

EQUELLA[®]

Discoverability Guide

Version 6.4

Document History

Document No.	Reviewed		Finalised		Published	
1	20/05/2015		20/05/2015		20/05/2015	

May 2015 edition.

Information in this document may change without notice. EQUELLA[®] and its accompanying documentation are furnished under a non-disclosure, evaluation agreement or licence agreement. Copying, storing, transmitting, or otherwise reproducing the software or this document in any form without written permission from Pearson is strictly forbidden.

All products, other than EQUELLA[®], named in this document are the property of their respective owners.

Property of:

Pearson
Level 1, 11 Elizabeth Street
Hobart, Australia, 7000

Copyright © 2015 Pearson

Contact documentation@equella.com for matters relating to documentation.

Table of Contents

Search Engine Optimisation	4
EQUELLA Configuration	4
Increasing Discoverability	7
Google Scholar	8
References	10
Contact Client Support	10

Search Engine Optimisation

Search engine optimisation, commonly referred to as SEO, is the process of increasing a website's search ranking so it appears more times in search results and thus increases the number of visitors to the site. Search engines are used to discover content and information, so ensuring that your content is correctly optimised for search engines is vital if you wish people to be able to discover your website.

This guide is divided into three sections: EQUELLA Configuration, Increasing Discoverability and Google Scholar. The first section provides information on how to configure your EQUELLA instance in order to optimise its discoverability through common search engines, such as Google. The second section contains guidelines on steps that can be performed to increase your EQUELLA instance's presence and discoverability on the web. The third section covers the requirements for content inclusion in Google Scholar.

EQUELLA Configuration

In order to ensure that your content is publically discoverable, it will need to have appropriate security permissions set in order to allow it to be discovered. That is, DISCOVER_ITEM and VIEW_ITEM privileges set to *'Everyone'*. See the *EQUELLA Security Administration Guide* for more information on configuring permissions.

Google cannot access your content if these permissions are not set, even if you have auto-login turned on, as the Googlebot does not store cookies and will always be redirected to the login page.

Use scripting to configure description meta tag information

To customise or ensure that Google displays the correct description for your resources in its search results, you can use the description meta tag to inform Google of the correct text. This can be achieved with scripting by going to the *Admin Console | Collection Definitions | Display*, and adding a new Scripting item summary display template. An example is shown in Figure 1.

