

Fondo CONACYT-INEGI

Propuesta de tema para integrar la Convocatoria 2016-2

Nombre del tema: 2	Diseño e implementación de métodos de búsqueda por similitud de usuarios de redes sociales.
Importancia y utilidad del tema para el INEGI:	Aprovechar la información proveniente de las redes sociales, para el conocimiento de ciertos aspectos de la población, es ahora, una tendencia adoptada por algunas oficinas de estadística. Mediante el análisis de esta información es posible medir y obtener indicadores útiles para la toma de decisiones en distintos ámbitos. Algunas de las ventajas de esta información, es que se genera constantemente y cada vez, es más representativa. En un esfuerzo por mantenerse a la vanguardia, el INEGI ha venido impulsando el apropiamiento y desarrollo de instrumentos tecnológicos y metodológicos, enfocados en el aprovechamiento de la información proveniente de medios sociales (<i>social media</i>). Esta propuesta aborda un problema identificado en los procesos de apropiamiento y desarrollo.
Descripción general del problema a resolver:	La actividad de usuarios de redes sociales dentro de entornos virtuales se manifiesta a partir de sus publicaciones (tuits, post, compartir archivos, videos, imágenes, etc.), reacciones a las publicaciones de otros usuarios (retuit, like, share, etc.) o acciones automáticas que se generan con previa autorización del usuario (check-in, notificación de actividad, etc.). Todas estas posibles “acciones” son susceptibles de ser registradas y almacenadas; y a su vez caracterizan la actividad del usuario. Una problemática importante a resolver es el diseño e implementación de modelos de representación de los usuarios, que capturen la información sobre su actividad en redes sociales y que al mismo tiempo sean útiles para la aplicación de métodos de búsqueda por similitud.
Objetivo general:	Construir un marco metodológico y computacional útil para la implementación de modelos de representación de usuarios de redes sociales y la aplicación de métodos de búsqueda por similitud, que se nutran de la actividad pública de los usuarios y que sean útiles para el INEGI en la generación de estadísticas.
Objetivos específicos:	<p>Diseño e implementación de un modelo de representación que permita transformar la información pública de los usuarios de redes sociales como una secuencia de objetos abstractos; esta representación se puede pensar como una huella digital.</p> <p>Utilizando estos modelos de representación de perfiles, implementar métodos para recuperar grupos de usuarios similares entre sí.</p>
Tiempo estimado de realización:	1 Año
Producto(s) esperado(s):	
<ol style="list-style-type: none"> 1. Modelo de representación de usuarios de redes sociales basado en su actividad pública. 2. Métodos de búsqueda de usuarios similares. 3. Prototipo de software que implemente los modelos y métodos desarrollados. 4. 2 artículos científicos donde se reporten las bondades de los modelos y métodos desarrollados. 5. Documento con el reporte global in-extenso de la investigación, que atienda a detalle el objetivo general y los objetivos específicos propuestos. 	

Fondo CONACYT-INEGI

Propuesta de tema para integrar la Convocatoria 2016-2

6. Un taller de al menos cuatro horas, dirigido a un grupo de trabajadores del INEGI donde se exponga el desarrollo y el resultado de la investigación, así como el uso de los materiales generados
7. Un artículo de investigación con calidad suficiente para ser dictaminado favorablemente para publicar en Realidad, Datos y Espacio Revista Internacional de Estadística y Geografía, en el que se presenten de manera clara y sintética los resultados de la investigación. <http://rde.inegi.org.mx/>

Observaciones:

Este proyecto requiere que la institución que lo ejecute tenga la capacidad de recolectar y analizar grandes cantidades de información proveniente de redes sociales, de manera automática y utilizando equipo de cómputo de alto rendimiento.