



UNIVERSIDAD
COMPLUTENSE
MADRID

Proyecto de Innovación y Mejora de la Calidad Docente

Convocatoria 2014

PIMCD 68-2014

Desarrollo de un módulo de análisis y explotación de datos de un Mooc mediante técnicas de Big Data y gramáticas de atributos, y su incorporación a un sistema de gestión de MOOC

Antonio Sarasa Cabezuelo

Facultad de Informática

Dpto. Sistemas Informáticos y Computación

Índice

1. **Objetivos propuestos en la presentación del proyecto**
2. **Objetivos alcanzados**
3. **Metodología empleada en el proyecto**
4. **Recursos humanos**
5. **Desarrollo de las actividades**
6. **Anexos**

1. Objetivos propuestos en la presentación del proyecto

Los objetivos que se plantearon en este proyecto fueron los siguientes:

- Estudiar el problema de cómo explotar la información que se genera en un sistema de gestión de MOOCs, estableciendo qué tipo de preguntas se quieren responder, y analizando de qué información se dispone y en qué formatos se encuentra.
- Definir una estrategia de aplicación de los resultados que se obtienen de los análisis de los datos: indicadores de calidad, mejoras del proceso de enseñanza-aprendizaje, adaptación de la oferta educativa, ...
- Definir una metodología para extraer, procesar e interpretar sistemáticamente los datos de interacción de los alumnos con un MOOC usando técnicas de “Big Data” y técnicas basadas en el uso de gramáticas de atributos.
- Diseñar un módulo software basado en la metodología definida en c), para el análisis y explotación de datos procedentes de la interacción con un MOOC , e integrarlo en un sistema de gestión de MOOCs concreto.
- Experimentar en un entorno real con un MOOC del área de filología/lingüística para evaluar el modulo implementado y analizar la información que se ha generado con su uso.
- Difundir los resultados en revistas y congresos especializados.

2. Objetivos alcanzados

De los objetivos propuestos inicialmente se han alcanzado con éxito los siguientes:

- Estudiar el problema de cómo explotar la información que se genera en un sistema de gestión de MOOCs, estableciendo qué tipo de preguntas se quieren responder, y analizando de qué información se dispone y en qué formatos se encuentra.
- Definir una estrategia de aplicación de los resultados que se obtienen de los análisis de los datos: indicadores de calidad, mejoras del proceso de enseñanza-aprendizaje, adaptación de la oferta educativa, ...
- Definir una metodología para extraer, procesar e interpretar sistemáticamente los datos de interacción de los alumnos con un MOOC usando técnicas de “Big Data” y técnicas basadas en el uso de gramáticas de atributos.
- Difusión de los resultados en ámbitos académicos: Celebración de un taller(ISELEAR 2014), y asistencia y presentación de un artículo en un congreso(TEEM 2014).

El objetivos referidos a implementación de los resultados teóricos no han sido alcanzados al no disponer de una suficiente financiación para su realización.

3. Metodología empleada en el proyecto

La forma de ejecución propuesta para el proyecto consistió en:

- Agrupar a los miembros del proyecto en diferentes grupos de trabajo especializados en desarrollar un conjunto de tareas concretas:
 - G1:Grupo de analizar la información, definir alcances y una estrategia de explotación.

