







Inteligencia Artificial en Redes Sociales: la IA 2.0

Carlos A. Iglesias
Univ. Politécnica Madrid

Santiago de Compostela, 2010

Índice

- Inteligencia Colectiva
- Recomendación colectiva
- Minería de Opiniones
- Semántica: SIOC, Linked Data
- Mercados predictivos





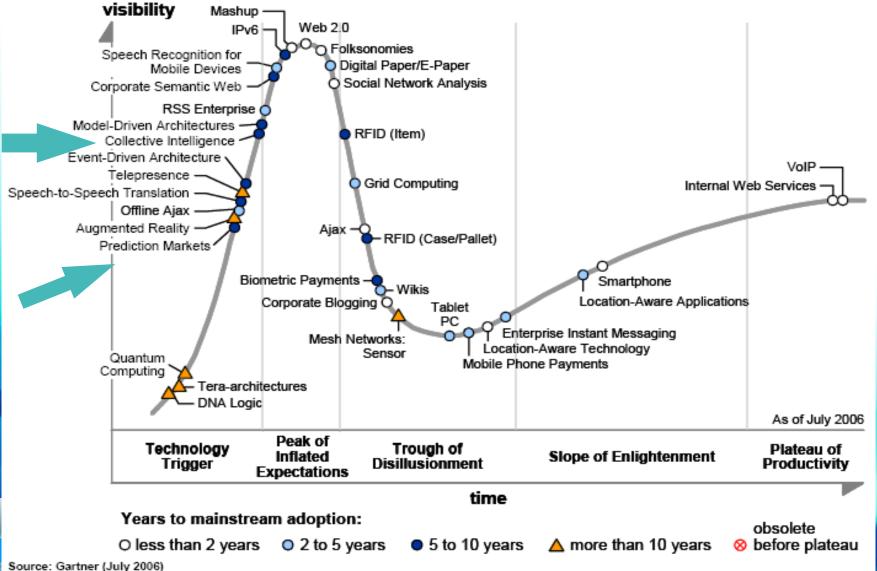
Índice

- Inteligencia Colectiva
- Recomendación colectiva
- O Minería de Opiniones
- Semántica: SIOC, Linked Data
- O Mercados predictivos



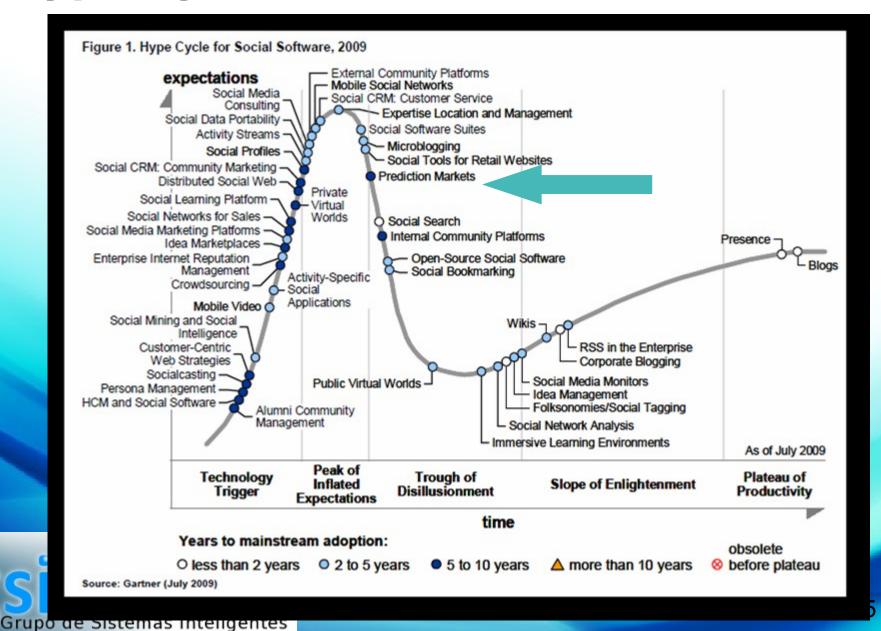


Hype Cycle for Emerging Technologies 2006



0

Hype Cycle for Social Software 2009



¿Qué es Inteligencia Colectiva?

- Grupos de individuos que realizan acciones colectivas que parecen inteligentes
- Concepto "antiguo"
 - Hormigas
 - Civilización humana
 - Naciones
 - Corporaciones
 - Tribus cazadores





¿Por qué es relevante hoy?

