

Academic SEO

Desde la perspectiva de las revistas científicas

Rodrigo Sánchez Jiménez

Departamento de Biblioteconomía y Documentación

Universidad Complutense de Madrid

Planteamiento

- La optimización para motores de búsqueda académicos es un campo en auge sobre el que habría muchos detalles que revisar
- Para acotar mi intervención
 - Search Engine / **Motor de Búsqueda** debe implicar recuperación de información y por tanto relevancia (no sólo recuperación de datos)
 - No todos los motores de búsqueda **funcionan** igual
 - Es interesante centrar el trabajo en los motores de búsqueda que usan nuestros **lectores/autores**

¿Cómo encuentran información científica los investigadores?

- Fuentes personales y otras menos mainstream
- Fuentes de información bibliográficas/documentales
- Cuáles son las fuentes preferidas??



Fuente	Visitas mensuales (similarweb.com)
Scopus	5.700.000
WoS	5.000.000
Dimensions	382.000
The Lens	123.000
Google Scholar	8.000.000
Pubmed (*)	75.000.000
Arxiv	18.000.000
IEEE Xplore	13.800.000
ERIC	4.000.000
DBLP	2.000.000

(*) Datos de la National Library of Medicine

¿Qué dice la literatura sobre Information Seeking Behaviour?

- Los **matemáticos**
 - **Google Scholar** 68.6%, arXiv.org. 61.5%, MathSciNet 40.4%, and Web of Science 13.8% as the most frequented search databases (Gordon et al, 2020)
- Los **investigadores multidisciplinares**
 - ...relied on **Google scholar** (100%), ToC email alerts (63%) and searching general databases such as Web of Science (58%) (Jamali & Nicholas, 2010)
- Los **investigadores nigerianos**
 - **Google Scholar** (75.4%), ResearchGate (64.9%), African Journals Online (57.0%), Web of Science (52.6%), Science Open (23.7%), Scopus (22.8%) (Makinde, Jiyane & Mugwisi, 2021)

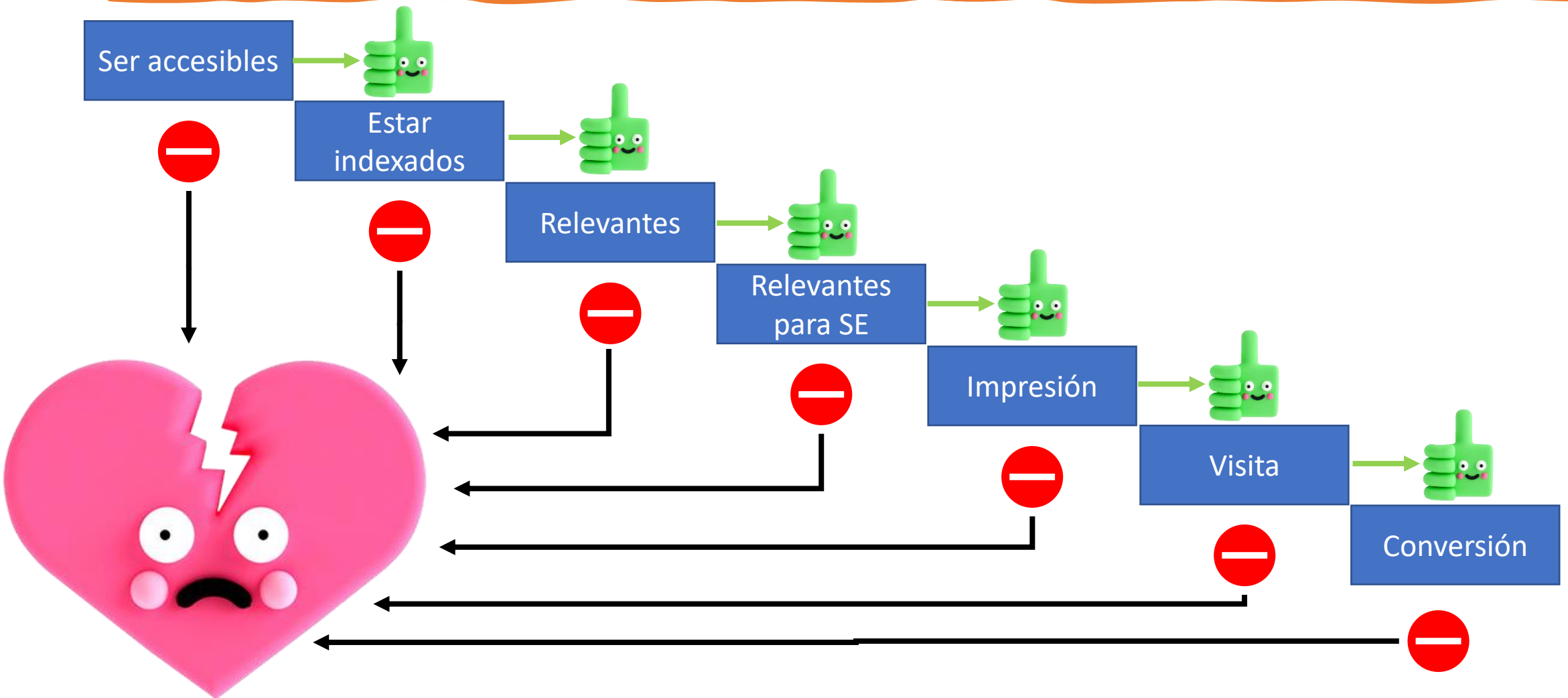
¿Qué dicen los expertos sobre Information Seeking Behaviour? II

