DSpace and Google Scholar Webinar for Ghana

DSPACE
(Michele Mennielli)

Google scholar
(Monica Westin)

CARLIGH
(Richard Bruce Lamptey)

eifl
KNOWLEDGE WITHOUT BOUNDARIES
(Iryna Kuchma)

WACREN
(Omo Oaiya)

Search   Log in

> Mailing Lists
∨ User Groups
  • DSpace-Konsortium Deutschl...
  › German DSpace User Group
  ∨ Ghanaian DSpace User Group
    • Webinar - DSpace and Goo...
  › Grupo Argentino de Usuarios ...
  › Grupo Brasileiro de Usuários ...
  • Grupo Mexicano de Usuarios ...
  › Grupo Peruano de Usuarios d...
  › North American User Group
  › UKUsergroup
› Meetings
• Slack
• IRC
› Contribution
› Development
› Project Teams
› RoadMap
› Strategic Planning
› Archived Pages

⚙ Space tools   «

Dashboard / Home / User Groups   ···

Ghanaian DSpace User Group

Created by Michele Mennielli, last modified on Jun 12, 2019

This space is thought to make it easy for DSpace users in Ghana to share information about their activities, projects, events and ideas and to make them known to the global DSpace Community. In Ghana there are 14 installations and a vibrant community of users. This page and its sub-pages are the platform to coordinate the collaboration among Ghanaian institutions in order to facilitate discussions on Country specific themes and topics of interest.

- 1. Who is using DSpace in Ghana?
- 2. How to communicate?
  - 2.a. Mailing Lists
    - General and global DSpace lists
    - National Mailing List
  - 2.b. Slack
  - 2.c. Events
    - Webinars
    - In person meetings

1. Who is using DSpace in Ghana?

In the Duraspace Registry there are information about the DSpace installations currently running in Ghana. Please take a moment to look for your Institution and check whether the information is correct. If your institution is missing, you can add it here: https://duraspace.org/registry/register-your-site/

2. How to communicate?

There are numerous ways to get in touch, both at a global level with the general DSpace Community and at the specific country level.

2.a. Mailing Lists

A new WIKI space for the Ghanaian DSpace User Group

A new Slack Channel for instant messaging #dspace-gh

#dspace-gh
☆ | 👥 2 | 🔒 0 | ✎ Add a topic

dspace-gh

You created this channel on June 12th. This is the very beginning of the # dspace-gh channel.
Purpose: This channel is for the Ghanaian DSpace User Group to allow members of the group to easily share ideas, information, questions and best practices (edit)

+ Add an app   👤 Add people to this channel

Wednesday, June 12th

Michele Mennielli 10:17 AM
joined #dspace-gh.

Michele Mennielli 10:17 AM
set the channel purpose: This channel is for the Ghanaian DSpace User Group to allow members of the group to easily share ideas, information, questions and best practices

Today

Richard Bruce Lamptey 8:53 AM
was added to #dspace-gh by Michele Mennielli. Also, Richard Bruce Lamptey joined.

+ | Message #dspace-gh

About #dspace-gh   ✕

ⓘ Channel Details   ▾

Purpose
This channel is for the Ghanaian DSpace User Group to allow members of the group to easily share ideas, information, questions and best practices

Created
Created by you on June 12, 2019

✨ Highlights   ▸

📌 Pinned Items   ▸

👥 2 Members   ▾
  • Michele Mennielli (you)
  ○ Richard Bruce Lamptey
Add more people...

Apps   ▸

Shared Files   ▸

Notification Preferences   ▸

Google Scholar Indexing for DSpace Repositories

Monica Westin, Google Scholar
mwestin@google.com

Overview

1. How the Google Scholar indexing system works
2. Common repository indexing problems
3. Suggested fixes
4. How to check your repository's coverage in Scholar
5. Resources and troubleshooting guidelines
6. Questions

How the Google Scholar indexing system works

- Google Scholar crawls the entire web looking for scholarly publications: articles, books, reports, theses, conference proceedings, preprints ...
- The indexing system identifies scholarly content, determines each item's bibliographic metadata, and groups all versions of an item together with this metadata in search results

We try to make repositories visible worldwide

- For uniquely held items (e.g. dissertations), repository content is the primary link

Primary link →

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 α-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).

