

RACO. Format per a la importació de metadades i dades

02-01-13

Característiques de l'arxiu

L'arxiu per a la importació ha de ser de contingut ASCII codificat amb el joc de caràcters UTF-8.

El nom de l'arxiu xml no pot contenir diacrítics i d'altres caràcters (ñ, ç...).

La llargada total de l'arxiu és limitada a 175 articles.

Estructura de l'arxiu

La descripció de cada sumari d'una revista correspon a un conjunt d'articles tal com es defineix més avall:

1. L'arxiu ha de començar amb `<issues>` i acabar amb `</issues>`.
2. Cada sumari, ha de començar amb `<issue>` i acabar amb `</issue>`, i es pot compondre de diversos articles, agrupats per seccions.
3. Cal agrupar els articles segons la secció a la que pertanyen, `<section>` i `</section>`. Si la vostra revista no té seccions utilitzeu la general, Articles, com a l'exemple:

```
<section>
  <title>Articles</title>
  <abbrev>ART</abbrev>
  <article>...</article>
</section>
```

4. Cada article ha de començar amb `<article>` i acabar amb `</article>`.
5. Ordeneu les seccions i els articles dins del número segons hagin d'aparèixer a RACO.

El format de l'arxiu serà el següent:

```
<issues>
  <issue>
    <section>
      <title>Nom secció1</title>
      <article>...</article>
      <article>...</article>
      ....
    </section>
    <section>
      <title>Nom secció2</title>
      <article>...</article>
    </section>
    ....
  </issue>
  <issue>
    <section>
      <title>Articles</title>
      <article>...</article>
      <article>...</article>
      ....
    </section>
  </issue>
</issues>
```

Una de les principals novetats en aquesta versió d'ojs que utilitza RACO és que molts camps poden ser introduïts en diverses llengües. Pels camps multilingües cal afegir la propietat per indicar la llengua a l'etiqueta corresponent utilitzant l'atribut `locale`. Els possibles idiomes són català (`ca_ES`), castellà (`es_ES`) i anglès (`en_US`). Per exemple:

`<title locale="ca_ES">Títol en català</title>`

`<title locale="es_ES">Título en español</title>`

`<title locale="en_US">English title</title>`

Definició dels camps del número (<issue></issue>)

Les marques d'inici i fi de les dades d'un número són <issue> i </issue>. Aquesta marca d'inici ha de tenir tres atributs:

- **identification:** indica quin camp del número es fa servir per identificar-lo. El seu valor pot ser title, num_vol_year, vol, vol_year o year (per defecte title). Es recomana utilitzar "title" amb el següent format:
 - Si només té any: AAAA
 - Si només té número: Núm.: NN
 - Si només té volum: Vol.: VV
 - Si té volum i número: Vol.: VV Núm.: NN
- **published:** indica si s'ha publicat o no. El seu valor pot ser true o false (per defecte false).
- **current:** indica si és el darrer número. El seu valor pot ser true o false (per defecte false).

Per exemple:

```
<issue identification="title" published="true" current="false">
```

Els camps que es poden incloure dins un número són:

Descripció	Codificació	Obligatori?	Repetible?	Format
Títol	<title></title>	Si	Si*	Text. És un camp multilingüe.
Descripció	<description></description>	No	Si*	Text. És un camp multilingüe.
Volum	<volume></volume>	No	No	Número
Número	<number></number>	No	No	Número
Any	<year></year>	Si	No	Número
Portada	<cover></cover>	Si	No	Veure el quadre "Camps portada"
Data de publicació	<date_published></date_published>	No	No	Data
Accés al text complet	<access_date></access_date> O bé <open_acces />	Si	No	Veure el quadre "Camps accés text complet"
Secció	<section></section>	Si	Si	Veure el quadre "Camps secció"

* Es pot repetir el camp sempre que s'indiqui l'idioma com a atribut. Per exemple:

```
<title locale="ca_ES">Títol en català del sumari</title>
<title locale="es_ES">Títol en castellà del sumari</title>
<title locale="en_US">Títol en anglès del sumari</title>
```

Camps portada (<cover></cover>)

Els camps ha de tenir la portada són:

Descripció	Codificació	Obligatori?	Repetible?	Format
Etiqueta	<caption></caption>	Si	No	Espai en blanc
Imatge	<image></image>	Si	No	href o embeded

La imatge es pot incloure amb dos formats:

- href: fent un lligam a la url on està disponible la imatge

```
<image>
<href src="http://www.raco.cat/imatges/portada.jpg" mime_type="image/jpeg" />
</image>
```

- embeded: incrustant la imatge en l'arxiu xml

```
<image>
<embed filename="portada.jpg" encoding="base64" mime_type="image/jpeg"></embed>
</image>
```

