



Recommended Practices for STEP File Compression

Release 1.2

August 15, 2016

Contacts

Jochen Boy
PROSTEP AG
Dolivostraße 11
64293 Darmstadt / Germany
jochen.boy@prostep.com

Phil Rosché
ACCR, LLC.
125 King Charles Circle
Summerville, SC 29485, USA
phil.rosche@accr-llc.com

Frederic Darré
CIMPA
frederic.darre.external@airbus.com

Table of Contents

1	Introduction.....	3
2	Scope.....	4
3	Fundamental Concepts	4
3.1	Compression Algorithm	4
3.2	STEP – ZIP File Structure	4
3.3	Compressed File Handling	5
4	Compression Process.....	5
4.1	Export.....	5
4.2	Import.....	6
4.3	Manual Handling / Compatibility	6
5	Extension for External References	7
5.1	File Name Reference.....	7
5.2	Reminder – Document Format Properties	7
6	STEP Compressed File Example.....	8
7	STEP File Compression Benefits	9
Annex A	References	10

List of Figures

Figure 1: Illustration of STEP interface options	3
Figure 2: Illustration of .STP / .STPZ File Structure	5
Figure 3: Export producing a set of compressed STEP files	6
Figure 4: Import of a set of compressed STEP files	6
Figure 5: Instantiation diagram for Document format Properties	7
Figure 6: Table of recommended Document Format Properties	8
Figure 7: STEP compressed file	8

Document History

Revision	Date	Change
1.0	2013-02-12	Initial release
1.1	2013-11-25	Change recommended file extension back to ".stpZ" (one dot)
1.2	2016-08-15	Added clarification for compression algorithm and file structure; Update file extensions for Business Object Model XML

1 Introduction

With the growing scope covered by STEP due to the latest developments, the average size of STEP files is increasing. The new capabilities included in AP242 “Managed Model Based 3D Engineering”, which will be the core data format in the aerospace and automotive, will accelerate this process even further. Hence, there is a need to decrease the size and optimize the performance of file exchange.

This requirement is generic and will be applicable for data exchange based on all STEP application protocols; foremost AP214, AP203e2, AP209e2, and AP242. The main types of information in focus are: 3D exact geometry, 3D PMI, 3D tessellated geometry.

This document provides a basic approach to compress STEP files, using the same approach as in the IFCZIP agreement of the Industry Foundation Class supported by the Implementer Support Group (see [1]).

The compression method can be applied to Part 21 files (ASCII format) as well as Business Object Model files (XML format). It shall not be used with the Part 26 files (binary format), as it would provide little to no benefit in this case.

In order to distinguish compressed STEP files as per this recommendation, the interface of creation must use the following files extensions:

- “*.stpZ” for compressed Part 21 (ASCII) files, or
- “*.stpxZ” for compressed Business Object Model XML files.

