

Analítica Social y de la Web *Social and Web Analytics*

1.2 Proceso de explotación de datos





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Introducción a los grafos sociales

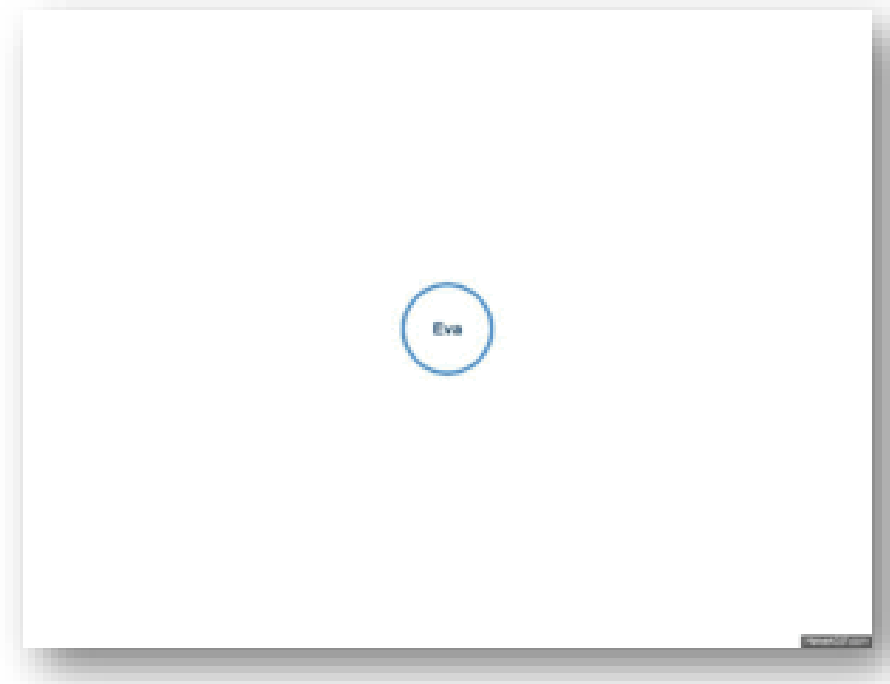


Introducción a los grafos sociales

Definición

El grafo social es un tipo de grafo que representa las relaciones sociales entre diferentes entidades. En síntesis, es un modelo o representación de una red social, sustentada en la teoría de grafos. Este término se ha descrito como «el mapa global de todo el mundo y cómo están relacionados».

El grafo social es un concepto abstracto que describe las relaciones entre las personas en línea, en contraposición a la idea de una red social, que describe las relaciones en el mundo real. Entonces, el gráfico social es digital y, más importante aún, se define explícitamente por todas las conexiones en cuestión.



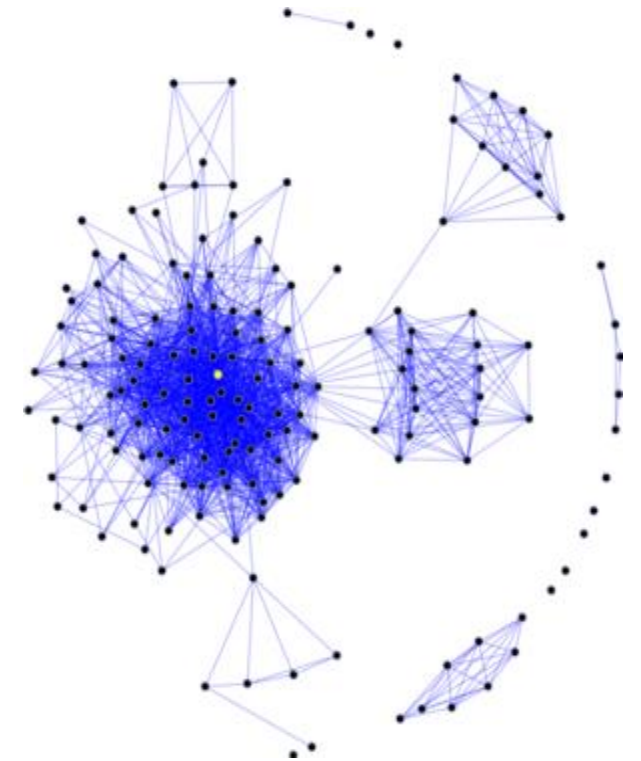
El grafo social

Origen

Fue popularizado cuando Mark Zuckerberg, fundador de Facebook, lo utilizó para referirse a la red social de relaciones entre los usuarios de Facebook.

