



European Network of  
Transmission System Operators  
for Electricity

---

# RDF-SYNTAX USER GUIDE

---

2024-02-08

---

VERSION 1.0  
ICTC APPROVED

1 Copyright notice:

2 **Copyright © ENTSO-E. All Rights Reserved.**

3 This document and its whole translations may be copied and furnished to others, and derivative  
4 works that comment on or otherwise explain it or assist in its implementation may be prepared,  
5 copied, published and distributed, in whole or in part, without restriction of any kind, provided  
6 that the above copyright notice and this paragraph are included on all such copies and  
7 derivative works. However, this document itself may not be modified in any way, except for  
8 literal and whole translation into languages other than English and under all circumstances, the  
9 copyright notice or references to ENTSO-E may not be removed.

10 This document and the information contained herein is provided on an "as is" basis.

11 **ENTSO-E DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT**  
12 **LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT**  
13 **INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR**  
14 **FITNESS FOR A PARTICULAR PURPOSE.**

15 **This document is maintained by the ENTSO-E CIM WG. Comments or remarks are to be**  
16 **provided at [cim@entsoe.eu](mailto:cim@entsoe.eu)**

17 **NOTE CONCERNING WORDING USED IN THIS DOCUMENT**

18 The force of the following words is modified by the requirement level of the document in which  
19 they are used.

- 20 • **SHALL:** This word, or the terms "REQUIRED" or "MUST", means that the definition is an  
21 absolute requirement of the specification.
- 22 • **SHALL NOT:** This phrase, or the phrase "MUST NOT", means that the definition is an  
23 absolute prohibition of the specification.
- 24 • **SHOULD:** This word, or the adjective "RECOMMENDED", means that there may exist valid  
25 reasons in particular circumstances to ignore a particular item, but the full implications must  
26 be understood and carefully weighed before choosing a different course.
- 27 • **SHOULD NOT:** This phrase, or the phrase "NOT RECOMMENDED", means that there may  
28 exist valid reasons in particular circumstances when the particular behaviour is acceptable  
29 or even useful, but the full implications should be understood and the case carefully weighed  
30 before implementing any behaviour described with this label.
- 31 • **MAY:** This word, or the adjective "OPTIONAL", means that an item is truly optional.

32

## Revision History

Version	Release	Date	Paragraph	Comments
0	0	2023-10-01		CGMES SG's first draft of the document.
0	1	2023-11-08		CIM WG agreement to continue the document and separate the JSON-LD specialisation.
0	0.2	2023-11-15		CGMES SG revision of the document and reset the numbering to 0.0.2. Name of the document changed from <i>RDF-syntax Data Exchange Specification</i> to <i>RDF-SyntaxUsageGuidelines</i> .
0	2	2023-12-18		Integrating additional feedback from CGMES SG
1	0	2024-01-17		Integrating additional feedback from CIM WG

34 **CONTENTS**

35 Copyright notice:..... 2

36 Revision History..... 3

37 CONTENTS ..... 4

38 1. Introduction ..... 5

39 2. Provided application profiles ..... 5

40 3. Combining different CIM versions ..... 6

41     Specifics on RDF/CIM-XML-syntax serialization: ..... 7

42 4. General differences between CIM XML (552) and RDF XML (W3C) ..... 7

43 5. Different CIM RDFS versions ..... 10

44 6. Datatypes and associated issues ..... 11

45 7. Available Tools ..... 11

46 8. SHACL based constraints and validation ..... 13

47     8.1. Validation of datasets ..... 13

48     8.2. Validation of multiple dependent datasets ..... 13

49

50 **List of figures**

51 **No table of figures entries found.**

52 **List of tables**

53 Table 1 ..... 12

54

## 55 1. Introduction

56 This document aims at providing technical information and guidance for software developers  
57 and power system engineers that are implementing RDF based data exchange using standards  
58 such as IEC 61970-600-1:2021, IEC 61970-600-2:2021 (CGMES) or ENTSO-E specifications  
59 such as Network Code (NC) Data Exchange Specification.

