User initiative for improving OOXML integration in LibreOffice/Apache Open Office

Dr. Matthias Stürmer, Ernst & Young

ApacheCon Europe, 8 November 2012

Short bio of Matthias Stürmer

Dr. Matthias Stürmer
Manager

Ernst & Young AG
Belpstrasse 23
3001 Bern
Switzerland

matthias.stuermer@ch.ey.com
Work: +41 58 286 61 97
Mobile: +41 58 289 61 97

- Matthias Stürmer is Manager at Ernst & Young based in Bern, Switzerland. He focuses on open source advisory, open data management, and social media governance.
- He studied business administration and computer science at University of Bern until 2005 and finished his doctoral dissertation at the Chair of Strategic Management and Innovation at ETH Zürich in 2009.
- His research focused on open source communities and firm involvement. The title of his doctoral thesis was "How Firms Make Friends: Communities in Private-Collective Innovation".
- He then worked for one year at Liip AG, a Swiss software company creating agile Internet solutions based on open source technologies.
- Matthias Stürmer is working group leader “Office Interoperability” of the Open Source Business Alliance, member of the board of Swiss Open System User Group /ch/open, member of the Swiss association Opendata.ch, secretary of the Swiss Parliamentarian Group for Digital Sustainability, and leader of www.opensource.ch and other open source initiatives.
- Since 2011 he is member of the parliament of the city of Bern.

Ernst & Young and Open Source

- Open source has one **major weakness: marketing and PR**
- Top management vendor-neutral brochure from Ernst & Young: **Why and how professionals use open source software**
- In English and German
- Content:
  - Benefits, risks and good practices
  - Professional application of open source software
  - Legal aspects of open source
  - Background information on open source software

Download as PDF on
www.opensource.ch

Open Source in Switzerland

- New study of October 2012
- Introduction by CIO of UBS
- Results of survey with 202 firms and public institutions
- Around 20 articles on open source topics and use cases
- Portrait of open source companies

Download as PDF on
www.opensource.ch

Agenda

- **Motivation for the initiative**
- Phase 1: Initialization
- Phase 2: Funding
- Phase 3: Implementation
- Discussion

Why promoting a proprietary format?

To increase adoption of
open source office suites!

Motivation of the initiative

- Currently Microsoft's Office Open XML **(OOXML) format is not well integrated** in LibreOffice/Apache Open Office
- However, this is necessary to successfully **complete and maintain migrations** towards open source office suites
- Today various **public institutions** have switched to LibreOffice/Apache Open Office on the desktop or use head-less open source office suites for business applications.
- Nevertheless citizens, businesses, and other public stakeholders continue to send OOXML documents, especially **.docx files.**
- These external stakeholders expect that the public institution are able to handle such files.

Major technical issues

- **Opening** OOXML formats is possible with current versions of LibreOffice/Apache Open Office
- Unknown if visualization is conform to Microsoft Office 20XX
- **Most relevant OOXML support issues:**
  - Layout of .docx documents regarding frames, images, tables, enumerations, header/footers etc.
  - Fonts embedding (resulting in wrong or different layout of documents)
  - Comments layout and editing in .docx and .xslx
  - Change tracking within .docx files

Collective action problem

- “Scratching a developer's itch” does not always work for public administrations because they have different itches.
- While open source projects bring great advantages for their users, they inherently bear the problem of **„collective action“.**
- Although it would be beneficial for all users to implement enhancements in open source projects, **no single stakeholder** is willing to invest sufficient resources to accomplish this considerable development effort alone. All the other beneficiaries would **free-ride** on the solution.
- Therefore often status quo does **not represent the optimum.**
- In order to solve this challenge **coordinated action** among the stakeholders is necessary.

Solving the collective action problem

Phase 1: Initialization

a) Mobilize attention and interest of several open source software users
b) Create clear and common understanding of the issues, ask the experts
c) Aggregate and specify the requirements, prioritize and cluster them

Phase 2: Funding

a) Publish Request for Proposal (RfP)
b) Evaluate and decide for best proposal(s)
c) Do fund-raising for the implementation project

Phase 3: Implementation

a) Organize project management, sign contracts, start implementing
b) Do testing among the open source software users, finalize development
c) Publish resulting source code, pass it upstream to the open source project

Agenda

- Motivation for the initiative
- **Phase 1: Initialization**
- Phase 2: Funding
- Phase 3: Implementation
- Discussion

Setup of the initiative

- Workshop in October 2011 in Zürich
- Tasks were:
  - Identification of common issues (users)
  - Technical challenges and innovative solutions regarding OOXML (developers)
  - Specification of use cases regarding requirements and technical aspects (users and developers in groups)
  - Discussion and prioritization of use cases (all)
  - Funding models of implementation phase (users)

Initiating organizations

Public Institutions

- City of Freiburg i.B.
- City of München
- City of Jena
- Swiss Federal Court
- Federal Steering Unit for IT (ISB)
- Canton of Vaud
- Another Swiss federal agency

Community organizations

- Association Swiss Open Systems User Group /ch/open
- Association Freies Office Deutschland e.V.
  (former association OpenOffice.org Deutschland e.V.)
- Open Source Business Alliance OSBA

Workshop participants

Public administration:

- Rüdiger Czieschla, City of Freiburg
- Andreas Kawohl, City of Freiburg
- Andor Ertsey, City of München
- Dr. Jutta Kreyss, City of München
- Dr. Martin Reinhardt, City of Jena
- Dr. Harro Rosner, City of Jena
- Marc von Weissenfluh, Swiss Federal Court
- Daniel Brunner, Swiss Federal Court

