



Sistemas de recuperación de información y de recomendación: Sistemas de Recomendaciones (Recommender Systems-RecSys)

Enrique Herrera Viedma

Dpto. De Ciencias de la Computación e I.A

Granada, Marzo de 2018



Outline

1. El problema de acceso a la información
2. RecSys y la Evolución de la Web
3. Qué son los RecSys ?
4. Estructura y funcionamiento de los RecSys: Tipos
5. Tendencias Futuras

● ● ● Problema de Acceso a la Información I

- ◆ **LA WEB:** Principales fuentes de generación y transmisión de información.
- ◆ Principal problema de la WEB: Crecimiento exponencial y descontrolado

Información



Usuarios



● ● ● Problema de Acceso a la Información II

- ◆ **Consecuencia:** Dificultad de los usuarios para acceder a la información
- ◆ **Solución:** Desarrollo de sistemas automáticos de búsqueda y acceso a la información en la Web



● ● ● | Problema de Acceso a la Información III

Dos categorías de sistemas:

1. Basados en RI tradicional-Consultas de términos
2. Basados en Filtrado de Información-Perfiles de usuario

● ● ● Problema de Acceso a la Información IV

1. Basados en RI tradicional-Consultas de términos

Características de la información: **Keywords**

BUSCADORES WEB
Search Engines



● ● ● Problema de Acceso a la Información V

2. Basados en Filtrado de Información-Perfiles de usuario

Características del usuario: **User Profiles**

SISTEMAS DE
RECOMENDACIONES
Recommender Systems
(RecSys)





Problema de Acceso a la Información VI

From Search to Recommendation

"The Web is leaving the era of search and entering one of discovery. What's the difference?"

Search is what you do when you're looking for something. **Discovery** is when something wonderful that you didn't know existed, or didn't know how to ask for, finds you." – CNN Money, "The race to create a 'smart' Google"



Problema de Acceso a la Información VII

The value of recommendations

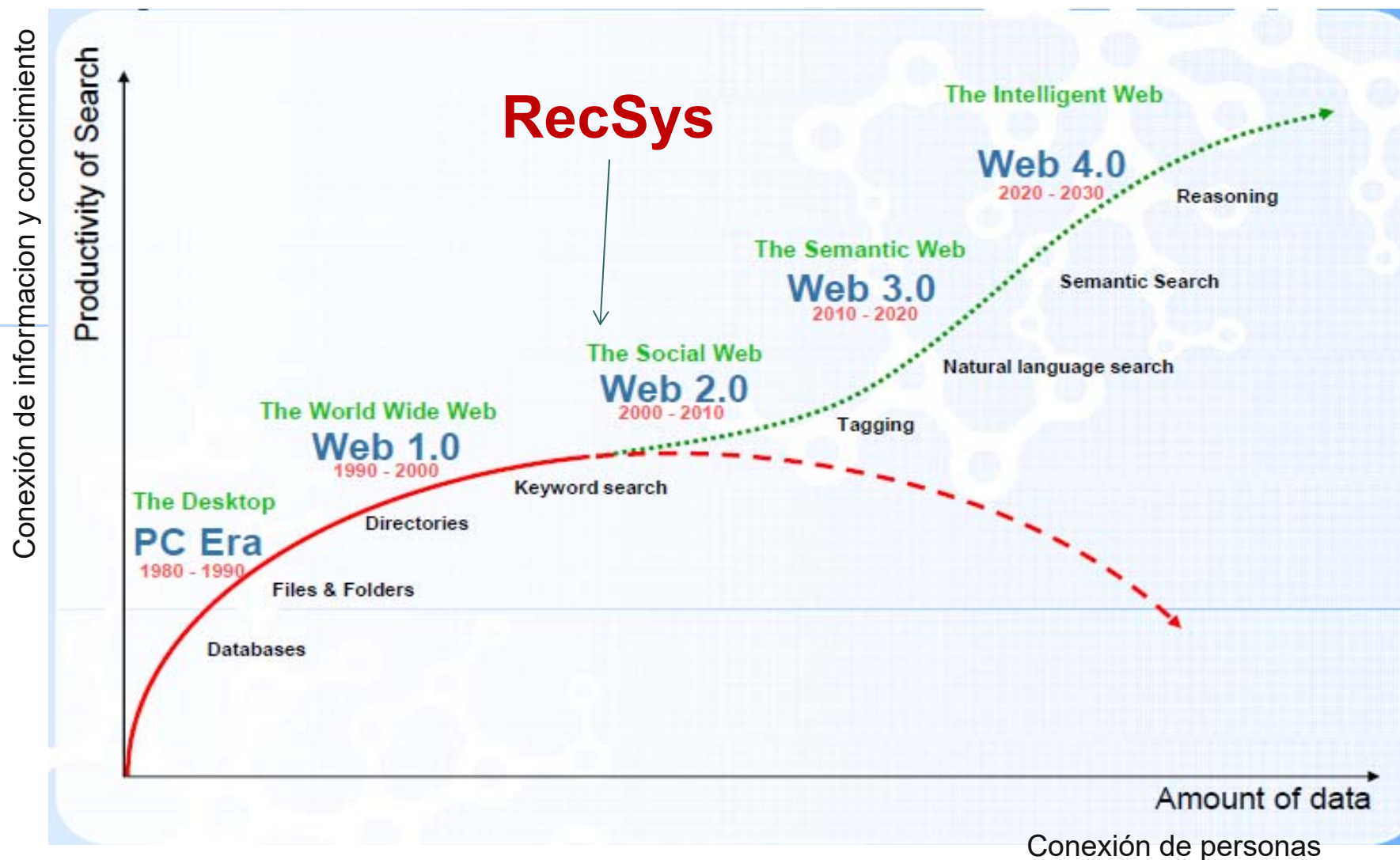
- Netflix: 2/3 of the movies watched are recommended
- Google News: recommendations generate 38% more click-throughs
- Amazon: 35% sales from recommendations