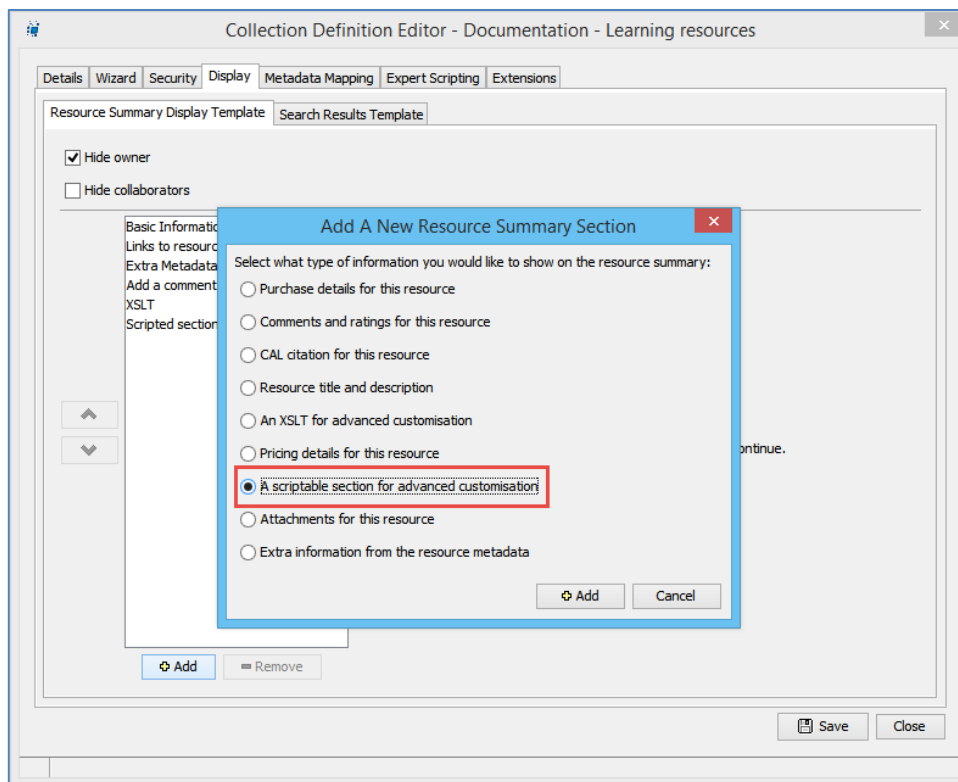


Figure 1 Add a scriptable display template

In the Pre-load Javascript section, you can use the 'meta' script object to add custom meta tags, for example:

```
meta.add("description", xml.get("/item/description"));
```

This will add a new description meta tag to the page with the content of the items /item/description field. Now when Google indexes your page it will use this meta tag for the description it shows to users. An example is shown in Figure 2.

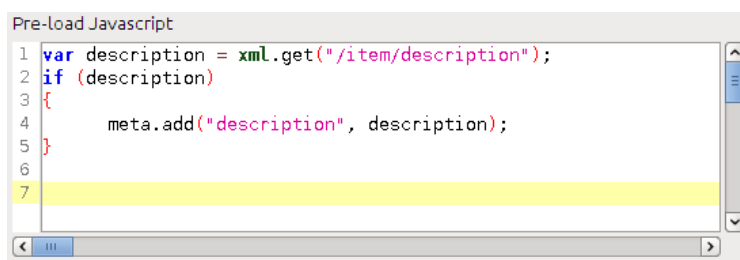


Figure 2 Adding a description meta tag with scripting

Consider adding a *Google +1* button or a *Facebook Like* button to allow users to easily share and recommend your content to other people. Both can be easily added using a scripting display template (described above), and the sample code provided by Facebook and Google, located at <http://developers.facebook.com/docs/reference/plugins/like/> and <http://www.google.com/intl/en/webmasters/+1/button/index.html>.

For example, the following will add both a *Google +1* button and a *Facebook Like* button to the resource summary when added to the HTML Markup section of the scripting display template:


```

<script type="text/javascript">
  (function() {
    var po = document.createElement('script'); po.type = 'text/javascript'; po.async =
true;
    po.src = 'https://apis.google.com/js/plusone.js';
    var s = document.getElementsByTagName('script')[0]; s.parentNode.insertBefore(po, s);
  })();
</script>

<div id="fb-root"></div>
<script>(function(d, s, id) {
  var js, fjs = d.getElementsByTagName(s)[0];
  if (d.getElementById(id)) return;
  js = d.createElement(s); js.id = id;
  js.src = "//connect.facebook.net/en_GB/all.js#xfbml=1";
  fjs.parentNode.insertBefore(js, fjs);
})(document, 'script', 'facebook-jssdk');</script>

<!-- Place this tag where you want the +1 button to render -->
<g:plusone annotation="inline"></g:plusone>
<br><br>

<!-- Place this tag where you want the Like button to render -->
<div class="fb-like" data-send="true" data-width="450" data-show-faces="true"></div>

```

An example is shown in Figure 3.



Figure 3 Simple Google+1 and Facebook Like integration with scripting

Try to ensure that your resources have titles that correctly identify their content to other users, and are short and to the point, otherwise they will be truncated or only a section of it will be shown in Google's search results.

Increasing Discoverability

In order for your site to have the greatest chance of being discovered, we recommend that you sign up and add your EQUILLA site to Google Webmaster Tools (<http://www.google.com/webmasters/>).

Once you have created an account, you will need to log in and add the site URL to Google Webmaster, then follow these recommendations in order to maximise the amount of content that Google will find.

- Google provides a number of ways for you to verify that you are the owner of the site that you are submitting. The easiest method to verify an EQUILLA site is via the 'Add a DNS record' method, assuming you have access to modify your DNS records. Google will walk you through the required steps in order to accomplish this. Once your

site is added and verified with Google Webmaster Tools, you can view a variety of statistics and information about your site.

