DTD Development and Maintenance Team
Elsevier Science
Sara Burgerhartstraat 25
1055 KV Amsterdam
Netherlands

Phone: +31 20 485 3480
Fax: +31 20 485 3266
Email: s.pepping@elsevier.nl
r.schrauwen@elsevier.nl

Release note for full length article DTD version 4.2.0

DTD00002 — version 0.1, 21 February 2000, by Simon Pepping and Rob Schrauwen

version 0.1	21 February 2000	First draft
version 0.2	28 February 2000	Incorporated comments by RS and JH

Release 4.2.0 of the full length article DTD of Elsevier Science is an almost backward compatible update of version 4.1.0. With backward compatible we mean that files that are valid under version 4.1.0 are still valid under version 4.2.0. With almost we mean that some files that are valid under version 4.1.0 are not quite valid under version 4.2.0 because a few more closing tags have become mandatory.

This release note provides a summary of the changes, and points out the consequences for *suppliers* of SGML files, both external and internal, and for *users* of SGML files, mainly external and internal maintainers of electronic publishing platforms. More information about each change, including an account of the considerations, can be found in a series of DTD notes, which accompany this release.

1 General

The release consists of three new files: the SGML declaration art420.dcl, the DTD art420.dtd and the Character Entities file chars402.ent. Suppliers and users should place the new files in suitable directories.

The DTD and the Character Entities file have new Public identifiers:

```
-//ES//ENTITIES special characters version 4.0.2//EN
-//ES//DTD full length article DTD version 4.2.0//EN
```

Suppliers and users should add the new Public Identifiers to the catalog.

2 Entities and the ES character grid

One new symbol was added: `&euro;`, with glyph € and Unicode code point `8364` = `20AC hex`.

A number of corrections resulted in new entities:

- `&iinfin;`, which is the proper entity associated with the symbol at grid position `Bkz`.

- `&smid;`, `&spar;`, `&nsmid;`, `&nspar;`, which are the proper entities associated with grid positions `Bd6`, `Bd7`, `Be6`, `Be7`.

One existing entity, `&z.mstpos;`, was associated with an incorrect grid position, `Bkz`. It is now associated with grid position `Bpz`; moreover, the glyph shown at this position was not quite correct, and has been modified.

Two more glyphs have been corrected:

- grid position `Bq3`, associated with entity `&cong;`, has been mirrored.
- grid position `CAy`, associated with entity `&z.shtsls;`, has been made smaller; it has the size of the stroke through the Polish ell.

The accent in the accent constructions for the lc and capital Polish ell, associated with the grid positions `CFl` and `CFm`, has been changed from `&z.xl;` to `&z.shtsls;`.

3 Front matter

Suppliers should use the new Public Identifier in the Doctype declaration and the new value 4.2.0 for the `version` attribute of the `art` element.

The `refers-to` attribute of the `art` element may now refer to multiple other articles. Together with the manuscript Elsevier Science should supply the PIIs of all articles referred to, and suppliers should enter them as a space separated string (`NMTOKENS`). Users should allow for more than one referenced article.

The Publication Item Type `req`, *Request for assistance*, has been added to the list of possible PITs (the `docsubty` attribute of the `art` element). Elsevier Science will provide a description of the expected content of articles of this PIT, and of the standard typographical representation. Suppliers should implement this PIT and its guidelines. Users should prepare to receive and present articles of this PIT.

DTD 4.2.0 makes it possible to capture a reference to the preprint version of the article if that has been published on a “recognized” preprint server. For this purpose there is a new element `preprint`, which may be inserted after the `copyright` element. Its sole content is an `inter-ref` element. The `inter-ref`’s `object-type` is always `preprint`. Its `locator-type` is one of a list that will be published by Elsevier Science; currently the only recognized value is `xxx-archive`. The content of the `inter-ref` may be empty.

As for the `copyright` element, the content of the `preprint` element, if any, should not be presented where it occurs in the SGML file. Instead, it should be saved, along with the information in the attributes, and presented elsewhere as specified for the publication in question. For example, in a web publication one might generate a link to the preprint in a footnote, where the text of the link is the content of the `preprint` element, or, if the element is empty, a default

text. In standardized output, the `preprint` element currently generates no text; however, this is liable to change.

See also the changes to the `inter-ref` element in section 4.1.

3.1 The author group `aug`

There are several changes to the author group `aug`.

