How HERE Extends the DITA-Open Toolkit

Jason Fox

Disclaimer

The views presented in this workshop are my own, and do not reflect the views of HERE.

HERE in numbers

200
Countries mapped

7,000+
Employees in 56 countries focused on delivering the world's best map and location technologies

HERE Maps on board of
100M
vehicles and counting

28
TB map data
collected per day

30
Years of experience transforming location technology

4 of 5
In-car navigation systems in Europe and North America use HERE maps

700,000
3D data points per second per car

400
HERE cars collecting data for our maps

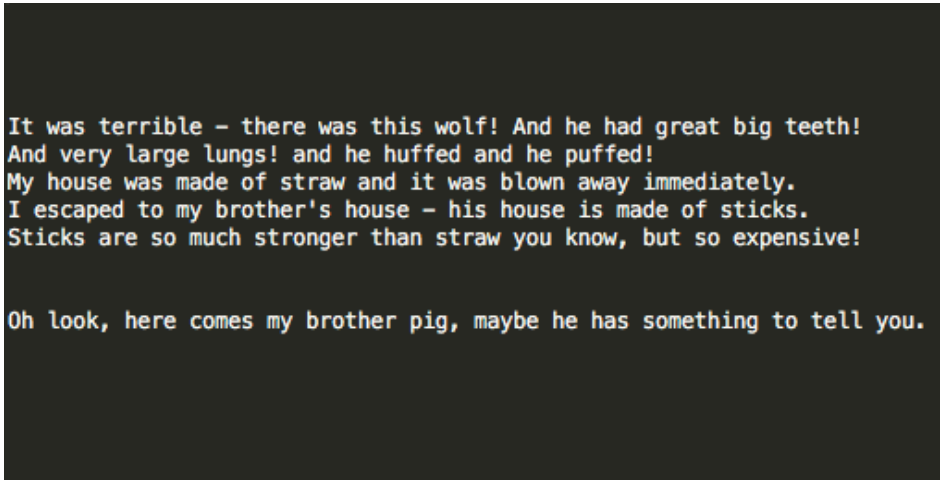
Software Development

- Thousands of engineers
- Hundreds of products
- Thousands of documents (including different versions)

Technical Writing

- 7 writers specializing on developer documentation
- Development team focusing on required tooling

Anti-pattern – Secretarial Role



- Unstructured - requires multiple review cycles
- Wrong tone - difficult to extract usable information.
- Output – not directly usable
  - Content needs further transformation into a usable format.
  - Content is not reusable
- Doesn't scale

Product Engineer – Blogger Role

```
<topic id="two-little-pigs">
   <title>The story of the two little pigs</title>
   <body>
        <p>Once upon a time there were two little pigs</p>
        <ul>
              <li><xref href="pig-1"/></li>
              <li><xref href="pig-2"/></li>
        </ul>
   </body>
   <topic id="pig-1">
        <title>First Little Pig</title>
        <body>
              <p>
                The first little pig made his house with straw
              </p>
        </body>
   </topic>
   <topic id="pig-2">
         <title>Second Little Pig</title>
         <body>
              <p>
                The second little pig made his house with sticks
              </p>
         </body>
   </topic>
</topic>
```

- Fully knowledgeable in subject matter, but limited knowledge of how to write documentation.
- Output - Partially structured -based on template
- Tone of voice - potentially incorrect based on:
  - training
  - templating
  - previous experience
- Scalable

Technical Writer - Editor Role

```
<topic id="building-materials">
   <title>Porcine Building Materials - Data reference</title>
   <body>
         <p>
            The following building materials are currently available:
         </p>
         <ul>
             <li><xref href="pig-1"/></li>
             <li><xref href="pig-2"/></li>
         </ul>
         <section id="pig-1">
              <title>Straw</title>
              <p>
                 Straw is a cheap and relatively abundant building
                 material that offers little to no resistance to
                 high-presssure wolf attack.
              </p>
         </section>
         <section id="pig-2">
               <title>Sticks</title>
               <p>
                  Sticks are a more expensive building material that offers
                  limited resistance to high-presssure wolf attack. This
                  material should only be used in areas of known wolf
                  infestation.
               </p>
         </section>
   </body>
</topic>
```

- Expert in the documentation field - limited resource.
- Fully knowledgeable in writing documentation – sufficient knowledge in subject matter
- Output - Corrected structure
- Corrected tone of voice

Documentation Engineer - Hot Metal

```
<!DOCTYPE topic PUBLIC "-//OASIS//DTD DITA Topic//EN" "topic.dtd">
<topic id="building-materials">
    <title>**************************</title>
    <body>
        <p>
            ******************
        </p>
        <ul>
            <li><xref href="pig-1"/></li>
            <li><xref href="pig-2"/></li>
        </ul>
        <section id="pig-1">
            <title>****</title>
            <p>
                ************************************************
                ************************************************
            </p>
        </section>
        <section id="pig-2">
            <title>*****</title>
            <p>
                ************************************************
                ************************************************
                ************************************************
            </p>
        </section>
    </body>
</topic>
```