- Los **químicos**
 - Google Scholar 61%, SciFinder 55.4%, Web of Science/Scopus 55.% Reaxys 23.8%, PubMed 17.3%... (Gordon, et al. 2018)
- Los **físicos**
 - ... arXiv database and Google Scholar were each selected by 71.4% of respondents, Web of Science was selected by 23.1%, Scopus was selected by 4.4%, and Inspec by 1.7% (Gordon et al. 2022)
- Sólo los **bibliotecarios** de las **ciencias del agua** le tienen tirria a GS
 - Aquatic Commons 54.8%, FAO Stat 44.68%, Ocean Docs 44.65%, **GS 4.6%!!** (Superio, Oliveros, Palcullo & Geromiano, 2020)

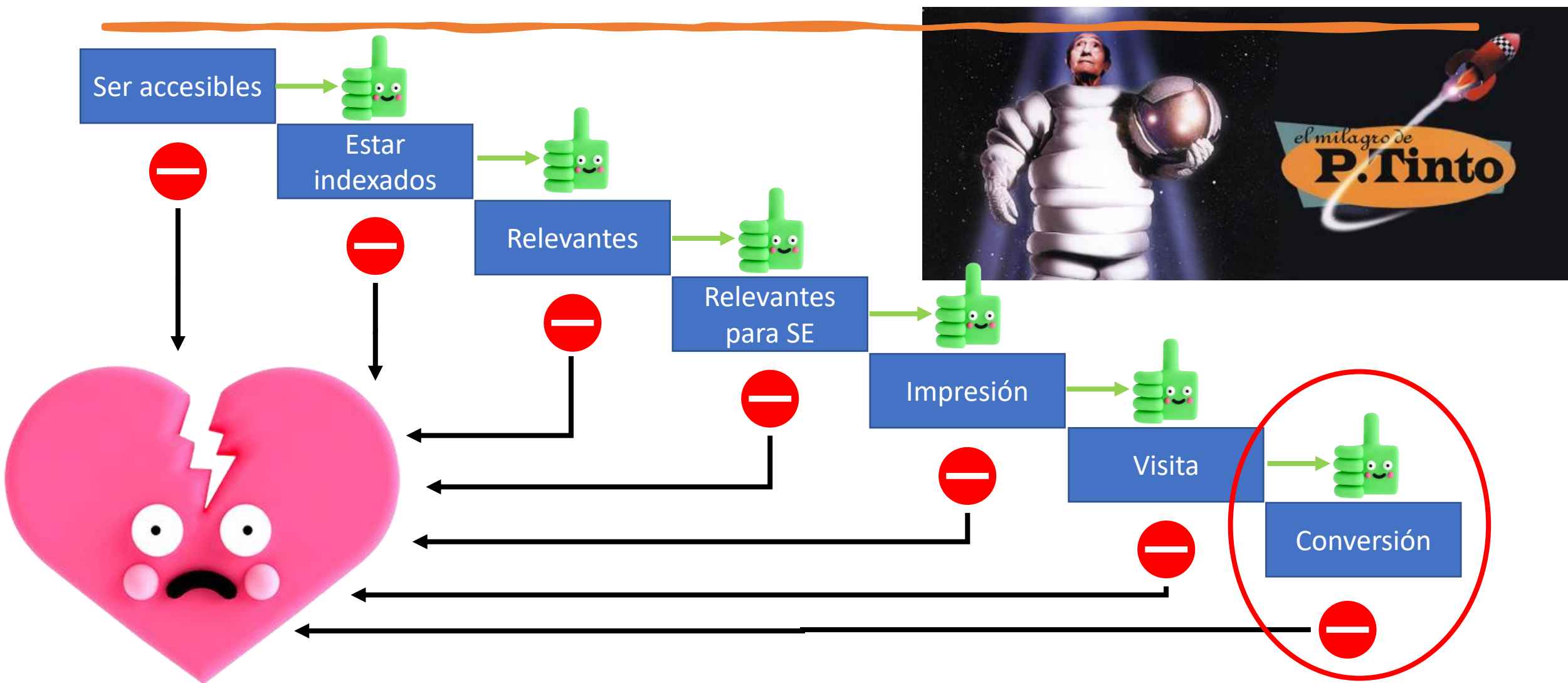
En conclusión...

