



## Content and Data Architecture

Elsevier  
Radarweg 29  
1043 NX Amsterdam  
Netherlands

Phone: +31 20 485 3433  
Fax: +31 20 485 3266  
Email: [t.bos@elsevier.com](mailto:t.bos@elsevier.com)

## Release note for Journal Article DTD 5.4.0 and Serial Issue DTD 5.4.0

TBo14006 — version 1.0, 3 October 2014, by Ton Bos and Jos Migchielsen

version 0.1	8 July 2014	First draft based on JA DTD v5.4.0r1, CEP v1.4.0r1, T-b-T 1.9.5
version 0.2	14 July 2014	Comments by JMi
version 1.0	3 October 2014	Final
version 1.1	18 December 2014	Final, including new SI DTD 5.4.0

### Contents

1. Introduction to the Journal Article DTD 5.4.0, Book DTD 5.4.0 and Serial Issue DTD 5.4.0 .....	1
2. Changes in Journal Article DTD 5.4.0 with respect to JA DTD 5.2.0 .....	2
2.1. New publication item types (PITs) .....	2
2.2. Article numbers: item-info/ce:article-number .....	3
2.3. Article history: change to ce:miscellaneous in top-level elements .....	3
3. Changes in Serial Issue DTD 5.4.0 with respect to SI DTD 5.2.0 .....	4
3.1. Cardinality of ce:pages in element issue-data .....	4
4. Changes in Common Element Pool 1.4.0 with respect to version 1.2.0 .....	4
4.1. Accessibility initiative: introduction of ce:alt-text .....	4
4.2. Accessibility initiative: enhanced table header functionality (rowheads) .....	6
4.3. Handling of content objects that reside outside the Electronic Warehouse (change in ce:link) .....	7
4.4. Direct linking of keywords to external databases (change in ce:text) .....	7
4.5. Article numbers: introduction of ce:article-number, sb:article-number .....	8
4.6. Unicode version 6: changes in use of ce:glyph .....	9
4.7. Better support of APA reference style: introduction of sb:ellipsis .....	9
4.8. Structured correspondence addresses: extended use of ce:correspondence .....	10
4.9. Nesting of ce:bibliography-sec and ce:further-reading-sec .....	11
5. References .....	11
6. Support .....	12

### 1. Introduction to the Journal Article DTD 5.4.0, Book DTD 5.4.0 and Serial Issue DTD 5.4.0

To support the Reed Elsevier Accessibility initiative [1] it has become necessary to update both the journal and book DTDs. At the same time a few missing features were added to the DTDS and their usability was extended.

Moreover, the journal and book DTDs were aligned in that from now on they will be using the same version of the common element pool (CEP). To emphasize this, both DTDs were assigned the same version number, 5.4.0; and the new CEP has version number 1.4.0. Note that version numbers 5.3.0 for the journal DTD and 1.3.0 for the CEP were skipped.

The Serial Issue DTD underwent a minor change to allow for journal articles without page numbers.

Implementation of the new DTDs is scheduled for the first quarter of 2015. The release comprises the following:

- the Journal Article DTD 5.4.0 (art540.dtd),
- the Serial Issue DTD 5.4.0 (si540.dtd),
- the Elsevier Book DTD 5.4.0 (book540.dtd),
- the Common Element Pool 1.4.0 (common140.ent),
- a complete set containing all relevant DTD files (dtd-2014.zip),
- two separate release notes, one for the Journal Article DTD (this document), one for the Book DTD,
- an updated Tag by Tag, version 1.9.5.

Additionally, separate implementation notes and updates of other documentation can be expected. Sample files will be made available and distributed separately.

A new feature of this release note is that an attempt was made to define the impact of the change on systems in general and on various steps of the end-to-end workflow, manuscript→XML→rendering to HTML/print PDF/web PDF.

## **2. Changes in Journal Article DTD 5.4.0 with respect to JA DTD 5.2.0**

The following sections list the changes with respect to JA DTD 5.2.0.

