
  </owl:Class>
  <owl:Class rdf:ID="Xsemana">
    <rdfs:subClassOf rdf:resource="#Chat" />
  </owl:Class>
  <owl:Class rdf:about="#Datos_Dinamicos">
    <rdfs:subClassOf rdf:resource="#Persona" />
  </owl:Class>
  <owl:Class rdf:ID="Laborales">
    <rdfs:subClassOf rdf:resource="#Datos_Estaticos" />
  </owl:Class>
  <owl:Class rdf:ID="Usuario">
    <rdfs:subClassOf rdf:resource="#Persona" />
  </owl:Class>
  <owl:Class rdf:ID="Personales">
    <rdfs:subClassOf rdf:resource="#Datos_Estaticos" />
  </owl:Class>
```

```

</owl:Class>
<owl:Class rdf:ID="Semana">
  <rdfs:subClassOf rdf:resource="#Foros" />
</owl:Class>
<owl:Class rdf:ID="Total">
  <rdfs:subClassOf rdf:resource="#Foros" />
</owl:Class>
<owl:Class rdf:ID="Academicos">
  <rdfs:subClassOf rdf:resource="#Datos_Estaticos" />
</owl:Class>
<owl:Class rdf:about="#Datos_Interaccion">
  <rdfs:subClassOf rdf:resource="#Persona" />
</owl:Class>
<owl:Class rdf:ID="total">
  <rdfs:subClassOf rdf:resource="#Chat" />
</owl:Class>
<owl:ObjectProperty rdf:ID="ObjectProperty_6" />
<owl:ObjectProperty rdf:ID="interac_foro_semana">
  <rdfs:domain rdf:resource="#Usuario" />
  <rdfs:range rdf:resource="#Semana" />
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="es_tutor">
  <owl:inverseOf>
    <owl:ObjectProperty rdf:ID="tutor" />
  </owl:inverseOf>
  <rdfs:range rdf:resource="#Equipo" />
  <rdfs:domain rdf:resource="#Usuario" />
</owl:ObjectProperty>
<owl:ObjectProperty rdf:about="#tutor">
  <rdfs:domain rdf:resource="#Equipo" />
  <owl:inverseOf rdf:resource="#es_tutor" />
  <rdfs:range rdf:resource="#Usuario" />
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="interac_usuario">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Total" />
        <owl:Class rdf:about="#total" />
        <owl:Class rdf:about="#Semana" />
        <owl:Class rdf:about="#Xsemana" />
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="#Usuario" />
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="interac_chat_semana">
  <rdfs:range rdf:resource="#Xsemana" />
  <rdfs:domain rdf:resource="#Usuario" />
</owl:ObjectProperty>
<owl:DatatypeProperty rdf:ID="sueldo">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string" />
  <rdfs:domain rdf:resource="#Laborales" />
</owl:DatatypeProperty>

```

```

<owl:DatatypeProperty rdf:ID="periodo-conexion">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdfs:domain rdf:resource="#Personales"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="asignaturas">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdfs:domain rdf:resource="#Academicos"/>
</owl:DatatypeProperty>
<owl:DatatypeProperty rdf:ID="trabajo_en">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdfs:domain rdf:resource="#Laborales"/>
</owl:DatatypeProperty>
<owl:FunctionalProperty rdf:ID="frecuencia_al_centro">
  <rdfs:domain rdf:resource="#Academicos"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="grado_colaboracion">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:domain rdf:resource="#Colaboracion"/>
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    >Máximo 1 y mínimo 0</rdfs:comment>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#float"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="interac_chat_total">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
  <rdfs:domain rdf:resource="#Usuario"/>
  <rdfs:range rdf:resource="#total"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="num_replicas_a_conversacion_iniciada">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Semana"/>
        <owl:Class rdf:about="#Xsemana"/>
        <owl:Class rdf:about="#total"/>
        <owl:Class rdf:about="#Total"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#int"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="owner">
  <rdfs:range rdf:resource="#Persona"/>
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Academicos"/>
        <owl:Class rdf:about="#Personales"/>
        <owl:Class rdf:about="#Laborales"/>
        <owl:Class rdf:about="#Preferencias"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>

```

```

</rdfs:domain>
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty" />
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="fecha">
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#date" />
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty" />
<rdfs:domain>
<owl:Class>
<owl:unionOf rdf:parseType="Collection">
<owl:Class rdf:about="#Semana" />
<owl:Class rdf:about="#Xsemana" />
</owl:unionOf>
</owl:Class>
</rdfs:domain>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="nivel_colaboracion">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty" />
<rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string"
>Los posibles valores son: muy colaborativo, colaborativo, poco
colaborativo</rdfs:comment>
<rdfs:domain rdf:resource="#Colaboracion" />
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string" />
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="nombre">
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string" />
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty" />
<rdfs:domain>
<owl:Class>
<owl:unionOf rdf:parseType="Collection">
<owl:Class rdf:about="#Equipo" />
<owl:Class rdf:about="#Usuario" />