Developers:

- Thorsten Behrens, Novell/SUSE
- Florian Reuter, self-employed developer
- Svante Schubert, self-employed developer
- Thomas Uhl, Open Source Business Alliance
- Dr. Frank Siebert, Freies Office Deutschland e.V.
- Lothar K. Becker, Freies Office Deutschland e.V.
- Mario Wendt, Microsoft
- Stefan Stüring, KO GmbH
- Tobias Hintze, KO GmbH
- Jos van den Oever, KO GmbH

Ernst & Young:

- Dr. Matthias Stürmer, workshop leader
- Mario Aeby, assistant

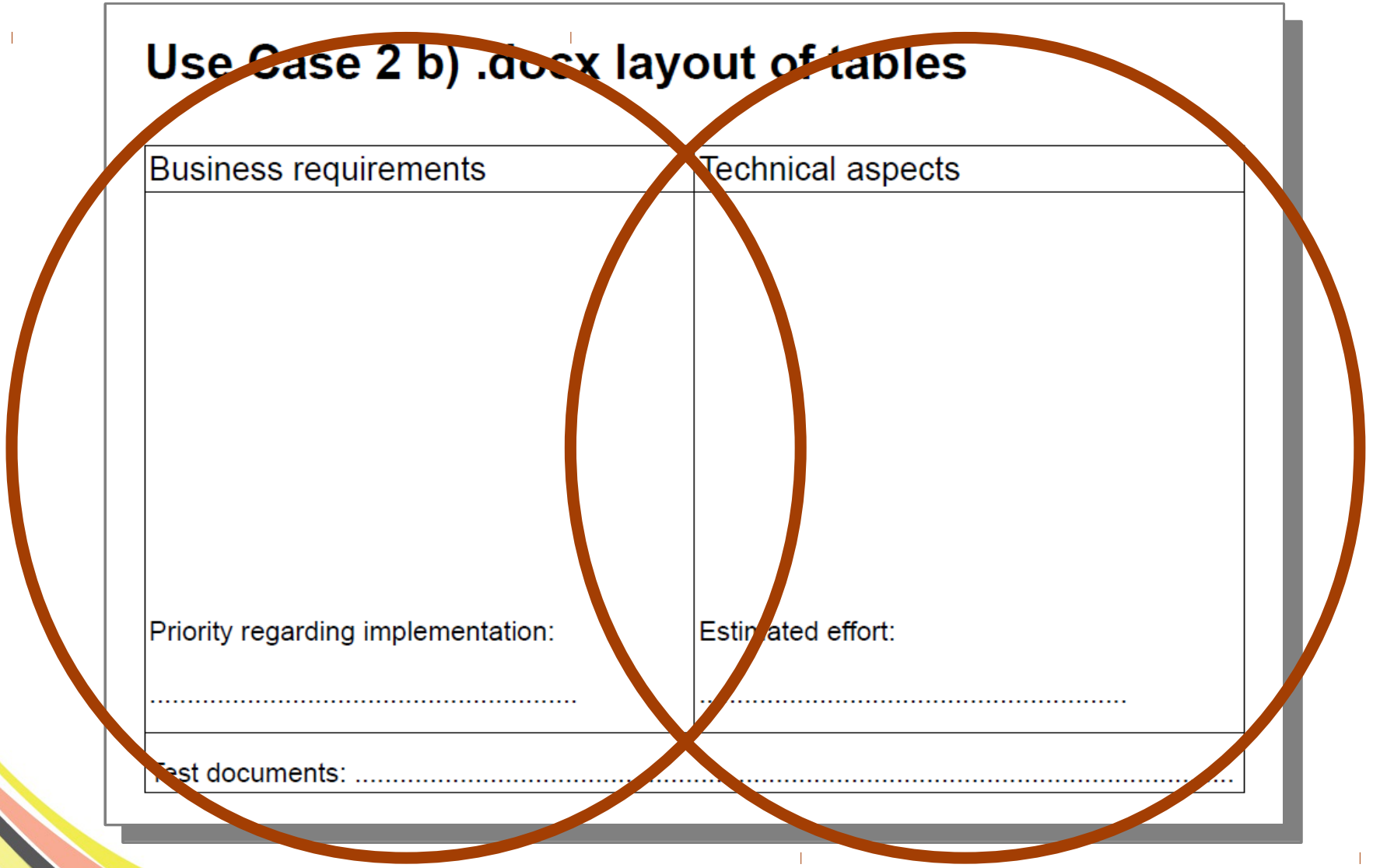
Challenges

- **Stakeholders**
  - Users: public institutions
  - Developers: LibreOffice and OOXML filter programmers
- **Challenges**
  - Huge knowledge gap: terminology, standard specification, structures and processes within public administrations etc.
  - Different perspectives: input oriented (=developers) vs. output oriented (=users)
  - Different interests: perfect implementation vs. solving problem
- **How to collaborate together?**

Example: filter vs. feature

- Important difference between
  - **error** in OOXML filter
  - missing OOXML **feature** within LibreOffice/ Apache Open Office
- What does that mean?
  - Little effort to correct OOXML filter
  - Large effort to implement new office feature
- Example of not correctly displayed OOXML standard: **Alternating row colors of tables within .docx files**
  - Little effort to 'hard format' colors, however loss of information in the .docx file
  - Large effort to program this feature in office application because of missing feature in ODF standard