60 The document intends to decrease the learning curve for people that are new to software  
61 implementations based on RDF technology or are looking for some necessary technical details  
62 to explain the reasoning of directions taken.

63 The information provided in the document that relates to CIMXML and RDFS does not replace  
64 or amend requirements and/or statements provided in other approved and published documents  
65 and it should be treated as technical guidance only. The disclaimer below which relates to  
66 application profiles should be noted.

### 67 Disclaimer

68 *The test configurations (models), documents and application profiles are owned by ENTSO-E  
69 and are provided by ENTSO-E “as it is”. To the fullest extent permitted by law, ENTSO-E shall  
70 not be liable for any damages of any kind arising out of the use of the test configurations  
71 (models), documents and application profiles (including any of their subsequent modifications).*

72 *ENTSO-E neither warrants, nor represents that the use of the test configurations (models),  
73 documents and application profiles will not infringe the rights of third parties. Any use of the  
74 test configurations (models), documents and application profiles shall include a reference to  
75 ENTSO-E. ENTSO-E web site is the only official source of information related to these test  
76 configurations (models), documents and application profiles.*

## 77 2. Provided application profiles

78 The application profiles are provided to facilitate the implementation of the CGMES profiles and  
79 related constraints as defined in IEC 61970-600-1:2021, IEC 61970-600-2:2021, IEC 61970-  
80 301 and other related 61970-45x series of profiles.

81 Note that the application profile serialization based on RDFS and RDF XML syntax is defined  
82 in IEC 61970-501:2006 (Ed1) and CIM XML serialization is defined in IEC 61970-552:2016.  
83 However, current implementations deviate from these standards due to various reasons  
84 addressed in this document.

85 For CGMES v3.0 the machine-understandable application profiles include the following  
86 packaged:

- 87 • RDFS2020 e.g., IEC61970-600-2\_CGMES\_3\_0\_0\_RDFS2020 for CGMES v3.0

88 A RDFS 2020 update export (see details on the update in the section “Different CIM RDFS  
89 versions”) of the RDFS augmented version that is based on IEC 61970-501:2006 (Ed1) and  
90 used for exporting the RDFS for CGMES v2.4. The only difference (compared with RDFS2019  
91 variant) is resolving export technical issues and the information from the abstract version class  
92 that is instantiated as part of the header of the RDFS instead as a version class with all details.  
93 No functional changes were made in RDFS2020 compared with RDFS2019. The notation “2020”  
94 does not refer to the year of generation, but it is the version of the augmented RDFS export by  
95 CimSyntaxGen.

- 96 • RDFSEd2Beta, e.g., IEC61970-600-2\_CGMES\_3\_0\_0\_RDFSEd2Beta for  
97 CGMES v3.0

98 This is a beta version of application profile based on RDFS specified in the draft IEC 61970-  
99 501:Ed2. The purpose of inclusion of the beta version in the distribution is to enable review  
100 process. Please use these files only for information on the direction where RDFS will evolve in  
101 that standard and provide feedback that will be discussed in the standardization process.  
102 Namely, the RDFS contains the vocabulary only, while the constraints (cardinalities, datatypes,  
103 etc.) are expressed by SHACL based constraints.

- 104 • SHACL, e.g., IEC61970-600-2\_CGMES\_3\_0\_0\_SHACL for CGMES v3.0

105 This is a package of all SHACL shapes/constraints applicable for CGMES v3.0. These are  
106 constraints for cardinalities and datatypes derived from the RDFS, constraints defined in the  
107 descriptions of the classes and attributes, constraints defined in IEC 61970-600-1:2021, IEC  
108 61970-600-2:2021, IEC 61970-301 and other related 61970-45x series of profiles and  
109 expressed there in plain English text. Note that SHACL based constraints in this folder are  
110 serialized in two RDF formats, Turtle and RDF XML plain (no nesting). Originally the constraints  
111 were developed in Turtle using Notepad++ as an editor and then converted to RDF XML using  
112 CimPal app. Because many constraints rely on SHACL SPARQL method, which is not covered  
113 in the draft IEC 61970-501:Ed2, the RDF XML may not represent the desired way of  
114 serialization. However, the resulted RDF XML version was not used or validated in terms of  
115 content and should be used with a caution.

