

Materiales de clase de la asignatura Búsqueda y Recuperación de Información - Grado en Información y Documentación (UCM)

Cita/referencia recomendada:

Lázaro-Rodríguez, Pedro (2024). Materiales de clase de la asignatura Búsqueda y Recuperación de Información - Grado en Información y Documentación (UCM). *Zenodo*. https://zenodo.org/doi/10.5281/zenodo.13326022





Nota importante 1: se trata de materiales de clase que uso en la asignatura "Búsqueda y Recuperación de Información" del Grado en Información y Documentación de la Universidad Complutense de Madrid. Son las diapositivas de los 4 temas de la asignatura. Estos materiales no son los únicos que utilizo, de hecho se refieren solo a la parte más teórica de la asignatura, y falta mucho -por no decir todo- de lo que tiene que ver con las prácticas. También, hay que entender estos materiales en el contexto de las clases de una asignatura durante todo un cuatrimestre. En lo personal, recomiendo usar estos materiales para crear otros propios, pero no recomendaría usarlos directamente tal y como están: son un guión personal.

Incido en el Share Alike (Compartir Igual) de la licencia CC BY-NC-SA si se usan. Por último, en las clases uso las diapositivas tema por tema y parte por parte. Por ello, cada parte de cada tema tiene un índice que en este pdf unido no funciona, pero la páginación en cada parte es independiente de las otras y sí que coinciden con las del índice de cada parte. Solo he añadido marcadores en el pdf para los 4 temas y cada parte: esto facilita la navegación por el pdf.



Nota importante 2: para asegurarse de que se está usando la última versión de estos materiales, recomiendo acceder al espacio en Zenodo. En este espacio es donde está disponible la última versión de los materiales:

Lázaro-Rodríguez, Pedro (2024). Materiales de clase de la asignatura Búsqueda y Recuperación de Información - Grado en Información y Documentación (UCM). *Zenodo*.

https://zenodo.org/doi/10.5281/zenodo.13326022





BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Temas 1 y 2 – Primera parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Diferencias entre búsqueda y recuperación de información (ByR)	3
La recuperación de información	10
La recuperación de información	
Los crawlers y el webcrawling	
Los crawlers: stemming y stopwords	
El crawler de Google	
Ejercicio de keywords	
Ejercicio de keywords (voyant-tools.org)	



Diferencias entre búsqueda y recuperación de información (ByR)

Algunas ideas del texto de Fadaie (2012):

"Information Seeking, Information Retrieval: Philosophical Points"



"La información en la **búsqueda de información** puede ser vaga, mientras que en la **recuperación de información** es clara"



"La información en la **búsqueda de información** se refiere a las personas usuarias que necesitan información, mientras que en la **recuperación de información** comprende a las personas usuarias y a las propietarias de la información"



"En la **búsqueda de información** se puede llegar al resultado, mientras que en la **recuperación de información**, por lo general se debe obtener el resultado, porque se supone que la información existe y se busca su lugar de almacenamiento"



"En la **búsqueda de información**, estar en camino de encontrar algo es importante por sí mismo, porque estar en camino de alcanzar algo aunque no se esté seguro del resultado es parte de la tarea de **búsqueda de información**, incluso si no se satisface su necesidad, mientras que en la **recuperación de información** alcanzar el resultado es importante y vital"



"En la **recuperación de información**, la clasificación se crea antes porque la parte responsable de la recuperación pone la información creada en algún lugar y sabe que las personas usuarias la necesitan y, por lo tanto, la pone conscientemente en el lugar correcto a mano o mediante un sistema"



"En la recuperación de información se asume que la información existe y que la persona bibliotecaria o especialista en información es responsable de recuperarla o de introducir un sistema organizado para que cada persona usuaria pueda buscar y encontrar su necesidad fácilmente"



Un ebook o manual muy bueno:

Blázquez-Ochando, **Manuel** (2013). *Técnicas avanzadas de recuperación de información: Procesos, técnicas y métodos*, mblazquez.es, ISBN 978-84-695-8030-1.

https://hdl.handle.net/20.500.14352/36646



"Conjunto de **procesos**..."



"desde la generación de las **colecciones**, su depuración, indexado, tratamiento textual, clasificación, almacenamiento..."

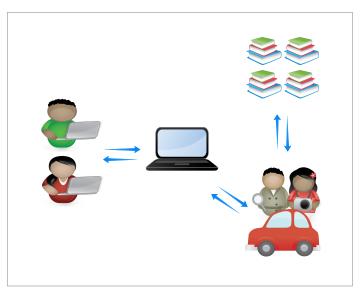


"[Pasando por la] **recuperación** mediante modelos booleanos, vectoriales, probabilísticos, basados en el lenguaje..."



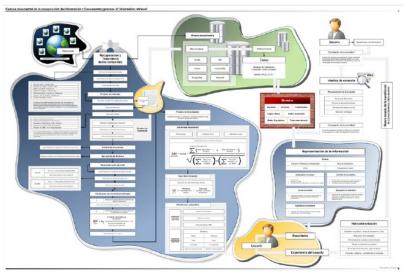
"Así como todos aquellos elementos que inciden en cualquier aspecto relacionado como por ejemplo el **interfaz** de consulta, el comportamiento del **usuario**, la **retroalimentación** de las consultas y la **representación** de la información"





Fuente: creación propia con iconos de Libreoffice





Fuente: <u>Blázquez-Ochando (2013)</u> - Open Access



Los crawlers y el webcrawling

- Software para el rastreo y análisis de sitios
- De web a web; de enlace a enlace...
- Stopwords o palabras vacías
- Proceso de stemming: supresión de prefijos, sufijos, etc.

To Policia

Los crawlers: stemming y stopwords

- El proyecto <u>snowball.tartarus.org</u>
 - Palabras vacías
 - Sample Spanish vocabulary con stemms
 - <u>Más extendido</u>: también en versión hoja de cálculo (<u>algunos</u> ejemplos)



El crawler de Google

Vídeo de hace 10 años:

https://www.youtube.com/watch?v=Md7K90FfJhg



El crawler de Google

Vídeo de hace 4 años:

https://www.youtube.com/watch?v=0eKVizvYSUQ

1 comentario: "I like how conveniently Google skipped the part about tracking your behavior on the web to modify search results" @WalidKhan



Ejercicio de keywords

Hay que acceder a algunos formularios desde el correo electrónico



Ejercicio de keywords (voyant-tools.org)

Ejercicio similar al de ayer pero extrayendo descriptores de 3 textos (o keywords con https://voyant-tools.org y trabajando con las *stopwords* el proceso de stemming

Los formularios llegan al correo



Ejercicio de Keywords: materias y tesauros

- Ediciones, Personas y Entidades de la BNE (está en beta)
- Control de autoridades (versión antigua) BNE
- Más que control de autoridades (se llamaban así antes), ahora se llaman
 "puntos de acceso". En <u>UCM y materias</u>, por ejemplo *su()* y
- La Biblioteca de la UCM, tiene <u>una guía</u> desde 2021 para controlar los puntos de acceso cuando se cataloga
- Tiene incluso <u>formularios</u> de propuestas para añadir nuevos



FIN

Pedro Lázaro Rodríguez







BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Temas 1 y 2 – Segunda parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Zotero	3
Recuperación de información: conceptos clave	9
Precisión	10
Exhaustividad	13
Pertinencia	16
Ruido (documental)	19
Relevancia	22
¿Cómo mejorar en exhaustividad?	26
¿Cómo mejorar en precisión?	28
Practicando en Google	
Campos y búsquedas en Google	30
Solucionando ecuaciones invertidas	
Un paréntesis.	51



Zotero

Zotero y sus funciones



Zotero: instalación

- Tres componentes:
 - Software (Zotero en sí)
 - Integrador (procesador de texto)
 - Conector (navegador)



Zotero: uso

Desde el navegador accedemos a documentos de muchos tipos → con el conector los mandamos a nuestro Zotero → En el procesador de texto los buscamos para las citas y las referencias



Zotero: uso

- Conviene ir clasificando todo en carpetas y subcarpetas (por temas, recursos de información, fuentes, etc.)
- Conviene revisar los metadatos en Zotero (software) antes de usarlos en el procesador de texto (es común que haya errores en apellidos, años ausentes, nombres de revistas, etc.)



Zotero: uso

 Es una forma de automatizar la creación de citas y referencias en nuestros textos



Zotero: versión web u exportación/importación

- Zotero ofrece 300 megas para una colección en línea. Es útil para no perder nada y que se quede guardada en la nube, pero es una opción limitada al tamaño
- En versión de escrito, no en línea, se puede tener una colección ilimitada en tamaño, y se puede exportar e importar sin dificultad



Recuperación de información: conceptos clave

Precisión | Exhaustividad | Relevancia | Pertinencia |
Ruido (documental)



Precisión

- Medida que define cuantitativamente la relación entre los documentos recuperados y su relevancia para satisfacer las consultas de las personas usuarias y sus necesidad de información (<u>Blázquez-Ochando</u>, 2013)
 - Es algo más propio del proceso de **recuperación de información**
 - Tips: ¿Los documentos recuperados tratan sobre el tema que estoy buscando? ¿Se desvían del mismo?



¿Precisión?

Hemos lanzado una búsqueda sobre "las olimpiadas" y hemos recuperado 80 documentos. Tras analizarlos, nos quedamos como interesantes y útiles con 40...

