Easy Access Rules XML

How to find and extract the images embedded in the OOXML flat file?

Answer

Automated approach

In the EASA XML export files, images (binary files) are stored in a Base64 ASCII string encoded format, following OpenXML standards. When programmatically accessing the file, Base64 is simple to convert back to binary, e.g.:

```
using System;
using System.Text;
    public class CLRTest
    {
        public static byte[] ConvertBase64ToBinary(string str)
        {
            if (str == null)
            {
                return null;
            }
            return Convert.FromBase64String(str);
        }
    }
```

But how do you locate the data? An example:

Here is a screenshot of a sample eRules XML export, opened in Word:

(1) Where a runway is not provided with a stopway or a clearway and the threshold is located at the extremity of the runway, the four declared distances should normally be equal to the length of the runway

Figure 1

And here is the XML, in pkg:part pkg:name="/word/document.xml":

```
<a:graphic
    xmlns:a="http://schemas.openxmlformats.org/drawingml/2006/main">
    <a:graphicData
      uri="http://schemas.openxmlformats.org/drawingml/2006/picture">
        <pic:pic xmlns:pic="http://schemas.openxmlformats.org/drawingml/2006/picture">
        <pic:nvPicPr>
        <pic:cNvPr id="638492513" name="adrAMCannexIVimage1.png"/>
        <pic:cNvPicPr/>
        </pic:nvPicPr>
        <pic:blipFill>
        <a:blip r:embed="rId18" cstate="print"/>
        <a:stretch>
        </pic:blipFill>
        <pic:spPr>
        </pic:pic>
    </a:graphicData>
</a:graphic>
```

The "rId18" value in the <a:blip> element refers to this relation in pkg:part pkg:name="/word/_rels/document.xml.rels":

```
<Relationship
    Type="http://schemas.openxmlformats.org/officeDocument/2006/relationships/image"
    Target="media/image3.png" Id="rId18"/>
```

And this finally gives you the location of the Base64 encoded data:

```
<pkg:part pkg:name="/word/media/image3.png" pkg:contentType="image/png" pkg:compression="store">
    <pkg:binaryData>iVBORw0KGgoAAAANSUhEUgAAARAAAAB5CAYAAACine0bAAAAAXNSR0IArs4c6QAAAARnQU1BAACx
        jwv8YQUAAAAJcEhZcwAADsMAAA7DAcdvqGQAAA1ySURBVHhe7d1pUBRnAsZxvqRqN192s8ZsTGmS
        UjeHUVNbZmvduLE84hExKIfCeAEiihCIOosRQRBEUcQ1EMZjQaOCioAcSkhpogkSFQURvDg8EDDx
        CCgeIJ7PTvc0OOhgoKeDzvTzq3o3rmbcHq5wM/3T39NtjAyIimRgQIpKNASi2RgQ
        IpKNASi2RgQIpKNASi2RgQIpKNASi2RgQIpLN4gJy9epVlJeXo6SkXo6SkASi2RgQ
        uSL9RjxbFheQ7Oxs9O/fHzY2Nhwcqqybc3Fxc3Fxs376dg0MVY+/evQgP
        D2dAzNEUkCqIGmGSD3i4uIYHMwIKImPDIjvAzkhXgbkPGSAMwBETA0IqxoCXiOEb
        D2dAzNEUkKCqIGmGSD3i4uIYHMwMKIKPDIgvQ9kNhwcqhy9evQgO5wM/MVY+/evQgP
```

Manual approach

For a very simple *manual* approach:

1. Open the XML-file in Word
2. Select the image you want to extract
3. RightClick the image and select "Save as picture":

(4)   Landing distance available (LDA).

The take-off run available (TORA), take-off distance avail[...]op dista[...]
available (ASDA) and landing distance available (LDA) s[...]ding to
following (all declared distances are illustrated for operat[...]

