


# Add License information to UMM - ECSE-171

Author(s): [Erich Reiter](#)

## Description of Problem

NASA management wants uniform license information for all collection level metadata. We need to first find or create a place where license information can be stored within the UMM collection metadata. Then we need to describe how we will map the license information from/to the UMM-C to/from the other supported specifications.

### JIRA Linkage



ECSE-171 - Jira project doesn't exist or you don't have permission to view it.

## Background:

NASA management wants uniform license information for all collection level metadata. Very few collections currently exist in the CMR where it looks like an attempt was made to add license information. Based on its definition, one element does exist in the UMM that can hold license information, but it is not adequate for NASA's needs.

## Analysis:

There have been several suggestions to use certain UMM-C elements for the license information and they are: access constraints, related URL, and use constraints. Looking at the definition of access constraints, it is clearly stated that this element is used to set permissions in order to access the metadata. Since license information doesn't pertain to permissions accessing the metadata this is not a proper fit. The related URL element is a general element that holds URLs. While this element could work for licenses that are located through a URL, it does not work for non URL license information that is to be inserted into the metadata. The UMM-C use constraints' definition states:

Designed to protect privacy and/or intellectual property by allowing the author to specify how the collection may or may not be used after access is granted. This includes any special restrictions, legal prerequisites, terms and conditions, and/or limitations on using the item. Providers may request acknowledgement of the item from users and claim no responsibility for quality and completeness. Note: Use Constraints describe how the item may be used once access has been granted; and is distinct from Access Constraints, which refers to any constraints in accessing the item.

This is the description for license information. Use Constraints is the proper top level element.

Currently, the use constraints element is defined as text that has a maximum length of 20,000 characters. In September of 2016 it was increased from 4,000 characters because a few records had very long use constraints. It looks like they made an attempt at putting license information into this element.

An option for NASA is to have a default license web page. This would allow data providers to just insert a URL into the metadata that points users to the license text on the web. This has several benefits:

1. NASA is able to change the license without having to re-edit and re-ingest most of the collections into the CMR.
2. NASA has more control over quality of the license text.
3. This keeps the size of each collection metadata record small, enabling the CMR to maintain fast performance.

The Project Open Data uses URLs for its licenses in its schema as seen here: <https://project-open-data.cio.gov/v1.1/schema/#license>

To be flexible we should be able to accommodate either a URL to where the license exists on the web, or to include the license text in the metadata itself. While I prefer the former the UMM-C should be able to handle license text. I don't think the UMM-C should allow both a URL and text at the same time. The reason for this is that one could contradict the other or one could be updated while the other isn't, producing conflicting information. I believe the UMM-C should allow a short description of the license or use of the product independent of the license text or URL.

## Recommendations:

Therefore the following structure is proposed:

```

UseConstraints - (0..1) - top-level element - If Use Constraints is to be used either Description or License
URL or License Text must be used.
  Description (0..1) - Description of the license or use constraint of the product.
  <Choice of 0 or 1>
    LicenseUrl - 0..1 - The License URL.
    LicenseText - 0..1 - The license text.
  </Choice>
</UseConstraints>

```

#### Code Block 1: High level UseConstraints

I contemplated using the already defined `OnlineResource` as the type for `LicenseUrl`. Basically giving the `LicenseUrl` 6 sub-elements. The structure then would look like the following:

```

UseConstraints (0..1) - top-level element - If Use Constraints is to be used either Description or License URL
or License Text must be used.
  Description (0..1) - Description of the license or use constraint of the product.
  <Choice of 0 or 1>
    LicenseUrl (0..1) - The URL and accompanying information about the URL.
      Linkage (1) - The URL.
      Protocol (0..1) The URL's protocol (usually http or https).
      ApplicationProtocol (0..1) - Not used for this purpose.
      Name (1) - A name such as "License URL".
      Description (1) - A description of the URL.
      Function (0..1) - Not used for this purpose.
    LicenseText (0..1) - The license text
  </Choice>
</UseConstraints>

```

#### Code Block 2: Complete Breakdown of UseConstraints using `OnlineResource`.

While this solution is more generic and follows the ISO 19115-2 and 1 standards, all of the sub-elements except for `Linkage` (the URL), are not needed. Furthermore, ISO didn't either anticipate or find it necessary to reference constraints such as licenses outside of the metadata through a URL since it is not included in the schema. To keep the `LicenseUrl` simple and easy to use we recommend using a URL directly as a value of `LicenseUrl`.