Outline

1. El problema de acceso a la información
2. RecSys y la Evolución de la Web
3. Qué son los RecSys ?
4. Estructura y funcionamiento de los RecSys: Tipos
5. Tendencias Futuras

Los RecSys y la Evolución de la Web I



● ● ● | Los RecSys y la Evolución de la Web II

◆ Web 1.0 (1990-2000): Informative Web



- La Web es un portal de información.
- Los usuarios son consumidores de información, sin interacción.
- Solamente la información está conectada.
- **Buscadores basados en RI tradicional:** Altavista



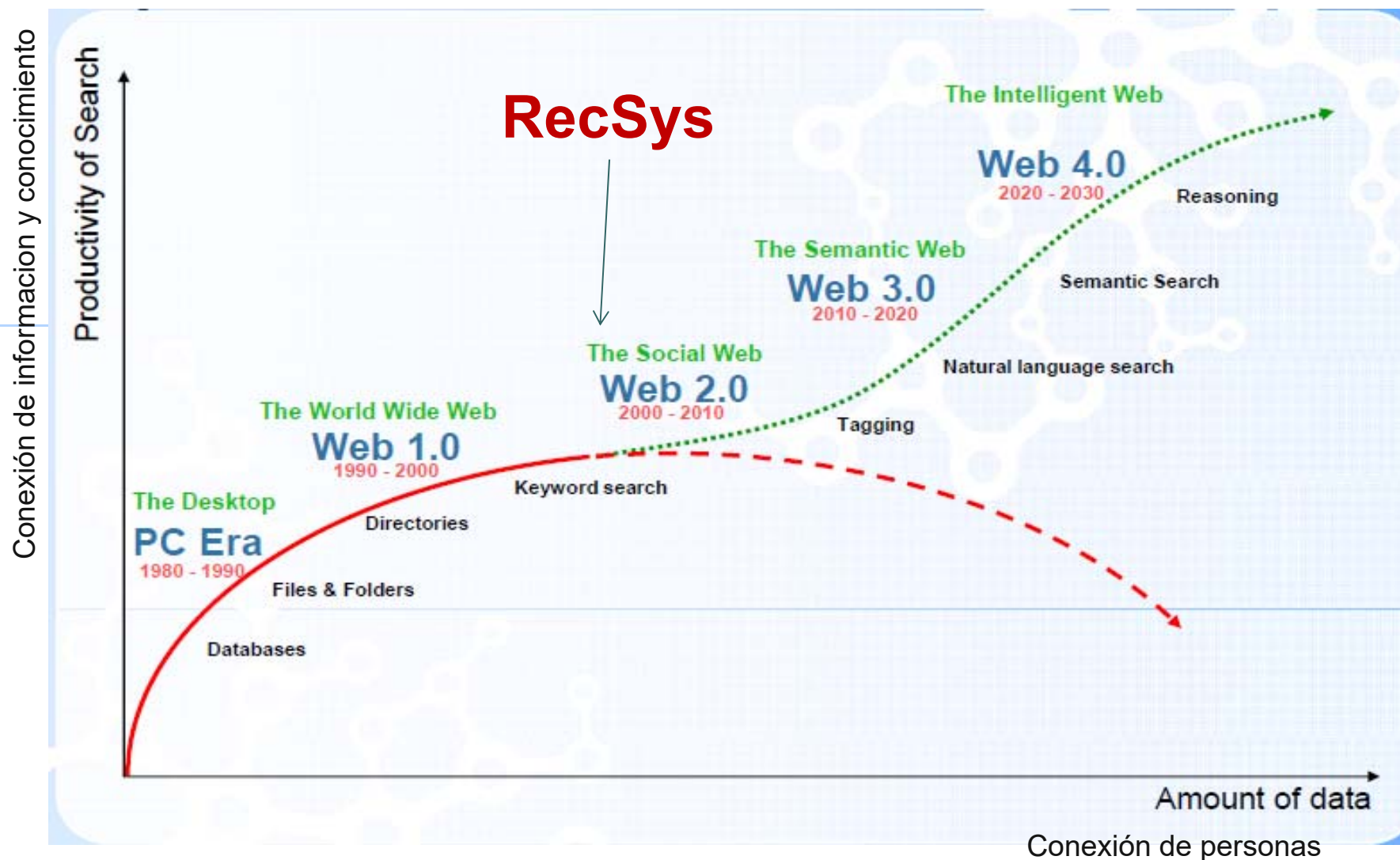
◆ Web 2.0 (2000-2010): Social Web



- Los usuarios es lo importante en la Web.
- Los usuarios son consumidores, proveedores y recomendadores de información.
- Las personas están conectadas: redes sociales.
- Cierta grado de personalización .
- **Buscadores y RecSys basados en Inteligencia Social:** Google , Amazon,
- **Herramientas Web 2.0:** wikipedia, Youtube, Flickr



