

Oxygen Script Validation

Octavian Nadolu, Syncro Soft

octavian.nadolu@oxygenxml.com

@OctavianNadolu

© 2023 Syncro Soft SRL. All rights reserved.

Agenda

- Overview of the Oxygen Validation Support
- Oxygen Scripting Tools
- Validation Script Command Line Script
- Reports Using Command Line Scripts
- Examples of Validation
- Transformation Command Line Script
- Comparison/Merge Command Line Script



Validation

Validation is the process of checking a document to confirm that it is both well-formed and also "valid" in that it follows a defined structure

- A well-formed document follows the basic syntactic rules
- A valid document also respects the rules dictated by a particular schema



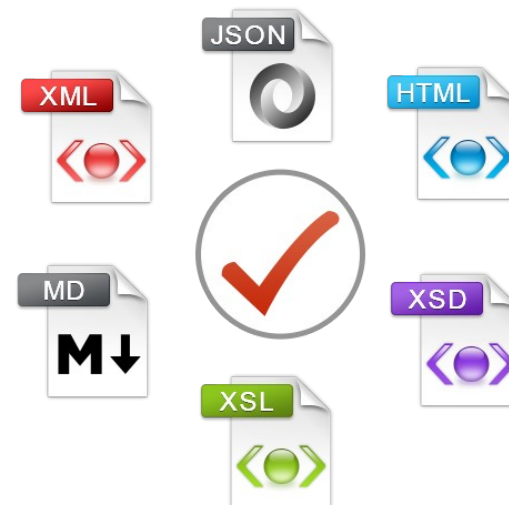
Validation Support in Oxygen

- Well-Formlessness Check
- Validate Documents Against a Schema
- Validation Scenarios
- Module Files Validation
- Batch Validation
- Validation As-You-Type



Validation File Type Support

- XML - validate with XSD, DTD, RelaxNG, NVDL, Schematron
 - Builtin schemas (Dita, DocBook, Tei ..)
- JSON/YAML - validate with JSON Schema, Schematron
 - Builtin schemas (OpenAPI, AsyncAPI, JSON-LD)
- XSD, DTD, RelaxNG, NVDL, Schematron, JSON Schema
- XSLT, XQuery, XProc
- HTML, XHTML, CSS
- Markdown
- WSDL



Validation Engines

- XML - Saxon, LibXML, Xerces
- JSON/YAML - Everit
- HTML, XHTML, CSS – W3C Nu Validator
- XSLT, XQuery – Saxon, MSXML, Xalan
- XSD - Saxon, Xerces
- DTD - Xerces
- RelaxNG, NVDL – Jing
- Custom Validator



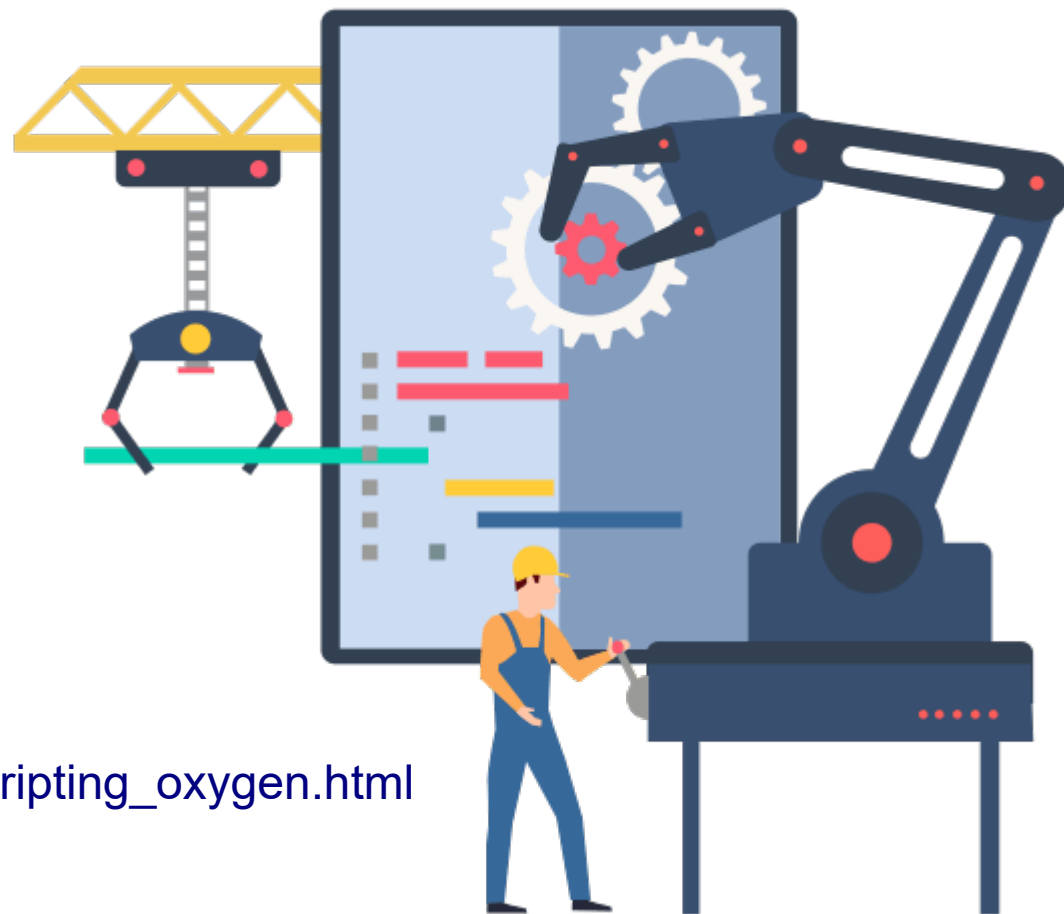
Question: What type of validation do you use or intend to use?