Camps accés text complet

Cal indicar el tipus d'accés que ofereix la revista al text complet dels articles inclosos en aquest número:

- Accés obert. Utilitzeu l'etiqueta <open_access/>
- Text embargat segons les "Condicions d'accés obert" de la revista.. Utilitzeu l'etiqueta <access_date> indicant la data (en format AAAA-MM-DD) a partir de la que serà accessible el text complet . Exemple :

```
<access_date>2007-11-01</access_date>
```

Camps secció (<section></section>)

Els camps que ha de tenir una secció són:

Descripció	Codificació	Obligatori?	Repetible?	Format
Títol	<title></title>	Si	Si*	Títol de la secció ja existent a la revista. És un camp multilingüe.
Abreviatura	<abbrev></abbrev>	Si	Si*	Abreujament de la secció. És un camp multilingüe.
Policy	<policy></policy>	No	Si*	Política de la secció. És un camp multilingüe.
Articles	<article></article>	Si	Si	Veure quadre "Camps article"

* Es pot repetir el camp sempre que s'indiqui l'idioma com a atribut. Per exemple:

```
<title locale="ca_ES">Títol en català de la secció</title>
<title locale="es_ES">Títol en castellà de la secció</title>
<title locale="en_US">Títol en anglès de la secció</title>
```

Exemple:

```
<section>
  <title>Articles</title>
  <abbrev>ART</abbrev>
  <article>
    ...
  </article>
  <article>
    ...
  </article>
</section>
```

Si la secció no existeix a la revista es crearà una nova amb la informació introduïda. Podeu definir seccions per a la revista utilitzant l'opció del menú "Seccions de la revista" de l'usuari amb rol administrador.

Exemple número:

```
<issues>
<issue identification="title" published="true" current="false">
  <title locale="ca_ES">Núm.: 1 </title>
  <volume>0</volume>
  <number>1</number>
```

```

<year>1997</year>
<cover>
  <caption> </caption>
  <image>
    <href src="http://cbuc.es/img/portada.jpg" mime_type="image/jpeg" />
  </image>
</cover>
<date_published>1997-01-01</date_published>
<open_access />
<section>
  <title locale="ca_ES">Articles</title>
  <abbrev locale="ca_ES">ART</abbrev>
  <article>
    ...
  </article>
  <article>
    ...
  </article>
  ...
</section>
</issue>
<issue identification="title" published="true" current="false">
  <title locale="ca_ES">Núm.: 2</title>
  <volume>0</volume>
  <number>2</number>
  <year>1998</year>
  <cover>
    <caption> </caption>
    <image>
      <href src="http://rev.cat/imatges/port.jpg" mime_type="image/jpeg" />
    </image>
  </cover>
  <date_published>1998-01-01</date_published>
  <open_access />
  <section>
    <title locale="ca_ES">Articles</title>
    <abbrev locale="ca_ES">ART</abbrev>
    <article>
      ...
    </article>
    <article>
      ...
    </article>
    ...
  </section>

```

```
</issue>
....
</issues>
```

Si algun camp obligatori per a un número no existeix, poseu les etiquetes de codificació i com a valor un espai. Per exemple el camp “caption” de la portada del número:

```
<caption> </caption>
```

Definició del camps dels articles (<article></article>)

Les etiquetes per identificar els articles que s’inclouen a les seccions són <article></article>. Si es vol identificar l’idioma en que està escrit, utilitzeu l’atribut language i com a valor cat per català, spa per castellà, eng per anglès, und si no s’especifica (per a codis addicionals consulteu <http://loc.gov/standards/iso639-2/langcode.html>). Per exemple:

```
<article language=”cat”>
```

Els camps que pot incloure un article són:

Descripció	Codificació	Obligatori?	Repetible?	Format
Títol	<title></title>	Si	Si *	Text
Data de publicació	<date_published> </date_published>	No **	No	AAAA-MM-DD
Autor	<author></author>	Si	Si	Veure el quadre “Camps autor”
Pàgines	<pages></pages>	Si	No	Text. Si hi ha més d’una separeu-les amb un guió. No cal que doneu l’abreviatura de “p.” ja que el sistema la genera automàticament.
Text complet	<galley></galley>	Si	Si*	Veure el quadre “Camps text complet”
Resum	<abstract></abstract>	No	Si *	Text
Accés al text complet	<open_access />	No	No	Només cal posar la etiqueta si volem indicar que és accés obert i en la definició del número hem

				indicat embargament..
Indexació	Indexing	No	No	Veure el quadre "Camps indexació"
Arxius addicionals	Supplemental_file	No	No	

* Es pot repetir el camp sempre que s'indiqui l'idioma com a atribut. Per exemple:

```
<title locale="ca_ES">Títol en català de l'article</title>
<title locale="es_ES">Títol en castellà de l'article</title>
<title locale="en_US">Títol en anglès de l'article</title>
```

** Aquest camp no és obligatori, excepte si l'article s'inclou a un número publicat (amb published="true"). Aleshores cal indicar en quina data s'ha publicat.

