ACER
Agency for the Cooperation of Energy Regulators

REMIT

Guidance on the implementation of web feeds for Inside Information Platforms

Version 1.0

15 November 2016

Agency for the Cooperation of Energy Regulators
Trg Republike 3
1000 Ljubljana, Slovenia

Version history

Version	Effective Date
Version 1.0	15 November 2016

Related Documents

- Regulation (EU) No 1227/2011 of the European Parliament and of the Council on wholesale energy market integrity and transparency, http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:326:0001:0016:en:PDF

- Commission Implementing Regulation (EU) No 1348/2014 on data reporting implementing Article 8(2) and (6) of Regulation (EU) No 1227/2011, http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:JOL_2014_363_R_0009&from=EN

- Updated 4th edition of ACER Guidance on the application of Regulation (EU) No 1227/2011 of the European Parliament and of the Council of 25 October 2011 on wholesale energy market integrity and transparency, 3 June 2015, www.acer.europa.eu/Official_documents/Other_documents/4th_Edition_ACER_Guidance REMIT.pdf

- ACER's Manual of Procedures on transaction data, fundamental data and inside information reporting, http://www.acer.europa.eu/en/remit/REMITATACER/Data_collection/Pages/default_ORIGINAL.aspx

- ACER's Public Consultation on the Common Standard for the Disclosure of Inside Information, 27 May 2015, http://www.acer.europa.eu/Official_documents/Public_consultations/Pages/PC_2015_R_03.aspx

- FAQs on REMIT fundamental data and inside information https://www.acer-remit.eu/portal/custom-category/remit_questions

Table of Contents

1. Introduction

1.1. Purpose

This document will serve as a technical guidance for the implementation of web feeds, namely RSS or Atom feeds containing Urgent Market Messages (UMMs) that will enable the Agency to collect inside information efficiently, as defined in Article 10(1) of Commission Implementing Regulation (EU) No 1348/2015 ('the Implementing Regulation').

1.2. Scope

The inside information collection by the Agency from inside information platforms starts as of January 2017. UMMs provided via web-feeds from Inside Information Platforms that publish UMMs on behalf of market participants - i.e. the platforms listed on the REMIT portal – will be collected.

1.3. Intended Audience

This document is targeted to Inside Information Platforms published on the REMIT Portal: https://www.acer-remit.eu/portal/list-inside-platforms. As it includes technical guidance on how the inside information will be collected it is expected that IT staff involved in technical implementation of web feeds will be familiar with the document.

This document does not apply to individual market participants.

1.4. Abbreviations

Acronym	Full version	Definition
UMM	Urgent Market Message	Published Inside Information shall be disclosed in the form of UMMs
RSS	Rich Site Summary (originally RDF Site Summary; often called Really Simple Syndication)	RSS FEED in compliance
ATOM	ATOM Syndication Format	ATOM FEED in compliance to RFC-4287 and RFC-5023
XML	eXtensible Markup Language	Markup language that defines set of rules to encode information in a specific format commonly called XML Document
XML Document	eXtensible Markup Language Document	Document written according to XML language
XSD	Xml Schema Definition	Document that describes in a specific language the format in which xml document must be written in order to comply with a structure of sort
SHA	Secure Hash Algorithm	Cryptographic Hash Function that produce a message digest based on an input content

2. UMM Reporting Choice

Two main and most widespread industry standards - RSS and ATOM feeds will be supported.

Any Inside Information Platform can use **RSS 2.0** feeds or **Atom** feeds at their own convenience. The Agency's UMM Collection System will autonomously detect which of the two supported feeds is being used and adapt accordingly to fetch the information.

3. UMM Information Reporting Mechanics

There are three major types of Inside Information:

- Electricity
- Gas
- Other

A single platform can report all the possible Inside Information data types at the same or via multiple URL(s) without providing additional indications.