- Internet facilita que las personas puedan trabajar de forma colectiva
- La web2.0 ha mostrado cómo la web puede canalizar esta inteligencia colectiva





Caso Inteligencia colectiva: Spam

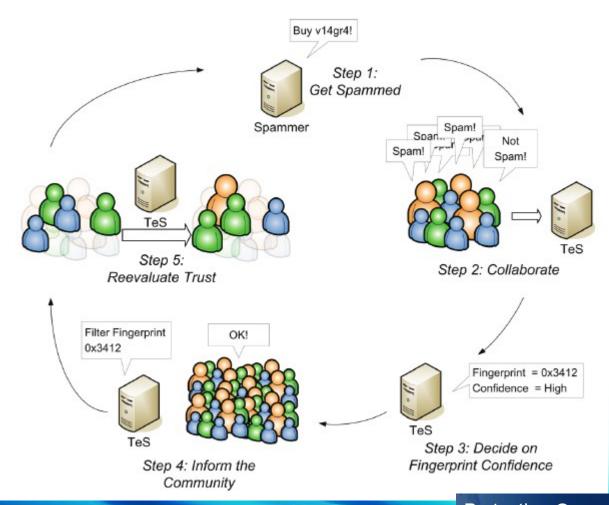
Cloudmark: uso de un sistema de reputación para filtrar el spam



spam

Grupo de Sistemas Inteligentes

Caso inteligencia colectiva: Spam





Messaging security solutions for the world's most demanding email, mobile and social networks





✓ Optimized network services



Fuentes de inteligencia colectiva

Contenido generado por los usuarios

Comentarios, recomendaciones, discusiones, etiquetado, puntuación, relaciones sociales Tareas conjuntas

Mercados predictivos, Etiquetado de imágenes con juegos Filtrado de spam Desarrollo de software libre Contenido conjunto, p.ej. Wikipedia

Análisis

Tareas inteligentes

Realizar recomendaciones

Recomendar libros, usuarios, discos, ...

Descubrir grupos

Usuarios parecidos para una campaña, errores reportados agrupados, ...

Clasificar información

Determinar las categorías de un contenido

Extraer conocimiento y consultarlo

Opiniones de los usuarios, consultas complejas en la wikipedia, ...

Entender actividades colectivas

Cómo funciona la psicología de los inversores, cómo reaccionamos ante una evacuación, ...

Realizar predicciones

Cómo nos comportaremos ante una oferta, cómo nos comportaremos de forma colectiva



Técnicas inteligentes

Realizar recomendaciones

Algoritmos de recomendación

Descubrir grupos

Algoritmos de agrupamiento (*clustering*)

Clasificar información

Minería de datos, técnicas de razonamiento

Extraer conocimiento y consultarlo

Minería de textos, tecnologías semánticas

Entender actividades colectivas

Tecnologías de agentes inteligentes, simulación con agentes

Realizar predicciones

Minería de datos, enfoques estadísticos, técnicas de simulación



A SHORT HISTORY OF MODERNIST PAINTING" (DETAIL) MARK TANSEY

Índice

- Inteligencia Colectiva
- Recomendación colectiva
- O Minería de Opiniones
- Semántica: SIOC, Linked Data
- O Mercados predictivos





Recomendación

 Un sistema de recomendación ofrece al usuario un conjunto de elementos (items) ordenados conforme a sus