Camps author (<author></author>)

Aquesta etiqueta té un atribut obligatori (primary_contact) que ens indica si l'autor és el contacte principal. Els valors que pot prendre són true o false. Si hi ha més d'un autor, un d'ells ha de ser el contacte principal.

Exemple:

```
<article>
<author primary_contact="true"> .... </author>
....
</article>
<article>
<author primary_contact="false"> .... </author>
<author primary_contact="true"> .... </author>
<author primary_contact="false"> .... </author>
....
</article>
```

Els camps que ha de tenir un autor són:

Descripció	Codificació	Obligatori?	Repetible?	Format
Nom	<firstname></firstname>	Si	No	Text
Cognoms	<lastname></lastname>	Si	No	Text
Correu	<email></email>	Si	No	Text. Si no el coneixeu deixeu un espai en blanc

Si no coneixeu qualsevol de les dades obligatòries de l'autor afegiu les etiquetes i com a valor un espai en blanc.

Si l'autor és una institució, afegiu l'etiqueta nom amb valor un espai en blanc i a l'etiqueta cognom escriviu el nom de la institució.

Exemple:

```
<author primary_contact="true">
    <firstname> </firstname>
    <lastname> </lastname>
    <email> </email>
</author>

<author primary_contact="false">
    <firstname> </firstname>
    <lastname>Generalitat de Catalunya</lastname>
    <email> </email>
</author>
```

Camps text complet (<galley></galley> o <htmlgalley></htmlgalley>)

Cal indicar l'idioma del text complet que

- PDF

Els camps que ha de tenir aquesta etiqueta per incloure un document pdf on es pot trobar el text complet de l'article són:

Descripció	Codificació	Obligatori?	Repetible?	Format
Etiqueta	<label></label>	Si	No	"Text complet"
Arxiu	<file></file>	Si	No	Href, embed

El document a text complet es pot incloure amb dos formats:

- href: fent un lligam a la url on està disponible el document en format PDF o HTML

```
<galley locale="ca_ES">
    <file>
        <href src="http://www.raco.cat/art/article.pdf" mime_type="application/pdf" />
    </file>
</galley>
```

- embed: incrustant el document pdf en l'arxiu xml, codificant-lo en base-64.

```
<galley locale=ca_ES>
  <file>
    <embed encoding="base64" filename="article.pdf" mime-type="application/pdf">
      PGJyPjxicj5IbyBzZW50aW0sIGFxdWVzdCBubyBkaXNwb3NlbSBkZWwgY29udGluZ3V0IGQnYXF1ZXN0IGFydGljbGUuPGJyPkxvIHNIbnRpbW9zLCBubyBkaXN
    </embed>
  </file>
</galley>
```

- HTML

Els camps que ha de tenir aquesta etiqueta per incloure un enllaç URL on es pot trobar el text complet de l'article són:

Descripció	Codificació	Obligatori?	Repetible?	Format
Etiqueta	<label></label>	Si	No	"Text complet"
Arxiu	<file></file>	Si	No	Href, embed

El text complet es pot incloure amb dos formats:

- href: fent un lligam a la url on està disponible el document en format PDF o HTML

```
<htmlgalley locale="ca_ES">
  <file>
    <href src="http://www.raco.cat/art/article.pdf" mime_type="application/pdf" />
  </file>
</htmlgalley>
```

- embed: incrustant el document redirect.html, codificant-lo en base-64.

```
<htmlgalley locale="ca_ES">
  <file>
    <embed encoding="base64" filename="redirect.html" mime-type="text/html">
      wb25lbW9zIGRlbCBjb250ZW5pZG8gZGUgZXN0ZSBhcnQmaWFjdXRlO2N1bG8uPGJyPlNvcnJ5LCB3ZSBkbyBub3QgaGF2ZSB0aGUgY29udGVudCBvZiB0aGlzI
    </embed>
  </file>
```

</htmlgalley>

On l'arxiu ha de ser el següent:
redirect.html

Per accedir al text complet de l'article, si us plau seguiu el següent enllaç:

Para acceder al texto completo del artículo, por favor siga el siguiente enlace:

To access the full text article, please click the link below:

 http://www.raco.cat/art/art.html

- Sense text complet

Si l'article no ofereix el text complet llavors heu d'utilitzar la etiqueta <htmlgalley></htmlgalley>.