- XML validation with DTD, XSD, RNG, NVDL
- JSON/YAML validation with JSON Schema
- Validation with Schematron
- XSD, DTD, RelaxNG, NVDL, Schematron, JSON Schema
- Other (use the questions panel)



Oxygen Scripting

Automate and Run Oxygen Utilities
from the Command-Line Interface



oxygenxml.com/oxygen_scripting.html

oxygenxml.com/doc/ug-editor/topics/scripting_oxygen.html

Oxygen XML Scripting Functions

Automate and run Oxygen utilities from the command-line Interface

- **Validation***
- Transformation*
- Compare and Merge Files or Directories*
- Conversion
- Generate Documentation
- Format and Indent
- and more

Webinar: Automate XML processing with Oxygen XML Scripting

Webinar: Oxygen Compare and Merge Scripts

Question: What Oxygen Scripting tools do you use or intend to use?

- Validation
- Transformation
- Compare and Merge
- Generate Documentation
- Other (use the questions panel)



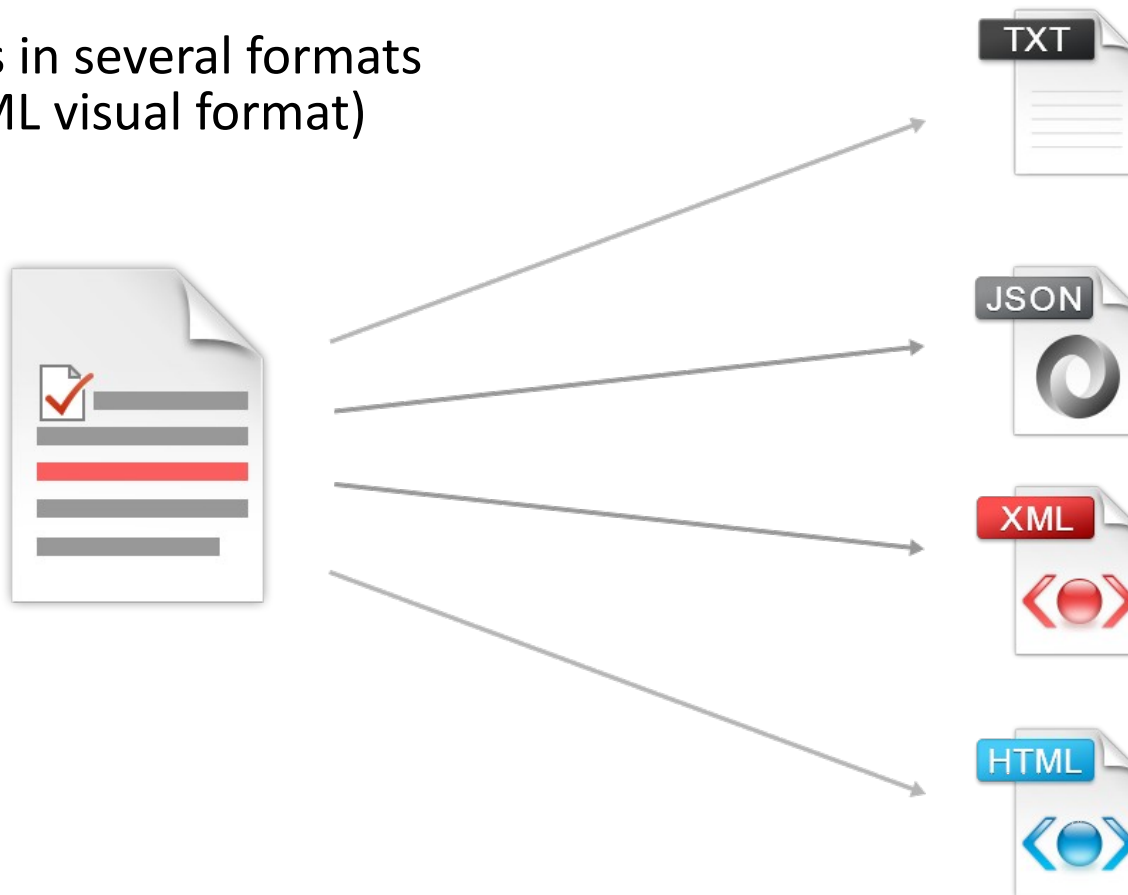
Validation Script

- Validate files or directories from a command line interface or from an integration server (CI/CD)



Validate Script Report

Generate validation reports in several formats
(Text, JSON, XML, or an HTML visual format)



Validation Script Options

- You can specify a **certain schema file** to validate against.
- You can validate using the **schema(s)** that are **internally associated** within files.
- You can validate using the **default** Oxygen **validation scenarios** for respective file types.
- You can **specify a validation scenario** name to be referenced from either a scenarios file or a project file.



Arguments for the Validate Script

```
sh scripts/validate.sh fileOrDirPath [-s schemaFilePath | -sn scenarioName]
[-sf scenariosFilePath] [-if includeFilesFilter] [-ef excludeFilesFilter]
[-ed excludeSubdirsFilter] [-rf reportFile] [-rft reportFormat] [-v] [-help
|--help | -h | --h]
```

fileOrDirPath - Mandatory argument that specifies the path of the file or directory to validate

-s schemaFilePath - Optional argument that specifies the file path of the schema to validate against

-sn scenarioName - Optional argument that specifies the name of the validation scenario to be applied

-sf scenariosFilePath - Optional argument that specifies the path of the file that stores the validation scenarios (either an Oxygen scenarios file or an Oxygen project file)

