

Technical Bulletin #16

This technical bulletin describes Astoria's support for the XML standard. Astoria has supported the XML standard since Astoria 3.0, but as the standard has been further defined, we have added special configuration settings that enable better support for some aspects of XML. In some cases, there are aspects that are not fully supported.

Note: In Astoria 3.4.3, we improved support for several aspects of the XML standard. The information in this technical bulletin applies directly to Astoria 3.4.3, but it is also applicable to 3.5 and 3.6.x; however, many of the settings recommended in this technical bulletin are now automatically included in these later releases. Further information can be found in the release notes included with each software version.

Case Sensitivity

The XML standard supports case-sensitive element names. Case sensitivity makes all aspects of element name management case sensitive. For example, databases created in this fashion require case-specific element name queries in Search (e.g., 'para' will not find instances of 'PARA' elements).

To enable case-sensitive element names in Astoria, you must specify this preference when the database is created. To create case-sensitive databases, you must ensure that your **astoria.ini** file contains the following setting when the database is created using the User Administration tool:

```
[Database]
InitialCaseSensitivity=1
```

If you use this setting and your database will contain both SGML and XML documents, you must change your SGML documents and DTDs to contain an SGML declaration with:

```
NAMECASE GENERAL NO
```

The default astoria declaration contains **NAMECASE GENERAL YES**, which will cause parsing errors for SGML documents.

Character Entities

In order to better handle character entity sets that define 16-bit character encodings, set the following entries in your **astoria.ini** file:

```
[SGML Import]
ReplaceInvalidCharRef = ?
MaintainCharRefs = 1
```

URI Syntax

Although support for the syntax will be included in the next release of Astoria, it does not yet support the **file:///mysystemid** URI syntax for system identifiers. Use a system identifier such as **c:\graphics\mygraphic.gif** or **\\servername\dtds\mydtd.dtd** instead.

Using Correspondent and ADEPT*Editor

If you use Correspondent and ADEPT*Editor to edit XML, you may get SGML errors during checkin. These errors occur because ADEPT*Editor requires you to use PUBLIC identifiers in your DOCTYPE statement, while XML requires you to use SYSTEM identifiers. Fortunately, as long as you use a SYSTEM identifier, XML permits you to include a PUBLIC identifier, too.

Note: Ensure that you actually edited the file and saved your changes; ADEPT*Editor makes various modifications to the file that Correspondent expects to see which may result in similar SGML errors.

To include a PUBLIC identifier:

- 1 Edit the DOCTYPE statement in your **.dcl** file to resemble the following:

```
<!DOCTYPE mydoc
PUBLIC "-//Your Company//DTD Your DTD Name//EN"
"full path to your DTD"
[]>
```

- 2 Edit the DOCTYPE statement in your **.xml** document to resemble the following:

```
<!DOCTYPE mydoc
PUBLIC "-//Your Company//DTD Your DTD Name//EN"
"full path to your DTD" >
```

- 3 Delete the old DTD and XML document.
- 4 Reimport the DTD and XML document.
- 5 Try editing an element using Correspondent.