- Hay que preocuparse mucho de **Google Scholar**
- Hay que preocuparse bastante de las **bases de datos especializadas** (localmente, incluso tanto como de Google Scholar)
- También hay que preocuparse algo de las **bases de datos generalistas**

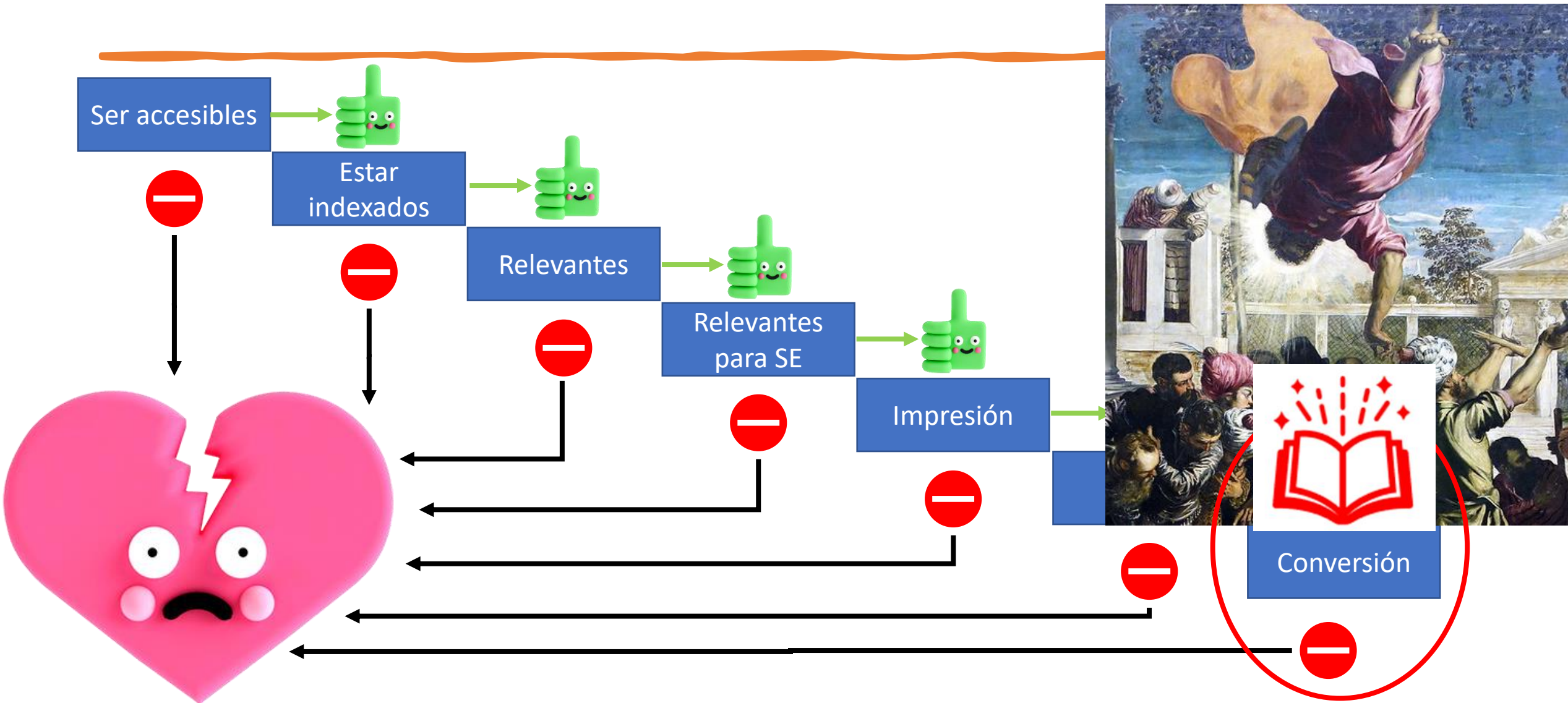
El proceso de la Recuperación de Información