Regardless of whether choosing RSS or Atom in terms of format and regardless of the nature of the information published there are some criteria that must be strictly adhered to:

1) Compliance with the information structure as detailed in paragraphs 3.1 and 3.2;
2) The feed will have to start as a blank feed (devoid of information, including though any information that was added in the last three hours before the new day, more information in the next paragraphs) every day at 00:00 UTC, meaning that at least for the address provided to the Agency for Inside Information publication, the expected behaviour is that at any given time the feed will contain only information published for a single calendar day. It has to be ensured that the updates to the feed made available just before the feed is reset are still included in the new feed for the next day as the Agency will not poll the feed continuously. As a general rule all Inside Information published less than **three** hours before the feed is reset should be included in the feed for the next day to ensure information is not lost. Please note that in exceptional circumstances when due to technical issues the feed would not be available for several days it is expected that after the issues are resolved the feed would include information from previous days that were not made available during the period of feed unavailability;
3) The feeds for the last 15 days **must** be kept and provided to the Agency on request. The channel for providing such data will be agreed on a case-by-case basis with the provider of the feed. Further URLs for those historical feeds would simplify the collection and are a preferred but not required solution. Only the final complete daily feeds need to be retained, there is no need to keep the intermediate updates within a day.

3.1. Data Provisioning

Based on each of the three XSD schemas a separate XML file per schema can be generated containing the information for 1 to n (unlimited) events tied to a specific event family depending on the specific XSD used.

These XML files should be validated against the XSD to ascertain the compliance of the xml, the relevant xsd can be found on the REMIT Portal under Documents --> Remit Reporting User Package inside the “ANNEX VIII - XML SCHEMA FOR INSIDE INFORMATION REPORTING“.

After that, the XML file(s) can be included inside the feed. While both RSS and ATOM feeds have fields in common and fields that are unique to one of the two feed types, the following should be considered:

- An ATOM feed has a repeating element called “*entry*” with sub-element(s) called “summary”, whereas an RSS feed has a repeating element called “*item*” with sub-element(s) called “description”.
- The automated collection relies on the concept of repeatable items with sub-field “*description*”//”summary” (respectively for RSS and for ATOM). It is expected to include within the CDATA declaration of the “description”//”summary” fields a valid XML that will contain the inside information relevant to the specific feed update.

  One single “description”//”summary” sub-element is dedicated for a single XML file (the XML file might hold multiple UMM messages).

The system does not have any expectations nor any restrictions to standard webfeed fields in the RSS (e.g. title, language, etc.) or the ATOM (e.g. title, subtitle, updated, etc.) standard, the only requirements that are made are relevant to the content of the fields that will hold the XML message embedded inside CDATA. The only fields that have to be considered are the following:

RSS:

<lastBuildDate> = datetime of the creation of the feed
<link> = URL where the feed is published
<pubDate> = datetime of the publication of the feed/item

ATOM:

<updated> = datetime of the publication of the feed/entry
<link> = URL where the feed is published

An example of how the RSS feed should be structured is below:

```
<?xml version="1.0" encoding="utf-8"?>
<rss version="2.0">
<channel>
<title>UMM for Platform ABC</title>
<description>This is a sample feed</description>
<link>http://www.someaddress.com</link>
<language>en-us</language>
<lastBuildDate>Tue, 19 Oct 2004 13:39:14 -0400</lastBuildDate>
<pubDate>Tue, 19 Oct 2004 13:38:55 -0400</pubDate>
<item>
<title>UMM Event1</title>
<description>
<![CDATA[ <?xml version="1.0" encoding="UTF-8"?> <ns1:REMITUrgentMarketMessages
xmlns:ns1="http://www.acer.europa.eu/REMIT/REMITUMMElectricitySchema_V1.xsd"
xmlns:ns2="http://www.acer.europa.eu/REMIT/REMITUMMCommonSchema_V1.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.acer.europa.eu/REMIT/REMITUMMElectricitySchema_V1.xsd
file:/D:/aris_stuff/UMM/REMIT_Inside%20Information_Electricity_Schema.xsd"> <ns1:UMM>
<ns1:messageId>messageIddfgertrertyetyer_001</ns1:messageId> <ns1:event>
<ns1:eventStatus>Active</ns1:eventStatus> <ns1:eventType>Production unavailability</ns1:eventType>
<ns1:eventStart>2006-05-04T18:13:51.0Z</ns1:eventStart> <ns1:eventStop>2006-05-
04T18:13:51.0Z</ns1:eventStop> </ns1:event> <ns1:unavailabilityType>Planned</ns1:unavailabilityType>
<ns1:publicationDateTime>2006-05-04T18:13:51.0Z</ns1:publicationDateTime> <ns1:capacity>
<ns1:unitMeasure>MW</ns1:unitMeasure> <ns1:unavailableCapacity>0</ns1:unavailableCapacity>
<ns1:availableCapacity>0</ns1:availableCapacity> <ns1:installedCapacity>0</ns1:installedCapacity>
</ns1:capacity> <ns1:unavailabilityReason>unavailabilityReason0</ns1:unavailabilityReason>
<ns1:biddingZone>00Taaaaaaaaaaaa</ns1:biddingZone>
```