☆ 🗩 Related articles   All 3 versions   »

Showing the best result for this search. See all results

[PDF] sissa.it   ← Access link

- For formally published articles, repository content appears as the access link and/or in "All XX versions"

◆ Articles   4 results (0.04 sec)

All versions

Directed complete poset models of T_1 spaces

D Zhao, X Xi - ... Proceedings of the Cambridge Philosophical Society, 2018 - cambridge.org   ← publisher version
A poset model of a topological space X is a poset P such that the subspace Max (P) of the
Scott space ΣP is homeomorphic to X, where 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 ...
☆ ⁇ Cited by 1 Related articles

Directed complete poset models of T1 spaces

D Zhao, X Xi - 2018 - repository.nie.edu.sg   ← repository version
A poset model of a topological space X is a poset P such that the subspace Max (P) of the
Scott space P is homeomorphic to X, where Max (P) is the set of all maximal points of P.
Every T1 space has a (bounded complete algebraic) poset model. It was, however, not ...
⁇

[CITATION] Directed complete poset models of T 1 spaces

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

Directed complete poset models of T 1 spaces

D ZHAO, X XI - Mathematical Proceedings of the Cambridge ..., 2018 - search.proquest.com   ← aggregator version
A poset model of a topological space X is a poset P such that the subspace Max (P) of the
Scott space [...] P is homeomorphic to X, where Max (P) is the set of all maximal points of P.
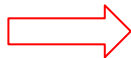Every T 1 space has a (bounded complete algebraic) poset model. It was, however, not ...
⁇

What Scholar needs for indexing

- Access to crawl the site
- Way to find all urls for articles-- sitemap or browse by date
- Bibliographic information in the form of machine-readable metadata tags ("metatags"), on by default for DSpace repositories past version 1.7

Bibliographic metatags tell the Scholar indexing system what the metadata for a repository item is: title, author, publication date, etc.

"Citation_pdf_url" metatag tells the indexing system which file to associate with this metadata

```
<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">
```

View source code from repository item page to view metatags

Right click or keyboard command to "View Page Source," depending on your browser
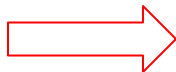
Search HTML source for "citation_" to view metatags

Metatags should match the version of record:

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

ADAPTAÇÃO E ESTUDO DA ESCALA DE VALOR DA SAÚDE

Filipa Pimenta[1], Isabel Leal[1,2] & João Maroco[1,2]