The Final UMM-C JSON specification for Use Constraint:

```

    "UseConstraintsDescriptionType": {
      "type": "object",
      "additionalProperties": false,
      "description": "This sub-element either contains a license summary or free-text description of the
constraint. In DIF, this field is called Access_Constraint. In ECHO, this field is called RestrictionComment.
Examples of text in this field are Public, In-house, Limited. Additional detailed instructions on how to access
the collection data may be entered in this field.",
      "properties": {
        "Description": {
          "description": "This sub-element either contains a license summary or free-text description
of the constraint. In DIF, this field is called Access_Constraint. In ECHO, this field is called
RestrictionComment. Examples of text in this field are Public, In-house, Limited. Additional detailed
instructions on how to access the collection data may be entered in this field.",
          "type": "string",
          "minLength": 1,
          "maxLength": 4000
        }
      }
    },
    "UseConstraintsType": {
      "type": "object",
      "additionalProperties": false,
      "description": "This element defines how the data may or may not be used after access is granted to
assure the protection of privacy or intellectual property. This includes license text, license URL, or any
special restrictions, legal prerequisites, terms and conditions, and/or limitations on using the data set. Data
providers may request acknowledgement of the data from users and claim no responsibility for quality and
completeness of data.",
      "oneOf": [{
        "properties": {
          "Description": {
            "$ref": "#/definitions/UseConstraintsDescriptionType"
          },
          "LicenseUrl": {
            "description": "This element holds the URL and associated information to access the
License on the web. If this element is used the LicenseText element cannot be used.",
            "type": "string",
            "minLength": 1,
            "maxLength": 1024
          }
        },
        "oneOf": [{
          "required": ["Description"]
        }, {
          "required": ["LicenseUrl"]
        }]
      }],
      "properties": {
        "Description": {
          "$ref": "#/definitions/UseConstraintsDescriptionType"
        },
        "LicenseText": {
          "description": "This element holds the actual license text. If this element is used the
LicenseUrl element cannot be used.",
          "type": "string",
          "minLength": 1,
          "maxLength": 20000
        }
      },
      "oneOf": [{
        "required": ["Description"]
      }, {
        "required": ["LicenseText"]
      }]
    }
  }
}

```

Code Block 3: Final UMM-C UseConstraint JSON Representation

The sections below detail the proposed changes based on the analysis done by the ECSE team:

## Changes to UMM-C elements

As shown in the Recommendations section, the Use Constraints element will be changed from a free text field to 3 sub-elements of which only 2 can be used at once: Description, LicenseUrl, or LicenseText. The UseConstraints Description element will be defined as its own definition so that it isn't defined twice. The resultant changes will look like what is described in Code Block 3.

## Changes to UMM-Common elements

No changes are proposed for UMM-Common elements since the license information is only needed for collections. The definitions can be moved at a later date to UMM-Common if other UMM profiles need the license information.

## Mappings to DIF, ECHO, ISO

DIF 9 and DIF 10:

Both DIF specifications include a Use\_Constraint element. It is currently defined as a free text element and is mapped to the UMM-C UseConstraint element. In order to add license information in the same manner that is proposed in this text, it is proposed that the DIF 9 and DIF 10 specifications adopt the same approach. If either specification can't be altered the Use\_Constraint element will be mapped to the UseConstraints/Description element.

ECHO 10:

Currently the ECHO 10 schema does not include an element to hold license information. It is proposed that the ECHO 10 schema adds the same elements as what will be contained in the UMM.