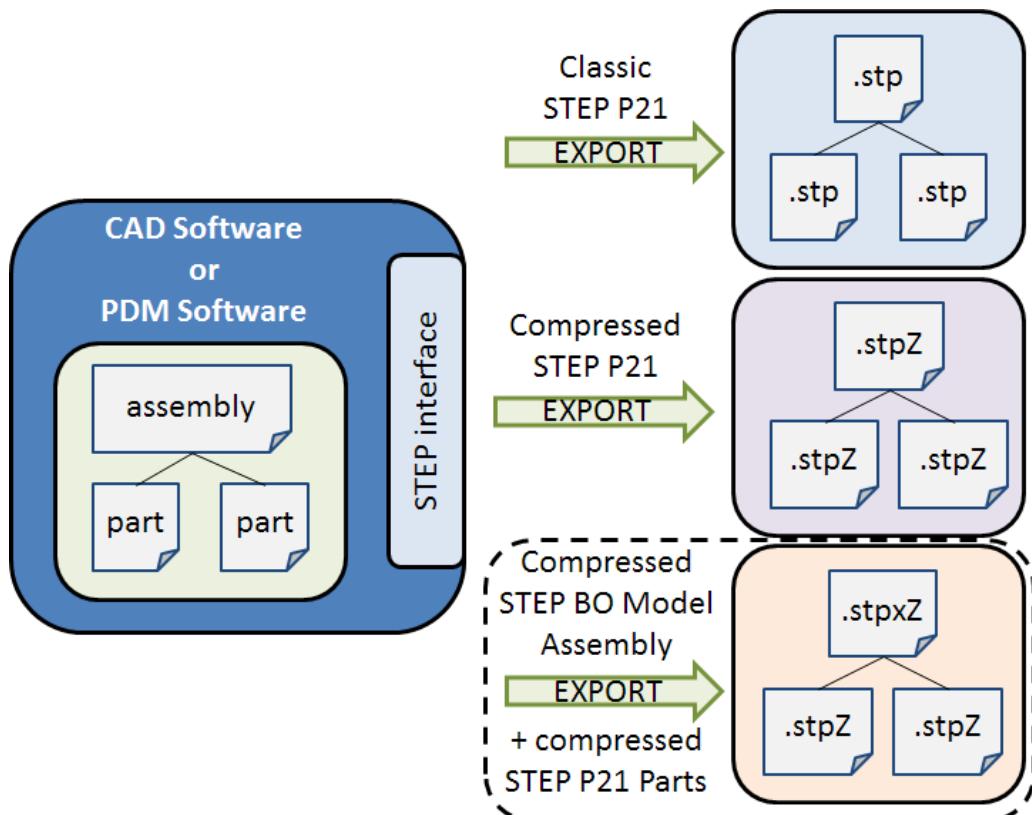


Figure 1: Illustration of STEP interface options

2 Scope

The following are within scope of this document:

- Definition of a compression algorithm to be used for STEP files
- Handling specification for import / export interfaces
- Extension for External References
- Example analysis of STEP file compression benefits

The following are outside the scope of this part:

- Description of STEP information related to a specific AP
- Compression of binary (e.g. Part 26) files.

3 Fundamental Concepts

Particular concepts of interest within the scope of this recommendation are described in the following sections.

3.1 Compression Algorithm

All compressed STEP files shall be written using the ZIP compression algorithm according to PKZIP version 2.05. The ZIP file format is being developed by PKWARE, who release the file format specification in the public domain (see [2]). This compression format is well-established and widely used. It is, among others, compatible with:

- pkzip/pkunzip
- Windows Compressed Folders
- winzip,
- info-zip
- zlib
- Mac OS X (10.3 and higher)

This algorithm is used in Part 21 Edition 3, as well as by the IFC format in the building industry.

The same compression mechanism shall be used for the compression of Part 21 as well as Business Object Model XML files.

Many CAX-IF members however use GZip / Zlib for compression. PKZip libraries can read GZip files; the differences are only in the file header. Both shall be supported on import. PKZip allows for additional features such as folder/file structures (see section 3.2 below). Note that in accordance with Part 21 Edition 3 Annex A.4, the compressed file shall not be encrypted.

3.2 STEP – ZIP File Structure

The most important rules for the STEP file compression concept are:

- There shall be only a single STEP file in each compressed file.
- The name of the STEP file, and the name of the STEP compressed file shall be the same, i.e. "sample_123.stp" will be compressed in "sample_123.stpZ".

Note: Initial draft versions of this document recommended using the file extension “*.stp.Z” (two dots), since that ensured preservation of the correct file name – including the “*.stp” extension –

of the compressed file. However, it was agreed to switch the recommendation to “*.stpZ” (one dot), since that will allow association of this file extension with the correct application to handle STEP files. The disadvantage of losing the file extension when using the basic GZip / Zlib algorithm and manually inflating such a file was agreed to be negligible, since whoever attempts this, knows what to expect.

Note: Part 21 Edition 3 Annex A.4 defines that entire file structures can be included in a compressed file using the PKZip algorithm, and that the root file shall be named “ISO-10303.p21”. The scope of these recommended practices, however, covers the specific needs of long-term archiving and for the transfer of 3D Tessellated Geometry. In these cases, there is always a 1:1 relationship, i.e. one STEP file per compressed file. Hence, the more basic GZip / Zlib can be used, and the “inner” and “outer” file names shall be identical, as defined above.

The compression is applicable to the full product structure, i.e. assembly STEP files (containing product structure information) and part STEP files (containing geometry) will be compressed using the same mechanism.