-rf reportFile - Specifies the path for the report file to save the validation results

-rft reportFormat - Specifies the format of the validation report. Possible values: text, xml, json, html

Arguments for the Validate Script

- if includeFilesFilter** - Use this argument to only validate the files that match the specified pattern (e.g. .xml,.json)
- ef excludeFilesFilter** - Excludes the files that match the specified pattern (e.g. test.wsdl,draft.xsl)
- ed excludeSubdirsFilter** - Excludes the sub-directories that match the specified pattern (e.g. .svn,_svn,.git)
- v** - Prints additional information to the console (Verbose mode)
- help | --help | -h | --h** - Displays help text

Examples of the Validate Script

Validate a File by Applying a Custom Validation Scenario

```
sh validate.sh xmlFile.xml -sn xmlFileScenario -sf xmlFile.scenarios
```

This command implies validating `xmlFile.xml` by applying the validation scenario named `xmlFileScenario`, described in the `xmlFile.scenarios` file. If you want to apply more than one validation scenario, you can use the `-sn scenarioName` construct multiple times.

Validation Report

- Validation report for one file, displayed in text format

```
--- SUMMARY ---  
Report created on: 1/6/2023 17:37:17  
1 file verified in total  
1 file reported with validation problems  
8 problems found on aggregate  
  
--- FILES WITH VALIDATION PROBLEMS ---  
FILE: D:\****\Samples\xml\schematron\xmlFile.xml (8 problems)  
  
System ID: D:\****\Samples\xml\schematron\xmlFile.xml  
Main validation file: D:\****\Samples\xml\schematron\xmlFile.xml  
Scenario name: xmlFileScenario  
Schema: D:\****\Samples\xml\schematron\rules.sch  
Engine name: ISO Schematron  
Severity: warning  
Problem ID: rules.sch:boldID  
Description: Bold element is not allowed in title.  
Start location: 4:21
```

Examples of the Validate Script

Validate a Directory by Applying an Oxygen Default Validation Scenario

```
sh validate.sh ditaFolder -sn DITA -rft xml
```

A scenario name is provided, but without specifying a scenarios file. This command implies validating all files from ditaFolder by applying the Oxygen default validation scenario named DITA (in accordance with the association made in the Document Type Configuration Dialog Box).