- ¿Cuál es la precisión aquí?
- Es una medida: 40/80 y podemos plantearla como porcentaje = 50 %



¿Precisión? De manera coloquial

"El hecho de haber encontrado el mayor número de documentos que traten sobre lo que necesitamos (necesidad de información) en el sistema de recuperación de información"



Exhaustividad

- Capacidad del sistema de información para recuperar todos los documentos relevantes con respecto a la totalidad de los existentes en la colección, de acuerdo a los condicionamientos y especificaciones de la consulta del usuario (<u>Blázquez-Ochando, 2013</u>)
 - Requeriría conocer toda la base de datos del sistema de recuperación de información.
 - Terreno que abarcamos
 - Tip: ¿Existe algún documento en la base de datos sobre el tema que estoy buscando y según esta ecuación que no se haya recuperado?



¿Exhaustividad?

Nos damos cuenta, analizando el sistema de recuperación a fondo y en concreto la base de datos, de que no solo había 80 documentos sobre "las olimpiadas", existen 20 más sobre el tema que podemos recuperar si unimos "las olimpiadas" OR "deportes olímpicos"

Ganamos en exhaustividad gracias al OR, y también en precisión 60/80
 = 75 %



¿Exhaustividad? De manera coloquial

"Tiene que ver con el terreno de documentos que conseguimos abarcar con nuestras búsquedas"



Pertinencia

- El conjunto pertinente de documentos recuperados puede definirse como el subconjunto de los **documentos** almacenados en el sistema que es apropiado para la necesidad de información del usuario (Blázquez-Ochando, 2013)
 - También nos haría falta conocer toda la base de datos para concluir con seguridad



¿Pertinencia?

En nuestro caso, podríamos considerar los 60 documentos como los pertinentes (hemos tenido acceso a la base de datos y sabemos que no hay más que esos)

- Importante apreciar que son 60 sí o sí, aunque se hubiesen recuperado 40 o los 60 finales.
- Se refiere a los documentos almacenados en el sistema que son apropiados para la necesidad de información



¿Pertinencia? De manera coloquial

"Se trata de un juicio de valor (opinión) sobre si lo que hemos encontrado trata o no realmente del tema que estamos buscando"



Ruido (documental)

 Situación en la que recuperamos documentos de la base datos de un sistema de recuperación de información que no son pertinentes ni precisos a nuestros criterios de búsqueda



¿Ruido (documental)?

Tras una revisión pormenorizada, nos percatamos de que en esos 60 o 40 documentos que pensábamos iban sobre nuestro tema, hay uno que va sobre un tema totalmente diferente (edición de libros antiguos) y que no nos explicamos cómo se recupera con nuestros criterios...

• Es un ejemplo de ruido documental, quizá por un error en la catalogación (asignación de materias, etc.)



¿Ruido (documental)? De manera coloquial

"Va más enfocado a posibles errores del sistema en el momento de la catalogación, asignación de materias, el lenguaje documental que utiliza...

No depende de las personas"



Relevancia

- Grado de importancia y significación que concede la persona usuaria a los resultados obtenidos en un sistema de información con una búsqueda (<u>Blázquez-Ochando, 2013</u>)
 - La relevancia la juzgamos las personas
 - Un mismo documento puede ser y no ser relevante a la vez para diferentes personas, e incluso para la misma (en momentos diferentes) →



¿Relevancia?: un ejemplo

- Un mismo documento puede ser y no ser relevante a la vez para diferentes personas, e incluso para la misma (en momentos diferentes)
 - Realizo una búsqueda en Scopus para recuperar documentos de <u>Pedro Lázaro Rodríguez</u> y recupero 13 documentos
 - Realizo una búsqueda en <u>Google Scholar</u> para lo mismo y hay 41 documentos incluidos los de la WOS
 - Esos 13 documentos, cuando busco en Google Scholar, son precisos pero ya no relevantes (porque ya los conocía)



¿Relevancia?

Finalmente nos hemos quedado con 59 documentos que son relevantes según nuestro criterio y en base a nuestra necesidad de información

 Ahora se podría lanzar otra búsqueda en otra base de datos para alcanzar nuevos y más documentos relevantes (si encontramos alguno de los 59 que ya tenemos, estos no serían relevante porque ya los conocemos)



¿Relevancia? De manera coloquial

"Por lo general, muchos documentos resultan relevantes en un primer momento, al conocerlos. Pasa que ciertamente algunos documentos van del tema que estamos buscando, pero al final no nos resultan cruciales para satisfacer nuestra necesidad de información"



¿Cómo mejorar en exhaustividad?

En ocasiones obtenemos pocos resultados para satisfacer nuestras necesidades de información. Para superarlo:

- OR con el mismo concepto en otros idiomas
- El operador OR con sinónimos o términos relacionados
 - Uso de tesauros para encontrarlos →



¿Cómo mejorar en exhaustividad?

- Tesauro de Derechos Humanos DADH (<u>link</u>)
- Tesauro de Biodanza (<u>link</u>)
- Campos de la Investigación y el Desarrollo OCDE (<u>link</u>)
- TESAURO EN FILOSOFÍA DE LA EDUCACIÓN (link)
- 777 tesauros hechos con el software abierto TemaTres (<u>link</u>)



¿Cómo mejorar en precisión?

En ocasiones obtenemos muchos resultados para satisfacer nuestras necesidades de información. Para superarlo:

 Operadores AND y NOT, y uso de campos como en título, en abstract, keywords, authors... según el lenguaje y las posibilidades de cada sistema



Practicando en Google

Practicando búsquedas en Google: AND, OR, NOT y sobre precisión, exhaustividad, relevancia, pertinencia y ruido documental (15/02/2024)

En campus

Tho fate light

Campos y búsquedas en Google

Para ganar en **precisión**:

- Muy útiles:
 - ∘ intitle: filetype: "" AND
- Útiles:
 - link: intext: inurl: related: site:

Para ganar en **exhaustividad**:

OR



Importante:

- En minúscula desde la primera letra
- Sin espacio tras los dos puntos
- El pegado al concepto que *queremos restar*
- El espacio en blanco actúa como AND
- Tema de tildes (afectan) y mayúsculas (no afectan)



- ¿Diferencias?
 - intitle:"leonardo dicaprio" intitle:"tom hanks"
 - o intitle: "leonardo dicaprio" AND intitle: "tom hanks"

- ¿Diferencias?
 - intitle:"leonardo dicaprio" intitle:"tom hanks"
 - o intitle: "leonardo dicaprio" AND intitle: "tom hanks"

No hay diferencias pues el espacio actúa como AND



- ¿Diferencias?
 - intitle:"crisis economica"
 - o intitle:"crisis económica"
 - o intitle: "Crisis Económica"



- ¿Diferencias?
 - intitle:"crisis economica"
 - o intitle:"crisis económica"
 - o intitle: "Crisis Económica"

Hay diferencias en el caso de la tilde en *económica*; pero no en el caso de las mayúsculas y minúsculas

- ¿Correcto o incorrecto?
 - intitle:leonardo dicaprio intitle:tom hanks
 - intitle: "leonardo dicaprio" intitle: "tom hanks"



- ¿Correcto o incorrecto?
 - intitle:leonardo dicaprio intitle:tom hanks
 - intitle: "leonardo dicaprio" intitle: "tom hanks"

Incorrectas: faltan comillas en los nombres, y hay espacio después de los dos puntos

¿Son iguales estas búsquedas?

- "inteligencia artificial" -ChatGPT
- "inteligencia artificial" -"ChatGPT"

*Colin Agreement Color Pool Pice

Campos y búsquedas en Google

¿Son iguales estas búsquedas?

- "inteligencia artificial" -ChatGPT
- "inteligencia artificial" -"ChatGPT"

Teóricamente, sí. En la práctica y en Google, se obtiene diferente número de resultados → ¿inconsistencia del motor de búsqueda?

Un ejemplo en Scopus: aquí no pasa



intitle:"la soledad de los números primos" intitle:libro -película - **cine**



inurl:repositoriorebiun.org "bibliotecas universitarias" "universidad privada"



inurl:repositoriorebiun.org ("universidad privada" OR "universidades privadas" OR "biblioteca privada" OR "bibliotecas privadas")



alquiler pisos "Marina d'Or" -inurl:idealista.com -vacaciones - **verano** -vacacionales



Intitle: "Marisa Mediavilla Herreros" "Biblioteca de Mujeres"



Intitle: "Marisa Mediavilla Herreros" "Biblioteca de Mujeres"



intitle: "Marisa Mediavilla Herreros" "Biblioteca de Mujeres"



intext:"Blanca Calvo" "canon por préstamo" filetype:pdf



intitle: "María Moliner" "expediente de depuración" filetype:pdf



intitle: "Alicia Girón" inurl:bne.es OR inurl:biblioteca.ulpgc.es



(intext:"Teresa Andrés Zamora" OR intext:"Ángela García Rives") filetype:pdf



Un paréntesis

DocuFem Semana de las Mujeres 2024

Incluye la exposición <u>Bibliotecarias Audaces</u>



FIN

Pedro Lázaro Rodríguez







BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Temas 1 y 2 – Tercera parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Lenguaje natural y lenguaje documental (I)	3
Lenguaje natural → Documental	4
Lenguaje Documental → Natural	8
Lenguaje natural y lenguaje documental (II)	11
Google Scholar	
5 búsquedas en Google Scholar	
Un reto al profesor	
1	



Lenguaje natural y lenguaje documental (I)

- Lenguaje natural:
 - o El que usamos en nuestro día a día
 - Para hablar con otras personas
- Lenguaje documental:
 - o El propio del ámbito de búsqueda y recuperación de información
 - Lo usamos para hablar con los sistemas de recuperación de información