(1)   Where a runway is not provided with a stopway or a[...]d is loca[...]
        at the extremity of the runway, the [...] declared [...]e equa[...]
        the length of the runway
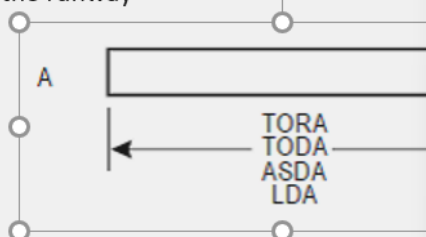
Figure 1

(2)   When a runway is provided with a clearway (CWY), [...] the len[...]
        of clearway.

Figure 2

Last updated:

28/06/2022

Link:

https://www.easa.europa.eu/en/faq/136269
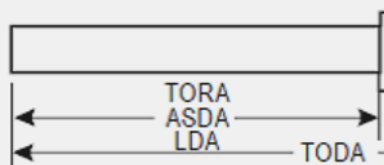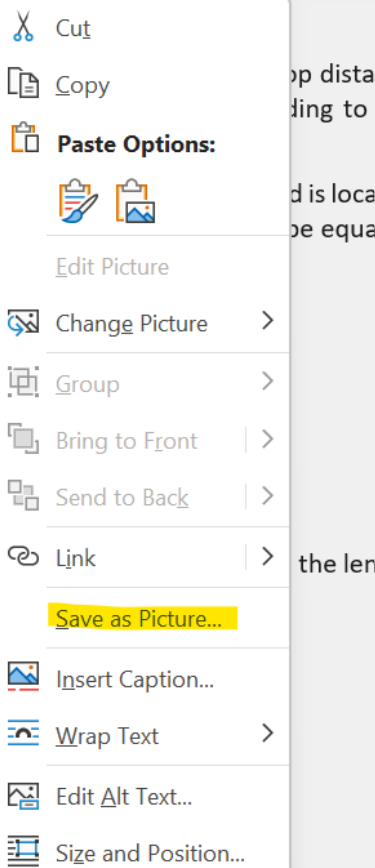
Since the format of the formulas inside OOXML is OMML, can you recommend a way to transform those to MathML (which is the standard for HTML)?

Answer

Automated approach

The OMML is found in the actual text content of the EASA XML files. E.g.:

```
<w:p w14:paraId="7D495C0B" w14:textId="1AD0660E" w:rsidR="00BC70!
    w:rsidRDefault="00355B60">
    <m:oMathPara>
        <m:oMath>
            <m:sSup>
                <m:sSupPr> [6 lines]
                <m:e>
                    <m:d>
                        <m:dPr> [7 lines]
                        <m:e>
                            <m:r>
                                <w:rPr> [3 lines]
                                <m:t>x+a</m:t>
                            </m:r>
                        </m:e>
                    </m:d>
                </m:e>
                <m:sup>
                    <m:r>
                        <w:rPr> [2 lines]
                        <m:t>n</m:t>
                    </m:r>
```

On a computer with a reasonably new version of Microsoft Word installed, you should be able to locate this XSLT file, used by Word to enable the manual process described below:

OMML2MML.XSL
C:\Program Files\Microsoft Office\root\Office16
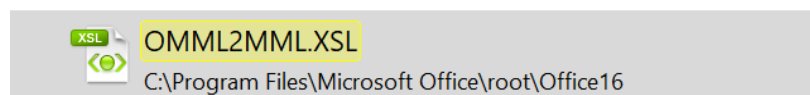
Using normal XML-DOM processing, you can extract the OMML content – and then apply this OMML2MML.XSL stylesheet to transform the OMML to MathML.

Manual approach

To manually transform the OMML to MathML:

A. One time preparation (settings):

1. Open the XML file in Word
2. Select the formula and the "Equations" tab
3. Select the little dropdown sign in the Conversions group on the ribbon (highlighted)
4. In the Equation Options dialog, be sure to turn the "Copy MathML to the clipboard" button on. This only needs to be done once – Word will persist the setting.
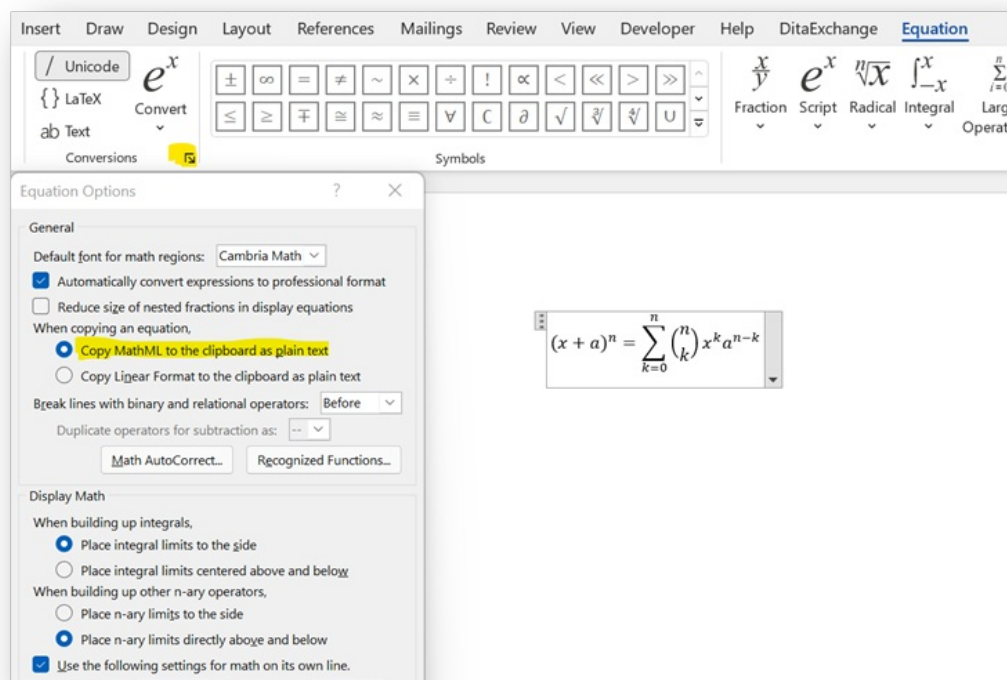
B. To get MathML for an equation (see the screenshot above for the sample equation):

1. Open the XML in Word
2. Select the equation (as you would select any text)
3. Copy (Ctrl+C)
4. Open an editor (for example an XML editor)
5. Create an empty XML file
6. Paste (Ctrl+V) to get this result:

Last updated:

28/06/2022

Link:

https://www.easa.europa.eu/en/faq/136270

How can we import the XML into an SQL database?

Answer

To our best knowledge, all commonly available SQL databases has one or more "import XML" features. Please see the description of your SQL database.
Depending on your particular use case, you could choose to:

1. For XML import, "clean up" the text content before the import, using an XSLT stylesheet
2. Convert to HTML before importing
3. Convert to JSON before importing

Last updated:

28/06/2022

Link:

https://www.easa.europa.eu/en/faq/136271

How can we use the eRules XML in a mobile application?

Answer

This can be done in countless ways. The following is just a few examples.

1. Use a PDF rendering of the file for the mobile application
2. Convert the content to JSON format for processing/use in the mobile app
3. Convert the content to HTML format for processing/use in the mobile app
4. Split the XML-file into topic modules, including topic metadata and then feed the modules into a database/search application (as XML, HTML, or JSON) – and then let the mobile app use the search API/web service of the database/search application to get just the right topic needed, based on the situation

…and there are probably countless other variations of this.

Last updated:

28/06/2022

How can we convert the eRules XML to JSON data?

Answer

Any XML file can easily be converted to JSON by applying a suitable XSLT stylesheet transformation to the XML.

For example, such a stylesheet can freely be downloaded from here:

https://github.com/bramstein/xsltjson

You can customize this XSLT ad libitum, to satisfy the needs of your application.

To get a more useful JSON rendition and before applying this transformation, you may want to use a custom XSLT to remove the formatting tags.

How can we convert the exported XML format to HTML?

Answer

Automated approach

You can develop your own XSLT to convert the XML to HTML. However, depending on the complexity of the content and your requirements, this could be a significant undertaking.

For an easier approach, please see this excellent article on how to programmatically transform the XML to HTML, using free, downloadable tools:
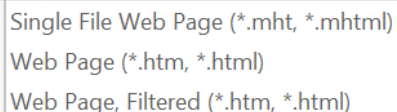
https://docs.microsoft.com/en-us/previous-versions/office/developer/office-2010/ff628051(v=office.14)?redirectedfrom=MSDN

…and see the "General Resources" below for download information.