</owl:unionOf>
</owl:Class>
</rdfs:domain>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="estrategia">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty" />
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string" />
<rdfs:domain rdf:resource="#Preferencias" />
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="laboral">
<rdfs:domain rdf:resource="#Usuario" />
<rdfs:range rdf:resource="#Laborales" />
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty" />
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="apellidos">
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty" />
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string" />
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="grado_rigidez">
<rdfs:domain rdf:resource="#Preferencias" />
<rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty" />
<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string" />
</owl:FunctionalProperty>

```

```

<owl:FunctionalProperty rdf:ID="sexo">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdfs:domain rdf:resource="#Personales"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="plan_estudio">
  <rdfs:domain rdf:resource="#Preferencias"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="es_alumno">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
  <rdfs:range rdf:resource="#Equipo"/>
  <owl:inverseOf>
    <owl:InverseFunctionalProperty rdf:ID="alumno"/>
  </owl:inverseOf>
  <rdfs:domain rdf:resource="#Usuario"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="preferencia">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
  <rdfs:domain rdf:resource="#Usuario"/>
  <rdfs:range rdf:resource="#Preferencias"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="email">
  <rdfs:domain rdf:resource="#Usuario"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="num_replicas_a_mensajes">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Semana"/>
        <owl:Class rdf:about="#Xsemana"/>
        <owl:Class rdf:about="#total"/>
        <owl:Class rdf:about="#Total"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#int"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="tipo_responsabilidades">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdfs:domain rdf:resource="#Personales"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="nacimiento">
  <rdfs:domain rdf:resource="#Personales"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="num_conversaciones_iniciadas">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>

```

```

<rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#int" />
<rdfs:domain>
  <owl:Class>
    <owl:unionOf rdf:parseType="Collection">
      <owl:Class rdf:about="#Semana" />
      <owl:Class rdf:about="#Xsemana" />
      <owl:Class rdf:about="#total" />
      <owl:Class rdf:about="#Total" />
    </owl:unionOf>
  </owl:Class>
</rdfs:domain>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="provincia">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Academicos" />
        <owl:Class rdf:about="#Personales" />
        <owl:Class rdf:about="#Laborales" />
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string" />
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty" />
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="colaboracion">
  <rdfs:domain rdf:resource="#Usuario" />
  <rdfs:range rdf:resource="#Colaboracion" />
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty" />
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="ciudad">
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Academicos" />
        <owl:Class rdf:about="#Personales" />
        <owl:Class rdf:about="#Laborales" />
      </owl:unionOf>
    </owl:Class>
  </rdfs:domain>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string" />
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty" />
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="interac_foro_total">
  <rdfs:range rdf:resource="#Total" />
  <rdfs:domain rdf:resource="#Usuario" />
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty" />
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="num_mensajes_enviados">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#int" />
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty" />
  <rdfs:domain>
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">

```

```

    <owl:Class rdf:about="#Semana"/>
    <owl:Class rdf:about="#Xsemana"/>
    <owl:Class rdf:about="#total"/>
    <owl:Class rdf:about="#Total"/>
  </owl:unionOf>
</owl:Class>
</rdfs:domain>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="personal">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
  <rdfs:range rdf:resource="#Personales"/>
  <rdfs:domain rdf:resource="#Usuario"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="jornada">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
  <rdfs:domain rdf:resource="#Laborales"/>
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="academico">
  <rdfs:domain rdf:resource="#Usuario"/>
  <rdfs:range rdf:resource="#Academicos"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
</owl:FunctionalProperty>
<owl:FunctionalProperty rdf:ID="cambios">
  <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
