ISO/IEC JTC 1 N 9046

2008-05-01

ISO/IEC JTC 1
Information Technology

Document Type:	**Proposed NP**
Document Title:	**SC 34 Proposal for New Work Item on OpenDocument Format (ISO/IEC 26300) / Office Open XML (ISO/IEC 29500) Translation**
Document Source:	**SC 34 Secretariat**
Reference:	
Document Status:	**This document is circulated to JTC 1 National Bodies for concurrent review. If the JTC 1 Secretariat receives no objections to this proposal by the due date indicated, we will so inform the SC 34 Secretariat.**
Action ID:	**ACT**
Due Date:	**2008-08-01**
No. of Pages:	**52**

Secretariat, ISO/IEC JTC 1, American National Standards Institute, 25 West 43rd Street, New York, NY 10036; Telephone: 1 212 642 4932; Facsimile: 1 212 840 2298; Email: lrajchel@ansi.org

ISO/IEC JTC 1/SC 34 N 1035

DATE: 2008-04-28

ISO/IEC JTC 1/SC 34
Document Description and Processing Languages
Secretariat: Japan (JISC)

DOC. TYPE		Text for NP
TITLE		Proposal for New Work Item on OpenDocument Format (ISO/IEC 26300) / Office Open XML (ISO/IEC 29500) Translation
SOURCE		National Body of Germany
PROJECT		
STATUS		This document is circulated to the SC 34 members for a three-month ballot and to JTC 1 for concurrent review. The working draft submitted by the proposer is appended for information. SC 34 members are requested to vote and comment by 2008-07-28 via the ISO committee internal balloting system on the ISO/TC server.
ACTION ID		LB
DUE DATE		2008-07-29
DISTRIBUTION		P, O and L Members of ISO/IEC JTC 1/SC 34 ; ISO/IEC JTC 1 Secretariat; ISO/IEC ITTF
ACCESS LEVEL		Open
ISSUE NO.		22
FILE	**NAME**	1035c.htm 1035.pdf
	SIZE (KB)	1 47
	PAGES	

Secretariat ISO/IEC JTC 1/SC 34 - IPSJ/ITSCJ (Information Processing Society of Japan/Information Technology Standards Commission of Japan)* Room 308-3, Kikai-Shinko-Kaikan Bldg., 3-5-8, Shiba-Koen, Minato-ku, Tokyo 105-0011 Japan *Standard Organization Accredited by JISC
Telephone: +81-3-3431-2808; Facsimile: +81-3-3431-6493; E-mail: kimura @ itscj.ipsj.or.jp

---

1035.pdf

New Work Item Proposal

February 2004

PROPOSAL FOR A NEW WORK ITEM

Date of presentation of proposal: 2008-04-06	Proposer: Germany DIN, NIA-34
Secretariat: JISC National Body of Japan	**ISO/IEC JTC 1 N** ISO/IEC JTC 1/SC 34 N 1035

A proposal for a new work item shall be submitted to the secretariat of the ISO/IEC joint technical committee concerned with a copy to the ISO Central Secretariat.

Presentation of the proposal - to be completed by the proposer. .

Title (subject to be covered and type of standard, e.g. terminology, method of test, performance requirements, etc.)

OpenDocument Format (ISO/IEC 26300) / Office Open XML (ISO/IEC 29500) Translation

Scope (and field of application)

The report will start by studying common use cases to identify how the most important functionalities of one document format can be represented in the other format. The technical report will do a thorough review of the concepts, architectures and various features of the two document formats in order to provide a good understanding of the commonalities and differences. We expect that functionalities will be able to be translated with different degrees of fidelity to the other format. We will provide, for each functionality, detailed information on the extent of which that functionality can be translated. This report is a necessary step to the goal of helping achieve interoperability and harmonization between the two formats.

Purpose and justification - attach a separate page as annex, if necessary

Programme of work

If the proposed new work item is approved, which of the following document(s) is (are) expected to be developed?

____ a single International Standard

___ more than one International Standard (expected number: .....)

____ a multi-part International Standard consisting of ..... parts

____ an amendment or amendments to the following International Standard(s) .....

__X_ a technical report , type ..3....

And which standard development track is recommended for the approved new work item?

____a. Default Timeframe

__X_b. Accelerated Timeframe

____c. Extended Timeframe

Relevant documents to be considered

Attached document: ISO-IEC_TR_Translation(E)_v081.doc

Co-operation and liaison

AFNOR (AFNOR is launching an initiative in order to work with the DIN on Office Open XML and ODF translation and interoperability) and Ecma TC45.

Preparatory work offered with target date(s)

Signature:

Will the service of a maintenance agency or registration authority be required? .....No.....
- If yes, have you identified a potential candidate? .....
- If yes, indicate name .....

Are there any known requirements for coding? .....No.....
-If yes, please specify on a separate page

Does the proposed standard concern known patented items? .....No.....
- If yes, please provide full information in an annex

Are there any known accessibility requirements and/or dependencies (see: http://www.jtc1access.org)?...No.

-If yes, please specify on a separate page

Are there any known requirements for cultural and linguistic adaptability?.....No.....

-If yes, please specify on a separate page

Comments and recommendations of the JTC 1 or SC 34 Secretariat - attach a separate page as an annex, if necessary

Comments with respect to the proposal in general, and recommendations thereon:
It is proposed to assign this new item to JTC 1/SC 34

Voting on the proposal - Each P-member of the ISO/IEC joint technical committee has an obligation to vote within the time limits laid down (normally three months after the date of circulation).

Date of circulation:	**Closing date for voting:**	**Signature of Secretary:**
2008-04-29	2008-07-29	Toshiko KIMURA

NEW WORK ITEM PROPOSAL - PROJECT ACCEPTANCE CRITERIA		
Criterion	**Validity**	**Explanation**
A. Business Requirement		
A.1 Market Requirement	Essential _X_ Desirable ___ Supportive ___	Identify in detail the differences between ODF and Open XML that can help harmonization and interoperability. The report will bring more clarity about the translatability of OOXML and ODF; the target group are users and template developers. Once the differences are found, further evolution of both formats can be done to improve interoperability.

A.2 Regulatory Context	Essential ___ Desirable ___ Supportive _X_ Not Relevant __	One outcome would be to understand at which level different subsets of functionalities used by ISO/IEC 29500 and ISO/IEC 26300 could be translated; governments as well as organisations like IDABC are looking for this information that is extremely important for interoperability and for the further evolution of both formats.
B. Related Work		
B.1 Completion/Maintenance of current standards	Yes ___ No_X_	
B.2 Commitment to other organisation	Yes _X_ No___	AFNOR (AFNOR is launching its initiative in order to work with the DIN on Office Open XML and ODF interoperability)
B.3 Other Source of standards	Yes ___ No_X_	
C. Technical Status		
C.1 Mature Technology	Yes ___ No_X_	
C.2 Prospective Technology	Yes ___ No_X_	
C.3 Models/Tools	Yes ___ No_X_	
D. Conformity Assessment and Interoperability		
D.1 Conformity Assessment	Yes ___ No_X_	
D.2 Interoperability	Yes _X__ No___	This is the central topic of the NWP
E. Adaptability to Culture, Language, Human Functioning and Context of Use		
E.1 Cultural and Linguistic Adaptability	**Yes**_____ **No__X___**	

E.2 Adaptability to Human Functioning and Context of Use	**Yes_____** **No___X___**	
F. Other Justification		

Notes to Proforma

A. Business Relevance. That which identifies market place relevance in terms of what problem is being solved and or need being addressed.

A.1 Market Requirement. When submitting a NP, the proposer shall identify the nature of the Market Requirement, assessing the extent to which it is essential, desirable or merely supportive of some other project.

A.2 Technical Regulation. If a Regulatory requirement is deemed to exist - e.g. for an area of public concern e.g. Information Security, Data protection, potentially leading to regulatory/public interest action based on the use of this voluntary international standard - the proposer shall identify this here.

B. Related Work. Aspects of the relationship of this NP to other areas of standardisation work shall be identified in this section.

B.1 Competition/Maintenance. If this NP is concerned with completing or maintaining existing standards, those concerned shall be identified here.

B.2 External Commitment. Groups, bodies, or for external to JTC 1 to which a commitment has been made by JTC for Co-operation and or collaboration on this NP shall be identified here.

B.3 External Std/Specification. If other activities creating standards or specifications in this topic area are known to exist or be planned, and which might be available to JTC 1 as PAS, they shall be identified here.

C. Technical Status. The proposer shall indicate here an assessment of the extent to which the proposed standard is supported by current technology.

C.1 Mature Technology. Indicate here the extent to which the technology is reasonably stable and ripe for standardisation.

C.2 Prospective Technology. If the NP is anticipatory in nature based on expected or forecasted need, this shall be indicated here.

C.3 Models/Tools. If the NP relates to the creation of supportive reference models or tools, this shall be indicated here.

D. Conformity Assessment and Interoperability Any other aspects of background information justifying this NP shall be indicated here.

D.1 Indicate here if Conformity Assessment is relevant to your project. If so, indicate how it is addressed in your project plan.

D.2 Indicate here if Interoperability is relevant to your project. If so, indicate how it is addressed in your project plan

E. **Adaptability to Culture, Language, Human Functioning and Context of Use**

NOTE: The following criteria do not mandate any feature for adaptability to culture, language, human functioning or context of use. The following criteria require that if any features are provided for adapting to culture, language, human functioning or context of use by the new Work Item proposal, then the proposer is required to identify these features.

E.1 Cultural and Linguistic Adaptability. Indicate here if cultural and natural language adaptability is applicable to your project. If so, indicate how it is addressed in your project plan.

ISO/IEC TR 19764 (Guidelines, methodology, and reference criteria for cultural and linguistic adaptability in information technology products) now defines it in a simplified way:

“ability for a product, while keeping its portability and interoperability properties, to:

- be internationalized, that is, be adapted to the special characteristics of natural languages and the commonly accepted rules for their se, or of cultures in a given geographical region;

- take into account the usual needs of any category of users, with the exception of specific needs related to physical constraints”

Examples of characteristics of natural languages are: national characters and associated elements (such as hyphens, dashes, and punctuation marks), writing systems, correct transformation of characters, dates and measures, sorting and searching rules, coding of national entities (such as country and currency codes), presentation of telephone numbers and keyboard layouts. Related terms are localization, jurisdiction and multilingualism.

