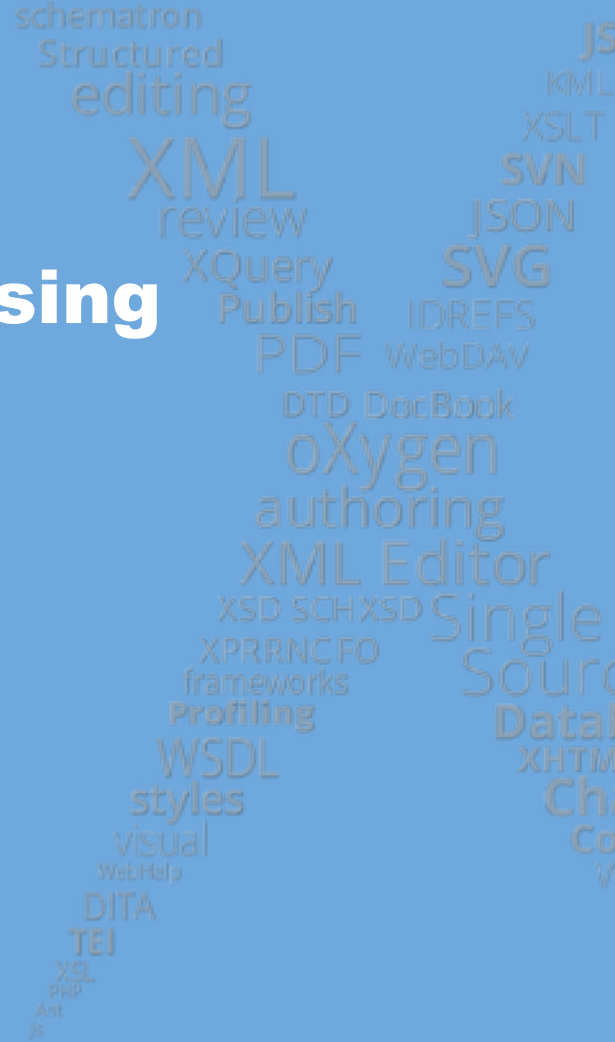


Providing content to AI engines using DITA OT plugins

Radu Coravu, Syncro Soft
radu_coravu@oxygenxml.com

© 2025 Syncro Soft SRL. All rights reserved.



Bio

- Software engineer working for Oxygen XML.
- DITA XML expert.
- Contributor in the DITA OT project.
- Lots of DITA XML related technical support.
- Small articles, blog posts, open source projects, documentation updates.
- Working on AI-based DITA XML editing tools.

[Radu Coravu - Bio](#)



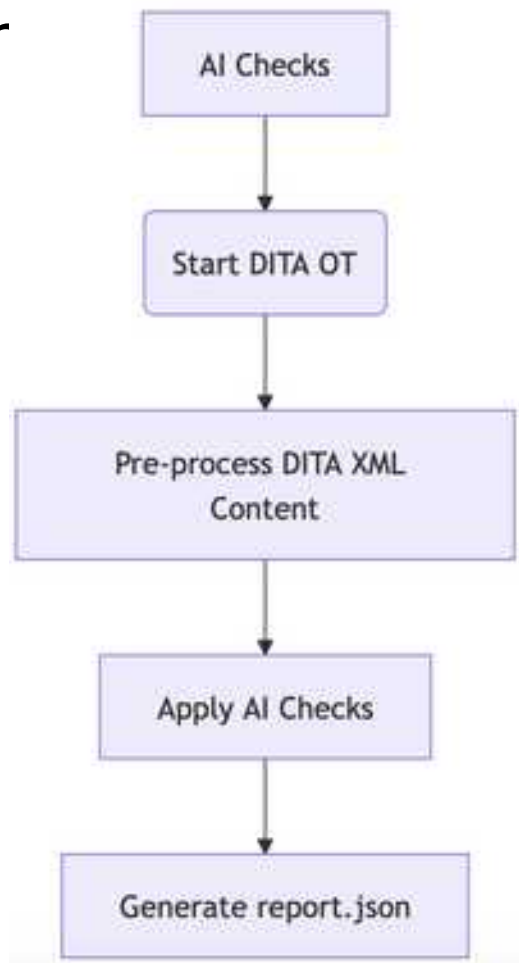
Overview

Proof of concept implementation of various AI-related features as DITA Open Toolkit plugins



(1) Apply AI Checks on DITA XML Center

- DITA OT Plugin which pre-processes the DITA XML content (resolves content and key references, applies filters).
- A prompt is applied on each pre-processed DITA XML file and a final report is generated.