### *2.1. New publication item types (PITs)*

In JA DTD 5.4.0 four new publication items are introduced, `dat`, `lst`, `mic` and `osp`.

PIT `dat` (for “data”) is intended for use in identifying articles that describe sets of data, sometimes focusing on the quality of said data. Such “data articles” strongly resemble regular full length articles with PIT=`f1a`, but differ in that there are no new scientific insights presented in them. Since assigning this PIT correctly to an article requires knowledge of the content of the article, use of this PIT requires collaboration with Publishing. Rendering and entitlements are as for PIT=`f1a`.

PIT `lst` (for “list”) was introduced primarily to meet needs in book publishing, but is also useful for journals, and was added to the JA DTD to keep both DTDs aligned. It is to be used for capturing all kinds of list that exist on their own, not as part of an article; an example would be a list of all videos that are part of the articles in a special issue.

PIT `mic` (for “micro-article”) is to be used for the identification of short articles dealing with a specific research object that can be directly accessed and interactively explored via this publication, and is therefore distinct from regular short articles that are assigned PIT=`sco`. This research object can be anything from a scientific dataset or computational model to a method, piece of code, algorithm, lab notebook, video, table or figure. As for PIT=`dat`, this PIT is not intended for general use; it may only be assigned to items of specific journals, to be determined in collaboration with Publishing. Rendering and entitlements are as for PIT=`f1a`.

PIT `osp` (for “original software publication”) is introduced to allow capturing articles that comprise computer code or software and a description thereof. As for PIT=`dat` or `mic`, this PIT is not for general use; it may only be assigned to items of specific journals, to be determined in collaboration with Publishing. Rendering and entitlements are as for PIT=`fla`.

DTD change requests CR85 (`lst`), CR88 (`mic`, `dat`), CR94/COS-22 (`osp`).  
Impact: systems (Evisé, PTS, EW), manuscript→XML, XML rendering.

## 2.2. *Article numbers: item-info/ce:article-number*

An optional element `ce:article-number` was added to the model of DTD element `item-info`. Element `ce:article-number` is new to the CEP 1.4.0, and is intended to capture a generic article number in case this feature is introduced in a journal. Note that this “generic” article number is in addition to the ones that are already being distinguished, and that should be handled as before: PII (captured using `ce:pii`), DOI (`ce:doi`), or journal acronym combined with PTS article identifier (`jid`, `aidd`).

An additional, generic article number is useful for allowing the citation of articles that are published without traditional volume/issue/page information. See Section 3.5 for use of generic article numbers in citations; a comparable new element, `sb:article-number`, was added to the structured bibliographic references namespace.

DTD change request CR89.  
Impact: manuscript→XML, XML rendering.

## 2.3. *Article history: change to ce:miscellaneous in top-level elements*

The occurrence indicator of element `ce:miscellaneous` was changed from ? (zero or one occurrences) to \* (zero, one or more occurrences) in top-level elements `head`, `simple-head` and `book-review-head`. Element `ce:miscellaneous` is used to capture additional article history information in the head of an article, next to the regular dates of receipt, review and acceptance. An example is the name of a communicating editor but also other data or dates in the history of the article can go into `ce:miscellaneous`. Allowing multiple occurrences of `ce:miscellaneous` offers flexibility to publishing and editorial in capturing information in the article heads in a journal-specific way.

*XML examples:*

```
<ce:miscellaneous id="misc0010">Published: July 3, 2014</ce:miscellaneous>
<ce:miscellaneous id="misc0015">Communicated by Jan Jansen</ce:miscellaneous>
<ce:miscellaneous id="misc0020">Artwork provided by Piet Pietersen</ce:miscellaneous>
```

DTD change request CR28.  
Impact: manuscript→XML, XML rendering.