> Edward R. Tufte

\$48.00 \$34.56

Hardcover

> Edward R. Tufte

\$52.00 \$35.10

Hardcover

Code Magnets Kit

Inc O'Reilly Media

Misc. Supplies

\$14.99 \$11.69

Techniques

> John Resig

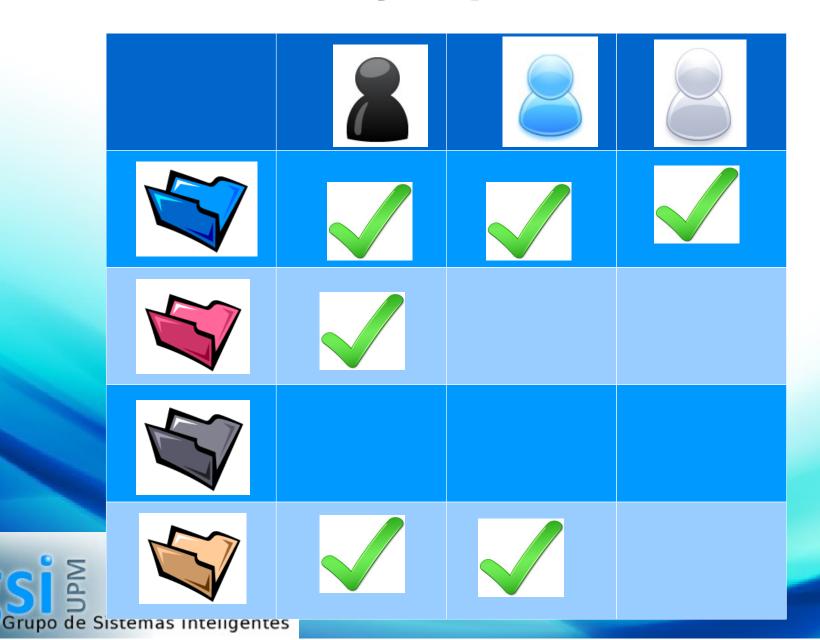
\$44.99 \$29.69

Paperback

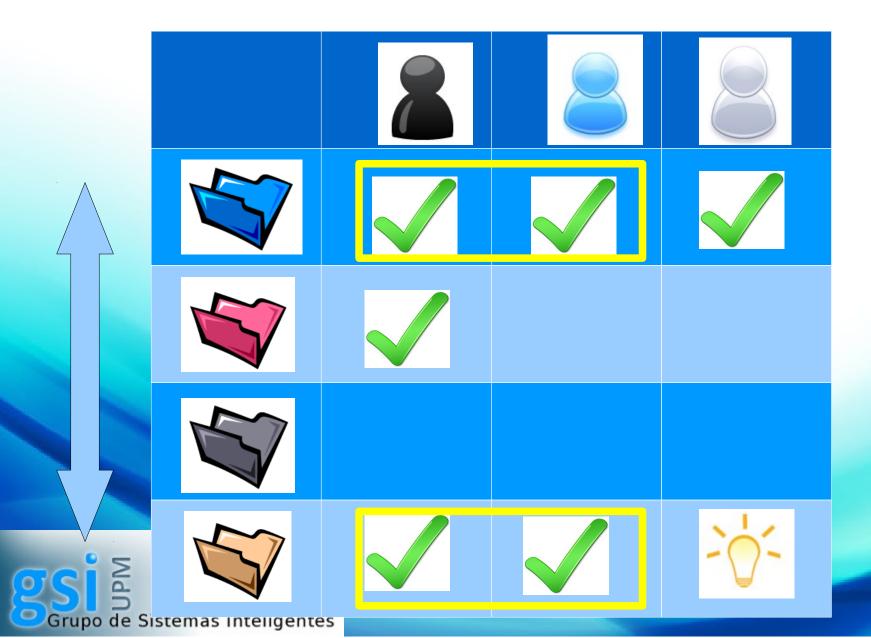
Filtrado colaborativo: idea

- Recomendamos "cosas" que les han gustado a personas con gustos parecidos
 - Item-a-Item
 - Persona-a-Persona

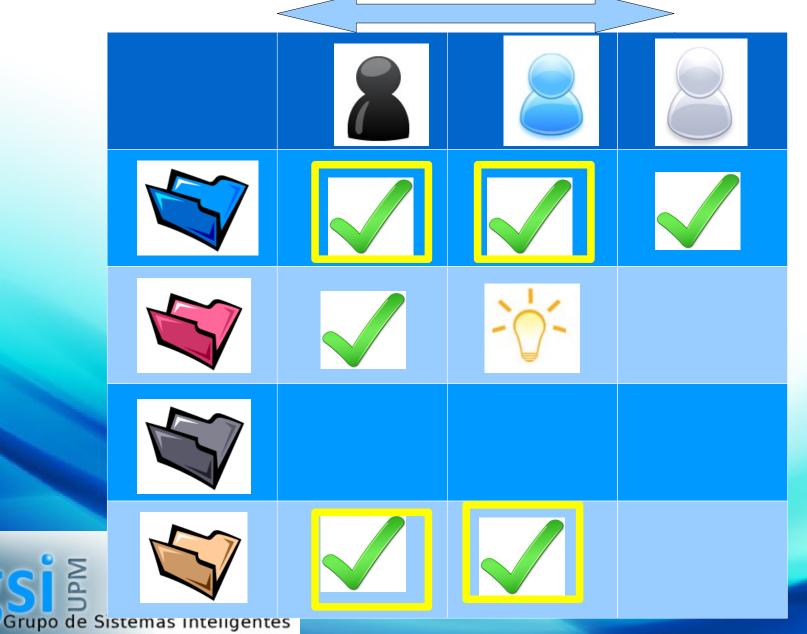
Ejemplo



Ejemplo Item-a-Item



Ejemplo Usuario-a-Usuario



Algunos detalles

- Normalmente normalizaremos los gustos de los usuarios. P. ej. 0..1
- Para calcular similitud de usuarios y/o items, calcularemos una distancia entre sus 'vectores de gustos' que mide la similitud