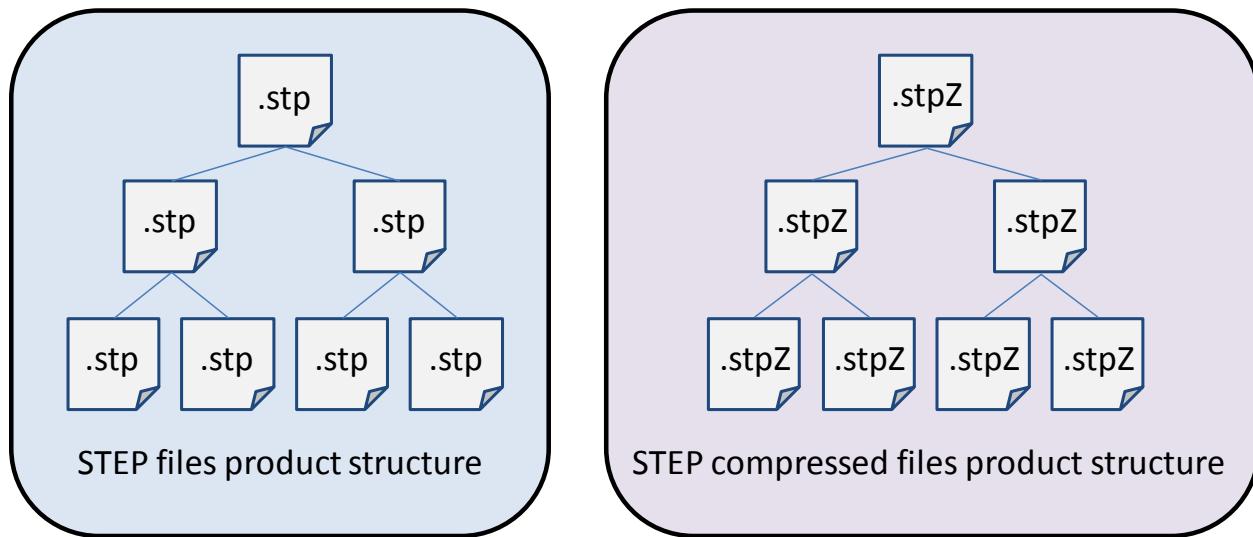


Figure 2: Illustration of .STP / .STPZ File Structure

3.3 Compressed File Handling

The file extension (“*.stpZ” / “*.stpxZ”) can be associated with any STEP consuming software that is capable of automatically decompressing the file upon import.

For STEP processors which are not capable of handling this STEP compressed format, there is the possibility to manually unpack the files using any available unzip tool, and then using the uncompressed STEP files directly with any STEP compatible software (see section 4.3).

4 Compression Process

4.1 Export

The export process is straightforward: The source system’s STEP interface creates in a first step the usual STEP files from the selected data, and then compresses each file to create a “*.stpZ”:

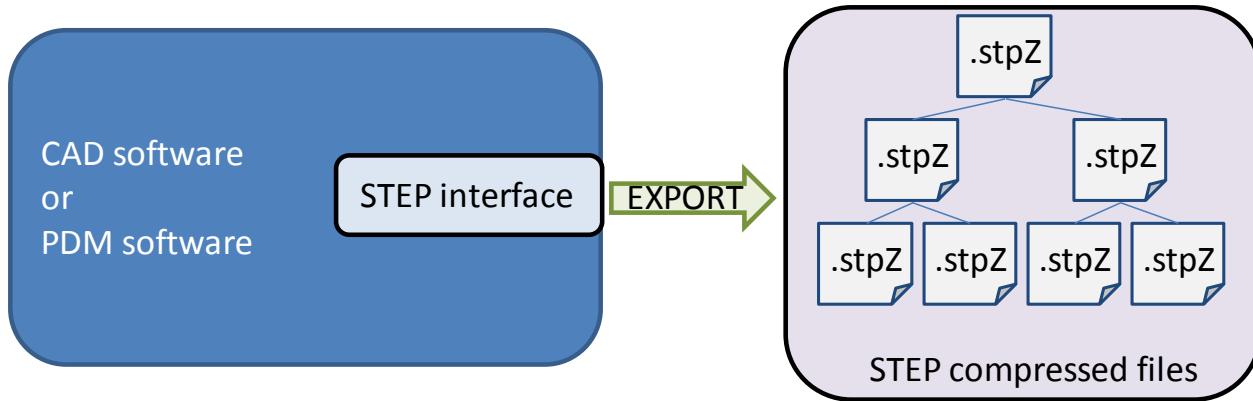


Figure 3: Export producing a set of compressed STEP files

4.2 Import

Just as for classic import, the root file needs to be selected for import. The STEP interface of the target system needs to decompress each “*.stpZ”, and then import and process the contained STEP file, as usual.

