

# User Search Behaviour y la Arquitectura de información:

Amigos para siempre

Jorge Serrano Cobos  
jorgeserrano@gmail.com



## Hablaremos de...

- Problemas de usabilidad, AI y Recuperación de Información
- Definición, beneficios y tipos de estudios de "User Search Behaviour"
- Indicadores de estudios, Zipf, Best bets...
- Relación con Findability, Posicionamiento en Buscadores y Arquitectura de Información
- Conducta de búsqueda del usuario antes y después del diseño web
- Ejemplos reales

## Buscar, explorar, tal vez encontrar...



## El contexto: diseño centrado en el usuario

- La usabilidad ya no está de moda
- De la heurística a la “usabilidad de boquilla”
- Todos siguen “opinando”
- Pocos usan las normas ISO
- Difícil de “evangelizar” entre equipos multidisciplinares

## Problemas de usabilidad

- Proceso lento y caro
- Poca investigación
- Muchas heurísticas, pocos datos empíricos
- Poca representatividad de ciertos datos
- Falta más investigación

## Problemas de Arq. de Información

- La gran desconocida
- ¿Tesauro? ¿Vocabulario controlado?
- Corto plazo vs. Largo plazo
- Pre-diseñar es difícil
- Sin tiempo para análisis post-diseño
- Si no se encuentra, “que use el buscador”

## Problemas de recuperación de información

- Lenguaje natural vs. Lenguaje documental
- Diversidad de algoritmos
- Un buscador para cada contexto
- Falta de contexto
- Falta de metadatos
- Localismos y variantes idiomáticas
- Sitios web multilingües: ¿traducción exacta?

## Problemas de recuperación de información

- Misspellings (*varco* en vez de *barco*)
- Desambigüación (*banco de pesca* versus *banco del parque*)
- Operadores booleanos (AND, +, Y...)
- Demasiados / demasiado pocos resultados
- Pertinencia vs. exhaustividad
- Pocas o demasiadas funcionalidades de búsqueda
- Usabilidad de las respuestas

¿Una solución?

Una aproximación

¡User search behaviour!

¿Qué es?

- Definición propia:  
“Estudio del lenguaje de recuperación y de los patrones de navegación de los usuarios de un sistema de información, en su búsqueda de información durante su interacción con el sistema.”

## Algunos nombres propios

- Gerard Salton
- Ingwersen
- Marchionini
- Belkin
- Carol Kuhlthau
- David Ellis
- Marcia J. Bates
- Donald A. Norman
- Avi Rappoport
- Lou Rosenfeld

## Dos tipos de estudios

- Basados en tests de usuarios
- Basados en análisis de search logs
  - Ambos pueden ser pre-diseño o post-diseño.
  - Ambos son compatibles

## Basados en tests de usuarios

- Conocemos los pasos que los usuarios dan para encontrar información
- Centrado en la navegación
- Se usan pocos usuarios
- Sirve para pre y rediseñar navegación
- Cubre buscador + “browsing” (information seeking)
- Aplicable a la arquitectura de información macro y micro.

## Ejemplo

- Famoso estudio “Eyetrack”.
- Ayuda a entender los procesos de investigación y descubrimiento de información, “seeking & browsing”



## Basados en análisis de search logs

- Se centra en el archivo log del buscador
- Aporta más datos cuantitativos
- Informa:
  - del lenguaje usado (“labelling” o etiquetado)
  - de quién y cuándo lo ha buscado (bajo registro)
  - del contenido encontrado,
  - y del no encontrado.
  - Conceptos relacionados
  - Categorías (+ buscado) > Subcategorías (- buscado)
  - Categorías (- palabras) > Subcategorías (+ palabras)

## Search logs

En crudo:

- **Keywords: focis; 0; 11/26/01 12:57 PM; XXX.XXX.XXX.2**
- **Keywords: focus; 167; 11/26/01 12:59 PM; XXX.XXX.XXX.2**
- **Keywords: focus pricing; 12; 11/26/01 1:02 PM; XXX.XXX.XXX.2**
- **Keywords: discounts for college students; 0; 11/26/01 3:35 PM; XXX.XXX.XXX.59**
- **Keywords: student discounts; 3; 11/26/01 3:35 PM; XXX.XXX.XXX.59**