116 The recommended serialization of SHACL constraints is Turtle as that was the primary  
117 serialization and it is well tested. There is a question in the standardization community if RDF  
118 XML will need to be supported as amended or new development will be done in JSON-LD. The  
119 tendency is that the JSON-LD will become the main serialization that semantic web tool vendors  
120 must support while the other (e.g. Turtle and RDF XML) becomes optional.

- 121 • OCL, e.g. IEC61970-600-2\_CGMES\_3\_0\_0\_OCL for CGMES v3.0

122 [deprecated, obsolete] This is a package of OCL based constraints that cover CGMES v3.0 in  
123 a similar way as SHACL shapes cover necessary validation scope. The RDFS Extracted  
124 subfolder contains OCL constraints derived from the RDFS. The XLSX Extracted subfolder  
125 contains constraints defined in the descriptions of the classes and attributes, constraints  
126 defined in IEC 61970-600-1:2021, IEC 61970-600-2:2021, IEC 61970-301 and other related  
127 61970-45x series of profiles and expressed there in plain English text. However, please note  
128 that this package was developed in Nov 2020 and may have deviations compared to the  
129 published version of CGMES v3.0. OCL is no longer maintained. Only SHACL constraints will  
130 be maintained.

131 Packaging for other profiles is different:

- 132 • CGMES v2.4 - do not include SHACL constraints; RDFS exports are done with  
133 earlier versions i.e., not RDFS2020; OCL constraints are provided
- 134 • NC profiles – follow the setup as in CGMES v3.0: SHACL constraints are  
135 provided both for derived from RDFS constraints and custom SHACL constraints;  
136 OCL is not provided; RDFS2020 is exported.

### 137 3. Combining different CIM versions

138 Historically every version of CIM canonical model<sup>1</sup> and related profiles<sup>2</sup> had different URI for  
139 the namespace. In addition, some implementations rely on namespace prefix, not on actual  
140 namespace URI. This makes it impossible to combine or support mix of versions / provenance  
141 in the instance file, which should technology-wise not be a problem (namespace concept serves  
142 this). These are reasons that software applications have difficulties in handling or combining  
143 data from different CIM version. Consequently, if there is data exchanged that is governed by  
144 different CIM versions – each version in own dataset – larger or smaller amount of custom

<sup>1</sup> Canonical model is published in standards like IEC 61970-301 and IEC 61970-302

<sup>2</sup> Profiles are published in standards like IEC 61970-452, 61970-453, 61970-456, 61970-600-1, 61970-600-2, etc.

145 pre/post-processing would be required where the different versions are compatible with each  
146 other, to fit the data for handling with off-the-shelf tools.

147 In order to support implementations, starting with CIM18, the CIM international standard  
148 development community agreed to keep the URI of the canonical CIM stable between different  
149 versions of CIM. This means that if a class is defined in CIM vocabulary its URI will not change.  
150 Semantic versioning is applied on profile level and different packages in CIM in order to be able  
151 to describe and explain CIM evolution.

152 Starting from CIM18, the following setup is planned:

- 153 • Namespace URI is stable
- 154 • Each package of the CIM canonical model is versioned with URI. The URI changes only when  
155 the package is modified. This allows tracing changes. This also enables a process in which a  
156 standard that defined a profile does not need to be updated if CIM version changes and if the  
157 profile is depended on canonical packages that have not changed compared to previous version
- 158 • Each profile has version URI that changes every time the profile changes.
- 159 • Semantic versioning is applied to all canonical model, profiles and all machine-readable  
160 artifacts, i.e., if RDFS and SHACL constrains can be updated independently of IEC standard if  
161 the standard document is not impacted.

