

# Indexación de Google Académico para Repositorios DSpace

Monica Westin, Google Académico

## Visión general

1. Cómo funciona el sistema de indexación de Google Académico
2. Problemas comunes de indexación en el repositorio
3. Correcciones sugeridas
4. Cómo verificar la cobertura de su repositorio en Google Académico
5. Recursos y pautas de solución de problemas
6. Preguntas

# Cómo funciona el sistema de indexación de Google Académico

- Google Académico rastrea toda la web en busca de publicaciones académicas: artículos, libros, informes, tesis, actas de congresos, preprints ...
- El sistema de indexación identifica el contenido académico, determina los metadatos bibliográficos de cada elemento y agrupa todas las versiones de un elemento junto con sus metadatos en los resultados de búsqueda.

"Directed complete poset models" of  $T_1$  spaces

Académico AÑO ▾ Las 4 versiones

Directed complete poset models of  $T_1$  spaces [PDF] nie.edu.sg

D. Zhao, X. Xi - ... Proceedings of the Cambridge Philosophical Society, 2018 - cambridge.org

A poset model of a topological space  $X$  is a poset  $P$  such that the subspace  $\text{Max}(P)$  of the Scott space  $\Sigma P$  is homeomorphic to  $X$ , where  $\text{Max}(P)$  is the set of all maximal points of  $P$ . Every  $T_1$  space has a (bounded complete algebraic) poset model. It was, however, not known whether every  $T_1$  space has a directed complete poset model and whether every sober  $T_1$  space has a directed complete poset model whose Scott topology is sober. In this paper we give a positive answer to each of these two problems. For each  $T_1$  space  $X$ , we ...

☆ 99 Citado por 1 Artículos relacionados Las 4 versiones

Mejor resultado para esta búsqueda. Ver todos los resultados

# Intentamos hacer visibles los repositorios a nivel mundial.

- Para los elementos con un único depósito (p.ej., tesis), el contenido del repositorio es el enlace principal

Enlace principal →

[THE MITOCHONDRIAL NATURE OF THE DNA REPAIR PROTEIN APE1](#)

[A Barchiesi - 2018 - iris.sissa.it](#)

Mitochondria arose around two billion years ago from the engulfment of an  $\alpha$ -proteobacterium by a precursor of the modern eukaryotic cell 1. Nowadays they are the principal source of energy production and metabolism that are fundamental processes for the survival and wellbeing of the cell. Mitochondria are organelles with a double membrane, the outer (OM) and the inner membrane (IM), and together they delimit two aqueous compartments: the matrix, in the inner side, and the intermembrane space (IMS)(Figure 1).

[☆](#) [99](#) Related articles All 3 versions [»»](#)

Showing the best result for this search. [See all results](#)

[PDF] [sissa.it](#)

← Enlace de acceso

- Para artículos publicados formalmente, el contenido del repositorio aparece como enlace de acceso y/o en "Las # versiones"

### Directed complete poset models of $T_1$ spaces

D Zhao, X Xi - ... Proceedings of the Cambridge Philosophical Society, 2018 - cambridge.org

A poset model of a topological space  $X$  is a poset  $P$  such that the subspace  $\text{Max}(P)$  of the Scott space  $\Sigma P$  is homeomorphic to  $X$ , where  $\text{Max}(P)$  is the set of all maximal points of  $P$ . Every  $T_1$  space has a (bounded complete algebraic) poset model. It was, however, not ...

☆ 99 Citado por 1 Artículos relacionados Las 4 versiones

[PDF] nie.edu.sg



## Artículos

4 resultados (0,05 s)

[Todas las versiones](#)[Directed complete poset models of  \$T\_1\$  spaces](#)

D Zhao, X Xi - ... Proceedings of the Cambridge Philosophical Society, 2018 - cambridge.org ← versión del editor

A poset model of a topological space  $X$  is a poset  $P$  such that the subspace  $\text{Max}(P)$  of the Scott space  $\Sigma P$  is homeomorphic to  $X$ , where  $\text{Max}(P)$  is the set of all maximal points of  $P$ . Every  $T_1$  space has a (bounded complete algebraic) poset model. It was, however, not ...

☆ 99 Citado por 1 Artículos relacionados

[Directed complete poset models of  \$T\_1\$  spaces](#)

D Zhao, X Xi - 2018 - repository.nie.edu.sg

A poset model of a topological space  $X$  is a poset  $P$  such that the subspace  $\text{Max}(P)$  of the Scott space  $P$  is homeomorphic to  $X$ , where  $\text{Max}(P)$  is the set of all maximal points of  $P$ . Every  $T_1$  space has a (bounded complete algebraic) poset model. It was, however, not ...