E.2 Adaptability to Human Functioning and Context of Use. Indicate here whether the proposed standard takes into account diverse human functioning and diverse contexts of use. If so, indicate how it is addressed in your project plan.

NOTE:

1. Human functioning is defined by the World Health Organization at http://www3.who.int/icf/beginners/bg.pdf as:
   <<In ICF (*International Classification of Functioning, Disability and Health), the term functioning refers to all body functions, activities and participation.>>*
2. Content of use is defined in ISO 9241-11:1998 (*Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability) as:*
   <<Users, tasks, equipment (hardware, software and materials), and the physical and societal environments in which a product is used.>>
3. Guidance for Standard Developers to address the needs of older persons and persons with disabilities).

F. Other Justification Any other aspects of background information justifying this NP shall be indicated here

ISO/IEC JTC 1/SC 34 N TODO

Date: 2008-01-23

ISO/IEC PDTR 2.2

ISO/IEC JTC 1/SC 34/ WG 1

Secretariat: DIN

OpenDocument Format / Office Open XML Translation

Convertion OpenDocument Format / Office Open XML

Warning

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

Document type: Technical Report
Document subtype:
Document stage: (20) Preparatory
Document language: E

ISO-IEC_TR_Translation(E)_v081.doc STD Version 2.1c2

Copyright notice

DIN German Institute for Standardization
DE-10772 Berlin
Tel. + 49 30 2601 2652
Fax + 49 30 2601 1723
E-mail Jan.Dittberner@din.de
Web http://www.ni.din.de/sc34 (public web site)

Contents

Page

Foreword

We, as the SC34 Mirror in Germany give our especial thanks to:

Fraunhofer FOKUS, Berlin, in person Christopher Taylor, Jan Ziesing, Philip Martin, Majid Gamsari, Gerd Schürmann; Dialogika GmbH, Saarbrücken in person Wolfgang Keber, Dirk Vollmar; Novell OpenOffice Entwicklergruppe, Bremen in person Florian Reuter; Innovimax, Paris, in person Mohamed Zeragaoui; Microsoft Deutschland GmbH, Hamburg, in person Mario Wendt;

for making this TR happen.

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

In exceptional circumstances, the joint technical committee may propose the publication of a Technical Report of one of the following types:

— type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;

— type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;

— type 3, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard (“state of the art”, for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC TR 2, which is a Technical Report of type [1/2/3], was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information Technology*, Subcommittee SC 34, *Document Description and Processing Languages*.

This second edition cancels and replaces the first edition, the formatting of which has been technically revised. Some parts start with the phrase “Text or Describe” or you find content like “TODO”, these are only general remarks done by the authors.

Introduction

OASIS ODF 1.0 (ISO/IEC 26300) and Office Open XML (ECMA-376) are both open formats for saving and exchanging documents, spreadsheets and presentations. Both formats are XML based but are different in design and scope.

OASIS ODF 1.0 has been accepted as standard by the ISO (ISO/IEC 26300) since 2006. Office Open XML was first approved as an ECMA-376 standard in December 2006 by the Ecma International General Assembly and in January 2007 was submitted to the ISO (ISO/IEC DIS 29500) for ISO/IEC standardization.

The report will start by studying common use cases to identify how the most important functionalities of one document format can be represented in the other format. The technical report will do a thorough review of the concepts, architectures and various features of the two document formats in order to provide a good understanding of the commonalities and differences. We expect that functionalities will be able to be translated with different degrees of fidelity to the other format. We will provide, for each functionality, detailed information on the extent of which that functionality can be translated. The goal is to help achieve harmonization and interoperability between the two formats.

Further drafts of the report will be public available using the following link:

http://www.nia.din.de/cmd?level=tpl-artikel&subcommitteeid=94340957&contextid=nia&cmstextid=74610&languageid=en

Text:

Describe: Who is this document targeted at? -> power-users, template designers

Guidelines rather than translation implementation (models / concepts / what features (not) to use, not XML-tags)

Based on ISO Standards (OpenDocument 1.0 and Office Open XML as proposed), not implementations (e.g. MS Word / OpenOffice.org)

Information about the evolution of office document formats (data and layout separating, etc.)

highlights of OOXML and ODF and design objectives

the complexity of mapping functionalities rather than simple tags

abbreviations: ODF → OpenDocument; OOXML → Office Open XML

OpenDocument Format / Office Open XML Translation

1 Use Cases

Top-level use cases can be categorized according to two major dimensions: the Type of the conversion and the functionality required of the conversion result. Some of these use cases can be used to limit the scope of this report, e.g. if the converted document does not need to be revisable but the graphical layout must be reproduced exactly, then conversion to another format, such as PDF or XPS might be more sensible. Use case descriptions will reference section 2 for the description of implied functionalities.

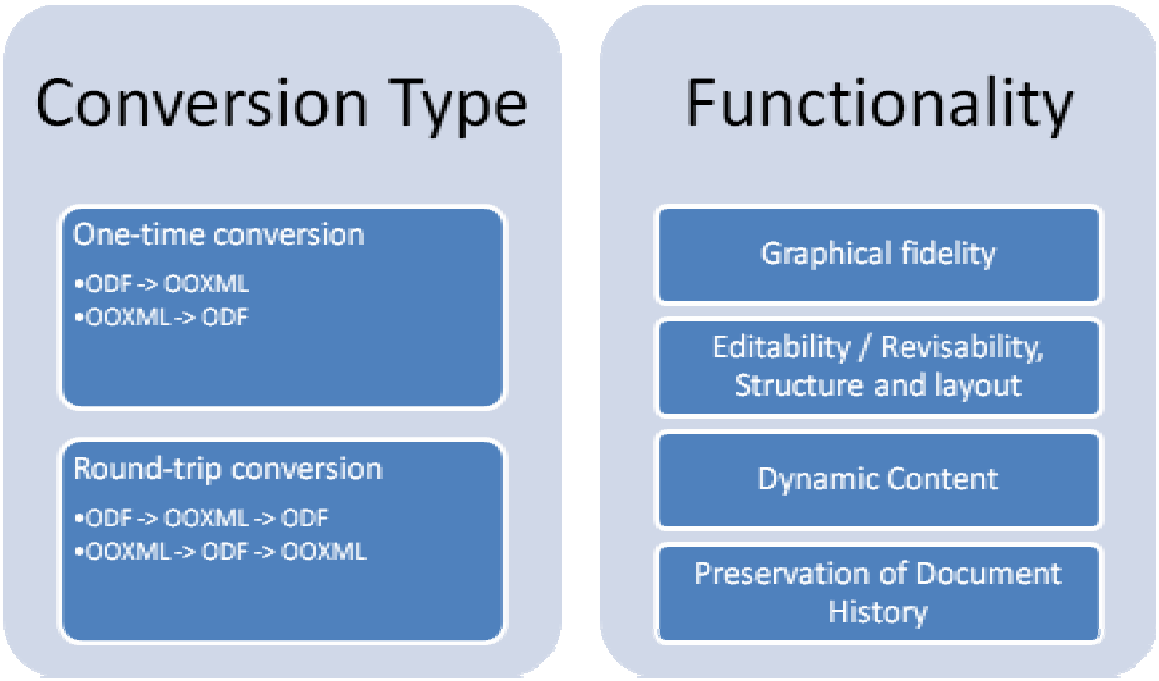


Figure 1 — Use-Case category overview

To classify the different features of office applications their abilities are structured into four different functionality types which are defined as follows.

Graphical fidelity includes all layout and design aspects in office documents and how they are displayed in the different applications. Converting office documents between the different standards must ensure that the original view is retained.

Editability, Revisability, Structure and Layout covers all aspects of structural document features e.g. headlines, tables, listings, captions or indexes. The focus of this functionality in the converting process is how these features are represented in the standards and how they are supported by the different office application. An example for the ideal support of these features is the “outline view” feature in Microsoft Office which makes use of the headlines to simplify the editing of the documents structure. Another example would be a “table wizard” which allows to easily add rows or columns to table structures.

Dynamic Content covers all aspects of automatically generated content, calculations or form functionalities. This could be “current date”-fields, table of contents or field calculations.

Preservation of meta-information includes all additional information besides the actual document content which is used for collaborative functions like change tracking or user notes and to ensure the documents accessibility.

1.1 Text Formatting Documents

1.1.1 Simple Text Formatting

A common sample for the appliance of basic text processing functionality is the issue of a formal business letter. A formal letter should strictly conform to a certain set of guidelines which can be divided into aspects of form and content. Regardless of the text editor used the layout or formal aspects of a business letters should always remain identical to leave a professional mark on the corporate communication.

The following use case describes the issue of a business letter and its conversion between the standards ODF and OpenXML with a special focus on formal aspects. The scenario starts with user John who likes to file a complaint to his preferred airline about a delayed flight which caused some trouble in his agenda. John works on his private laptop with Open Office (ODF Standard) and starts to write the letter which looks like the following sample:

Figure 2 — Use-Case Simple Text Formatting - screenshot of the sample letter

This sample letter makes use of all typical text formatting features. There is centralized text on the top and the date information which is positioned on the right. The receiver address is aligned on the left. The letters body paragraph is in block format. Furthermore the text formatting contains a bold paragraph as the subject line and embedded italic characters in the body text. In the end of this document an image is embedded which represents the subscription of the editor.

After finishing the letter John sends it to his secretary Mary via mail. Mary imports the document to Microsoft Word 2007 (Open XML Standard) to check layout and spelling. Then she mails it to the Customer Complaints Center (CCC) of GoFast Air in London. The receiving agent in the CCC in London works with Star Office (ODF Standard).

This scenario requires layout consistency during multiple conversions of a formal business document. If the actual functionalities of a proper conversion engine could match these requirements is regarded in the tables "text formatting" and "paragraph formatting" in section 2.3 comparing selected aspects. As one can find there simple text formatting like bold or italic characters and paragraph formatting like standard text alignments could be easily converted between the different formats with the exception of so called theme fonts which are not supported in ODF.

1.1.2 Tables and Field Functions

More advanced features of text processing is the usage of tables and predefined field functions for example in invoice summary documents. Regarding the conversion of those documents between the two standards ODF and OpenXML the result has to meet structural, functional and workability related requirements in addition to the fidelity of the graphical representation as seen in use case 1. The following scenario should exemplify these requirements.