- Natural: "necesito webs sobre atajos de teclado en microsoft word"
- Documental:
 - intitle:"atajo de teclados" intitle:"microsoft word"
 - intitle:"atajo de teclados" OR intitle:shortcuts intitle:"microsoft word"
 - (intitle:"atajo de teclados" OR intitle:shortcuts) intitle:"microsoft word"



- Natural: "necesito webs sobre atajos de teclado en microsoft word"
- Documental:
 - intitle:"atajos de teclados" intitle:"microsoft word"
 - intitle:"atajos de teclados" OR intitle:shortcuts intitle:"microsoft word"
 - (intitle:"atajos de teclados" OR intitle:shortcuts) intitle:"microsoft word"



- Natural: necesito conocer que premios se suelen dar en el ámbito de las bibliotecas públicas en España
- Documental:
 - o intitle: "premio" intitle: "bibliotecas públicas"
 - o intitle: "premio" intitle: "bibliotecas públicas" "España"



- Natural: "Tengo que hacer un trabajo sobre la nueva Ley Orgánica del Sistema Universitario y necesito consultar algunas webs sobre la misma"
- Documental:
 - o intitle: "Ley Orgánica del Sistema Universitario" OR intitle: "LOSU"
 - intitle:"Ley Orgánica del Sistema Universitario" OR intitle:"LOSU" filetype:pdf



Lenguaje Documental → **Natural**

- Documental: intitle:"premio" intitle:"bibliotecas públicas" "España"
 - Natural: aquí tenemos a una persona que está buscando información sobre premios en bibliotecas públicas. Estas sueles recibir pero, y también, dar premios. Esta búsqueda sería difícil de precisar, porque las palabras vacías en recuperación de información muchas veces no se tienen en cuentan. No obstante, las resultados que se alcancen seguro son relevantes. Además, esta persona ha matizado España en cualquier parte de las webs. Quizá no haría falta, pues la búsqueda la plantea en español. Pero viene bien para sacar de los resultados a otros países que usan el español



Lenguaje Documental → **Natural**

- Documental: (intitle:"atajos de teclados" OR intitle:shortcuts)
 intitle:"microsoft word"
 - Natural: "aquí hay 2 conceptos clave: atajos de teclado, y Microsoft Word. La necesidad de esta persona es conocer cómo son los atajos de teclado (teclas rápidas) en el procesador de texto Microsoft Word. Ha hecho muy bien porque se ha dado cuenta que "atajo de teclado" quizá es poco usado. La palabra *shortcut* se usa mucho también en este contexto en España. Como ha unido uno y otro concepto, encontrará, o bien de uno, o de otro, o de los dos"



Lenguaje Documental → **Natural**

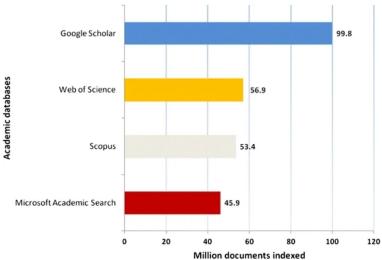
- Documental: intitle:"Ley Orgánica del Sistema Universitario" OR intitle:"LOSU" filetype:pdf
 - Natural (con todo detalle): "me atrevo a decir que el concepto principal de esta búsqueda es la Ley Orgánica del Sistema Universitario". Esta persona está buscando información sobre esta ley. Como la ley también se conoce por sus siglas, la persona ha añadido LOSU como sinónimo de la Ley. Ha hecho bien, porque mucha gente nombra a la Ley como LOSU. El OR sirve para unir sinónimos o conceptos relacionados. Además, ha añadido una clausula más para que los resultados estén en pdf. En definitiva, el tema central es la LOSU, con la condición de que los resultados sean en PDF. La necesidad de información sería encontrar información en pdf sobre la LOSU"



Lenguaje natural y lenguaje documental (II)

- En el día a día de un servicio de documentación (biblioteca pública, universitaria, etc.)
- Las necesidades de información llegan planteadas de una forma muy abierta
- Tenemos que traducirlas al lenguaje documental de cada recurso de información





Tamaño en 2014-2015. Fuente: https://doi.org/10.1007/s11192-015-1614-6



"We can assume that Google Scholar, with **389 million records**, provides by far the greatest volume of scholarly information..."

Fuente: Gusenbauer (2019)

Tolo Valeria do Carro

- Sintaxis de las búsquedas similar a Google
- Campos clave:
 - ∘ **intitle:** → en el título
 - o author: → para buscar autores/as concretos/as
 - ∘ **source:** → fuentes (revistas científicas) concretas
 - ∘ Además de: "" **AND** (espacio en blanco)



- filetype: → se puede obviar, pues los trabajos académicos se suelen publicar en pdf
- intitle: → para búsquedas exploratorias, por temas, no sabemos quién publica sobre el tema, ni dónde se suele publicar
- author: y source: → búsquedas no tan exploratorias, más analíticas.
 De alguna forma, implican conocer a autores/as y saber que publican sobre un tema concreto; implica conocer las revistas



Para buscar autores/as es poco habitual tener la necesidad de recuperar documentos de personas que se llaman, por ejemplo, Javier, o María, pero con **author:** podríamos hacerlo:

- <u>author:Javier</u>
- En realidad, solemos buscar por nombre y apellidos concretos:
- author: "Javier López Gijón"



Para buscar revistas con algún concepto clave, sí que es más habitual usar el campo **source:**

- source:"dopamine"
- Esta búsqueda recupera documentos académicos cuyas *revistas* incluyen en su título el concepto de *dopamine*. Nos ayuda a encontrar revistas sobre el tema, en las que se publican trabajos sobre lo que queremos encontrar



- Búsquedas exhaustivas
- Si quiero buscar un concepto formado por más de un término, por ejemplo, traducción automática, escribiré: "gamificación educativa", con comillas.



- Búsquedas más precisas
- intitle: "gamificación educativa"



- Búsquedas más precisas y a la vez más exhaustivas
- <u>intitle:"gamificación educativa" OR intitle:"educational gamification"</u>

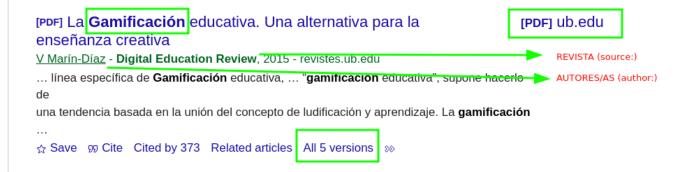


- Búsquedas más analísticas (conocemos autores/as y revistas)
- intitle: "gamificación educativa" author: "Ana Manzano León"
 - Requiere saber el nombre de la autora

To Politic

- Búsquedas más analísticas (conocemos autores/as y revistas)
- intitle: "gamificación" source: "Digital Education Review"
 - Requiere saber que hay una revista que se llama "Digital Education Review"







Algunos ejemplos de búsquedas:

- intitle:"Quentin Tarantino" -"pulp fiction"
- author: "Juan Antonio Martínez Comeche"
- <u>author:"Juan Antonio Martínez Comeche" source:"Investigación bibliotecológica"</u>



Algunos ejemplos de búsquedas:

- (intitle:comics OR intitle:comic OR intitle:manga) (intitle:"bibliotecas públicas" OR intitle:"bibliotecas públicas" OR intitle:"public library"

 OR intitle:"public libraries")
 - Por un lado: (intitle:comics OR intitle:comic OR intitle:manga)
 - AND
 - Por otro: (intitle:"bibliotecas públicas" OR intitle:"bibliotecas públicas" OR intitle:"public library" OR intitle:"public libraries")



5 búsquedas en Google Scholar

[ByR-23-24] Practicando en Google Scholar (29/02/2024)

En el campus



Un reto al profesor



FIN

Pedro Lázaro Rodríguez







BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Temas 1 y 2 – Cuarta parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Búsquedas en sitios desde Google con site:	3
Un cuestionario con 5 necesidades de información	10



El operador/campo → **site:**



- Permite delimitar las búsquedas a un sitio web concreto
 - + Precisión
- Es como si tuviésemos un buscador propio en cada sitio web
- Requiere conocer el sitio web en el que queremos buscar



Pueden plantearse búsquedas en base a intereses particulares:

site:wikipedia.org intitle:"Elvira Lindo"

Recupera páginas de Wikipedia que van sobre Elvira Lindo

site:wikipedia.org intext:"Elvira Lindo"

Recupera páginas de Wikipedia que mencionan a Elvira Lindo



Pueden plantearse búsquedas en base a intereses particulares:

site:wikipedia.org intitle:"Fahrenheit 451"

Recupera páginas de Wikipedia que van sobre el libro Fahrenheit 451

site:wikipedia.org intext:"Fahrenheit 451"

Recupera páginas de Wikipedia que mencionan al libro Fahrenheit 451



Pueden plantearse búsquedas en base a intereses particulares:

site:marca.com intitle:"Xavi Hernández"

site:mundodeportivo.com intitle:"Xavi Hernández"

Se puede analizar cómo y cuánto tratan un mismo tema diferentes medios, en este caso, del ámbito deportivo



Pueden plantearse búsquedas en base a intereses particulares:

site:okdiario.com intitle:casoplón

site:elmundo.es intitle:casoplón

site:publico.es intitle:casoplón

Se puede analizar cómo y cuánto tratan un mismo tema diferentes medios, en este caso, del ámbito ¿político?



