Central Application Management, Operations
Elsevier
Radarweg 29
1043 NX Amsterdam
Netherlands

Phone: +31 20 485 2316
Fax: +31 20 485 3266
Email: j.migchielsen@elsevier.com

Contrast-out schemas

jmi07009 — version 1.0, 20 August 2007, by Jos Migchielsen

version 0.1	26 July 2007	Draft, comments from RSc
version 0.2	9 August 2007	Draft, comments from SPe, RSc
version 1.0	20 August 2007	Final, comments from RCa

Contents

1. Introduction

The Electronic Warehouse delivers material to its customers in so-called datasets. These datasets can be structured according to the old standard EFFECT or according to the new standard CONTRAST (content transport standard). Each Contrast dataset contains a dataset package file, dataset.xml, which is an XML file structured according to an XML schema. This file authoritatively describes the content of the dataset. Customers are expected to access all content via this file (instead of relying on conventions which are subject to change). There are separate schemas for journal deliveries and for book deliveries.

In this document we describe how Contrast schemas are expected to be used by the EW customers and how they will be notified about new versions.

2. Contrast-out schemas

The Contrast-out schemas are updated regularly, in principle with every EW release, i.e. 2–3 times per year. The version number of the schemas EW 1.10 uses for journal deliveries is 1.4; for book deliveries it is 1.2. The EW will in principle only use one version at any one time.

A journal output dataset contains, as from EW 1.10, the following in the dataset package file (dataset.xml):

```
<dataset
  xmlns="http://www.elsevier.com/xml/schema/transport/ew-xcr/journal-
1.4/issues"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation=
```

```
    "http://www.elsevier.com/xml/schema/transport/ew-xcr/journal-1.4/issues
     http://www.elsevier.com/xml/schema/transport/ew-xcr/journal-
1.4/issues.xsd"
  schema-version="1.4">
```

Here, `http://www.elsevier.com/xml/schema/transport/ew-xcr/journal-1.4/` `issues` is the namespace URI of version 1.4 of the EW output schema for serial publications. As usual in the XML world this is not an actual URL. The location of the XML schema file itself is to be based on the schema location which is captured in the `xsi:schemaLocation` attribute. The convention is that the schema location is almost identical to the namespace. The schema location is not an existing URL on the Elsevier website either.

It is expected that applications use XML catalogs in order to retrieve an instance of the schema. The catalog should e.g. contain

```
<!-- Contrast-out Schemas -->
<group xml:base="file:/path_to_contrast_out_schemas/">
    <rewriteSystem
systemIdStartString="http://www.elsevier.com/xml/schema/transport/ew-xcr/"
rewritePrefix=""/>
    <rewriteURI
uriStartString="http://www.elsevier.com/xml/schema/transport/ew-xcr/"
rewritePrefix=""/>
</group>
```

Here `path_to_contrast_out_schemas` is the path to the directory that contains the schemas.

An XML-catalog aware application will find the schema that is used in the dataset.xml file. It is the second part of top element `dataset`'s attribute `xsi:schemaLocation`. It will then replace part of that URL by an empty string as described in the catalog. Finally it will prepend the result with the base as described in the catalog. In our example the application will end up with `file:/path_to_contrast_out_schemas/journal-1.4/issues.xsd`.

3. Versions

As stated above the output schemas are updated regularly, in principle with every EW release. Using an XML catalog as described above makes it easy to introduce new schemas. For example:

With the next EW release (1.11) new journal out-schemas will be used, version 1.5. Before the final version is published so-called review versions are created. These schemas have r1, r2, ... appended to their version number, so 1.5r1, 1.5r2, etc. This version number is only visible in the schema itself. It is not used in the dataset package files.

Datasets that use one of these review versions of the schema will contain the following:

```
<dataset
  xmlns="http://www.elsevier.com/xml/schema/transport/ew-xcr/journal-
1.5/issues"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation=
    "http://www.elsevier.com/xml/schema/transport/ew-xcr/journal-1.5/issues
```

```
      http://www.elsevier.com/xml/schema/transport/ew-xcr/journal-
1.5/issues.xsd"
  schema-version="1.5">
```

Note that the r-version is not mentioned.

To be able to process a dataset that uses one of these review versions of the new schema one only has to add a directory *path_to_contrast_out_schemas*/journal-1.5 and fill it with the new schema(s). It should be possible to do this without a release of the application.

Every review version is replaced by the next review version (by overwriting it). The last review version is replaced by the final version.

Because new versions have to be finalised a long time before they are actually used by the EW it is possible that some changes have to be made after finalisation. This is done in the form of patches. The patched schemas will have p1, p2, ... appended to their version number. Again, the new version number (e.g. 1.5p1) is only visible in the schema itself. The patched version should replace the final version of the schema. A second patch version replaces the first one, etc.

A flexible way of depositing a set of schema files in a new or existing directory is a requirement for customers who receive Contrast-out material.

4. Namespaces

It should be noted that with the introduction of a new version of the schema also a new namespace is introduced. Any XML application that is used to process Contrast-out material should be able to handle a new namespace. The namespace can be found in top element dataset's attribute xmlns.

5. New versions

As stated above Contrast will be subject to a version release schedule similar to that of the EW. Systems receiving Contrast output are expected to be able to seamlessly integrate updated versions of the Contrast schemas in their workflow by installing them in such a way that they can be replaced easily. In this respect Contrast is different from its predecessor, Effect, which has remained more or less unchanged during its lifetime.

The release schedule of Contrast will run in parallel with that of the EW, each new version of the Contrast schemas becoming available for testing purposes at approximately the same time the previous versions of EW and Contrast go live.

New versions of the Contrast schemas will retain backward compatibility with their predecessors.

Two situations may occur, (1) the tags newly introduced in Contrast are not required by the platform receiving the dataset, or (2) the newly introduced tags are required. In the first case the platform should be able to simply ignore the new tags, in the second case the involvement of the receiving platform during testing will be expected as part of an end-to-end implementation.

In case a resupply of a dataset takes place it is very well possible that in the period between receipt of the original dataset and the resupply a new Contrast version was introduced. All receiving platforms are expected to be able to handle this situation.

6. Contrast-in

The Contrast standard also describes the way material is delivered to the EW. There also the dataset package files are XML files structured to according to XML schemas. These schemas are called the Contrast-in schemas while the schemas used for material from the EW are called the Contrast-out schemas. Note that "in" and "out" are used with respect to the EW, the Contrast-out material is "platform-in" material.

7. Notifications

The Contrast schemas are created by the Elsevier DTD Development & Maintenance Team (DMT). The latest version of the schemas and their documentation are publicly available on NonSolus [1] and on the Elsevier corporate website [2]. If you would like to be notified when a new version (review, final or patched) is published please send an email to Jos Migchielsen (j.migchielsen@elsevier.com).

The EWOps department will also send a notification to their customers announcing each new EW release. This will include the date on which a new Contrast version will be used (usually the EW release date).

8. References

[1] NonSolus Contrast page, http://nonsolus/dtd/contrast.htm.
[2] Elsevier Corporate Website Contrast section, http://www.elsevier.com/wps/find/librariansinfo.librarians/dtds_htm#sec6.