El camps que cal posar-hi són:

Descripció	Codificació	Obligatori?	Repetible?	Format
Etiqueta	<label></label>	Si	No	""
Arxiu	<file></file>	Si	No	Href, embed

On l'arxiu ha de ser el següent:
sense_text_complet.html

<html>

<head></head>

<body>

Ho sentim, no disposem del contingut d'aquest article.

Lo sentimos, no disponemos del contenido de este artículo.

Sorry, we do not have the content of this article.

</body>

</html>

Aquest arxiu és per indicar que l'article no s'ofereix a text complet.

Per incloure l'arxiu utilitzeu qualsevol dels dos mètodes explicats anteriorment: enllaç (href) o incrustació (embed).

Si els enllaços al text complet estan disponibles en una pàgina segura (https://), poseu-vos en contacte amb l'oficina del CBUC.

Indexació

Les etiquetes que defineixen la indexació són `<indexing></indexing>`. Els camps que pot tenir són:

Descripció	Codificació	Obligatori?	Repetible?	Format
Disciplina acadèmica i subdisciplines	<code><discipline></discipline></code>	No	Si*	Text Les diferents opcions se separen amb comes. És un camp multilingüe.
Classificació per matèria	<code><subject_class></subject_class></code>	No	Si*	Text Les diferents opcions se separen amb comes. És un camp multilingüe.
Paraules Clau	<code><subject></subject></code>	No	Si*	Text Les diferents opcions se separen amb comes. És un camp multilingüe.
Abast	<code><coverage></coverage></code>	No	No	Veure quadre "Camps abast".

* Es pot repetir el camp sempre que s'indiqui l'idioma com a atribut. Per exemple:

```
<subject locale="ca_ES">ciència</subject>
<subject locale="es_ES">ciencia</subject>
<subject locale="en_US">science</subject>
```

Camps abast

Els camps que es poden incloure a l'abast són:

Descripció	Codificació	Obligatori?	Repetible?	Format
Termes geogràfics o geoespacials	<code><geographical></geographical></code>	No	Si*	Text. És un camp multilingüe.
Termes històrics o cronològics	<code><chronological></chronological></code>	No	Si*	Text. És un camp multilingüe.
Característiques de la mostra	<code><sample></sample></code>	No	Si*	Text. És un camp multilingüe.

* Es pot repetir el camp sempre que s'indiqui l'idioma com a atribut. Per exemple:

<geografical locale="ca_ES">Títol en català del sumari</geografical >
 <geografical locale="es_ES">Títol en castellà del sumari</geografical >
 <geografical locale="en_US">Títol en anglès del sumari</geografical >

Exemple

```
<article language="cat">
  <title locale="ca_ES">OMG, i aixó què és?</title>
  <date_published>1997-01-011</date_published>
  <author primary_contact = "true">
    <firstname>Rafael</firstname>
    <lastname>Tauler</lastname>
    <email> </email>
  </author>
  <pages>7-13</pages>
  <htmlgalley locale="ca_ES">
    <label></label>
    <file>
      <embed
        encoding="base64"
        filename="redirect.html"
        mime_type="text/html">PGJyPjxicj5IbyBzZW50aW0sIGFxdWVzdCBubyBkaXNwb3NlbS
        BkZWwgY29udGluZ3V0IGQnYXF1ZXN0IGFydGljbGUuPGJyPjxicj5IbnRpbW9zLCB
        ubyBkaXNwb25lbW9zIGRlbCBjb250ZW5pZG8gZGUgZXN0ZSBhcnQmaWFjdXRlO2N1
        bG8uPGJyPlNvcnJ5LCB3ZSBkbyBub3QgaGF2ZSB0aGUgY29udGVudCBvZiB0aGlzIGFy
        dGljbGUuPGJyPjxicj4=</embed>
      </file>
    </htmlgalley>
    <open_access />
  </article>

<article language="cat">
  <title locale="ca_ES">La radiació ionitzant en la indústria alimentària</title>
  <date_published>2000-01-011</date_published>
  <author primary_contact = "true">
    <firstname>Josep</firstname>
    <lastname>Calderón</lastname>
    <email> </email>
  </author>
  <pages>8-11</pages>
  <galley locale="ca_ES">
    <label>Text complet</label>
    <file>
      <href
        src="http://sumaris.cbuc.es/bbdd/ISSN11377976%2FA2000N000005V000000/ISSN11377
        976A2000N000005V000000article0.pdf" mime_type="application/pdf" />
      </file>
```

```
</galley>  
<open_access />  
</article>
```

Si necessiteu consultar el DTD poseu-vos en contacte amb l'oficina del CBUC.