Pueden plantearse búsquedas en base a intereses particulares:

site:elpais.com intext:"elmundo.es"

site:elmundo.es intext:"elpais.com"

site:okdiario.com intext:"eldiario.es"

En este caso, analizamos cómo se citan unos medios de comunicación a otros



Un cuestionario con 5 necesidades de información

En el campus:

[ByR-23-24] Búsquedas en sitios desde Google (06/03/2024)



FIN

Pedro Lázaro Rodríguez





BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Temas 1 y 2 – Cierre de los temas 1 y 2

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Buscadores más allá de Google: DuckDuckGo	≾
Buscadores más allá de Google: Searx	
Cerrando los temas 1 y 2	



https://duckduckgo.com/



- #1- Google te rastrea. Nosotros no.
- #2 Bloquea los rastreadores de Google que acechan por todas partes.
- #3 Obtén resultados imparciales, fuera de la burbuja de filtros.
- #4 Te escuchamos.
- #5 No intentamos atraparte en nuestro "ecosistema".

Fuente: Why Should I Use DuckDuckGo Instead of Google?



- #6 Tenemos !bangs.
- #7 Nos esforzamos por conseguir un mundo en el que tengas control sobre tu información personal.
- #8 Nuestros resultados de búsqueda no están cargados de anuncios.
- #9 Busca sin miedo.
- #10 Google es simplemente demasiado grande y demasiado poderoso.

Fuente: Why Should I Use DuckDuckGo Instead of Google?



Instalación de DuckDuckGo en Chrome, Firefox, etc:

- En Chrome: enlace (y ajustes / buscadores / DuckDuckGo)
- Firefox: <u>enlace</u> (lo mismo; aunque viene por defecto instalado)
- Edge: enlace

Ejemplos	Resultados
cats dogs	Results about cats or dogs
"cats and dogs"	Results for exact term "cats and dogs". If no results are found, we'll try to show related results $$
cats <mark>-dogs</mark>	Fewer dogs in results
cats <mark>+</mark> dogs	More dogs in results
cats filetype:pdf	PDFs about cats. Supported file types: pdf, $doc(x)$, $xls(x)$, $ppt(x)$, $html$
dogs site:example.com	Pages about dogs from example.com
cats -site:example.com	Pages about cats, excluding example.com
intitle:dogs	Page title includes the word "dogs"
inurl:cats	Page url includes the word "cats"

Fuente: enlace



- Los <u>!bangs</u> → son atajos de buscadores y hay más de 13.000
- Habitualmente buscamos en Google la base de datos / accedemos / y buscamos en ella
- Con los !bangs no ahorramos el paso de buscar la base de datos y acceder

web de !bangs



!Bangs	Busca directamente en
!wes concepto	Wikipedia en español
!cc concepto	Creative Commons
!worldcat <i>concepto</i>	Catálogo Worldcat
!pm o !pubmed concepto	Pubmed (info médica)
!rae <i>concepto</i>	RAE
!rag concepto	RAG de Galicia
!dialnet <i>concepto</i>	Dialnet
!wosc concepto	Web of Science
!deepl concepto o !deepleen concepto	En el traductor Deepl (es - en)



Algunas pruebas: lanza en duckduckgo una búsqueda de tu tema del trabajo (o alguno similar o adaptado al idioma que mejor resulte) con los !bangs de la tabla de la diapositiva anterior. Por ejemplo:

- !wes "cine mudo"
- !cc "cine mudo"
- !worldcat "cine mudo"
- ...



Breve ejercicio con algunos !bangs de DuckDuckGo:

Conociendo los !bangs de DuckDuckGo (07/03/2024)



Buscadores más allá de Google: <u>Searx</u>

- Es en realidad un metabuscador
- Es un software y existen diferentes instancias
- Lista de <u>instancias</u>:
 - Un ejemplo
 - Con opciones avanzadas para seleccionar buscadores



Cerrando los temas 1 y 2

Búsquedas en Google:

- De lo más exhaustivo a lo más preciso → intitle: y filetype:
- En sitios concretos: site:

Búsquedas en Google Scholar:

• intitle: y en menor medida author: y source:

Tiempo para el trabajo



FIN

Pedro Lázaro Rodríguez







BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Tema 3 – Primera parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Dialnet	3
Dialnet: historia	5
Dialnet: ¿Que revistas se incluyen?	6
Dialnet: documentos indexados	7
Dialnet: búsqueda y recuperación de información	8
Registro en Dialnet (muy recomendado)	9
Alta de usuarios y usuarias (con cuenta de la UCM)	9
Pasos para lanzar búsquedas y recupera información con éxito	12
Practicando en Dialnet	13



Dialnet

Es uno de los mayores portales bibliográficos del mundo, cuyo principal cometido es dar mayor visibilidad a la literatura científica hispana. Incluye:

Base de datos | Servicio de alertas bibliográficas | Hemeroteca virtual |
 Depósito o repositorio de acceso

Fuente: <u>Dialnet</u>



Dialnet

Es una de las mayores bases de datos de contenidos científicos en lenguas iberoamericanas y cuenta con diversos recursos documentales:

Artículos de revistas | Artículos de obras colectivas | Libros | Actas de Congresos | Reseñas bibliográficas | Tesis doctorales

Fuente: Dialnet



Dialnet: historia

El Proyecto se inició en el año 2001 en la Universidad de La Rioja, y en estos momentos, colaboran las siguientes instituciones:

- <u>Bibliotecas Universitarias</u>
- Bibliotecas Públicas
- <u>Instituciones y Bibliotecas Especializadas</u>

Fuente: <u>Dialnet</u>



Dialnet: ¿Que revistas se incluyen?

Las revistas son incluidas en Dialnet por las instituciones colaboradoras, que deciden qué revistas indexar

Algunos criterios:

 Acceso Abierto | Contenido de los sumarios en formato electrónico accesibles por OAI-PMH Priorizando OJS | España, espacio iberoamericano y revistas editadas en cualquier otro país de ámbito hispano | y con autoría mayoritaria de investigadores de universidades españolas o revistas de gran interés para la propia institución

Fuente: Dialnet



Dialnet: documentos indexados

Dialnet supera los 9 millones de documentos indexados (08/02/2024)

Fuente: <u>Dialnet</u>



Dialnet: búsqueda y recuperación de información

- Se recomienda registrarse en Dialnet
- Las opciones de filtros y búsqueda aumentan
- Para las búsquedas, se reconoce el AND y OR, " ", y paréntesis
- Pero cuidado: no reconoce el NOT y el llamado comodín * "sólo funciona aplicado a la búsqueda de términos individuales, no para frases"

Fuente: Dialnet



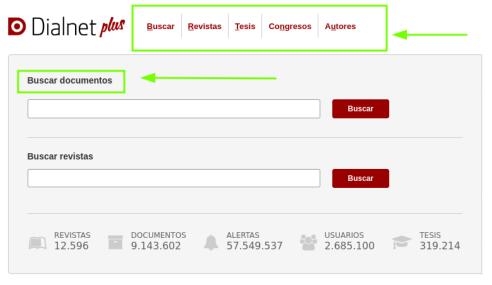
Registro en Dialnet (muy recomendado)

Alta de usuarios y usuarias (con cuenta de la UCM)

https://dialnet.unirioja.es/servlet/alta



Dialnet: búsqueda y recuperación de información

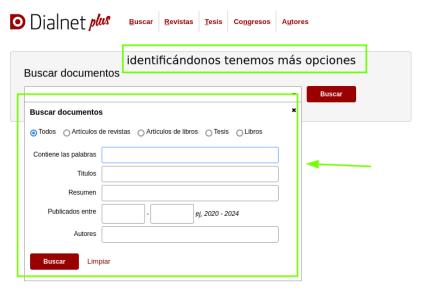


Captura de Dialnet





Dialnet: búsqueda y recuperación de información



Captura de Dialnet

- <u>Pedro Lázaro Rodríguez</u> – Creative Commons <u>Atribución-NoComercial-CompartirIgual 4.0 Internacional</u> -



Pasos para lanzar búsquedas y recupera información con éxito

En el campus



Practicando en Dialnet

Búsqueda y recuperación de información en Dialnet (13/03/2024):

En el campus



FIN

Pedro Lázaro Rodríguez







BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Tema 3 – Segunda parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Web of Science (WOS)	3
Scopus	
Proquest	
Sobre WOS y Scopus y Proquest	
Breve ejercicio sobre WOS, Scopus y Proquest	



Web of Science (WOS)

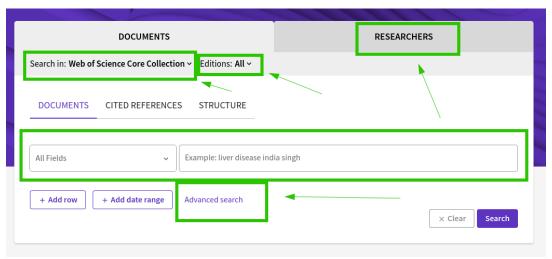
- Múltiples bases de datos, índices (index) o editions
- Clarivate tiene todavía más servicios
- Posibilidad de buscar documentos o investigadores/as
- Motor de búsqueda por cuadros (búsquedas más simples) y motor de búsqueda avanzado (tipo google expertise)



Web of Science (WOS)

- En realidad, WOS -históricamente hablando- es una base de datos bibliográfica: para buscar información sobre documentos
- Aunque es muy importante en evaluación de la ciencia, cada vez se discute más sobre su papel en ello
- Más información en la FECYT: enlace
- Suscripción de pago





Captura del motor de búsqueda de la WOS



En la búsqueda avanzada tenemos los campos y operadores disponibles (¿en verde los más relevantes?):

Booleans: AND, OR, NOT Field Tags:

- TS=TopicTI=Title
- AB=Abstract
- AU=[Author]AI=Author
- Identifiers

 AK=Author
- Keywords

 GP=[Group Author]
- ED=Editor
- KP=Keyword Plus [®]
- SO=[Publication Titles]
- o DO=DOI
- PY=Year Published
- CF=Conference

- AD=Address
- OG=[Affiliation]
- OO=Organization
 SG=Suborganization
- SA=Street Address
- CI=City
- PS=Province/State
- CU=Country/Region
 ZP=Zip/Postal Code
- FO=Funding Agency
- FG=Grant Number
- FD=Funding Details
 FT=Funding Text
- SU=Research Area
- WC=Web of Science

- Categories ☑

 IS=ISSN/ISBN
- UT=Accession
 Number
- PMID=PubMed ID
- DOP=Publication Date
- LD=Index Date
- PUBL=Publisher
- ALL=All Fields
 FPY=Final
- publication year
- EAY=Early Access Year
- SDG=Sustainable Development Goals

Captura de los campos disponibles en WOS



Algunos ejemplos:

TI=("public libraries")



Algunos ejemplos:

TI=("public libraries" OR "bibliotecas públicas")

TI=("public libraries") OR TI=("bibliotecas públicas")



Captura de los resultados de 2 búsquedas en WOS



Algunos ejemplos:

TI=("Quentin Tarantino")



Algunos ejemplos:

TI="Quentin Tarantino" AND AB="Django Unchained" AND WC=Communication

TI=("Quentin Tarantino") AND AB=("Django Unchained") AND WC=(Communication)

aro Laterial docente

Web of Science

Algunos ejemplos:

TI=("Bob Dylan") OR AK=("Bob Dylan")

Para recuperar documentos con un concepto en el título o con ese concepto en las "palabras clave" de autores/as*

*Aclaración sobre Author Keyword y Keyword Plus



Extrapolando a tu tema:

TI=("tu tema en español" OR "tu tema en inglés") OR AK=("tu tema en español" OR "tu tema en inglés")



Operadores:

- AND
- OR
- NOT



Aunque diferente, motor de búsqueda y posibilidades muy similares:



Captura del motor de búsqueda de Scopus



- Aunque pueden considerarse competencia (Scopus VS WOS), son muy similares también
- Su esencia histórica era como base de datos bibliográfica
- Más amplia que Web of Science, al menos hasta hace poco
- Suscripción de pago como la WOS



- How can I best use the Advanced search?
 - Operators
 - Field codes
 - Find exact or approximate phrases and words
 - Advanced search error messages
- Hay en torno a 50 campos para buscar

Los principales en nuestro contexto:

- TITLE() → en título
- ABS() → en resumen
- KEY() → en keywords (palabras clave)
- AUTH() → autores/as
- AUTHKEY() → en keywords dadas por autores/as



Operadores:

- AND
- OR
- AND NOT



Algunos ejemplos:

TITLE("public libraries")

TITLE("public libraries" OR "bibliotecas públicas")

TITLE("public libraries") OR TITLE("bibliotecas públicas")



Algunos ejemplos:

TITLE("Bob Dylan") OR AUTHKEY("Bob Dylan")

Para recuperar documentos con un concepto en el título o con ese concepto en las "palabras clave" de autores/as



Extrapolando a tu tema:

TITLE("tu tema en español" OR "tu tema en inglés") OR AUTHKEY("tu tema en español" OR "tu tema en inglés")



- ProQuest Researh Library es una base de datos multidisciplinar
- Incluye cientos de revistas de distinto rango (académicas y no académicas) gran parte de ellas a texto completo
- Abarca un amplio espectro de materias: economía, ciencia política, literatura, psicología, artes, educación, sociología, salud, etc.

Fuente: <u>biblioguias.ucm.es</u>



Proquest UCM



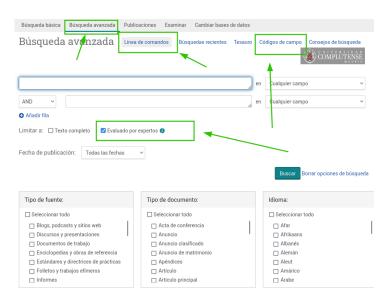
Captura de Proquest desde el acceso UCM



- Acceso desde este enlace
- También requiere suscripción
- Se considera multidisciplinar
- Para nuestra área, tiene algunas bases de datos muy relevantes:
 - <u>Library & Information Science Collection (1969 actualidad)</u>
 - <u>Library Science Database (1970 actualidad)</u>



La búsqueda avanzada tiene muchas posibilidad (¿sentido negativo?)



Captura de la búsqueda en Proquest



- Nos interesa la búsqueda por "Línea de comandos"
- Las opciones de campos son de nuevo múltiples:
 - Common Field Codes
- A resaltar: no hay un campo propio de palabras clave, pero sí de SUBJECT, KEYWORDS/IDENTIFIERS, ABSTRACT y otros relevantes
- Funcionan el AND, OR, NOT y otros operadores



Algunos ejemplos:

TITLE("public libraries")

TITLE("public libraries" OR "bibliotecas públicas")

TITLE("public libraries") OR TITLE("bibliotecas públicas")



Algunos ejemplos:

TITLE("Bob Dylan") OR SUBJECT("Bob Dylan")

TITLE("Bob Dylan") OR ABSTRACT("Bob Dylan")

Para recuperar documentos con un concepto en el título o con ese concepto en materias o en el resumen



Extrapolando a tu tema:

TITLE("tu tema en español" OR "tu tema en inglés") OR ABSTRACT("tu tema en español" OR "tu tema en inglés")



Sobre WOS y Scopus y Proquest

- Las posibilidades son numerosas y podríamos estar toda una asignatura con una sola base de datos
- WOS, Scopus y Proquest requieren suscripción. Es decir: no son gratuitas
 - o Podemos usarlas gracias a la IP y suscripción de la UCM
- WOS y Scopus se utilizan (cada vez más discutidas) para evaluar la producción científica (bibliometría) y a investigadores/as



Sobre WOS y Scopus y Proquest

- Reflexión sobre el precio de las suscripciones
 - Ver <u>estadísticas</u> REBIUN (9.1.6.Gasto dedicado a bases de datos de pago)
 - Cuestión: ¿Tiene sentido aprender recursos de recuperación de información con suscripción de pago?
 - Como documentalistas
 - Para el futuro profesional



Breve ejercicio sobre WOS, Scopus y Proquest

Búsqueda y recuperación de información en Web of Science, Scopus y Proquest (20/03/2024)

En campus



FIN

Pedro Lázaro Rodríguez







BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Tema 3 – Tercera parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Sci-Hub a debate: ¿Sci-Hub sí o Sci-Hub no?	3
Sci-Hub como tema en la ciencia	14
Sci-Hub a debate: ¿Sci-Hub sí o Sci-Hub no?	
Extensiones útiles para el navegador	
Tiempo para el trabajo	
1 1 J	



Un recorrido por las noticias recopiladas por Julio Alonso Arévalo y entradas en <u>universoabierto.org</u>

Universo Abierto (2024). <u>Sci-hub | Resultados de la búsqueda | Universo Abierto</u>.



- La piratería en línea se extiende a los libros y materiales académicos
- Problemática en torno a SciHub, la web de descarga gratuita de artículos científicos
- El dilema de acceso a la información sanitaria a través de SciHub en países con bajos recursos económicos e importantes necesidades médicas



- Hacia un mercado competitivo y sostenible del Acceso Abierto en Europa
- El bueno, el malo y el feo: Cómo recuperar artículos científicos de Internet
- Los ingresos por ventas generados por artículos de acceso abierto crecieron más que los de revistas de suscripción



- La guerra de la ciencia abierta
- Los bibliotecarios amotinados ayudan a impulsar el cambio en la política de la multinacional Elsevier
- Plataformas piratas de Acceso Abierto como desobediencia civil: ¿Es ético violar los muros de pago de las publicaciones académicas de pago?



- El acceso abierto pirata como desobediencia civil electrónica: ¿Es ético violar las barreras de pago de las publicaciones académicas?
- Aumenta la piratería de artículos académicos Japón
- Library Genesis demandada por cuatro importantes editoriales estadounidenses por presunta infracción de derechos de autor



- La directora de investigación de Australia se enfrenta a los editores de publicaciones periódicas por el control de los conocimientos
- ¿Quién está descargando artículos científicos piratas? TODO EL MUNDO. En los países ricos y pobres, los investigadores acuden al sitio web Sci-Hub.
- Elsevier contra Sci-Hub: una sentencia obliga a Alexandra Elbakyan a indemnizar con 15 millones de dólares al gigante de la edición científica

Fuente: Universo Abierto (2024)



- Sci-Hub proporciona acceso al 85% de la literatura científica publicada
- <u>Las descargas desde la web pirata de artículos de revistas Sci-Hub se triplicaron en 2017</u>
- Sci-Hub proporciona acceso a más de dos tercios de todos los artículos científicos publicados



- El proveedor de servicios de Internet sueco Bahnhof bloquea a Elsevier por el intento de bloqueo de Sci-Hub en ese país
- <u>El Tribunal Superior de París ordena a los proveedores de servicios de Internet en Francia bloquear Sci-Hub y LibGen</u>
- <u>Efectos potenciales de Sci-Hub sobre la publicación académica</u>



- ¿Sci-hub es bueno o malo para la comunicación científica?
- El efecto Sci-hub: las descargas de Sci-hub conducen a más citas de artículos
- Conferencia magistral de Alexandra Elbakyan creadora de SCI-HUB



- Elsevier en lugar de negociar con las universidades propone un spywere contra Sci-Hub
- Caso Sci-Hub en la India: la justicia debería proteger la ciencia de editores académicos codiciosos
- Twitter cierra la cuenta de Sci-Hub