The Open Office user John plans to provide an invoice summary draft for his marketers to inform them about their monthly trading results. The following paragraph is a brief excerpt from the documents design:

Real Estate Information Bulletin *Invoice summary 2007* *www.johnmarketer.com/reports*		*Wednesday, 31 October 2007*	
Division & Properties		**Results 2007**	
		1st Quarter	*2nd Quarter*
Division 1	*Rentals*	200.000 $	220.000 $
	Sells	430.000 $	440.000 $
Division 2	*Rentals*	135.000 $	175.000 $
	Sells	390.000 $	376.000 $
Division 3	*Rentals*	310.000 $	340.000 $
	Sells	612.000 $	685.000 $

Figure 3 — Use-Case Tables and predefined functions - screenshot of table formatting

This excerpt shows a table with joined cells and common text formatting. Cells are joined spanning multiple rows and columns. Different cell alignment appears as left, center and right aligned text. Additionally a hyperlink was inserted into the header row of the table.

After filling out the sales report John sends it to his marketers via email. The marketers using different text processing applications are awaiting the figures to spread the news to their own division staff. The text processing applications are Microsoft Word 2007, Microsoft Word 2003, Open Office and Star Office.

As mentioned this scenario requires structural and functional workability to enable the user to edit hyperlinks, table cells and even complex nested tables after converting the documents format. If the actual functionalities of a proper conversion engine could match these requirements is regarded in section 2.3 comparing selected

aspects. According to the statements there the conversion of table structures between OOXML and ODF is supported in most cases. Problems appear using table background patterns, background images or sub-tables which are not supported in OOXML. See section 2.3.3 for detailed information.

1.1.3 Itemization and Numeration

Beside table functionality one can find another important feature for office documents which is of common use regarding the presentation of structured information. Numberings and Listings are often adopted in non-fictional texts like technical documentations. Let us imagine John gets a documentation paper from his technical department about how to logon to his new workstation. The document was created in Microsoft Office 2007 and is describing the related tasks in a few steps. As John receives the paper he is going to open it using his Open Office application.

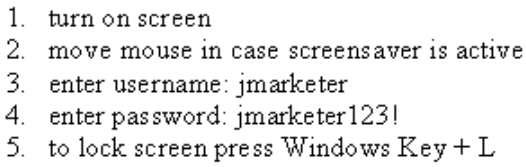

1. turn on screen
2. move mouse in case screensaver is active
3. enter username: jmarketer
4. enter password: jmarketer123!
5. to lock screen press Windows Key + L

Figure 4 — Screenshot of numbered items

A conversion engine between ODF and OOXML should be able to keep the structural order in the numbering and listing parts of the document. If these features could be fully matched between the different standards is regarded in section 2.3.5 “Itemization and Numeration”.

The issue of converting Itemization and Numeration between the standards ODF and OOXML is regarded in section 2.3.5 more in detail. Generally could be determined that there is a high translatability of these structural features between ODF and OOXML.

1.1.4 Indexes

Additionally to continuous text, structural and layout features large documents contain indexes and registers to ease their readability and to make them human searchable at all. Indexes should show up the documents structure based on headlines and page numbers as well as figures and tables based on their captions. Further on all indexes should be generated automatically by the certain text processing application and kept updated as well if changes are applied regarding headline, caption or page number. The following screenshot is an example index of a market report John downloaded from the internet in OOXML format.

Figure 5 — Screenshot of auto-generated document index

After opening it with his Open Office application he likes to cut out different chapters to generate a condensed version of the document to spread it to his colleagues. The table of content should be adapted automatically which means deleted chapters should not appear anymore in the index and the new page numbers of the remaining parts should be updated.

The main question in this scenario is if the table of content from the OOXML document could be converted into the ODF format while fully keeping its structure and meta-information.

Text: Generally could be determined that … A more detailed view is presented in section … TODO

1.1.5 Metadata and Accessibility

To ensure the accessibility of office documents certain additional information must be kept as metadata. One of the important features is the textual description of embedded multimedia data in case the user's application is not able to display it.

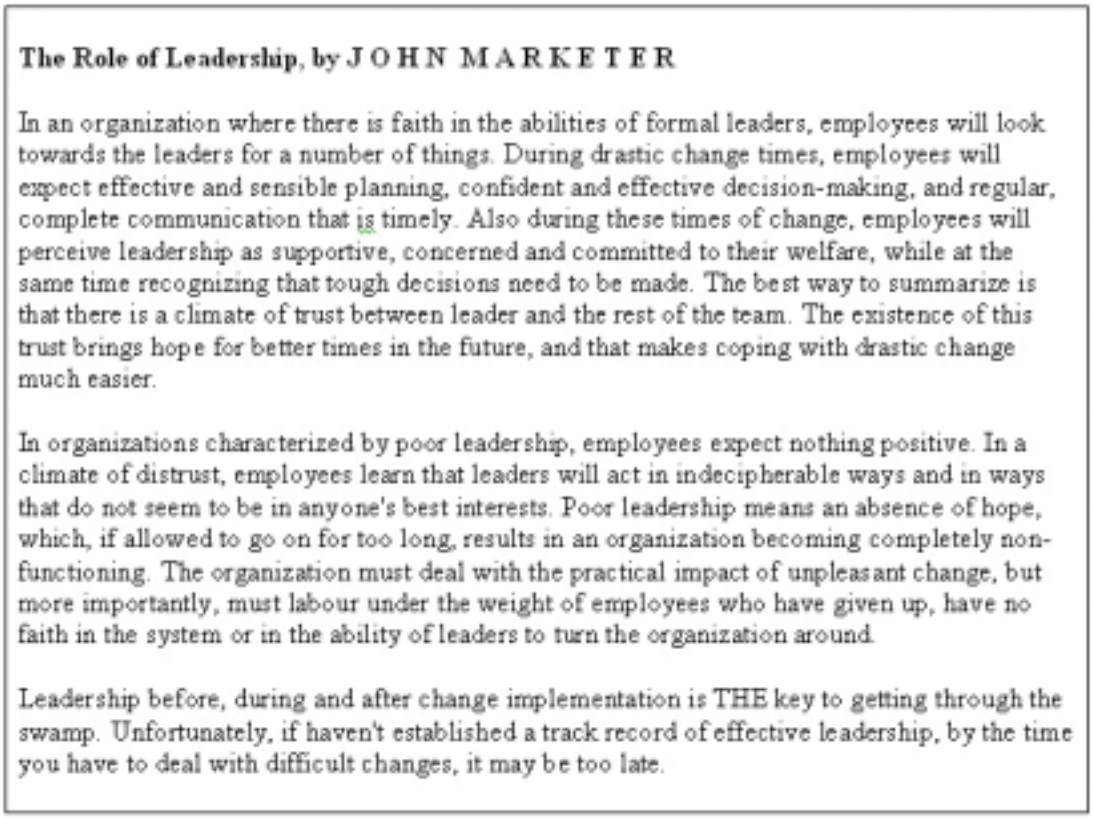
Text: Sprache EN / DE in Dokumenten

Text: Generally could be determined that … A more detailed view is presented in section … TODO

1.1.6 Change Tracking and Collaborative Functions

One of the most important features for editing large documents with multiple authors are called collaborative functions including user specific commenting and tracking of changes. These functions enable collaborative workflows allowing document editing and reviewing with multiple participants. The information including user data, notes or tracked changes which is required for this workflow is embedded within the document as meta-data. Beside the preservation of graphical and structural fidelity while converting between ODF and OpenXML the proper adoption of meta-information plays an important role in collaborative authoring processes. The following scenario illustrates how the collaboration between different authors using different text processors should work.

The OpenOffice User John Marketer and some of his partners are planning to launch an article in the online magazine "OpenBusinessMag". Besides John Marketer another person will be involved into the authoring process. The first draft of the document will be provided by John himself. The following screenshot illustrates the initial version of the article:



The Role of Leadership, by J O H N M A R K E T E R

In an organization where there is faith in the abilities of formal leaders, employees will look towards the leaders for a number of things. During drastic change times, employees will expect effective and sensible planning, confident and effective decision-making, and regular, complete communication that is timely. Also during these times of change, employees will perceive leadership as supportive, concerned and committed to their welfare, while at the same time recognizing that tough decisions need to be made. The best way to summarize is that there is a climate of trust between leader and the rest of the team. The existence of this trust brings hope for better times in the future, and that makes coping with drastic change much easier.

In organizations characterized by poor leadership, employees expect nothing positive. In a climate of distrust, employees learn that leaders will act in indecipherable ways and in ways that do not seem to be in anyone's best interests. Poor leadership means an absence of hope, which, if allowed to go on for too long, results in an organization becoming completely non-functioning. The organization must deal with the practical impact of unpleasant change, but more importantly, must labour under the weight of employees who have given up, have no faith in the system or in the ability of leaders to turn the organization around.

Leadership before, during and after change implementation is THE key to getting through the swamp. Unfortunately, if haven't established a track record of effective leadership, by the time you have to deal with difficult changes, it may be too late.

Figure 6 — Use-Case Collaborative functions and metadata - screenshot of continuous text

John sends this document to his co-author with request for comments. The co-authors office suite is Microsoft Office 2007 using the conversion feature to import the ODF text file. Using the Microsoft Word commenting and change tracking feature the co-authors reviews the document. The comments which are shown as colored boxes on the right margin can be applied on paragraphs, words and even on single characters. The comments are user specific containing the users initials and appearing in different colors for different users. The change tracking function highlights added, edited or deleted text parts and shows the obsolete text parts in red comment boxes. The following screenshot shows how the revised document looks like in Microsoft Word 2007:

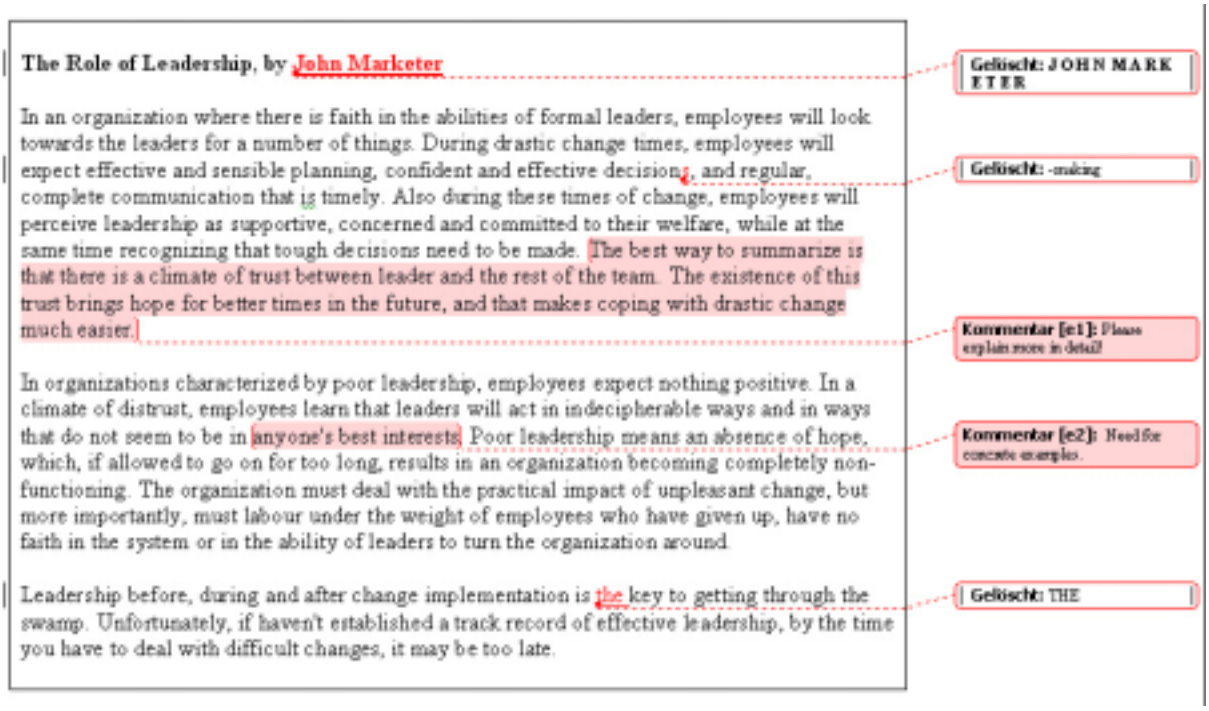


Figure 7 — Use-Case Collaborative functions and metadata II - screenshot of continuous text with markup

After the co-author had sent back the revised version of the article John can start to adapt his text regarding comments and proposed changes. The comments are simple text notes and have to be applied manually. The tracked changes could be applied automatically for the whole document if John agrees with the revision.

This kind of application supported workflow highly alleviates the revision of documents with multiple authors. The foundation for this document lifecycle is the proper conversion of metadata from one standard into the other to keep comments and proposed changes.

The table “change tracking and document revision” in section 2.3.9 goes into the details of how collaborative functions could be used between the different document formats OOXML and ODF. As one can see there both document formats offer support for revision for example in the change tracking feature. While OpenDocument only records the fact that a change has occurred OOXML records the full history of changes to ensure the reconstruction of previous text version. Another minor difference is in the understanding of text comments, while Office Open XML allows adding comments to arbitrary text ranges this feature is not supported by OpenDocument. However similar functionality may be provided by inserting Notes (which are associated with a point in the text, not a range).

1.1.7 Forms

Modern office documents are more and more directly integrated into electronic workflows. On the one hand they are serving as static output formats for reports or certificates. But due to extended form functionalities they can be integrated as dynamic, data driven front-ends as well. The following scenario illustrates some simple features which are commonly in use.

John Marketer's company has optimized its internal processes for years. Paper based workflows are rare now and almost all corporate forms are digitalized each in different formats gathered on an internal website. The application for leave is one simple example and illustrated by the following screenshot:

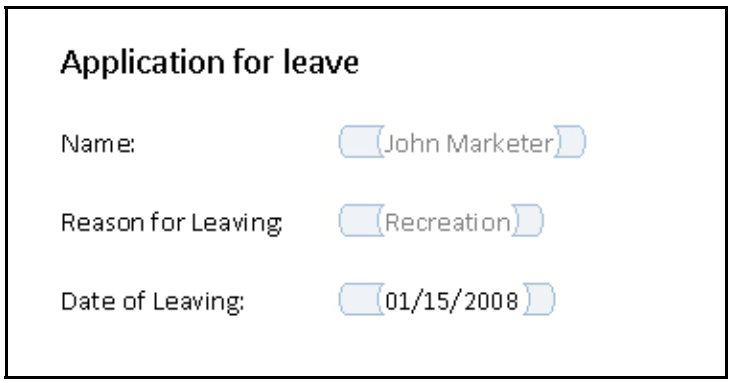


Figure 8 — Screenshot of office form

As one can see here the form contains different textboxes and a pre-selected date field. The form is supposed to be filled by the applicant, to be saved and then to be sent via email for approving. To pass this between different office applications based on ODF or OOXML the forms functionality has to be preserved.

Text: XForms vs. Custom Schemas

Orientierung in Unterschiedliche aspekte des form handlings

SOA kompatibilität

Smart docs

1.1.8 Vector Graphic Formats

Betrifft die verschiedenen unterstützen vektorgarfikformate

1.1.9 Font Metrics and MS C-Fonts

Infos von Florian

1.1.10 Generic Fields

Kommt von Florian

1.1.11 Equations

MathML, etc. kommt von Mohamed

1.2 Spreadsheet Documents

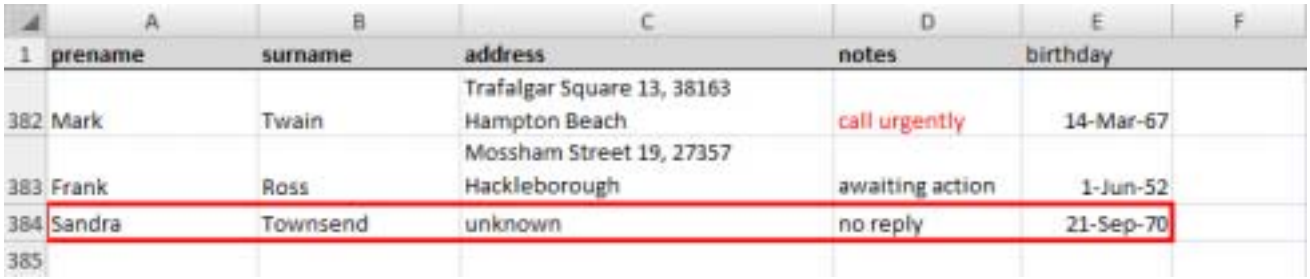
1.2.1 Simple Text Formatting and Graphics

Weiterer use case zu text formatting und Einbindung von Grafiken, diese features basieren aus benutzersicht nicht auf dem text processing teil

1.2.2 Listing, Structural Features and Layout

One of the main applications for spreadsheet documents is the simple listing and structuring of large amounts of data in sortable tables. The conversion of these tables must preserve beside the structural fidelity the graphical representation as well. Graphical features could be frames, shading and colors used for highlighting and structuring certain parts of the document. The following scenario should illustrate the most important functionalities used in listing documents.

John Marketer makes use of a spreadsheet document to store contact information of his personal clients using Open Office on his private laptop. The table has 4 columns and about 400 entries with names, addresses and birthdays. The top row contains the title of the columns called prename, surname, address and notes. To make the sheet navigable easier the top row is fixed and will not move while scrolling down the rows. The screenshot shows an excerpt from the spreadsheet.



	A	B	C	D	E	F
1	prename	surname	address	notes	birthday	
382	Mark	Twain	Trafalgar Square 13, 38163 Hampton Beach	call urgently	14-Mar-67	
383	Frank	Ross	Mossham Street 19, 27357 Hackleborough	awaiting action	1-Jun-52	
384	Sandra	Townsend	unknown	no reply	21-Sep-70	
385						

Figure 9 — Screenshot of spreadsheet address list

The graphical characteristics of this excerpt are the fixed top row, the grey shade of the top row, the colored text in a single cell and the highlighting colored frame on a complete row. Additionally the last column has a date formatting which formats any entry into the date pattern. Open Office user John Marketer sends this document to his secretary Mary to update the customer database manually. Mary is using Microsoft Office on her workstation so she has to import the ODF document.

Text: Aussagen aus der Tabelle …

1.2.3 Formulas and Calculation

Spreadsheet documents were not only designed for the purpose of listing data, it is also a powerful tool for complex and dynamic calculations. Within a spreadsheet document each cell could be provided with a formula able to include other cells by referencing them through row and column numbers. The conversion of calculation spreadsheets should preserve the correct behavior of the provided formula logic and as well the correct presentation of the layout itself. The following use case provides an example to illustrate the most common application.

John is an Open Office user working for a big marketing services company. The IT department of John's company provides spreadsheet templates to the employees enabling them to place orders for their demand on new computer equipment. The main template is designed using Microsoft Office 2007 and looks like the following screenshots shows:

Product ID	Description	Quantity	Unit Price	Line Total
A1234	Flatscreen	1	275.00	275.00
A1235	Optical Mouse	1	15.00	15.00
A1236	Keyboard Standard	1	17.00	17.00

	SUBTOTAL	307.00
PST	6,50%	-
GST	3,20%	-
SHIPPING & HANDLING		-
	TOTAL	307.00
	PAID	-
	TOTAL DUE	307.00

Figure 10 — Screenshot of spreadsheet based invoice template

As one can see the most important part of this document consists of a table for the invoice line items and a self accumulating field for the total costs of the ordered items. As soon as a new line tem is added, the total due field is updated automatically. This functionality should be preserved while John is working on his order using his Open Office application.