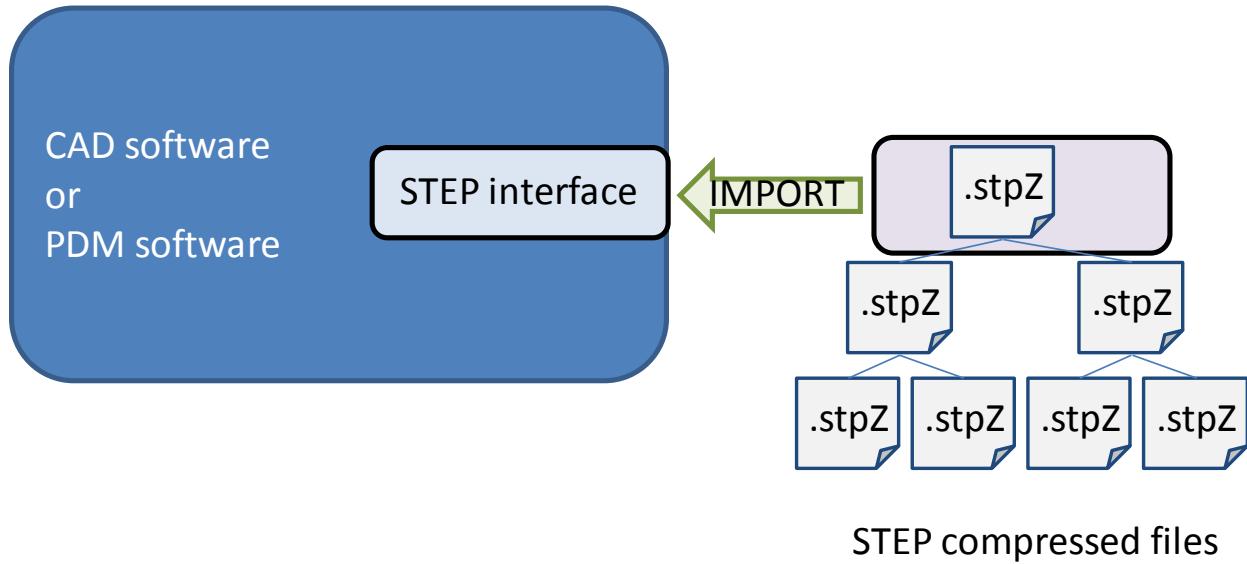


Figure 4: Import of a set of compressed STEP files

The distinction between compressed and uncompressed STEP files will have to be made by the importing processor, based on file extension (“*.stp” / “*.stpZ”) and file type (ASCII / binary).

4.3 Manual Handling / Compatibility

Compression of exported files, and decompression of files to be imported can also be done manually (or as part of a batch process) in case the involved STEP interface cannot handle the compressed files as per this recommendation directly.

After export, using a ZIP tool compatible with the compression algorithm stated above, create an archive for each STEP file. Make sure that there is only one STEP file per archive, and no sub-folders. Then rename the resulting file from “sample_123.zip” or “sample_123.stp.zip” to “sample_123.stpZ” to obtain a compressed file compliant to this recommendation.

For import, open the “*.stpZ” file with a compatible archiving tool, and unpack the enclosed STEP file. Handle this file as usual.

This method can be used as a backup solution, but a direct integration of the compress/uncompress method to the STEP interface is the preferred approach.

5 Extension for External References

This paragraph describes how to use compressed STEP files in the context of External References, which are described in the “Recommended Practices for External References” (see [3]). In this case, it is required to clearly identify when a reference links to a plain STEP file, or a compressed STEP file.

5.1 File Name Reference

It is important to note that, when creating the external references, the name of the referenced external file stored in `applied_external_identification_assignment.assigned_id` in the referencing file (cp. section 2.1 in [3]) shall always be the name of the plain STEP (Part21/ 28) file, and not the name of the compressed file, i.e. the external reference will always point to “sample_123.stp”, regardless of whether it is compressed or not.