<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone> <ns1:affectedAsset>
<ns2:name>name0</ns2:name> </ns1:affectedAsset> <ns1:marketParticipant>
<ns2:name>name1</ns2:name> <ns2:lei>lei1111111111111111</ns2:lei> </ns1:marketParticipant>
<ns1:marketParticipant> <ns2:name>name2</ns2:name> <ns2:lei>lei33333333333333333</ns2:lei>
</ns1:marketParticipant> </ns1:UMM> <ns1:UMM>
<ns1:messageId>messageId2fgertrertyetyer_001</ns1:messageId> <ns1:event>
<ns1:eventStatus>Active</ns1:eventStatus> <ns1:eventType>Production unavailability</ns1:eventType>
<ns1:eventStart>2006-05-04T18:13:51.0Z</ns1:eventStart> <ns1:eventStop>2006-05-
04T18:13:51.0Z</ns1:eventStop> </ns1:event> <ns1:unavailabilityType>Planned</ns1:unavailabilityType>
<ns1:publicationDateTime>2006-05-04T18:13:51.0Z</ns1:publicationDateTime> <ns1:capacity>
<ns1:unitMeasure>MW</ns1:unitMeasure> <ns1:unavailableCapacity>0</ns1:unavailableCapacity>
<ns1:availableCapacity>0</ns1:availableCapacity> <ns1:installedCapacity>0</ns1:installedCapacity>
</ns1:capacity> <ns1:unavailabilityReason>unavailabilityReason1</ns1:unavailabilityReason>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone> <ns1:affectedAsset>
<ns2:name>name3</ns2:name> </ns1:affectedAsset> <ns1:marketParticipant>
<ns2:name>name4</ns2:name> <ns2:bic>bic11111111</ns2:bic> </ns1:marketParticipant>
<ns1:marketParticipant> <ns2:name>name5</ns2:name> <ns2:lei>lei55555555555555555</ns2:lei>
</ns1:marketParticipant> </ns1:UMM> </ns1:REMITUrgentMarketMessages>]]>
</description>
<link>http://www.somefeed.com/alink</link>
<pubDate>Thu, 07 Jul 2016 11:09:11 -0400</pubDate>
</item>
<item>
<title>UMM Event2</title>
<description>
<![CDATA[<?xml version="1.0" encoding="UTF-8"?> <ns1:REMITUrgentMarketMessages
xmlns:ns1="http://www.acer.europa.eu/REMIT/REMITUMMElectricitySchema_V1.xsd"
xmlns:ns2="http://www.acer.europa.eu/REMIT/REMITUMMCommonSchema_V1.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.acer.europa.eu/REMIT/REMITUMMElectricitySchema_V1.xsd
file:/D:/aris_stuff/UMM/REMIT_Inside%20Information_Electricity_Schema.xsd"> <ns1:UMM>
<ns1:messageId>messageIddfgertrertyetyer_002</ns1:messageId> <ns1:event>
<ns1:eventStatus>Active</ns1:eventStatus> <ns1:eventType>Production unavailability</ns1:eventType>
<ns1:eventStart>2006-05-04T18:13:51.0Z</ns1:eventStart> <ns1:eventStop>2006-05-
04T18:13:51.0Z</ns1:eventStop> </ns1:event> <ns1:unavailabilityType>Planned</ns1:unavailabilityType>
<ns1:publicationDateTime>2006-05-04T18:13:51.0Z</ns1:publicationDateTime> <ns1:capacity>
<ns1:unitMeasure>MW</ns1:unitMeasure> <ns1:unavailableCapacity>0</ns1:unavailableCapacity>
<ns1:availableCapacity>0</ns1:availableCapacity> <ns1:installedCapacity>0</ns1:installedCapacity>
</ns1:capacity> <ns1:unavailabilityReason>unavailabilityReason0</ns1:unavailabilityReason>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone> <ns1:affectedAsset>
<ns2:name>name0</ns2:name> </ns1:affectedAsset> <ns1:marketParticipant>
<ns2:name>name1</ns2:name> <ns2:lei>lei1111111111111111</ns2:lei> </ns1:marketParticipant>
<ns1:marketParticipant> <ns2:name>name2</ns2:name> <ns2:lei>lei33333333333333333</ns2:lei>
</ns1:marketParticipant> </ns1:UMM> <ns1:UMM>
<ns1:messageId>messageId2fgertrertyetyer_002</ns1:messageId> <ns1:event>
<ns1:eventStatus>Active</ns1:eventStatus> <ns1:eventType>Production unavailability</ns1:eventType>
<ns1:eventStart>2006-05-04T18:13:51.0Z</ns1:eventStart> <ns1:eventStop>2006-05-
04T18:13:51.0Z</ns1:eventStop> </ns1:event> <ns1:unavailabilityType>Planned</ns1:unavailabilityType>
<ns1:publicationDateTime>2006-05-04T18:13:51.0Z</ns1:publicationDateTime> <ns1:capacity>
<ns1:unitMeasure>MW</ns1:unitMeasure> <ns1:unavailableCapacity>0</ns1:unavailableCapacity>
<ns1:availableCapacity>0</ns1:availableCapacity> <ns1:installedCapacity>0</ns1:installedCapacity>
</ns1:capacity> <ns1:unavailabilityReason>unavailabilityReason1</ns1:unavailabilityReason>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone> <ns1:affectedAsset>
<ns2:name>name3</ns2:name> </ns1:affectedAsset> <ns1:marketParticipant>