Ello ocurrió en el marco de la conferencia Facebook F8 (2007), cuando fue utilizado para explicar que la plataforma de Facebook se beneficiaría del gráfico social, tomando ventaja de las relaciones entre los individuos que Facebook ofrece para ofrecer una rica experiencia en línea.

Aquella ha sido la definición más amplia para referirse a un gráfico social de todos los usuarios de Internet.



El grafo social

¿Qué tamaño tiene?

**Puede que haya pasado de moda, pero
Facebook sigue siendo el rey: ya tiene
3.000 millones de usuarios**

El grafo social

Problemas

Se han presentado varios problemas con respecto a la implementación existente del gráfico social propiedad de Facebook. Por ejemplo, actualmente, un servicio de redes sociales desconoce las relaciones que se forjan entre personas en un servicio diferente.

Esto crea una experiencia en línea que no es perfecta y, en cambio, proporciona una experiencia fragmentada debido a la falta de un gráfico disponible abiertamente entre servicios. Además, los servicios existentes definen las relaciones de manera diferente.

El mismo problema, a otra escala, plantea dificultades en la **emergencia de un metaverso integrado** que ofrezca persistencia de la identidad digital entre “miniversos”.

El grafo social

Tipos

- **Grafo de usuarios:** Muestra las relaciones entre individuos interconectados
- **Grafo de contenidos:** Muestra las relaciones entre tipos de contenidos (texto, imágenes, vídeos, multimedia), basándose en la semántica del contenido u otras variables que puedan relacionarlos
- **Grafo de intereses:** va un paso más allá, relacionando a individuos basándose en sus intereses mutuos (aunque no se conozcan ni hayan conectado)



COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Pregunta:

¿Qué es un influencer?