1. Correspondence addresses, `cor`, are now cross-referenced explicitly; therefore they have an `id` attribute and a `no` element, like other elements that can be the target of a cross-reference. This makes it possible to have several corresponding authors with one correspondence address.

   The correspondence address may have a `cty` and a `cny` element, like affiliations. Different from affiliations this is optional; it is not possible to have a `cty` element without a `cny` element.

2. Footnotes, `fn`, are now allowed in the content of an `aug` element. This makes it possible to place footnotes to authors and the cross-references to them after the `au` element, where they more properly belong.

3. As a consequence of the above changes, all `cross-ref` elements are now placed after the `au` element. The standardized markup is as follows:

   - First a `cross-ref` element to the correspondence address.
   - Then a `cross-ref` element to the affiliations.
   - Finally a `cross-ref` element to the footnotes.

The `cross-ref` elements are followed by the footnotes `fn`, the correspondence address `cor`, and the electronic address `ead`.

```
<aug><au><fnm>Simon<snm>Pepping</au>
     <cross-ref refid="cor1"><sup>&ast;</sup></cross-ref>
     <cross-ref refid="aff1 aff2"><sup>a,b</sup></cross-ref>
     <cross-ref refid="fn1 fn2"><sup>1,2</sup></cross-ref>
     <fn id="fn1"><no>1</no><p>Partially supported...</fn>
     <fn id="fn2"><no>2</no><p>Acknowledges grants...</fn>
     <cor id="cor1"><no>*</no>Corresponding author. Present address:
         Sara Burgerhartstraat 25, 1055 KV <cty>Amsterdam</cty>,
         <cny cny-code="nl">Netherlands</cny>.
     <au><fnm>Rob<snm>Schrauwen</au>
     <cross-ref refid="cor1"><sup>&ast;</sup></cross-ref>
     <cross-ref refid="aff2 aff3"><sup>b,c</sup></cross-ref>
     <cross-ref refid="fn3"><sup>3</sup></cross-ref>
     <fn id="fn3"><no>3</no><p>Supported by...</fn>
```

Suppliers should change their markup to this model. While markup according to the model of DTD 4.1.0 remains valid as SGML, it will no longer be accepted as a deliverable.

Users should prepare to accept the above model. Its rendering is expected to be less complex that that of the old model, because no cross-reference symbols have to be generated, and all `cross-ref` elements are rendered at the position at which they occur in the SGML file. But a new feature is, that the series of `cross-ref` elements should be joined in presentation by a suitable symbol. In standardized output, they should be joined by superior commas.

In addition, users should be able to accept files that still have the model of DTD 4.1.0. Such files may have footnotes and their `cross-ref` elements in the `au` element. These should be treated as under DTD 4.1.0. `cor` elements without `id` attribute and `no` element are `cor` elements according to the old model; they should generate a * as under DTD 4.1.0.

3.2 Other changes to the author group `aug`

1. Relating an author to his/her vitae in the backmatter, `vt`, is no longer done via an empty cross-reference. Instead, the author has a new attribute `vtid`, which is a reference to the `id` attribute of the `vt` element.
2. In addition to `jr`, `degs`, and `roles` elements an author can now also have a `ranking` element, which can have a superior * as content, to denote an important author. No footnote or correspondence address needs to be associated with the ranking *.
3. The electronic address `ead` has a `type` attribute. Its default value is `email`, which was the only type under DTD 4.1.0; the other possible value is `url` for a URL.

```
<au vtid="vt1"><fnm>Simon<snm>Pepping<ranking><sup>&ast;</sup></au>
<ead type="url">www.elsevier.nl/~spepping
<au vtid="vt2"><fnm>Rob<snm>Schrauwen<ranking><sup>&ast;</sup></au>
<ead type="email">r.schrauwen@elsevier.nl
```

Suppliers should change their markup according to the above model. Especially, a * without an associated correspondence address should be marked up as a `ranking` element. Other electronic addresses than standard email addresses may be expected.

Users should now relate an author to his vitae `vt` via the `vtid` attribute instead of an empty `cross-ref` element.

The content of the `ranking` element should be joined with the `cross-ref` elements following the author. A joining superior comma should be added in standardized output.

When the electronic address is turned into a hyperlink, attention should be paid to its `type`.