*Louisrosenfeld.com*

## Search logs

Un search log elaborado:

<b>Período:</b> No	0 resultados para el 5,29% de	
<b>Palabra Clave:</b>	las búsquedas:	
<b>Nº búsquedas:</b>	¡PROBLEMAS!	
<b>% del total:</b> 5,29%	<b>vacías:</b> no	
<b>PC Superior:</b> Teatro / Infancia / Educación	<b>Resultados búsqueda:</b> 0	
<b>PC Inferior:</b> talleres teatro niños Gandía	<b>Misspelling:</b> no	
<b>PC Relacionadas:</b> cursos / niños / Animación a la lectura	<b>Usuarios:</b> Pepe, Juan, ...	

## Para qué sirven

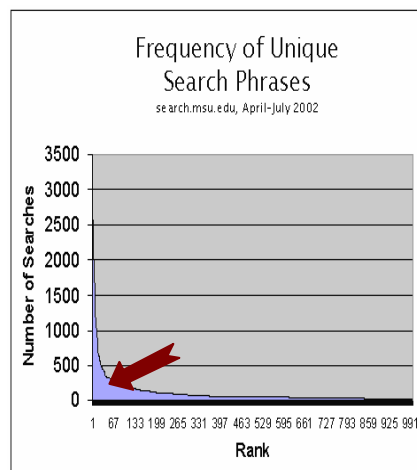
- Aportar métricas exhaustivas
- Evaluar el lenguaje documental
- Saber cómo definen un concepto ("modelo mental")
- Valorar el sistema de recuperación de información, el éxito (o no) del buscador (exhaustividad, pertinencia...)
- Grado de visibilidad del contenido

## Beneficios

- Sistema no intrusivo para el usuario
- Sistema a priori barato
- Aporta datos reales
- Datos exhaustivos (de todos los usuarios del sitio)
- Los datos son del propio contexto del sitio
- Permite muchos análisis sincrónicos y diacrónicos

## Punto de partida

- Las búsquedas siguen una distribución según Zipf, con búsquedas más comunes
- De ahí salen *best bets*, contenidos nuevos...



lourosenfeld.com

## Estudios posibles

- Sinónimos y conceptos relacionados
- Variantes idiomáticas, jergas y equivalentes multilingües
- Determinación de “best bets”
- Falta de contenido (“Nº de respuestas: 0”)
- Relevancia por retroalimentación
- Estudios diacrónicos (evolución temporal)

## Indicadores

- Nº de búsquedas por período en el año
- Top de búsquedas
- Top temáticas (clusters del Top búsquedas)
- Top de 0 respuestas
- Top de pocas respuestas
- Cambios de búsquedas más frecuentes durante el año
- Correlación de datos con cambios en diseño, etc.

## Indicadores

- Búsquedas cada vez más frecuentes
- Patrones de movimiento expresiones más/menos frecuentes
- Top de páginas más visitadas tras consultas y a través de qué búsqueda

## Ejemplo: [comunitatvalenciana.com](http://comunitatvalenciana.com)

mapa	17 -> 15.6%
mapa de valencia	14 -> 12.84%
MAPA VALENCIA	7 -> 6.42%
mapa de benidorm	5 -> 4.59%
MAPAS	4 -> 3.67%
mapa de la comunidad valenciana	3 -> 2.75%
mapa fisico	3 -> 2.75%
mapa de la ciudad de Valencia	3 -> 2.75%
mapa de comarcas	2 -> 1.83%
mapa provincia valencia	2 -> 1.83%
mapa de orihuela	2 -> 1.83%
fallas mapa	2 -> 1.83%
mapa de la ciudad	2 -> 1.83%
mapa de carreteras	2 -> 1.83%
mapa de llanuras y montañas	1 -> 0.92%
mapas Comunidad Valenciana	1 -> 0.92%
MAPA CALLEJERO DE VALENCIA	1 -> 0.92%
MAPAS CALLEJEROS	1 -> 0.92%