Otros algoritmos de recomendación

Basados en contenido

- Se emplea una descripción de perfil del usuario y del item (p.ej. Categoría, etc.)
- Se busca afinidad entre estas descripcioens

Contextuales

La recomendación cambia según el contexto

Sociales

- La red social del usuario influye
- Se usa la reputación de los usuarios y la confianza del usuario en otros



Índice

- Inteligencia Colectiva
- Recomendación colectiva
- Minería de Opiniones
- Semántica: SIOC, Linked Data
- O Mercados predictivos





Minería de Opiniones / Análisis de Sentimientos

 Área de investigación que intenta crear sistemas automáticos que procesen las opiniones humanas expresadas en textos



Minería de Opiniones

- Normalmente los motores de búsqueda buscan hechos
- Los hechos pueden ser expresados con palabras clave
- Sin embargo, los usuarios a menudo expresan opiniones
- ¿Qué tecnología podemos emplear para conocer estas opiniones colectivas?
- <u>o ¿Qué aplicaciones pueden tener?</u>



Aplicaciones de Minería de Opiniones

Empresas de Producto

- Conocer gustos de consumidores para vender más
- Impacto de una campaña
- Imagen de marca

Individuos

- Conocer opiniones de otros consumidores para comprar
- Opiniones políticas
 - Imagen en Internet

Anunciantes

Poner anuncio si hablan bien de ti o mal de tu competencia





Caso practico: Swotti (I)

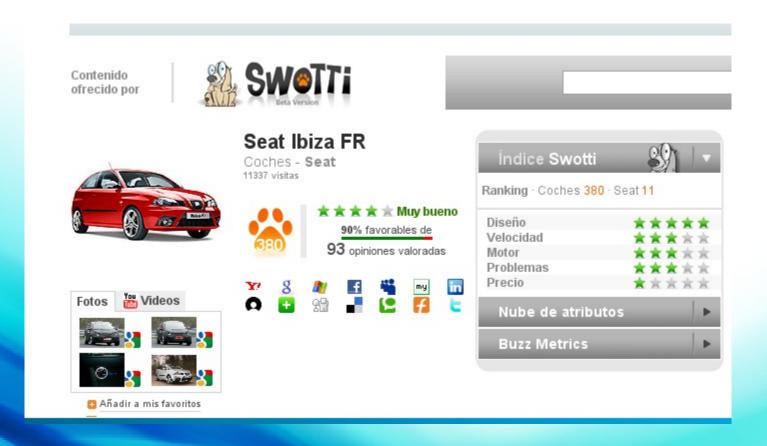
Coogle Q buscar

Tecnología Motor Destinos Juegos cine Música Literatura Personalidades Negocios

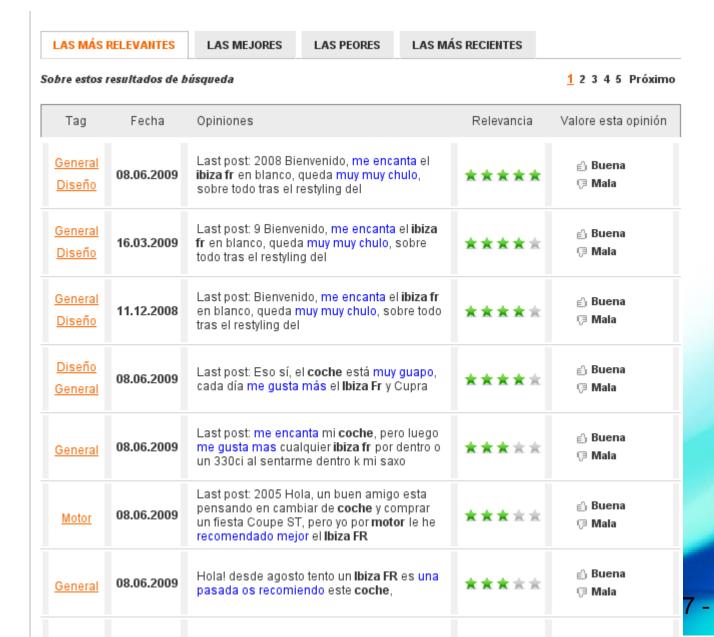




Caso práctico: Swotti (II)