Validation Report

- Validation report for directory, displayed in XML format

```

<report>
  <summary>
    <totalNumberOfFilesVerified>29</totalNumberOfFilesVerified>
    <numberOfFilesWithValidationProblems>2</numberOfFilesWithValidationProblems>
    <problemsFoundOnAggregate>3</problemsFoundOnAggregate>
  </summary>
  <filesWithValidationProblems><file filePath="D:\***\Samples\ dita \taskbook.ditamap" numberOfProblems="1">
    <incidents>
      <incident>
        <engine>DITA Validation</engine>
        <severity>error</severity>
        <problemID>dmv.file.not.found</problemID>
        <description>File not found: D:\***\Samples\ dita \tasks\unscrewcove.dita.</description>
        <systemID>D:\***\Samples\ dita \taskbook.ditamap</systemID>
        <operationDescription>
          <mainValidationFile>D:\***\Samples\ dita \taskbook.ditamap</mainValidationFile>
          <scenario>DITA</scenario>
          <documentType>DITA 2.x</documentType>
        </operationDescription>
        <type>Reference</type>
        <location>
          <start>
            <line>66</line>
            <column>28</column>
          </start>

```

...

Examples of the Validate Script

Validate a File by Applying Associated Scenarios Stored in an Oxygen Project File

```
sh validate.sh json -sf validateScript.xpr -rft json
```

This command line implies that if validation scenario associations for the files are found in `validateScript.xpr`, then those scenarios are identified and applied. Otherwise, the validation first considers the schema associations declared in files (if any), or default Oxygen validation scenarios are applied in accordance with the type of the file to validate.

Validation Report

- Validation report for directory, displayed in JSON format

```
{
  "report": {
    "summary": {
      "totalNumberOfFilesVerified": 9,
      "numberOfFilesWithValidationProblems": 4,
      "problemsFoundOnAggregate": 4
    },
    "filesWithValidationProblems": [
      {
        "filePath": "file:/D:/***/Samples/json/JSON-LD/person.json",
        "numberOfProblems": 1,
        "incidents": [
          {
            "engine": "JSON Validator",
            "severity": "error",
            "description": "Missing value",
            "systemID": "file:/D:/***/Samples/json/JSON-LD/person.json",
            "mainValidationFile": "file:/D:/***/Samples/json/JSON-LD/person.json",
            "location": {
              "start": {
                "line": 4,
                "column": 12
              }
            }
          }
        ]
      }
    ]
  }
}
```

...

Examples of the Validate Script

Directory Default Validation and Custom Formatted Report Saved to a Specific Location

```
sh validate.sh xmlFolder -rft html -rf validation_rep.html
```

No validation scenario name, no scenario file, and no schema provided. This command line involves validating all files from the xmlFolder. Each file is validated against the schema(s) internally associated (if any). Otherwise, the default Oxygen validation scenarios for the respective file type are applied. Also, the validation report is formatted in HTML and is saved to the validation_rep.html file.

HTML Report for Validation

SUMMARY

16 files verified in total

7 files reported with validation problems

11 problems found on aggregate

FILES WITH VALIDATION PROBLEMS

D:***\Samples\xml\dtd\personal.xml (2 problems)

| | |
|-----------------------------|---|
| Description | Element type "email" must be declared. |
| Severity | Error |
| Problem ID | MSG_ELEMENT_NOT_DECLARED |
| Engine name | Xerces |
| System ID | D:***\Samples\xml\dtd\personal.xml |
| Main validation file | D:***\Samples\xml\dtd\personal.xml |
| Schema | D:***\Samples\xml\dtd\personal.dtd |
| Start location | line: 19, column: 10 |
| End location | line: 19, column: 16 |
| Description | Unexpected element "email". The content of the parent element type must match "(name,email*,link?,url*)". |
| Severity | Error |
| Problem ID | MSG_CONTENT_INVALID_UNEXPECTED_ELEMENT |
| Engine name | Xerces |
| System ID | D:***\Samples\xml\dtd\personal.xml |
| Main validation file | D:***\Samples\xml\dtd\personal.xml |
| Schema | D:***\Samples\xml\dtd\personal.dtd |
| Start location | line: 19, column: 10 |
| End location | line: 19, column: 16 |

Validation Script GitHub Template

- Automatically validate files on each commit
- Generate report and publish it on [GitHub pages](#)
- Run validation on specific folder from project