## Ejemplo: comunitatvalenciana.com

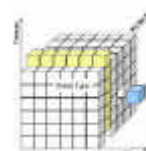
webcam	52 -> 32.5%
web cam	22 -> 13.75%
webcams	15 -> 9.38%
camara web	7 -> 4.38%
web cams	5 -> 3.13%
web cam denia	4 -> 2.5%
webcamera's	3 -> 1.88%
web cam comunidad valenciana	2 -> 1.25%
web cameras	2 -> 1.25%
WEBCAM TORREVIEJA	2 -> 1.25%
red web cams	2 -> 1.25%
web cam de elx	1 -> 0.63%
web cam valencia	1 -> 0.63%
webcamera albir	1 -> 0.63%
WEBS CAMS	1 -> 0.63%
webcam en directo	1 -> 0.63%
camera web	1 -> 0.63%

## El futuro del análisis de search logs

Business intelligence:

- Cubo de datos con log de visitas
- Coordenadas XYZ páginas, usuarios, tiempo
- Permite segmentar por XYZ
- Permite realizar estimaciones y simulaciones a futuro
- Mezclado con search log:

*“¿Cuántos usuarios mayores de 30 años varones de Valencia buscaron “pañales” a las 12 de la noche de ayer y no encontraron contenido satisfactorio?”*



evolucy.com

## ¿Y antes de comenzar la Arquitectura de Información?

### Findability

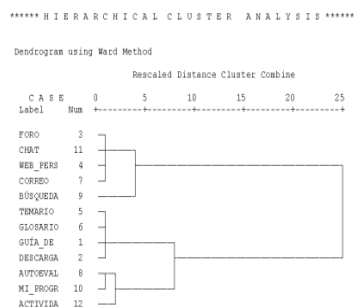
- ¿Qué términos o palabras clave son aquellas por las que queremos ser encontrados en internet?  
(Posicionamiento en buscadores)
- ¿Qué términos queremos que sean encontrados dentro de nuestro sitio web? (Arquitectura de Información)

## Construcción de categorías

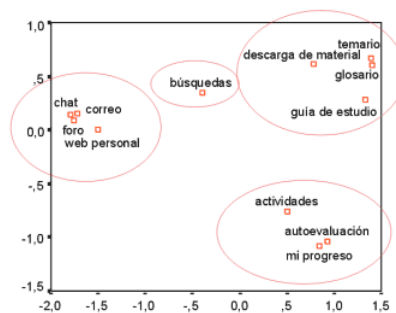
- Dos aproximaciones:
  - Top-down (card sorting)
  - Bottom-up (keywords clustering)

## Arquitectura de Información Top - Down

### Dendogramas



### Escalamiento Multidimensional (MDS)



Nosolousabilidad.com

# Arquitectura de Información Bottom-up

## Herramientas de ayuda en la selección de keywords

- Herramienta de sugerencias de Google Adwords

<https://adwords.google.com/select/KeywordSandbox>

- Herramienta de sugerencias de Overture:

<http://inventory.overture.com/searchinventory/suggestion>

- Wordtracker: [www.wordtracker.com](http://www.wordtracker.com)

The image shows a screenshot of a Google search results page for the query "planes pensiones". The search bar at the top contains the text "planes pensiones" and "Búsqueda en Google". Below the search bar, there are navigation links for "La Web", "Imágenes", "Grupos", "Directorio", "News", and "Páginas". The search results are displayed in a list format. The first result is "Pensión Nº 1 Rentabilidad" from www.canaldirect.com. The second result is "Plan de Pensiones Naranja" from www.ingdirect.es. The third result is "Planes de pensiones y seguros en Segurosbroker". The fourth result is "Caixa Tarragona : Planes de Pensiones". The fifth result is "Planes Pensiones - Fiscalidad". The sixth result is "40.000 empresas en España : SEGUROS Y PLANES PENSIONES, NO...". The seventh result is "40.000 empresas en España : Seguros y planes pensiones, excp...". On the right side of the page, there are several sponsored links (Enteses Patrocinadas) for "Planes Pensiones", "Planes de pensiones", "ING Rotorale Nederland", "Plan de Pensiones Renta 4", and "Pensión Plus". Annotations in red text and brackets are present: "Concepto buscado" points to the search bar; "Google AdWords (CPC)" is written next to the first two results; "Resultados de búsqueda basados en Algoritmo Google (PageRank)" is written next to the third, fourth, and fifth results; and "Google AdWords (CPC)" is written next to the sixth, seventh, and eighth results. A vertical dashed line separates the organic search results from the sponsored links.