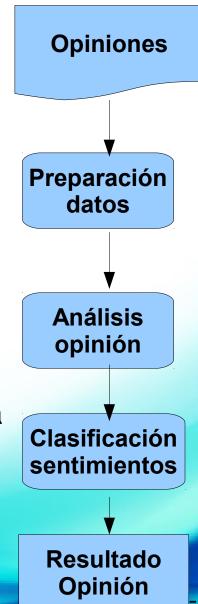
Caso práctico: Swotti (III)





Tecnología

- Paso 1. Recuperación de opiniones
 - Técnicas de crawling
 - Buscadores verticales y horizontales
- Paso 2. Extracción de opiniones
 - Parsers ad-hoc
- Paso 3. Extracción de rasgos y sentimiento de cada opinión
 - Etiquetado de partes de habla (función)
 - Etiquetado de sentimiento de cada palabra
 - Combinación de 'polaridades'





Opine

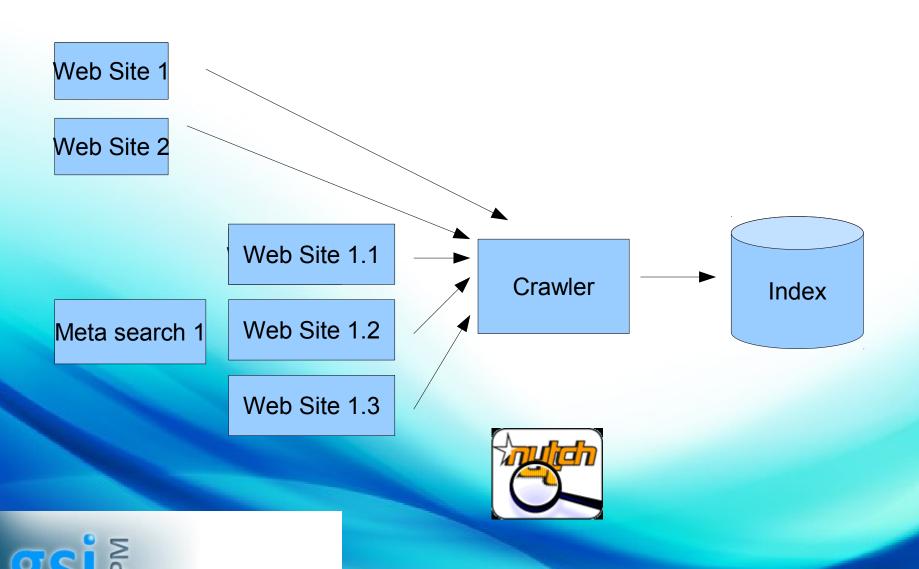
- Entrada: clase de producto(P), críticas(C)
- Salida: conjunto de tuplas [rasgo, ranking de opiniones]
- o analizaCríticas(C) → C'
- buscaRasgosExplícitos(P,C') → E (temperatura era alta)
- buscaOpiniones(C',E) → O
- agrupaOpiniones(O) → CO
- buscaRasgosImplícitos(CO,E) → I (hacía calor)
- ordenaOpiniones(CO) → RO
- Tuplas de salida(RO, E ∪ I)





Opine, Mining Product Reviews, Popescu et. al. 2005

Recuperación opiniones



Grupo de Sistemas Inteligentes

Preparar los datos

- "Limpiar los datos"
- Quitar etiquetas / datos que sobran

Separar las opiniones



Analizar cada opinión

- Etiquetar con diccionarios las opiniones
 - [ART] El [N] coche [V]consume [ADV]muy [ADV] poco.
- Etiquetar negaciones
- Extraer rasgos
 - Extraer nombres y sintagmas nominales (e.j. Velo
- Extraer "subjetividad"
 - Extraer adjetivos (también verbos, nombres, etc.)
 - Positivos: honesto, barato, importante,
 - Negativos: deshonesto, ladrón, caro, inseguro



Analizar sentimientos

Asignar polaridad



- Consumir poco (+)
- Correr poco (-)
- Combinar polaridad
- Existen diccionarios de polaridad: SemWordNet
 - Técnicas de aprendizaje automático
 - Basadas en textos previamente etiquetados, aprendizaje supervisado



Problemática

- Negaciones
- Subjetividad de las opiniones
- Diferentes enfoques de aprendizaje automático para clasificar revisiones, combinados con enfoques heurísticos
- Tratamiento de ironía, dobles negaciones, ...