- Miembros: Ana Fernández-Pampillón, Antonio Pareja y José María Ruiz Ruiz.
 - G2:Grupo de procesamiento de la información.
Miembros: José Luis Sierra y Antonio Sarasa
 - G3:Grupo de desarrollo tecnológico encargado de desarrollar el modulo e integrarlo.
Miembros: Daniel Rodriguez Cerezo, Bryan Temprado y Joaquín Gayoso.
 - G4:Grupo de expertos en el dominio encargados de experimentar con un MOOCs en filología/lingüística en un entorno real.
Miembros: Ángel Luis Encinas y M^a Carmen Horno.
 - G5:Grupo de coordinación encargado de coordinar al resto de grupos y dirigir el desarrollo del proyecto.
Miembros: Antonio Sarasa.
- Para cada objetivo planteado se definió un conjunto de tareas que permitieran cubrir dicho objetivo. Estas tareas fueron asignadas a los grupos de trabajos responsables de llevarlas a cabo. Se definió una planificación temporal disponiendo para su realización de unos tiempos prefijados. Cada tarea tenía asociada un conjunto de entregables. Así mismo en cada grupo se eligió un responsable de grupo encargado de mantener la comunicación con el grupo de coordinación.
 - Se realizó un análisis del riesgo del proyecto de manera que se dispusiera de alternativas ante situaciones no previstas en el desarrollo del mismo, y que permitieran asegurar la finalización correcta del proyecto.
 - Para conseguir el aseguramiento de la calidad del proyecto se definió un conjunto de directivas a cumplir por todos los grupos de trabajo en cuanto a formatos de los entregables, revisión de los resultados, pruebas a realizar,...
 - El desarrollo de cada tarea se realizó siguiendo una metodología de desarrollo incremental de varias iteraciones.

4. Recursos humanos

Los componentes del proyecto han sido:

Universidad Complutense de Madrid

Dpto. Inteligencia Artificial e Ingeniería del Software:

- José Luis Sierra Rodríguez(Profesor Titular)
- Joaquín Gayoso Cabada(Personal de Investigación)
- Bryan Temprado Battad(Personal de Investigación)
- Daniel Rodríguez Cerezo(Personal de Investigación)

Dpto. Filología Románica, Filología Eslava y Lingüística General

- Ana María Fernández Pampillón(Profesora Contratado Doctor)
- Ángel Luis Encinas Moral(Profesor Contratado Doctor)

Dpto. Didáctica y Organización Escolar

- José María Ruiz Ruiz(Profesor Titular)

Dpto. Sistemas Informáticos y Computación

- Antonio Pareja Lora(Profesor Colaborador)
- Antonio Sarasa Cabezuelo(Profesor Contratado Doctor)

Universidad de Zaragoza

Dpto. Lingüística General

- María del Carmen Horno Chéliz. (Profesor Titular)

5. Desarrollo de las actividades

El desarrollo del trabajo se realizará en 3 fases:

Fase 1: Estudio del problema del análisis y explotación de información de MOOCS. Se ha realizado un análisis de la información disponible en los MOOCs, se ha definido el alcance y los resultados que se desean obtener y se ha definido una estrategia de explotación de resultados.

Fase 2: Técnicas de procesamiento de información. Se ha analizado la aplicabilidad de técnicas de "Big Data" y de gramáticas de atributos para el procesamiento de la información, y se ha definido un modelo de procesamiento de la información.

Fase 3: Publicación de resultados. Se ha realizado el workshop ISELEAR 2014 durante los días 12-14 de Noviembre de 2014, y se ha asistido a la presentación de una ponencia al congreso TEEM 2014 del 1 al 3 de Octubre de 2014.

En este sentido los resultados de este proyecto han sido:

- Estudio del problema del análisis y explotación de información de MOOCS .
- Estudio sobre técnicas de procesamiento de la información procedente de los MOOCs.
- Organización del taller "5º Taller sobre Ingeniería del Software en eLearning (ISELEAR'14)" que se ha realizado como un track del congreso internacional SIIE 2014 durante los días 12-14 de Noviembre en la Universidad de La Rioja.
- Presentación de ponencia en el congreso TEEM 2014 durante los días 1-3 de Octubre en la Universidad de Salamanca (<http://dx.doi.org/10.1145/2669711.2669894>)

6. Anexos

Relación de gastos para los que se ha utilizado la subvención del Vicerrectorado

Material Informático

1 USB Kingston 32 GB

Total con 21% IVA-->14.50€

Asistencia a congreso

Congreso Teem 2014-->200 €

Total gastado-->200 €

Material Bibliográfico

Total gastado-->156,15 €

Total-->370,65€