Los RecSys y la Evolución de la Web I





Los RecSys y la Evolución de la Web III



Web 3.0 (2010-2020): Semantic Web



- La Web es incorpora semántica en los datos.
- Todo en la Web está conectado: [Internet of the Things](#)
- Información y conocimiento personalizado.
- [Buscadores y Recsys semánticos](#): Gnoss, Sirpross

GNOSS



Sirpross



Web 4.0 (2020-2030): Intelligent Web



- La Web con habilidades de razonamiento
- Computación ubicua o omnipresente: [Things that Think](#)
- Grado máximo de personalización .
- [Buscadores y RecSys basados en agentes inteligentes](#)





Outline

1. El problema de acceso a la información
2. RecSys y la Evolución de la Web
3. Qué son los RecSys ?
4. Estructura y funcionamiento de los RecSys: Tipos
5. Tendencias Futuras

Qué son los Sistemas de Recomendaciones?

- ◆ **Definición:** Un RecSys es un tipo de sistema de filtrado de información que asiste a los usuarios en los procesos de acceso a la información, mediante **la predicción** y **recomendación** de items de información que pueden ser de interés.



Qué son los Sistemas de Recomendaciones?

◆ Fundamento de los RecSys:

Los RecSys tratan de ser una alternativa al *proceso social de recomendación*, es decir, ese acto habitual que todos practicamos al recurrir a las opiniones de conocidos o expertos cuando tenemos que tomar una decisión para adquirir algo sin tener la suficiente información para ello, generalmente decisiones sencillas en el entorno de la vida cotidiana como qué libro leer, qué película ver o qué música comprar.



● ● ● Qué son los Sistemas de Recomendaciones?

The “Recommender problem”

Estimate a utility function
to predict how
a user will like an item.

Qué son los Sistemas de Recomendaciones?

The “Recommender problem”

- $C := \{\text{users}\}$
- $S := \{\text{recommendable items}\}$
- $u :=$ utility function, measures the usefulness of item s to user c ,

$$u : C \times S \rightarrow \mathbb{R}$$

where $R := \{\text{recommended items}\}$.

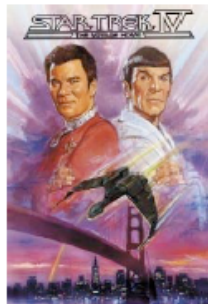
- For each user c , we want to choose the items s that maximize u .

$$c \in C \quad s'_c = \operatorname{argmax}_s u(c, s)$$

Qué son los Sistemas de Recomendaciones?

- ◆ Concepto básico: **La personalización** es un concepto clave de los RecSys.

A good recommendation



is relevant to the user: *personalized*



● ● ● | Qué son los Sistemas de Recomendaciones?

- ◆ Concepto básico: **La diversidad** es un concepto clave de los RecSys.

A good recommendation

- is diverse:

it represents all the possible interests of one user



Qué son los Sistemas de Recomendaciones?

- ◆ Concepto básico: **La diversidad** es un concepto clave de los RecSys.

Diversity in Recommendation (II)

- Some good music recommendations:

						
Wrecking Ball	Not your Kind of People	Like a Prayer	Choice of Weapon	Sweet Heart Sweet Light	The Light the Dead See	Little Broken Hearts
B. Springsteen	Garbage	Madonna	The Cult	Spiritualized	Soulsavers	Norah Jones

- Different artists and genres.
- Not similar between them.
- These are much better recommendations!

Qué son los Sistemas de Recomendaciones?

- ◆ Más allá de la Web búsqueda: Los RecSys nos adentran en la **Web del descubrimiento**.

“los usuarios ya no buscan por algo en concreto, si no que esperan a descubrir cosas que ni siquiera sabían que existían, o no sabían cómo realizar la consulta para encontrar”

- ◆ **SERENDEPIA**: Encontrar cosas que te interesan buscando algo no relacionado.



Qué son los Sistemas de Recomendaciones?

A good recommendation

- Does not recommend items the user already knows or would have found anyway.
- Expands the user's taste into neighboring areas.

Serendipity = Unsought finding



● ● ● | Los RecSys y el E-Commerce

◆ OBJETIVOS:



Selecciona el producto que maximiza el valor, tanto para el comprador como para el vendedor en un momento determinado