<ns2:name>name4</ns2:name> <ns2:bic>bic11111111</ns2:bic> </ns1:marketParticipant>
<ns1:marketParticipant> <ns2:name>name5</ns2:name> <ns2:lei>lei55555555555555555</ns2:lei>
</ns1:marketParticipant> </ns1:UMM> </ns1:REMITUrgentMarketMessages>]]>
</description>
<link>http://www.somefeed.com/alink</link>
<pubDate>Thu, 07 Jul 2016 11:09:11 -0400</pubDate>
</item>
</channel>
</rss>

An example of how the ATOM feed should be structured is below:

<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">
<title>UMM for Platform ABC</title>
<subtitle>This is a sample feed</subtitle>
<link href="http://www.someaddress.com" rel="self" />
<id>urn:uuid:60a76c80-d399-11d9-b91C-0003939e0af6</id>
<updated>2003-12-13T18:30:02Z</updated>
<entry>
<title>UMM Event1</title>
<link href="http://example.org/2003/12/13/atom03" />
<id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344efa6a</id>
<updated>2003-12-13T18:30:02Z</updated>
<summary><![CDATA[<?xml version="1.0" encoding="UTF-8"?>
<ns1:REMITUrgentMarketMessages
xmlns:ns1="http://www.acer.europa.eu/REMIT/REMITUMMElectricitySchema_V1.xsd"
 xmlns:ns2="http://www.acer.europa.eu/REMIT/REMITUMMCommonSchema_V1.xsd"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xsi:schemaLocation="http://www.acer.europa.eu/REMIT/REMITUMMElectricitySchema_V1.xsd
file:/D:/aris_stuff/UMM/REMIT_Inside%20Information_Electricity_Schema.xsd">
 <ns1:UMM><ns1:messageId>messageIddfgertrertyetyer_001</ns1:messageId><ns1:event>
 <ns1:eventStatus>Active</ns1:eventStatus>
 <ns1:eventType>Production unavailability</ns1:eventType>
 <ns1:eventStart>2006-05-04T18:13:51.0Z</ns1:eventStart>
 <ns1:eventStop>2006-05-04T18:13:51.0Z</ns1:eventStop>
 </ns1:event>
 <ns1:unavailabilityType>Planned</ns1:unavailabilityType>
 <ns1:publicationDateTime>2006-05-04T18:13:51.0Z</ns1:publicationDateTime>
 <ns1:capacity>
 <ns1:unitMeasure>MW</ns1:unitMeasure>
 <ns1:unavailableCapacity>0</ns1:unavailableCapacity>
 <ns1:availableCapacity>0</ns1:availableCapacity>
 <ns1:installedCapacity>0</ns1:installedCapacity>
 </ns1:capacity>
 <ns1:unavailabilityReason>unavailabilityReason0</ns1:unavailabilityReason>
 <ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
 <ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
 <ns1:affectedAsset>
 <ns2:name>name0</ns2:name>
 </ns1:affectedAsset>
 <ns1:marketParticipant>
 <ns2:name>name1</ns2:name>
 <ns2:lei>lei11111111111111111</ns2:lei>
 </ns1:marketParticipant>
 <ns1:marketParticipant>