- Extensión de Sci-Hub para encontrar artículos científicos
- Estadísticas sobre el uso de Sci-Hub por Alexandra Elbakyan
- Alexandra Elbakyan y la resistencia a los cercamientos del conocimiento científico (Sci-Hub)
- Sci-Hub presenta una paradoja para la publicación en acceso abierto



Sci-Hub como tema en la ciencia

- (intitle:"sci-hub" OR intitle:scihub)
 - 155 resultados (sin citas)
- (intitle:"sci-hub" OR intitle:scihub) OR (intext:"sci-hub" OR intext:scihub)
 - 20.500 resultados



"En mi opinión, lo que es absolutamente antiético, es que haya ciudadanos que deban pagar por acceder a los resultados de investigaciones financiadas por dinero público. Y pagar cantidades importantes. Es que si alguien se para a pensar en el modelo de negocio de esas publicaciones, es demencial: No pagan los contenidos, que es lo que venden, a veces incluso cobran por publicar, si quieres el acceso abierto, casi siempre cobran, no pagan a los pares por revisar, muchas veces ni siquiera a los editores por el «prestigio» que ello conlleva... Y encima imponen medidas draconianas como tener que suscribirse a paquetes hipercaros para poder acceder a lo que interesa, embargos al autoarchivo (o sea, no puedo compartir mi propio artículo!!!!), etc... En fin, lo dicho. Demencial. Lástima que el sistema permita (y obligue) a publicar en esos medios para poder acceder a financiación, sexenios, etc..." (Un comentario de Amg en una entrada de UniversoAbierto; Amg agosto 19, 2016)



Extensiones útiles para el navegador

- https://unpaywall.org/
- https://click.endnote.com/

Tiempo para el trabajo

Algunas grandes búsquedas en bases de datos:

• Ejemplo <u>1</u>, <u>2</u>, <u>3</u> y <u>4</u>



FIN

Pedro Lázaro Rodríguez







BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Tema 3 – Cuarta parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

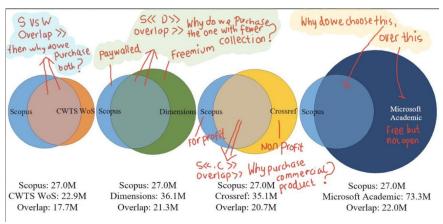
Bases de datos sin suscripción de pago	3
Dimensions	6
Lens	
Un gráfico de Lens (22/03/2014)	8
PubMed	
Algunos datos comparativos	12
Breaking research	13
Conociendo Dimensions y Lens:	18
Búsquedas en Dimensions	
Búsquedas en Lens	



Bases de datos sin suscripción de pago

- Surgen como *alternativas* muy potentes a Scopus y Web of Science
- No requieren suscripción de pago en el todo o en partes
- Algunos ejemplos:
 - Dimensions
 - Lens
 - OpenAlex

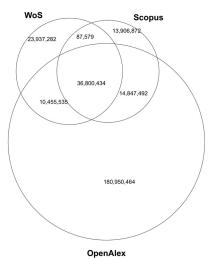
Bases de datos sin suscripción de pago



Fuente: visto en Twitter – hilo de Ludo Watlman @RINarxiv



Bases de datos sin suscripción de pago



Fuente: Culbert et al. (2024)



Dimensions

- Requiere registro con correo electrónico
- Búsquedas por campos algo limitadas
- Posibilidad de alertas
- Utilizada como instrumentos para la última convocatoria de sexenios (importante)
- Un recorrido por su interfaz: https://www.dimensions.ai/

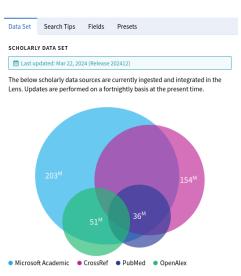


Lens

- Interfaz muy sobrecargada (quizá negativo)
- Data Partners: Microsoft Academic, CrossRef, ORCID, PubMed, Impactstory, CORE, European Patent Office (EPO), United States Patent and Trademark Office (USPTO), IP Australia y World Intellectual Property Organization (WIPO) (fuente: <u>about.lens.org</u>)
- Un recorrido por su interfaz: https://www.lens.org/



Un gráfico de Lens (22/03/2014)



Fuente: https://www.lens.org/...



PubMed

- PubMed es un recurso gratuito que facilita la búsqueda y recuperación de bibliografía biomédica y de ciencias de la vida con el objetivo de mejorar la salud, tanto a nivel mundial como personal
- Contiene más de 36 millones de citas y resúmenes de literatura biomédica. No incluye el texto completo de los artículos de revistas; sin embargo, suele haber enlaces al texto completo cuando está disponible en otras fuentes

Fuente: https://pubmed.ncbi.nlm.nih.gov/about/



PubMed

Como se puede ver en el gráfico incluido 2 diapositivas antes, Lens indexa mucho de Pubmed (<u>fuente</u>). También OpenAlex (<u>fuente</u>)

Otras bases de datos como Web of Science, permiten buscar en

MEDLINE



Fuente: captura del buscador de la WOS

PubMed

- Podemos decir que Medline es una base de datos de la U.S. National Library of Medicine (NLM)
- PubMed incluye a Medline



Algunos datos comparativos

	Number of works	Open Access works	Citations	Price	Data Openness	Org structure
<u>OpenAlex</u>	243M	48M	1.9B	Freemium	Fully open, CC0 license	Non-profit
<u>Scopus</u>	87M	20.5M (<u>ref</u>)	1.8B	Subscription	Closed	For Profit
Web of Science (core)	87M (<u>ref</u>)	12M (<u>ref</u>)	1.8B	Subscription	Closed	For Profit
<u>Dimensions</u>	135M	29M (<u>ref</u>)	1.7B	Freemium	Partly open, personal use	For Profit
Google Scholar	389M (<u>estimated</u>)	?	?	Free	Closed	For Profit
Crossref	145M	20M	1.45B	Free	Fully open, CC0 license	Non-profit

Fuente: https://openalex.org/about



Del día 2 de abril, visto el día 3:

Preprint: "Open Access Coverage of OpenAlex, Scopus and Web of Science"

<u>Simard et al. (2024)</u>



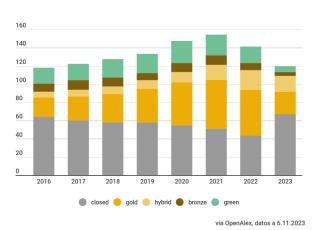
Tipos de Open Access (acceso abierto):				
Vía dorada (gold OA)	Publicado en revistas con acceso abierto, cobran APCs			
Vía diamante (diamond OA)	Gratuitos para autores/as y lectores/as			
Vía verde (green OA)	Acceso de pago en la página del editor, pero con versión abierta en repositorio institucional			
Vía híbrida (hybrid OA)	Autores/as pagan por publicar en abierto en revistas que también publican sin ser abierto			
Vía bronce (bronze OA)	Gratuito en página de editores, pero sin licencia abierta			
Cerrado (closed)	No hay Open Access			

Fuentes: biblioguias.ucm.es y Simard et al. (2024)

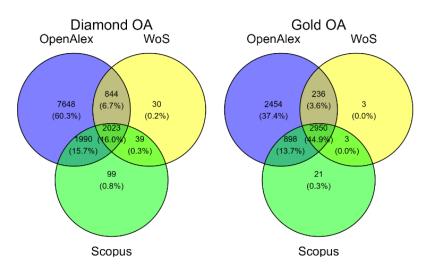


Publicación en acceso abierto en España





Fuentes: biblioquias.ucm.es



Fuente: Simard et al. (2024)



- Más OA en OpenAlex que en WOS y Scopus
- Un recurso para ver qué marcan las editoriales académicas sobre el Open Access de versiones finales, postprints y preprints:
 - https://v2.sherpa.ac.uk/romeo/



Conociendo Dimensions y Lens:

[ByR-23-24] Conociendo bases de datos sin suscripción de pago (3/04/2024) En el campus



Conociendo Dimensions y Lens:

Consejo sobre conceptos de más de una palabra:

- Hay que intentar que no excedan las 2 palabras o 3 si incluyen alguna preposición (es un consejo, no una regla)
- Por ejemplo:
 - "Estrés en estudiantes universitarios" → demasiado grande
 - Mejor: estrés AND "estudiantes universitarios"



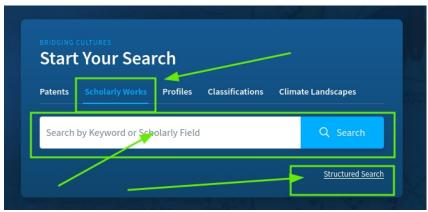
Búsquedas en Dimensions

Resolviendo ejercicios 1 y 2:

- Ejercicio 1
- Ejercicio 2

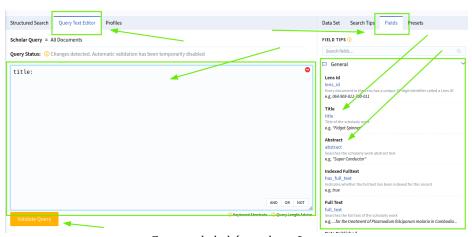


Búsquedas en Lens



Captura de la búsqueda en Lens

Búsquedas en Lens



Captura de la búsqueda en Lens



Búsquedas en Lens

Ejemplos de búsquedas:

- title: (stress AND "university students")
- title: (libraries AND "information literacy")
- title: (bibliotecas AND "alfabetización informacional")
- title: (libraries AND "information literacy") OR title: (bibliotecas AND "alfabetización informacional")



FIN

Pedro Lázaro Rodríguez



- Pedro Lázaro Rodríguez – Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional





BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Tema 3 – Quinta parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Recursos totalmente Open Access	3
DOAJ (Journals)	4
DOAB (Books)	5
Conociendo DOAJ	
Conociendo DOAB	15
Búsquedas en DOAJ y DOAB	16



Recursos totalmente Open Access

- <u>Directory of Open Access Journals: DOAJ</u>
 - 10,000,699 article records
- Directory of Open Access Books: DOAB
 - Over 80,000 academic peer-reviewed books



DOAJ (Journals)

Misión:

 Aumentar la visibilidad, la accesibilidad, la reputación, el uso y el impacto de las revistas académicas de investigación de calidad, revisadas por pares y de acceso abierto en todo el mundo, independientemente de la disciplina, la geografía o el idioma.