Ventajas para el eCommerce



Aumento de ventas



Aumento de la fidelización

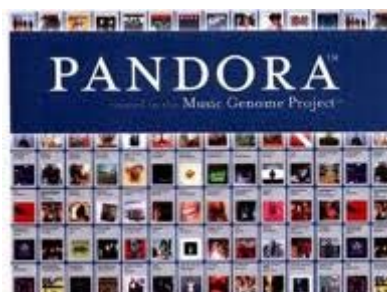


Diferenciación con la competencia

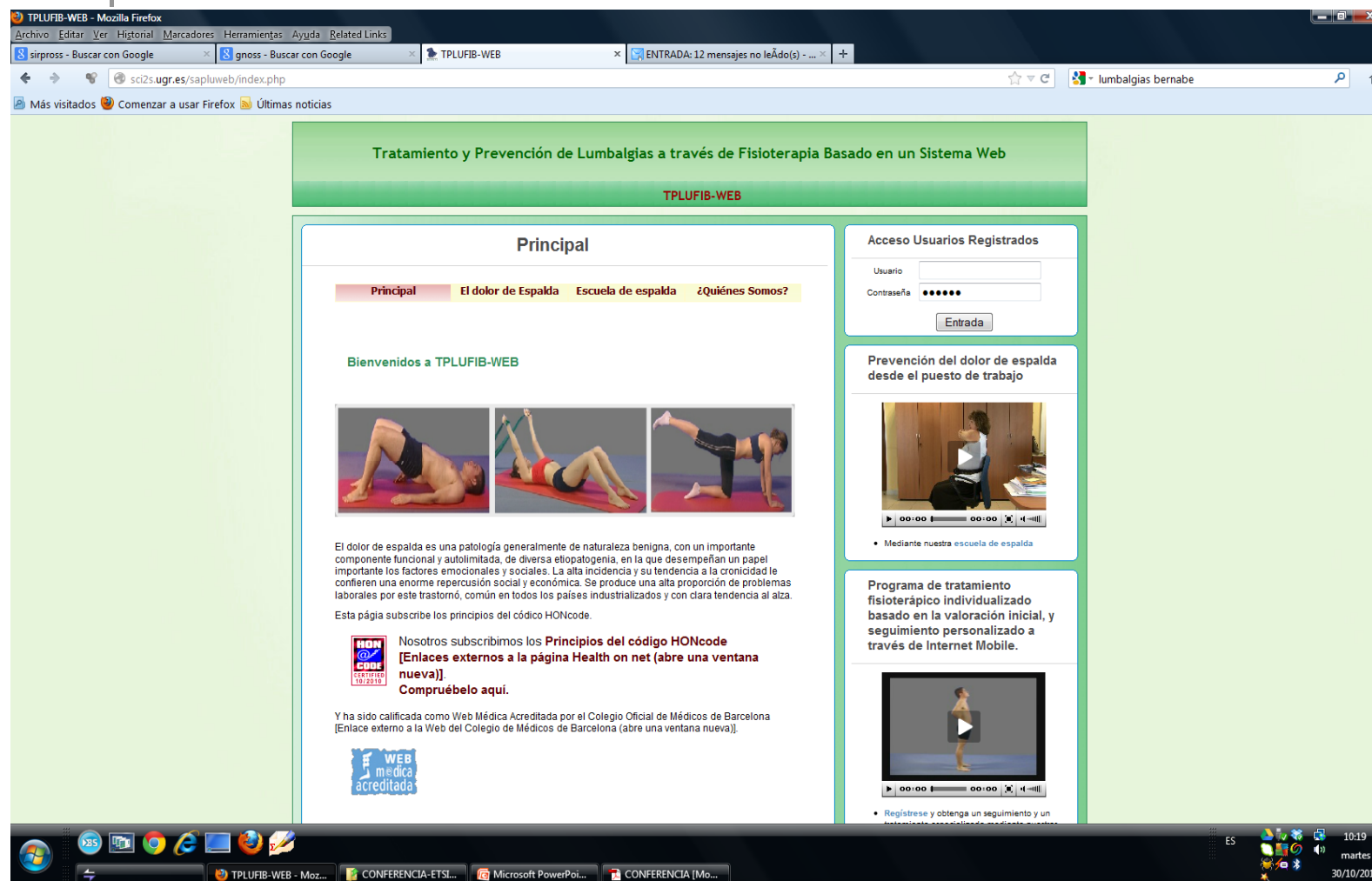
Según Brent Smith, director de personalización de Amazon:

“las recomendaciones personalizadas son la base del por qué las ventas online ofrecen tantas oportunidades”

● ● ● Ejemplos de populares RecSys



Ejemplo propio de RecSys en e-Salud



<http://sci2s.ugr.es/sapluweb/index.php>

Diferencias entre los SRI y los RecSys

	SRI	RecSys
<i>Funcionamiento</i>	On-line	Off-line
<i>Representación de las necesidades de los usuarios</i>	Mediante consultas puntuales	Mediante perfiles
<i>Conocimiento sobre los usuarios</i>	No dispone de conocimientos sobre ellos	En los perfiles se almacenan sus características



Outline

1. El problema de acceso a la información
2. RecSys y la Evolución de la Web
3. Qué son los RecSys ?
4. Estructura y funcionamiento de los RecSys: Tipos
5. Tendencias Futuras

