

Actualización de materiales docentes: una perspectiva tecnológica y ética

Manuel Palomo Duarte*, Antonio Balderas Alberico, Federico Carrillo Chaves

*Departamento de Ingeniería Informática, Escuela Superior de Ingeniería

manuel.palomo@uca.es

RESUMEN: Cada vez es mayor la presencia de iniciativas de transparencia en administraciones públicas e iniciativas privadas. Esos datos permiten obtener una visión de la realidad libre de sesgos subjetivos, prejuicios, falsas noticias que ocultan intereses, etc. Entre las competencias que se desarrollan en la asignatura Bases de Datos se incluye "... la capacidad de reunir e interpretar datos relevantes para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética". Es por ello que consideramos de interés que los alumnos realicen sus prácticas de la asignatura usando datos reales y respondiendo preguntas sobre ellos que aumenten su espíritu crítico. Para ello se han analizado noticias en medios de comunicación tradicionales (papel y digital) y redes sociales, buscando cuáles de ellas estaban basadas en datos objetivos y disponibles públicamente. Esos datos se descargaron y se pasaron a formato relacional normalizado que se estudia en la asignatura. Por último se preparó una máquina virtual basada en el software libre de referencia en bases de datos, con una base de datos distinta para cada sesión de prácticas junto a un lote de ejercicios alineados con las competencias de la asignatura. El resultado ha sido recibido muy positivamente por los alumnos.

PALABRAS CLAVE: proyecto de innovación, materiales, bases de datos, datos abiertos, ética.

INTRODUCCIÓN

Cada vez es mayor la presencia de iniciativas de transparencia en administraciones públicas e iniciativas privadas, incluida la Universidad de Cádiz [1]. Los datos proporcionados permiten obtener una visión crítica de la realidad libre de sesgos subjetivos, prejuicios o tener que creer falsas noticias que ocultan determinados intereses. Sin embargo, para facilitar el tratamiento de los datos estos deben estar disponibles en un formato fácilmente manipulable, lo que se conoce como Datos Abiertos [2]. Aunque hay instituciones que ofrecen los datos en dicho formato [3], lamentablemente no es la norma general.

La asignatura Bases de Datos del Grado en Ingeniería Informática desarrolla la competencia CB3, "Que los estudiantes tengan la capacidad de reunir e interpretar datos relevantes (normalmente dentro de su área de estudio) para emitir juicios que incluyan una reflexión sobre temas relevantes de índole social, científica o ética". Es por ello que consideramos de interés que los alumnos de la asignatura realicen sus prácticas usando datos reales y respondiendo preguntas sobre ellos que aumenten su espíritu crítico. Para ello se han analizado noticias en medios de comunicación tradicionales (papel y digital) y redes sociales, buscando cuáles de ellas estaban basadas en datos objetivos y disponibles públicamente. Esos datos se descargaron y se pasaron a formato relacional normalizado que se estudia en la asignatura.

Por último se ha preparado una máquina virtual basada en el software libre de referencia en gestión de bases de datos, con un conjunto de datos distinto para cada sesión de prácticas, junto

a un lote de ejercicios alineados con las competencias de la asignatura para cada una. El resultado ha sido recibido muy positivamente por los alumnos.

TRABAJO REALIZADO

Se planificó un trabajo en 6 fases:

1. Consultar fuentes de datos abiertos sobre temas actuales de interés del alumnado.
2. Descargar conjuntos de datos de dichas fuentes.
3. Diseñar una base de datos relacional adecuada para cada conjunto de datos.
4. Normalizar cada conjunto de datos para incorporarlos a su correspondiente base de datos relacional.
5. Plantear ejercicios sobre el contenido de cada base de datos intentado despertar en el alumnado interés por aspectos ético que puedan mostrar los datos.
6. Montar un sistema basado en una máquina virtual con todo el software y datos instalado y configurado, de forma que se simplifique al máximo su uso. Estará basado exclusivamente en software libre.

A continuación se detalla los aspectos más importantes de cada una.

Búsqueda de fuentes

Se consultaron fuentes abiertos de interés para el alumnado. Para ello se buscaron noticias en ediciones digitales de periódicos y redes sociales que hicieran referencia a datos. Algunas preguntas que seleccionaron fueron: ¿cuál es mejor país del

mundo para nacer? ¿Cuáles son los factores para ser feliz en un puesto de trabajo? ¿Qué hace que una ciudad sea sostenible?

Cuando estaban disponibles, se descargaron conjuntos de datos de las fuentes encontradas.

Adaptación a base de datos relacional

Se adaptó el formato de los datos a una base de datos relacional, que es el modelo más extendido en la industria y el que se estudia en la asignatura.

Este proceso se complementó, cuando fue necesario, con una normalización. Esta normalización de bases de datos relacionales, que se estudia en la asignatura, reduce la redundancia de los datos, minimizando los problemas que esta provocaría y facilitando la integridad de datos.

De este modo los alumnos tuvieron a su disposición estructuras de datos modélicas que además de facilitar la realización de ejercicios, les sirven de modelo cuando diseñen sus bases de datos en la asignatura.