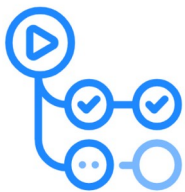


github.com/oxygenxml/oxygen-script-validation-template



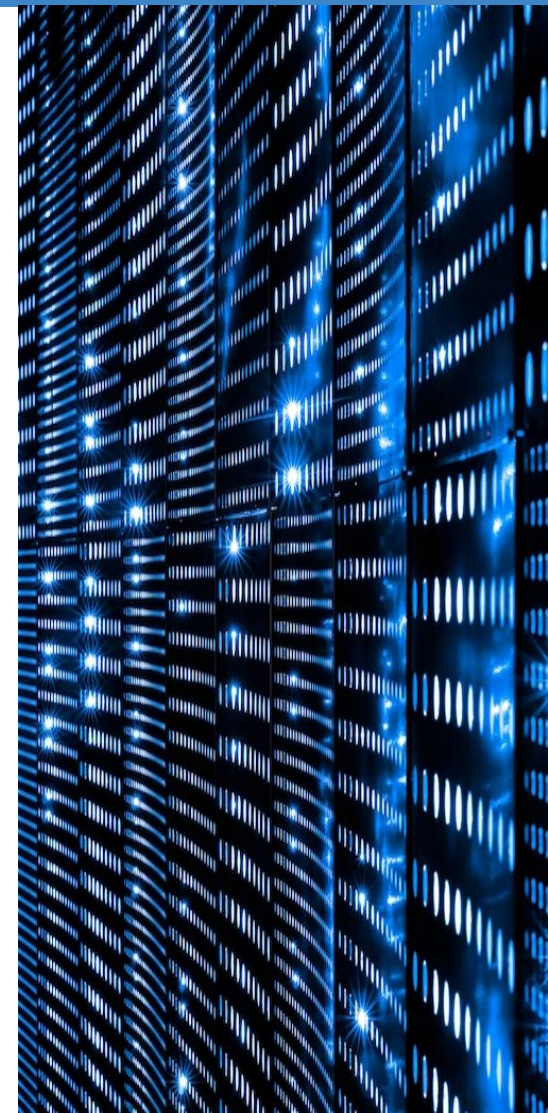
Validation Script GitHub Action

- Action that triggers Oxygen Scripting
- Performs validation on your repository files
- Can be integrated within an exiting workflow



GitHub Actions

github.com/marketplace/actions/oxygen-validation



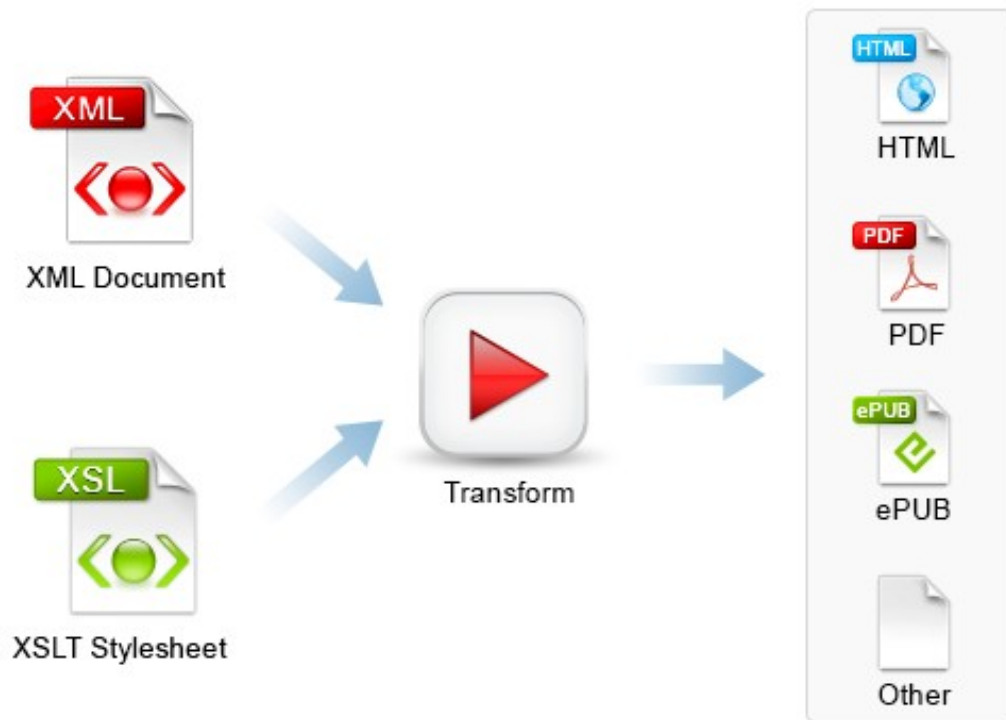
DITA Validate and Check For Completeness

- Validate a DITA map or a DITA Open Toolkit project
- Specify a settings file and a context id

```
sh scripts/validateCheckDITA.sh -i inputFile [-c contextId] [-s settingsFile] [-r reportFile]
```

Transform Script

Execute a transformation scenario from command line or from an integration server (CI/CD)



Transform Script

- Run the transformation scenarios for the existing document types
- Specify a scenarios file or a project file that contains scenarios

```
sh transform.sh -i inputFile -sn scenarioName [-s scenariosFile] [-v]
```

oxygenxml.com/doc/ug-editor/topics/scripting_oxygen_transform.html

Transformation Script GitHub Template

- Automatically transform files on each commit
- Generate report and publish it on [GitHub pages](#)
- Run transformation on specific file from project

