

Social-Univ 2.0: Monitorización del entorno digital social universitario

Social-Univ 2.0: Monitoring the Digital and Social Environment of the University

Isabel Moreno^{1,2}, Javi Fernández², Yoan Gutiérrez¹

¹Instituto Universitario de Investigación Informática, Universidad de Alicante

²Departamento de Lenguajes y Sistemas Informáticos, Universidad de Alicante
Apdo. de Correos 99 E-03080, Alicante, Spain
{imoreno, javifm, ygutierrez}@dlsi.ua.es

Resumen: Social-Univ 2.0 es una aplicación web que monitoriza el ecosistema digital de universidades con presencia en redes sociales y, más concretamente, se centra en analizar el entorno de la Universidad de Alicante. Usa las tecnologías del lenguaje humano para extraer meta-datos sobre comentarios en Twitter para representar perfiles de actores sociales, así como su relación social en un período de tiempo. Esta aplicación, en su primera versión: facilita tanto el estudio de interacciones sociales y sus matices, como el análisis de grupos sociales y sus comportamientos; identifica a actores sociales relevantes e influyentes, así como los temas comunes en determinados grupos sociales; y favorece la medición del impacto de campañas y eventos.

Palabras clave: Tecnologías del lenguaje humano, universidad, relaciones sociales

Abstract: Social-Univ 2.0 is a web application that monitors the university digital ecosystem in social networks and, specifically, it focuses on analysing the University of Alicante environment. It uses human language technologies to extract meta-data from tweets and represents actors' profiles, as well as their social relations in a time period. This application, in its first version: eases the study of social interactions and its nuances, as well as the analysis of social groups and their behaviour; identifies relevant and influential social actors as well as common topics in groups; and helps to measure impacts from campaigns and events.

Keywords: Human language technologies, university, social relationships

1 Introducción

Millones de personas de todo el mundo participan en las redes sociales. Las interacciones de los usuarios ya no pasan únicamente por consumir información, si no que también expresan su opinión sobre dicha información. Si analizamos Twitter, una de las redes sociales más conocidas, observamos que dispone de más de 344 millones de usuarios activos¹, que generan 500 millones de tweets al día².

Existen infraestructuras analíticas, como Atribus³ o SumAll⁴, que automáticamente recuperan y procesan información de las redes sociales para ser analizada por usuarios. Sin embargo, si nos enfocamos en las universidades, a día de hoy, no encontramos infraestructuras que analicen el impacto que es-

tas tienen en las redes sociales. Por tanto, si deseáramos identificar los roles de todos aquellos actores y eventos que pudieran incidir en la reputación de una universidad a nivel social y global, necesitaríamos hacerlo de forma (semi-)manual y emplear para ello cuantiosos recursos, humanos y temporales.

Ante la falta de tecnologías que resuelven esta problemática, la herramienta *Social-Univ 2.0* tiene como objetivo principal el uso de Tecnologías del Lenguaje Humano (TLH) para extraer comentarios de las redes sociales, así como sus meta-datos, y representar tanto perfiles de actores sociales (empresas, instituciones, personas, etc.) como la relación social que existe entre ellos y una universidad en un determinado período de tiempo.

2 Descripción del Sistema

Social-Univ 2.0 analiza comentarios de la red social Twitter y relaciona actores mediante los mensajes en los que intervienen. Pa-

¹<https://bit.ly/1pIwiJ8> (Marzo 2019)

²<https://bit.ly/1pJzREY> (Marzo 2019)

³<http://www.atribus.com> (Febrero 2019)

⁴<https://sumall.com> (Febrero 2019)

ra iniciar el estudio social, el sistema requiere a un actor de partida (i.e. una universidad) que se debe proporcionar según la forma autóctona de identificador en redes sociales (ej. @UA_Universidad) o nombre (ej. Universidad de Alicante).