3.3 Graphical abstracts

Graphical abstracts summarize the contents of the paper in a concise pictorial form. Usually, they contain a structure formula as a picture, although graphical abstracts without a picture do occur.

DTD 4.2.0 makes it possible to capture graphical abstracts in the SGML file. They are to be tagged as an abstract with a new value of the `class` attribute: `graphical`. The picture is included in one of the paragraphs as a `fig` element with a new value of the `loc` attribute: `abs`. Since such a figure is not expected to be cross-referenced, it need not have a value for the `id` attribute nor a `no` element.

```
<abs class="graphical">
  <p>Enantiomerically-enriched &alpha;-silyl aldehydes may be prepared
     from 2,3-epoxy alcohols via ring-opening and oxidative cleavage.
     <fig loc="abs">
       <caption><p>R=Me, <it>i</it>-Bu, <it>t</it>-Bu; Y=O,
          CHCH<inf>2</inf>CH=CH<inf>2</inf>,
          C(CH<inf>2</inf>CH=CH<inf>2</inf>)<inf>2</inf>
       <link locator="gr1">
     </fig>
</abs>
```

In standardized output all graphical abstracts are grouped together and printed in the front of the issue.

3.4 Stereochemistry abstracts

Some journals ask their authors to submit for each important chiral compound a stereochemistry abstract, which contains details about the compound. Although they are called abstracts, stereochemistry abstracts are in nature more like “keywords” or “nomenclature” than “abstracts”. There may be several stereochemistry abstracts per article.

DTD 4.2.0 makes it possible to capture stereochemistry abstracts in the SGML file. They are to be tagged as a new element `stereochem`. They are very structured, and contain the following elements:

- the (pictorial) structure `compound-struct`, consisting only of a `link` element;
- the name `compound-name`;
- the chemical formula `compound-f`;
- stereochemical information `compound-info`, consisting of one or more list items `li`.

```
<stereochem>
  <compound-struct><link locator="fx1">
  <compound-name>(<it>S</it>)-2-<it>t</it>-Butyldimethylsolylpent-4-enal
  <compound-f>C<inf>12</inf>H<inf>22</inf>OSi
  <compound-info>
    <li><p>E.e. &ge; 95%
    <li><p>[&alpha;]<sup>25</sup><inf><scp>D</scp></inf>= ...
    <li><p>Source of chirality: Sharpless AE
    <li><p>Absolute configuration: <it>S</it>
<stereochem>
  <compound-struct><link locator="fx2">
  <compound-name>(<it>S</it>)-2-<it>t</it>-Butyldimethylsilhex-5-enal
  <compound-f>C<inf>12</inf>H<inf>24</inf>OSi
  <compound-info>
    <li><p>E.e. &ge; 95%
    <li><p>[&alpha;]<sup>25</sup><inf><scp>D</scp></inf>= ...
    ...
<stereochem>
  <compound-struct><link locator="fx3">
  ...
```

Stereochemistry abstracts may be used to build an index, or, on electronic platforms, a database of stereochemistry compounds.

3.5 A few small changes

The list of country codes published with DTD 4.1.0 contained an error, which is corrected in DTD 4.2.0. The country code `zr` (used inapproprately for Zambia) has been removed and the country code `za` (to be used for South-Africa) has been added. The country code `za` now stands for South-Africa, and the country code `zm`, which in DTD 4.1.0 was associated with South-Africa, now stands for Zambia. Section 3 of the DTD Implementation Note *The CNY element and the CNY-CODE attribute* is herewith obsolete.

The abstract class `inspec` has been removed. Since it was not really a valid abstract class, and was erroneously included in the list of abstract classes of DTD 4.1.0, we do not expect that this removal has practical consequences.

Three new keyword classes have been added:

- `stma`, *Statistical Theory & Method Abstracts* (from ISI);
- `astronomy` and `geo`.

The `stma` class should be presented as: ‘Statistical Theory & Method Abstracts’. The classes `astronomy` and `geo` should be presented as: ‘Keywords’; they are controlled keywords that replace `class="kwd"` in some journals that Elsevier Science will specify.

4 Body

4.1 The element `inter-ref`

The attributes `locator-type` and `object-type`, which in DTD 4.1.0 could only take one of a limited list of values, can now formally take any string (`CDATA`) as a value. In practice, their value should be one of a documented list of values. This enables us to update the list more flexibly as the need of additional entries arises.