99

← versión del repositorio

[\[CITAS\] Directed complete poset models of  \$T\_1\$  spaces](#)

← cita

D Zhao, X Xi - Mathematical Proceedings of the Cambridge ..., 2018 - adsabs.harvard.edu

99

[Directed complete poset models of  \$T\_1\$  spaces](#)

D ZHAO, X XI - Mathematical Proceedings of the Cambridge ..., 2018 - search.proquest.com

A poset model of a topological space  $X$  is a poset  $P$  such that the subspace  $\text{Max}(P)$  of the Scott space [...]  $P$  is homeomorphic to  $X$ , where  $\text{Max}(P)$  is the set of all maximal points of  $P$ . Every  $T_1$  space has a (bounded complete algebraic) poset model. It was, however, not ...

99

← versión de un agregador

# Lo que necesita Google Académico para la indexación

- Acceso para rastrear el sitio.
- Una manera de encontrar todas las direcciones URL de los artículos: mapa del sitio o navegación por fecha
- Información bibliográfica en forma de etiquetas de metadatos legibles por máquina ("metaetiquetas") (activado de forma predeterminada para los repositorios DSpace desde la versión 1.7)

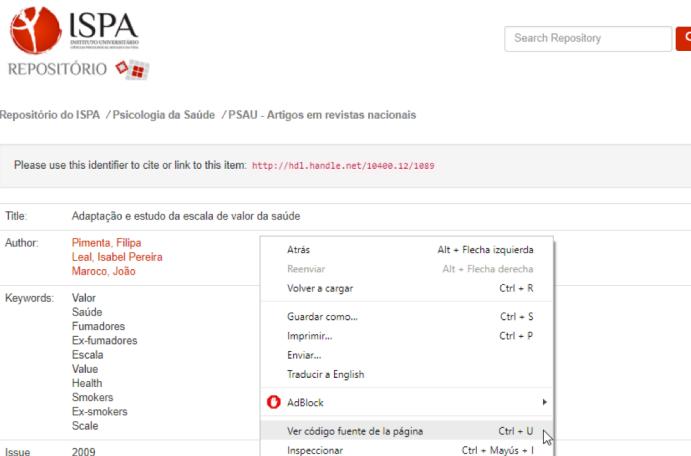
Las metaetiquetas bibliográficas indican al sistema de indexación de Google Académico cuáles son los metadatos para un elemento del repositorio: título, autor, fecha de publicación, incluyendo soporte para metadatos ORCID.

La metaetiqueta "citation\_pdf\_url" le dice al sistema de indexación qué archivo asociar con estos metadatos

```
<meta name="citation_title" content="The testis isoform of the phosphorylase kinase catalytic subunit (PhK-T) plays a critical role in regulation of glycogen mobilization in developing lung">
<meta name="citation_author" content="Liu, Li">
<meta name="citation_author" content="Rannels, Stephen R.">
<meta name="citation_author" content="Falconieri, Mary">
<meta name="citation_author" content="Phillips, Karen S.">
<meta name="citation_author" content="Wolpert, Ellen B.">
<meta name="citation_author" content="Weaver, Timothy E.">
<meta name="citation_publication_date" content="1996/05/17">
<meta name="citation_journal_title" content="Journal of Biological Chemistry">
<meta name="citation_volume" content="271">
<meta name="citation_issue" content="20">
<meta name="citation_firstpage" content="11761">
<meta name="citation_lastpage" content="11766">
<meta name="citation_pdf_url"
      content="http://www.example.com/content/271/20/11761.full.pdf">
```

# Vea el código fuente de la página del elemento para ver las metaetiquetas

Haga clic con el botón derecho para "Ver código fuente de la página", según su navegador