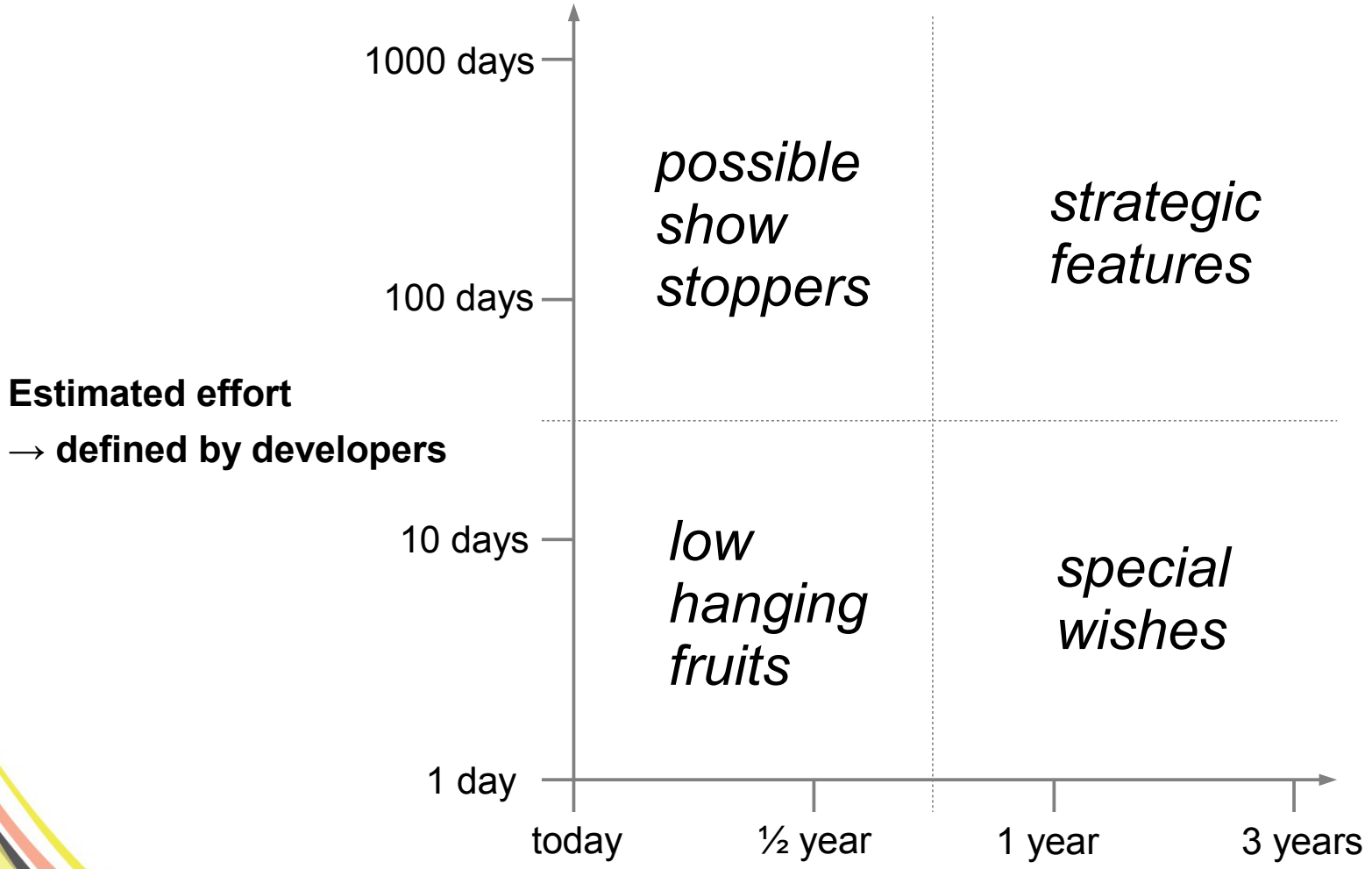
Structure of a use case



Use Case 2 b) .docx layout of tables

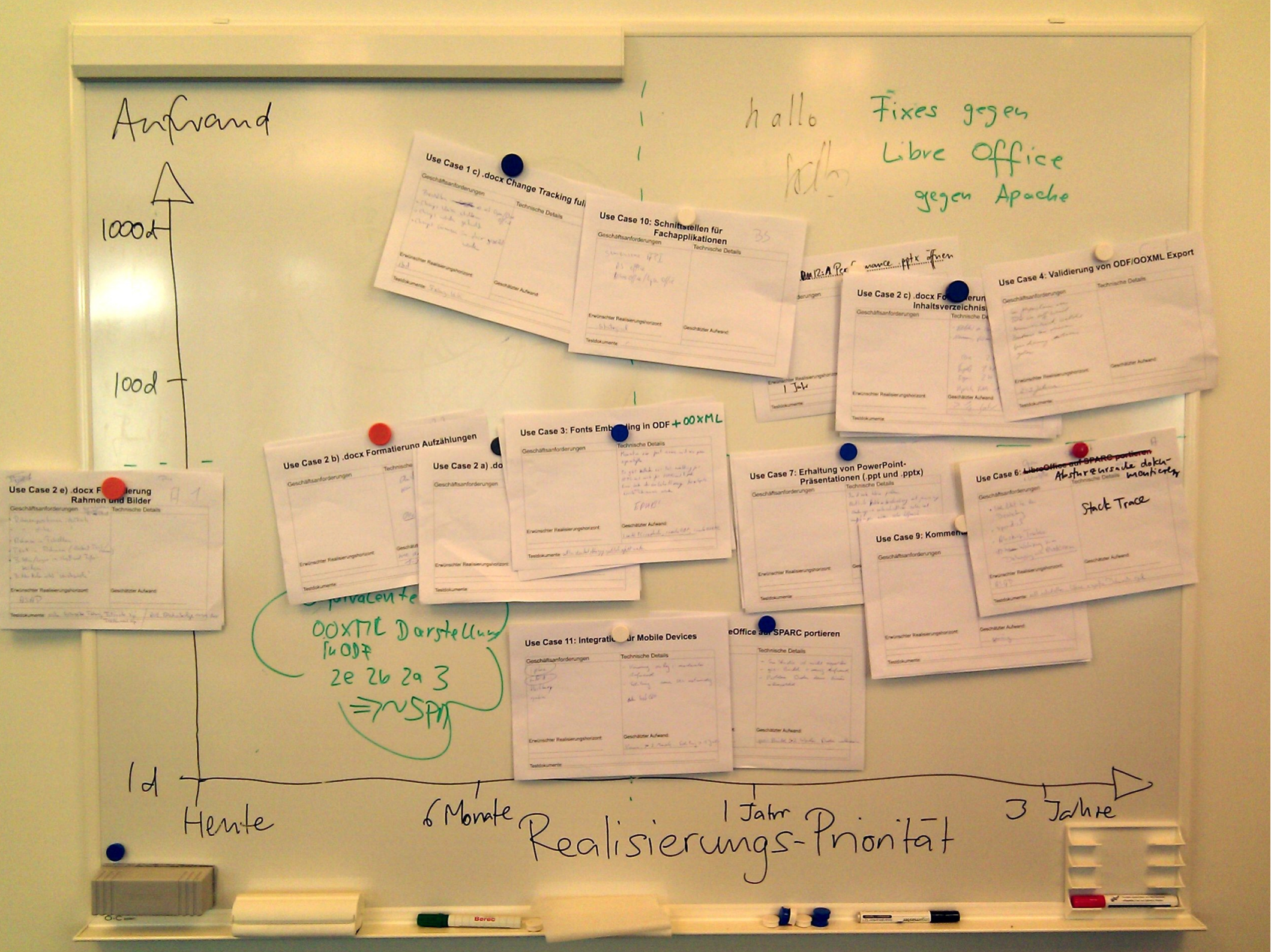
Business requirements	Technical aspects
Priority regarding implementation:	Estimated effort: .....
Test documents:	

Evaluation of use cases



Estimated effort
→ defined by developers

Priority regarding implementation → defined by users (public institutions)


Anfwand
1000d
100d
1d
hallo
Fixes gegen
Libre Office
gegen Apache
Use Case 1 c) .docx Change Tracking full
Use Case 10: Schnittstellen für Fachapplikationen
Use Case 2 c) .docx Formatierung Inhaltsverzeichnis
Use Case 4: Validierung von ODF/OOXML Export
Use Case 2 b) .docx Formatierung Aufzählungen
Use Case 2 a) .do
Use Case 3: Fonts Embedding in ODF + OOXML
Use Case 2 e) .docx Formatierung Rahmen und Bilder
Use Case 7: Erhaltung von PowerPoint-Präsentationen (.ppt und .pptx)
Use Case 6: LibreOffice auf SPARC portieren
Stack Trace
Use Case 9: Kommen
Use Case 11: Integrati ür Mobile Devices
eOffice auf SPARC portieren
OOXML Darstellu
in ODF
2e 2b 2a 3
Heute
6 Monate
1 Jahr
3 Jahre
Realisierungs-Priorität

Agenda

- Motivation for the initiative
- Phase 1: Initialization
- **Phase 2: Funding**
- Phase 3: Implementation
- Discussion

Bundle of use cases

- Use case bundle is called **“Layout-preserving view of OOXML documents in open source office applications”**