<ns2:name>name2</ns2:name>
<ns2:lei>lei33333333333333333</ns2:lei>
</ns1:marketParticipant>
</ns1:UMM>
<ns1:UMM>
<ns1:messageId>messageId2fgertrertyetyer_001</ns1:messageId>
<ns1:event>
<ns1:eventStatus>Active</ns1:eventStatus>
<ns1:eventType>Production unavailability</ns1:eventType>
<ns1:eventStart>2006-05-04T18:13:51.0Z</ns1:eventStart>
<ns1:eventStop>2006-05-04T18:13:51.0Z</ns1:eventStop>
</ns1:event>
<ns1:unavailabilityType>Planned</ns1:unavailabilityType>
<ns1:publicationDateTime>2006-05-04T18:13:51.0Z</ns1:publicationDateTime>
<ns1:capacity>
<ns1:unitMeasure>MW</ns1:unitMeasure>
<ns1:unavailableCapacity>0</ns1:unavailableCapacity>
<ns1:availableCapacity>0</ns1:availableCapacity>
<ns1:installedCapacity>0</ns1:installedCapacity>
</ns1:capacity>
<ns1:unavailabilityReason>unavailabilityReason1</ns1:unavailabilityReason>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
<ns1:affectedAsset>
<ns2:name>name3</ns2:name>
</ns1:affectedAsset>
<ns1:marketParticipant>
<ns2:name>name4</ns2:name>
<ns2:bic>bic11111111</ns2:bic>
</ns1:marketParticipant>
<ns1:marketParticipant>
<ns2:name>name5</ns2:name>
<ns2:lei>lei55555555555555555</ns2:lei>
</ns1:marketParticipant>
</ns1:UMM>
</ns1:REMITUrgentMarketMessages>]]>
</summary>
</entry>
<entry>
<title>UMM Event2</title>
<link href="http://example.org/2003/12/13/atom03" />
<id>urn:uuid:1225c695-cfb8-4ebb-aaaa-80da344efa6a</id>
<updated>2003-12-13T18:30:02Z</updated>
<summary><![CDATA[<?xml version="1.0" encoding="UTF-8"?>
<ns1:REMITUrgentMarketMessages
xmlns:ns1="http://www.acer.europa.eu/REMIT/REMITUMMElectricitySchema_V1.xsd"
xmlns:ns2="http://www.acer.europa.eu/REMIT/REMITUMMCommonSchema_V1.xsd"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.acer.europa.eu/REMIT/REMITUMMElectricitySchema_V1.xsd
file:/D:/aris_stuff/UMM/REMIT_Inside%20Information_Electricity_Schema.xsd">
<ns1:UMM>
<ns1:messageId>messageIddfgertrertyetyer_002</ns1:messageId>
<ns1:event>
<ns1:eventStatus>Active</ns1:eventStatus>
<ns1:eventType>Production unavailability</ns1:eventType>
<ns1:eventStart>2006-05-04T18:13:51.0Z</ns1:eventStart>
<ns1:eventStop>2006-05-04T18:13:51.0Z</ns1:eventStop>

```
</ns1:event>
<ns1:unavailabilityType>Planned</ns1:unavailabilityType>
<ns1:publicationDateTime>2006-05-04T18:13:51.0Z</ns1:publicationDateTime>
<ns1:capacity>
<ns1:unitMeasure>MW</ns1:unitMeasure>
<ns1:unavailableCapacity>0</ns1:unavailableCapacity>
<ns1:availableCapacity>0</ns1:availableCapacity>
<ns1:installedCapacity>0</ns1:installedCapacity>
</ns1:capacity>
<ns1:unavailabilityReason>unavailabilityReason0</ns1:unavailabilityReason>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
<ns1:affectedAsset>
<ns2:name>name0</ns2:name>
</ns1:affectedAsset>
<ns1:marketParticipant>
<ns2:name>name1</ns2:name>
<ns2:lei>lei11111111111111111</ns2:lei>
</ns1:marketParticipant>
<ns1:marketParticipant>
<ns2:name>name2</ns2:name>
<ns2:lei>lei33333333333333333</ns2:lei>
</ns1:marketParticipant>
</ns1:UMM>
<ns1:UMM>
<ns1:messageId>messageId2fgertrertyetyer_002</ns1:messageId>
<ns1:event>
<ns1:eventStatus>Active</ns1:eventStatus>
<ns1:eventType>Production unavailability</ns1:eventType>
<ns1:eventStart>2006-05-04T18:13:51.0Z</ns1:eventStart>
<ns1:eventStop>2006-05-04T18:13:51.0Z</ns1:eventStop>
</ns1:event>
<ns1:unavailabilityType>Planned</ns1:unavailabilityType>
<ns1:publicationDateTime>2006-05-04T18:13:51.0Z</ns1:publicationDateTime>
<ns1:capacity>
<ns1:unitMeasure>MW</ns1:unitMeasure>
<ns1:unavailableCapacity>0</ns1:unavailableCapacity>
<ns1:availableCapacity>0</ns1:availableCapacity>
<ns1:installedCapacity>0</ns1:installedCapacity>
</ns1:capacity>
<ns1:unavailabilityReason>unavailabilityReason1</ns1:unavailabilityReason>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
<ns1:biddingZone>00Taaaaaaaaaaaaa</ns1:biddingZone>
<ns1:affectedAsset>
<ns2:name>name3</ns2:name>
</ns1:affectedAsset>
<ns1:marketParticipant>
<ns2:name>name4</ns2:name>
<ns2:bic>bic11111111</ns2:bic>
</ns1:marketParticipant>
<ns1:marketParticipant>
<ns2:name>name5</ns2:name>
<ns2:lei>lei55555555555555555</ns2:lei>
</ns1:marketParticipant>
</ns1:UMM>
</ns1:REMITUrgentMarketMessages>]]>
</summary>
```