The screenshot shows a web page from the ISPA Repository. At the top is the ISPA logo and a search bar. Below it is a breadcrumb navigation: 'Repositório do ISPA / Psicologia da Saúde / PSAU - Artigos em revistas nacionais'. A text box contains the identifier: 'Please use this identifier to cite or link to this item: <http://hdl.handle.net/10400.12/1089>'. The main content area displays article details: Title ('Adaptação e estudo da escala de valor da saúde'), Author ('Pimenta, Filipa; Leal, Isabel Pereira; Maroco, João'), Keywords ('Valor; Saúde; Fumadores; Ex-fumadores; Escala; Value; Health; Smokers; Ex-smokers; Scale'), Issue Date ('2009'), Publisher ('Sociedade Portuguesa de Psicologia da Saúde'), Citation ('Psicologia, Saúde & Doenças, 10 (2), 217-225'), and Abstract ('O presente estudo pretende averiguar se pessoas com diferentes comportamentos de consumo de tabaco valorizam a sua saúde de forma discrepante; é igualmente objetivo desta investigação adaptar a Escala de Valor da Saúde (Lau, Hartman, & Ware, 1986) à população Portuguesa. Para tal aplicou-se o instrumento a uma amostra de 380 estudantes universitários (fumadores regulares, ocasionais e exfumadores). Os resultados demonstram que pessoas com diferentes padrões de consumo tabágico atribuem um valor...'). A context menu is open over the author's name, with the option 'Ver código fuente de la página' highlighted.

Busque en la fuente HTML "citation\_" para ver las metaetiquetas



The screenshot shows the browser's developer tools with the 'view-source' tab selected. The URL is <http://repositorio.ispa.pt/handle/10400.12/1089>. The source code is displayed, highlighting several meta tags under the 'citation\_' prefix:

```
<meta name="citation_author" content="Pimenta, Filipa" />
<meta name="citation_author" content="Leal, Isabel Pereira" />
<meta name="citation_author" content="Maroco, João" />
<meta name="citation_isbn" content="1645-0086" />
<meta name="citation_journal_title" content="Psicologia, Saúde & Doenças" />
<meta name="citation_date" content="2009" />
<meta name="citation_abstract_html_url" content="https://repositorio.ispa.pt/handle/10400.12/1089" />
<meta name="citation_lastpage" content="225" />
<meta name="citation_language" content="por" />
<meta name="citation_pdf_url" content="https://repositorio.ispa.pt/bitstream/10400.12/1089/1/PSD%202009%2010%282%29%202017-225.pdf" />
<meta name="citation_keywords" content="Valor; Saúde; Fumadores; Ex-fumadores; Escala; Value; Health; Smokers; Ex-smokers; Scale; article" />
<meta name="citation_publisher" content="Sociedade Portuguesa de Psicologia da Saúde" />
<meta name="citation_title" content="Adaptação e estudo da escala de valor da saúde" />
```

# Las metaetiquetas deben coincidir con la versión del registro:

PSICOLOGIA, SAÚDE & DOENÇAS 2009, 10 (2), 217-225

## ADAPTAÇÃO E ESTUDO DA ESCALA DE VALOR DA SAÚDE

Filipa Pimenta<sup>1</sup>, Isabel Leal<sup>1,2</sup> & João Maroco<sup>1,2</sup>

<sup>1</sup>Unidade de Investigação em Psicologia e Saúde, I&D

<sup>2</sup> ISPA

**RESUMO:** O presente estudo pretende averiguar se pessoas com diferentes comportamentos de consumo de tabaco valorizam a sua saúde de forma discrepante; é igualmente objectivo desta investigação adaptar a Escala de Valor da Saúde (Lau, Hartman, & Ware, 1986) à população Portuguesa. Para tal aplicou-se o instrumento a uma amostra de 380 estudantes universitários (fumadores regulares, ocasionais e ex-fumadores). Os resultados demonstram que pessoas com diferentes padrões de consumo tabágico atribuem um valor elevado à sua saúde, não existindo diferenças significativas entre os três grupos ( $F=1,594$ ;  $p=0,205$ ). O estudo das características psicométricas do

→ C view-source:<http://repositorio.ispa.pt/handle/10400.12/1089>

```
<meta name="citation_author" content="Pimenta, Filipa" />
<meta name="citation_author" content="Leal, Isabel Pereira" />
<meta name="citation_author" content="Maroco, João" />
<meta name="citation_issn" content="1645-0086" />
<meta name="citation_journal_title" content="Psicologia, Saúde & Doenças" />
<meta name="citation_date" content="2009" />
<meta name="citation_abstract_html_url" content="https://repositorio.ispa.pt/handle/10400.12/1089" />
<meta name="citation_lastpage" content="225" />
<meta name="citation_language" content="por" />
<meta name="citation_pdf_url" content="https://repositorio.ispa.pt/bitstream/10400.12/1089/1/PSD%202009%2010%282%29%20217-225.pdf" />
<meta name="citation_keywords" content="Valor; Saúde; Fumadores; Ex-fumadores; Escala; Value; Health; Smokers; Ex-smokers; Scale; article" />
<meta name="citation_publisher" content="Sociedade Portuguesa de Psicologia da Saúde" />
<meta name="citation_title" content="Adaptação e estudo da escala de valor da saúde" />
```

