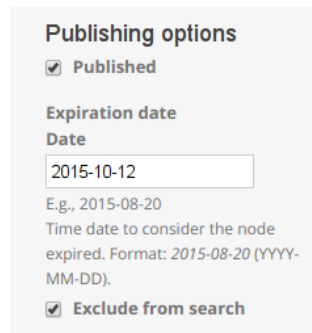


## 9. Restricted Searching

Restricted searching is controlled on PECE through user roles – as explained in the section of this document on “Data Access” – and on an individual artifact-basis. Administrator, Researcher, and Collaborator roles, having different levels of access to content, have also different levels of access to the search functionalities of the system, being only allowed to search and find content that is available to them through the permission system. Administrators and Researchers have the ability to exclude particular nodes from search results, for instance. This option is available when creating or editing a particular type of content as shown below:



**Publishing options**

☒ Published

**Expiration date**  
Date  
2015-10-12  
E.g., 2015-08-20  
Time date to consider the node expired. Format: 2015-08-20 (YYYY-MM-DD).

☒ Exclude from search

Following the permission settings on the platform, content is only visible through the built-in search function to the authenticated users with specific roles:

Artifact Permissions	Authorized Roles	Access through Search API
Open	All	Non-restricted to all
Restricted	Researchers, Administrators	Restricted to other roles
Private	Individual user and artifact-based	Restricted to all

In order to achieve our mission of promoting data exchange and enhance collaboration among ethnographers, we encourage users to release their data as often and open as possible (while being quite observant of the privacy and ethical issues when doing so). For this purpose, all the artifacts with “open” permissions are available to indexing and searching through our Open API and the pluggable extensions for Elasticsearch server back-end via Drupal Search API.

In terms of technical capability, PECE is shipped with scalable search server extensions in addition to the built-in restricted searching functionality of Drupal. In specific, the platform distribution comes with an extension for the Elasticsearch search server back-end. Search servers are key for our web framework because they allow for powerful discovery capabilities in big corpus of texts and across different corpora of texts. It is a known limitation of the native search capability of Drupal to under-perform with a SQL database with more than 50k documents/nodes. Another important benefit of having a search database back-end is the ability to perform searching across different PECE instances for identifying ethnographic content as well as for asking research questions across several ethnographic collections.

We have tested alternatives such as ApacheSolr and ElasticSearch and planned but not yet configured our scalable searching back-end. However, PECE comes ready to interface with ElasticSearch (if you are planning to use your own back-end instead of ours).

Administrators and users have the option of using **three search back-ends**: one is **Drupal's native search mechanism** already explained above; another is a connector from our platform to an **ElasticSearch back-end** (which can be used with in the future with our ElasticSearch cluster or your own infrastructure); and, finally, we will provide a **SPARQL endpoint** to communicate with a dedicated Semantic Web search server infrastructure. We will use the ElasticSearch and the SPARQL back-ends for searching content in the platform following the RDA policy for restricted content as well, but mostly for content that is open for non-restricted distribution under flexible copyright licenses. Both, the ElasticSearch and the SPARQL back-end will be used to index and query across several PECE instances in the near future.

---

**Instructions for Developers:** the aforementioned search server extensions depend on the following modules: **“elasticsearch\_connector”** v. 7.x-1.0-alpha7; **“entity”** 7.x-1.6; **“search\_api”** 7.x-1.14; **“search\_api\_multi”** v. 7.x-1.2. For the restricted search capability, we rely on the core Search API configuration of Drupal core plus **“search\_config”** v. 7.x-1.1. Under “Search API”, configure new server: 'PECE Search' and select under “Elasticsearch Connector Service settings, cluster: 'pecesearch’”. Default configuration for “Elasticsearch connector” on “Clusters” is: [cluster\_name: pecesearch; server\_URL: worldpece.org:9200; make\_this\_cluster\_default\_connection ON; status Active; connection\_timeout: 200]. For the configuration of “search\_config” module, under general system permissions mark: “authenticated user: exclude particular nodes from search results”.