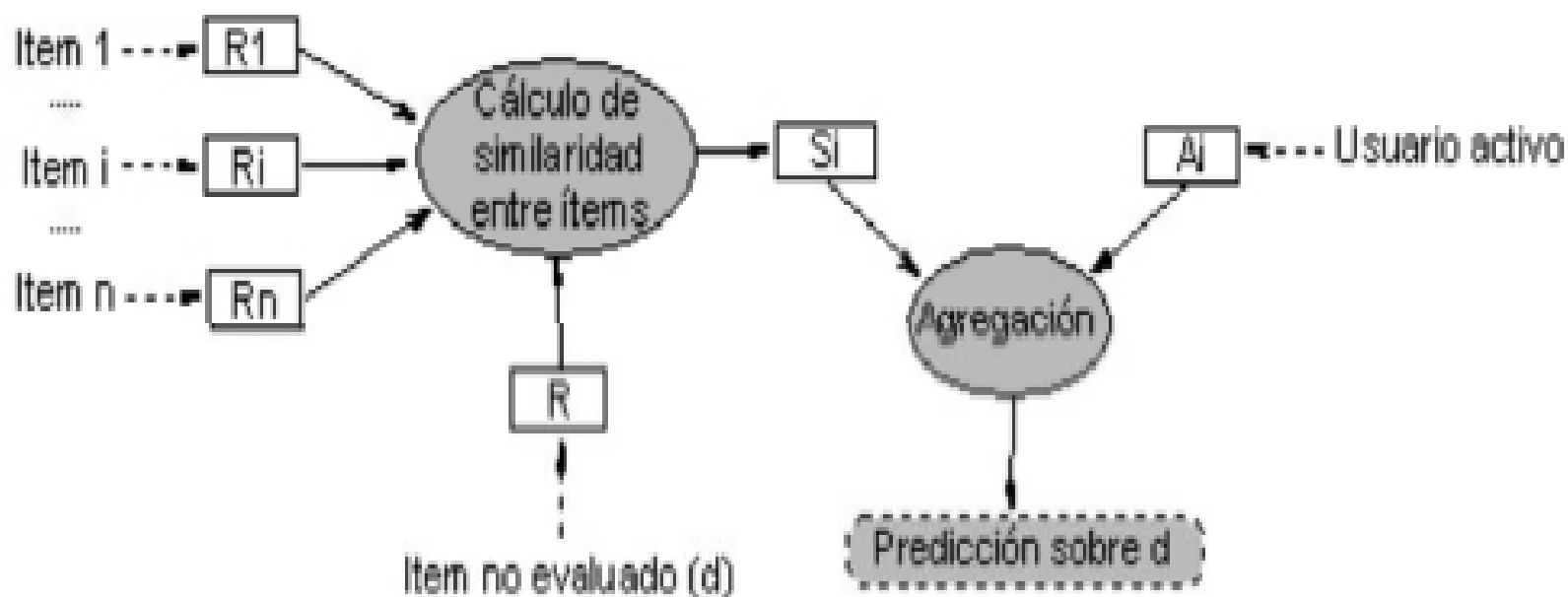
Estructura de los RecSys: Componentes

- ◆ INPUT:
 - ❖ Perfil del usuario activo
 - ❖ Perfiles de los colaboradores
 - ❖ Representación de los ítems a recomendar
 - ❖ Ratings dados a los ítems (user feedback)
- ◆ OUTPUT:
 - ❖ Sugerencia o lista de sugerencias
 - ❖ Valor de predicción del grado de satisfacción.
 - ❖ Visualizar valoraciones individuales
- ◆ TECNICA DE PERSONALIZACIÓN:
 - ❖ Sin personalización
 - ❖ Personalización efímera
 - ❖ Personalización persistente
- ◆ METODO DE GENERACIÓN DE RECOMENDACIONES:
 - ❖ RecSys basados en Contenido
 - ❖ RecSys colaborativos
 - ❖ RecSys híbridos

● ● ● | Tipos de RecSys: Basados en Contenido

- ◆ Si al usuario U le gusta el ítem A, y este ítem es similar al ítem B, entonces al usuario U es probable que le guste el ítem B.
- ◆ Perfil del usuario activo es representado en función de los ítems que previamente dicho usuario a probado/comprado y valorado mediante ratings.
- ◆ La recomendación o valor de predicción se obtiene usando las características de la representación de los ítems y mediante el proceso de cálculo de similaridad + agregación.

Tipos de RecSys: Basados en Contenido



● ● ● Tipos de RecSys: Colaborativos

- ◆ Si al usuario U le gusta el ítem A, y al usuario V le gustan los ítems A B, entonces al usuario U es probable que le guste también el ítem B.
- ◆ Perfil de un usuario es representado mediante el **patrón de su comportamiento**, el cual puede ser dado por sus ratings, compras o comentarios o por características propias del individuo.
- ◆ Normalmente, los perfiles de usuario e ítems son representados mediante patrones de valoraciones numéricos: Matriz de valoraciones:

	Item 1	Item N
User 1	-	4
.....
User M	5	6

- ◆ La recomendación o valor de predicción se obtiene **usando los patrones de comportamiento** y mediante el proceso de **cálculo de similaridad + agregación**.

Tipos de RecSys: Colaborativos



Tipos de RecSys: Colaborativos

Hay dos posibilidades para generar la recomendación: Identificar grupos de usuarios o ítems afines: se mide el grado de similitud entre usuarios o ítems atendiendo a las valoraciones suministradas (clasificación kNN – k Nearest Neighbors):

- Colaborativo basado en usuario: considera los usuarios con patrones de evaluación similares.

	Ítem1	Ítem2	Ítem3	Ítem4	Ítem5	Ítem6
Usuario activo	4	2	3	3	5	?
Usuario1	5	2	2	4	5	3
Usuario2	1	4	2	4	2	2
Usuario3	2	5	5	5	2	4
Usuario4	5	2	4	3	4	4

- Colaborativo basado en ítem: considera los ítems con patrones de evaluación similares.