  <rdfs:domain rdf:resource="#Preferencias"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#DatatypeProperty"/>
</owl:FunctionalProperty>
<owl:InverseFunctionalProperty rdf:about="#alumno">
  <owl:inverseOf rdf:resource="#es_alumno"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty"/>
  <rdfs:range rdf:resource="#Usuario"/>
  <rdfs:domain rdf:resource="#Equipo"/>
</owl:InverseFunctionalProperty>
</rdf:RDF>

```

<!-- Created with Protege (with OWL Plugin 3.3.1, Build 430) <http://protege.stanford.edu> -->



## **Apéndice III: Summary and Conclusions**

### **Summary**

The Distance Education (DE) is an education model has a number of relevant features, such as breaking the barriers of time and place. The teacher can instruct the student without having to be both on the same site and at the same time. This advantage brings the disadvantage that the communication between them has to use other communication means less rich and intense than face to face. The great advantage of the ED is paradoxically also the main constraint (Garcia Aretio, 2009).

The motivation of this thesis has been the improvement of learning in an environment of DE over the Internet through the strategy of collaborative learning. Although the advantages seem obvious use of collaboration in distance education, as noted, a study of collaboration is necessary to try to ensure that collaborative learning is performed.

The design of web collaborative learning environments must primarily ensure collaborative learning that is the same or better than individual or competitive learning. According to Johnson & Johnson (1989), offering students in these environments a set of communication tools and a set of tasks to complete is not enough. These authors establish five conditions to make collaborative learning better than individual or competitive learning (Johnson & Johnson, 2004): (1) positive interdependence (everyone shares the goals), (2) individual accountability/personal responsibility (everyone is in charge of oneself), (3) promote interaction (mainly face to face interaction), (4) interpersonal and small group skills (everyone works effectively with each other and functions as part of a group), and (5) frequent and regular processing of the group's functioning to improve its effectiveness in the future should all be clearly perceived.

In e-learning environments, the use of collaboration strategy is highly recommendable to minimize disadvantages inherent to distance learning. In this context communication between students can be asynchronous, where students are free to perform communication acts when they want (Santos & Boticario, 2004b), i.e. they control the collaboration process. As for managing collaborative experiences in

these contexts, the tutor has a very complicated task if s/he is in charge of analyzing regular and frequent collaboration, especially if the courses are distant teaching and involve over a hundred students.

Given the previous conditions, the design of the collaborative learning environment should therefore use a method that analyses student collaboration regularly and frequently with little or no intervention by the tutor of the learning experience. Learner modeling techniques are appropriate in these contexts because they have been used to analyze the learners' behavior with the objectives to model the user to use this model so that the system or the learner improves the learning, in an educational environment (Brusilovsky & Millan, 2007). So that the modeling techniques are automatic or they can be used with little or no intervention by the tutor of the learning experience, we can require data mining methodology (Romero & Ventura, 2007) and use machine learning technology (Russell & Norvig, 1995).

Motivation and problem definition of this thesis have addressed the same objectives, in addition to the assumptions used for solving the problem defined. To carry out research have postulated a set of hypotheses:

1. it is possible to characterize the collaboration in open environments using technology instead of an expert analysis,
2. increasing control over the collaborative process encourages collaboration and improves the management process itself,
3. increasing control to be effective must be done while the process of collaboration is taking place, and
4. an analysis of collaboration without domain knowledge is possible and which allows repeated in other educational settings.

The hypotheses require evaluation of research in a real collaborative learning and establish a set of objectives. The overall objectives were:

1. characterize their students on their collaboration,
2. use a method of analysis to get results on a regular and frequent
3. characterization to provide students and assess their learning, and
4. offer a general solution to the work of this thesis are applicable in other collaborative learning environments.

## Framework

The work of improving the collaborative learning modeling the learners' collaboration is a research, which can be divided in (Romero and Ventura, 2007; Kobsa, 2001; Soller et al., 2005): the data acquisition methods, the inference method, the model and the process that they apply to their results and model.