Manual approach

Follow these steps:

1. Open the XML file in Word
2. Select "File > Save as
3. And choose between:

Single File Web Page (*.mht, *.mhtml)
Web Page (*.htm, *.html)
Web Page, Filtered (*.htm, *.html)

You can experiment with the three options to see what works best for your app.

Last updated:

28/06/2022

Link:

https://www.easa.europa.eu/en/faq/136274

How can we remove the formatting information from the XML?

Answer

If the target format is HTML, you should start by taking a look at Answer 6 above (How can we convert the exported XML format to HTML?).

Otherwise, it is quite easy to use XSLT processing to remove formatting information, by simply "ignoring it" in your XSLT.

Here is a very small example that would clean out a lot of formatting and various other attributes from a paragraph, leaving only a <p> tag and the text itself:

```xslt
<xsl:template match="w:p">
    <p><xsl:value-of select="normalize-space(string(.)"/></p>
</xsl:template>
```

Last updated:
28/06/2022

Link:
https://www.easa.europa.eu/en/faq/136275

How can we merge our own operating procedures with the EASA content?

Answer

This can be done in many, many ways.

At the one extreme – a very basic example, you can simply edit the XML file in Microsoft Word, adding or removing content as needed (note there would be legal waivers involved here!).

Maintaining this solution when newer versions of the eRules are published would be entirely manual, and risky too.

At the other extreme, you could import the EASA XML eRules into your preferred component-based content management solution (CCMS) – and then use the components as integrated components in you own solution.

This could involve just referencing the components from your proprietary components – or even cloning the EASA components for customisation.

In case of the use of a CCMS, it would normally be possible to automate the updating the content as new releases are released by EASA.

Last updated:
28/06/2022

Link:
https://www.easa.europa.eu/en/faq/136276

When an amended version of a rule is published in the XML format, is there a way to determine what are the changes compared to a previous version?

Answer

There are several ways to make a comparison. One is to simply - programmatically - compare the two XML files and extract the changes. In addition to that or as an alternative to this method, you can use the attribute topic-metadata/@RegulatorySource. When a topic which appears in one version of a publication is modified in a subsequent version, the value of its attribute topic-metadata/@RegulatorySource changes.

For example: The topic with the identifier ERulesId="ERULES-1963177438-2548" was present in one version of a rule with the value of the attribute
topic-metadata/@RegulatorySource ="ED Decision 2014/012/R"
If the topic's content is modified in the next version of the rule, then the attribute topic-metadata/@RegulatorySource will receive a new value corresponding to the decision that

approved the change, for example topic-metadata/@RegulatorySource ="ED Decision 2018/009/R"

Last updated:
28/06/2022

Link:
https://www.easa.europa.eu/en/faq/136277

I'm interested in extracting the data structure inside a topic, because my application is used for checking compliance with criteria that are listed (e.g. as bulleted items) inside topics. How can I do this?

Answer

The current version of eRules XML specification regards a topic (a rule paragraph) as the lowest level of content that can be identified and extracted from the rule.
This means that the structure of the text inside a topic is not guaranteed to be unambiguously identified by automated processing.

In practice, however, a processing party can use the formatting (e.g. styling) information available as part of OOXML tagging to extract the necessary data. However, cautious needs to be exercised before using the data as part of a completely automated process. This is because the styling information associated with parts of text is not guaranteed to remain the same between different versions of rules.

Last updated:
28/06/2022

Link:
https://www.easa.europa.eu/en/faq/136278

In the pdf version of Easy Access Rules sometimes one topic presents two versions of the same rule or of the same paragraph (see the figure below). What does it mean and how is this represented in the XML format?

Answer

Some of the Easy Access Rule have topics containing rules which are applicable currently as well as rules which are applicable at a later date.
When viewed in PDF or in XML opened in Microsoft Word, such a topic containing such content

looks like this:

Note the special colouring (magenta) of the content applicable at a later date.