[1]Unidade de Investigação em Psicologia e Saúde, I&D
[2] 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 &amp; 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" />
```

Repository indexing errors

- Commonly: Incorrect bibliographic metatags
    - Sites with widespread metadata errors can't be indexed

- Occasionally: Site outages that occur while the indexing system is looking for publications in your repository

- Occasionally: crawler issues, including blocking the Googlebot crawler, slow responsiveness or errors to crawlers, or limiting the crawl speed.

Incorrect publication dates in citation_date metatag

- Occurs most commonly when upload/online date is given as publication date when no publication date exists, often via batch uploads, e.g.
  - `<meta name="citation_date" content="2014"/>`

  versus publication date in version of record:

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

ADAPTAÇÃO E ESTUDO DA ESCALA DE VALOR DA SAÚDE

Filipa Pimenta[1], Isabel Leal[1,2] & João Maroco[1,2]

[1]Unidade de Investigação em Psicologia e Saúde, I&D

[2] ISPA

Fix: Incorrect publication dates in citation_date metatag

- Test by comparing date citation_date tag with date in the version of record. If date listed in citation_date tag is later than the version of record, likely online date is being used as publication date.

- Fix: patch for DSpace repositories available at https://github.com/DSpace/DSpace/pull/2294.patch

   - More information here: https://jira.duraspace.org/browse/DS-4104

Incorrect author order in citation_author metatags

- Occurs when the author metatags are listed out of order in the source code, e.g.
  - <meta name="citation_author" content="Leal, Isabel" />
  - <meta name="citation_author" content="Pimenta, Filipa" />
  - <meta name="citation_author" content="Maroco, João" />

  versus author order in version of record:

  Filipa Pimenta[1], Isabel Leal[1,2] & João Maroco[1,2]

Fix: Incorrect author order in citation_author metatags

- Fix: If using DSpace versions 5.0, 5.1, 5.2, and 5.3, patch for DSpace repositories available at https://github.com/DSpace/DSpace/pull/999
  - More information here: https://jira.duraspace.org/browse/DS-2679
  - Upgrading to DSpace v5.4 or above will also fix this problem

- Fix: Adjust author metatag order to match author order in the version of record
  - <meta name="citation_author" content="Pimenta, Filipa" />
  - <meta name="citation_author" content="Leal, Isabel" />
  - <meta name="citation_author" content="Maroco, João" />

Filipa Pimenta[1], Isabel Leal[1,2] & João Maroco[1,2]

Missing authors in metatags

- Most commonly only including authors from institution or only listing the first author of a publication, e.g.
    - <meta name="citation_author" content="Leal, Isabel" />

    vs.   Filipa Pimenta[1], Isabel Leal[1,2] & João Maroco[1,2]

- Fix: Include all authors listed in the version of record in author metatags, not just authors from your own institution. (And again, list authors in the order that they appear in the published version.)

Too many authors in metatags

- Most commonly listing the advisor as an author for a thesis/dissertation, e.g.
    - <meta name="citation_title" content="Jungian approaches to underwater basketweaving"/>
    - <meta name="citation_author" content="Sara Student"/>
    - <meta name="citation_author" content="Professor Patricia"/>

- Fix: Only include one author of a thesis or dissertation in author metatag— the student who wrote it

    - <meta name="citation_title" content="Jungian approaches to underwater basketweaving"/>
    - <meta name="citation_author" content="Sara Student"/>

Trailing information in metatags

- Most commonly repository name or article type appended to title metatag, e.g.

  - <meta name="citation_title" content="Jungian approaches to underwater basketweaving Northern California College Repository/>

  - <meta name="citation_title" content="Jungian approaches to underwater basketweaving Thesis" />

- Fix: (1) Avoid adding any elements other than article bibliographic information, (2) Remove non-bibliographic information from metatags

  - <meta name="citation_title" content="Jungian approaches to underwater basketweaving" />

Multilinguals in metatags

- Combining different scripts in metatags results in mixed bibliographic information that prevents the item from being ranked in Scholar search results — and is misleading for users
- ex: including the translated version of the title in title metatags, e.g.
    - `<meta name="citation_title" content="War and Peace == Война и мир" />`
- ex: listing authors in native script of home institution when it is not the language in which the article was written, e.g.
    - `<meta name="citation_author" content="Толстой, Лев Николаевич " />`
    - `<meta name="citation_author" content="Tolstoy, Lev Nikolayevich "/>`

Fix: Multilinguals in metatags

- Fix: use the language of the full text/abstract for all metatags. Don't duplicate metatag information in multiple scripts