El milagro de la Recuperación de Información

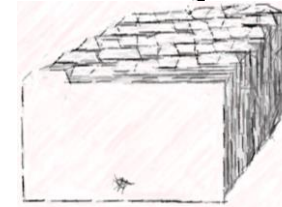
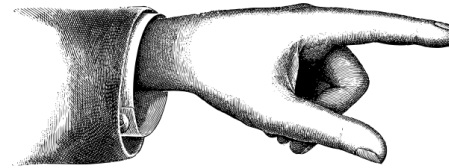


El milagro de la Recuperación de Información



¿Qué puede hacer una revista para propiciar el milagro de la RI?

- Preocuparse por la **correcta indexación** de sus trabajos
 - Algorítmica
 - Editorial



- Preocuparse por que se **encuentren** con la mayor facilidad posible



- Ser **atractivos**



- Muchos dolores de cabeza!



Estar indexados...

IEEE Xplore[®]
Digital Library



ERIC
Institute of Education Sciences

PubMed



dblp
computer science bibliography



LENS.ORG

Solving The Problem Of Problem Solving™

Google Académico

arXiv



Dimensions




Scopus



Clarivate
Web of Science[™]

Indexación, vigilancia y persecución

- Darse de **alta, solicitar inclusión, cumplir criterios**
- **Monitorizar** cobertura en buscadores y bases de datos
 - Medir la cobertura y sus **cambios**
 - Corregir problemas de **trabajos no indexados**
 - Perseguir **citas no indexadas**
- ¿Cómo afrontamos esto?
- Leyendo fuentes  que describen bien el problema y las soluciones



Indexación, vigilancia y persecución

- Darse de alta, solicitar inclusión, cumplir criterios
- Monitorizar cobertura en buscadores y bases de datos
 - Medir la cobertura y sus cambios
 - Corregir problemas de trabajos no indexados
 - Perseguir citas no indexadas
- ¿Cómo afrontamos esto?
- Poniendo velas a Santo Tomàs!



Indexación de los trabajos de la revista

- Situación de EPI en Scopus / WoS

Años	Trabajos en OJS	Trabajos Scopus	Trabajos WoS	% Index. Scopus	% Index. WoS
2013	72	71	71	99%	99%
2014	73	73	56	100%	77%
2015	89	87	88	98%	99%
2016	94	93	94	99%	100%
2017	115	115	113	100%	98%
2018	122	122	111	100%	91%
2019	128	127	128	99%	100%
2020	192	191	189	99%	98%
2021	124	122	123	98%	99%
2022	39	19	19	49%	49%



En Google Scholar
100% de trabajos
indexados en los
últimos 3 años

Duplicados y Versiones Principales (ajenas)

- Google Scholar es muy propenso a los dos fenómenos
- Las versiones principales indexadas en otros servidores son frustrantes

Servidores	% Puntos de entrada
EPI	43,7%
Dialnet	25,6%
Rep. Universidades	11,8%
Ebscohost	6,3%
Proquest	4,9%
Otros	2,8%
Rep. Eprints	2,4%
ResearchGate	1,4%
Academia.edu	0,6%
Instit. Investigación	0,4%

Años	% Puntos entrada directos EPI
2020	41%
2021	44%
2022	62%

Años	Duplicados detectados	Trabajos OJS
2020	83	192
2021	24	124
2022	4	39



Autores, esperad un poco!