### 3. Changes in Serial Issue DTD 5.4.0 with respect to SI DTD 5.2.0

#### 3.1. Cardinality of `ce:pages` in element `issue-data`

As of JA DTD 5.4 it is possible for journal article to be published with an article number replacing the traditional page numbers (see section 4.5). This has necessitated a minor change in the SI DTD, in which the page range of an issue is now optional; this was done by making element `ce:pages` as child of `issue-data` become optional. (Note that page numbers for items themselves were already optional in earlier versions of the SI DTD.)

### 4. Changes in Common Element Pool 1.4.0 with respect to version 1.2.0

The new JA DTD 5.4.0 uses the latest version of the common element pool, CEP 1.4.0. Four new elements were introduced in CEP 1.4.0, `ce:alt-text`, `ce:article-number`, `sb:article-number` and `sb:ellipsis`. Moreover, the models of various elements that are already in existence were updated. Changes in CEP 1.4.0 with respect to its predecessor CEP 1.2.0 are described in the following sections.

#### 4.1. Accessibility initiative: introduction of `ce:alt-text`

In order to comply with the Reed Elsevier Accessibility initiative [1] a new element `ce:alt-text` was introduced. This element makes it possible to include descriptive text alternatives for non-text content such as figures or videos, for example a description of a figure which can be used with screen reading software to make the content accessible for the visually impaired. The most appropriate way of rendering alt-text is left to the various web platforms, taking into account the standards described in [1].

`ce:alt-text` was added as an optional element to the models of `ce:figure`, `ce:table`, `ce:textbox`, `ce:inline-figure` and `ce:e-component`. Element `ce:alt-text` has a role attribute with three allowed values:

- `short` for a description of less than 30 words,
- `long`, for a more verbose text alternatives,
- `summary`, which is to be used for table summaries only.

There must be one `ce:alt-text` with role value `short` and only one `ce:alt-text` for every role value. The attribute role must be populated.

In case text alternatives for any of these five elements are available in source content, it is expected that these alternatives are tagged as `ce:alt-text`; how these text alternatives are actually obtained is outside the scope of this document.

*XML and rendering sample for `ce:figure`:*

```
<ce:figure id="f055">
...
<ce:alt-text id="at070" role="short">Painting by John William Waterhouse, 'The lady of Shalott', 1888.</ce:alt-text>
```

...  
</ce: figure>



ce: alt-text role="short" translated to HTML image alt attribute.

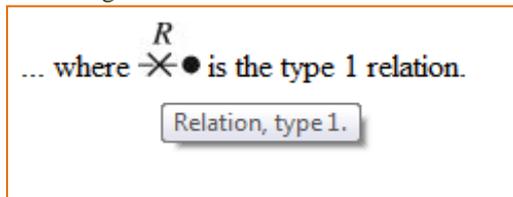


ce: alt-text translated to link.

*XML and rendering sample for ce:inline-figure:*

```
<ce:inline-figure baseline="0.0">  
  <ce:link locator="fx1"/>  
  <ce:alt-text id="at346"  
    role="short">Relation,  
    type 1.</ce:alt-text>  
</ce:inline-figure>
```

Rendering in HTML:



DTD change request CR87.

Impact: systems, manuscript→XML, XML rendering.

#### 4.2. Accessibility initiative: enhanced table header functionality (rowheads)

One of the Web Accessibility Guidelines Working Group [1] recommendations to improve accessibility of scientific articles is to clearly indicate the row and column headings of tables in HTML. Tables are a very visually oriented way of organizing data; simply looking at a table in print or on screen generally suffices to understand the table's structure, and it is easy to get the meaning of the content of any particular table cell. But as anyone who is familiar with the way tabular content is represented in the text-based context of native HTML or XML will know, making sense of a table without visual clues is far more difficult.

For a visually impaired person to make use of tabular material, it is essential that the contents of each table cell can be related to the content of the column head, the row head, or usually both. This can then be used in conjunction with screen reading software to get a full description of the content of any table cell. In the basic Elsevier CALS tables specification, column head cells are already identified as such since their entry elements are children of element `thead`. CEP 1.4.0 adds a way of identifying row head cells as well by specifying the use of a `role` attribute `rowhead` in the entry elements capturing the contents of the first cell in each row.