# Errores de indexación del repositorio

- Comúnmente: Metaetiquetas bibliográficas incorrectas.
  - Los sitios con errores de metadatos generalizados no pueden ser indexados
- Ocasionalmente: caídas/interrupciones del sitio mientras el sistema de indexación está buscando publicaciones en su repositorio
- Ocasionalmente: problemas con los rastreadores (*crawlers*), incluido el bloqueo del rastreador Googlebot, respuesta lenta o errores a los rastreadores, o limitaciones a la velocidad de rastreo.

## Fechas de publicación incorrectas en la metaetiqueta citation\_date

- Ocurre con mayor frecuencia cuando la fecha de subida del archivo o de puesta en línea se da como fecha de publicación cuando no existe una fecha de publicación. Pasa a menudo cuando se ha hecho cargas por lotes, por ejemplo:
  - <meta name="citation\_date" content="2014"/>

versus la fecha de publicación en el registro:

PSICOLOGIA, SAÚDE & DOENÇAS, 2009, 10 (2), 217-225

### ADAPTAÇÃO E ESTUDO DA ESCALA DE VALOR DA SAÚDE

Filipa Pimenta<sup>1</sup>, Isabel Leal<sup>1,2</sup> & João Maroco<sup>1,2</sup>

<sup>1</sup>Unidade de Investigação em Psicologia e Saúde, I&D

<sup>2</sup> ISPA

**Solución:** Fechas de publicación incorrectas en la metaetiqueta citation\_date

- Pruebe comparando la fecha de la etiqueta citation\_date con la fecha en el registro. Si la fecha que figura en la etiqueta citation\_date es posterior a la versión del registro, es probable que la fecha de puesta en línea se esté utilizando como fecha de publicación.
- Solución: parche para repositorios DSpace disponible en  
<https://github.com/DSpace/DSpace/pull/2294.patch>
  - Más información aquí: <https://jira.duraspace.org/browse/DS-4104>
  - El parche fue realizado en diciembre de 2018
  - Las versiones 5.11, 6.4 y 7.0 ya están corregidas.

# Orden de autor incorrecto en las metaetiquetas citation\_author

- Ocurre cuando las metaetiquetas del autor se enumeran fuera de orden en el código fuente, por ejemplo:
  - <meta name="citation\_author" content="Leal, Isabel" />
  - <meta name="citation\_author" content="Pimenta, Filipa" />
  - <meta name="citation\_author" content="Maroco, João" />

versus orden de autor en la versión de registro:

**Filipa Pimenta<sup>1</sup>, Isabel Leal<sup>1,2</sup> & João Maroco<sup>1,2</sup>**

## **Solución:** Orden de autor incorrecta en las metaetiquetas citation\_author

- Solución: si utiliza las versiones 5.0, 5.1, 5.2 y 5.3 de DSpace, aplique el parche disponible en: <https://github.com/DSpace/DSpace/pull/999>
  - Mas informacion aqui: <https://jira.duraspace.org/browse/DS-2679>
  - La actualización a DSpace v5.4 o superior también solucionará este problema
- Solución: ajuste la orden de metaetiqueta del autor para que coincida con el orden del autor en la versión de registro
  - <meta name="citation\_author" content="Pimenta, Filipa" />
  - <meta name="citation\_author" content="Leal, Isabel" />
  - <meta name="citation\_author" content="Maroco, João" />

Filipa Pimenta<sup>1</sup>, Isabel Leal<sup>1,2</sup> & João Maroco<sup>1,2</sup>

# Autores desaparecidos en las metaetiquetas

- Lo más común es que solo se incluya autores de la institución o solo se enumere al primer autor de una publicación, por ejemplo,
  - `<meta name="citation_author" content="Leal, Isabel" />`
- vs. **Filipa Pimenta<sup>1</sup>, Isabel Leal<sup>1,2</sup> & João Maroco<sup>1,2</sup>**
- Solución: incluya a todos los autores enumerados en la versión del registro en las metaetiquetas del autor, no solo a los autores de su propia institución. (Y nuevamente, enumere los autores en el orden en que aparecen en la versión publicada).