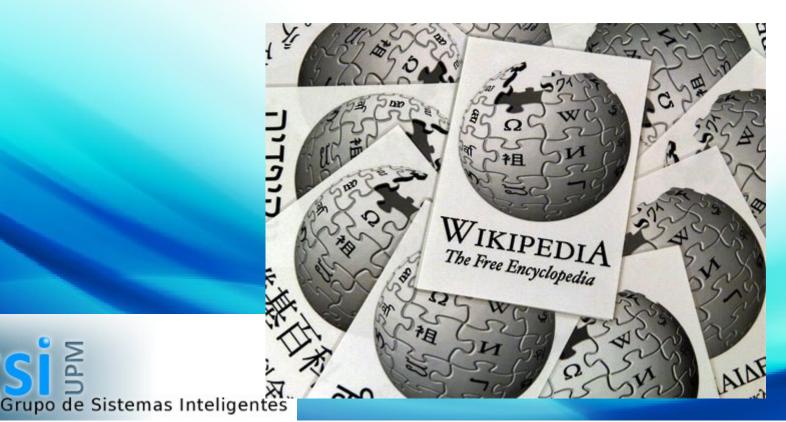
Índice

- Inteligencia Colectiva
- Recomendación colectiva
- O Minería de Opiniones
- Semántica: SIOC, Linked Data
- O Mercados predictivos



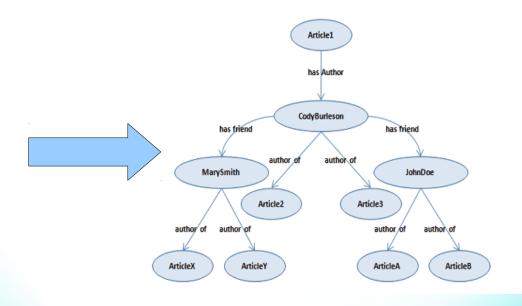
Wikipedia

- Ejemplo de inteligencia colectiva
- ¿Cuántos futbolistas que han ganado el mundial lo ganaron de juveniles?
- O → No está automatizado



