

ODF Advocacy Open Project @ OASIS

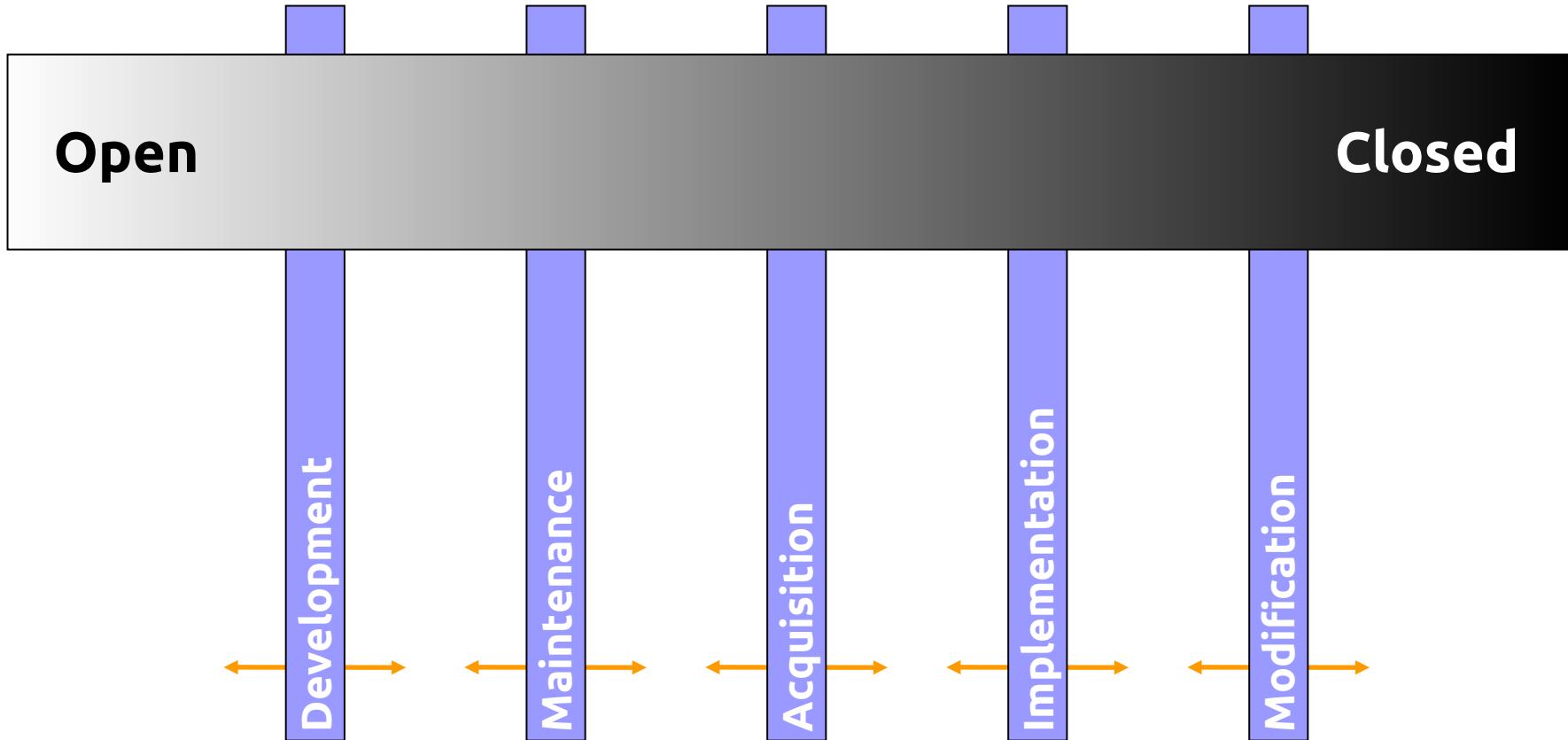
Open Document Format

the true document standard: freedom
of choice without any vendor lock-in

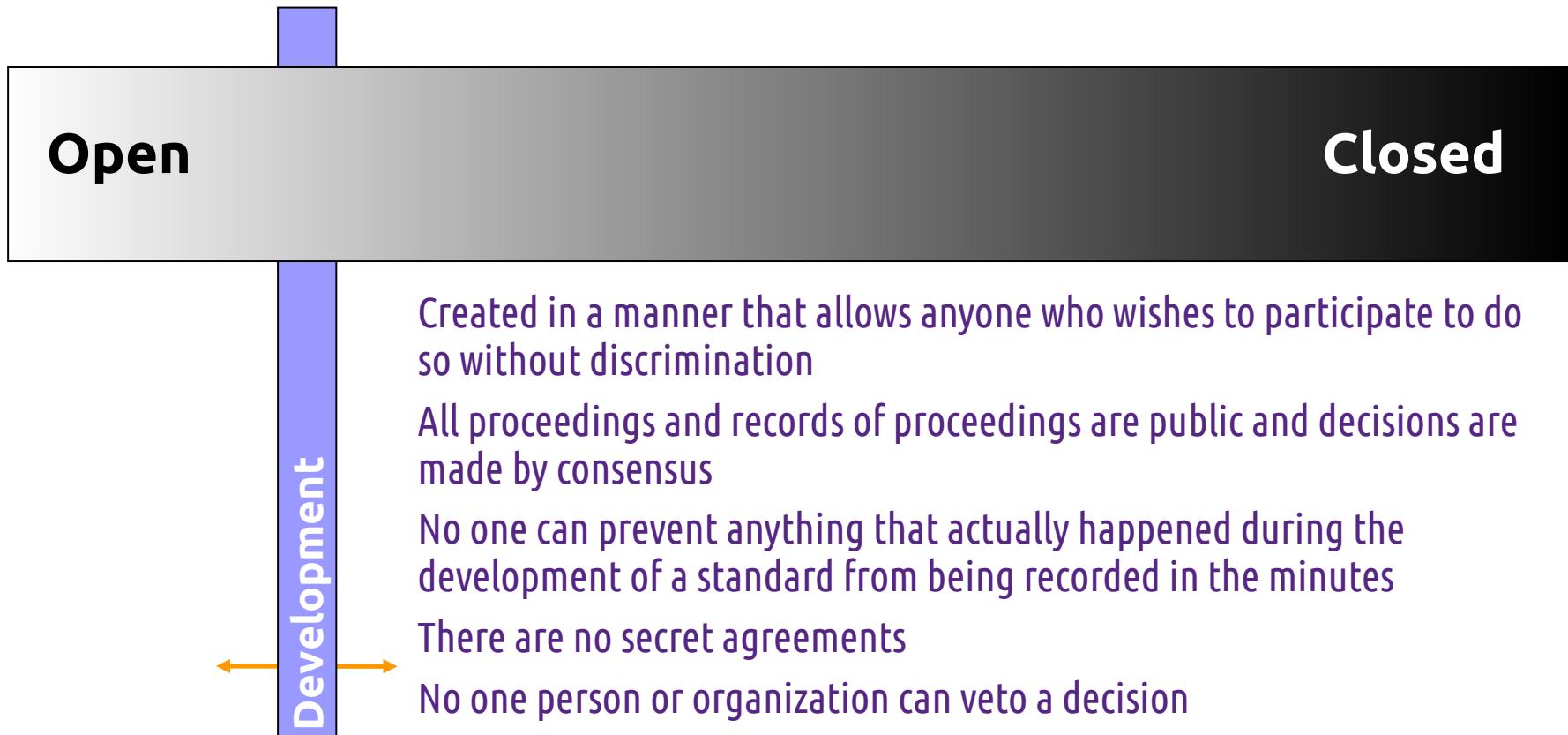