- **Targeted use cases:**
  1. Formatting of frames and images within .docx
  2. Formatting of tables within .docx
  3. Formatting of lists and numbers within .docx
  4. Formatting of comments within .docx and .xlsx
  5. Font embedding for both OOXML and ODF

- **Every use case is specified by test documents**

ÜBER UNS
MITGLIEDER
WORKING GROUPS
RICHTLINIEN DER WORKING GROUPS
OFFICE INTEROPERABILITÄT
DEUTSCHE WOLKE
OPEN THIN CLIENT ALLIANCE (OTCA)
PUBLIC AFFAIRS
BITKOM
VERANSTALTUNGEN
EDUCATION

OFFICE INTEROPERABILITÄT

Die Arbeitsgruppe "Office Interoperability" geht aus einem Workshop hervor, der im Oktober 2011 IT-Verantwortliche aus Behörden in Deutschland und der Schweiz zu dem Thema OOXML-Interoperabilität von Libre/OpenOffice zusammengebracht hat. Die konkrete Umsetzung des initiierten Prozesses wird in der Working Group der OSB Alliance erfolgen.

Die Ziele der Working Group können somit wie folgt zusammengefasst werden:
Verbesserung der OOXML-Interoperabilität bei Libre/OpenOffice
Priorisierung und Spezifikation von Bug/Feature-Requests aus Anwendersicht
Koordination der gezielten Finanzierung und Weiterentwicklung von Libre/OpenOffice (insbesondere Public Sector)

Als erstes Projekt verfolgt die Working Group die Umsetzung der Spezifikation „Layout-getreue Darstellung von OOXML-Dokumenten in Open Source Office Applikationen".