```
<meta name="citation_title" content="Война и мир" />

<meta name="citation_author" content="Толстой, Лев Николаевич " />
```

Site outages

- Outages result in dead pages in search results
- Extended or repeated site outages obstruct the indexing system's crawler from finding the items in your repository

- Extended outages can also result in the site dropping out of the index
- **Fix:** Don't keep your repository site down for extended periods of time, e.g. multiple days

- **Fix:** If you are migrating to a new site, keep the old site functional while developing the new version. When the new version is ready, change DNS lookup from the old server to the new one.

Example: site outage for large DSpace repository in Europe

DSpace repository items indexed by Google Scholar:

— Indexed items

Outage
occurs

Items are re-added to Scholar
index after site is restored
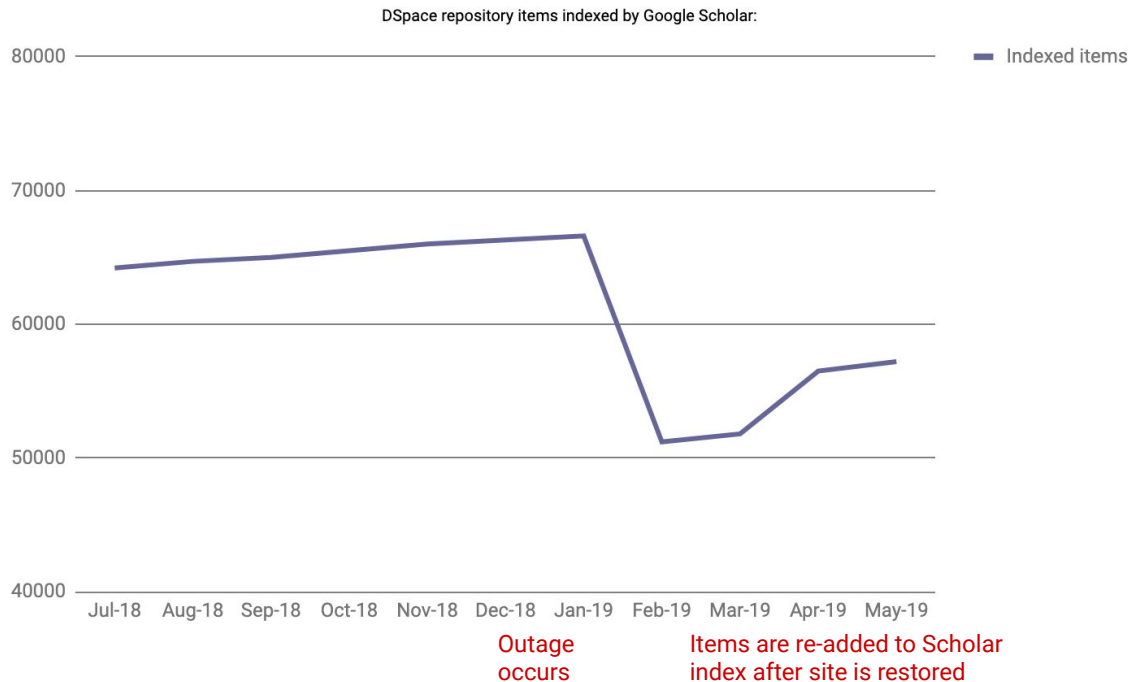
Other repository errors

- Occasionally crawler access is blocked by other site configurations
    - **Fix:** Keep default crawler settings for DSpace sites

- Repository moved without redirects, or items renumbered without redirects.
    - **Fix:** Set up redirects whenever URLs for publications in the repository change.

- Adding an interstitial for accessing fulltext
    - **Fix:** Skip interstitials for users clicking on search results

- Adding a cover page to fulltext articles
    - **Fix:** Host PDFs articles as-is

How to do a Scholar coverage check for your repository: What **doesn't** work

- The result count of searching your repository site ("site:XXX") is not an accurate indicator of Scholar coverage
- This is because the number listed in Scholar search results for a site only applies to the primary links — and as described earlier, the repository content is often not the primary link (in "All XX versions")

Google Scholar

site:dspace.myuniversity.org

Not effective!
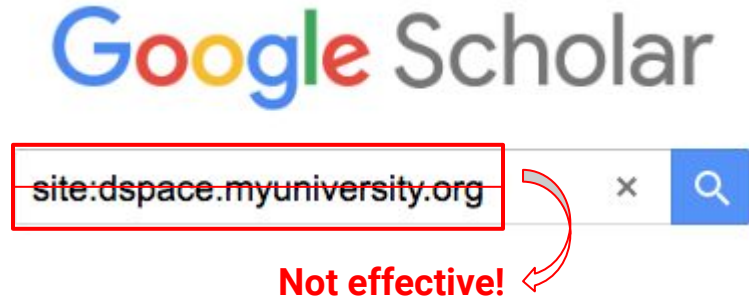
How to do a coverage check for your repository: what **does** work

- Search in Scholar for titles of several dozen **randomly** selected items across the repository and see if these papers are included
- Be sure to check the "All XXX versions" link as well, as often the repository version will not be the primary link

Google Scholar guidelines and resources for repositories

1. Google Scholar inclusion guidelines & troubleshooting guidelines

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

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

2. "Indexing Repositories: Pitfalls & Best Practices" presentation from 2015 Open Repositories conference

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

3. **Additional resources for your country**

Thank you for joining us! Questions?