Italo Vignoli



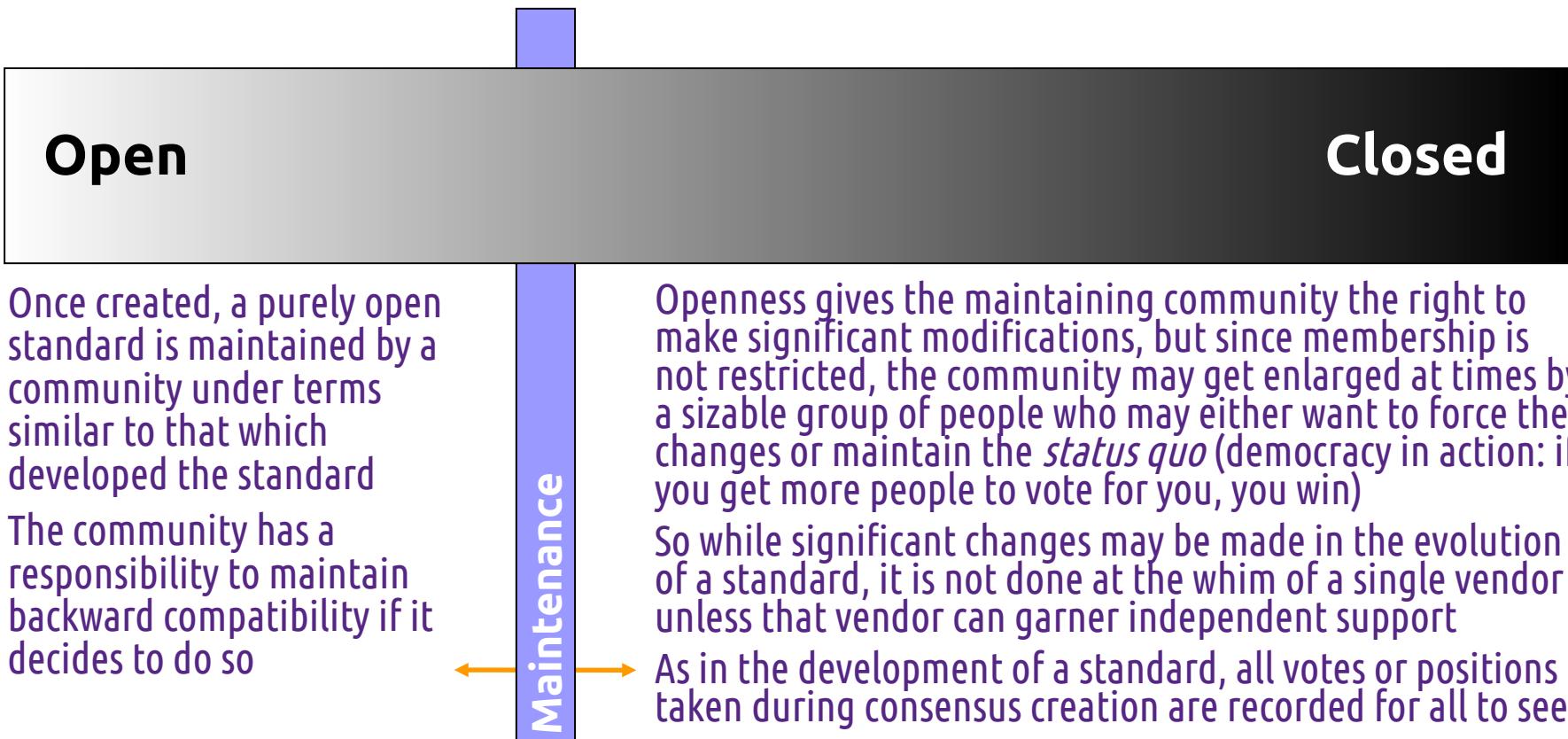
Open Standards. What Does Open Mean?



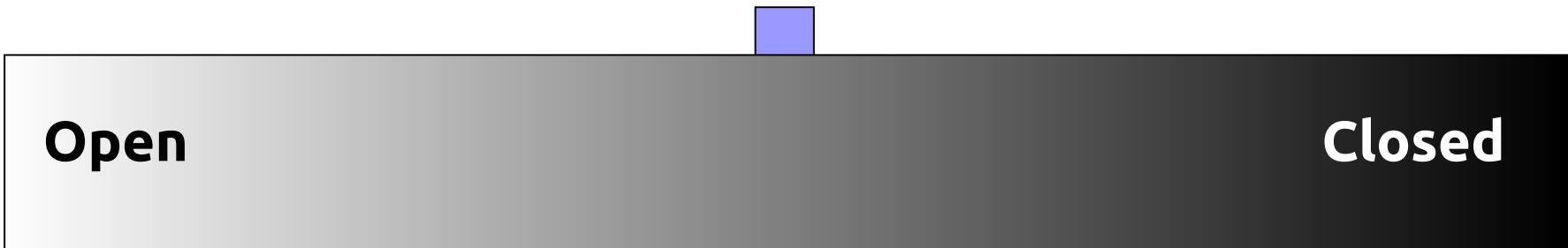
What Does Open Mean? Development



What Does Open Mean? Maintenance

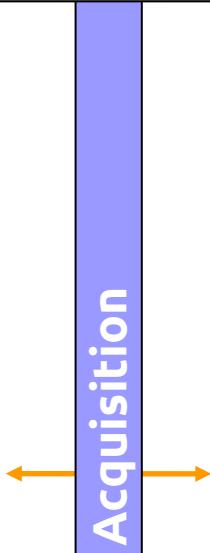


What Does Open Mean? Acquisition



Anyone who wishes to have a copy of a purely open standard may have one at no cost