Data acquisition methods are: a) Qualitative (Meier et al., 2006), where the individuals participating in the research are directly questioned, or experts assess participants' activities; b) Quantitative (Talavera & Gaudioso, 2004; Redondo et al., 2003; Hong, 2001; Bratitsis et al., 2008), where statistical information on participants' activities is collected; and c) Mixed (Collazos et al., 2007; Daradoumis et al., 2006; Martínez et al., 2006; Perera et al., 2007), the use of both methods simultaneously.

Systems can be characterized according to the inference method used to deduce the value of certain characteristics, such as collaboration, which has occurred or is occurring. The methods may be: 1) an expert's analysis (Meier et al., 2006); 2) comparison with a pre-existing model (Redondo et al., 2003); 3) different statistical or interaction analysis techniques (Hong, 2001; Daradoumis et al., 2006; Martínez et al., 2006; Bratitsis et al., 2008) ML techniques, like decision trees, clustering, sequence pattern mining, etc., (Talavera & Gaudioso, 2004; Redondo et al., 2003; Perera et al., 2007); 4) even characterizing systems by not using any inference system (Collazos et al. 2007;).

After the inferring method and before of using the results of the method, some researches have model the learners' collaboration. Two main strategies have been identified. There has been research that has proposed a model for collaboration (Redondo et al., 2003; Baghaei & Mitrovic, 2007; Martínez et al., 2003; Vidou et al., 2006; Barros et al., 2001; Barros et al., 2002; Durán, 2006), while others have proposed a number of indicators related to collaboration (Collazos et al., 2007; Baldiris et al., 2007; Soller, 2001; Park & Hyun, 2006; Baeza-Yates & Pino, 2006; Meier et al., 2007; Kahrimanis et al., 2009; Bayón et al., 2007; Martínez et al., 2006; Daradoumis et al., 2006; Bratitsis et al., 2008). As many have noted (Park and Hyun, 2006; Strijbos and Fischer, 2007), analysis or modeling of the collaboration is a field still open and it needs a deep work to standardize the field and build methodologies.

It has been said that to ensure that collaborative learning takes place, it is necessary to analyze student collaboration regularly and frequently. In order to achieve this, an analysis method and a tool that uses the inferred information have to be planned. In collaborative environments these tools can be classified according to their function (Soller et al., 2005): I) Monitoring tools that collect data on student interaction automatically and show this information (Meier et al., 2007; Kahrimanis et al., 2009; Collazos et al., 2007; Daradoumis et al., 2006; Martínez et al., 2006; Bratitsis et al., 2008); II) Metacognitive tools that show, as well as the interactions, the information inferred from processing the collaboration analysis, i.e., they propose judgments (Redondo et al., 2003; Talavera & Gaudioso, 2004; Perera et al., 2007); III) Guide tools, which propose corrective measures to help the student, once the right information has been inferred (Baghaei & Mitrovic, 2007; Santos & Boticario, 2008).

In other words, there are tools (I) whose aim is to show the data on students, so the priority is to discover how these data are acquired, filter them and display them. For other tools (II) the priority is to infer useful information from the data collected, and others (III) guide or recommend students thanks to the results inferred.

A useful strategy on the educational field is the self-regulation. A self-regulation tool uses metacognitive techniques to discover what has been learnt (Hartman, 2001). It has been shown that helping students use their self-regulation skills improves their learning (Pintrich, 2000) and dividing self-regulation into three characteristics or tasks (Zimmerman, 1990) has even been proposed: i) self-observation (tracking of own activities), j) self-judgment (self-assessment of performance) and, k) self-reactions (reactions to performance results).

Moreover, (Stiffens, 2001) establishes that self-regulation is more than the regulation of own cognitive activities (metacognition), since it also involves motivational and emotional aspects. It is obvious, according to the "self-regulation handbook" (Beakers et al., 2000), that self-regulation is not only used in monitoring learning processes of oneself, but it also plays an important role in managing social activities.

From what has been said about self-regulation it is deduced that a metacognitive tool fulfils the conditions to be considered as a self-regulation tool. A metacognitive tool monitors students, provides them with an independent judgment that may change

according to their reactions, so they can assess themselves and be aware of their reactions.