- Expert in publication - limited resource.
- Considers the source as structured data only - text is irrelevant
- Knowledge of subject matter or document structure is irrelevant

Existing DITA-OT Error Handling

DITA-OT Errors

- Generated within DITA-OT code
- Verbose output only
- Error codes assume prior DITA-OT knowledge

```
[gen-list] [DOTJ031I][INFO] No specified rule for 'otherprops=generated'
was found in the ditaval file. This value will use the default action, or
a parent prop action if specified. To remove this message, you can specify
a rule for 'otherprops=generated' in the ditaval file.

 [gen-list] [DOTX008E][ERROR] File 'file/../lorem.dita' does not exist or
cannot be loaded.

[gen-list] Failed to read DITAVAL file: /···/ditavals/ditaval.dtd (No such
file or directory)

[gen-list] [DOTX008W][WARN] File 'file:/···/graphics/lorem.jpg' cannot be
loaded, and no navigation title is specified for the table of contents.
```

FOP Processing

- Java Stack trace for unhandled FOP errors
- Error occurs on topic.fo
- No context for End User

```
[fop-error] Anttask - Error rendering fo file: /···/topic.fo
<org.apache.fop.apps.FOPException: The current document is unable to
create an element of the requested type (namespace:
http://www.w3.org/2000/svg, name: flowRoot).
[fop] org.w3c.dom.DOMException: The current document is unable to create
an element of the requested type (namespace: http://www.w3.org/2000/svg,
name: flowRoot).>org.apache.fop.apps.FOPException: The current document is
unable to create an element of the requested type (namespace:
http://www.w3.org/2000/svg, name: flowRoot).
[fop] org.w3c.dom.DOMException: The current document is unable to create
an element of the requested type (namespace: http://www.w3.org/2000/svg,
name: flowRoot).
[fop]  at
org.apache.fop.tools.anttasks.FOPTaskStarter.renderInputHandler(Fop.java:6
```

XSLT Processing

- XSL Transform engine error handling
- Errors traced to XSLT file
- No context for End User

```
/plugins/org.dita.pdf2/build.xml:276: Fatal error during transformation
using /../dita_ot/plugins/org.dita.pdf2.fop/xsl/fo/topic2fo_shell_fop.xsl:
An empty sequence is not allowed as the value of variable $entryNumber;
SystemID: file:/···/dita_ot/plugins/org.dita.pdf2/xsl/fo/tables.xsl; Line#:
431; Column#: -1
```

HERE Validator Plug-in Error Handling

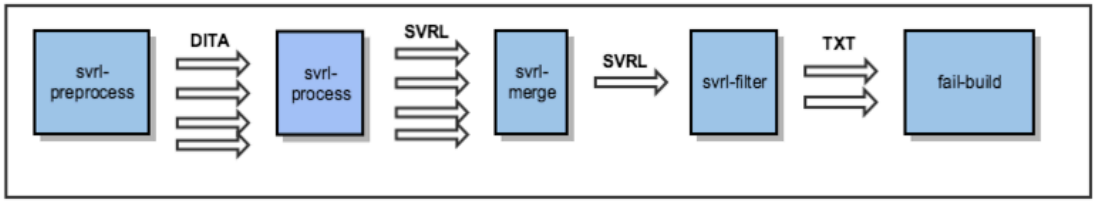
```
[validate] [FATAL]  [/topics/filename.dita]
[validate]   Line 33: colspec - [colspec-colnum-not-unique]
[validate] The key/value pair colnum="1" is not unique. Make sure there are no duplicates.
[validate] [WARN]   [/topics/filename.dita]
[validate]   Line 15: section[id="bad"] - [comment-fixme]
[validate] Found 'FIXME' comments within the <section> element - fix as requested and delete the comment.
[validate]
[validate]  FIXME: This comment requires action
[validate] [WARN]   [/topics/filename.dita]
[validate]   Line 15: section[id="bad"] - [comment-todo]
[validate] Found 'TODO' within the following text of the <section> element - fix as requested and delete the comment.
[validate]
[validate]  TODO: We need to delete this section prior to release
[validate] [FATAL]  [/topics/filename.dita]
[validate]   Line 16: p[conref="./unknown-file.dita#href"] - [conref-external-file-not-found]
[validate] The linked file does not exist. Check the related folder for the missing file and make sure the file name is correct.
[validate] [FATAL]  [/topics/filename.dita]
[validate]   Line 17: p[conref="./conref-internal-id-not-found.dita#bad"] - [conref-external-id-mismatch]
[validate] The conref source <> and the destination <p> are not the same type. Check to make sure the referenced
[validate] [ERROR]  [/topics/href-not-lower-case-image.dita]
[validate]   Line 15: image - [href-not-lower-case]
[validate] The value provided in href="../graphics/SAMPLE-JPEG.jpg" is invalid, allowed characters are: lowercase, a-z only, words separated by hyphens.
[validate] [ERROR]  [/topics/fig-id-invalid.dita]
[validate]   Line 20: fig[id="bad"] - [fig-id-invalid]
[validate] ID values must start with 'fig-' in all <fig> elements.
```