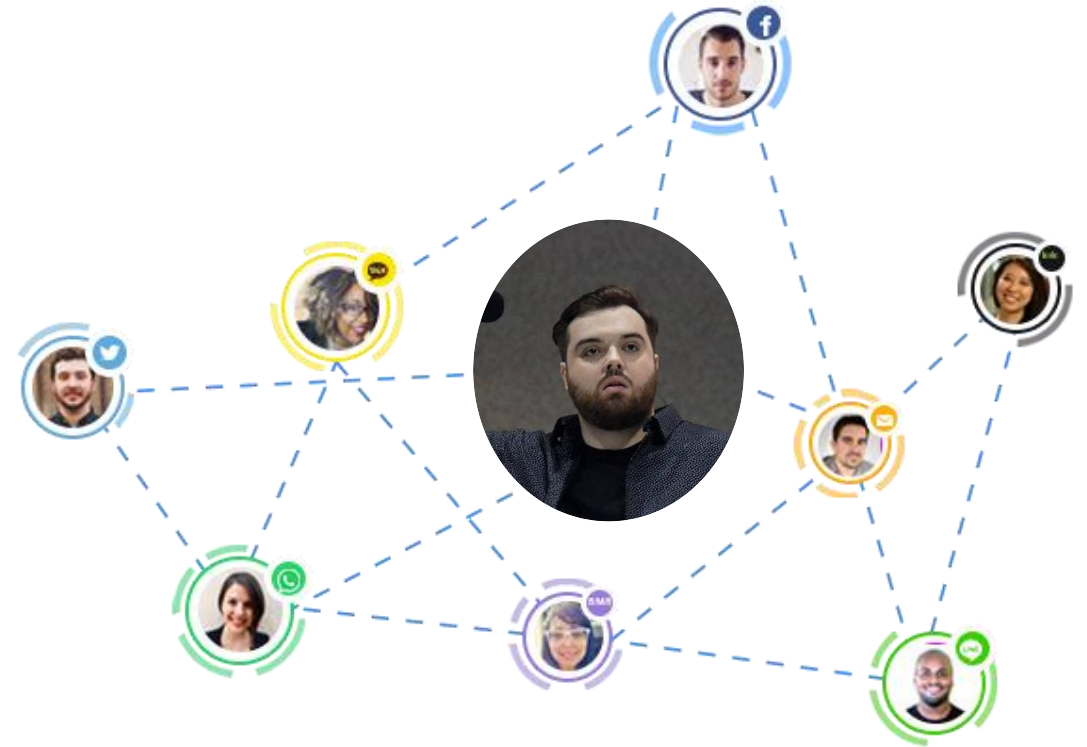


El grafo social

Influencia

Como ocurre en la vida real, algunos nodos del grafo social tienen más “peso” que otros. Se ha desarrollado una ciencia completa en torno al concepto de la identificación de influyentes o estrellas.

El asunto es complicado, ya que la influencia puede depender de la temática o subgrupo social, por lo que es muy relativa, y se puede medir por cantidad o por calidad de la influencia.



El grafo social

Klout



Servicio que se enfocaba en medir tu habilidad para generar acción en las personas, es decir, la influencia que tenías sobre alguien para que éste reaccione.

Klout medía tu influencia calificándola de 1 a 100; cuanto más alta fuera la puntuación más influyente eras. Para hacer esta medición Klout tomaba datos de sitios como Twitter, Facebook, Google+, LinkedIn, Foursquare, Youtube, WordPress, Instagram, Tumblr, Blogger o Last.fm y medía la influencia de la red de una persona; el contenido creado y cómo otras personas interactuaban con ese contenido, cómo interaccionaba el usuario con otros usuarios o el alcance y amplificación de sus mensajes.



El grafo social

Klout

Para realizar este cálculo se miden tres variables:

- True Reach (Alcance real) Es la cantidad de personas reales a las que tus publicaciones llegan. Se eliminan cuentas con poca actividad, robots y spam.
- Amplification (Amplificación) Mide la probabilidad de que el contenido que publicas sea compartido o comentado.
- Network Score (Puntuación en la red) Si las personas que reaccionan a tus publicaciones son consideradas influyentes, tu Network Score será más alto.

¿Por qué es importante el Klout y para qué se utiliza?

- Entrevistas de empleo: ¿Un Social Media Manager con poca influencia online? No, gracias.
- Invitaciones a actividades: 50+ Klout party son los nuevos parties elites.
- Descuentos: ¿Tu Klout es más de 55? Pues tienes 20% de descuento.
- El premio a la Influencia

El grafo social

Controversia y venta

Uno de los problemas planteados por Klout era la dependencia en “métricas de vanidad”, por lo que corría el riesgo de convertirse en un juego o incluso ser vulnerable a trampas, y en consecuencia terminar siendo inútil.

En 2014 fue vendido a Lithium Technologies por 200 millones de euros, y en 2018 se publicó su desaparición definitiva

Consejos para elevar al cielo una Puntuación en Klout

- Crear y compartir contenidos.
- Iniciar conversaciones.
- Crear y conectar tus redes.
- Enfócate con fuerza en Twitter.
- Haz Auto-Follow a todos.
- A Klout no le interesa de dónde proviene el contenido, sólo cuántos Retweet tienes.
- Involúcrate en cada chat Twitter en que sea posible.
- Sigue las etiquetas de conferencias.
- Lanza decenas de citas inspiradoras.
- Troll Reddit o Digg para los enlaces. Encuentra los links de batalla de otras fuentes y twitealos por tu cuenta.
- Toma parte en las etiquetas más relevantes (virales) de la semana.
- *No te vayas jamás de vacaciones. ¡Si dejas de twitear y dejas de seguir estas reglas por unos pocos días, tu “influencia” se irá al suelo!*





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Los datos en las Redes Sociales



Los datos en las Redes Sociales

Plataforma vs aplicación

Comenzando por Facebook y siguiendo con el resto de las redes sociales mayoritarias, marca un punto de inflexión el momento en que empiezan a venderse como plataformas (se permite crear aplicaciones por encima que exploten los datos de distintas maneras) en lugar de aplicaciones (el uso de los datos es fijo).

Ya sea gratis o con un coste, los datos amasados por las redes sociales pueden ser empleados de forma automatizada por terceros para dar lugar a nuevas aplicaciones o estudios.