Studying the literature on the topic we have observed that two types of tools behave like self-regulation tools: 1) those that show the information to students, and 2) those that allow students to manage their own data. These second tools are known as scrutable (Kay, 1999).

Some research studies pose scrutable systems to improve learning. They have noticed improved understanding and control of learning in educational environments that use scrutability and open models (Bull & Gardner, 2009), which are student models that can be accessed and managed by the student (Bull & Kay, 2007). Scrutable systems are gaining interest in educational environments, as some research studies show (Verpoorten et al., 2009).

### **Critical Analysis**

When establishing a framework for the improvement of collaborative learning, we have studied the literature that addresses the problem or any of its parts. With the above, the design of the collaborative learning environment must be considered analyzing the behavior of participants on a regular and frequent with less intervention of tutors, because of the context of the DE. The characterizations of the researches in the field can be divided into: how to obtain data (preprocessing), which method of analysis used (inference), what type of modeling is applied (modeling), what type of tool used (tool).

Due to the lack of methodology, standards and empirical studies (Strijbos & Fischer, 2007, Perera et al., 2007; Bratitsis & Dimitracopoulou, 2006), there are several open issues in the analysis of collaboration. The first is the data source. Due to the educational context and objectives of the investigation, it was thought that thing was used as a source of data, quantitative information about the interactions of students relating to collaboration. In this way promotes regular and frequent collection of data, and the lowest possible intervention of tutors and students.

Another open issue in the analysis of collaboration is the method of inference. Since the analysis should be performed regularly and frequently with the least

possible intervention of tutors or experts, it is necessary to use machine learning techniques.

The analysis of collaboration has been the main part of the modeling of the student, but not the only one. Modeling the student is a major challenge for the personalization of learning environments (McCalla et al., 2000). A very interesting strategy from the pedagogical point of view is the open model (Bull and Kay, 2008), where the result of modeling is shown to the student and where he can manage. Therefore the model obtained from the modeling process should be understandable by the student.

In collaborative environments thus information on the process and in the context of cooperation are of interest (Muehlenbrock, 2005). We conclude that useful information for the student to work is due to the context of collaboration, the process of collaboration, and inferred by the analysis as it provides judgments about student collaboration.

To close the cycle of improvement of collaborative learning, we have discussed various pedagogical tools used. We have focused on the benefits of self-regulatory tools in the educational context and are associated with collaboration, as the social processes involved in self-regulation (Boekaerts et al., 2000). Characteristics have been established self-regulatory tools and compared with metacognitive tools that can provide a collaborative environment. We have opted in this investigation by a self-regulatory metacognitive tool.

Following the thread of the models open to students, we have studied the tellers (Kay, 1999). In defining the searching has found affinity with the searching tools and metacognitive self-regulation. It has been reported the theoretical advantages of the teller to increase accountability and, therefore, motivation and active learning (Hummel et al., 2005, Burlison, 2005; Boticario and Gaudioso, 2000a).

In summary, we conclude that a collaborative learning environment via the web within the paradigm of the ED requires the system to perform analysis of collaboration on a regular and frequent. Through the analysis can be modeled to students. Once the modeling result, ie, the model can become a self-regulatory metacognitive tool with which to provide the student with a means of helping manage the collaborative process. This tool can add the benefits of teller to increase student responsibility in the entire collaborative process.

## Research

This research was done at UNED (National University for Distance Education, Spain). UNED students are adults with responsibilities other than learning. Thus, it is difficult for them to participate in collaborative environments with the typical time restrictions (Santos et al., 2003), since they must control their learning processes (Gaudioso et al., 2003). In order to improve the learning of students on the fourth year Artificial Intelligence and Knowledge-based Engineering (AI-KE) subject at UNED Computer Engineering School, during the academic years 2006-07, 2007-08 and 2008-09 they were offered a long-term collaborative learning experience where collaboration was controlled and managed by the students. Over 100 students participated in each collaborative learning experience over the three consecutive years.