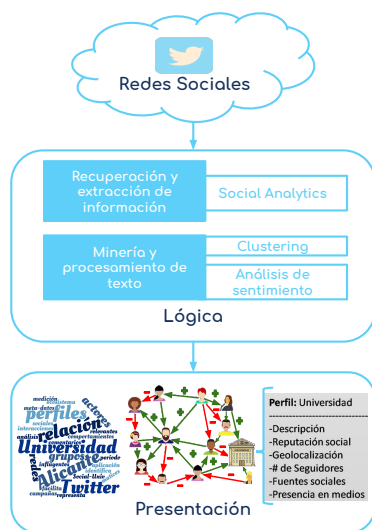


Figura 1: Arquitectura de *Social-Univ 2.0*

La Figura 1 muestra como *Social-Univ 2.0* se divide en dos capas. Por un lado, la capa lógica, detallada en la Sección 2.1, se encarga de descargar comentarios de Twitter y analizarlos semánticamente para valorar las diferentes posturas que cada actor tiene respecto a otros, así como su influencia en el entorno digital-social. Por otro lado, la capa de presentación, detallada en la Sección 2.2, muestra al usuario la información social analizada mediante grafos, perfiles y nubes de palabras en una interfaz que permite filtrar el análisis por periodos inferiores a 6 meses.

2.1 Capa de lógica de la aplicación

Esta capa hace uso de las TLH y, como puede verse en la Figura 1, se divide en dos módulos:

El **módulo Recuperación y extracción de información** se encarga de descargar el contenido generado por el usuario en Twitter, en inglés y español, que mencione a la universidad a estudiar. Este proceso se realiza reutilizando la plataforma Social Analytics (Fernández et al., 2017) que permite recuperar el contenido de la Web 2.0 y extraer meta-datos sobre cualquier entidad digital.

El **módulo Minería y procesamiento de texto** analiza el contenido generado por el usuario para determinar: (i) Cuáles son los términos relevantes y contabilizarlos con

técnicas de clusterización usando el algoritmo *Lingo*, con los parámetros por defecto, de la herramienta Carrot2⁵ (Stefanowski y Weiss, 2003). (ii) Cuál es la geolocalización de contenido y del autor, buscando la población de origen de los usuarios en la base de datos de Geonames⁶, de forma similar a Fernández et al. (2017). Y (iii) si es una opinión, cuál es su polaridad, basándonos en la aproximación de Fernández, Gómez, y Martínez-Barco (2014) y Fernández et al. (2017). Esto último es clave para valorar las diferentes posturas (valoración, opinión, etc.) que cada actor tiene de otros y su influencia en un entorno social digital. Finalmente, se obtienen modelos de representación de perfiles de actores sociales y su relación en un período de tiempo.

2.2 Capa de presentación

Esta capa genera un informe mediante HTML, PHP y Javascript que visualiza el análisis de la capa lógica (Sección 2.1) de un modo útil y sencillo para ser interpretado por usuarios finales (gestores universitarios). El informe contiene:

1. Selector de fechas para analizar periodos de tiempo (día, semana o mes).
2. Nube de palabras relevantes clusterizadas que se muestra con D3-Cloud⁷.
3. Grafos sociales y conjuntos, donde se representan los actores sociales y sus relaciones mediante la herramienta Vis⁸. Las relaciones se establecen a través de los mensajes recuperados, que se clasifican y ponderan como positivas o negativas según el contenido del mensaje.
4. Perfil de cada actor con: descripción, geolocalización, páginas web relacionadas, número de seguidores en Twitter (de su API pública⁹) y reputación social en una escala de 1 a 10 para el periodo seleccionado.

3 Caso de Estudio: Universidad de Alicante

Una vez elegida la universidad a monitorizar, comienza a realizarse el seguimiento en

⁵<https://project.carrot2.org> (Marzo 2019)

⁶<https://www.geonames.org/> (Marzo 2019)

⁷<https://github.com/jasondavies/d3-cloud> (Marzo 2019)

⁸<https://github.com/almende/vis> (Marzo 2019)

⁹<https://developer.twitter.com> (Marzo 2019)