- The best way to provide Google with a complete list of available resources in your repository is to:
 - a. Go to the **Diagnostics** page, then the **Fetch as Googlebot** tool. This allows you to enter in a URL that Google can use as a starting point.
 - b. EQUELLA provides a URL located at <http://your.equella.com/harvest.do> which will produce a list of all the resources that are accessible. Enter this URL and press **Fetch**.
 - c. After a short wait the *Fetch Status* will change to **Success**, and there will be a **Submit to index** link. Click this then select the '**URL and all linked pages**' option to submit the URLs. Google will now use the *harvest.do* page as a starting point, and from there it will be able to find all of the available content in your EQUELLA repository.
- The search rankings for your site are influenced by how many other websites are promoting or linking to it (called backlinks). Increasing these will ensure more people know about your site and can help increase search rankings. Blog postings, social media links, and other well-renowned sites pointing to your site will help direct people to your site and improve your search rankings.
- You can submit your site to the Open Directory Project (<http://www.dmoz.org/>), which is the largest free directory of sites on the internet. It is run by volunteers and provides people with an excellent way of discovering sites that they are interested in.

These are just a few basic guidelines to help you start getting your EQUELLA instance discoverable on the web, but are by no means an exhaustive or guaranteed method of driving traffic to your site. The resources listed below contain further information and suggestions for optimising your site for search engines and increasing discoverability.

Google Scholar

Google has published a list of guidelines that must be met in order for articles to be included in the Google Scholar search results. EQUELLA can adhere to these guidelines and this document will demonstrate how this can be achieved. The full guidelines are available at <http://scholar.google.com/intl/en/scholar/inclusion.html>.

To be included in Google Scholar, the following must be met:

- The articles must be in HTML or PDF format and have searchable text.
- Each article cannot be larger than 5 MB.
- Bibliographic data must be provided in the form of meta tags.
- Each article must have a clearly visible abstract.

The bibliographic data has to be provided in the form of meta tags which can be added in EQUELLA via scripting (see [Use scripting to configure description meta tag information](#) on page 4 for more information).

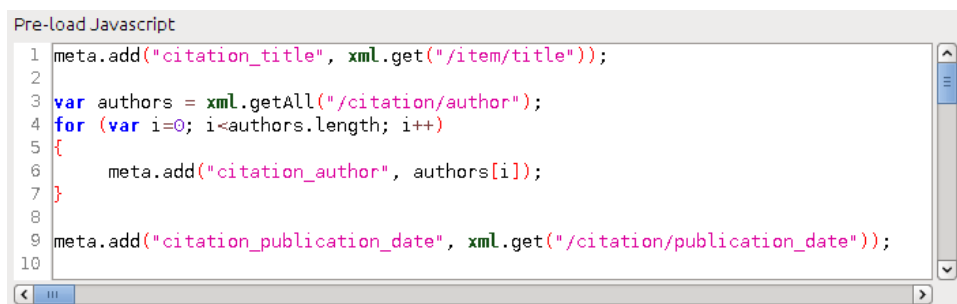
Google scholar can read a variety of different tag schemas including Highwire Press tags (citation_title), Eprints tags (eprints.title), BE Press tags (bepress_citation_title) and

PRISM tags (`prism.title`). Dublin core tags can also be used (`DC.title`) although these can be ambiguous and are not recommended for journal papers. In this guide we will be configuring EQUELLA to output Highwire Press tags as this is the schema that Google uses for its examples.

The following tags are required and if they are not present then your content will not be included by Google Scholar:

- **citation_title**—the title of the paper;
- **citation_author**—the author of the paper (can be used multiple times); and
- **citation_publication_date**—date of publication, format of YYYY/MM/DD or YYYY.

An example is shown in Figure 4.