Hier finden Sie die englische Version Specification of "Layout-true Representation of OOXML Documents in Open Source Office Applications".

Weiterführende Informationen zu der Working Group Office Interoperability erhalten Sie bei dem designierten Working Group Sprecher Dr. Matthias

AKTUELLES

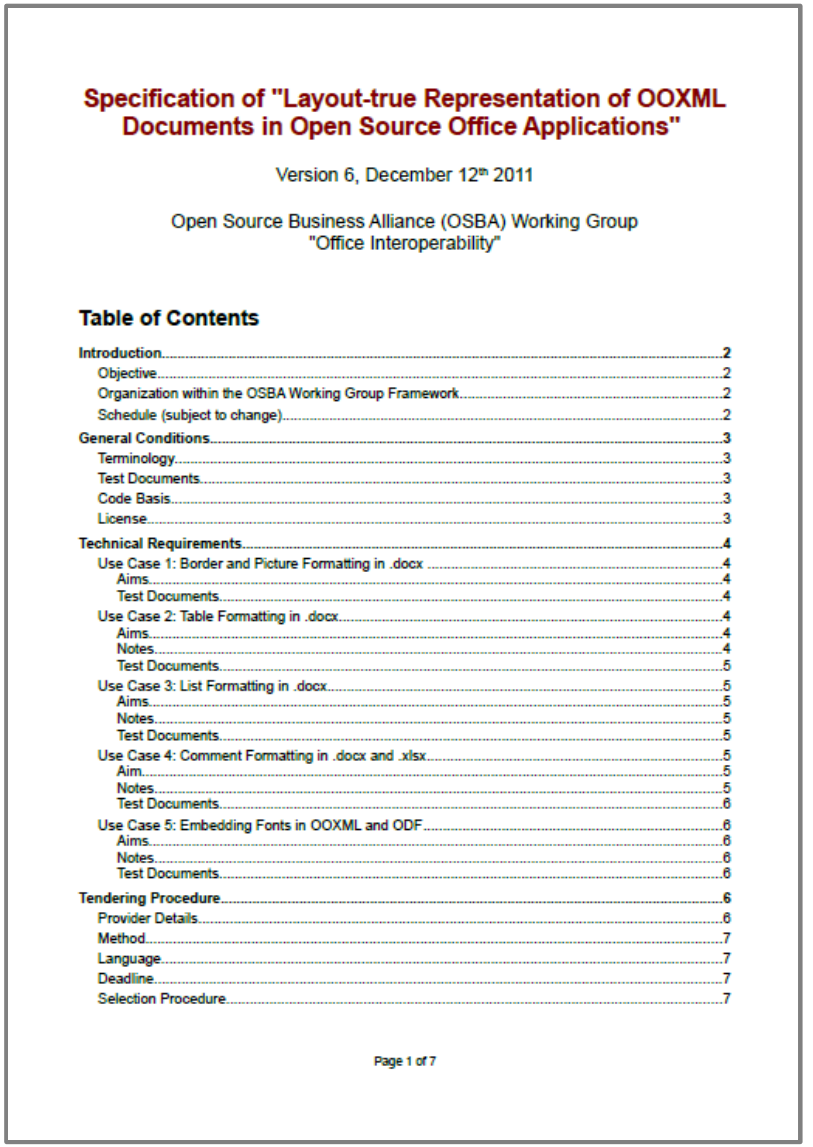
Open Source Day 2012
20 November 2012
12:15 - 17:30 Uhr
Veranstaltungsort: IHK München
Informationen zur Veranstaltung

OSB TWEETS

Herzlich willkommen! Wir haben ein neues Mitglied: Allrutz Consulting - und freuen uns auf gemeinsame Aktivitäten!

Specification and RfP

- Specification and Request for Proposals (RfP) published in December 2011
- Press releases by the OSBA working group Office Interoperability
- Proposals submitted by Lanedo and SUSE in February/March 2012
- Contracts split up along the five use cases



Specification of "Layout-true Representation of OOXML Documents in Open Source Office Applications"