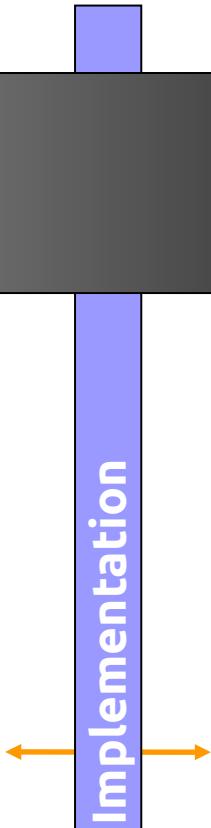
The standard is usually available online and the document format used to describe the specs is an open document standard



What Does Open Mean? Implementation



A true open standard can be implemented by anyone in any way desired with no payment of royalties



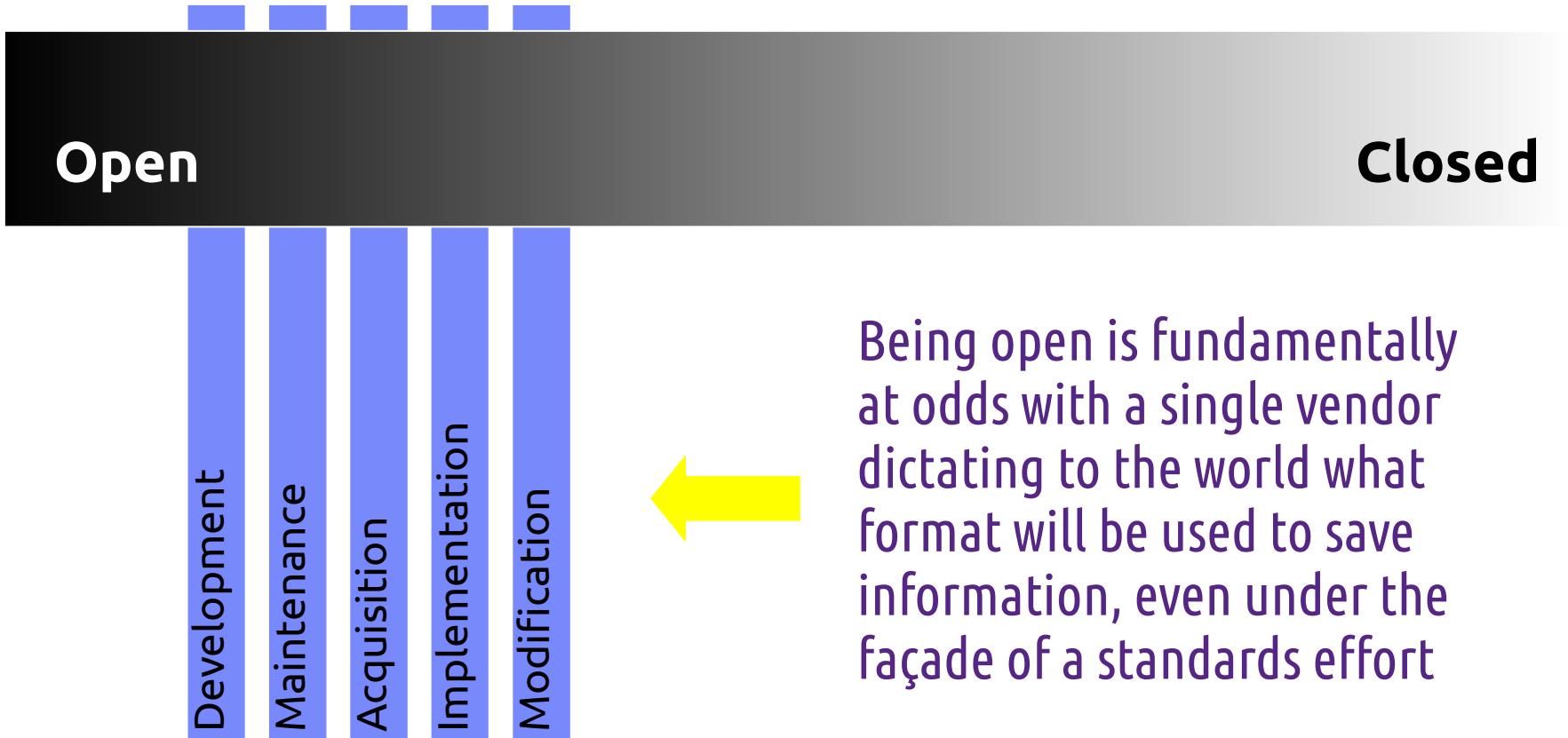
What Does Open Mean? Modification



A true open standard can be mined for its good ideas and used in part or in whole in other standards