DEMONSTRATION

- DITA VALIDATION
- SPELL CHECKING
- AUTO-CORRECTION
- FIXING ATTRIBUTES

Deep Dive: HERE Validator Plug-in Workflow

```
../dita_ot/bin/dita  -input ../document.ditamap -format svrl-echo
```




```
[validate] [WARN]   [/topics/filename.dita]
[validate]   Line 15: section[id="bad"] - [comment-fixme]
[validate] Found 'FIXME' comments within the <section> element - fix as requested and delete the
comment.
```

1. **Preprocess** - setup defaults and run sanity checks
2. **Process** - Run ruleset over every DITA file
3. **Merge** - Consolidate result to one file
4. **Filter** - Process results to readable text
5. **Fail-build** - Fail if errors found

Deep Dive: Preprocess

- Set default root based on **`args.input`**
- Use default **`cfg`** rulesets if not overridden
- Use default **`Customization`** XSL files if not overridden
- Set defaults for command line parameters
  - `args.validate.ignore.rules`
  - `args.validate.check.case`
  - `args.validate.blacklist`
  - `args.validate.mode`
- Windows preprocessing
- DTD validation

Deep Dive: Process in Detail

- Read current ruleset (**FATAL, ERROR, WARNING**)
  - **`ruleset2regex.xsl`**
- Process XSLT on **`*.ditamap`** files (XML -> SVRL)
  - **`ditamap2svrl.xsl`**
  - **`ditamap-rules.xsl`**
  - **`schematron.xsl`**
- Process XSLT on **`*.dita`** files (XML -> SVRL)
  - **`dita2svrl.xsl`**
  - **`schematron.xsl`**
  - Common Attribute Rules: **`common-rules.xsl`**
  - Rules for Text Elements: **`textual-rules.xsl`**
  - Rules for FOP: **`fop-rules.xsl`**
  - Rules for Specific DITA Elements: **`codeblock.xsl, figure.xsl, section,xsl etc.`**

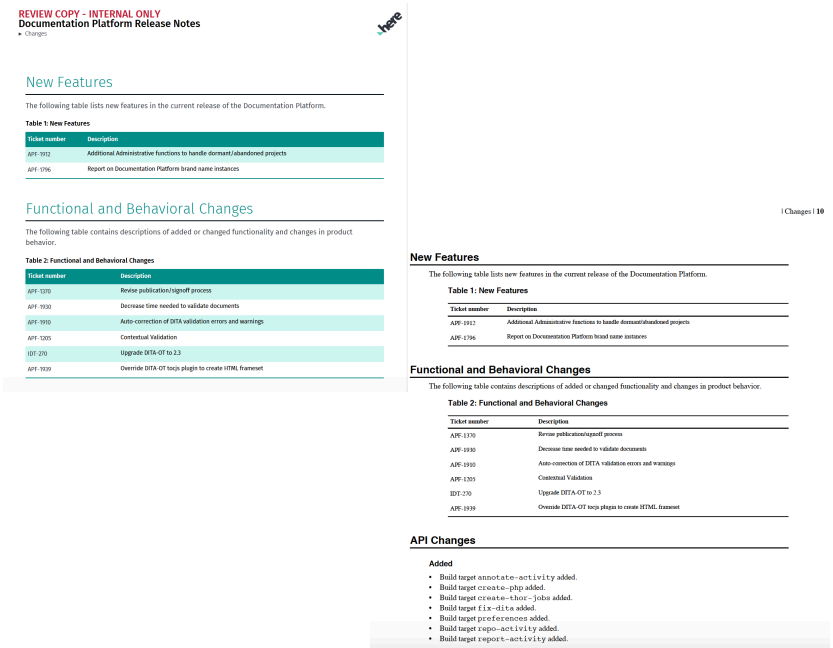
Deep Dive: Merge, Filter and Result

- Read all files and merge to single SVRL
  - **`schematron-merge.xsl`**

- Read merged SVRL and transform to text (colorize is post-processing)
  - **`schematron-output.xsl`**

- Read merged SVRL and output count of Errors
  - **`schematron-result.xsl`**

Customizing the base ruleset



REVIEW COPY - INTERNAL ONLY
Documentation Platform Release Notes
▸ Changes

New Features

The following table lists new features in the current release of the Documentation Platform.