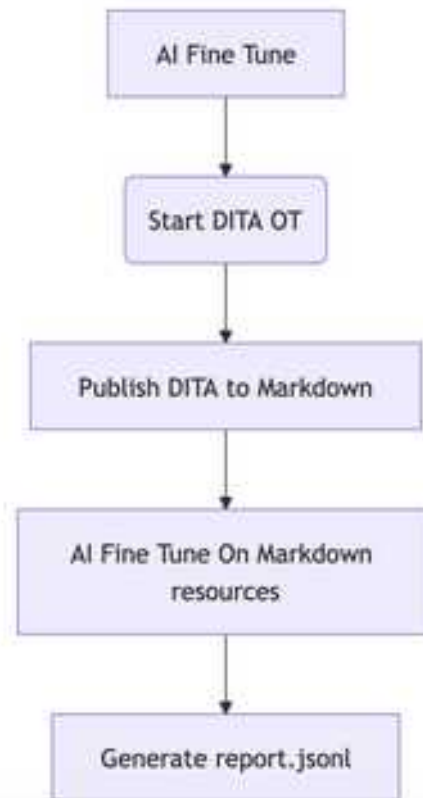


AI Checks Example – Grammar and Spell Check

- file:/Users/raducoravu/eclipse-workspace/eXml/samples/dita/flowers/topics/flowers/lilac.dita:
 - errors:
 - line: 29
column: 94
error: use of 'fallacy' in context may require a clearer explanation or could be a misuse of the term given the context.
 - errors:
 - line: 43
column: 53
error: Typographical error: 'encurved' should be 'curved' or revised for clarity, as 'encurved' is not a standard English term.
- file:/Users/raducoravu/eclipse-workspace/eXml/samples/dita/flowers/topics/care.dita:
 - errors:
 - type: spelling
message: Misspelled word: Preparaton
 - suggestions:

(2) Produce Fine-tune Dataset from DITA Style Guide

- DITA OT Plugin which publishes DITA to Markdown and then creates of question/answer pairs with AI generated examples based on the Markdown content.
- The question/answer pairs are grouped inside a JSONL file which can be used to fine tune an OpenAI model.



Fine-tune

Use of Specific Words (Style Rules)

The following *Style Rules* provide guidelines for the use of particular words:

» » "Authors" vs. » » "Writers"

» » **Rule:** Simply as a matter of company preference, use » » "content authors" instead of » » "technical writers".

» » "Combo Box" vs. » » "Combobox"

» » **Rule:** Use » » "combo box" instead of » » "combobox".



```
JSON {
  messages:
    role: system
    content: You are a technical writing expert who applies the style guide rules.
  messages:
    role: user
    content: Should I use 'content authors' or 'technical writers'?
  messages:
    role: assistant
    content: You should use 'content authors'.
    Examples:
    Invalid usage: Technical writers create the documentation.
    Valid usage: Content authors create the documentation.
}
```

Fine-tune

Chat

[Clear](#)[History](#)

gpt-4o-2024-08-06

ft:gpt-4o-2024-08-06:syncro-soft-development:oxygen-style-guide-new:AvdyPTd

System message Answer the questions below based on the style g...

System message Answer the questions below based on the style g...

User

Should I use 'content authors' or 'technical writers'?

User

Should I use 'content authors' or 'technical writers'?

Assistant

The choice between 'content authors' and 'technical writers' depends on the specific role and context you are referring to. Use 'content authors' when discussing individuals responsible for creating general content, such as blog posts, articles, or marketing ma 3,351ms 33t 87t Request ID erring to

Assistant

You should use 'content authors' instead of 'technical writers'.

(0)

2,262ms 33t 14t Request ID

Fine-tune

Chat

[Clear](#)[History](#)

gpt-4o-2024-08-06

User

<shortdesc>With just a little bit of care and preparation, any flower garden can be a vibrantly colored environment. Flowers can be selected for specific blooming seasons, colors and shapes.</shortdesc>

Assistant

With some thoughtful care and preparation, any flower garden can become a vibrant oasis of color. Choose flowers based on their blooming seasons, colors, and shapes for a stunning display all year round.