While the eRules XML Export version 1.0.0 does not provide EASA-specific XML elements that would allow a processor to identify the two types of content within a topic, it is possible to use the OOXML formatting tags for such a goal.

This is because inside the (OO)XML file the text applicable at a later date is formatted using a special style with the value "GeneralAviation"

More concretely, a text with a later applicability date will have as a Property (<w:pPr> or <w:rPr>) a Run Style (<w:rStyle>) with the attribute w:val="GeneralAviation" as in the following snippet which describes the rendering to OOXML of the text

(b)   The design management system shall:


```
<w:p  w:rsidR="00741909"

    w:rsidRPr="00741909"

    w:rsidP="00741909"

    w14:paraId="1DFCCDAA"

    w14:textId="1B30FEA1">

    <w:pPr>

<w:pStyle w:val="ListLevel0"/>

<w:rPr>

<w:rStyle w:val="GeneralAviation"/>

</w:rPr>

    </w:pPr>

<w:r w:rsidRPr="00741909">

<w:rPr>

<w:rStyle w:val="GeneralAviation"/>
```

```
</w:rPr>

<w:t>(b)</w:t>

</w:r>

<w:r w:rsidRPr="00741909">

<w:rPr>

<w:rStyle w:val="GeneralAviation"/>

</w:rPr>

<w:tab/>

<w:t>The design management system shall:</w:t>

</w:r>

</w:p>
```

Note: This special style was introduced as part of Easy Access Rules publication in PDF format as a visual aid helping humans understand how a topic will be changed in the future. As such, it is not included as part of the first release of the eRules XML Export specification. Future releases will include EASA specific XML tags allowing a more consistent identification and extraction of data structures inside a topic without relying on formatting elements.

Last updated:
28/06/2022

Link:
https://www.easa.europa.eu/en/faq/136653

When a topic is open in pdf/Word or OOXML, the regulatory source of the content is displayed as indicated by the arrow in the picture below. Where can I retrieve this information in the XML?

Answer

Article 7 Permit to fly

Regulation (EU) No 748/2012

The conditions determined before 28 March 2007 by the Member States for permits to fly or other airworthiness certificate issued for aircraft which did not hold a certificate of airworthiness or restricted certificate of airworthiness issued under this Regulation, are deemed to have been determined in accordance with this Regulation, unless the Agency has determined before 28 March 2008 that such conditions do not provide for a level of safety equivalent to that required by Regulation (EC) No 216/2008 or this Regulation.

A topic content in XML will contain the text identifying the regulatory document introducing the topic or the last amendment thereto. However, the recommended way to retrieve that data is to extract it from the corresponding metadata 'Regulatory source' (please refer to EASA eRules XML Export Specification, Chapter 5.2 'Attribute topic-metadata — business description of the metadata')

```
<er:topic sdt-id="-1455166071"
source-title="Article 7 Permit to fly"
ERulesId="ERULES-1963177438-3631"
Domain="Initial airworthiness;"
ActivityType=""
AircraftUse=""
AircraftCategory=""
AmendedBy=""
ApplicabilityDate=""
EntryIntoForceDate=""
EquivalentForeignRegulation=""
ICAOReference=""
Keywords=""
RegistryState=""
RegulatedEntity=""
RegulatorySource="Regulation (EU) No 748/2012"
RegulatorySubject="Part-21;Cover regulation;"
TechnicalSubjectMatter=""
TypeOfContent="IR (Implementing rule);"
ParentIR="Powers and recitals"
EASACategory=""/>
```

Last updated:

28/06/2022

Link:

https://www.easa.europa.eu/en/faq/136654

In the XML Schema, there are two elements, "frontmatter" and "backmatter". We have not seen these elements in the actual XML-files, what is their purpose?

Answer

These two elements are reserved for future use (for example to hold document elements as legal disclaimers, table of contents, copyright notices, indices, etc.).

Last updated:

28/06/2022

Link:

https://www.easa.europa.eu/en/faq/136655

I'm trying to export the content of these xml files to an excel file, and I did not succeed. What can I do?