Table 1: New Features

Ticket number	Description
APF-1912	Additional Administrative functions to handle dormant/abandoned projects
APF-1796	Report on Documentation Platform brand name instances

Functional and Behavioral Changes

The following table contains descriptions of added or changed functionality and changes in product behavior.

Table 2: Functional and Behavioral Changes

Ticket number	Description
APF-1370	Revise publication/signoff process
APF-1930	Decrease time needed to validate documents
APF-1910	Auto-correction of DITA validation errors and warnings
APF-1205	Contextual Validation
IDT-270	Upgrade DITA-OT to 2.3
APF-1939	Override DITA-OT tocjs plugin to create HTML frameset

New Features

The following table lists new features in the current release of the Documentation Platform.

Table 1: New Features

Ticket number	Description
APF-1912	Additional Administrative functions to handle dormant/abandoned projects
APF-1796	Report on Documentation Platform brand name instances

Functional and Behavioral Changes

The following table contains descriptions of added or changed functionality and changes in product behavior.

Table 2: Functional and Behavioral Changes

Ticket number	Description
APF-1370	Revise publication/signoff process
APF-1930	Decrease time needed to validate documents
APF-1910	Auto-correction of DITA validation errors and warnings
APF-1205	Contextual Validation
IDT-270	Upgrade DITA-OT to 2.3
APF-1939	Override DITA-OT tocjs plugin to create HTML frameset

API Changes

Added

- Build target `annotate-activity` added.
- Build target `create-php` added.
- Build target `create-thor-jobs` added.
- Build target `fix-dita` added.
- Build target `preferences` added.
- Build target `repo-activity` added.
- Build target `report-activity` added.

- Validator plugin structure is based on the PDF plugin and is customizable in the same fashion
  - **`cfg`** – Changing severity and adding language support
  - **`Customization`** – Adding rules
- The default ruleset demonstrates a boilerplate use case. Unlikely to meet your own requirements
- Customize the rules by creating your own plugin extension
- https://github.com/heremaps/dita-ot-plugins for examples

Ignoring rules – Part One

- Command line invocation
  - **`-Dargs.validate.ignore.rules=table-id-missing`**

- Add code linter style comments
  - `<!– ignore-rule:rule-id -->` - For a specific instance
  - `<!– ignore-all-warnings -->` - For all warnings within a file
  - `<!– ignore-all-errors -->` - For all errors within a file

- Alter rule severity
  - **`cfg/rulesset/default.xml`**

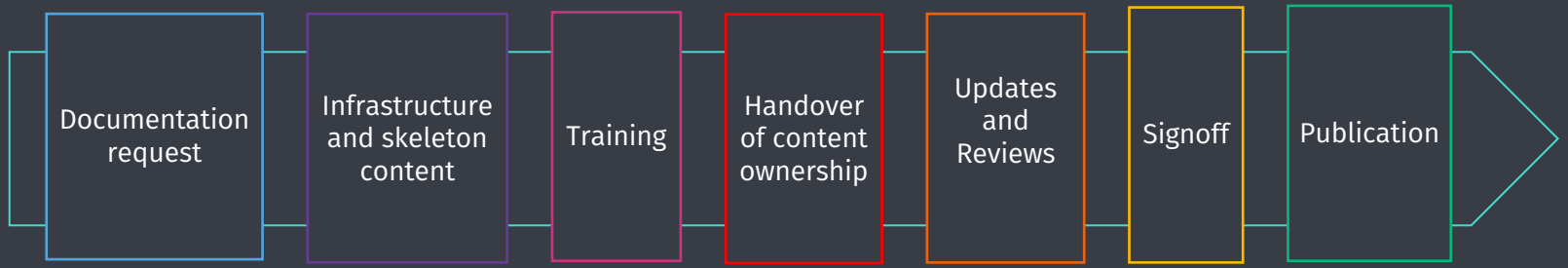
Ignoring rules – Part Two

- Plugin XSLT customization
  - Override using **`xsl:template match`**

```
<!--
    Ignore the common Style Rules for id attributes  when the topic is generated
-->
<xsl:template match="topic[@otherprops = 'generated']" mode="common-id-checks"/>

<!--
    Ignore the Common Style Rules for href attributes for generated topics
    and links to generated topics
-->
<xsl:template match="*[ancestor::topic[@otherprops = 'generated']]" mode="common-href-checks" />
<xsl:template match="*[contains(@href, 'topics_api')]" mode="common-href-checks"/>
```

The Documentation Process at HERE



Thank you

Contact

Jason Fox

HERE Technologies
Invalidenstraße 116, 10115
Berlin

jason.fox@here.com

https://github.com/heremaps/dita-ot-plugins

here