	Ítem1	Ítem2	Ítem3	Ítem4	Ítem5	Ítem6
Usuario activo	4	2	3	3	5	?
Usuario1	5	2	2	4	5	3
Usuario2	1	4	2	4	2	2
Usuario3	2	5	5	5	2	4
Usuario4	5	2	4	3	4	4



Tipos de RecSys: Híbridos

- ◆ Combinan ambas estrategias para mejorar el rendimiento del RecSys.
- ◆ Se suelen diseñar para atajar el problema del COLD STAR: RecSys basados en contenido cuando no hay muchos datos y RecSys colaborativos cuando hay más datos.
- ◆ Los RecSys híbridos son los más comunes

Tipos de RecSys: Híbridos

- Enfoque híbridos
 - Combinan distintas estrategias para explotar las ventajas y reducir los inconvenientes que cada una de ellas presenta por separado.
 - Se pueden combinar de distintas maneras:
 - *Mixed hybrid RecSys*: presentan al usuario una única lista con sugerencias de varios RecSys.
 - *Weighted hybrid RecSys*: combinan ponderando según importancia las predicciones de varios RecSys.
 - *Switched hybrid RecSys*: alternan entre un esquema y otro según un cierto criterio.
 - *Cascade hybrid RecSys*: enlazan secuencialmente varios enfoques, donde la salida de uno es la entrada del siguiente.

Tipos de RecSys: Actualización

“Core” Recommendation Techniques

U is a set of users

I is a set of items/products

Technique	Background	Input	Process
Collaborative	Ratings from U of items in I .	Ratings from u of items in I .	Identify users in U similar to u , and extrapolate from their ratings of i .
Content-based	Features of items in I	u 's ratings of items in I	Generate a classifier that fits u 's rating behavior and use it on i .
Demographic	Demographic information about U and their ratings of items in I .	Demographic information about u .	Identify users that are demographically similar to u , and extrapolate from their ratings of i .
Utility-based	Features of items in I .	A utility function over items in I that describes u 's preferences.	Apply the function to the items and determine i 's rank.
Knowledge-based	Features of items in I . Knowledge of how these items meet a user's needs.	A description of u 's needs or interests.	Infer a match between i and u 's need.

[Burke, 2002]

Tipos de RecSys: Evaluación

	Selected	Not selected	Total
Relevant	N_{rs}	N_{rn}	N_r
Irrelevant	N_{is}	N_{in}	N_i
Total	N_s	N_n	N

Table 4.1. Contingency table.

$$P = \frac{N_{rs}}{N_s} \quad R = \frac{N_{rs}}{N_r} \quad F1 = \frac{2 \times R \times P}{R + P}$$



Tipos de RecSys: Evaluación

	User1	User2	User3	User4	User5	User6	User7	User8	User9	User10
Nrs	5	6	3	4	5	6	5	3	4	5
Nrn	3	2	2	2	3	2	2	1	3	2
Nis	4	3	6	5	4	3	4	6	5	4
Nr	8	8	5	6	8	8	7	4	7	7
Ns	9	9	9	9	9	9	9	9	9	9

Table 4.2. Experimental contingency table.