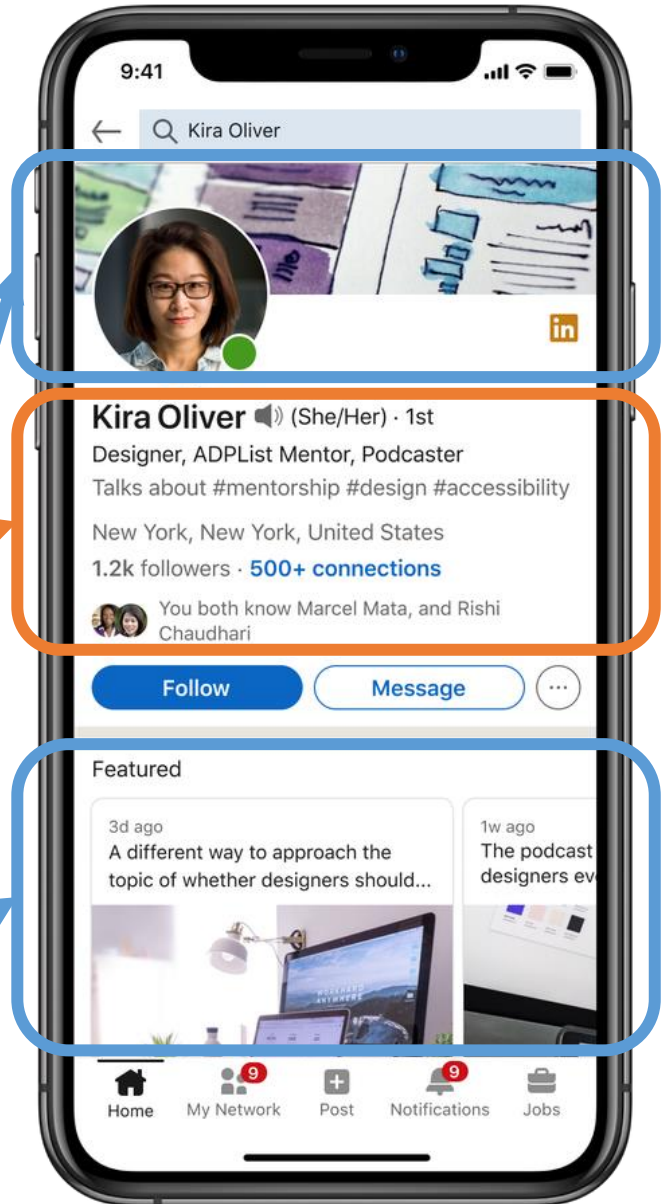
Esta es la piedra angular de la minería de datos de redes sociales, y por tanto de esta asignatura.

Los datos en las Redes Sociales

Datos sociales

Como su nombre indica, se trata de los datos adquiridos de las redes sociales. Los datos de los usuarios que son numéricos o cuantificables (nombre, ciudad, intereses, hábitos, etc.) se llaman **datos estructurados**.

Sin embargo, gran parte de los datos en redes sociales se publica en forma de textos, imágenes o vídeos, que a pesar de ser muy ricos en información, no pueden analizarse de forma directa. Por ello se llaman **datos no estructurados**.



Los datos en las Redes Sociales

Datos sociales

El proceso de aplicación de métodos rigurosos para obtener información relevante de los datos sociales se llama **análisis de datos sociales** (***social data analytics***).

Puesto que hay pocas limitaciones en la publicación en redes sociales, muchos datos son simplemente irrelevantes (cuentas, contenidos, interacciones...)

Por ello, gran parte del esfuerzo de esta disciplina tendrá por objetivo **separar la información del ruido**. Solo tras limpiarlos podremos emplear métodos estadísticos y algorítmicos para obtener conclusiones relevantes:
VALOR PARA EL NEGOCIO.



Los datos en las Redes Sociales

Semántica

Las imágenes y los vídeos también son símbolos, no obstante en esta asignatura no profundizaremos en su análisis.

La **semántica** (proveniente de la ciencia lingüística y de la filosofía) trata sobre el **significado de los datos**, que se obtiene al comprender las relaciones entre palabras, frases y otros tipos de símbolos (en redes sociales, podríamos hablar de emoticonos, memes, etc. que no siendo un lenguaje formal como tal, y sin embargo pueden contener tanto o más significado que las propias palabras).