Intentando ser relevantes

- Intentar ser relevantes **para los lectores**
 - Acierto en la política editorial
 - Capacidad par atraer trabajos y autores interesantes
- Intentar ser relevantes **para los buscadores**
 - Hay varias cosas escritas sobre **factores intrínsecos** y **extrínsecos**
 - Creo que subestimamos la importancia de los **resúmenes** y los **títulos**

¿Sobre qué se calcula la relevancia?

- Texto **completo**
 - Google Scholar, arxiv
- Sólo **título, abstract** y palabras clave:
 - WoS y Scopus, pero también, Dimensions o Lens*
 - Pubmed, dblp, ERIC, IEEE Xplore y muchas otras bases de datos especializadas, que localmente son muy importantes en las comunidades científicas correspondientes
- El título, resumen y palabras clave, son **super importantes**
- Incluso en **Google Scholar**, son importantes:
 - Como **condición de indexación** (resumen y título)
 - Como factor de intrínseco (campos **bien ponderados**)
 - Como medio para conseguir pasar de la **impresión** a la **visita**

¿Cómo se ofrecen los resultados de búsqueda?

Snippets dinámicos para capturar el contexto en el que aparecen los términos de búsqueda

The image shows a screenshot of the Google Académico search interface. At the top, the search bar contains the text "fake news" and a search icon. Below the search bar, the results are displayed under the heading "Artículos" with a subtext "Aproximadamente 312.000 resultados (0,10 s)".

On the left side, there are several filters and sorting options:

- Cualquier momento**: Desde 2019, Desde 2018, Desde 2015, Intervalo específico...
- Ordenar por relevancia**: Ordenar por fecha (highlighted with a red box and the word "snippet" written inside it).
- Cualquier idioma**: Buscar sólo páginas en español
- Checkboxes for "incluir patentes" and "incluir citas", both of which are checked.

The search results list several articles:

- [LIBRO] Fake News: La verdad de las noticias falsas**
MA García - 2018 - books.google.com
¿ No sabes qué son las **fake news**? Lee este libro.¿ Lo sabes y crees que no van contigo? Lee este libro.¿ Te crees capaz de diferenciar una noticia falsa de una noticia verdadera? Lee este libro.¿ Compartes noticias en Internet sin importarte si son o no verdad? Lee este ...
☆ 📄 Citado por 26 Artículos relacionados
- Social media and fake news in the 2016 election**
H Allcott, M Gentzkow - Journal of economic perspectives, 2017 - aeaweb.org
Following the 2016 US presidential election, many have expressed concern about the effects of false stories ("**fake news**"), circulated largely through social media. We discuss the economics of **fake news** and present new data on its consumption prior to the election ...
☆ 📄 Citado por 1854 Artículos relacionados Las 24 versiones Web of Science: 303 📄
- The science of fake news**
DMJ Lazer, MA Baum, Y Benkler, AJ Berinsky... - ..., 2018 - science.sciencemag.org
The rise of **fake news** highlights the erosion of long-standing institutional bulwarks against

¿Cómo se ofrecen los resultados de búsqueda?

- El **snippet/abstract** cuando es el contexto más apropiado o se menciona a los autores



The screenshot shows the Google Académico search interface. The search bar contains the text "cristòfol rovira academic seo" and a blue search button. Below the search bar, the results are categorized under "Artículos" with a subtext "Aproximadamente 419 resultados (0,04 s)". The first result is titled "Ranking by relevance and citation counts, a comparative study: Google Scholar, Microsoft Academic, WoS and Scopus" and is from the journal "Future Internet" (2019) by authors C. Rovira, L. Codina, F. Guerrero-Solé, and C. Lopezosa. The snippet describes search engine optimization (SEO) as methods to increase website visibility and visits. The result includes a PDF link from mdpi.com and various interaction options like "Guardar", "Citar", and "Citado por 32".

Google Académico

cristòfol rovira academic seo

Artículos

Aproximadamente 419 resultados (0,04 s)

Cualquier momento

Desde 2022

Desde 2021

Desde 2018

Intervalo específico...