Text:

Formeln sind Appliaktionsabhängig

Optionen:

Eigene Sprache verwenden

Auf spezifische Sprachen mappen

Auf einheitlichen standard warten

Problem liegt beim Benutzer

1.2.4 Embedded Spreadsheet Documents

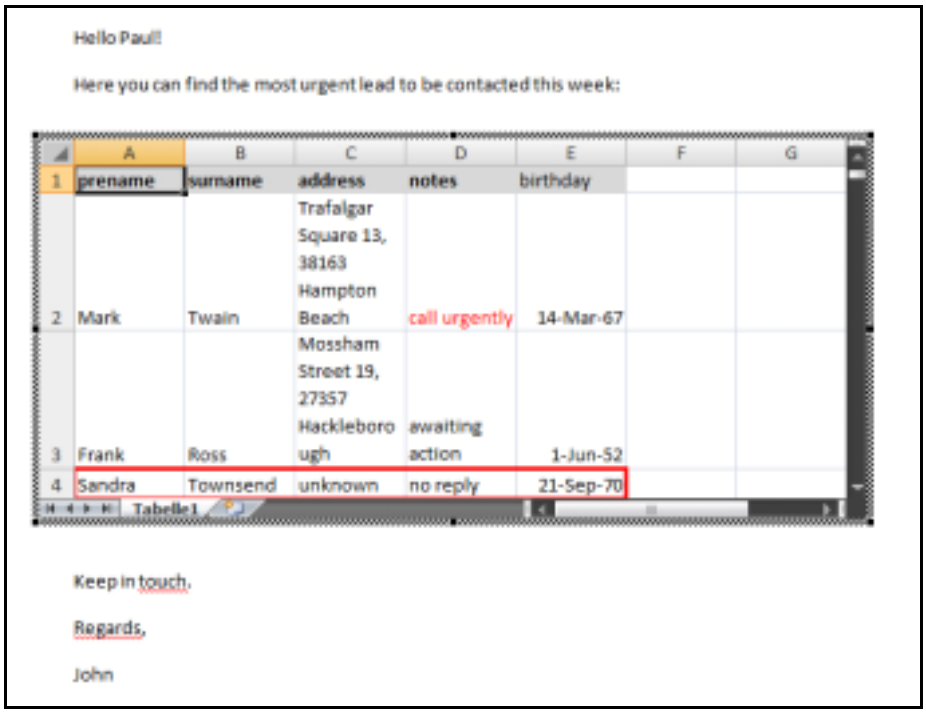


Figure 11 — Screenshot of embedded spreadsheet in text document

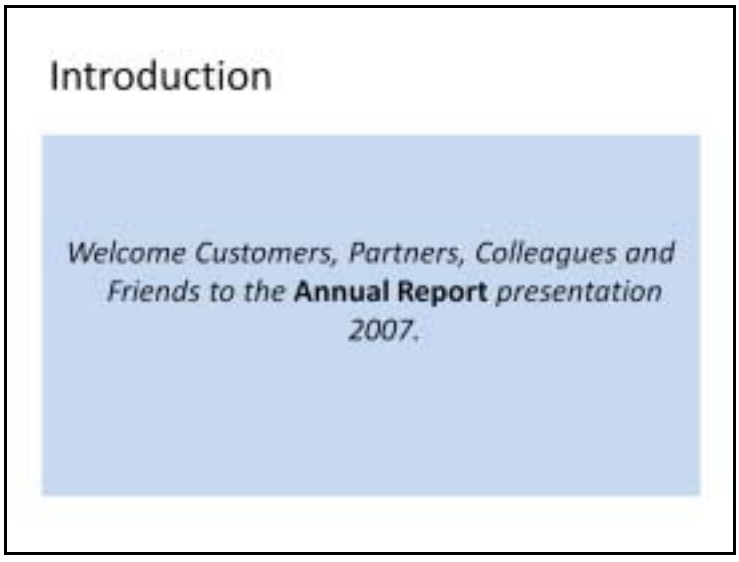
Weiterer use case, wie sich embedded spreadsheet in word dokumenten oder presentationen verhalten

Ob eine round trip möglich ist, rückwärtskompatibilität

Wie werden die spreadsheet docs referenziert

1.3 Presentation Documents

1.3.1 Simple Text Formatting



1.3.2 Itemization and Numeration

1.3.3 Positioning and Layout

1.3.4 Slide Blending and Effects

1.3.5 Animations

1.3.6 Diagrams

1.3.7 Multimedia Content and Vector Graphics

Text: Flash , MP3 und verschiedene Grafikformate

1.3.8 Masterlayout

2 Functionalities

2.1 Introduction

This section contains descriptions of the functionalities needed for the Use Cases described in section 1. The descriptions may reference section 3 for detailed descriptions of differences between ODF and OOXML in the implementation of the functionality and/or solutions to the problems.

The tables in the following subsections as well as those in section 0 summarize the availability of various functionalities for each of the two document formats as well as an estimate of the functionalities' "translatability", defined as follows:

Low: Either one of the formats does not contain this functionality at all, or the way the functionality is realized differs a lot, so that conversion of this feature is impossible without information loss.

Medium: Functionalities categorized as having a translatability of "Medium" usually are supported in both formats, although some aspects might differ and workarounds may be required. Features marked as "Medium" may support a single conversion but will result in information loss during round-trip conversions. The "Notes" column provides more details.

High: These functionalities are supported by both formats and should cause no problems during round-trip conversion.

It is important to note that the focus of this section is to describe the translatability of various office document features, not to engender discussion about the relevance of certain features or to make recommendations for the addition or removal of features from one of the standards.

2.2 Representation vs. interpretation

Both standards OpenDocument and Office Open XML focus on specifying the syntax or representation. However, to be able to define a mapping between the two standards a knowledge about the underlying interpretation or semantic of the different XML token and attribute is important. For illustration consider the following example:

Both OpenDocument and Office Open XML allow the definition of tab stops for a paragraph. In OpenDocument this is done by the <style:tab-stop> element. The position of a tab stop is defined by the style:position attribute. Its documentation reads: “The style:position attribute specifies the position of a tab stop.”

In Office Open XML the tab stop is defined by the <w:tab> element and the position is specified by the w:pos attribute: “Specifies the position of the current custom tab stop with respect to the current page margins.”.

The problem here is that OpenDocument does not specify whether the tab-stop position is relative to the margin or relative to the paragraph indent. (Please note that OpenDocument differentiates between tab-stops relative to margin or paragraph indent in the table of contents - but is silent about general tab stops. Also note that e.g. OpenOffice.org Writer treats style:position to be relative to the paragraph indent.).

These kinds of “unspecified behavior” make a precise mapping definition hard.

2.3 Word-Processing Documents

Text

2.3.1 Text Formatting

This section contains properties that may be applied to text in word-processing documents. Both office suites support formatting text per paragraph as well as on a more fine-grained level (Office Open XML calls this a “run”, OpenDocument a “span”). The following table summarizes the features that appear in the use cases in section 1. A complete listing of all text formatting settings and their translatability is available in section 3.2.1.1.

Table 1 — Text Formatting

Functionality	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
Bold text (font weight)		Yes	Yes	Medium	In addition to bold, OpenDocument allows font weight to be specified numerically (100-900). However e.g. OpenOffice.org only supports bold/non-bold.
Text color					
	RGB	Yes	Yes	High	
Font selection					
	By font name	Yes	Yes	High	

Functionality	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
	By font family	Yes	Yes	High	
	Theme fonts	Yes	No	Low	OpenDocument does not support the concept of document themes
Italic text		Yes	Yes	Medium	OpenDocument supports italic and oblique text, Office Open XML makes no such distinction
Kerning		Yes	Yes	High	
Text language		Yes	Yes	High	
Raised / lowered text		Yes	Yes	Medium	Office Open XML uses absolute values, OpenDocument uses percentages. This may lead to conversion problems.
Underline		Yes	Yes	Medium	Office Open XML allows single and double underline. OpenDocument offers more fine-grained control of underline options and styles.

2.3.2 Paragraph Formatting

In the context of word-processing documents, a paragraph is the smallest unit of text upon which layout is performed. Both document formats support applying the text formatting properties above on a per-paragraph basis, in fact Office Open XML simply embeds a run-properties element within the paragraph format whereas OpenDocument paragraph styles may contain paragraph and text properties.

The table below summarizes the paragraph formatting features that appear in the use cases. A complete listing of all supported paragraph features can be found in section 3.2.1.2.

Table 2 — Paragraph Formatting

Feature	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
Text alignment (left / right / centered / justified)		Yes	Yes	Medium	Office Open XML supports a range of additional values for arabic and thai text.
	for last line in paragraph	No	Yes	Low	
	justify single word	No	Yes	Low	
Margins					
	absolute, relative	No	Yes	Medium	Office Open XML only supports absolute values for paragraph margins.

Feature	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
	Left / right / top / bottom	Yes	Yes	Medium	Office Open XML supports contextual spacing where top/bottom spacing is ignored for identically formatted paragraphs.

2.3.3 Tables

Office Open XML and OpenDocument support the insertion of tables inside a document. Both formats allow table cells to span across multiple rows and / or columns and provide detailed control over the display of table elements. The table below encompasses the table features from the use case in section 1.1.2 and highlights further areas where functionality varies between the document formats.

Table 3 — Table Functionalities

Feature	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
Table properties					
	right-to-left layout	Yes	No	Medium	OpenDocument does not support rtl layout for tables. However the functionality can be emulated by reversing the cell order appropriately.
	alignment of whole table (left, center, right, auto, indented)	Yes	Yes	Medium	OpenDocument has no support for floating tables. However, this functionality may be emulated by placing a table inside a frame.
	background color	Yes	Yes	Medium	Again, OpenDocument does not support document themes, so information may be lost on conversion.
	background pattern	Yes	No	Low	
	background image	No	Yes	Low	
Data alignment	Horizontal / vertical	Yes	Yes	High	
Column Settings					
	adjust column width	Yes	Yes	High	
Row Settings					
	adjust row height	Yes	Yes	High	

Feature	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
Cell Settings					
	span multiple columns	Yes	Yes	High	
	span multiple rows	Yes	Yes	High	
Sub-Tables		No	Yes	Low	OpenDocument contains the concept of Sub-Tables, e.g. tables embedded seamlessly within a table cell. While the same effect may be reproduced by splitting and rejoining cells in the containing table, this would require a converter to be able to "render" the complete table internally.
Borders					
	Color / width / style	Yes	Yes	Medium	Both formats allow the same values as for paragraph borders.

2.3.4 Page Formatting

Text

Describe: …..

2.3.5 Itemization and Numeration (Lists and numbered Headings)

Since OpenDocument and Office Open XML differ in the way that numbering (e.g. of lists or headings) is handled, the following two subsections contain a short discussion of each document format's approach. The table in subsection 1.1.1.3 contains more a detailed functionality-based comparison based on the functionalities found in the use case in section 1.1.3.

Numbering in this context includes the handling of bulleted (itemized) lists, as both document formats handle them the same way as numbered lists.