Los datos en las Redes Sociales

Plataforma vs aplicación

Para la extracción del significado de los datos sociales, las tecnologías semánticas emplean algoritmos de inteligencia artificial y de *machine learning*.

Estos algoritmos permiten encontrar patrones que luego puedan ser interpretados por humanos.



Netflix te crea portadas personalizadas según tu perfil de usuario

Netflix conoce que personajes te gustan más para mostrar portadas que llamen tu atención

Por Manu Contreras  @mcontreras, hace 5 años



<https://netflixtechblog.com/artwork-personalization-c589f074ad76>

Los datos en las Redes Sociales

La velocidad a la que avanzamos

make them accessible to everyday users. So, today we are able to search our questions and select websites or pages that are linked to the answer. Even social media content is more and more accessible via search engines. You may find a tweet that you created two years back suddenly showing on a Google result. The problem of organization of web content is almost a solved problem. However, wouldn't it be more exciting if you asked a question on Google, Bing, or another search engine and it directly gave you the answer, just like a friend with the required knowledge? This is exactly what the future web would look like and would do. Already, for example, if you put the query on Google about *what's the temperature*

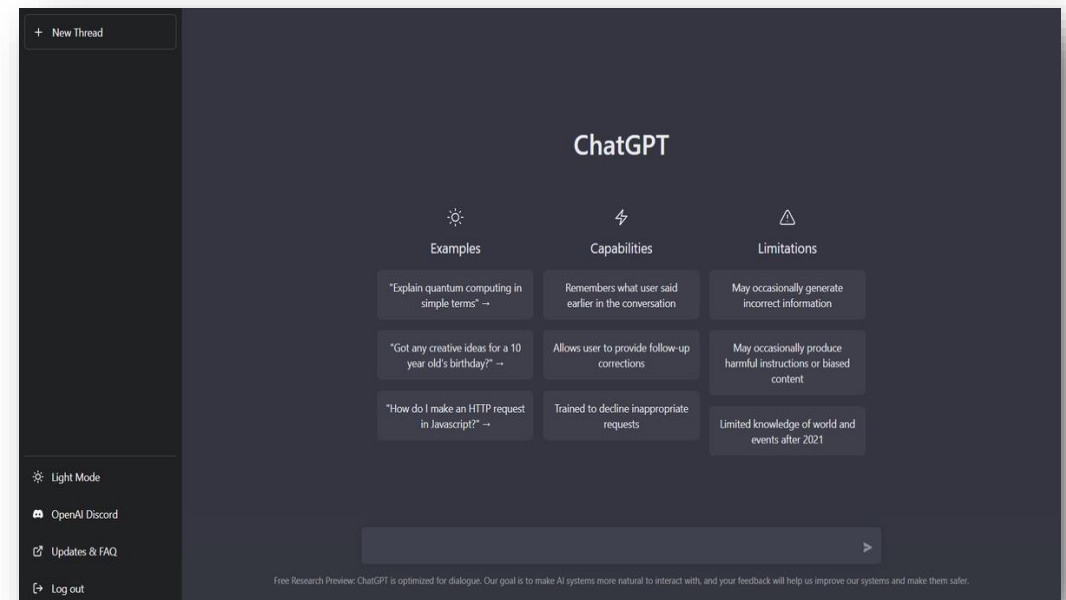
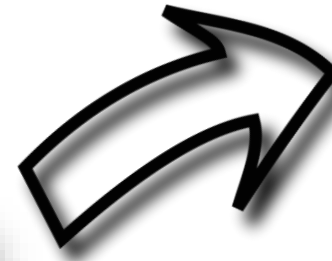
Los datos en las Redes Sociales

La velocidad a la que avanzamos

However, today only a fixed number of questions can be answered, as there is big risk of inferring wrong answers on multiple questions. The future of the internet will be an extension to the World Wide Web, which is the semantic web. The term was coined by the creator of the World Wide Web, Tim Berners-Lee. The semantic web is a complex concept on a simple idea of connecting entities (URLs, pages, and content) on the web through relations, but the underlying implementation is difficult at scale, due to the sheer volume of entities present on the internet. New markup languages called **Resource Description Framework (RDF)** and **Web Ontology Language (OWL)** will be used to create these links between pages and content, based on relations. These new languages will allow creators of content to add meaning to their documents, which machines could process for reasoning or inference purposes, allowing automating many tasks on the web. Our book is not about

Los datos en las Redes Sociales

De búsqueda... a conversación





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Ejercicio por grupos:

Búsqueda manual...