Tecnología semántica



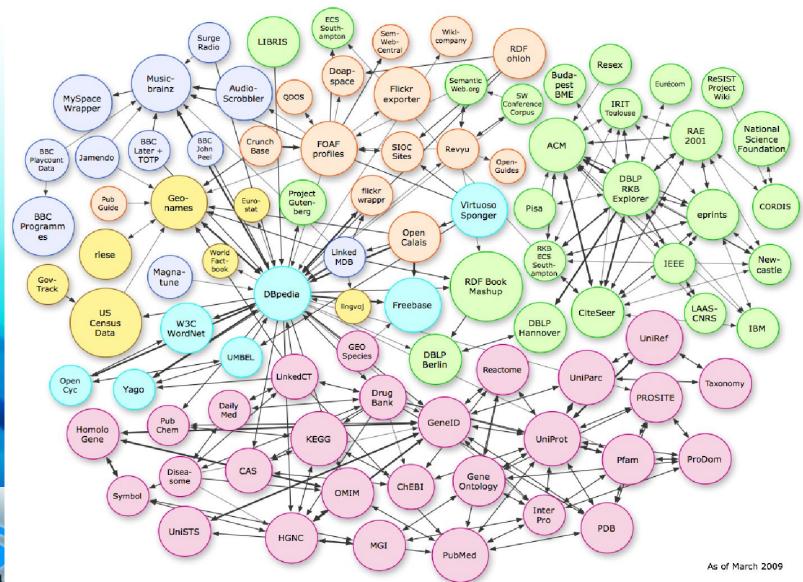


No estructurado → para personas

Estructurado → para máquinas



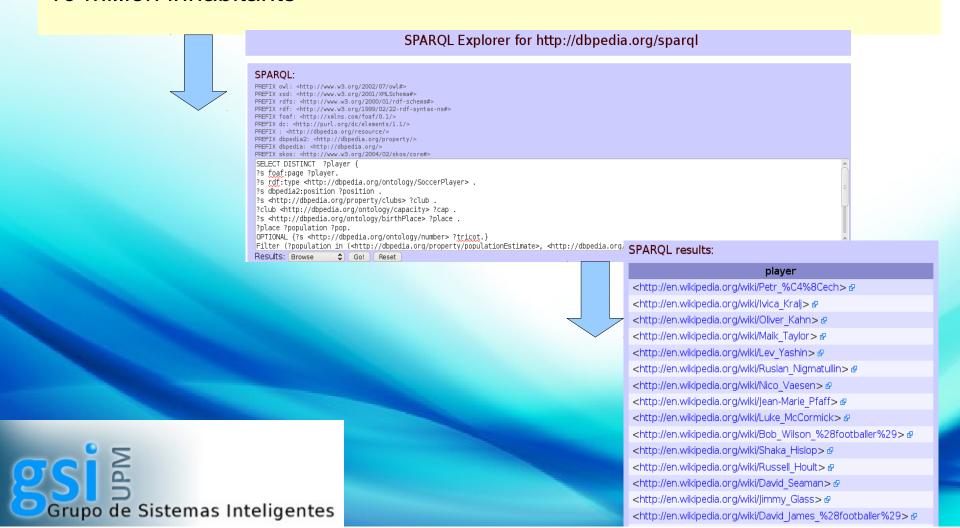
Linked Data



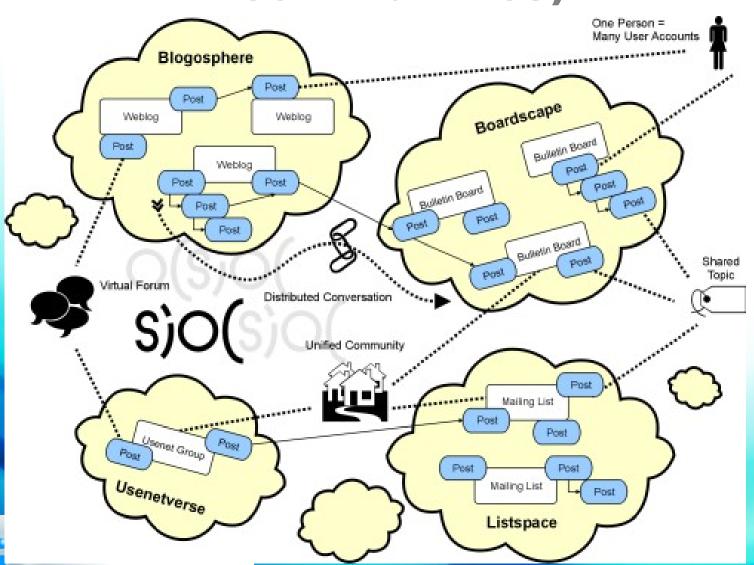


Ej. DBPedia

All soccer players, who played as goalkeeper for a club that has a stadium with more than 40.000 seats and who are born in a country with more than 10 million inhabitants



SIOC (Semantic Interlinked Online Communities)



Grupo de Sistemas Inteligentes

Índice

- Inteligencia Colectiva
- Recomendación colectiva
- O Minería de Opiniones
- Semántica: SIOC, Linked Data
- Mercados predictivos



Mercados predictivos

- Aprovechar las predicciones de los usuarios para predecir el futuro
- Mejores resultados que los dados por los expertos





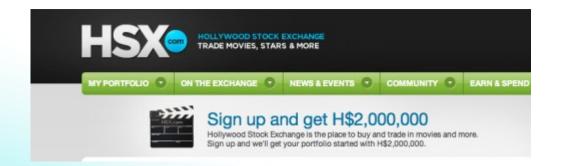
predictalot.yahoo.com agrees with Paul the octopus: #ESP has 62% chance of winning #WorldCup

¿Qué es un mercado predictivo?

- Las acciones son los posibles resultados de un suceso futuro(España gane el mundial, Obama gane elecciones, ...)
- Se utiliza un mecanismo de mercado de subasta
- El precio de una acción representa la probabilidad de que suceda ese evento futuro. P. ej. si el valor de 'España gana' es 0.75 céntimos (rango 0-100 céntimos)→ probabilidad de ganar del 75%
- El mercado agrega las creencias de los participantes



Mercados para éxito de Películas



Hollywood Stock Exchange – 96.5% de precisión



Twitter Buzz – 97.3% de precisión



Twitter Buzz - 97.3% de precisión!

Twitter Buzz Predicts Box-Office Success Better Than Hollywood Stock Exchange

Social media predictive power might also extend beyond box office success

By Jeremy Hsu Posted 04.02.2010 at 11:47 am 3 Comments



#Nowplaying

Happy Good Friday

#followfriday

Tapping into the wisdom of the crowds to forecast future trends has served prediction markets well for years, but Twitter might be even more effective than even the biggest and most widely used market, the Hollywood Stock Exchange. In a recent study, watching tweets among Twitter users allowed HP Labs researchers to predict box office figures better than the Hollywood Stock Exchange, Fast Company reports.

