

Difundiendo el conocimiento a través de redes sociales con Carlos Piña García

¿Sabías que las redes sociales son un gran medio para comunicar temas relacionados con la ciencia? El experto en recolección y análisis de datos del IIMAS, Carlos Piña García, estuvo en Aldea Digital 2016 con su ponencia "Divulgación de la ciencia por medio de twitter", donde habló sobre cómo la información que compartimos a través de Twitter puede utilizarse para investigar un sin número de cosas sobre nosotros mismos y sobre el mundo. Explicó que "cada vez hay más personas de distintas disciplinas que buscan involucrarse en estos estudios. Este tipo de análisis de información se ha hecho mucho a nivel internacional, pero queremos lograr que México no se quede atrás".

Antes de entrar de lleno al tema central de su ponencia, nos dió un preámbulo sobre el flujo de diversas redes sociales, las interacciones y por qué esto importa. "Todos los días se producen 2.5 trillones de bytes de información digital"; para hacer de estos datos abstractos algo más concreto, Carlos, a través de un programa de computadora, nos mostró la cantidad abismal de información que se está subiendo cada segundo en distintas redes sociales, páginas web, canales y distintas plataformas digitales. Habló, por ejemplo, sobre un estudio de las conversaciones de Twitter sobre la Eurocopa: "Este mapa muestra puntos de luz en las zonas geográficas en las que se está hablando de la Eurocopa. Podemos ver que Europa está completamente iluminado; sin embargo, si nos fijamos en México, sólo la Ciudad de México se encuentra iluminada. Muchas veces cuando pensamos que toda la República se está uniendo a hablar de un tema o una causa, no nos damos cuenta de que se trata sólo de un punto pequeño en el mapa". Continuó explicando: "lo mismo sucedió con el tema del zika, pues a través de las conversaciones de Twitter, se muestra que en México el tema no fue tan relevante, como en otros países". Estos mapas de información desmitifican muchas creencias falsas sobre diversos temas. El científico de datos explicó que no sólo se pueden medir textos, localizaciones y temas, sino también el sentir de las personas hacia distintos tópicos.

Después Carlos Piña habló de la ciencia de datos y el big data: "Ahora todo mundo habla de big data, minería de datos, ciencia de datos, y demás. Todo mundo dice que hace big data; sin embargo éste es como el sexo en la adolescencia, todo mundo dice que lo hace, nadie sabe hacerlo y todos quieren hacerlo porque escuchan que otros lo están haciendo. Es mejor hablar de ciencia de datos, pues hacer big data es algo muy complejo".

Para hablar de la relación de la ciencia con Twitter, nos ofrece un panorama general del movimiento de esta red social: "Cada día se generan 58 millones de tuits, y muchos de estos son basura para la investigación. En Twitter hay alrededor de 650 millones de personas registradas; pero sólo 135 mil tuitean diariamente. De acuerdo con Google Trends, Twitter tuvo su pico en 2012. En México el porcentaje de la sociedad que usa Twitter es muy bajo, por lo que la información que obtenemos de éste no representa la percepción de todo el país. No podemos ver el panorama completo, sólo una pequeña porción."

A través de un experimento, nos mostró cómo es que se llevan a cabo los mapeos que él hace: "Saquen sus celulares y escriban en Twitter #AldeaDigital. Tuitéenlo cuando yo les diga (dice

mientras programa la búsqueda de ese hashtag en el programa Python) y verán cómo en un momento Twitter enviará toda su información al programa”. Y así fue. Una vez que Piña nos mostró esto, explicó que esa información venía sucia, y que para poder hacer un estudio de ésta era necesario limpiarla: “Como verán la info no viene limpia, se debe filtrar. Mientras más relevante sea ésta y más gente hable del tema, más difícil será obtener información legítima”.

Explicó que los bots (cuentas administradas por inteligencia artificial) simplemente hacen ruido sobre la información que se pretende analizar; por lo que si se desea, por ejemplo, saber cuál es la influencia de algún deporte, habrá que descartar dicha información, pues dado que es spam, no reflejará los intereses que se buscan. “El 24% de los tuits los generan bots. Éstos llegaron para quedarse, por lo que hay que aprender a filtrarlos, pues no podemos eliminarlos”

Finalmente nos relató sobre el proyecto de investigación científica que está haciendo en el IIMAS de la UNAM: “analizamos qué tanto la gente habla sobre ciencia en Twitter. Utilizamos palabras clave como: inteligencia artificial, biotecnología, nanotecnología y demás; exploramos qué dispositivos se utilizan para hablar de dichos temas, en qué países se tuitea más sobre esos temas, qué tipo de personas son las que hablan sobre esos temas, en fin... es increíble la cantidad de información que puedes extraer de un solo tuit”. Explicó que durante 92 días se recolectó una muestra de cerca de 300 mil tuits ya limpios; estos se categorizaron por tema para ver cual de los temas es el que se tocaba más. ‘Neurociencia’ fue la más popular en el periodo seguido por ‘genómico’, ‘biotecnología’ y ‘nanotecnología’. Donde más se publicaba sobre ciencia fue en Estados Unidos y Europa. Fue notable que en China y Rusia no se habló de estos temas, lo que refleja que no divulgan información científica a través de Twitter, pues China, por ejemplo, esta red está prohibida”.