*Rendering notes:*

The XML transform to HTML needs to ensure that the XML fragment

```
<entry role="rowhead">Content of row head cell 1</entry>
```

translates to the following HTML:

```
<td scope="row">Content of row head cell 1</td>
```

And similarly that

```
<thead>
  <row>
    <entry>Content of column head cell 1</entry>
  ...
</row>
</thead>
```

translates to:

<th scope="col">Content of column head cell 1</th>

DTD change request CR35.

Impact: manuscript→XML, XML rendering to HTML.

#### 4.3. *Handling of content objects that reside outside the Electronic Warehouse (change in ce:link)*

In its Virtual Total Warehouse, Elsevier is moving toward a content handling model in which figures, video files and other components of an article can be addressed individually as so-called content objects rather than as assets that are an inseparable part of the article. In this line of thinking these component files can be stored separate from the article rather than combined with it in a dataset in the Electronic Warehouse (EW); in case of video files for example in the DAM (the Digital Asset Management system) or even on a non-Elsevier system such as Akamai. The `ce:link` element is used for pulling in assets from the EW; the exact location of these files is not prescribed in the XML.

To allow content objects that reside outside EW datasets to be pulled in from external locations, `ce:link` was updated by the introduction of three optional attributes: `xlink:type`, `xlink:role` and `xlink:href`. These attributes work the same way as they do in the existing element `ce:inter-ref`: the `xlink:href` attribute contains the link to the external location where the content object resides, `xlink:type` and `xlink:role` describe properties of the link. Detailed specifications of the `href` linking scheme used in `ce:link` can be found in the VTW documentation.

*XML examples:*

```
<ce:link id="lk00360" locator="gr23"
  xlink:type="simple" xlink:href="pii:M2013402009010228X">
```

```
<ce:link id="lk00361" locator="gr1"
  xlink:type="simple" xlink:href="pii:S9999999414207732/gr1"/>
```

[Note that the model for element `ce:link` now is identical to that of `ce:inter-ref` with the addition of the locator attribute that specifies the name of the file that is to be retrieved.]

DTD change request CR22.

Impact: manuscript→XML, XML rendering.

#### 4.4. *Direct linking of keywords to external databases (change in ce:text)*

As of the introduction of CEP 1.4.0, it is possible to link keywords to an external database; previously, they could only be captured as plain text. This was done by changing element `ce:text` to allow it to contain `ce:inter-ref` as well. Detailed specifications of the `href` linking scheme used in `ce:inter-ref` can be found in a separate document [2].

*XML example:*

```
<ce:keywords id="kwds0010">
```

```

<ce: keyword id="kwd0010">
  Alpha Centauri
</ce: keyword>
<ce: keyword id="kwd0020">
  <ce: inter-ref id="ir0010" xlink:type="simple"
  xlink:href="ascl:1201.001">1201.001</ce: inter-ref>
</ce: keyword>
...
</ce: keywords>

```

*Rendered example:*

Keywords:  
Alpha Centauri  
[1201.001](#)

DTD change request CR84.  
Impact: manuscript→XML.

#### 4.5. Article numbers: introduction of *ce:article-number*, *sb:article-number*

The traditional, print-based way of citing journal article references makes use of the volume number, issue number and page number of an article (VIP citation). To some extent, the scientific community is moving away from VIP citation to a system in which articles are cited using an article number, sometimes in addition to (part of) the VIP information. This article number could be the DOI of the article, or a “generic” article number introduced by a journal or an STM publisher for citation purposes.

Two new elements were introduced to support the use of article numbers, *ce:article-number* and *sb:article-number*. Element *ce:article-number* allows articles to be tagged with an article number, and was added to the model of *item-info* as described in Section 2.2. Element *sb:article-number* allows articles to be cited using article numbers and was added to the model of element *sb:host*.