Sitaram Asur and Bernardo Huberman of HP Labs kept track of movie mentions on among 2.9 million tweets from 1.2 million users for three months. Their sample focus included 24 movies such as box office king *Avatar* and *Twilight*: New Moon.

For opening weekend, their computer model monitored the rate of tweets near a movie's release date and also factored in the number of theaters showing the flick. That allowed the model to predict the opening weekend revenues with 97.3 percent accuracy, compared to the Hollywood Stock Exchange's 96.5 percent accuracy.

Twitter Box Office Prediction Will titans clash for "Clash of the Titans?" Let's ask Twitter

Tapping into the wisdom of the crowds to forecast future trends has served prediction markets well for years, but Twitter might be even more effective than even the biggest and most widely Grupo de Sistemas inteligentes

IOWA Market

Iowa market heavily favors Obama



By ANNA LOTHSON - THE DAILY IOWAN | 7/16/08 11:15 AM EST

Text Size - + reset







Investors believe there is a 64 percent probability that Obama will win the popular vote in the general election.

Photo: AP

The Iowa Electronic Markets' presidential contest gives Barack Obama a 28-point lead over John McCain, while a Gallup tracking poll released Monday has Obama up by only 3 percent.

The Democratic contract on the winner-take-all market traded at 64.3 cents, according to its website. This figure shows that investors believe there is a 64 percent probability that Obama will win the popular vote in the general election.

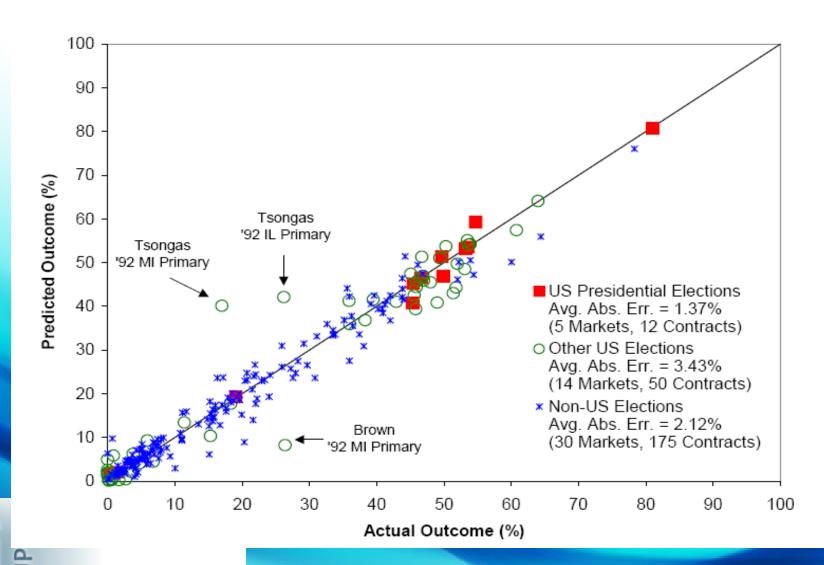
The Gallup poll had Obama up 46 to 43 percent.

Polls and, increasingly, markets have been used to judge the outcomes of elections. This year's

difference means that investors seem to like Obama, while polls of likely voters show they are considering McCain more.



IOWA Market



Grupo de Sistemas Inteligentes

Mercados para empleados

HP, IBM, Google...

- Microsoft: cuándo fundinará una funcionalidad
- Google: valor de acciones de la compañía, fecha de lanzamiento de un producto
- HP: Previsión de ventas





ABMS (Agent Based Modeling and Simulation)

Un agente es

- Un individuo con un conjunto de características
- Se comporta conforme a un conjunto de reglas predefinidas, y tiene capacidad para tomar decisiones, comunicarse con otros agentes y responder al entorno
- Los agentes pueden ser heterogéneos



Aplicaciones

- Flujos: evacuación, tráfico, clientes
- Mercados: bolsa, compra en internet, simulación estratégica
- Organizaciones: riesgo de operación y diseño organizativo
- Difusión: difusión de la innovación y dinámica de las opiniones



Conclusiones

- La inteligencia colectiva emerge de las acciones de los individuos
- Los sistemas inteligentes ofrecen diversas líneas de investigación para explotar esta inteligencia colectiva en nuestras aplicaciones
 - Sistemas de recomendación Minería de opiniones
 - Tecnologías semánticas
 - Tecnologías de agentes



Gracias por la atención ¿Preguntas?