Ranking by relevance and citation counts, a comparative study: Google Scholar, Microsoft **Academic**, WoS and Scopus

[C Rovira](#), [L Codina](#), [F Guerrero-Solé](#), [C Lopezosa](#) - Future Internet, 2019 - mdpi.com

Search engine optimization (SEO) constitutes the set of methods designed to increase the visibility of, and the number of visits to, a web page by means of its ranking on the search ...

☆ Guardar Citar Citado por 32 Artículos relacionados Las 12 versiones

[PDF] mdpi.com

Algunas ideas de investigación

- El **papel** de los **abstracts** y **títulos** en el **impacto** de las publicaciones tiene una literatura en **cienciometría** (por mencionar unos pocos)
 - The impact of title length and punctuation marks on article citations
 - The effect of keyword repetition in abstract and keyword frequency per journal in predicting citation counts
 - The effect of a country's name in the title of a publication on its visibility and citability
 - Automatic prediction of citability of scientific articles by stylometry of their titles and abstracts
 - Research article titles in written feedback on English as a second language writing
 - Non-alphanumeric characters in titles of scientific publications: An analysis of their occurrence and correlation with citation impact
 - Factors affecting number of citations: a comprehensive review of the literature
 - What makes a 'good' title and (how) does it matter for citations? A review and general model of article title attributes in management science
 - The advantage of simple paper abstracts
 - Short and amusing: The relationship between title characteristics, downloads, and citations in psychology articles
- El nivel de **preocupación** por la capacidad que tengan **resúmenes** y **títulos** para mejorar el **ratio de conversion** o para **mejorar la posición** de los artículos en la lista de resultados de los buscadores parece ser **mucho menor en ASEO** → Lo anterior puede funcionar de **proxy**

Algunas recomendaciones (probablemente discutibles)

- Adoptar un **papel activo** (desde la revista) en la confección de **abstracts** y **títulos**
- Privilegiar abstracts más largos
 - Nada de resúmenes **indicativos**
 - Resúmenes **estructurados** (cuando sea costumbre en el campo)
 - Resúmenes **informativos** o incluso **analíticos**
- Estructura de **pirámide invertida** (GS sólo despliega una parte pequeña de los abstracts)
- Incentivar un **léxico variado** pero compuesto de **términos bien aceptados** → Sinónimos con alguna esperanza de convertirse en puntos de entrada

Bibliografía

- Gordon, I. D., Meindl, P., White, M., & Szigeti, K. (2018). Information seeking behaviors, attitudes, and choices of academic chemists. *Science & Technology Libraries*, 37(2), 130-151.
- Gordon, I. D., Cameron, B. D., Chaves, D., & Hutchinson, R. (2020). Information seeking behaviors, attitudes, and choices of academic mathematicians. *Science & Technology Libraries*, 39(3), 253-280.
- Gordon, I. D., Chaves, D., Dearborn, D., Hendriks, S., Hutchinson, R., Popovich, C., & White, M. (2022). Information Seeking Behaviors, Attitudes, and Choices of Academic Physicists. *Science & Technology Libraries*, 1-31.
- Jamali, H. R., & Nicholas, D. (2010). Interdisciplinarity and the information-seeking behavior of scientists. *Information processing & management*, 46(2), 233-243.
- Makinde, O. B., Jiyane, G. V., & Mugwisi, T. (2021). Information-seeking behaviour of science and technology researchers in Nigeria: A survey of the Federal Institute of Industrial Research Oshodi. *IFLA journal*, 47(1), 20-36.
- Orduña-Malea, E., Martín- Martín, A., Ayllón, J.M., Delgado López-Cózar, E. (2016). La revolución Google Scholar: destapando la caja de Pandora académica. *La revolución google scholar*, 1-268.
- Superio, D. L., Oliveros, M. G. H., Palloculo, V. E. V., & Geromiano, J. F. (2020). The information-seeking behavior of aquatic science librarians in response to a query. *Public Services Quarterly*, 16(4), 280-291.
- Saleh, A., Beck, T., Galke, L., & Scherp, A. (2018). Performance Comparison of Ad-hoc Retrieval Models over Full-text vs. Titles of Documents. In *International Conference on Asian Digital Libraries* (pp. 290-303). Springer, Cham.