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

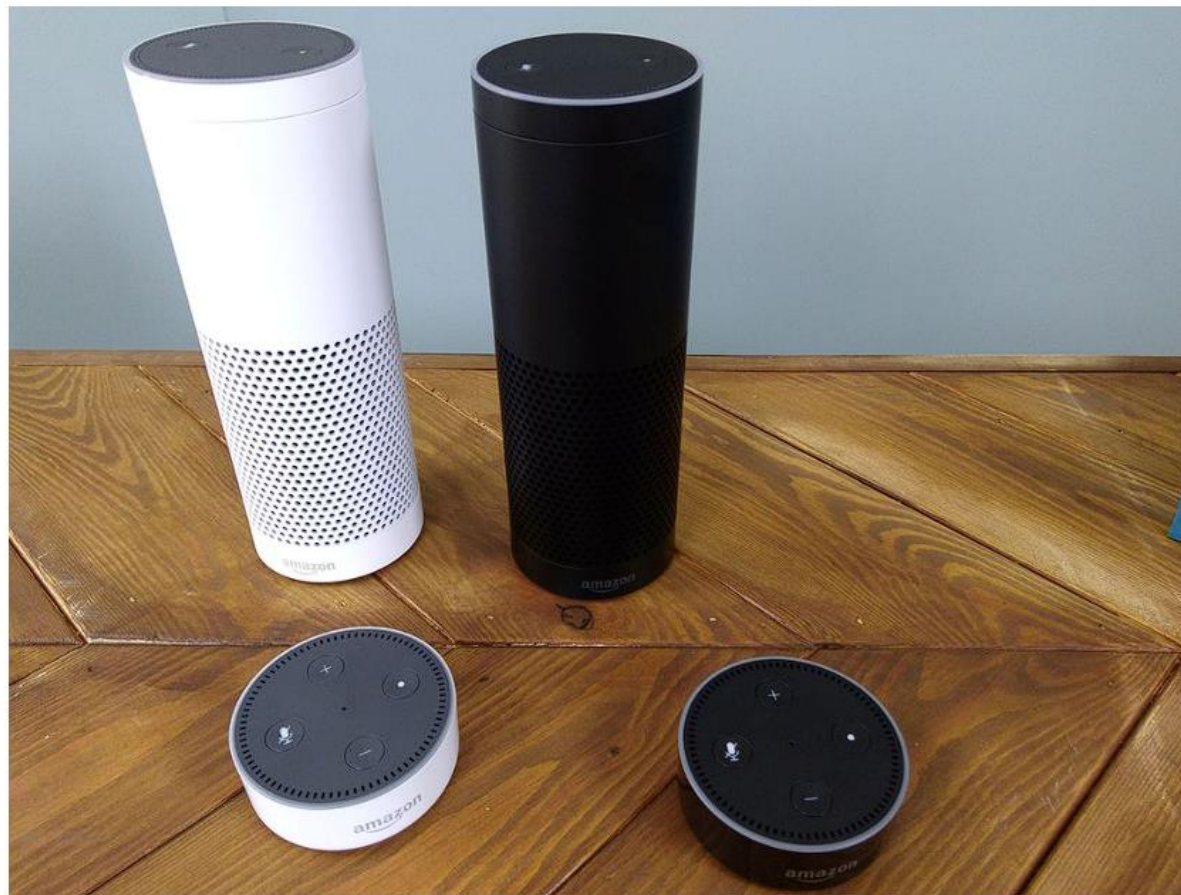
“Aquella vez que Alexa fue llamada como testigo para resolver un asesinato...”



UN CASO, DE MOMENTO, SIN RESOLVER

'Alexa, ¿quién la ha matado?': el altavoz de Amazon, último testigo de un asesinato

Una mujer muerta en su casa y sin pistas de qué ha podido pasar. ¿Puede el asistente inteligente de Amazon echar luz a un caso, aparentemente, irresoluble?