## Selección de keywords con AdWords

Ejemplo: comercialización de un Plan de Pensiones

Palabra clave	Clics / día	
ahorro	12,0	<a href="#">palabras clave sugeridas</a> / <a href="#">suprimir</a>
cuenta	33,0	<a href="#">palabras clave sugeridas</a> / <a href="#">suprimir</a>
deposito	6,2	<a href="#">palabras clave sugeridas</a> / <a href="#">suprimir</a>
pensiones	32,0	<a href="#">palabras clave sugeridas</a> / <a href="#">suprimir</a>
plan	190,0	<a href="#">palabras clave sugeridas</a> / <a href="#">suprimir</a>
plan pensiones	2,6	<a href="#">palabras clave sugeridas</a> / <a href="#">suprimir</a>
<b>Total</b>	<b>275,8</b>	

## Selección de keywords con AdWords

Ejemplo: comercialización de un Plan de Pensiones

### Plan

### Pensiones

#### Sugerencias de palabras clave en AdWords de Google para plan

Sus anuncios se mostrarán por las búsquedas siguientes (a menos que introduzca sus palabras clave como concordancia exacta [1]). Si alguna de estas búsquedas no resulta relevante para su campaña, por favor seleccione palabras clave más específicas, como las que se muestran a continuación, y elimine las más generales de su lista de palabras clave. También debería considerar la opción de añadir palabras clave negativas [2].

- "plan general"
- "plan nacional"
- "plan general contable"
- "plan de marketing"
- "plan general de contabilidad"
- "plan hidrológico"
- "plan renova"
- "plan de empresa"
- "plan estrena"
- "plan hidrológico nacional"
- "plan de estudios"
- "plan fp"
- "plan vivienda"
- "plan estratégico"
- "plan de pensiones"
- "plan hidrológico"
- "plan prever"

Puede que desee añadir estas palabras clave similares a su lista para acompañar a plan .

- planes
- ordenacion
- ordenación
- plans
- planning
- mappy
- floorplan
- maporama
- planeamiento
- eplan
- contingencia
- grada
- plu
- planm
- michelin
- cuentas
- pla

#### Sugerencias de palabras clave en AdWords de Google para pensiones

Sus anuncios se mostrarán por las búsquedas siguientes (a menos que introduzca sus palabras clave como concordancia exacta [1]). Si alguna de estas búsquedas no resulta relevante para su campaña, por favor seleccione palabras clave más específicas, como las que se muestran a continuación, y elimine las más generales de su lista de palabras clave. También debería considerar la opción de añadir palabras clave negativas [2].

- "planes de pensiones"
- "hostales y pensiones"
- "plan de pensiones"
- "pensiones madrid"
- "fondos de pensiones"
- "hostales pensiones"
- "pensiones en madrid"
- "pensiones barcelona"
- "pensiones no contributivas"
- "planes pensiones"
- "hoteles y pensiones"
- "pensiones hostales"
- "pensiones y hostales"
- "pensiones en barcelona"
- "hoteles pensiones"

Puede que desee añadir estas palabras clave similares a su lista para acompañar a pensiones .

- hostales
- jubilacion
- pension
- alojamiento
- albergues
- alojamientos

## Selección de keywords con AdWords

Ejemplo: comercialización de un Plan de Pensiones

### Plan Pensiones

#### Sugerencias de palabras clave en AdWords de Google para plan pensiones

Sus anuncios se mostrarán por las búsquedas siguientes (a menos que introduzca sus palabras clave como concordancia exacta [2]). Si alguna de estas búsquedas no resultan relevante para su campaña, por favor seleccione palabras clave más específicas, como las que se muestran a continuación, y elimine las más generales de su lista de palabras clave. También debería considerar la opción de añadir palabras clave negativas [2].