2,628ms ↑ 61t ↓ 39t Request ID

ft:gpt-4o-2024-08-06:syncro-soft-development:oxygen-style-guide-new:AvdyPTdl

User

<shortdesc>With just a little bit of care and preparation, any flower garden can be a vibrantly colored environment. Flowers can be selected for specific blooming seasons, colors and shapes.</shortdesc>

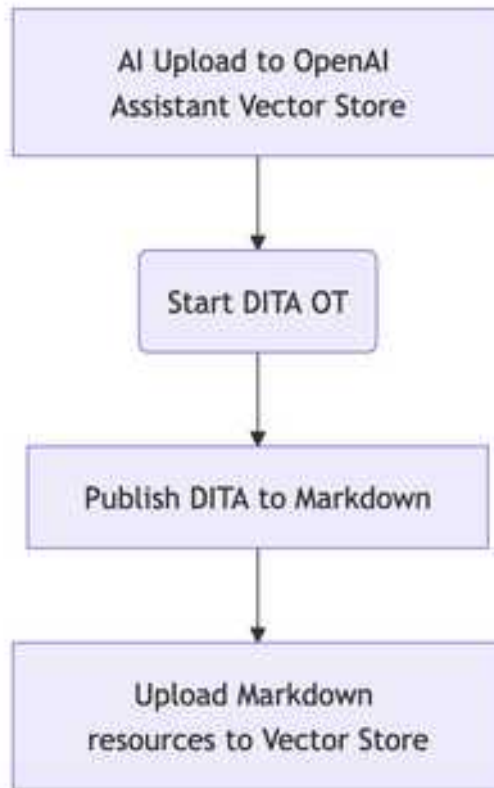
Assistant

<shortdesc>Transform a flower garden into a vibrant oasis with thoughtful plant selection based on blooming seasons, colors, and shapes.</shortdesc>

1,757ms ↑ 61t ↓ 30t Request ID

(3) Upload Data to OpenAI Assistant Vector Store

- DITA OT Plugin which publishes DITA to Markdown a then uploads the content to an AI Assistant Vector Store.



Vector Store Example – Assistant based on User’s Guide

The screenshot displays the Oxygen AI Assistant interface. On the left, the 'Assistants' panel shows the selected assistant 'Oxygen User's Guide Assistant' with a dropdown arrow. Below this, the 'Model' is set to 'gpt-4o'. The 'TOOLS' section includes 'File search' (enabled) with a '+ Files' button, and 'Code interpreter' (disabled) with a '+ Files' button. The 'FUNCTIONS' section has a '+ Functions' button. At the bottom, 'MODEL CONFIGURATION' is visible with a 'Clone' button and an 'Updated 1/29, 8:31 AM' timestamp.

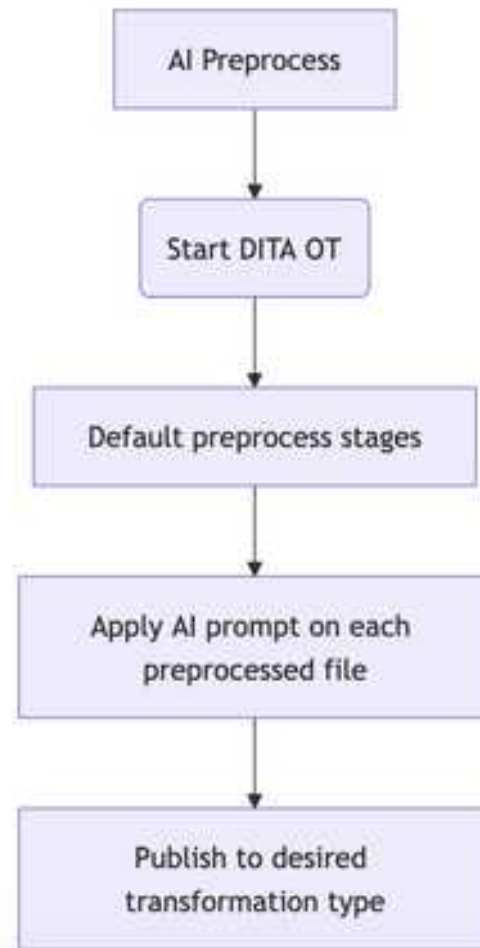
The main chat area shows a 'THREAD' with ID 'thread_5wucMhzC6cMj2Vxe8Dq8z14b' and '15058 tokens'. The user's message is 'How do I extend a framework?'. The assistant's response is: 'To extend a framework, you can follow these general steps, which involve the use of tools and settings available in the Oxygen XML Editor:'. The response includes a numbered list:

- 1 **Preferences Configuration:** Open the **Preferences** dialog in Oxygen XML Editor and navigate to the **Document Type Association** settings. In the **Locations** tab, add the directory path where your custom framework will reside [1] .
- 2 **Select Existing Framework:** On the **Document Type Association** page...

At the bottom of the chat area, there is an input field 'Enter your message...' with a '+ Run' button and a 'x ++' icon. A footer note states: 'Plain text messages can be viewed by anyone at your organization using the API'.

(4) AI Preprocess

- DITA OT Plugin which applies modifications using the pre-processing stage.



AI Preprocess Example – Translate at Publishing Time

Setting up a conference call

Short Description: You can easily set up a conference call with your family, friends, co-workers. Call the first participant and then join calls together as described in the following procedure.

Metadata

Q: How to make a conference call?

About this task:

Before starting, make sure that your SIM card⁴ is enabled for conference calls service.



Configuration d'une conférence téléphonique

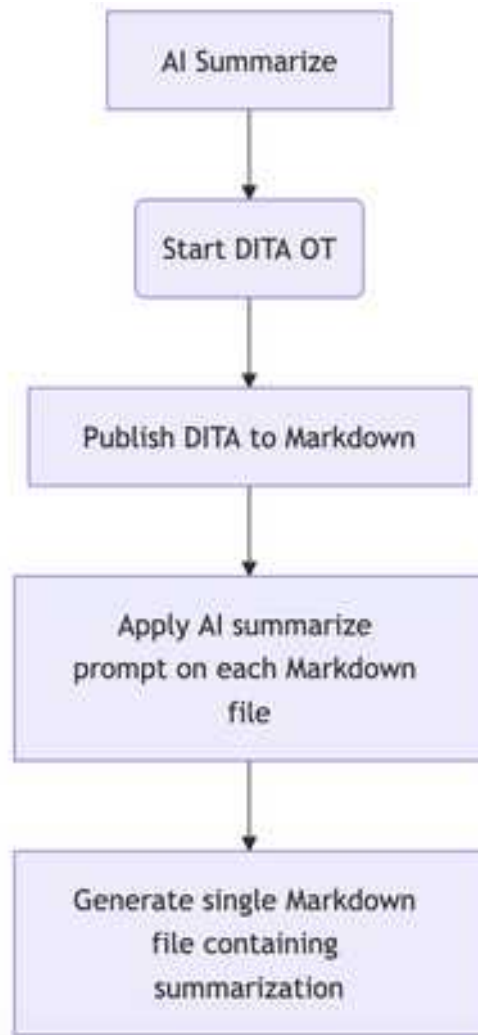
Vous pouvez facilement mettre en place une conférence téléphonique avec votre famille, vos amis, vos collègues. Appelez le premier participant et puis rejoignez la procédure suivante.

Avant de commencer, assurez-vous que votre carte SIM est activée pour le service de conférence téléphonique.

1. Appelez le premier participant. Voir [Passer un appel](#)
2. Pendant l'appel, appuyez sur la touche pour ajouter un participant. De cette manière, le premier participant est mis en attente.
3. Saisissez le numéro de téléphone du participant ou recherchez-le dans les contacts.
4. Appuyez sur la touche .
5. Pour également ajouter le premier contact à la conférence téléphonique, appuyez à nouveau sur la touche .
6. Pour ajouter d'autres contacts à la conférence, répétez les étapes 2 à 5.

(5) AI Summarize

- DITA OT Plugin which generates a summary of the entire publication, in the order specified in the DITA Map.



AI Summarize - Examples

- Suggest additional links to add between topics.
- Suggest place in DITA Map where to add a new DITA topic.
- Suggest ways to re-organize the user manual's DITA Map.
- Generate a graph showing relations between topics.

THANK YOU!

**Questions
time...**

Radu Coravu

radu_coravu@oxygenxml.com