Fuente: https://doaj.org/about/



DOAB (Books)

Misión:

• El principal objetivo de DOAB es aumentar la visibilidad de los libros en acceso abierto. El directorio está abierto a todos los editores que publican libros académicos revisados por pares en acceso abierto y debería contener tantos libros como sea posible, siempre que estas publicaciones sean en acceso abierto y cumplan los estándares académicos.

Fuente: https://www.doabooks.org/en/doab/purpose-of-doab





Captura de la búsqueda en **DOAJ**



Se pueden buscar:

- Revistas
- Documentos (artículos, reviews, etc) en esas revistas

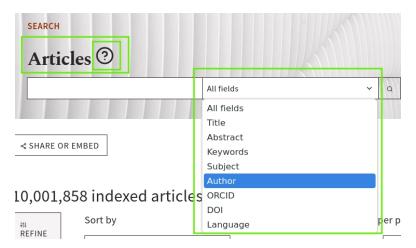


Campos y operadores disponibles para buscar <u>revistas</u>:





Campos para <u>buscar artículos</u>:





- No se pueden cruzar campos
 - Es decir: no se puede lanzar una búsqueda con XXX concepto en título AND con XXX concepto en materia.
- Pero con un mismo campo, se pueden usar operadores
- →



Operadores:

- Funcionan las comillas, y no los paréntesis
- Funcionan el OR
- Funciona el AND (y también el espacio en blanco funciona por defecto como AND)
- Aunque en las instrucciones o tips oficiales no hablan del NOT, a mí me funciona
- Se puede usar * como comodín: pero en conceptos de un palabra



Algunos ejemplos en <u>búsqueda de artículos</u>:

- <u>"software libre" OR "free software"</u> | en título
- <u>librar* AND "third place"</u> | en resumen
- (librar* OR bibliot*) AND ("third place" OR "tercer lugar") | en resumen → **no funciona**
- <u>bibliot* AND "tercer lugar"</u> | en resumen
- <u>stress AND "university students"</u> | en título
- <u>stress AND "university students" NOT "covid-19"</u> | en título



En resumen, se recomienda:

- Plantear búsquedas más simples o directas que en otros recursos
- Si requieren conceptos paralelos, mejor que sean monotemáticas o de un solo tema
- Si son de 2 temas, mejor no usar distintos idiomas en la misma búsqueda → lanzar dos búsquedas diferentes cada una en un idioma
- Sin que se necesiten paréntesis

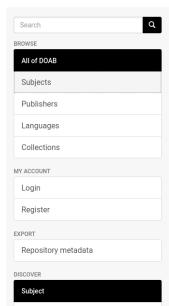


Algunos consejos en <u>búsqueda de revistas</u>:

- Lanzar búsquedas por grandes temas:
 - Capital social en bibliotecas mejor: bibliotecas
- Si son conceptos de una palabra, mejor utilizar <u>bibliot*</u>
 - Recuperaremos revistas sobre bibliotecas, bibliotecarios/as, biblioteconomía, bibliotecología, etc.
- Concepto en inglés y español: <u>bibliot* OR librar*</u>
- Uso de filtros tanto *a priori* como *a posteriori*



- Buscador menos profesional
- Es más un navegador por materias, publishers, idiomas y colecciones



Captura de **DOAB**



Búsquedas en DOAJ y DOAB

Para la semana que viene



FIN

Pedro Lázaro Rodríguez







BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Tema 4 – Primera parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Literatura gris	3
Literatura gris: patentes	
Literatura gris: patentes - modelos de utilidad	
Literatura gris: patentes – recursos de búsqueda	11
Literatura gris: normas	13
Repositorios institucionales	18
Repositorios o bancos de datos y materiales complemnetarios	21
Noticias en medios digitales	22



Literatura gris

- García Santiago, Lola (1999). Manual básico de literatura gris : el lado oscuro de la documentación. Trea.
- https://ucm.on.worldcat.org/oclc/ 1024403763



Captura del libro de Lola García Santiago

Literatura gris

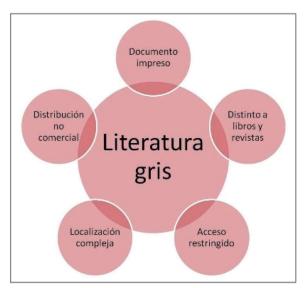


Figura 1 en Martínez-Méndez y López-Carreño (2011)



Figura 2 en Martínez-Méndez y López-Carreño (2011)

- <u>Pedro Lázaro Rodríguez</u> – Creative Commons <u>Atribución-NoComercial-CompartirIgual 4.0 Internacional</u> -



Literatura gris – ejemplos en 2004

Actas de congreso, Acuerdos de intercambio, Bibliografías, Blogs, Boletines, Catálogos, Conferencias, Directorios, Documentos de trabajo, Ensayos, Entrevistas, E-prints, Estadísticas, Guías, Informes de investigación, Informes técnicos, Legislación, Libros blancos, Libros verdes, Manuales, Materiales de cursos, **Normas, Patentes**, Pósters, Protocolos científicos, **Revistas [magazines]**, Software, Tesis doctorales, Traducciones...

Fuente: Document Types in Grey Literature (Greynet International) 2004



Literatura gris: patentes

"Una patente es un título de propiedad industrial que reconoce el derecho exclusivo sobre una invención, impidiendo a otros su fabricación, venta o utilización sin consentimiento del titular. Como contrapartida, la patente se pone a disposición del público para conocimiento general"

Fuente: OEPM (Oficina Española de Patentes y Marcas)



Literatura gris: patentes

La patente tiene carácter territorial y puede ser de tipo:

- Nacional
- Europeo
- Internacional:



Literatura gris: patentes - modelos de utilidad

Modelos de utilidad:

 La modalidad de Modelo de Utilidad protege invenciones en las que el requisito de actividad inventiva tiene un menor grado de exigencia que en el caso de una Patente. A cambio, la duración de la protección es también menor. Los procedimientos de concesión presentan también diferencias importantes

Fuente: <u>administracion.gob.es</u>



Literatura gris: patentes - modelos de utilidad

- Patente: un procedimiento, un método de fabricación, una máquina o aparato o un producto
- Modelos de utilidad: un utensilio, un instrumento, una herramienta, un aparato, un dispositivo o una parte del mismo, así como un producto químico o un alimento

Fuente: <u>administracion.gob.es</u>



Literatura gris: patentes - modelos de utilidad

- Ni por Patente ni por modelo de utilidad: descubrimientos, teorías científicas, métodos matemáticos, obras literarias, artísticas y científicas, métodos económico-comerciales, programas de ordenador, formas de presentación de información, razas animales y variedades vegetales... (fuente: administracion.gob.es)
- Un folleto de la OEPM: ¿Patentar la vida? Normas y usos europeos para patentar invenciones biotecnológicas (<u>fuente</u>)



Literatura gris: patentes – recursos de búsqueda

- En España:
 - Oficina Española De Patentes y Marcas (OEPM)
 - Buscador <u>INVENES</u>
 - o Folletos divulgativos muy útiles
- En Europa:
 - <u>European Patent Office</u> (EPO); <u>Patent search Espacenet</u>
- A nivel Internacional:
 - Organización Mundial de la Propiedad Intelectual (OMPI / WIPO);
 Patentscope

Literatura gris: patentes – recursos de búsqueda

Pero, y también:

- <u>Google Patents</u>: en título → TI=() | Por ejemplo: <u>TI=(guitarra)</u>
- Lens.org (<u>patent search</u>)
- ...



Literatura gris: normas

"ISO, la Organización Internacional de Normalización, reúne a expertos mundiales para que se pongan de acuerdo sobre la mejor manera de hacer las cosas, para todo, desde fabricar productos hasta gestionar procesos"

Fuente: iso.org



Literatura gris: normas

"Como una de las organizaciones internacionales no gubernamentales más antiguas, ISO ha hecho posible el comercio y la cooperación entre personas y empresas de todo el mundo desde 1946. Las Normas Internacionales publicadas por ISO sirven para hacer que la vida sea mejor, más fácil y más segura"

Fuente: <u>iso.org</u>



Literatura gris: normas – recursos de búsqueda

- ISO:
 - International Organization for Standardization
 - Normas ISO
- AENOR:
 - Asociación Española de Normalización y Certificación
 - Entidad Española
 - Normas UNE



Literatura gris: normas – algunos ejemplos ByD

- ISO 11620:2023 Information and documentation Library performance indicators (enlace)
- ISO 2789:2022 Information and documentation International library statistics (enlace)
- UNE-ISO 16439:2022 Información y documentación. Métodos y procedimientos para evaluar el impacto de las bibliotecas (enlace)



A medio camino entre la literatura no convencional y convencional: repositorios institucionales



Repositorios institucionales

- Máxima relevancia en el marco actual de la ciencia abierta
- Un recordatorio: Sherpa Romeo
- A día de hoy: universidades, centros de investigación, etc., tienen su repositorio; y también los hay temáticos
- Por ejemplo, en el ránking de repositorios de <u>webometrics.info</u>, en España y en Acceso Abierto listan:
 - 132 repositorios



Repositorios institucionales: Docta Complutense

El repositorio de la UCM:

- https://docta.ucm.es/home
 - Mi perfil en Docta
 - O el <u>Departamento de ByD</u>
 - O por <u>tipo de documento</u>



Repositorios institucionales: otros ejemplos

- <u>DIGIBUG</u> (Universidad de Granada)
- <u>RE-UNIR</u> (Universidad Internacional de La Rioja)
- Repisalud (ISCIII, CNIO y CNIC)



Repositorios o bancos de datos y materiales complemnetarios

- Importantes también en el marco de la ciencia abierta
 - https://figshare.com/
 - https://zenodo.org/
 - https://osf.io/
 - 0 ...



