

4Science proposal to improve authority support in DSpace7

This page is outdated / obsolete

 This page represented a temporary discussion area. As agreement has been achieved it has moved forward into this REST Contract PR: <https://github.com/DSpace/Rest7Contract/pull/128>.

Therefore, this page is now OBSOLETE and the REST CONTRACT should be considered the latest version.

The current proposal have two goals:

1. modify the 7.0beta2 implementation to **revert changes in how the information is stored**, avoiding to force the use of the authority value where not necessary/liked. This mean that NO changes in the existing metadata and related configuration (browse, search) will be needed for Institution that comes from a previous version;
2. use a single infrastructure to provide support for all the different controlled values features available in DSpace, i.e.
 - a. dropdown in submission, aka value pairs
 - b. authority lookup and suggestion, i.e. ORCID, LoC, Sherpa
 - c. controlled-vocabulary, i.e. srsc.xml

The following REST contract will be changed
<https://github.com/DSpace/Rest7Contract/blob/master/authorities.md> more specifically the

/api/integration/authorities/<:authority-name>/entries

will be renamed in

/api/integration/authorities/<:authority-name>/choices

and will be changed to return options to present in the UI not directly related with an authority value, i.e.

example from a value pair

```
{
  "_embedded": {
    "authorityChoices": [
      {
        "display": "Animation",
        "value": "Animation",
        "type": "authorityChoice"
      },
      {
        "display": "Article",
        "value": "Article",
        "type": "authorityChoice"
      },
      {
        "display": "Book",
        "value": "Book",
        "type": "authorityChoice"
      },
      ...
    ]
  }
}
```

example from the ORCID Authority

```
{
  "_embedded": {
    "authorityChoices": [
      {
        "display": "Bollini, Andrea",
        "value": "Bollini, Andrea",
        "authority": "will be generated::orcid::0000-0002-9029-1854",
        "otherInformation": {
          "first-name": "Andrea",
          "last-name": "Bollini",
          "orcid": "0000-0002-9029-1854"
        },
        "type": "authorityChoice",
        "_links": {
          "authorityValue": {
            "href": "https://dspace7.4science.cloud/server/api/integration/authorityValues/ORCID:will be generated::orcid::0000-0002-9029-1854"
          }
        }
      },
      ...
    ]
  }
}
```

please note that the above authority value from the Orcidv2AuthorityValue is exactly as it is today when the value is not yet stored in SOLR

example from the srsc controlled vocabulary

```
{
  "_embedded": {
    "authorityChoices": [
      {
        "display": "History of religion",
        "value": "Research Subject Categories::HUMANITIES and RELIGION::Religion/Theology::
History of religion",
        "otherInformation": {
          "note": "Religionshistoria"
        },
        "type": "authorityChoice",
        "_links": {
          "authorityValue": {
            "href": "https://dspace7.4science.cloud/server/api/integration
/authorityValues/srsc:VR110102"
          }
        }
      },
      {
        "display": "Church studies",
        "value": "Research Subject Categories::HUMANITIES and RELIGION::Religion/Theology::
Church studies",
        "otherInformation": {
          "note": "Kyrkovetenskap"
        },
        "type": "authorityChoice",
        "_links": {
          "authorityValue": {
            "href": "https://dspace7.4science.cloud/server/api/integration
/authorityValues/srsc:VR110103"
          }
        }
      },
      ...
      {
        "display": "Religion/Theology",
        "value": "Research Subject Categories::HUMANITIES and RELIGION::Religion/Theology",
        "type": "authorityChoice",
        "_links": {
          "authorityValue": {
            "href": "https://dspace7.4science.cloud/server/api/integration
/authorityValues/srsc:SCB110"
          }
        }
      },
      ...
    ]
  }
}
```

the display value is what we expect that the Angular UI (or any other REST client) will show to the user. The value and eventually the authority are what the client should store in the metadata value, i.e. should use in the subsequent PATCH request.

the relevant changes here are

there is not anymore an ID attribute, as this object is not directly addressable in the URL

the _link section eventually contains a reference to an authorityValue but there is not self link

Please note that the existence of an authority will not imply that an authority id must be stored. If the authority value must be stored or not when available will be indicated by a new attribute in the authority endpoint

/api/integration/authorities/<:authority-name>

```
{
  "id": "srsc",
  "name": "srsc",
  "scrollable": false,
  "hierarchical": true,
  "storeAuthority": false,
  "type": "authority"
}
```

but the REST client can also guess this directly from the authorityChoice response that could eventually include an authorityValue link without an authority attribute.

Moreover, we want to add an extra parameter to the `/api/integration/authorities/<:authority-name>/choices` endpoint

`strict = true`

to indicate that the search must use the `getBestMatch` method of the authority framework and only return the first Choice that matches exactly with the authority if provided or with the value. This parameter will be used by angular to retrieve the Choice related to an existing metadata value (i.e. when a submission is resumed) as the display value must be shown to the user instead of the stored value. In the case that no result is found the angular application will be free to prompt an alert or keep the current value displaying it as is.

A new endpoint

`/api/integration/authorityValues/<authority-name>:<authority-ID>`

will be implemented supporting only direct GET request to a specific ID.

This endpoint is needed to provide a persistent direct access to authority value and support the Controlled-Vocabulary hierarchical visualization via the additional search methods included here <https://github.com/DSpace/Rest7Contract/pull/120> (to be updated according to the new proposed endpoint `/integration/authorityValues`)

Please note that the ID will be a combination of the authority pluginname and the authority value usually stored in the metadata. There is not plan to change what is eventually stored (as the authority ID is not necessarily stored) in the metadata value, the authority pluginname will be retrieved when necessary via a new search method.

<https://dspace7.4science.cloud/server/api/integration/authorities/search/byMetadataAndCollection>

It should be noted that the above search method means that an authority plugin can be now associated with a specific metadata eventually for a specific collection. This is a bit more flexible than what was feasible in DSpace < 7 as the authority plugin was associated at level of metadata and it was not possible to associate different plugin for the same metadata depending on the collection.

The reason of this change can be found in the way that value-pairs are managed, it is possible to define a collection where the `dc.type` comes from a list and another collection that uses a different list or either just free text. This is currently partially broken in DSpace 7 beta 2.

To be backward compatible, the current way to configure an authority over a metadata will automatically imply that the same configuration applies to all the collections.

Finally, the hierarchical controlled vocabulary feature will be supported thanks to the search methods proposed in <https://github.com/DSpace/Rest7Contract/pull/120> (to be updated according to the new endpoints structure). There will be not need to store the authority ID in the metadata value also for controlled vocabulary but we will always auto generate a temporary ID to allow the navigation of the controlled vocabulary tree if the ID is not present in the controlled vocabulary xml definition file.

Estimation: all the above changes are expected to have a limited impact on the Angular side, both for existing features in the master (value-pairs, authority lookup & suggestion) than on the ongoing development (controlled-vocabulary). To make these changes we estimate **4-8 hours** of work **on the Angular side**. **On the REST side** more work is required, we estimate an effort **in the range of 3-5 days**