*XML example:*

```

<ce: bib-reference id="br0120">
  <ce: label>[12]</ce: label>
  <sb: reference id="sbr13">
    <sb: contribution>... </sb: contribution>
    <sb: host>
      <sb: issue>
        <sb: series><sb: title><sb: maintitle>Phys. Rev. Lett.</sb: maintitle></sb: title>
          <sb: volume-nr>90</sb: volume-nr>
        </sb: series>
        <sb: date>2003</sb: date>
      </sb: issue>
      <sb: article-number>194101</sb: article-number>
    </sb: host>
  </sb: reference>
</ce: bib-reference>

```

DTD change request CR89.

Impact: manuscript→XML, XML rendering.

#### 4.6. Unicode version 6: changes in use of *ce:glyph*

Nearly all conceivable symbols used in publishing are covered by the Unicode grid that has been part of the Elsevier DTDs since DTD version 5.0; however, a small number of essential symbols are not yet represented in Unicode, and have to be tagged using *ce:glyph* instead. Unicode version 6 has introduced ten Unicode code points for symbols that were previously only available through the Elsevier grid of glyphs. For reasons of backward compatibility these glyphs are not removed from the grid, but in DTD 5.4 files Unicode code points rather than *ce:glyph* elements have to be used to represent these symbols.

Symbol	Symbol name	Unicode code point	Obsolete glyph name	Elsevier grid position
ɔ̣	curly-tail d (phonetic symbol)	00221	dcurt	Pid
ɟ̣	j, undotted (phonetic symbol)	00237	jnodot	Pfj
◆	lozenge, filled	029EB	lozf	Bgi
ɱ̣	curly-tail n (phonetic symbol)	00235	ncurt	Phn
⬠	pentagon	02B20	pent	Bo1
◼	square, bottom filled	02B13	sqfb	Bfw
◼	square with filled N-E-corner	02B14	sqfne	Bfp
◼	square with filled S-W-corner	02B15	sqfsw	Bfr
◼	square, top filled	02B12	sqft	Bfv
ɮ̣	curly-tail t (phonetic symbol)	00236	tcurt	Pht

DTD change request CR75.

Impact: manuscript→XML.

#### 4.7. Better support of APA reference style: introduction of *sb:ellipsis*

In the sixth version of the so-called APA reference style, in case of eight or more authors only the first six and the last one are shown, as follows: "au1, au2, au3, au4, au5, au6, ... au37". To support this feature, a new, empty element *sb:ellipsis* was introduced to the model of *sb:authors*. This element captures the ellipsis typographical symbol (...) used in APA-style references to represent the omitted author names.

*XML example:*

```
<sb: authors>
  <sb: author>
    <ce: given-name>C. P. </ce: given-name>
    <ce: surname>Black</ce: surname>
  </sb: author>
  <sb: author>
    <ce: given-name>S. T. </ce: given-name>
    <ce: surname>Arlo</ce: surname>
  </sb: author>
  <sb: author>
    <ce: given-name>R. </ce: given-name>
    <ce: surname>Recht</ce: surname>
  </sb: author>
```

```

<sb: author>
  <ce: given-name>J. P. </ce: given-name>
  <ce: surname>Machlen</ce: surname>
</sb: author>
<sb: author>
  <ce: given-name>K. </ce: given-name>
  <ce: surname>Sempson</ce: surname>
</sb: author>
<sb: author>
  <ce: given-name>A. L. </ce: given-name>
  <ce: surname>Bee</ce: surname>
</sb: author>
<sb: ellipsis/>
<sb: author>
  <ce: given-name>S. P. </ce: given-name>
  <ce: surname>Clark</ce: surname>
</sb: author>
</sb: authors>

```

*Rendered example:*

Black, C. P., Arlo, S. T., Rechit, R., Machlen, J. P., Sempson, K., Bee, A. L., ... Clark, S. P.