A screenshot of a code editor window titled "Pre-load Javascript". The code is as follows:

```
1 meta.add("citation_title", xml.get("/item/title"));
2
3 var authors = xml.getAll("/citation/author");
4 for (var i=0; i<authors.length; i++)
5 {
6     meta.add("citation_author", authors[i]);
7 }
8
9 meta.add("citation_publication_date", xml.get("/citation/publication_date"));
10
```

Figure 4 Adding citation meta tags with scripting

The other tags are not required but provide more information for Google Scholar to correctly index your content:

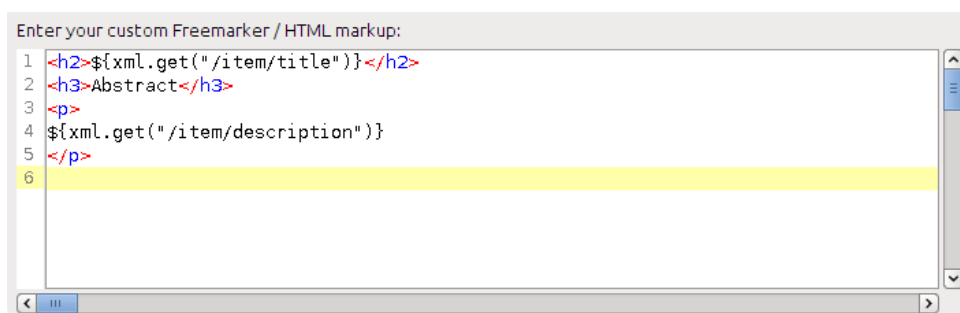
- **citation_online_date**—the online publication date (e.g. into the repository)
- **citation_pdf_url**—the location of the full paper
- **citation_conference_title**—the conference name or the proceedings title (for conference and workshop papers)
- **citation_journal_title**—the journal name (for journal papers)
- **citation_volume**—the volume (for journal papers)
- **citation_issue**—the issue number (for journal papers)
- **citation_issn**—the journal ISSN (for journal papers)
- **citation_isbn**—ISBN number
- **citation_firstpage**—the first page of the article
- **citation_lastpage**—the last page of the article
- **citation_dissertation_institution**—the university name (for Master and Ph.D. thesis)
- **citation_technical_report_institution**—the institution name (for technical reports)
- **citation_technical_report_number** – the technical report number (for technical reports).

Once added to the display template, EQUELLA will generate the meta tags and add them to the item summary in the correct format and position, as shown in Figure 5.

```
<meta name="citation_title" content="PILOT puts R&D in true perspective">
<meta name="citation_author" content="Robertson, PL">
<meta name="citation_publication_date" content="2005/12/12">
<meta name="citation_journal_title" content="The Australian Financial Review">
<meta name="citation_firstpage" content="36">
<meta name="citation_lastpage" content="36">
<meta name="citation_pdf_url" content="http://research.equella.com/items/a6c388a8-93ef-4c7d-b279-7ecba235a712/1/RobertsonAFR2005.pdf">
```

Figure 5 Generated meta tags from scripting

Your articles should also adhere to the recommended layout, which includes having the title as the largest font on the page, and making the abstract clearly visible (e.g. a paragraph of text underneath a heading of "Abstract"). This can also be easily achieved using the Markup section of a scripting display template. An example is shown in Figure 6.



```
Enter your custom Freemarker / HTML markup:
1 <h2>${xml.get("/item/title")}</h2>
2 <h3>Abstract</h3>
3 <p>
4 ${xml.get("/item/description")}
5 </p>
6
```

Figure 6 A sample custom layout

Once your repository adheres to the Google Scholar guidelines it can take up to six months for the changes to be reflected in the search results. If your repository is not automatically picked up by Google Scholar you can manually submit it by going to the following URL: <http://support.google.com/scholar/bin/request.py?hl=en> and filling out the required forms.

References

Adding a site to Google

<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=34397>

Webmaster Guidelines

<http://support.google.com/webmasters/bin/answer.py?hl=en&answer=35769>

Search Engine Optimization Starter Guide

<http://www.google.com/webmasters/docs/search-engine-optimization-starter-guide.pdf>

Google Scholar Inclusion Guidelines

<http://scholar.google.ca/intl/en/scholar/inclusion.html>

Contact Client Support

We are always happy to help.

If your organisation has a support agreement with EQUELLA then help is available at <http://equella.custhelp.com>.