Noticias en medios digitales

En la UCM tenemos suscripción a la base de datos MyNews:

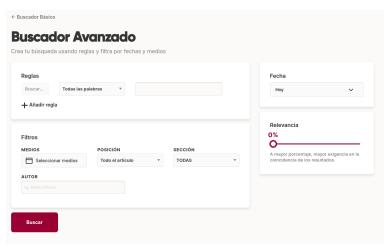
- Acceso con VPN o desde ordenadores de la facultad:
 - https://ucm.mynews.es/hu/
 - 3 tipos de búsqueda:
 - Simple, Avanzado, Profesional





Noticias en medios digitales — <u>MyNews</u>

Búsqueda Avanzada:



Captura de la búsqueda avanzada de MyNews



Noticias en medios digitales — <u>MyNews</u>

Búsqueda Profesional:

• En título:



(biblioteca OR bibliotecas)&xoptions=contentfields=title

En título y subtítulo:

("parálisis del sueño")&xoptions=contentfields=title:subtitle



FIN

Pedro Lázaro Rodríguez



- Pedro Lázaro Rodríguez – Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional





BÚSQUEDA Y RECUPERACIÓN DE INFORMACIÓN (Grupo A – 23/24):

Tema 4 – Segunda parte

#ByR

Pedro Lázaro Rodríguez





Índice / guión

Alertas	3
Alertas: Google Scholar	4
Alertas en Dialnet	
Alertas en Dimensions	12
Alertas en otras bases de datos	16
Consejos generales	17
Ejercicio sobre alertas	



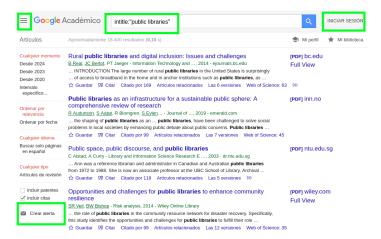
Alertas

- Se trata de guardar búsquedas en bases de datos
- Cuando se publica algo nuevo sobre ese tema, nos llega al correo
- La mayoría de recursos que hemos visto dispone de esta opción



Alertas: Google Scholar

Hay que estar registrado/a. Esta es la interfaz sin estar logeado/a:



Una búsqueda en Google Scholar

- <u>Pedro Lázaro Rodríguez</u> – Creative Commons <u>Atribución-NoComercial-CompartirIgual 4.0 Internacional</u> -



Alertas: Google Scholar

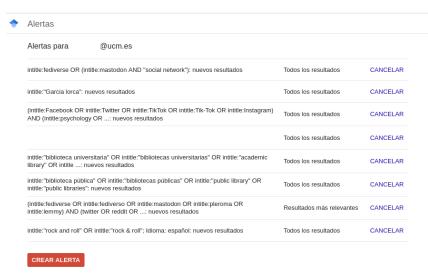
- En ese ejemplo, es una búsqueda de *intitle: "public libraries"*
- Si pinchamos en crear alerta, podemos configurar la búsqueda y la alerta en sí →
- Y ya con "crear alerta", nos irán llegando nuevos resultados al correo



Creando alertas en Google Scholar



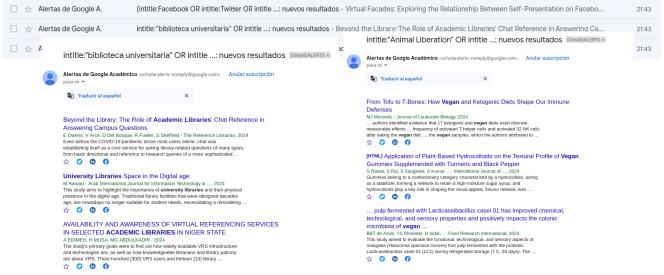
Alertas: Google Scholar – ejemplo



Alertas en Google Scholar



Alertas: Google Scholar – ejemplo



Las alertas en Gmail



Alertas: Google Scholar – recomendaciones

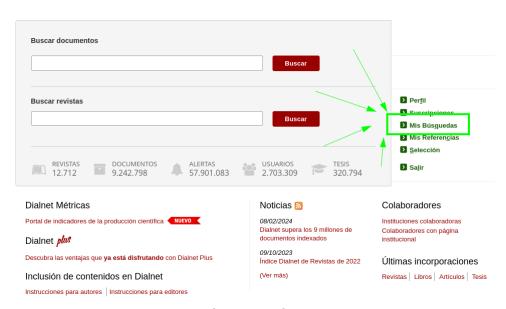
- Base de datos muy grande, quizá la que más
- Mejor crear alertas de búsquedas precisas
 - Me suelen llegar 2 o 3 correos cada semana con nuevas publicaciones
 - Para no saturar
- Crear una etiqueta/carpeta en el correo y un filtro, y dirigirlo todo a esa carpeta

[Gmail]/_PERSONAL
 [Gmail]/_REDIRIS
 [Gmail]/_SERVER
 [Gmail]/_SINDICAL
 [Gmail]/ALERTS

Captura de etiqueta



Alertas en Dialnet



Alertas en Dialnet

- <u>Pedro Lázaro Rodríguez</u> – Creative Commons <u>Atribución-NoComercial-CompartirIgual 4.0 Internacional</u> -



Alertas en Dialnet

Búsquedas guardadas

Actualmente tiene 2 búsquedas guardadas. Todavía puede crear 18 más.



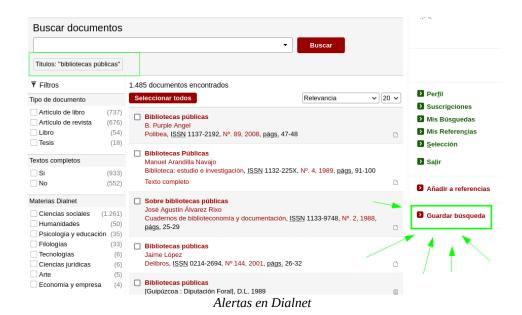
Para dejar de recibir alertas de una búsqueda guardada se deberá seleccionar la frecuencia desactivada. Para más información consulte la ayuda: dejar de recibir alertas de búsquedas guardadas.

Alertas en Dialnet



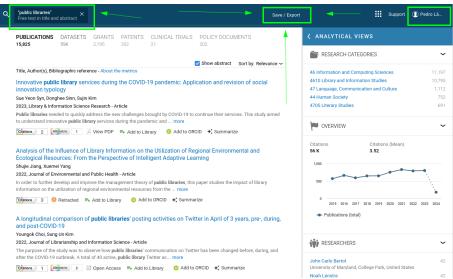


Alertas en Dialnet



- <u>Pedro Lázaro Rodríguez</u> – Creative Commons <u>Atribución-NoComercial-CompartirIgual 4.0 Internacional</u> -





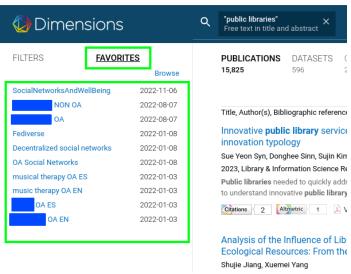
Alertas en Dimensions



Save as favorite	×
Name Public libraries	
Send me email updates for new results related to this favorite: Publications	
Cancel	Save

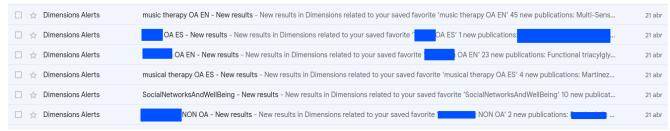
Alertas en Dimensions





Alertas en Dimensions





Captura de cómo quedan las alertas en Gmail

To la de conte

Alertas en otras bases de datos

- Lens.org, Scopus, WOS, PubMed, etc., tienen servicio de alertas
- En esta asignatura las planteamos por temas
- Pero son muy útiles en bibliotecas y centros de investigación para trackear y tener controlado la producción científica de investigadores/as



Consejos generales

- En lo personal, uso principalmente alertas de Google Scholar y de Dimensions.
- A veces me llegan de Dialnet, de Lens y de Scopus, pero no atiendo a ellas con atención
- Razón → los documentos nuevos se solapan, es decir: los que llegan, suelen llegar por muchas vías, y controlando una o dos vías, es suficiente. Por ello → alertas en Google Scholar y Dimensions
- Crucial → carpeta/etiqueta en Gmail y filtro



Ejercicio sobre alertas

Alertas en bases de datos y recursos de la asignatura (25/04/2024) En campus



FIN

Pedro Lázaro Rodríguez

