



We shall call it...
preprocess... 2!

Jarno Elovirta @ DITA-OT Day 2017



Vanity slide

I write programs

Started working on DITA-OT internal years ago

Consultant at WunderDog



What is a DITA map

```
<map>  
  <title>Title</title>  
  <topicref href="intro.dita"/>  
  <topicref href="reference.dita"/>  
  <topicref href="tutorial.dita"/>  
</map>
```



What is a DITA map

```
<map>  
  <title>Title</title>  
  <topicref href="intro.dita"/>  
  <topicref href="reference.dita"/>  
  <topicref href="tutorial.dita"/>  
</map>
```

```
<topic>  
  <title>Title</title>  
  <related-links>  
    <link href="intro.dita"/>  
    <link href="reference.dita"/>  
    <link href="tutorial.dita"/>  
  </related-links>  
</map>
```



What is a DITA map

```
<mapref href="keys.ditamap"/>
<topicref href="tutorial.dita">
  <ditavalref href="novice.ditaval">
    <ditavalmeta>
      <dvrKeyscopePrefix>novice-</dvrKeyscopePrefix>
    </ditavalmeta></ditavalref>
  <ditavalref href="novice.ditaval">
    <ditavalmeta>
      <dvrKeyscopePrefix>novice-</dvrKeyscopePrefix>
    </ditavalmeta>
```



"This is my map.
There are many like it, but this one is mine.
Without me, my map is useless.
Without my map, I am useless."

Fictional topic



DITA map is a processing context

Maps change how topics are processed, but not vice versa

Topics require the map context to be processed—unless topics are self-contained

Without map context topics are left partially processed, exposing DITA constructs

Map is also a hierarchy structure for topic relations



Original preprocessing

Read and process both maps and topics at the same time

Original implementation approach from DITA-OT 1.0 and IBM

DITA 1.3 requires many and more complex steps than earlier versions

Interaction between steps and their order creates a fragile system




Map-first preprocessing

Construct context first by processing maps before topics

Processes topics with complete context available as a second step

Makes processing easier to reason about

Enables optimizations, cleaner code, and makes DITA 2.0 easier to implement



"for each desired change,
make the change easy (warning: this may be hard),
then make the easy change"

@KentBeck



Map-first processing now and next

DITA-OT 3.0 has taken preprocess2 into use in PDF and HTML Help

Rest of the built-in transtypes will follow in future releases

- preprocess,

+ preprocess2,

Original preprocess will become a legacy feature



Future work

Make preprocessing more robust

Add extension points only when actually needed

Optimize for performance where possible

Try and do something about... chunking 🤖



Ask me more

jelovirt@gmail.com

[@jelovirt](#)

[slack.dita-ot.org](#)

[wunder.dog](#)