Answer

The files containing EASA's Easy Access Rules in XML format can be opened for visualization and used directly with many different tools, including:

- Microsoft Word (where it would appear as a normal Word document)
- any XML editor
- any text editor (like Notepad, WordPad) that is able to open large text files

Normally, it is also possible to open an XML file in Microsoft Excel, but if the XML file is very big, Excel will fail or generate errors. Microsoft has not currently published information about the exact size limit, and the complexity of the XML structure could also play a role.

This means that, for the files containing EASA's Easy Access Rules in XML format, Microsoft Excel is not a recommended tool.

If your objective is to just have a look at the XML content, Notepad would be a readily available choice although the lack of XML syntax comprehension does not make it practical for understanding the structure. An XML editor would be more useful for that purpose.

Otherwise, if the objective is to be able to extract, process and use the content in other applications, we would recommend investigating the use of a processing software that could transform the export XML into exactly the format you need for your other applications.

The processing software could be, for example

- an XSLT parser using a customized XSLT transformation
- a custom made software written in your language of choice

Last updated:

08/12/2022

Link:

https://www.easa.europa.eu/en/faq/137331

I tried to open one of the XML files in my text/XML editor but it is displayed incorrectly and I think the XML file may be corrupted. What can I do?

Answer

Please note that some of the XML files can be quite large and not all text/XML editors can handle and display correctly such large files.

Therefore If the XML file is not displayed properly when you try to view it in text/XML editor then:

- please make sure your text or XML editor is able to handle large XML files.
- try opening it in MS Word. As it was mentioned in the documentation the eRules XML files, being in OOXML format, can be opened in MS Word. If you can open the file in MS Word then it is very likely that the file is well-formed
- use an XML parser to check the syntactical correctness (in other words check if the XML file is well-formed). For ensuring the file is not corrupted you don't need to check its compliance with the referenced namespaces.

Last updated:

08/12/2022

Link:

https://www.easa.europa.eu/en/faq/137332

Are the Easy Access Rules topic IDs stable over the lifecycle of a topic, including when a topic is deleted?

Answer

The EAR content published from the eRules platform in the XML format is divided into topics (the smallest units of information). Each topic has a unique identifier, the ERulesId (e.g., ERULES-1963177438-14838), and this is stable over time. This means that even if the topic title and metadata are changed in new versions, the ERulesId will uniquely identify the topic throughout the lifecycle.
In EASA's internal system, topics will not be deleted but will be marked obsolete. Should the topic be re-instated in any way, it will have the same ERulesId as before.

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137674

Will the changes to topic metadata be part of the version history of a topic and will EASA publish timeline versions of topics?

Answer

In EASA's internal system, the historical versions of each topic are preserved, and the version history includes the changes to metadata for the topic. For the time being we do not export the history of changes.

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137675

Are the sdt-id values stable over time?

Answer

The value of the sdt-id attribute is only valid as an internal pointer within the same XML file, linking a topic's metadata to its actual content, and is not guaranteed to remain the same in different versions of the same XML file. Therefore, it cannot be used as a unique identifier for a topic across several XML files.

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137676

Are there any stable IDs available on the sub-topic level, for example for paragraphs, sub-paragraphs and list items?

Answer

With respect to the internal structure of the content inside a topic, please see the question: I'm interested in extracting the data structure inside a topic, because my application is used for

checking compliance with criteria that are listed (e.g. as bulleted items) inside topics. How can I do this?

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137677

Can Easy Access Rules in XML be imported into other applications, for example spreadsheets or databases?

Answer

The EAR XML format can be transformed (possibly with loss of some information) to any character-based format, for example JSON, HTML, other XML formats or the simple CSV (comma-separated values) format.

The transformation can be done using an XSLT stylesheet or other software approaches for manipulating the XML files.

Many of these target formats can be then imported directly into spreadsheets or databases, please see the documentation for the application in question.

See also the answer to the question: How can we import the XML into an SQL database?

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137678

The XML files of Easy Access Rules (EAR) are very large and not easy to open and navigate. Have you considered exporting smaller units/files?