Version 6, December 12th 2011

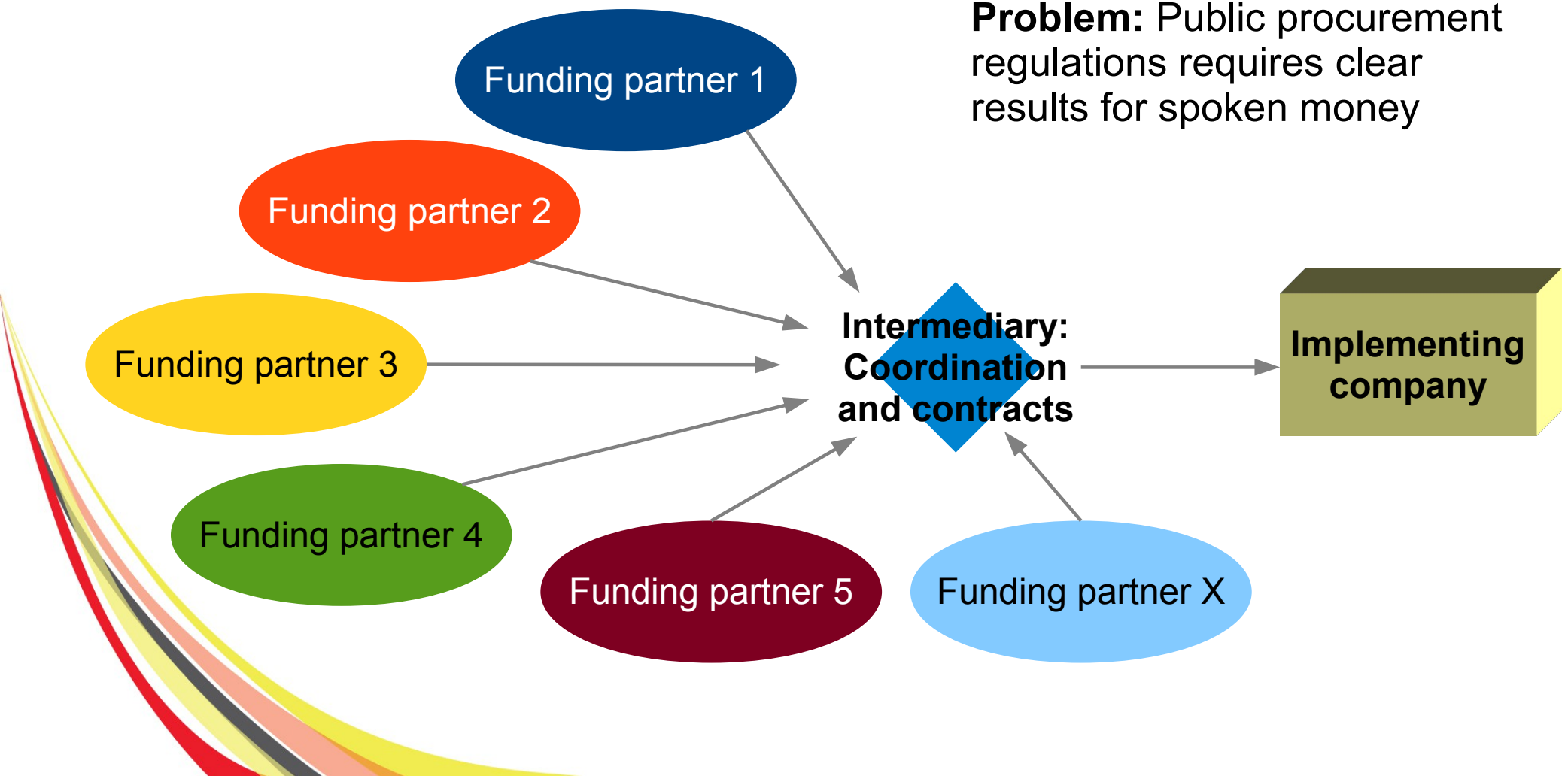
Open Source Business Alliance (OSBA) Working Group
"Office Interoperability"

Table of Contents

Introduction ... 2
Objective ... 2
Organization within the OSBA Working Group Framework ... 2
Schedule (subject to change) ... 2
General Conditions ... 3
Terminology ... 3
Test Documents ... 3
Code Basis ... 3
License ... 3
Technical Requirements ... 4
Use Case 1: Border and Picture Formatting in .docx ... 4
Aims ... 4
Test Documents ... 4
Use Case 2: Table Formatting in .docx ... 4
Aims ... 4
Notes ... 4
Test Documents ... 5
Use Case 3: List Formatting in .docx ... 5
Aims ... 5
Notes ... 5
Test Documents ... 5
Use Case 4: Comment Formatting in .docx and .xlsx ... 5
Aim ... 5
Notes ... 5
Test Documents ... 6
Use Case 5: Embedding Fonts in OOXML and ODF ... 6
Aims ... 6
Notes ... 6
Test Documents ... 6
Tendering Procedure ... 6
Provider Details ... 6
Method ... 7
Language ... 7
Deadline ... 7
Selection Procedure ... 7

Public procurement challenge

Funding model "intermediary":