#### 162 4. Specifics on RDF/CIM-XML-syntax serialization: General differences between 163 CIM XML (552) and RDF XML (W3C)

164 The CIM XML is defined in the IEC 61970-552:2016. This version of the standard is based on  
165 a much earlier edition in which some serialization assumptions were made. Important: When  
166 the initial version of IEC 61970-552 was developed, the W3C recommendations on RDF XML  
167 were not released. Therefore, there was a growing gap during the last two decades. The latest  
168 RDF XML was standardized by W3C in 2014 ([RDF 1.1 XML Syntax \(w3.org\)](#)) and IEC 61970-  
169 552 did not align with this due to existing implementations objecting changes in CIM XML.

170 Many experts complain that RDF is difficult to read due to the references and the flat structure  
171 of the file, i.e., no nesting as present in XSD<sup>3</sup>-governed XML. Such complaints should not be  
172 addressed to RDF in general, but rather to IEC 61970-552 CIM XML. The W3C RDF XML can  
173 be serialized in different forms and many open libraries as Apache Jena support these natively.  
174 For instance, the abbreviated version of RDF XML is very much mirroring nested structure of  
175 XSD-governed XML.

176 Another important point to note is that RDF as a general framework is not bound to a given  
177 serialization. RDF based dataset can be serialized in different forms – CIM XML (IEC 61970-  
178 552), [RDF XML \(W3C\)](#), [Turtle \(W3C\)](#), [JSON-LD \(W3C\)](#), [N-Triples \(W3C\)](#), etc. Each of these  
179 different serializations have their advantages and disadvantages. In general Turtle, due to its  
180 human readability, is a preferred serialization to provide example datasets when explaining  
181 concepts. JSON-LD, which is not the JSON, but a special JSON for linked data, is targeted  
182 serialization for the future.

183 CIM XML, used for CIM based data exchanges, is based on RDF XML serialization specification  
184 and restricts it to simplify processing of instance data. However, it also introduces some  
185 changes or special assumptions, which have evolved with time and have been difficult to change  
186 (due to current implementations for CIM based data exchanges), but which require special  
187 pre/post processing when using off-the-shelf RDF tools. Known differences between CIM XML  
188 and the RDF XML are listed below:

- 189 - CIM XML defined in IEC 61970-552 is unclear regarding the exchange of the datatypes (float,  
190 integer, etc.). The approach taken by the community is that datatypes are not exchanged with  
191 the instance data assuming receiving party is aware of expected datatypes. Therefore, when  
192 consuming data one needs to know the expected datatype from the information in the Schema

<sup>3</sup> W3C XML Schema, used to validate XML instance data. [W3C XML Schema Definition Language \(XSD\) 1.1 Part 1: Structures](#)

193 (RDFS), e.g., in the RDFS we have information that an attribute has ActivePower as a datatype  
194 and that value is float, unit W and multiplier M. Note that there are differences in profiling  
195 approaches and for example if the profile is generated using CIMTool, a separate profile  
196 (separate schema) needs to be prepared for the Domain package.