The attribute `locator` has been made mandatory. Formally this violates the requirement of backward compatibility. In practice `inter-ref` elements without a `locator` attribute make no sense, and the change will not cause any problem. `locator-type` remains formally optional, but in practice we require it to be present.

The QC tool will perform two checks on this element. It will check the presence of the attribute `locator-type`, and it will check that the attribute `locator-type` has one of the documented values, and that the attributes `object-type` and `refid` have one of the corresponding values or are absent.

Currently, the documented values are as follows.

- `locator-type="aoi"`: The locator is an astronomical object identifier. No object types are to be specified.

```
<inter-ref locator="LC 123" locator-type="aoi">LC 123</inter-ref>
```

- `locator-type="doi"`: The locator is a DOI. No object types are to be specified.

```
<inter-ref locator="10.1016/S12384107697000225"
    locator-type="doi">...</inter-ref>
```

- `locator-type="fiz"`: The locator addresses a document in the FIZ database. The attribute `object-type` is mandatory here; it must have the value `inspec` (an inspec record).

```
<inter-ref locator="85:2535122" locator-type="fiz"
    object-type="inspec">...</inter-ref>
```

- `locator-type="genbank"`: The locator is a Genbank address. No object types are to be specified.

```
<inter-ref locator="AB026824" locator-type="genbank">AB026824</>
```

- `locator-type="pii"`: The locator is a valid PII number of an article that is guaranteed to come from the same publisher as the article in which the `inter-ref` appears. No `object-type` is specified. The attribute `refid` can be an ID in the target document.

```
<inter-ref locator="S0167839697000186" locator-type="pii">...</>
```

- `locator-type="url"`: The locator is a URL. Optionally an `object-type` can be specified: it must be a MIME type. The optional `refid` is a named location in the target object.

```
<inter-ref locator="http://www.elsevier.com" locator-type="url"
    object-type="text/html">Elsevier Science homepage</inter-ref>
```

- `locator-type="xxx-archive"`: The locator is an address of the `xxx-archive` database. The attribute `object-type` is mandatory here; currently the only allowed value is `preprint`.

```
<inter-ref locator="hep-th/9112009" locator-type="xxx-archive"
    object-type="preprint">hep-th/9112009</inter-ref>
```

Thus `inter-ref` allows one to capture a large array of locator types. This imposes upon users the burden of representing all those types. In print, only the content of the element needs to be reproduced. Some publications may require that the locator is printed in some form, depending on the locator type. For example, a journal may require that URLs are shown in a footnote. In electronic media the element may be represented as a hyperlink, depending on the locator type and on the options of the publication site. For example, the FIZ locator allows one to request the abstract from the FIZ database, using the protocol to access that database; but this only makes sense if readers can gain access to the abstract from the publication site.

Applications should ignore `inter-ref` for locator types they do not support and not complain if unknown locator types are used.

4.2 The element `intra-ref`

`intra-ref` is used to refer to objects "under control of the publisher". The difference between `inter-ref` and `intra-ref` is that the latter does not require the connection to a foreign host.

The `intra-ref` element has no `locator-type` attribute, because its `locator` always denotes an entity. The type of the `object-type` attribute has been changed to `CDATA` like for the `inter-ref` element.

There is currently no documented value for `object-type` (except for the values `embase` and `geobase`, which are reserved but not used). The currently documented usage for `intra-ref` is to cross-reference between documents and their linked textboxes or between chapters of a book.

An `intra-ref` can now refer to more than one place in the referenced document. To this end the type of the attribute `refid` has been changed to `NAMES`. The value of `refid` should be one or more `ID`s of referable elements in the referenced document. Note that referable elements should contain an element that can be used as a text representation of the link, such as a `no` element.

```
See <intra-ref locator="main"
 refid="bib17 bib18 bib19 bib22">[17-19,22]</intra-ref>
See <intra-ref locator="ch2">Chapter 2</intra-ref>
```

4.3 Other changes

Versions 4.0.0 and 4.1.0 of the DTD did not have a table legend, so that work arounds had to be used. Fortunately, the table legend is back in the DTD. It is to be captured in the `legend` element, which is the last element in a `tbl`, right after the `tblbdy` or `link`.

In DTD 4.2.0 it is possible to register a separate copyright status for `fig` and `upi` elements. If the copyright status of a `fig` or a `upi` is different from that of the whole document, one may insert a `copyright` element after the `caption`. The rules for this `copyright` element are identical to those for the `copyright` element of the whole document.