ISO 19115-2 (MENDS/SMAP):

The ISO schema does have an element that can include license information, it is called resourceConstraints. The schemas between SMAP and MENDS differ slightly in the case of resourceConstraints where the SMAP implementation includes /gmd:DS\_Series/gmd:seriesMetadata/ before the MENDS gmd:MI\_Metadata tag. Following is an example of each:

MENDS: /gmi:MI\_Metadata/gmd:identificationInfo/gmd:MD\_DataIdentification/gmd:resourceConstraints

SMAP: /gmd:DS\_Series/gmd:seriesMetadata/gmi:MI\_Metadata/gmd:identificationInfo/gmd:MD\_DataIdentification/gmd:resourceConstraints

the xpath after gmd:resourceConstraints is the same both both implementations.

By substituting the MD\_LegalConstraints element for the MD\_Constraints element, there are four sub-elements for resourceConstraints: useLimitation, accessConstraints, useConstraints, and otherConstraints.

Looking at the ISO 19115-1 documentation:

1. useLimitation describes limitations affecting the fitness for use of a resource or its metadata. The example provided for this element is "Not to be used for Navigation".
2. accessConstraints describes the constraints of obtaining the resource or metadata.
3. useConstraints describes the constraints or limitations on using the resource.
4. otherConstraints describes other constraints used in obtaining or using the resource.

(Note: The 19115-2 schema doesn't include any documentation of these elements, but I have the 19115-1 documentation and both schemas use the same elements.)

Using the above definitions, it is our recommendation to use the following for the UMM UseConstraints leaf elements:

1. For all three leaf elements use the following path
  - a. use gmd:resourceConstraints/gmd:MD\_LegalConstraints/gmd:useConstraints/gco:CharacterString
2. Prefix the following text before the value of the elements.
  - a. For Description use a string prefix of: "Description: "
  - b. For LicenseUrl use a string prefix of: "LicenseUrl: "
  - c. For LicenseText use a string prefix of: "LicenseText: "

For example, if a UMM-C UseConstraints elements looks like the following:

```
UseConstraints: {
  "Description": "This collection is protected by the NASA Public Domain Collection license."
  "LicenseText": "By exercising the Licensed Rights (defined below), You accept and agree to be bound by the
terms and conditions of this Creative Commons Attribution 4.0 International Public License ("Public License").
To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in
consideration of Your acceptance of these terms and conditions,..."
}
```

Then the ISO translation would look like:

```

<gmi:MI_Metadata>
  <gmd:identificationInfo>
    <gmd:MD_DataIdentification>
      <gmd:resourceConstraints>
        <gmd:MD_LegalConstraints>
          <gmd:useConstraints>
            <gco:CharacterString>Description: This collection is protected by the NASA Public Domain
Collection license.</gco:CharacterString>
          </gmd:useConstraints>
          <gmd:useConstraints>
            <gco:CharacterString>LicenseText: By exercising the Licensed Rights (defined below), You
accept and agree to be bound by the terms and conditions of this Creative Commons Attribution 4.0 International
Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are
granted the Licensed Rights in consideration of Your acceptance of these terms and conditions,...</gco:
CharacterString>
          </gmd:useConstraints>
        </gmd:MD_LegalConstraints>
      <gmd:resourceConstraints>
        </gmd:MD_DataIdentification>
      </gmd:identificationInfo>
    </gmi:MI_Metadata>

```

Currently the translations for UMM-C AccessConstraints uses the ISO useLimitation element which we think isn't the best choice. We recommend changing the translations for access constraints to use the ISO accessConstraints element.

## Changes to CMR

The translation of the legacy formats into UMM will take into account the translation process described above.

## What should the MMT forms look like

The structure of the forms shouldn't change, but the entry form for Use Constraints should require entry of a description and either the license URL or license text.

## How should the values in these fields be presented to the user on the EDSC

Use constraints along with license information should be available to the user when they view a collection.



Unknown macro: 'pageapproval'

## Next Steps after approval

1. Write tickets and ask for changes in the supporting specifications (DIF 9, DIF 10, ECHO 10).
2. Update JAMA documentation and mapping spreadsheet.
3. Write issues to update the UMM schema.
4. Write issues to update the CMR ingest and its translations.