197 The example below is from RDFS of EQ profile and illustrates how ActivePower datatype is  
198 defined. The property cims:isFixed<sup>4</sup> is used to define that the values for multiplier and unit. The  
199 ActivePower.value attribute has cims:dataType property which defines that the datatype is the  
200 primitive Float which maps to xsd:float.

```

201     <rdf:Description rdf:about="#ActivePower">
202     <rdfs:label xml:lang="en">ActivePower</rdfs:label>
203     <rdfs:comment
204     rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Product of RMS value
205     of the voltage and the RMS value of the in-phase component of the
206     current.</rdfs:comment>
207     <cims:stereotype>CIMDatatype</cims:stereotype>
208     <cims:belongsToCategory
209     rdf:resource="http://iec.ch/TC57/ns/CIM/CoreEquipment-
210     EU#Package_CoreEquipmentProfile"/>
211     <rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
212     </rdf:Description>
213     <rdf:Description rdf:about="#ActivePower.value">
214     <cims:stereotype
215     rdf:resource="http://iec.ch/TC57/NonStandard/UML#attribute"/>
216     <rdfs:label xml:lang="en">value</rdfs:label>
217     <rdfs:domain rdf:resource="#ActivePower"/>
218     <cims:dataType rdf:resource="#Float"/>
219     <cims:multiplicity rdf:resource="http://iec.ch/TC57/1999/rdf-schema-
220     extensions-19990926#M:0..1" />
221     <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-
222     ns#Property"/>
223     </rdf:Description>
224     <rdf:Description rdf:about="#ActivePower.multiplier">
225     <cims:stereotype
226     rdf:resource="http://iec.ch/TC57/NonStandard/UML#attribute"/>
227     <rdfs:label xml:lang="en">multiplier</rdfs:label>
228     <rdfs:domain rdf:resource="#ActivePower"/>
229     <rdfs:range rdf:resource="#UnitMultiplier"/>
230     <cims:multiplicity rdf:resource="http://iec.ch/TC57/1999/rdf-schema-
231     extensions-19990926#M:0..1" />
232     <cims:isFixed
233     rdf:datatype="http://www.w3.org/2001/XMLSchema#string">M</cims:isFixed>
234     <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-
235     ns#Property"/>
236     </rdf:Description>
237     <rdf:Description rdf:about="#ActivePower.unit">
238     <cims:stereotype
239     rdf:resource="http://iec.ch/TC57/NonStandard/UML#attribute"/>
240     <rdfs:label xml:lang="en">unit</rdfs:label>
241     <rdfs:domain rdf:resource="#ActivePower"/>
242     <rdfs:range rdf:resource="#UnitSymbol"/>
243     <cims:multiplicity rdf:resource="http://iec.ch/TC57/1999/rdf-schema-
244     extensions-19990926#M:0..1" />
245     <cims:isFixed
246     rdf:datatype="http://www.w3.org/2001/XMLSchema#string">W</cims:isFixed>
247     <rdf:type rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-
248     ns#Property"/>
249     </rdf:Description>
250

```

---

<sup>4</sup> *cims* is the prefix for namespace <https://iec.ch/TC57/1999/rdf-schema-extensions-19990926#> which is the CIM-specific extension to RDFS.



251 In accordance with current versions of standards the information of the datatypes is not  
252 exchanged and in the instance data, the following will be serialized

```
253 <cim:SynchronousMachine rdf:about="#_3a3b27be-b18b-4385-b557-6735d733baf0">
254   ...
255   <cim:RotatingMachine.p>-90</cim:RotatingMachine.p>
256   ...
257 </cim:SynchronousMachine>
258
```

259 Looking at the instance data it is not possible to say if the value is kW or MW. It is also not  
260 possible to validate if the value is float as a common parser will parse the value as a string.  
261 This is why the parser needs to use the information from the RDF in order to assign the expected  
262 datatype, e.g. xsd:float and then SHACL validator can validate if the value conforms to the  
263 declared datatype for this property.

264 If the datatype were instantiated it would look like this below. However, this would increase  
265 information in the instance file.

```
266 <cim:SynchronousMachine rdf:about="#_3a3b27be-b18b-4385-b557-6735d733baf0">
267   ...
268   <cim:RotatingMachine.p
269   rdf:datatype="http://www.w3.org/2001/XMLSchema#float">-
270   90</cim:RotatingMachine.p>
271   ...
272 </cim:SynchronousMachine>
273
```

274 Still this does not solve the problem with multipliers. Currently the validation of this is part of  
275 the conformity process related to CGMES. This is yet another reason why conformity is  
276 important.