The element `inline-fig` has been furnished with an attribute `baseline`, whose default value is `0.0`. It denotes the fraction of the height of the graphic that must appear below the baseline. For this feature a separate test will be conducted; until further notice its usage is not required in SGML files that are delivered to Elsevier Science.

5 Back matter

In DTD 4.2.0 it is possible to group several `bb` or `other-ref` elements in a single `bib` element. To make these contained `bb` or `other-ref` elements individually referable, they may contain an `id` attribute and a `no` element.

```
<bib id="bib7">
<no>[7]</no>
<bb id="bib7a">
  <no>(a)</no>
  <contribution> ...
  <host> ...
<other-ref id="bib7b">
  <no>(b)</no>
  A. Author, ...
</other-ref>
```

In DTD 4.2.0 there is more room for adding comments to a structured reference `bb`. One can insert a `comment` element before the `contribution`, between the `contribution` and the `host`, and after the `host`. In order to maintain backward compatibility, the `comment` elements *within* a `contribution` or a `host` are still possible, but deprecated. This means that SGML files delivered by suppliers should use the new positions of the `comment` element. In order to tag the position of the `comment` element unequivocally, the closing tag of the `contribution` and `host` elements has been made mandatory. This is a small, non-backward compatible change, which is only noticeable with files that use tag minimization.

For users the consequences of this change are small. In DTD 4.2.0 files there may be an additional `comment` element at the start of a structured reference `bb`. A `comment` element between the `contribution` and the `host` is almost equivalent to a `comment` element at the old position at the end of the `contribution`; likewise a `comment` element after the `host` is almost equivalent to one at the old position at the end of a `host`.

In the new DTD it is possible to add explanatory text to a `bib` element. The explanatory text is tagged as a `note` element, which contains one or more `p` elements, and is added at the end of the `bib` element.

Users should observe that the `note` need not necessarily be presented at the end of the `bib` element. A journal style, e.g., might require that it be presented at the start of the `bib` element.

A `bib` element may also consist of a `note` alone. Thus the so-called end notes, i.e. notes within the bibliography, should now be captured in the new `note` element within a `bib` element, instead of in an `other-ref` element. This makes it clear that they are notes, not references.

```
<bib id="bib49">
<no>49.</no>
<bb><comment>See the references in
    <contribution>...
```

```
    <comment>first published in
    <host>...
    <comment>also available electronically as
    <host>...
    <comment>(in Japanese)
</bb>
<note><p>This reference explains the usage of the comment and note
    elements. Comments and the other components of bb together
    form one text. The note may contain details about the
    reference.
    <p>It might be necessary to be more careful with punctuation.
</note>
<bib id="bib50">
<no>50.</no>
<note><p>This is a note in between the references; in DTD 4.1.0 it would
    be tagged as other-ref, even though it is not a reference at all.
    Such notes are often seen in Tetrahedron.
</note>
</bib>
```

We have added the possibility to give an `id` attribute to `bibl` and `further-reading` elements. This makes it possible to make a cross-reference to these elements.

```
<p>Finally, see the books listed in the <cross-ref
   refid="prog">further reading</cross-ref> section on
   programming.</p>
```

In a `further-reading`, the new DTD makes it possible to intersperse paragraphs of text, `p`, between the `bib` elements. The relation between the text and the references is not defined: the text may refer to one or more references above it, or to those following it, or have any other relevance to the references.

```
<further-reading id="prog">
  <st>Useful books on programming</st>
  <p>One of the most famous works on programming is the following.</p>
  <bib id="fr1">
    <no>1.</no>
    <bb>
      <contribution>
        <authors>...
        <title>Fundamental Algorithms</title>
      </contribution>
      <host>
        <book>...
      </host>
  </bib>
  <p>The following is <it>the</it> theoretical treatise on object
  oriented programming.</p>
  <bib id="fr2">
```

```
    <no>2.</no>
    <bb>
      <contribution>
        <authors>...
        <title>Object-Oriented Software Construction</title>
      </contribution>
      <host>
        <book>...
      </host>
  </bib>
  <p>There are many excellent text books on these topics, too many
  to be listed here.</p>
</further-reading>
```

The DTD note 99003 contains further explanations and an extensive example.