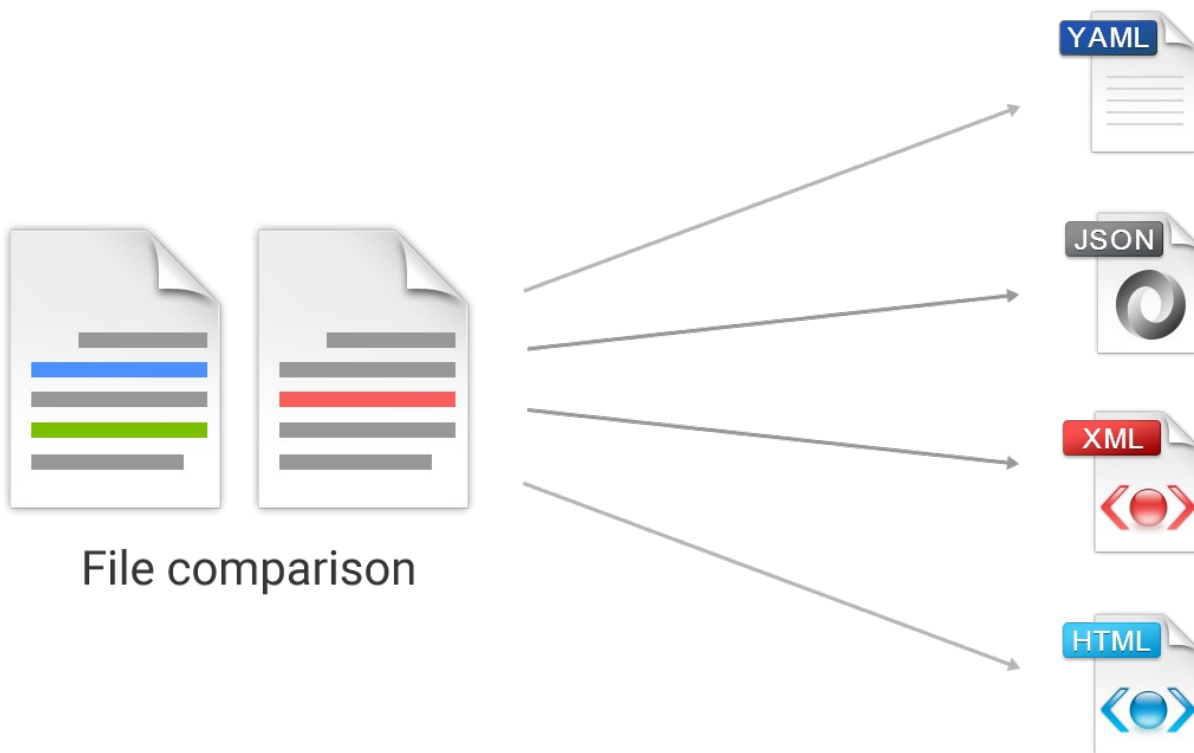


github.com/oxygenxml/oxygen-script-transformation-template



Compare and Merge Files Script

Compare and merge files and get the comparison results in various formats



Compare Files Script Report

Example of generating a file comparison report in HTML format

```
compareFiles file1 file2 -out html
```

Report information:

- Total number of modifications
- Number of modifications by type
- Location of the modifications

Differences: 5 difference blocks, 8 differences in total

Comparison details by difference blocks: all (5) incoming (3) outgoing (2)

| Base file: D:/Sample1/dita-flowers/flowers-base/topics/flowers/gerbera.dita | |
|---|---|
| File 1: D:/Sample1/dita-flowers/flowers-by-John/topics/flowers/gerbera.dita | File 2: D:/Sample1/dita-flowers/flowers-by-Mary/topics/flowers/gerbera.dita |
| 8 is a genus of ornamental plants from the daisy family (Asteraceae). It was named in honor of the German naturalist Traugott Gerber (1710-1743) who travelled extensively in Russia and was a friend of Carl Linnaeus. </p> | + - is a genus of ornamental plants from the sunflower family (Asteraceae). It was named in honor of the German naturalist Traugott Gerber. </p> |
| 12 <!--Maybe we can add more pictures here.--> | + - <p>It has approximately 30 species in the wild, extending to South America, Africa and tropical |
| 18 also known as Transvaal daisy or Barberton Daisy </p> | - + also known as Transvaal daisy or Barberton Daisy. </p> |
| 19 <p>Gerbera species bear a large capitulum with striking, two-lipped ray florets in yellow, | + - <p>Gerbera species bear a large capitulum with striking, two-lipped ray florets in yellow, |
| 25 The domesticated <xref keyref="cultivar" format="dita">cultivars</xref> are mostly a result of a | - + The domesticated <xref format="dita" keyref="cultivar">cultivars</xref> are mostly a result of a |