277 Starting from CIM18, there is an agreement to change the multipliers in all profiles to "none".  
278 The result of this will be that the value for active power will be exchanged in W. Engineering  
279 notation is used to help serializing values with the right precision as shown below.

```
280 <cim:SynchronousMachine rdf:about="#_3a3b27be-b18b-4385-b557-6735d733baf0">
281   ...
282   <cim:RotatingMachine.p>-90E6</cim:RotatingMachine.p>
283   ...
284 </cim:SynchronousMachine>
285
```

286 - In CIM XML, there is special treatment of rdf:ID vs. rdf:about. In CIM XML:

- 287 ○ rdf:ID is used for all objects that are serialized first time in the instance data (semantic
- 288 of "create"), while
- 289 ○ rdf:about is used when an object (the instance of the class) is updated. This should not
- 290 be confused with an update of a value of a property. For example, in the equipment
- 291 (EQ) profile all classes have rdf:ID as the EQ is the base profile, but in Steady State
- 292 Hypothesis (SSH) all objects are with rdf:about as only attributes are added to existing
- 293 classes already exchanged in the EQ. In CGMES profiling style such classes are
- 294 marked with stereotype "Description".

295 Rule MVAL5 in IEC 61970-600-1:2021 provides some example of this.

296 - In CIM XML, there is no declaration of xml:base, which means that parsed data will get local  
297 URI if the parser do not impose specific xml:base at the time of the parsing.

298 If the CIM XML the following and xml:base is not declared:  
 299 <cim:ACLineSegment rdf:ID="\_ffbabc27-1ccd-4fdc-b037-e341706c8d29">  
 300 For example, Apache Jena library will produce this, something comparable is to expect for other  
 301 standard RDF parsing tools. The identifier of the object is  
 302 [file:\\C:\\Temp\\test.xml#\\_ffbabc27-1ccd-4fdc-b037-e341706c8d29](file:\\C:\\Temp\\test.xml#_ffbabc27-1ccd-4fdc-b037-e341706c8d29) if the instance file was located  
 303 in temp folder in C drive.  
 304

305 If xml:base<sup>5</sup> is declared are the time of parsing and if the base is <http://iec.ch/TC57/CIM100> the  
 306 result is  
 307 [http://iec.ch/TC57/CIM100#\\_ffbabc27-1ccd-4fdc-b037-e341706c8d29](http://iec.ch/TC57/CIM100#_ffbabc27-1ccd-4fdc-b037-e341706c8d29)  
 308 Different implementations can import with different xml:base if xml:base is not declared but  
 309 defined by the implementation at the time of the parsing. Implementations can eventually ignore  
 310 xml:base declaration, but this is not defined for specific reason and certain base is required in  
 311 an exchange.  
 312  
 313  
 314

315 - In the IEC 61970-552 we have the header definition embedded with the serialization  
 316 instructions, which makes it complex to transition between different versions of the header  
 317 information. Both CGMES v2.4 and CGMES v3.0 refer to IEC 61970-552 which required to have  
 318 md:FullModel class as a header. This is why when ENTSO-E had to cover additional  
 319 requirements and align with W3C DCAT 3, it was necessary to just add W3C DCAT attributes  
 320 to md:FullModel. In the future, it is expected that the serialization of the header and the rest of  
 321 the instance file are decoupled, which will allow that instance files are using dcat:Dataset as a  
 322 header class instead of md:FullModel. However, in order to realize this, CGMES standards need  
 323 to be updated through the lengthy IEC standardization process.  
 324

## 325 5. Different CIM RDFS versions

326 Some information is already provided above in the section that explains the application profiles.  
 327 Due to historical reasons, the definition of RDF Schema (RDFS) exported for each of the profiles  
 328 also deviates from W3C. There is a strong influence of profiling techniques used by different  
 329 CIM communities, e.g. IOP vendors discussions reflected in EA Add-ins, IEC WG13, IEC WG16,  
 330 etc. There are also cims (CIM scheme) extensions, as defined in IEC 61970-501 standard (in  
 331 2006). Finally, and in addition to the above, the RDF Schema used for profiles generated since  
 332 2010 do not fully follow IEC 61970-501 either, but the implementation has been industry-driven.  
 333 In 2019-2020 there was an effort to prepare draft for the next Edition of 61970-501, but due to  
 334 lack of resources, this work has not been completed yet. Currently CimSyntaxGen<sup>6</sup> supports  
 335 the following exports:

- 336 - RDFS2019 export of profiles relate to RDFS that was used by 2019 (this is industry driven  
 337 implementation which is not documented in either specification or a standard);
- 338 - RDFS2020 export of profiles have change in the Schema header (in 2019 version, the version  
 339 information is serialized as instances of the version class, while in 2020 version this information  
 340 is translated to a header). This is also industry driven implementation – there is no approved  
 341 specification.
- 342 - The beta version of RDFS edition 2 is still a work in progress. This version was a prototype of a  
 343 draft version related to IEC 61970-501 Ed2, which was not finished. The main changes are  
 344 separation of vocabulary description and constraints as well as aligning with W3C of RDF  
 345 scheme definitions. The objective was to eliminate the usage of proprietary CIM namespace

<sup>5</sup> this is what in XSD-governed XML one would call namespace URI; “cim” prefix associated with this namespace URI would be the namespace prefix.

<sup>6</sup> The tool used to generate the RDFS of the profiles.

346 (cims) in the RDFS. This version is available for the purpose of collecting feedback that will be  
347 integrated in the final version.

348 RDFS2020 version is the recommended version to use. It is exported for CGMES v3 profiles,  
349 Network Code profiles and metadata and header profiles. CGMES v2.4 used RDFS version  
350 which was implemented before RDFS 2019

## 351 6. Datatypes and associated issues

352 In Canonical CIM, datatypes are specified with classes marked with stereotype “Primitive”,  
353 “CIMDatatype” and “Compound”. These classes are profiled and assigned to the attributes that  
354 use them. In the RDFS derived from profile definition, there is complete information on the  
355 datatypes and their multipliers<sup>7</sup>. However, in the CIM XML where data is serialized there is no  
356 information on the datatypes. When the data is parsed a standard parser would most probably  
357 assume all attributes’ values as strings and if this data is validated, non-string datatype will be  
358 reported as invalid. Therefore, when implementing data import, developers need to know this  
359 and apply some rules considering the information provided in the RDFS.

360 The same is valid for the units. In RDFS, there is information if, for example, active power value  
361 should be in MW. This information is not explicitly exchanged in the instance data and RDFS  
362 needs to be consulted when mapping/converting the data to the internal data model.

363 Detailed example is provided in Section 4 of this document.

## 364 7. Available Tools

365 There are multiple tools available either for free or under specific license conditions when used  
366 in production enterprise environment. Most of them come with some maintenance support.

### 367 Profiling tools:

- 368 • CimContextor, CimSyntaxgen
- 369 • [CIMTool](#)

370

### 371 W3C RDF Test Suites:

372 Two test suites published by the W3C, a W3C RDF Validation Tool Service  
373 (<https://www.w3.org/RDF/Validator/>) that the W3C School on RDF  
374 ([https://www.w3schools.com/XML/xml\\_rdf.asp](https://www.w3schools.com/XML/xml_rdf.asp)). These test suites are only recommended for  
375 non-confidential data for testing purposes.

- 376 • RDF 1.0 Test Suite <https://www.w3.org/TR/rdf-testcases/>
- 377 • RDF 1.1 Test Suite [https://www.w3.org/TR/2014/NOTE-rdf11-testcases-  
378 20140225/#test-suites-and-implementation-reports](https://www.w3.org/TR/2014/NOTE-rdf11-testcases-20140225/#test-suites-and-implementation-reports)

379

### 380 RDF Tools and Libraries for data processing and validation:

---

<sup>7</sup> Note that this is specific to the profiling technique and tooling used. For instance, this is not the case with CIMTool where there is reduction to just corresponding primitive type and info about Datatype is lost. The same is valid for the units.