1.1.1.1 Numbering in OpenDocument

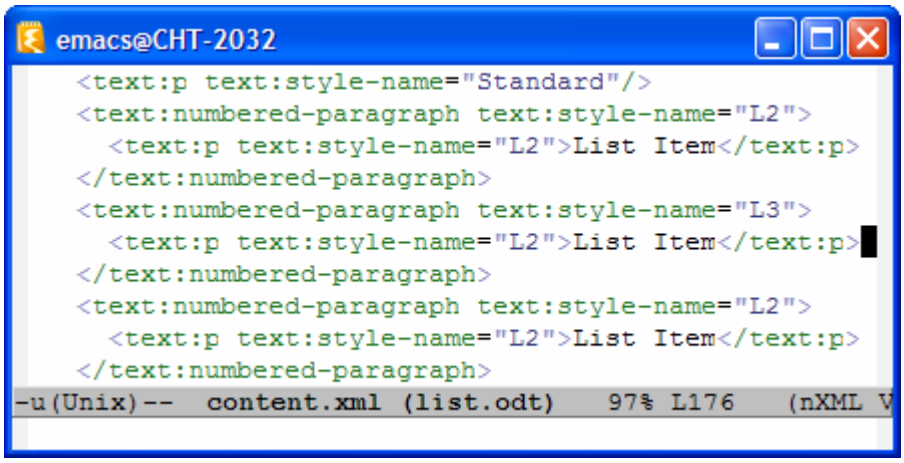
OpenDocument contains two ways of expressing lists: an approach based on the nesting of the individual XML tags used to define the list (structural approach) and another one in which regular paragraphs are marked as belonging to a list (attribute approach). The numbering and list formatting which is applied to a list item or heading is determined by a list style associated to the list (or numbered paragraph).

The structural approach is reminiscent of the way lists are constructed in XHTML[1] with specialized tags denoting lists and list items and the list level being determined by the nesting of list tags in the XML

[1] http://www.w3.org/TR/xhtml1/

representation of the document content. The attribute approach, on the other hand, simply annotates regular paragraphs with attributes identifying them as items of a specific list style at a certain list level. Both approaches are functionally equivalent, however only the latter approach can be used to apply numbering information to headings.

Unfortunately, the OpenDocument standard is worded ambiguously and allows for different interpretations of the attribute approach described above. It is unspecified whether the numbering logically resides with the list style or if there is a global counter for each list level which needs to be restarted manually. For example, the XML code in FIXME may be rendered as in FIXME when the numbering resides with the list style. However, when a global counter is used, the list would show up as in FIXME.




```
<text:numbered-paragraph text:style-name="L2">
  <text:p>List Item</text:p>
</text:numbered-paragraph>
<text:numbered-paragraph text:style-name="L3">
  <text:p>List Item</text:p>
</text:numbered-paragraph>
<text:numbered-paragraph text:style-name="L2">
  <text:p>List Item</text:p>
</text:numbered-paragraph>
```

Figure 12 — Ambiguous List formatting in OpenDocument: XML

TODO

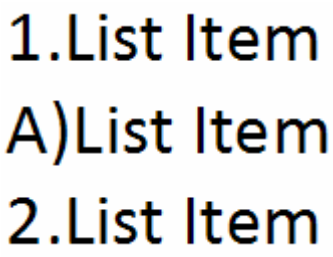


Figure 13 — Ambiguous List formatting in OpenDocument: Counter associated with list style

TODO

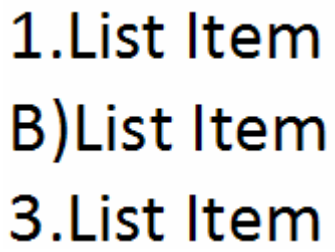


Figure 14 — Ambiguous List formatting in OpenDocument: Global counter

1.1.1.2 Numbering in Office Open XML

Office Open XML has no distinct concept of lists. Instead, a similar approach to OpenDocument's "attribute approach" as explained above is used: List items (and headings) are simply regular paragraphs to witch special properties are attached that contain information on the list structure (an identifier for the list the paragraph belongs to and the list level) and a reference to the formatting information for the list. Headings are treated the same way, except that they contain additional information on the outline level of the heading within the document.

A detailed explanation of the concepts used for numbering information in Office Open XML is contained in Part 3, section 2.10 of the Office Open XML standard. There exist three different ways how numbering information is applied to a paragraph.

In the most simple case, the paragraph is annotated with a reference to a *numbering definition*, which in turn inherits the actual numbering settings from an abstract numbering definition, as illustrated in Figure 15 — .

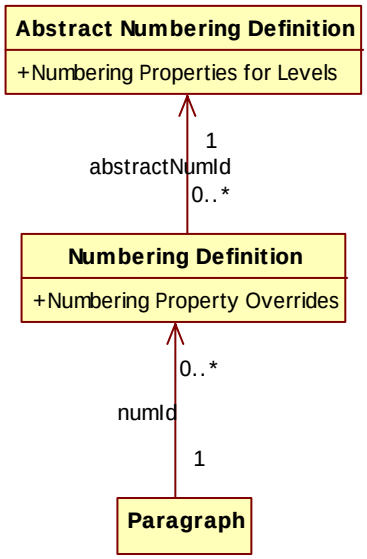


Figure 15 — Direct list formatting

Alternatively, a numbering style may be applied to the paragraph via one of two distinct yet equivalent approaches. In both cases, the numbering style is not referenced directly; instead, a numbering definition which references the style via its associated abstract numbering definition is applied as explained above. Figure 16 — illustrates the simple case in which the numbering style references the same numbering definition as the paragraph. However, the numbering style may reference a separate numbering definition, resulting in a structure as shown in Figure 17 — .

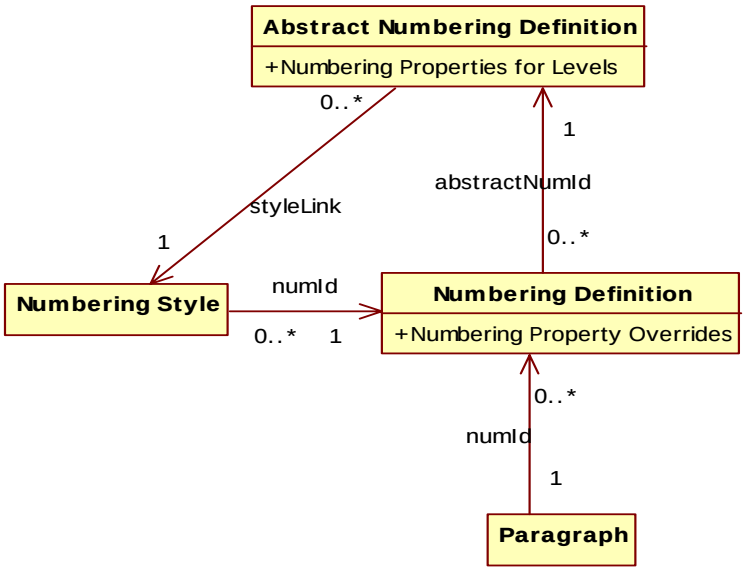


Figure 16 — Numbering Styles

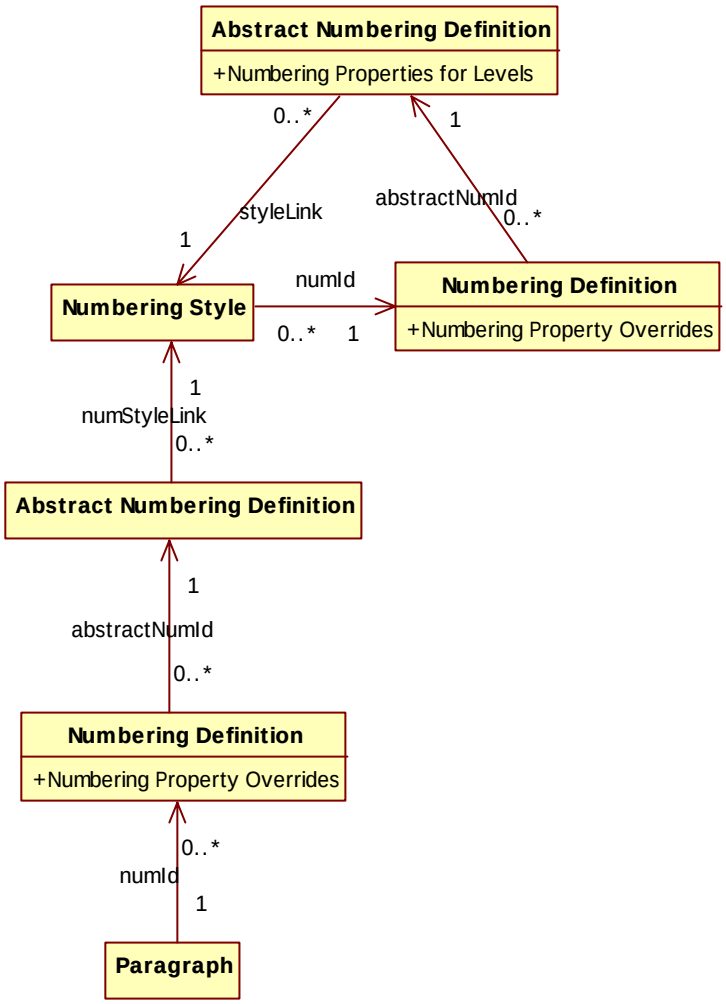


Figure 17 — Complex Numbering Styles

1.1.1.3 Comparison of Numbering and Enumeration in OpenDocument and Office Open XML

Both document formats offer a comparable level of support for numbered and/or bulleted lists. Office Open XML allows for more flexibility when specifying the formatting of nested numbering – e.g. the third-level heading numbered "1.2.3" may be represented as "Section I, 2.b)", whereas OpenDocument only allows specifying a prefix and a suffix for the numbering, while the levels always are separated by periods, resulting in "Section I.2.b)" (Prefix: "Section ", Suffix: ")").

Since both Formats offer multiple ways of applying numbering information to text segments, an implementation will most likely require fairly complex processing in order to retain the best possible graphical fidelity.

2.3.6 Indexes

Office Documents may contain various types of indexes, including the table of contents, but also indexes of figures, tables, etc. Since the two document formats follow different approaches in the way indexes are represented, this section will give an overview over the approaches in subsections 2.3.6.1and 2.3.6.2; while subsection FIXME offers a comparison of the functionalities included in the use case in section 1.1.4

2.3.6.1 Indexes in OpenDocument

OpenDocument supports three different types of indexes: tables of content, alphabetical indexes and user-defined indexes. Each index in turn is composed of two parts: an *index template* specifying all information necessary to generate the index and an *index body* containing a rendition of the index.

The information contained within the index template varies according to the type of index. It specifies the source material for the index, as well as an optional title and a template specifying how the title and each index entry should be rendered.