Merge Files Script

Example of merging the changes between two files:

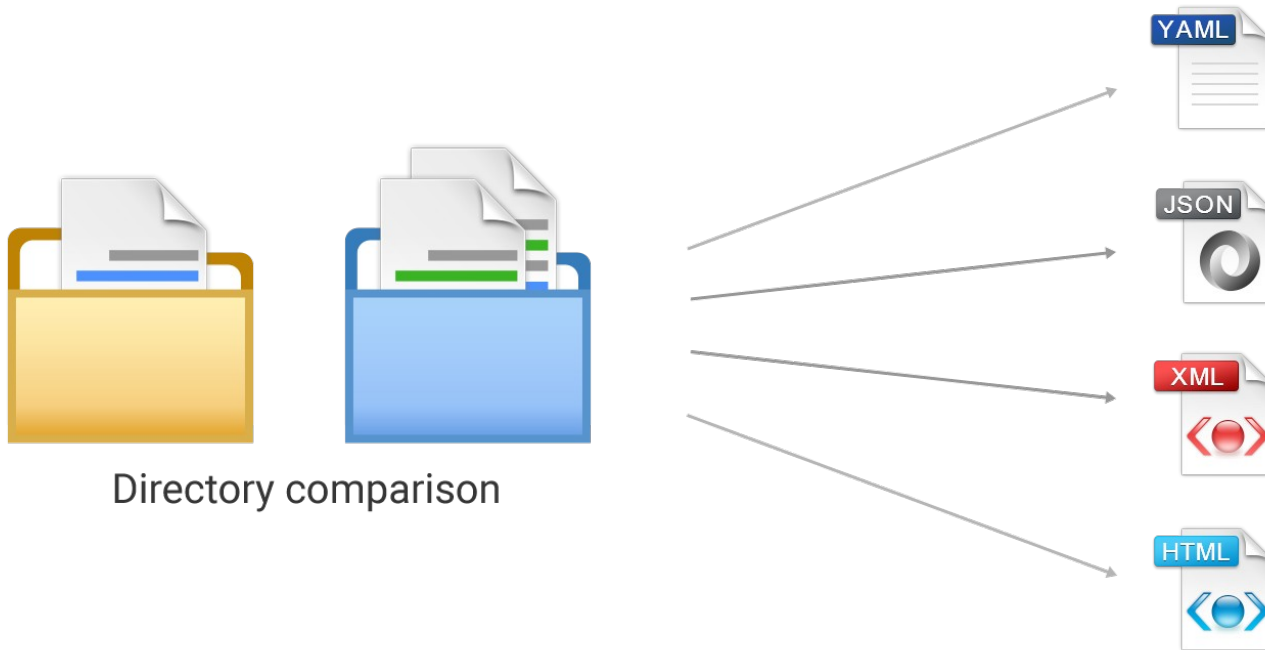
```
compareFiles file1 file2 baseFile -mergeout destFolder
```

```
mergeReport:  
  operationStatus: "OPERATION COMPLETED"  
  changesFound: 8  
  modificationsPerformed: 8
```

oxygenxml.com/doc/ug-editor/topics/scripting-compare-files.html

Compare and Merge Directories Script

Compare and merge directories and get the comparison results in various formats



Compare Directories Script Report

Example of generating a directories comparison report in HTML format

```
compareDirs dir1 dir2 baseDir -out html
```

Report information:

- Total number of modifications
- Number of modifications by type
- Modified files

Differences: 13

Comparison details: all differences (13) outgoing (5) incoming (5) conflicts (3)

Base folder: D:/Sample1/dita-flowers/flowers-base/

Folder 1: D:/Sample1/dita-flowers/flowers-by-John/ Folder 2: D:/Sample1/dita-flowers/flowers-by-Mary/