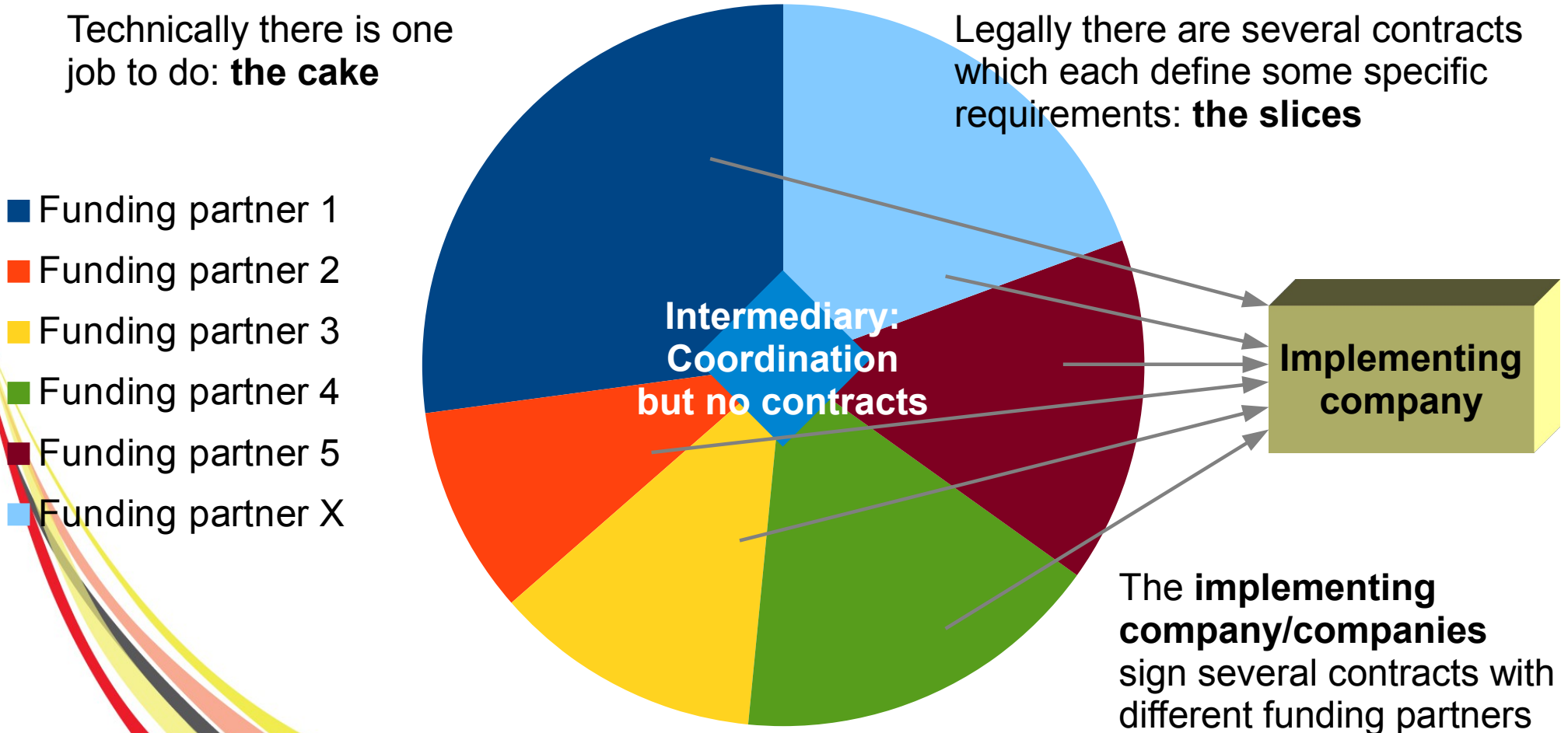
Problem: Public procurement regulations requires clear results for spoken money

Public procurement solution

Funding model “cake slices”:

Technically there is one job to do: **the cake**

Legally there are several contracts which each define some specific requirements: **the slices**



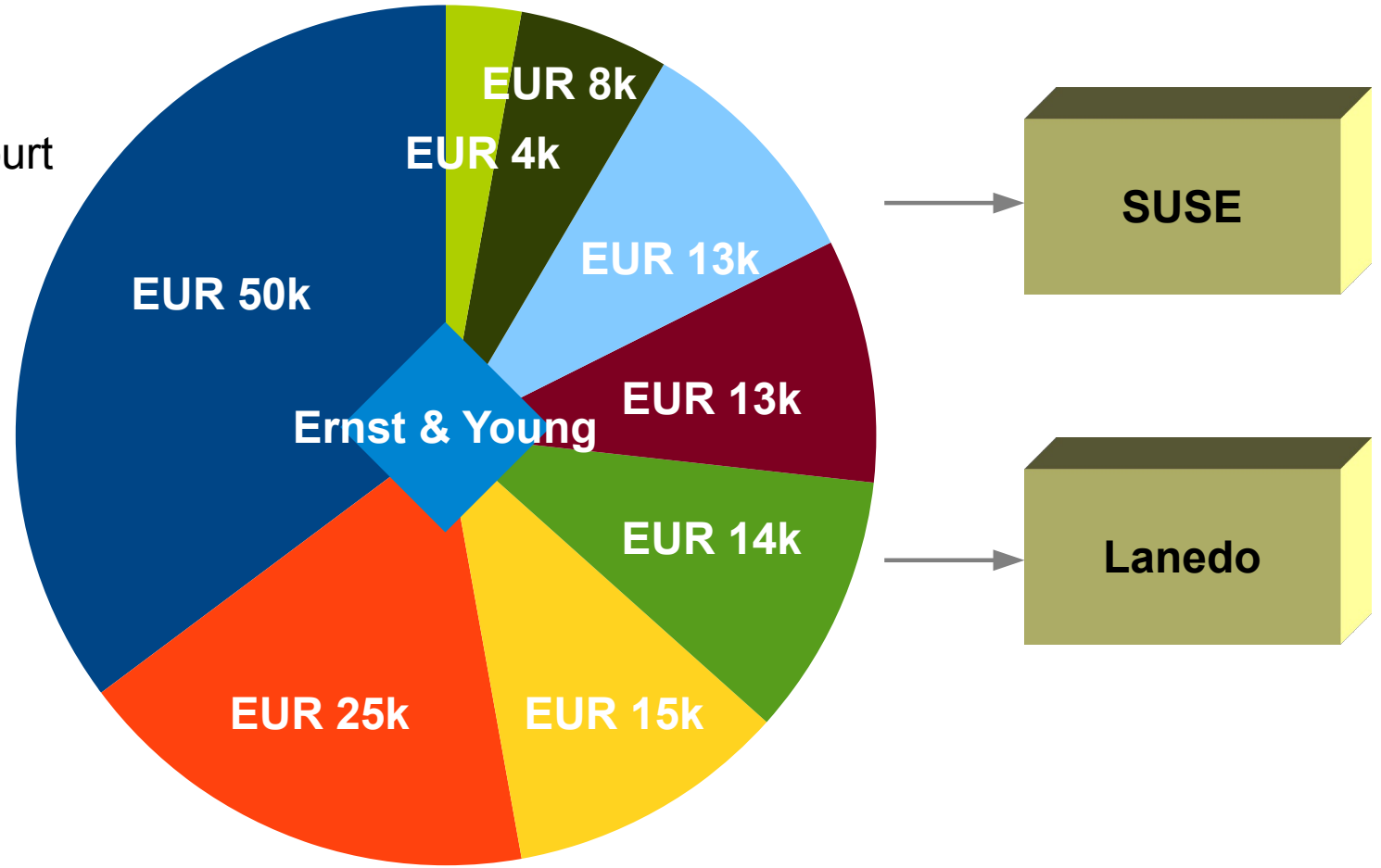
The **implementing company/companies** sign several contracts with different funding partners