Depending on the type of index, the source material may be derived from

— The document structure (headings and subheadings)

— Index marks: explicit markers placed in the document

— Sequences, such as those used for the captions of tables or illustrations

— The paragraph styles applied to an entry

For example, the table of contents described in the use case 1.1.4 is built from the document's headings. Since the index has no title, the template would not specify one. Each entry is built from:

— The item's Title

— A tab stop

— The page number of the heading

The template specifies a paragraph style which is applied to each index entry. For more information about the supported formatting options, see section 3.2.1.2.

2.3.6.2 Indexes in Office Open XML

Indexes in Office Open XML are handled as part of the concept of document fields or dynamic data. This most closely resembles the second approach in the list for OpenDocument above. Index marks are placed

2.3.7 Page Layout

Text

2.3.8 Dynamic Content (Forms/fields/FIXME)

Text

2.3.9 Change Tracking and Document Revision

Both document formats offer support for revision tracking in word processing documents. In addition to the common operations, Office Open XML allows highlighting text regions with a limited set of colors (for more information, see section 2.3.1). OpenDocument's change tracking support is more coarse-grained than that of Office Open XML in that formatting changes, including those in tables, are recorded, but no information about the previous state is kept, so that the previous state cannot be reconstructed by rejecting the change.

Table 4 — Change Tracking and Document Revision

Functionality	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes

Functionality	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
Text insertion		Yes	Yes	High	
Text deletion		Yes	Yes	High	
Formatting changes		Yes	Yes	Medium	OpenDocument only records the fact that a change has occurred. However, no further information is recorded, so that it is impossible to reconstruct the previous state.
Comments		Yes	No	Medium	Office Open XML allows adding comments to arbitrary text ranges. This is not supported by OpenDocument, however similar functionality may be provided by inserting Notes (which are associated with a point in the text, not a range).
Text highlighting		Yes	No	High	Although OpenDocument does not support text highlighting, the functionality may be emulated by setting the text background color (see the section on text formatting)
Meta Data					
	Name	Yes	Yes	High	
	Date / Time	Yes	Yes	High	
	Author shorthand for comments	Yes	Yes	High	

2.4 Spreadsheet Documents

Text

Describe: ….

2.4.1 Formatting

Text

Describe: ….

2.4.2 Calculation

Text

Describe: ….

2.5 Presentation Documents

Text

Describe: ….

2.5.1 Slides

Text Describe: ….

2.5.2 Text Formatting

Text

Describe: ….

2.6 Common Aspects

Text

Describe: ….

2.6.1 Document Meta-Data

Text

Describe: ….

3 Detailed Information

Text

Describe: ….

3.1 Document Format Architecture

Text

Describe: ….

European Commission Paper -> Wolfgang Keber / Dialogika

3.1.1 Common Aspects

From the perspective of the user, both file formats manifest themselves as single files. From a developer's perspective, however, they consist of a ZIP[2] container containing various data streams, primarily in XML format. Both formats specify the way in which the data is organized as well as an index which enumerates the data streams and their content types.

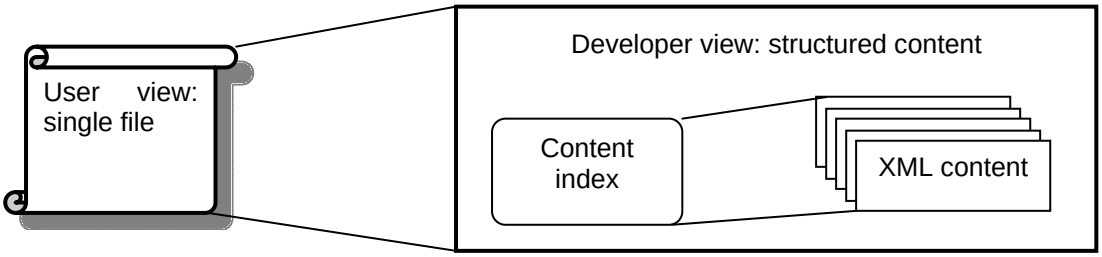


Figure 18 — User View versus Developer view

The file formats employ established methods for specifying the content allowed inside the XML data parts. While OpenDocument uses RelaxNG, an OASIS/ISO standard, Office Open XML uses the W3C's XML-Schema language as normative data model.

In Office Open XML, each document type's content is essentially represented by its own markup language. In addition, the specification contains a number of shared markup languages used by all document types, as follows:

Table 5 — Markup languages used by OOXML

Name	Description
WordprocessingML	Word processing content.
SpreadsheetML	Spreadsheet content.
PresentationML	Presentation content.

[2] Although the ZIP file format is not formally standardized, a well maintained specification is available at: http://www.pkware.com/documents/casestudies/APPNOTE.TXT

DrawingML	Vector drawings
VML	Vector drawings (marked as deprecated)
OMML (Office Math Markup Language)	Math formulas.
Extended Properties	
Custom Properties	
Custom XML Data Properties	
Bibliography	
Relationships	Used by Open Packaging Convention (OPC)

OpenDocument's design with respect to markup languages is more modularized. Each markup language realizes a certain functionality, which then is incorporated into the different document types (e.g. a table in a word processing document uses the same syntax as one within a spreadsheet). Since the individual markup languages have no names, the table below lists the default (XML-) namespace prefix instead:

Table 6 — Markup languages used by OpenDocument

Namespace Prefix	Description
office	Common information shared by all document types.
text	Text documents and text parts of other document types (e.g., a spreadsheet cell).
table	Spreadsheets, tables in a text document.
presentation	Presentations.
style	Style and formatting information.
drawing	Vector graphics.
dr3d	3D graphics.
chart	Charts.
anim	Animation information.
smil	Attributes defined in SMIL20 (supplements anim).
math	Mathematical equations (uses MathML).
meta	Meta information
dc	Meta information (uses the Dublin Core Namespace).
config	Application-specific settings.
form	Forms and controls.
script	Scripting and event handling.
number	Data type information (e.g. for spreadsheets).
manifest	The package manifest.
fo	Attributes defined in XSL:FO.
svg	Elements or attributes defined in SVG.
xlink	The XLink namespace.
xforms	The XForms namespace.
dom	The WWW Document Object Model namespace.

3.1.2 Office Open XML

The following diagram illustrates the relationships between the technologies Office OpenXML is based upon:

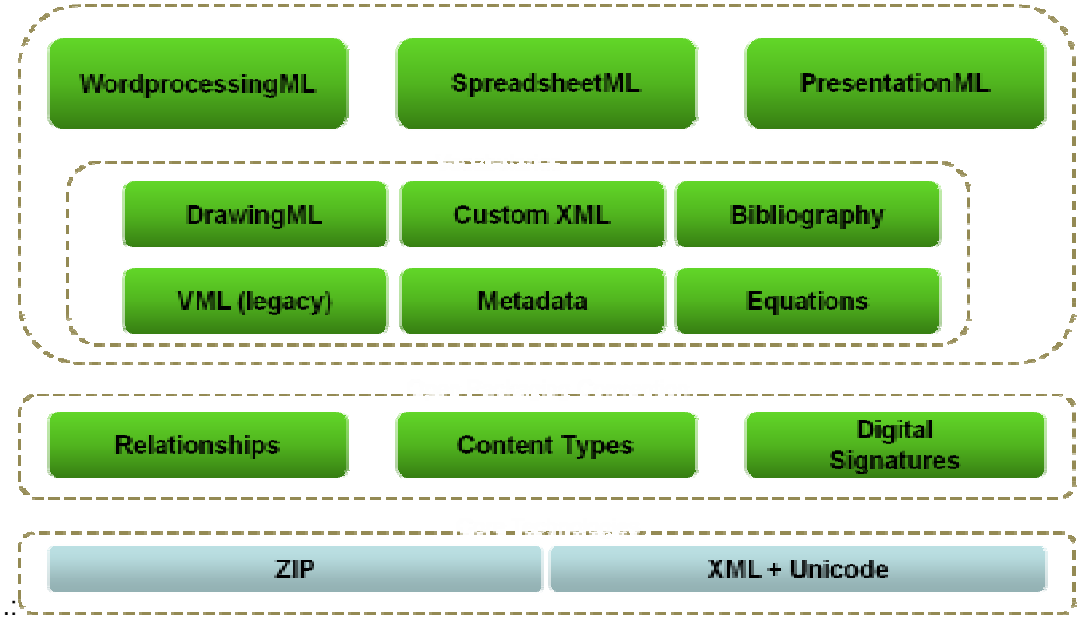


Figure 19 — OOXML architecture

Office Open XML contains specifications for the following three document types:

- Word processing
- Spreadsheets
- Presentation

Each document type is defined via its own markup language and uses the shared languages for functionalities common to all three document types (e.g. drawings, metadata …).

The structure of the ZIP container is defined via the Open Packaging Conventions (OPC), an abstraction layer between the physical file / directory layout inside the ZIP file and the document structure. OPC defines the concepts of *Parts* containing data and *Relationships* connecting the Parts.

At the root lies the so called “Content Type Stream” which identifies the type of the overall document as well as the content type of the individual Parts. The root Relationship defines the location in the ZIP file of the main document Part. Depending on the document type and the contents of the document, the main Part will be connected to further Parts and/or external documents via Relationships.

3.1.3 OpenDocument

Documents following the OpenDocument Format (ODF) have an internal structure similar to that of Java-archive files (JAR-files), in that they, too, contain a manifest file which declares the type and location of the files contained in the archive.

Unlike Office Open XML, ODF contains no abstraction layer above the physical file layout within the ZIP archive; instead, fixed file names are used for document content (content.xml), style information (styles.xml), meta information (meta.xml) and application settings (settings.xml). These files are placed in the root directory of the archive; the manifest (manifest.xml) resides in the subdirectory META-INF. For unencrypted documents,

the ODF standard mandates the presence of a thumbnail representation of the document, in PNG format, at the location Thumbnails/thumbnail.png.

In order to allow easier content-type recognition by operating systems and processing applications, ODF documents may contain a file “mimetype” containing nothing but the mime type in the root directory. If this file is present, the standard specifies that this type must not be compressed and must be the first entry in the ZIP archive, so that the information it contains is located at a fixed offset within the binary representation of the document.