381 There are a number of libraries, editors, and databases that are fairly common. This list is not  
 382 extensive or complete but does provide a cross section of tools that a knowledge engineer or  
 383 developer might use in conjunction with CIM Profiling tools.

384

385

**Table 1**

Tool	Type	Comments
<a href="#">Protégé</a> v5.5	Ontology editor	
<a href="#">Apache Jena</a> 4.9.0	Java Library	
<a href="#">ValiMate</a> (put link)	SHACL validation	Free GUI version available. Other versions provided under license conditions.
<a href="#">CimPal</a> (open source) –	RDF conversion and some basic manipulations around RDF; generation of SHACL	Java based, using Apache Jena
PyPi <a href="#">rdflib</a> v7.0	Python Library	Can process exactly 1 RDFS, 1 RDFXML and 1 SHACL at a time; not multiple, which is normally needed.
<a href="#">GraphDB</a> Desktop v10.2.1 (Free)	Triplestore	It can be used for testing, but for enterprise use, license is needed.
<a href="#">Blazegraph</a> v2.1.4	Triplestore	
<a href="#">Neptune Serverless</a> 1.2.0.2	Triplestore	
<a href="#">Topbraid</a> Composer 6.0.1	Ontology editor	Not maintained anymore; they moved to a cloud solution. Still, TopBraid SHACL validation code available from GitHub.
<a href="#">Easy RDF</a>	Web service	For testing purpose.
<a href="#">W3C Validation Service</a>	Web service	For testing purpose.
<a href="#">StarDog</a>	Enterprise Knowledge Graph platform	

<a href="#">RDF4J</a>	Java library	
<a href="#">SESAME</a>	Open source RDF database	
<a href="#">PySHACL</a>	Python library on SHACL	
<a href="#">Neo4J/NeoSemantics</a>	Graph Database	

## 386 8. SHACL based constraints and validation

387 SHACL is a W3C recommendation: [Shapes Constraint Language \(SHACL\) \(w3.org\)](#)

388 SHACL constraints that are published for CGMES v3.0 include the following types of  
389 constraints:

- 390 - Constraints derived from RDFS
  - 391 ○ Cardinality of associations and attributes
  - 392 ○ Datatypes
  - 393 ○ Association ends within profile and cross profile
- 394 - Constraints developed based on descriptions of classes, attributes and associations
- 395 - Constrains derived from the test in the standards such as IEC 61970-301, IEC 61970-  
396 452, etc.
- 397 - Constraints already numbered in standards like IEC 61970-600-1 and IEC 61970-600-  
398 2, etc.

399 The constraints derived from RDFS are created in an automated way using CimPal. All the rest  
400 are pretty much a manual effort. The constraints are maintained under Apache 2 license. There  
401 is still work to be done on ensuring the maintenance process is robust enough.

402 Work on specification to export SHACL constraints from EnterpriseArchitect is planned.

403 The main serialization for SHACL constraints is Turtle as this was tested in the initial  
404 development. Turtle is the most human readable RDF serialization. RDF XML and JSON-LD as  
405 possible serialization but RDF XML was not well tested for SHACL and JSON-LD is under  
406 development.

### 407 8.1. Validation of datasets

408 It is recommended that SHACL constraints are used for RDF data validation. In order to validate  
409 you need to have the instance data, the SHACL constraints and the validation engine. As  
410 explained above SHACL constraints already include constraints derived from the scheme and  
411 custom constraints. Therefore, the RDFS is not needed for the validation.

### 412 8.2. Validation of multiple dependent datasets

413 Validation of multiple datasets if necessary for most business processes. However, it is not  
414 efficient to exchange everything every time and to validate everything every time. It is  
415 recommended that each business process describes a data validation strategy/framework in  
416 order to describe what is validated where. This approach provides direct input to the design of  
417 the SHACL constraints that can be applied on a portion of data. For example, for the CGMES  
418 conformity assessment scheme, DNV as an Assessment Body defined the necessary subset of

419 constraints that are active when validating different use cases and test steps part of the test  
420 use cases.

421