- "plan de pensiones"
- "rescate plan de pensiones"
- "rescate plan pensiones"
- "que es un plan de pensiones"
- "plan pensiones telefonica"
- "plan de pensiones empleados de telefonica"
- "plan de pensiones de telefonica"
- "plan de pensiones garantizado"
- "simulador plan de pensiones"
- "planes pensiones"

Puede que desee añadir estas palabras clave similares a su lista para acompañar a *plan pensiones*.

**Tener en cuenta:**

- Singulares y Plurales
- Acentos
- Y todas las variaciones y extensiones de cada término.

## Construcción de categorías

- Agrupar las palabras clave según su afinidad.
- Dividir las en grupos que compartan la misma temática

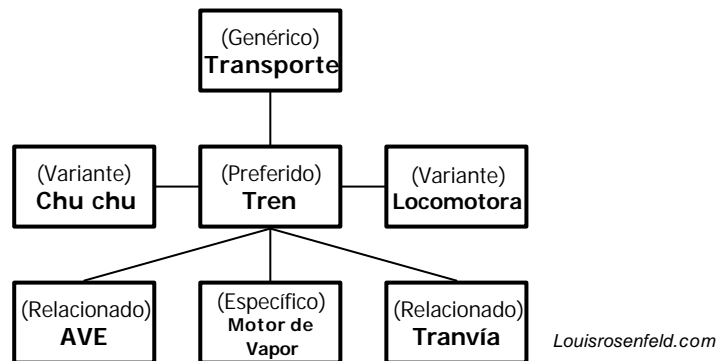
planes de pensiones
fondos de pensiones
calculo pensiones
pension compensatoria
calcular pension
calcular pension jubilacion
calculo de la pension
calculo pension viudedad
fondo de pension
fondo de pensiones
ley de pensiones
pension asistencial
pension contributiva
pension de horfandad
pension de invalidez
pension de jubilacion
pension de jubilación
pension de orfandad
pension de vejez
pension de viudedad
pension maxima
pension maxima de jubilacion



planes de pensiones
pension contributiva
pension compensatoria
ley de pensiones
pension asistencial
pension de invalidez
pension de horfandad
pension de orfandad
pension maxima
pension maxima de jubilacion
pension de jubilacion
pension de vejez
pension de viudedad
calcular pension jubilacion
calculo pensiones
calcular pension
calculo de la pension
calculo pension viudedad
fondos de pensiones
fondo de pension
fondo de pensiones

## Construcción de categorías

La agrupación de sinónimos y subcategorías por palabras más/menos buscadas puede indicarnos también la **navegación**



Nota: clics, no búsquedas

Seleccione la moneda y el coste por clic

EUR €

Estimador de tráfico

Palabra clave	Clics / día	Promedio de coste por clic	Coste / día	Posición promedio	
ahorro	12,0	€0,17	€1,94	1,1	<a href="#">palabras clave sugeridas / suprimir</a>
cuenta	33,0	€0,17	€5,64	1,2	<a href="#">palabras clave sugeridas / suprimir</a>
deposite	6,2	€0,14	€0,84	1,1	<a href="#">palabras clave sugeridas / suprimir</a>
pensiones	32,0	€0,47	€14,94	1,3	<a href="#">palabras clave sugeridas / suprimir</a>
plan	190,0	€0,13	€22,86	1,1	<a href="#">palabras clave sugeridas / suprimir</a>
plan pensiones	2,6	€0,23	€0,59	1,3	<a href="#">palabras clave sugeridas / suprimir</a>
<b>Total</b>	<b>275,8</b>	€0,17	€46,68	1,2	<a href="#">Modificar palabras clave</a>

CPC Máximo o Puja Techo de cada grupo de Palabras Clave

- Estimación de Clics diarios para este listado de Palabras Clave.

## Ejemplo comunidadvalenciana.com (datos OJD)

**Marzo 2004:**  
**1.742.038 páginas vistas**



+ 53 %



**Febrero 2005:**  
**2.656.849 páginas vistas**



# ¡GRACIAS!



Consultoría  
y desarrollo  
para Internet

C/ Álvaro de Bazán 10  
46010 Valencia  
Tel. 96 369 41 23

C/ Tejo 22  
28045 Madrid  
Tel. 91 530 89 03

  
www.masmedios.com