Given the research aims and the educational context, ML techniques are advisable to analyze student interaction. With this approach one difficulty is the lack of methodology and standards to analyze collaboration (Strijbos & Fisher, 2007), as well as not being able to identify beforehand the most appropriate ML technique for the problem. Therefore, an empirical study is necessary. This work presents two approaches to analyze collaboration regularly and frequently using different ML techniques; the results are analyzed and compared.

The two approaches use student interactions in forums as a source of data, like other related research studies (Bratitsis et al., 2008; Gómez-Sánchez et al. 2009), and their main difference is: 1) classifying students according to their collaboration using unsupervised classification techniques (clustering) (Anaya & Boticario, 2009a); 2) constructing collaboration metrics, using supervised classification techniques (decision tree algorithms), which assign each student with a collaboration value so that students can be compared. Using the statistics of forum interactions is justified because forums are a communication service widely used in e-learning environments, because they were the main means of communication in our collaborative learning experience, and because these statistics are an effective way of obtaining information on forum users, such as their level of activity or frequency of use (Dringus & Ellis, 2005).

After the collaboration analysis a model of collaboration has been built. The model structures information on collaboration and it is divided in: information on the collaboration context (personal data, academic data, work data, collaboration facilities data), information on the collaboration process (statistical indicators of learners' interactions in forums) and, evaluations on learners' collaboration (obtained by inferring method). The model uses the scrutability strategy and shows the data so that learners can understand and use to improve the collaboration learning (Bull & Kay, 2008).

Four tools were provided during the academic year 2008-09, because we wanted to research what type of information was the most useful to improve the collaboration process, and how to present it. Two presentation approaches were used. On the one hand, the simplest way for students to use and understand information (Barkley et al., 2004) and, on the other, the scrutable strategy was used.

Once the collaborative learning experience had finished, the tools were evaluated bearing in mind the students' opinions, the collaborative work rating and the learning rating. The results show that the information from the interaction analysis is useful to improve collaboration, but the most useful information is from the inference method. The results also showed that, in this educational context, the tools that only show or display the information collected or inferred help more than the tool that shows it scrutably.

## **Results**

The results obtained with the data from the three collaborative learning experiences show that students can be classified according to their collaboration using clustering techniques, although their inferences are approximate.

They also show that collaboration metrics can be constructed to give a value to each student's collaboration. The collaboration results are also approximate in this second approach.

The research ended by comparing both approaches. Accordingly, two variables were proposed: the difference ( $\Delta$ ) of average values of the metrics between the different collaboration levels and error (E%) as the average percentage of variance for the metric value in each instance. Obviously, as the difference ( $\Delta$ ) is greater between



levels it is assumed that the approach could better distinguish student collaboration, and less error (E%) would make the result more certain. Comparative analysis has told us that the approach using clustering techniques provides inference with less accuracy than the metric approach.

During the 2008-09 academic year experience the interactions were analyzed and the inference method was done that provided information on student collaboration. Four tools were offered that provided the former information to check: 1) the type of information on the most useful collaboration, and 2) the type of strategy for presenting the information. Accordingly, a tool was designed as a web application that showed all the information on the collaboration scrutably (Kay, 1999), and three portlets were created, small windows in the platform, which showed information on the collaboration process (Portlet I), on student collaboration inference (Portlet II), and another that showed both types of information (Portlet III).

The results in this article indicate that: 1) the most useful information to improve collaborative learning is that inferred on student collaboration, i.e., student levels of collaboration, and 2) the students participating in the research made the most of that information that was shown in a more simple form, so the scrutable strategy, used in the web application, has not achieved the aims.

## **Conclusions**

The research achievements are divided in: the process of analysis of collaboration, the proposed model of collaboration and the tools that have been proposed to improve the collaborative process.

1. The analysis of collaboration is divided into data acquisition and method of inference. To achieve the objectives defined for this part, data acquisition method must be quantitative to get the data regularly and frequently, evaluations of collaboration are necessary to ensure that collaborative learning occurs, machine learning techniques can be used to infer assessments collaboration. A method of analysis that will follow these arguments can be automated and has great potential to be transferred to other settings. In the thesis two approaches are described which have been validated in Chapter V. Therefore, an achievement of the thesis was: to find instead of one, two inferring methods of collaboration with which

assessments of the collaboration are obtained regularly and frequently, and they can be automated and transferred to other environments.