Answer

Many different tools are available to manually work on the (large) XML files, ranging from simple text editors to professional XML Editor. However, the EAR XML files are intended to be used as input to an automated transformation process executed by a toolchain. This process should transform the input file into the format required by your application. This transformation could be used to split the large file into smaller units, for example one file containing the structure and the metadata (the elements inside the "er" namespace) and then one file per

topic.

It is also possible to manually reduce the size of files, by deleting the unwanted parts of the XML content, for example removing the pkg:part elements containing the base64-encoded images. This can be easily done using a simple XSLT program.

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137679

How can we identify changes in the eRules, down to the paragraph level? Will you export any "delta" files, specifying only the changes?

Answer

For identifying changes down to the topic level, please refer to the question:When an amended version of a rule is published in the XML format, is there a way to determine what are the changes compared to a previous version?

Otherwise, there are many tools (with customisable sensitivity, so only important changes are included) available that will allow you to programmatically determine changes. Currently, we do not export 'delta' files.

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137680

Would you provide a specific DTD file with the XML file provided, in order to interpret and manage this XML file format in a consistent way? And why not using the ATA standard S1000D format?

Answer

The eRules XML format used is fully documented using the Office Open XML standard (XML Schemas available) with the EASA Erules XML Schema added. DTDs (the XML Schema's predecessor) are not available. Using the schemas you can validate the XML, both the "er" namespace and the Office Open XML parts. In most of the cases, though, you probably do not need to validate against the schema in order to extract the data from the file.

There are many partially competing standard formats available, and while the ATA S1000D is certainly interesting, we did find that using OOXML as a basis provided us with a good coverage of the stakeholder use cases, as well as future flexibility.

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137681

Could you please explain how the elements work? Do they have their own IDs?

Answer

The toc container elements illustrate how the topics are arranged hierarchically – they illustrate the table of content structure that the topics are placed in. All the topic and child toc elements for a particular toc element are contained inside a toc element.

They do not have an ID on their own – as opposed to the topics which all have IDs.

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137682

The smallest "text" unit is an Implementing Rule (IR) paragraph, as I understand. Are there plans to develop the XML scheme to also handle sub-paragraphs (a), (b), (c) etc., like ADR.OR.D.005(a)? The reason for this question is, that EASA tends to combine several compliance topics/subjects in one IR. But one often has to handle these compliance topics separately (e.g. in manuals).

Answer

The topic is currently the smallest content unit with a unique identifier. However, we are working on enhancing the content architecture, with sub-topic level elements and semantic tagging.

With respect to a book-like structure for Easy Access Rules (chapter/section/paragraph/ etc.), we use a more dynamic model where the <heading> and <toc> elements, which can contain

either a child <toc> and/or a topic, represent all levels above the topics. We think this structure is more suitable to be used in different applications e.g. in portals.

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137683

Are there any plans to provide API access?

Answer

The suggestion to provide REST-based API access to the eRules is very interesting. We are in the process of identifying the use cases that will require the implementation of such an API.

Last updated:

08/03/2023

Link:

https://www.easa.europa.eu/en/faq/137684

What is the legal status of Easy Access Rules in the XML format?

Answer

Easy Access Rules are issued by EASA to provide its stakeholders with an updated, consolidated, and easy-to-read publication. It has been prepared by putting together the officially published EU regulations/EASA certification specifications with the related EASA acceptable means of compliance (AMC) and guidance material (GM) (including the amendments) adopted so far. Easy Access Rules (in any of the available formats) are not an official publication. Only European Union documents published in the Official Journal of the European Union are deemed authentic. (For the rules governing authenticity of the Official Journal, see Council Regulation (EU) No. 216/2013 of 7 March 2013). If errors are brought to our attention, we will try to correct them. However, EASA accepts no responsibility or liability whatsoever with regard to the information contained in Easy Access Rules. Metadata associations are not part of the legal text of official publications, but they are ancillary features to help users explore and extract data.

Last updated:

13/03/2023

Link:

https://www.easa.europa.eu/en/faq/137699

Are there any intellectual copyright rules to be followed when using the XML content in a process support application?

Answer

Please refer to the copyright notice on the Easy Access Rules XML export page

Last updated:

13/03/2023

Link:

https://www.easa.europa.eu/en/faq/137700