Tipos de RecSys: Evaluación

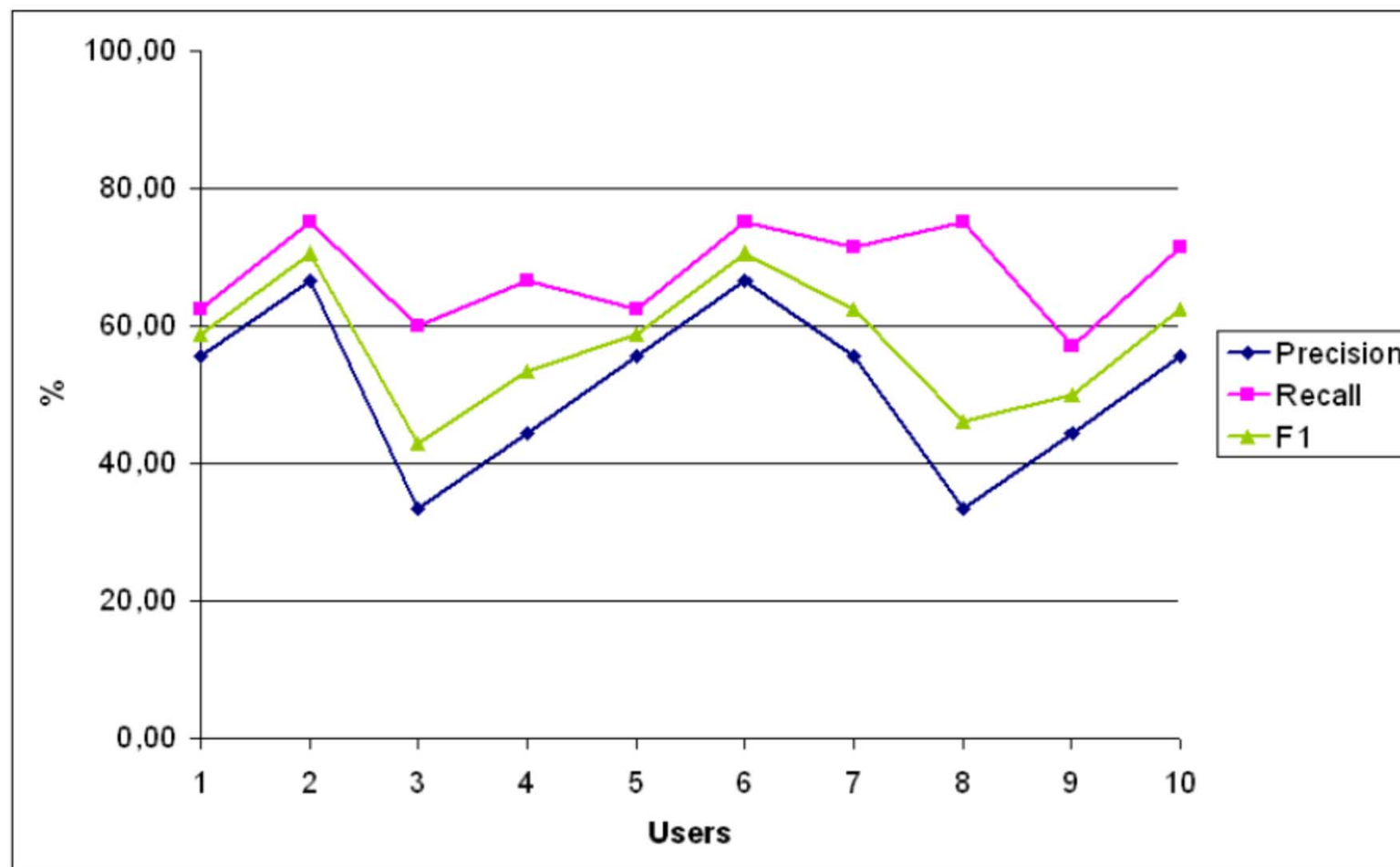
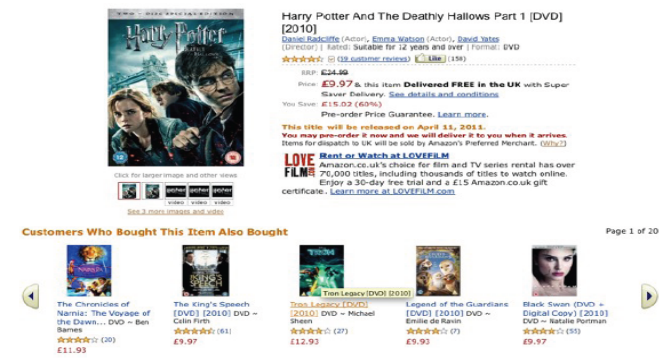


Figure 8: Experiment result

RecSys Basados en Contenido: Problemas

- ◆ Sobre-especialización de la recomendación: Al usuario se le recomiendan ítems similares a los que ya le han gustado, por tanto generan recomendaciones obvias y es difícil descubrir ítems inesperados-**SERENDEPIA**



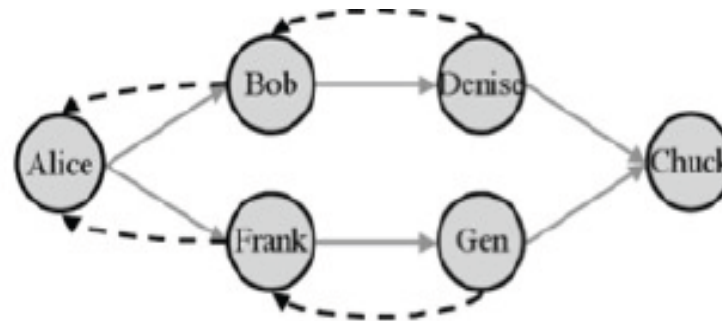
- ◆ Problema del Cold Star con los nuevos usuarios: No se puede recomendar nada a alguien del que no se conoce nada.
- ◆ Feedback del usuario: obtener los ratings de usuario



RecSys Colaborativos: Problemas

- ◆ Problemas de confianza (Trust): Porqué ha de confiar un usuario en la recomendación de alguien que no conoce.

Trust RecSys.



- ◆ Problema del Cold Star con los nuevos ítems: No se puede generar una recomendación para un ítem que no ha sido evaluado previamente por algún usuario .
- ◆ Problema de cálculo de similitud entre usuarios: Cuando no hay suficiente ítems valorados comunes.
- ◆ Feedback del usuario: obtener los ratings de los usuarios



RecSys Colaborativos: Problemas

- ◆ SPARSITY: Problema de cálculo de similaridad entre usuarios. Cuando no hay suficiente ítems valorados comunes.

