Best Practice for Using the DITA CONREF Attribute for Translation

By Andrzej Zydron
For OASIS DITA Translation Subcommittee

4 March 2008

OASIS

OASIS (Organization for the Advancement of Structured Information Standards) is a not-for-profit, international consortium that drives the development, convergence, and adoption of e-business standards. Members themselves set the OASIS technical agenda, using a lightweight, open process expressly designed to promote industry consensus and unite disparate efforts. The consortium produces open standards for Web services, security, e-business, and standardization efforts in the public sector and for application-specific markets. OASIS was founded in 1993. More information can be found on the OASIS website at http://www.oasis-open.org.

The purpose of the OASIS DITA Technical Committee (TC) is to define and maintain the Darwin Information Typing Architecture (DITA) and to promote the use of the architecture for creating standard information types and domain-specific markup vocabularies. The Translation Subcommittee defines best practices and guidelines for DITA authoring, translation and localization, and recommends solutions for industry requirements for consideration by the OASIS DITA TC. The group recommends widespread adoption of these concepts through liaisons with industry, other standards, and providers of commercial and open source tools.

Table of Contents

1. Statement of the Problem

Following the DITA specification and best practices, all standard text that can be reused in many places, such as a warranty disclaimer, should be handled via the conref mechanism. Using conref ensures that all translatable text can be handled with one XML translation process.

However, great care should be taken regarding the use of the conref attribute in DITA topics that will be translated. The conref attribute may appear to be a very efficient artifice to allow for the simple substitution of product names and other variables, allowing for the runtime resolution of given parameters. However, runtime resolution can cause grammatical problems with highly inflected and gender sensitive languages such as most Slavonic and Germanic languages, where nouns can change (both root and ending changes are possible) depending on case/context.

Runtime resolution poses less of a problem in languages such as English and French, where only the genitive case requires some modification—or a leading vowel demands a different form of indirect article. Therefore, writers in English and French are tempted to use single noun substitution. This substitution will potentially lead to ungrammatical text in both source and target languages.

What works extremely well for English or French source text can result in ungrammatical text for other target languages. Such an outcome is generally undesirable because it reflects badly on the quality of the documentation.

2. Recommended Best Practices

This section recommends the best practices to use in carefully applying the conref attribute in text that is to be translated into inflected languages.

We understand that many software teams are actively pursuing conref'ing all uicontrols and wintitles in their content to allow for single sourcing the text strings used by the software and the references to those strings in content. This approach ensures that the text in the documentation matches what is on the screen.

Conref is also commonly used for other kinds of volatile content, such as product names. Some teams may use conref for audience names or roles so that they can be changed where they occur in text for customer-specific terminology. If that becomes broadly adopted, we may see a common scenario in which nouns and noun phrases are conref'ed.

As a general rule, the safest way to use conref is to reference only grammatically complete sentences, phrases, or complete blocks of text. This text can stand alone and does not rely on the surrounding text in the sentence for its context). For example, the note in the following example is completely independent text that stands alone.

_____

…

</steps>

<result>The call is ended</result>

<postreq>

<note conref="reuse.dita#reusablephrases/hangup"/>

</postreq>

_____

Common nouns

If you are translating into inflected languages, avoid using conref to substitute common nouns or noun phrases. For example, do not use 'conref" to add individual words to a sentence, such as 'hammer' or 'screwdriver,' and so on. Avoid using conrefs in situations such as the following:

_____

<p>

Using a <keyword conref="tools.dita#tools/ClawHammer"/>, remove the CPU from its mount.

</p>

_____

Proper nouns

Where the conref is used for proper nouns, such as *Ford Focus,* the safest solution is to ensure that the proper noun is always the subject of the sentence and in the nominative case.

Do not use:

——————————————————————————

<p>

Driving the <keyword conref="models.dita#ford/FordFocus"/> is an exhilarating experience.</p>

——————————————————————————

Rather use:

——————————————————————————

<p>

The <keyword conref="models.dita#ford/FordFocus"/> provides an exhilarating driving experience.</p>

——————————————————————————

The subject of this sentence is in the nominative case and therefore does not normally require inflection.

Ensure that the translators receive appropriate instructions to be certain that the substituted proper nouns are translated as the subject of the sentence.

In the case of a sentence written as a command and containing a conref'ed element, it may not be possible to change the noun to the nominative case.

For example, in the sentence, "Click the OK Button" , the noun is the object of the verb. In the event that this noun is conref'ed from a list of user interface controls, the conrefs are best resolved prior to translation. In this way, the translator can select the proper grammatical construction for the noun in the target language.

Note: The topic IDs in the examples above are illustrative only. The topic ID need not match the file name.

Instruction to processors

It is recommended that some conrefs, especially those involving inflected languages, be resolved prior to translation. For example, all conref targets to "uistrings.dita" would be resolved in the content before it is sent to translation, giving the maintainers of the source content the full advantages of conref for volatile content without affecting translation capabilities.

> For example, consider managing product name volatility:
> If we know that the content ships with three products and if product names are treated the same in languages x, y, and z, then we defer the conref resolution for those languages so that topics only get translated once.

> But for languages where the product name is treated differently depending on context, we resolve the conrefs prior to translating. Those languages require that all three versions of each topic that mentions the product name must be translated separately.

Instruction to the translators

It is recommended that the translator be provided with a composed version of the source text to review and understand the context of the text in which the conref appears. The translator will need to review a composed version of the translation with the conref resolved so that the grammatical implications of the translation can be understood.