[Note that in Elsevier XML files there are no spaces between initials in `ce: given-name`; these have to be added when the XML file is rendered in print PDF, web PDF or HTML for full compliance with the APA standard. The same holds true for the space that follows the ellipsis in the rendered example.]

DTD change request CR91.

Impact: manuscript→XML, XML rendering.

#### 4.8. *Structured correspondence addresses: extended use of ce:correspondence*

In the previous release of the CEP (v. 1.1.6) structuring of author affiliations was made possible by the introduction of element `sa: affiliation` (see Section 3.1.3). In the current release, this element was added to the data model of `ce: correspondence`, allowing any correspondence address present in the source content to be structured as well. All correspondence addresses should be tagged using `sa: affiliation` in a similar way as affiliations:

```

<ce: correspondence id="cor1">
  <ce: label>&#x0240E;</ce: label>
  <ce: text>Correspondence to: R. Schrauwen, Chief Content Architect,
    Elsevier, Radarweg 29, 1043 NX Amsterdam, The Netherlands
</ce: text>
  <sa: affiliation>
    <sa: organization>Elsevier</sa: organization>
    <sa: address-line>Radarweg 29</sa: address-line>
    <sa: city>Amsterdam</sa: city>
    <sa: postal-code>1043 NX</sa: postal-code>
    <sa: country>The Netherlands</sa: country>
  </sa: affiliation>
</ce: correspondence>

```

DTD change request CR99/COS-21.

Impact: manuscript→xml.

#### 4.9. Nesting of *ce:bibliography-sec* and *ce:further-reading-sec*

As of CEP 1.4.0, sections within the bibliographic references or further reading lists may be nested. This allows better formatting of extended reference lists. To enable this nesting, an optional element *ce: bibli ography- sec* was added to the data model of *ce: bibli ography- sec* and similarly an optional element *ce: further- readi ng- sec* to that of *ce: further- readi ng- sec*. Nesting is allowed up to one level deep.

*XML example (id and view attributes removed for improved legibility):*

```
<ce: bibli ography>
  <ce: secti on- tit le>References</ce: secti on- tit le>
  <ce: bibli ography- sec" >
    <ce: secti on- tit le>Human Papi l l oma Vi ruses</ce: secti on- tit le>
    <ce: bibli ography- sec>
      <ce: secti on- tit le>Preval ence and Pathogenesi s</ce: secti on- tit le>
      <ce: bi b- referenc e id="bi b0005"> ... </ce: bi b- referenc e>
      ...
      <ce: bi b- referenc e id="bi b00x0"> ... </ce: bi b- referenc e>
    </ce: bibli ography- sec>
  <ce: bibli ography- sec>
    <ce: secti on- tit le>Genotypi ng of Human Papi l l oma Vi ruses</ce: secti on- tit le>
    <ce: bi b- referenc e id="bi b00x5"> ... </ce: bi b- referenc e>
    ...
    <ce: bi b- referenc e id="bi b00xx"> ... </ce: bi b- referenc e>
  </ce: bibli ography- sec>
</ce: bibli ography- sec>
</ce: bibli ography>
```

*Rendered example:*

#### **References**

*Human Papilloma Viruses*  
*Prevalence and Pathogenesis*  
...  
*Genotyping of Human Papilloma Viruses*  
...

Impact: manuscript→XML, XML rendering.  
DTD change request CR97.

## **5. References**

[1] Section 508 of the US Rehabilitation Act, <http://www.section508.gov/>, and Web Accessibility Guidelines Working Group (WCAG 2.0), <http://www.w3.org/TR/WCAG20/>.

[2] External Object Linking - Supplier specification

## **6. Support**

The Elsevier DTDs and schemas are developed by the DTD Maintenance & Development Team, who will be happy to answer queries about the new DTDs. For this release, please contact [David Kuilman](#), [Jos Migchielsen](#) or [Rob Schrauwen](#).