# Demasiados autores en las metaetiquetas

- Es común el error de enumerar al asesor como autor en una tesis, p.ej.:
  - <meta name="citation\_title" content="Jungian se acerca a tejer canastas bajo el agua"/>
  - <meta name="citation\_author" content="Estudiante Estella"/>
  - <meta name="citation\_author" content="Profesora Patricia"/>
- Solución: En la metaetiqueta de autor incluya como autor de la tesis solo al alumno que la escribió:
  - <meta name="citation\_title" content="Jungian se acerca a tejer canastas bajo el agua"/>
  - <meta name="citation\_author" content="Estudiante Estella"/>

# Información añadida al final de las metaetiquetas

- Es común el error de añadir el nombre del repositorio o el tipo de artículo a la metaetiqueta del título, p.ej.:
  - `<meta name="citation_title" content="Aproximaciones jungianas al tejido de canastas submarinas  
Repositorio de la Universidad de California del Norte" />`
  - `<meta name="citation_title" content="Aproximaciones jungianas al tejido de canastas submarinas Tesis"  
/>`
- Solución: (1) Evite agregar elementos que no sean información bibliográfica del artículo, (2) Elimine información no bibliográfica de las metaetiquetas
  - `<meta name="citation_title" content="Aproximaciones jungianas al tejido de canastas submarinas" />`

# Información multilingüe en metaetiquetas

- La combinación de diferentes alfabetos en una metaetiqueta da como resultado información bibliográfica mixta que evita que el elemento aparezca en los resultados de búsqueda de Google Académico, y además es confuso para los usuarios, p.ej.:
  - Cuando se incluye la versión traducida del título en las metaetiquetas del título:
    - `<meta name="citation_title" content="War and Peace == Война и мир" />`
  - Cuando se lista autores en el alfabeto nativo de la institución además del idioma en el que se escribió el artículo, p. ej.:
    - `<meta name="citation_author" content="Толстой, Лев Николаевич" />`
    - `<meta name="citation_author" content="Tolstoy, Lev Nikolayevich" />`

# Solución: Información multilingüe en metaetiquetas

- Solución: Utilice el idioma del texto completo / resumen para todas las etiquetas meta. No duplique la información de las metaetiquetas en múltiples alfabetos

Л. Толстой. «Война и мир».

Часть 1

Прочтите приведенный ниже фрагмент текста и выполните задания В1-В7, С1-С2.

Вследствие этого страшного гула, шума, потребности внимания и деятельности, Тушин не испытывал ни малейшего неприятного чувства страха, и мысль, что его могут убить или больно ранить, не приходила ему в голову. Напротив, ему становилось все веселее и веселее. Ему казалось, что уже очень давно, едва ли не вчера, была та минута, когда он увидел неприятеля и сделал первый выстрел, и что клочок поля, на котором он стоял, был ему давно знакомым, родственным местом. Несмотря на то, что он все помнил, все соображал, все делал, что мог делать самый лучший офицер в его положении, он находился в состоянии, похожем на лихорадочный бред или на состояние пьяного человека.

Из-за оглушающих со всех сторон звуков своих орудий, из-за свиста и ударов снарядов неприятеля, из-за вида вспотевшей, раскрасневшейся, торопящейся около орудий прислуки, из-за вида крови людей и лошадей, из-за вида дымков неприятеля на той стороне (после которых всякий раз прилетало ядро и било в землю, в человека, в орудие или в лошадь), — из-за вида этих предметов у него в голове установился свой фантастический мир, который составлял его наслаждение в эту минуту.

<meta name="citation\_title" content="Война и мир" />

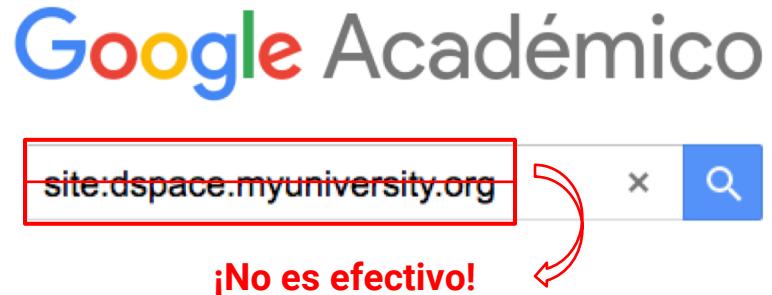
<meta name="citation\_author" content="Толстой, Лев  
Николаевич" />