Figura 3: Perfil de la Universidad de Alicante

las redes sociales. En este caso, se ha creado una instancia de la aplicación para la Universidad de Alicante¹⁰(UA). A medida que se van encontrando diferentes publicaciones mencionando a la UA, se genera una nube de palabras que recoge la forma de expresarse de los actores involucrados durante un periodo de tiempo. Por ejemplo, en la Figura 2a vemos que, del 4 al 10 de marzo de 2019, el “*Semanario Internacional de Jóvenes Investigadores*” recibió atención los días previos a su celebración (2 de abril de 2019). Otro ejemplo, durante el periodo 11 al 17 de marzo de 2019, lo encontramos en la Figura 2b y vemos que se produjo el “*congreso Can Robots Invent and Create*”, organizado por la Facultad de Derecho de la UA, del 14 al 15 de marzo.

También se construye un grafo social que relaciona a los actores relevantes en un lapso de tiempo. Continuando con la Figura 2b, se observa como se produce una influencia positiva (flecha verde) con todos los actores sociales que interactúan con la UA; o como el Equipo de fútbol sala femenino de la UA (usuario UF) tiene tanto influencias negativas (flechas rojas) con 2 usuarios, como positivas con 3, como se aprecia en la Figura 2a. Por último, si hacemos clic sobre un actor del grafo, accedemos a su perfil. La Figura 3 muestra el perfil de la UA con 97.2 mil seguidores y una reputación de 8.5 sobre 10.

Todo ello permite analizar a la UA, a sus órganos, y cualquier otro actor que interactúe con ella. Además, posibilita observar la evolución social digital de la UA para que, como actores sociales o representantes de dicha entidad, consigamos visualizar grafos cada vez más poblados (i.e. más actores debaten so-

bre la UA) y donde el color dominante sea el verde (i.e. mejor reputación social).

4 Conclusiones

Hemos presentado la primera versión de *Social-Univ 2.0*, una herramienta que extrae comentarios de la red social Twitter sobre una universidad y sus meta-datos. Con esta tecnología, hemos estudiado el entorno social de la UA y las entidades sociales digitales con las que se relaciona. Mediante el grafo social generado por *SocialUniv*, hemos analizado las interacciones sociales entre grupos de usuarios y sus comportamientos en diferentes momentos o detectado a los actores sociales relevantes e influyentes. Además, la nube de palabras nos ha permitido identificar la expresión social en los comentarios analizados. Por último, el perfil ha permitido caracterizar al actor principal de este estudio, la UA.

El próximo paso es evaluar extrínsecamente *Social-Univ 2.0* para medir su utilidad como un todo. En un futuro, podría adaptarse a otras redes sociales, a aplicaciones móviles y al campus virtual de universidades españolas.

Agradecimientos

Aplicación desarrollada por el Grupo de Procesamiento del Lenguaje Natural y Sistemas de Información (GPLSI)¹¹ de la UA. Financiada por la UA a través del proyecto “*Social-Univ 2.0: Tecnologías del Lenguaje Humano, aplicación para la monitorización omnicanal del entorno social de la Universidad de Alicante (ENCARGOINTERNO5-19EN)*”, conjuntamente con la Generalitat Valenciana (PROMETEU/2018/089) y el gobierno de España (RTI2018-094653-B-C22 y RTI2018-094649-B-I00).

Bibliografía

- Fernández, J., J. M. Gómez, y P. Martínez-Barco. 2014. A supervised approach for sentiment analysis using skipgrams. En *Proceedings of the Workshop on Natural Language Processing in the 5th Information Systems Research Working Days (JISIC)*, páginas 30–36. Association for Computational Linguistics.
- Fernández, J., F. Llopis, P. Martínez-Barco, Y. Gutiérrez, y Á. Díez. 2017. Analizando opiniones en las redes sociales. *Procesamiento del Lenguaje Natural*, 58:141–148.
- Stefanowski, J. y D. Weiss. 2003. Carrot 2 and language properties in web search results clustering. En *International Atlantic Web Intelligence Conference*, páginas 240–249. Springer.

¹⁰<https://www.ua.es/> (Marzo 2019)

¹¹<https://gplsi.dlsi.ua.es/> (Marzo 2019)