</entry>
</feed>

Important Note

When performing inclusion of an XML inside CDATA tags please be aware that:

1) What is being embedded within the CDATA tags is an XML file, so any non UTF-8 characters will cause parsing issues.
2) After the **<![CDATA[** tag no additional character should be present before the starting header of the XML file (**<?xml version="1.0" encoding="UTF-8"?>**) which must always be explicitly declared. If this is not respected the extracted XML file will be considered as not compliant with the relevant XSD schema and the validation of the file will fail.
3) Avoid putting entire content of the XML file in one line. Use line breaks to ensure the file is not parsed as one very long line of text which may cause issues with parsing and processing.

3.2. Updating the daily information

Each tag (item//entry) represents an information update compared to the previous item//entry that was provided in the previous version of the RSS//ATOM feed, with a new set of UMMs.

It is expected that different items/entries hold different information. This means that if the first RSS item field contains the description of certain Urgent Market Messages it is expected that other item fields present in the same RSS feed contain different information (e.g. other UMMs).

As a single item/entry can contain an XML file that can include one or more UMMs it is important to ensure that each update of the feed is done by adding new items/entries that contain different XML files with different UMMs. No overlapping of information is allowed.

A specific UMM message should ONLY be inside one XML (and thus in one specific description//summary CDATA tag), so this means that no single UMM message should be repeated inside the XML file across the various item/entry tags.

To further elaborate on this point please consider the following example.