# Otros errores de repositorio

- Las interrupciones prolongadas o repetidas del sitio bloquean el rastreador (*crawler*) del sistema de indexación y dan como resultado resultados de búsqueda que llevan a los usuarios a las páginas incorrectas
  - **Solución:** no permita que el sitio quede inactivo durante largos períodos de tiempo, p.ej., varios días
- Ocasionalmente, el acceso del rastreador está bloqueado por otras configuraciones del sitio
  - **Solución:** mantener la configuración predeterminada de DSpace para los rastreadores
- Mover el repositorio sin redirecciones, o renumerar elementos sin redirecciones.
  - **Solución:** configure redirecciones cada vez que cambien las URL de las publicaciones en el repositorio.
- Agregar un formulario intersticial antes previa al acceso al texto completo
  - **Solución:** omita formularios intersticiales para los usuarios que hacen clic en los resultados de búsqueda
- Agregar una página de portada a los artículos de texto completo
  - **Solución:** almacene las versiones PDF de los artículos sin modificaciones

# Cómo verificar la cobertura de Google Académico para su repositorio: Lo que **no** funciona

- El número de resultados de la búsqueda por el sitio de su repositorio ("site:XXX") no es un indicador preciso de la cobertura de Google Académico
- Esto se debe a que el número que aparece en los resultados de búsqueda de Google Académico para un sitio solo se aplica a los enlaces principales, y como se describió anteriormente, el contenido del repositorio a menudo no es el enlace principal (en "Las # versiones")



# Cómo verificar la cobertura de Google Académico para su repositorio: Lo que **sí** funciona

- Busque en Google Académico el título de algunas docenas de elementos de su repositorio seleccionados **al azar** y vea si estos documentos están incluidos.
- Asegúrese de verificar también el enlace “Las # versiones”, ya que a menudo la versión del repositorio no será el enlace principal

The screenshot shows a Google Scholar search results page for the title "3D sound scattering by rigid barriers in the vicinity of tall buildings". The results are listed vertically, each with a title, author(s), source, a brief abstract, and a link to "Citado por 53 Artículos relacionados". A red oval highlights the "Citado por 53 Artículos relacionados" link for the first result, and a red arrow points from this oval to the corresponding link in the second result below it. The third result has a similar structure but lacks the highlighted annotation.

"3D sound scattering by rigid barriers in the vicinity of tall buildings  
L Godinho, J António, A Tadeu - Applied Acoustics, 2001 - Elsevier  
The boundary element method (BEM) is used to evaluate the acoustic scattering of a three-dimensional (3D) sound source by an infinitely long rigid barrier in the vicinity of tall buildings. The barrier is assumed to be non-absorbing and the buildings are modeled as an infinite barrier. The calculations are performed in the frequency domain and time signatures are obtained by means of inverse Fourier transforms. The 3D solution is obtained by means of Fourier transform in the direction in which the geometry does not vary. This requires ...  
☆ 99 Citado por 53 Artículos relacionados [Las 7 versiones](#)

"3D sound scattering by rigid barriers in the vicinity of tall buildings  
L Godinho, J António, A Tadeu - Applied Acoustics, 2001 - Elsevier  
The boundary element method (BEM) is used to evaluate the acoustic scattering of a three-dimensional (3D) sound source by an infinitely long rigid barrier in the vicinity of tall buildings. The barrier is assumed to be non-absorbing and the buildings are modeled as an ...  
☆ 99 Citado por 53 Artículos relacionados

"3D sound scattering by rigid barriers in the vicinity of tall buildings  
L Godinho, J António, A Tadeu - 2001 - estudogeral.sib.uc.pt  
The boundary element method (BEM) is used to evaluate the acoustic scattering of a three-dimensional (3D) sound source by an infinitely long rigid barrier in the vicinity of tall buildings. The barrier is assumed to be non-absorbing and the buildings are modeled as an ...  
99

# Pautas y recursos de Google Académico para repositorios

1. Pautas de inclusión de Google Académico y pautas de solución de problemas

<https://scholar.google.com/intl/en/scholar/inclusion.html#indexing>

<https://scholar.google.com/intl/en/scholar/inclusion.html#troubleshooting>

1. “Indexing Repositories: Pitfalls & Best Practices” (Presentación de la conferencia Open Repositories 2015)

<https://www.or2015.net/wp-content/uploads/2015/06/or-2015-anurag-google-scholar.pdf>

¡Gracias por estar con nosotros!  
¿Preguntas?