The reason is that the compression procedure shall have no impact on the file contents at all. If the STEP files are uncompressed manually first (as described in 4.3), the external references still have to work.

5.2 Reminder – Document Format Properties

In section 2.3 of the aforementioned Recommended Practices, document format properties for the referenced external file are defined as follows:

The Document Format is an additional property which completes the structure displayed in Figure 3 [in [3]], and is attached to the `document_file` entity. It provides information about the data format of this referenced file.

The format of the referenced file is stored in the `descriptive_representation_item.description` (see Figure 5):

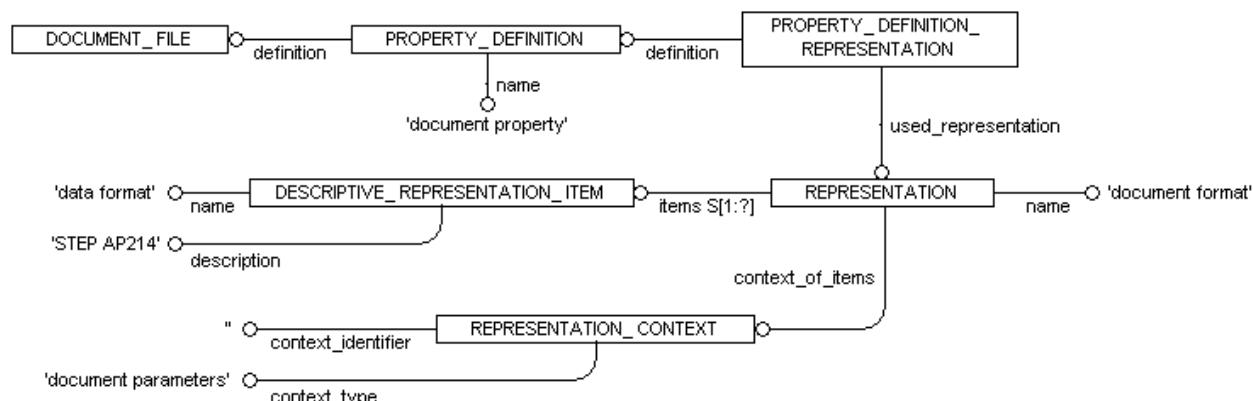


Figure 5: Instantiation diagram for Document format Properties

The list given in Figure 6 extends the list of recommended values for the document format property as defined in [3] with applicable values for the new AP242. Unless noted otherwise, the document format property makes no distinction between assembly and geometry files.

Note also that, as the stored name of the referenced file will always be that of the plain (uncompressed) file as mentioned in section 5.1 above. The same applies for the document format properties; no distinction will be made between compressed and uncompressed STEP File. The reason is that manual handling of the files on either end of the process would lead to inconsistencies.

The importing STEP processor can still unambiguously detect the file format by the file extension ("*.stp" or "*.stpZ") together with the fact the file is either an ASCII or binary file.

Format property for plain file	Description of referenced file
'ISO 10303-203', 'STEP AP203'	Eventually followed by the release number (E2, E3): The document contains data in ISO 10303-203 Part21 format
'ISO 10303-214', 'STEP AP214'	Eventually followed by the release number (E2,E3): The document contains data in ISO 10303-214 Part21 format
'STEP AP214 CC06'	The document contains data in ISO 10303-214 Part21 format according to Conformance Class 06 (product structure only, the file contains no geometry, but references to external geometry files)
'ISO 10303-242', 'STEP AP242'	The document contains data in ISO 10303-242 Part 21 format
'ISO 10303-242 BO Model XML'	The document contains data in ISO 10303-242 XML format

Figure 6: Table of recommended Document Format Properties

6 STEP Compressed File Example

This section presents a STEP compressed file using AP214. The AP214 file in Part 21 format is compressed inside the STEP compressed file:

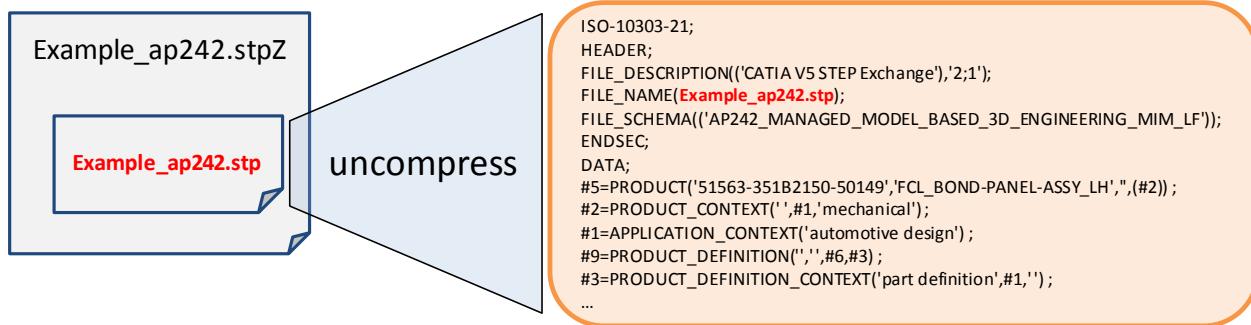


Figure 7: STEP compressed file

The following Part 21 excerpt is taken from the Recommended Practices for External References, and highlights the references to the file illustrated above:

```
#18 = SHAPE REPRESENTATION('',( #17 ),#8 );
#26 = PRODUCT('10001','L-BRACKET','NOT SPECIFIED',( #25 ));
#28 = PRODUCT_DEFINITION_FORMATION('','/ANY',#26 );
#29 = PRODUCT_DEFINITION('design','','#28,#24 );
#30 = PRODUCT_DEFINITION_SHAPE('','SHAPE FOR L-BRACKET',#29 );
#31 = SHAPE_DEFINITION REPRESENTATION(#30,#18 );
[...]
#48 = DOCUMENT_TYPE('');
#49 = DOCUMENT_FILE('l-bracket_prt.stp','','','',#48,'',$);
#51 = PROPERTY_DEFINITION('external definition','','#49 );
#52 = PROPERTY_DEFINITION REPRESENTATION(#51,#18 );
#53 = EXTERNAL_SOURCE(IDENTIFIER(''));
#54 = IDENTIFICATION_ROLE('external document id and location',$);
#55 = APPLIED_EXTERNAL_IDENTIFICATION_ASSIGNMENT
    ('Example_ap214.stp',#54,#53,( #49));
[...]
#428 = DOCUMENT_REPRESENTATION_TYPE('digital',#49 );
#433 = PRODUCT RELATED_PRODUCT_CATEGORY('part',$, (#26,#72,#117,#159,
    #213,#245,#300,#374,#402));
#455 = APPLIED_DOCUMENT_REFERENCE(#49,'',( #29));
#456 = DOCUMENT('','',$,#458);
#458 = DOCUMENT_TYPE('configuration controlled document version');
#459 = DOCUMENT_PRODUCT_EQUIVALENCE('equivalence',$,#456,#440);
#460 = ROLE_ASSOCIATION(#461,#455);
#461 = OBJECT_ROLE('mandatory',$);
[...]
#856 = PROPERTY_DEFINITION('document property','','#49 );
#857 = PROPERTY_DEFINITION REPRESENTATION(#1856,#1855 );
#855 = REPRESENTATION('document format',( #1853 ),#1854 );
#853 = DESCRIPTIVE REPRESENTATION_ITEM('data format','STEP AP214' );
#854 = REPRESENTATION_CONTEXT('','document parameters');
```

7 STEP File Compression Benefits

Based on experience and depending on the file contents, a compressed STEP file is between 15% and 40% the size of the uncompressed original STEP file.

The following table represents test cases from CATIA V5 R20 converted to STEP AP214 with Geometric Validation Properties.

Test Case	Native File Size	Plain STEP File Size	Compressed STEP File Size
Low complexity	1.73 MB	2.2 MB	311 kB
Medium complexity	37.9 MB	57.1 MB	8.1 MB
High complexity	60.2 MB	172.3 MB	26.7 MB
Assembly (12 elements)	82 MB	188 MB	37.6 MB

Annex A References

- [1] Agreement on the compression mechanism for ifcZIP:
<http://www.buildingsmart-tech.org/implementation/ifc-implementation/ifc-impl-agreements/cv-2x3-154>
- [2] ZIP Application Note (File format specification):
<http://www.pkware.com/support/zip-app-note/>
- [3] CAx-IF References for External References, v2.1:
http://www.cax-if.de/documents/rec_prac_ext_ref_v21.pdf