| File name | Size | Modified | | File name | Size | Modified |
|-----------------------------------|-------|---------------------|-----|---|-------|---------------------|
| concepts/autumnFlowers.dita | 1151 | 2021-07-06 01:49:12 | * × | concepts/autumnFlowers.dita | 1143 | 2021-07-16 06:52:21 |
| concepts/glossaryGenus.dita | 571 | 2021-07-16 08:54:17 | * | concepts/glossaryGenus.dita | 577 | 2021-07-06 01:49:12 |
| concepts/glossaryPanicle.dita | 483 | 2021-07-15 06:53:34 | * | concepts/glossaryPanicle.dita | 495 | 2021-07-06 01:49:12 |
| images/Gerbera.jpg | 10134 | 2021-07-06 01:49:12 | * | images/Gerbera.jpg | 22776 | 2021-07-16 05:55:25 |
| | | | | + publishing/flowers/resources/images/flower_logo.png | 6178 | 2021-07-06 01:49:12 |
| | | | | + publishing/flowers/README.txt | 127 | 2021-07-06 01:49:12 |
| publishing/flowers/README.txt | 127 | 2021-07-06 01:49:12 | - | | | |
| tasks/gardenPreparation.dita | 2275 | 2021-07-06 01:49:12 | * | tasks/gardenPreparation.dita | 2291 | 2021-07-15 12:21:02 |
| topics/flowers/chrysanthemum.dita | 2949 | 2021-07-15 07:48:25 | * | topics/flowers/chrysanthemum.dita | 2932 | 2021-07-06 01:49:12 |
| topics/flowers/gerbera.dita | 2432 | 2021-07-16 06:18:28 | * × | topics/flowers/gerbera.dita | 2456 | 2021-07-16 06:25:13 |
| topics/flowers/snowdrop.dita | 2883 | 2021-07-15 13:57:59 | * | topics/flowers/snowdrop.dita | 2806 | 2021-07-06 01:49:12 |
| topics/test/ | | 2021-07-15 12:36:56 | + | | | |
| topics/introduction.dita | 770 | 2021-07-16 06:58:21 | * × | topics/introduction.dita | 758 | 2021-07-15 12:12:46 |

Merge Directories Script

Example of merging the changes between two directories:

```
compareDirs dir1 dir2 baseDir -mergeout destFolder
```

oxygenxml.com/doc/ug-editor/topics/scripting-compare-directories.html

```
mergeReport:
  operationStatus: "OPERATION COMPLETED"
  added:
    filesNo: 2
    files:
      - publishing\flowers\Resources\images
      - publishing\flowers\READ-ME.txt
  deleted:
    filesNo: 1
    files:
      - publishing\flowers\README.txt
  updated:
    filesNo: 2
    files:
      - images\Gerbera.jpg
      - tasks\gardenPreparation.dita
  merged:
    filesNo: 1
    files:
      - topics\flowers\gerbera.dita
  conflicts:
    filesNo: 3
    files:
      - concepts\autumnFlowers.dita
      - topics\flowers\gerbera.json
      - topics\introduction.dita
```

Conclusion

- Complete Support for **File and Directory Validation**
- Scripts to Generate File and Directory **Validation Reports**
- Options to **Filter the Validated Files**
- Support to Validate **Different Type of Files**
- **Multiple Command-Line Scripts**: Transformation, Compare, Merge, ...

Future Plans

- Improve Validation Script
- GitHub Templates and Actions for Scripting
- JSON Schema and OpenAPI Documentation Script
- XSD to JSON Schema Script
- Other (feedback is welcome)



Question: What features are the most important for you?

- ❑ Improve Validation Script
- ❑ GitHub Templates and Actions for Scripting
- ❑ JSON Schema and OpenAPI Documentation Script
- ❑ XSD to JSON Schema Script
- ❑ Other (feedback is welcome)



Resources

- <https://www.oxygenxml.com/demo/validating-xml-and-json.html>
- https://www.oxygenxml.com/demo/DITA_Map_Completeness_Check.html
- https://www.oxygenxml.com/demo/Schematron_Validation.html
- https://www.oxygenxml.com/demo/Validation_Scenario.html
- https://www.oxygenxml.com/demo/Working_With_XML_Modules.html

- [oxygenxml.com/demo/generating_file_comparison_reports_using_command_line.html](https://www.oxygenxml.com/demo/generating_file_comparison_reports_using_command_line.html)
- [oxygenxml.com/demo/generating_directory_comparison_reports_using_command_line.html](https://www.oxygenxml.com/demo/generating_directory_comparison_reports_using_command_line.html)
- [oxygenxml.com/events/2021/webinar_the_new_oxygen_compare_and_merge_scripts.html](https://www.oxygenxml.com/events/2021/webinar_the_new_oxygen_compare_and_merge_scripts.html)
- [oxygenxml.com/events/2020/webinar_automate_xml_processing_with_oxygen_xml_scripting.html](https://www.oxygenxml.com/events/2020/webinar_automate_xml_processing_with_oxygen_xml_scripting.html)



Questions?

Octavian Nadolu
Product Manager at Syncro Soft

octavian.nadolu@oxygenxml.com

Twitter: [@OctavianNadolu](https://twitter.com/OctavianNadolu)

LinkedIn: [octaviannadolu](https://www.linkedin.com/in/octaviannadolu)