Amazon Echo, el altavoz inteligente que podría resolver un asesinato. (Reuters)

Por **Rubén Rodríguez**

05/11/2019 - 10:01 Actualizado: 05/11/2019 - 17:39



Los datos en las Redes Sociales

¿Podrías responder a estas preguntas?

- ¿Cuándo y dónde ocurrió? ¿Cómo se llamaba la víctima?
- ¿Se ha resuelto el crimen?
- ¿Cuántas veces ha ocurrido esto?
- ¿Graba Alexa todo lo que ocurre en una casa?
- ¿Es legal que Amazon entregue las grabaciones ocurridas en mi casa?
- ¿Funcionalidad o *bug*?
- **¿Qué estrategias podría derivar Amazon u otras empresas para crear o mejorar su negocio?**





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

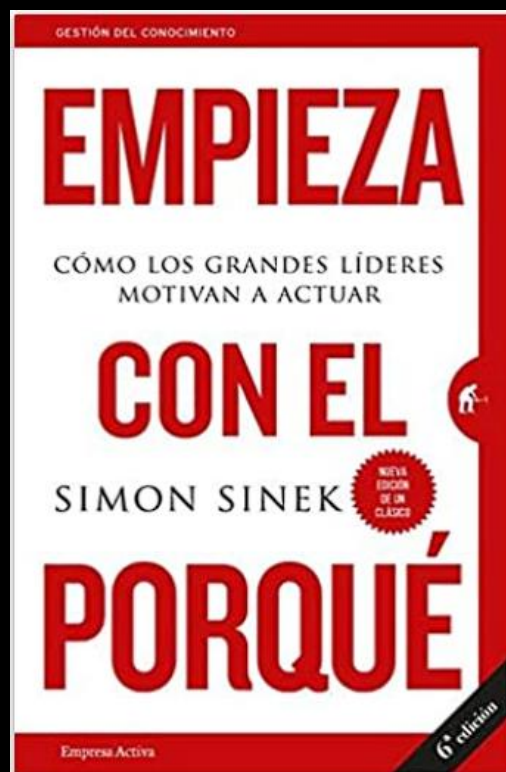
Comprensión del proceso



Comprensión del proceso

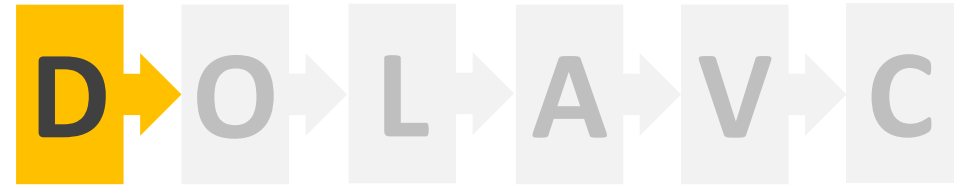
Proceso de Explotación de Datos Sociales





Comprensión del proceso

Definición del problema

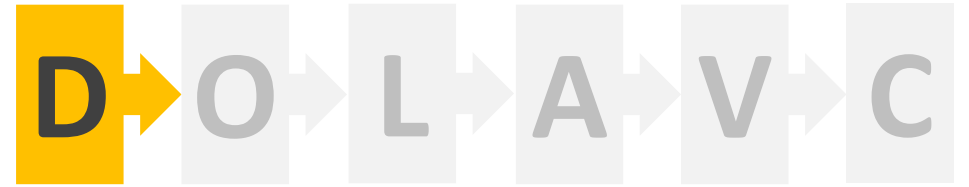


Comprender el problema que quieres resolver es clave para realizar un análisis de datos sociales:

- Seleccionar los datos adecuados y sus fuentes
- Limpiar con un propósito
- Generalizar vs concretar (pregunta general → conclusión no accionable)
 - “Quiero encontrar tendencias en Twitter”, o
 - “Quiero saber la reacción más común de los clientes de habla inglesa ante mi marca en Twitter a lo largo de la última semana, y quiero saber si ha sido positiva o negativa.”
- Estrategia equivocada implica enormes gastos de tiempo y recursos

Comprensión del proceso

¿Cómo se define un objetivo?



SMART





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

Ejercicio:

*Definición de objetivos
(marca + 3 objetivos)*





COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

¡Muchas gracias!