The following diagram illustrates the minimal structure of an ODF document stored in a ZIP container:

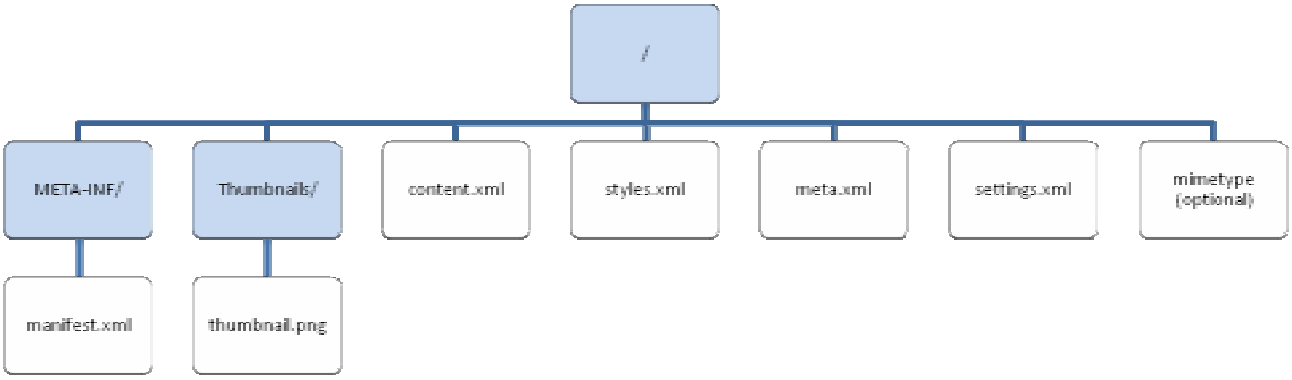


Bild 20 — Minimal OpenDocument structure

3.2 Detailed comparison of selected functionalities

Since an exhausting comparison of all functionalities would exceed the scope of a technical report, this section contains *complete* comparisons only of certain aspects of the functionalities implemented by the document formats. The sub-heading names in this section mirror those of section 2 in order to make it easier to locate the corresponding information in the sections referenced by the use-cases.

3.2.1 Word-Processing Documents

Text

3.2.1.1 Text Formatting

Text

Table 7 — Text formatting in general

Functionality	Sub-functionality (if applicable)	Office Open XML	Open Document	Translatability	Notes
Bold text (font weight)		Yes	Yes	Medium	In addition to bold, OpenDocument allows font weight to be specified numerically (100-900). However e.g OpenOffice.org only supports bold/non-bold.
Text borders		Yes	No	Low	OpenDocument only supports borders on whole paragraphs.
Capitalization					
	All upper case	Yes	Yes	High	
	Small caps	No	Yes	Low	
	All lower case	No	Yes	Medium	
	First-letter	No	Yes	Medium	First letter of each word is capitalized
Text color					
	RGB	Yes	Yes	High	
	Background color	Yes	Yes	High	
	Based on theme	Yes	No	Medium	OpenDocument has no concept of a “document theme”
	Blinking text	No	Yes	Low	(FIXME) OOXML supports blinking backgrounds, but no blinking AFAIK. Maybe flipping the colors can simulate the behaviour...
	Text highlighting	Yes	No	Medium	Only a limited range of colors is available for text highlighting

Functionality	Sub-functionality (if applicable)	Office Open XML	Open Document	Translatability	Notes
Complex script support		Yes	Yes	FIXME	The formats differ in how complex scripts (east-asian, right-to-left scripts) are supported. See FIXME. Right. There is a request in the ECMA TC to clarify the relationship to the Unicode BIDI Alg.
East-Asian text					
	Packing two lines into one	Yes	Yes	High	
	Brackets around two-lined text	Yes	Yes	Medium	In OpenDocument, left and right brackets can be specified independently
	Vertical text	Yes	Yes	Medium	OpenDocument supports rotating text by 0, 90 and 270 deg., Office Open XML only 0 and 90 deg.
	Emphasis marks	Yes	Yes	Medium	OpenDocument offers more fine-grained support. Marks may be placed above or below text.
Font selection					
	By font name	Yes	Yes	High	
	By font family	Yes	Yes	High	
	Theme fonts	Yes	No	Low	OpenDocument does not support the concept of document themes
Font effects					
	Emboss	Yes	Yes	High	
	Imprint	Yes	Yes	High	
	Outline	Yes	Yes	High	
	Shadow	Yes	Yes	Medium	OpenDocument allows for fine-grained control of text-shadow parameters, Office Open XML only allows turning shadow on or off
Manual specification of run/span width		Yes	Yes	Medium/low (FIXME)	Office Open XML uses absolute values, OpenDocument uses percentages. This may lead to conversion problems
Italic text		Yes	Yes	Medium	OpenDocument supports italic and oblique text, Office Open XML makes no such distinction
Kerning		Yes	Yes	High	
Text language		Yes	Yes	High (FIXME)	

Functionality	Sub-functionality (if applicable)	Office Open XML	Open Document	Translatability	Notes
Enable/disable spell checking for run/span		Yes	No	Low	
Raised/lowered text		Yes	Yes	FIXME	Office Open XML uses absolute values, OpenDocument uses percentages. This may lead to conversion problems.
Strikethrough		Yes	Yes	Medium	Office Open XML allows single and double strikethrough. OpenDocument offers more fine-grained control of strikethrough options and styles.
Underline		Yes	Yes	Medium	The note on strikethrough applies equally to text underlining.

3.2.1.2 Paragraph Formatting

Text

Table 8 — Paragraph formatting in general

Feature	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
Line height			Yes		
	fixed	Yes	Yes	High	
	minimum	Yes	Yes	High	
	line distance	Yes	Yes	High	
	line spacing ("durchschuss")	No	Yes	Low	
	font-independent line spacing	No	Yes	Low	
	automatic	Yes	No	Low	Office Open XML provides a (boolean) option that specifies "HTML-like" line spacing.
Text alignment (left/right/centered/justified)		Yes	Yes	Medium	Office Open XML supports a range of additional values for arabic and thai text.
	for last line in paragraph	No	Yes	Low	
	justify single word	No	Yes	Low	
Do not split paragraph into multiple pages		Yes	Yes	High	
Keep paragraph on same page as following paragraph		Yes	Yes	High	
Widows/orphans			Yes		
Tab stops		Yes	Yes		
	position	Yes	Yes		
	type (left, center, right, decimal)	Yes	Yes	Medium	Office Open XML does not support specifying the decimal character.
	type (bar, clear, list)	Yes	No	Low	These tab stop styles are supported in Office Open XML but their use is discouraged.

Feature	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
	leader properties	Yes	Yes	Medium	The formats support different kinds of leader styles. OpenDocument reuses the same styles it allows for underline and strikethrough. Office Open XML supports a fixed list of styles.
	leader text, text style	No	Yes	Low	OpenDocument allows for an arbitrary styled character as leader text.
	default tab stop	Yes	Yes	High	
Hyphenation					Office Open XML only allows suppressing automatic hyphenation on a per-paragraph basis.
	last word on page	No	Yes	Low	
	max. consecutive hyphenated lines	No	Yes	Low	
Drop Caps		Yes, via text frames	Yes	Medium	Office Open XML handles drop caps via specialized text frames. OpenDocuments approach is more straight-forward
register true text (same text baselines across mult. Pages /columns)			Yes		
Margins					
	absolute, relative	No	Yes	Medium	Office Open XML only supports absolute values for paragraph margins.
	left/right/top/bottom	Yes	Yes	Medium	Office Open XML supports contextual spacing where top/bottom spacing is ignored for identically formatted paragraphs.
first line indent			Yes		
	absolute, relative	Yes	Yes	Medium	Office Open XML only supports absolute values for first-line indentation.
	based on font size	No	Yes		FIXME: ODF auto-text-indent, ???

Feature	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
Page/Column Break					
	before paragraph	Yes	Yes	Medium	Office Open XML does not support column breaks as paragraph properties
	after paragraph	No	Yes	Low	
Background color		Yes	Yes	Medium	Office Open XML allows using theme color attributes. OpenDocument does not support the concept of a "document theme".
Background Pattern		Yes	No	Low	
Background image			Yes		FIXME: Does Office Open XML support background images?
	tiled(repeated)/stretched		Yes		
	position		Yes		
	filter??		Yes		
	opacity (percent)		Yes		
Background transparency			Yes		
Borders		Yes	Yes		
	top/bottom/left/right	Yes	Yes	High	
	between/bar	Yes	No	Low	Office Open XML a paragraph may have a "bar" (a border on the "inner" side of the paragraph when a book-like layout is used). Additionally a "between" border can be specified for paragraphs with identical border formatting. OpenDocument allows merging the borders of consecutive, identically formatted paragraphs.
	color	Yes	Yes	Medium	Office Open XML allows using theme color attributes. OpenDocument does not support the concept of a "document theme".
	frame effect	Yes	No	Low	
	shadow effect	Yes	Yes	Medium	OpenDocument offers more fine-grained control of shadow parameters

Feature	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
	spacing	Yes	Yes	High	FIXME: the Office Open XML-standards description seems to confuse page- and paragraph borders?!?
	width	Yes	Yes	High	
	type	Yes	Yes	Medium	Office Open XML documents can specify "art borders", a concept not supported by OpenDocument. While both document formats support a wide range of border styles, the sets differ. However, common styles (single/double/dotted lines) are supported by both formats.
Padding			Yes		
Shadow			Yes		
Keep on same page as following paragraph			Yes		
line numbering		No	Yes	Low	Office Open XML only supports line numbering on a per-section level, not as a paragraph setting. Individual paragraphs can be excempted from line numbering.
	(re-)set start value	No	Yes	Low	
vertical alignment (top,middle,bottom,baseline)			Yes		
Asian / complex text layout properties					
	add space between asian, ctl and western text	Yes	Yes	Medium	Office Open XML allows specifying extra spacing between Asian and Roman text as well as Asian Text and Numbers. OpenDocument allows spacing between asian, ctl and western text (but not numbers)
	allow punctuation to hang into margin	Yes	Yes	High	

Feature	Sub-Feature (if applicable)	Office Open XML	Open Document	Translatability	Notes
	line breaking behavoiur (strict / auto)	Yes	Yes	Medium	FIXME: Office Open XML allows more specific settings (kinsoku).
	writing mode (lr/rl/tb)	Yes	Yes	Medium	Office Open XML only supports setting paragraph properties to right-to-left or left-to-right.
	snap to layout grid	Yes	Yes	High	
	define page number if paragraph starts new page		Yes		
FIXME: Text Frames		2.3.1.11			
	suppress overlap	2.3.1.36			
FIXME: Lists					

4 Conclusion

Text