Realización de boletines de ejercicios

Se plantearon ejercicios para todas las sesiones de prácticas de la asignatura, usando una base de datos distinta en cada una. En concreto:

1. En la primera sesión se trabaja la manipulación básica de datos con una base de datos de la huella ecológica de supermercados en España. Los datos los recopiló Greenpeace España en 2018, y permiten responder a preguntas referentes a la eliminación de plásticos de un solo uso para 2025, el fomento de envases reutilizables, la venta de fruta y verdura a granel, etc.
2. En la segunda sesión se trabaja la manipulación de datos con una base de datos de la Comisión Europea sobre la edad media de emancipación de hombres y mujeres en cada país de la Unión Europea. De este modo, los alumnos pueden realizar consultar referente a la edad de emancipación en países mediterráneos respecto al resto de Europa, o diferencias por sexos según el país, etc.
3. En la tercera sesión se trabajan las funciones y expresiones con una base de datos de geolocalización de miradores en la provincia de Cádiz. Estos datos, descargados del portal de Reutilización de Información Pública del Gobierno de España, fueron enriquecidos con información sobre la comarca en la que se encuentra cada mirador. De esta forma los alumnos pudieron consultar información sobre municipios con miradores, miradores en una determinada latitud, comarcas con más y menos miradores, etc.
4. En la cuarta sesión se trabajan las consultas anidadas con una base de datos

de personal empleado en I+D (jornada completa y parcial) por sector de ejecución, ocupación y sexo en 2017. Esta información se descargó de la sección pública la página web del Instituto Nacional de Estadística. De esta forma los alumnos pudieron consultar información sobre empleo en I+D por sector público/privado, comparación de empleo en dicho sector por sexos o según cualificación.

5. En la quinta sesión se siguen trabajando las consultas anidadas con una base de datos de los libros de texto utilizados en centros de enseñanza andaluces privados y públicos en el curso 2018/2019 (en nivel Primaria y E.S.O.). De esta forma los alumnos pudieron consultar información sobre los centros públicos y privados por provincia, los libros más usados en determinado municipio, provincia o curso, etc.
6. En la sexta sesión se trabajan las consultas a múltiples tablas con una base de datos de toreros, cantantes y bodas entre ellos. Esta información se descargó usando la interfaz del proyecto Dbpedia, que consume datos de Wikipedia. De esta forma los alumnos pudieron consultar información sobre toreros ya fallecidos que se habían casado con cantantes de copla, matrimonios entre parejas de una determinada edad, etc.
7. En la sexta sesión se siguen trabajando las consultas a múltiples tablas con una base de datos de Censo Electoral en 2019. Esta información se descargó de la sección pública la página web del Instituto Nacional de Estadística. De esta forma los alumnos pudieron consultar información sobre municipios por provincias, extranjeros residentes por comunidad autónoma, evolución de la población entre dos censos electorales, etc.

Creación de máquina virtual

Las bases de datos anteriores se instalaron en un sistema basado en una máquina virtual VirtualBox lista para usarse (sin necesidad de instalar software ni configuraciones), redactando un documento sobre su uso. Como el espacio que ocupaba es un poco alto, esto ralentizaba la ejecución en algunos portátiles con pocos recursos. Por lo tanto se realizó otra versión con el sistema de virtualización Vagrant. Tanto la virtualización como el software usado (MySQL) son software libre.

DISCUSIÓN Y CONCLUSIONES

La valoración del proyecto es muy positiva. Consideramos que la disponibilidad de bases de datos reales con ejercicios ayuda a aumentar la motivación de los alumnos. Hay que tener en

cuenta que la asignatura suele tener más de 150 alumnos cada curso, por lo que la dedicación con cada uno es muy limitada. Y, aunque es cierto que la información que hemos usada es toda pública y podemos pensar que la podrían usar los alumnos, ha sido necesario procesarla para que esté en el formato (relacional) que se trabaja en la asignatura.

Hay que destacar que para la consecución de los objetivos fue clave poder contar con un becario, que se encargó de tareas de búsqueda y clasificación de información, así como asuntos técnicos todos ellos dentro de las competencias de su titulación. Además, el contar con un alumno ha permitido detectar las áreas de interés de los alumnos (para buscar bases de datos relacionadas) y para solucionar los problemas de rendimiento que se observaron en la máquina virtual.

De cada a próximos cursos se tiene como trabajo futuro actualizar las bases de datos con datos más recientes que se recopilen. De este modo, además obtener resultados actuales, se podrá comparar la evolución estos en el tiempo.

REFERENCIAS

1. Portal de Transparencia de la Universidad de Cádiz. <https://transparencia.uca.es> .Último acceso 24 de septiembre de 2019.
2. Portal de Datos Abiertos 5 estrellas. <https://5stardata.info/es>. Último acceso 24 de septiembre de 2019.
3. Portal de Datos Abiertos de la Universidad de Granada. <https://opendata.ugr.es> .Último acceso 24 de septiembre de 2019.
4. Memoria del Grado en Ingeniería Informática de la Universidad de Cádiz de febrero de 2018.

AGRADECIMIENTOS

Los autores agradecen la financiación recibida por parte de la Unidad de Innovación Docente de la UCA que ha permitido la contratación de un becario para el proyecto.