2. There is a lack of agreement and the lack of comparative studies in the collaborative modeling field (Bratitsis and Dimitracopoulou, 2006). We have followed a strategy aimed at expanding the learners control over their processes, and use is given directly to the user. This strategy is called the open model (Bull and Kay, 2008). We have followed this approach because the information structure in a model, which can be understood as an ontology (Mizoguchi, 2005), the learners uses the model, thus the model needs to be used and understood by the learner. These ideas provide pedagogical characteristics, since its objective is to improve learning, and reusability in other environments and contexts. Both features are appropriate to the objectives of the research. We have proposed a collaborative model divided into three parts to receive background information on the process and assessments of the collaboration. The achievement has been to create a model of collaboration to report on relevant aspects of the learners' collaboration, which is understandable and easily transferable to other environments.
3. To close the cycle of research in educational settings, we studied the possible corrective actions to be undertaken to ensure that collaborative learning takes place and also providing more control over the collaboration process. Studying the problem we have noted the benefits that a self-regulatory tool provides an educational and collaborative environments (Steffens, 2001). This idea is closely related to the open model (Bull and Kay, 2008) and metacognitive tools (Soller et al., 2005). We have suggested to students four self-regulatory tools in order to check the type of information more useful and how or tools. For these reasons it has developed a tool scrutable (Kay, 1999) that displayed all the collaborative model, and three other self-regulatory tools, one monitoring tool, which informs on the process of collaboration, and two that have already been metacognitive, which report the evaluation of the collaboration. An evaluation process was conducted. The achievements of this part are: check the advantages in working as a self-regulatory metacognitive tool in collaborative learning environments.

This thesis has as general achievement, which has proposed a methodology with the motivation to improve collaborative learning increasing control over the collaborative process. The collaboration has been analyzed and modeled the

collaboration, has proposed a model which has been used as the basis for a multitude of self-regulation that have been evaluated to see if targets are met. It has been shown that a relationship exists between the work (analysis model) with the increase of collaborative learning. We must say it has succeeded in achieving the improvement of collaborative learning, where students, tutors or experts have the least intervention possible, automation has been proposed that makes possible the transfer of the method to other environments, something that has not yet been tested.

## **Future works**

The empirical study done in this research has to continue. Two ML techniques that were thought to be more appropriate for the problem were chosen. However, other ML techniques cannot be excluded, like Bagging. With this technique the statistical indicators most related to collaboration could be identified instead of using decision tree algorithms. From applying Bagging, other collaboration metrics could be constructed.

The results obtained are approximates. In both the approaches researched we have to continue to reduce the percentages of error and increase the difference between the levels of collaboration. The source of data used cannot be forgotten. Statistical indicators of students' active interactions in forums were used. The source of data could be increased by also using passive interactions (Burr & Spennemann, 2004), understood as the reading or visualisation of forum messages, so that they reinforce the collaboration related tasks. The use of some domain-related indicators can also be studied (Ravi et al., 2007; Lin et al., 2009). These indicators could add much information to the analysis, although we would have to consider that the analysis might then not be able to be transferred to other environments.

The results obtained indicate that collaboration analysis, as has been explained, provides useful information to the students on their own collaboration and that of their classmates, and students used it to improve collaboration. Learning has increased but to a less extent according to the students' subject exam results. This is a point open to research. It is very difficult to achieve results that unequivocally show the correlation between learning and access to the information provided, although the evidence obtained is encouraging. Other many dependent variables should be considered, like

for example the time devoted, which can depend on personal situations that were not controlled because of the inherent difficulty and so as not to disturb the students any further. It should be borne in mind that they are students following a distance education model. In future experiences this point will have to be clarified. Another point that this research leaves open is the usefulness of the scrutable strategy. The results were negative, but the limited use of the web application makes it impossible to pronounce judgment, other than that the tool in itself is not motivating. In future experiences the use of the web application is going to be encouraged improving usability integrating it into the dotLRN platform. Once it has been integrated and reasonably used by students, it will be possible to know whether in the educational domain posed scrutability provides students with greater responsibility and whether this improves the collaboration and learning.