3.2.1. High-level notation used in an example

```
<rss>
        <usualRssField>...</usualRssField>
        <item>
                <description cdata></description>   <--- XML content
        </item>
        <item>
                <description cdata></description>   <--- XML content
        </item>
</rss>
```

3.2.2. Progressive Update Example

At 09:00 a certain RSS feed is published:

```
<rss>
        <usualRssField>...</usualRssField>
        <item>
                <description cdata>*UMMEvent1-START*,*UMMEvent2-START*</description>
        </item>
</rss>
```

This means that the feed at 9:00 contains two UMM events that have started on that particular date at a certain time.

At 11:30 there is an update to the previously published feed and the new feed should look like this:

```
<rss>
        <usualRssField>...</usualRssField>
        <item>
                <description cdata..>*UMMEvent1-START*,*UMMEvent2-START*</description>
        </item>
        <item>
                <description cdata..>*UMMEvent3-START*,*UMMEvent2-UPDATE*</description>
        </item>
</rss>
```

Important Note

1) The feed now has a new "*item*" and the information is updated in a way to preserve the way the old information was represented.
2) There is no informational overlap between the events as the new item contains no information about the old events but only the new updates of previous events (update for UMMEvent2) or addition of new events (UMMEvent3).

3.2.3. Duplicate item recognition Example

Consider another example (below the feed at 11:30):

```
<rss>
        <usualRssField>...</usualRssField>
        <item>
                <description cdata..>*UMMEvent1-START*,*UMMEvent2-START*</description>
        </item>
        <item>
                <description cdata..>*UMMEvent3-START*,*UMMEvent2-UPDATE*</description>
        </item>
</rss>
```

At 15:30 the feed is updated:

```
<rss>
        <usualRssField>...</usualRssField>
        <item>
                <description cdata..>*UMMEvent1-START*,*UMMEvent2-START*</description>
        </item>
        <item>
                <description cdata..>*UMMEvent3-START*,*UMMEvent2-UPDATE*</description>
        </item>
        <item>
                <description cdata..>*UMMEvent1-START*,*UMMEvent2-START*</description>
        </item>
        <item>
                <description cdata..>*UMMEvent1-START*,*UMMEvent2-START*,*UMMEvent7-
        START*</description>
        </item>
</rss>
```

In this situation, when the new RSS feed is pulled and provided that the previous one at 11:30 was correctly processed, the UMM Collection Mechanism will automatically recognise that the third item is completely the same as the first, and will discard its content as duplicated, whereas the content of the fourth item is different from the first and the data will be fully processed though a duplicate information for UMMEvent1-START is present. The mechanism is quite simple: each time a feed is polled and parsed the SHA value is calculated for each string embedded within CDATA tags for all such items present in the feed. The system will check if there is an entry//item from the same platform on the same day that has the same SHA for the CDATA content. If yes, the further processing of the individual item//entry is skipped and the information will be considered as a duplicate.

A similar approach is used to determine whether the informational content of the feed has changed since the last polling. The system will poll the available feed on a scheduled interval and it will use the SHA of the whole feed as an indication of whether or not there was an update to the feed since the last polling.

3.3. Data duplication

Apart from the mechanism described in the previous section there is currently no other mechanism implemented that could allow identification of duplicates. It is therefore expected that a significant portion of duplicate data will be processed so it is even more important that the providers of the feeds make sure duplication does not occur already at the level of the single feed.

4. Registration of URLs embedding the RSS/Atom feed with ACER

In order to bring the content of the feed into the Agency's REMIT Information System (ARIS), a connector (automated polling process) is established between ARIS and the Inside Information Platform's website embedding the RSS/ATOM feed.

The connector will be established based on the feed URL(s) of the platforms.

If an Inside Information Platform publishes UMMs on multiple websites (using multiple feeds) or uses multiple URLs for embedding the RSS/ATOM feeds different from those for the UMM publication on the platform, then all the URLs have to be registered in ARIS which are relevant for UMM data collection.

Web feeds can be made available over http or https without any preference as far as no authentication mechanism is required.

It is expected that Inside Information Platforms will inform the Agency (via inside.information@acer.europa.eu) by 30 November 2016 about the URLs they will use to publish webfeeds according to this guidance. It is also expected that the feeds will not be protected by a password. Access to the feeds can be limited to The Agency's IP addresses if necessary, but this has to be clearly requested so that the Agency can provide a list of IP addresses that will allow to poll the feeds.