Typically large product sets & few user ratings
e.g. Amazon:

- ◆ in a catalogue of 1 million books, the probability that two users who bought 100 books each, have a book in common is 0.01
 - ◆ in a catalogue of 10 million books, the probability that two users who bought 50 books each, have a book in common is 0.0002
- ◆ Feedback del usuario: obtener los ratings de los usuarios



RecSys Propios

◆ BIBLIOTECAS DIGITALES

J.M. Morales-del-Castillo, R. Pedraza-Jiménez, A.A. Ruíz, E. Peis, [E. Herrera-Viedma](#), **A Semantic Model of Selective Dissemination of Information for Digital Libraries**. *Information Technology and Libraries* 28(1) (2009) 22-31

◆ OTRI: RECURSOS DE INVESTIGACIÓN

[C. Porcel](#), [A.G. López-Herrera](#), [E. Herrera-Viedma](#), **A Recommender System for Research Resources based on Fuzzy Linguistic Modeling**. *Expert Systems with Applications*, 36 (2009) 5173-5183

◆ MEDICINA: TRATAMIENTO DE LUMBARGIAS

Un Sistema Web para Prevención de Lumbalgias Crónicas. *Fundación MAPFRE*. Periodo: Febrero 2009 - Febrero 2010.

Coordinador: [Dr. Carlos Porcel Gallego](#)



Outline

1. El problema de acceso a la información
2. RecSys y la Evolución de la Web
3. Qué son los RecSys ?
4. Estructura y funcionamiento de los RecSys: Tipos
5. Tendencias Futuras



Tendencias de Futuro: RecSys x Web 2.0

1. Social RecSys (RecSys basado en Redes Sociales)
2. Trust RecSys (RecSys basados en Confianza)
3. Context RecSys (RecSys basado en Contexto)
4. RecSys en Plataformas de E-commerce

● ● ● Tendencias de Futuro: RecSys x Web 2.0

1. Social RecSys (RecSys usados en Redes Sociales):

- ❖ Usar la gran cantidad de datos que se genera en los entornos Web 2.0 para generar recomendaciones sociales: resuelve el cold star. ([Big Data](#))
- ❖ Minar las opiniones dadas en las redes sociales:
 - ◆ Por Twitter y por Facebook vamos a encontrar montones de referencias a marcas y productos comerciales, tanto positivas como negativas.
- ❖ Analizar sentimientos: detectar gente apasionada o decepcionada
- ❖ Recomendar amigos, etiquetas enlaces....

● ● ● Tendencias de Futuro: RecSys x Web 2.0

2. Trust RecSys (RecSys basados en Confianza):

- ❖ Computar confianza en otros usuarios
- ❖ Generar recomendaciones desde los usuarios en los que confiamos (no desde los similares)



● ● ● Tendencias de Futuro: RecSys x Web 2.0

Building RS Using Trust

- Trust for CF
 - Use trust to give more weight to some users
 - Use trust in place of (or combined with) similarity
- Trust for sorting & filtering
 - Prioritize information from trusted sources

● ● ● | Tendencias de Futuro: RecSys x Web 2.0

3. Context RecSys (RecSys basados en Contexto):

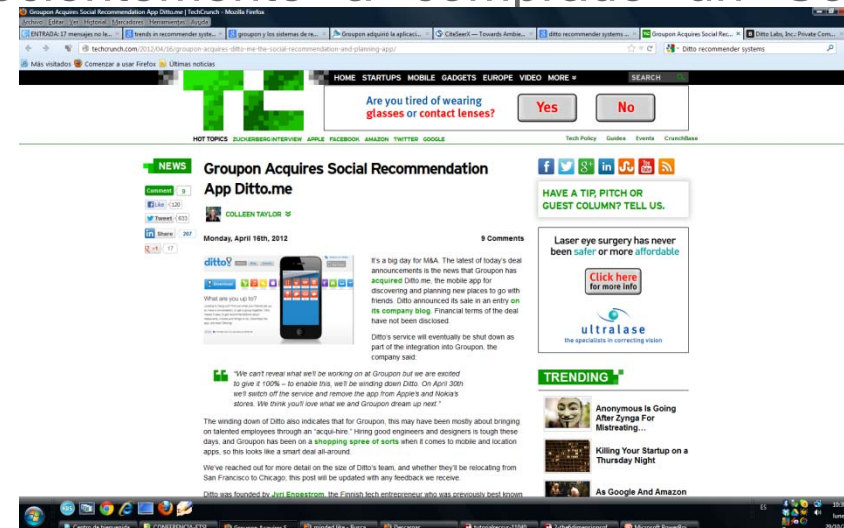
- ❖ Hay dominios donde considerar solo ítems y usuarios no es suficiente.
- ❖ **Contexto** engloba ubicación, tiempo, momento (trabajo o casa), juegos recientes....
- ❖ Nos interesan recomendaciones en función del contexto:
 - ◆ Si estoy en Granada, no me recomiendes comer en Madrid
 - ◆ Si últimamente juego a la PS3 no me recomiendes juegos de PC



● ● ● Tendencias de Futuro: RecSys x Web 2.0

4. RecSys en Plataformas de E-commerce:

- ❖ Introducir RecSys Sociales en las plataformas de E-commerce.
- ❖ Ayudar a la gente a conectar y compartir sus experiencias:
 - ◆ Groupon presenta ofertas diarias de minoristas y restaurantes a lo largo del mundo. Recientemente a comprado un Social RecSys llamado Ditto.





Conclusión



John Riedl, University of Minnesota (2010)

“El efecto de los sistemas de recomendaciones será uno de los mayores cambios en la próxima década. La web social va a ser dirigida por los RecSys”



Sistemas de Recomendaciones (Recommender Systems-RecSys)

Enrique Herrera Viedma

Dpto. De Ciencias de la Computación e I.A

Granada, Marzo de 2018