Current OOXML improvements

Funding by

- City of Freiburg i.B.
- City of München
- City of Jena
- Swiss Federal Court
- Federal Steering Unit for IT (ISB)
- Canton of Vaud
- Another Swiss federal agency
- French ministry of culture and communication

Total: approx. EUR 140k (excl. VAT)



18 July 2012, 10:54

« previous | next »

Improved OOXML support for LibreOffice and OpenOffice

Developers from a project hosted by the Open Source Business Alliance are working to improve the compatibility of LibreOffice and OpenOffice with Microsoft Office. The German municipalities of Munich, Jena and Freiburg, and the Federal Supreme Court of Switzerland, the Swiss Federal IT Steering Unit (FITSU) and the Swiss Canton of Vaud – who together use OpenOffice on around 18,000 workstations – have jointly raised €140,000 (approximately £109,000) funding for the project.

The goal is to improve the way in which Microsoft Word, Excel and PowerPoint documents that have been saved in the Microsoft-developed OOXML format can be imported, displayed and edited. Among other things, the developers plan to enhance the formatting of frames, images, tables and lists in Word documents. They are also working to fix the display of comments in Word and Excel files and add the option to embed fonts in ODF

Agenda

- Motivation for the initiative
- Phase 1: Initialization
- Phase 2: Funding
- **Phase 3: Implementation**
- Discussion

Agenda

Phase 1: Initialization

a) Mobilize attention and interest of several open source software users
b) Create clear and common understanding of the issues, ask the experts
c) Aggregate and specify the requirements, prioritize and cluster them

Phase 2: Funding

a) Publish Request for Proposal (RfP)
b) Evaluate and decide for best proposal(s)
c) Do fund-raising for the implementation project

Phase 3: Implementation

a) Organize project management, sign contracts, start implementing
b) Do testing among the open source software users, finalize development
c) Publish resulting source code, pass it upstream to the open source project

Relevant for Apache Open Office

The specification explicitly says:

“License - The contractor is obliged to publish the new OpenOffice/LibreOffice software developed for this project under **Apache License Version 2.0.** This license permits the widest possible distribution of the development services.”

Current state of the work

Some background details:

- Use case 1 and 2 implemented by Lanedo
- Use case 3 to 5 implemented by SUSE
- Functionality should go into LibreOffice 3.7
- All changes will go as source code tarballs (Apache License) to the OSBA working group
- Technical contacts:
  - Thorsten Behrens from SUSE
  - Lionel Dricot from Lanedo

Current state:

- Use case 1 fully implemented
- Use case 2 fully implemented
- Use case 3 fully implemented
- Use case 4 partially implemented
- Use case 5 started

Work in progress in daily builds

Index of /daily/Linux-x86_10-Release-Configuration/libreoffice-3-6

Name	Last modified	Size	Description
Parent Directory		-	
2012-10-24_16.01.55/	25-Oct-2012 17:40	-	
2012-10-25_15.26.14/	26-Oct-2012 17:15	-	
2012-10-26_15.00.56/	27-Oct-2012 16:48	-	
2012-10-27_14.34.39/	28-Oct-2012 15:24	-	
2012-10-28_14.09.49/	29-Oct-2012 15:02	-	
2012-10-30_12.29.35/	31-Oct-2012 13:01	-	
2012-10-31_11.46.49/	01-Nov-2012 12:38	-	
2012-11-01_11.24.00/	02-Nov-2012 12:25	-	
2012-11-02_11.10.06/	03-Nov-2012 12:35	-	
2012-11-03_11.21.12/	04-Nov-2012 14:02	-	
2012-11-04_12.48.17/	05-Nov-2012 13:43	-	
current/	05-Nov-2012 13:43	-	

Apache/2.2.12 (Linux/SUSE) Server at dev-builds.libreoffice.org Port 80

Agenda

- Motivation for the initiative
- Phase 1: Initialization
- Phase 2: Funding
- Phase 3: Implementation
- **Discussion**

Questions

- How can OOXML filter patches programmed by the SUSE and Lanedo be integrated in Apache Open Office?
- In what way do Apache Open Office and LibreOffice collaborate today?
- What could be done to improve collaboration between Apache Open Office and LibreOffice?

Thank you for your attention!

For any feedback, questions,
ideas, suggestions please contact:

Dr. Matthias Stürmer
Manager at Ernst & Young AG
OSBA Working Group Leader Office Interoperability

matthias.stuermer@ch.ey.com
Work: +41 58 286 61 97
Mobile: +41 58 289 61 97