In particular, a profile can be created that states which parts of multiple standards should be used together to achieve a certain end (for example, a subset of a set of standards for use on mobile phones)

Push to Open





Open Document Format

the true document standard
which offers freedom of choice



ODF @ OASIS

- ODF Technical Committee
 - Manages the technical evolution of the standard
 - Development oriented, focused on specifications
- ODF Advocacy Open Project
 - Just relaunched, announcement in May 2019
 - Education, communication & marketing oriented



Basic Concepts

- ODF is solid and robust
- ODF is consistent across OS
- ODF is truly interoperable
- ODF is predictable
- **ODF is a better standard file format
for users of personal productivity SW**



ODF Design Goals

Statement of Purpose

The purpose of this TC is to create an open, XML-based file format specification for office applications.

The resulting file format must meet the following requirements:

1. it must be suitable for office documents containing text, spreadsheets, charts, and graphical documents,
2. it must be compatible with the W3C Extensible Markup Language (XML) v1.0 and W3C Namespaces in XML v1.0 specifications,
3. it must retain high-level information suitable for editing the document,
4. it must be friendly to transformations using XSLT or similar XML-based languages or tools,
5. it should keep the document's content and layout information separate such that they can be processed independently of each other, and
6. it should 'borrow' from similar, existing standards wherever possible and permitted.



What is ODF? (1)

- OpenDocument Format (ISO / IEC 26300)
 - ISO: International Organization for Standardization
 - IEC: International Electrotechnical Commission
- Open file format based on XML, to create, view, edit and store Office documents
 - Text documents, spreadsheets, and presentations



What is ODF? (2)

- Defined with an open and transparent process by OASIS
- Approved by the Joint Technical Committee 1 (JTC 1) of the IEC as International Standard (IS) in May 2006
- Available for deployment and use with no license, royalty payments or other restrictions



ODF Advocacy on GitHub

 OASIS OPEN PROJECTS

Welcome to the home of the ODF Advocacy Open Project.

The ODF Advocacy project works to create awareness and educate the world about the benefits of using the OpenDocument Format OASIS Standard (also published as ISO/IEC 26300).

ODF is an XML-based file format for personal productivity applications such as office suites, word processors, text/document editors, spreadsheets, and presentation software. Use of ODF guarantees access to your data forever, ensuring that data can be transferred between different computers and operating systems, without having to worry about vendor lock-in or license fees.

The ODF Advocacy project develops non-commercial, informational materials for a sustained communication campaign about the technical advantages and cost-savings of using standard-based document interoperability over proprietary formats.

ODF Advocacy is an [OASIS Open Project](#).



Charter

- The objective of the ODF Advocacy Open Project at OASIS is to create awareness and educate the world about the benefits of using ODF
- The group is developing advocacy materials for a communication campaign about the importance of standard-based document interoperability versus proprietary formats



Purpose and Scope

- **Awareness:** Dedicated Website, Social Media Campaign, Media Outreach Campaign, Conference Program, Plugfests
- **Education:** What is a Document Standard?, Importance of Document Standards, Advantages of Document Standards, De Jure vs De Facto Standards
- **Marketing:** Impact of Standards on Productivity, Economic Value of Interoperability, Competitive Advantages of ODF, Standards & Innovation



Business Benefits

- Users of personal productivity software will benefit from the ODF awareness and education campaign, because they will learn how true document interoperability can protect their data into the future and ensure freedom-of-choice
- Organizations deploying personal productivity software will discover the advantages and savings made possible by ODF, and the increased security based on the lower number of vulnerabilities affecting files based on standard file formats



Purpose and Scope

- Educate users of personal productivity applications such as office suites about the advantages of adopting the international, ISO-published, ODF standard
- Help users migrate from pseudo-standard formats to ODF to ensure interoperability and save costs



Digital Document

- Can be used only by those who have access to the decoder
- Primary purpose of a digital document is to use it in the future
- It should be readable and interpretable as long as possible, and ideally forever



Open Document Format

- **Independent** from a single product: anyone can write a software that handles an open format
- **Interoperable**: allows the transparent sharing of data between heterogeneous systems
- **Neutral**: it does not force the user to adopt – and often buy – a specific product, but leaves a wide choice based on features/quality vs price ratio
- **Perennial**: protects user developed contents from the “evolution” based obsolescence of technology



When the Decoder is Proprietary

- Your own ideas, encoded in a digital document, are at the mercy of the owner of the decoder
- You have lost your right to access and read your own documents in the future
- This possibility is dangerous for a digital society

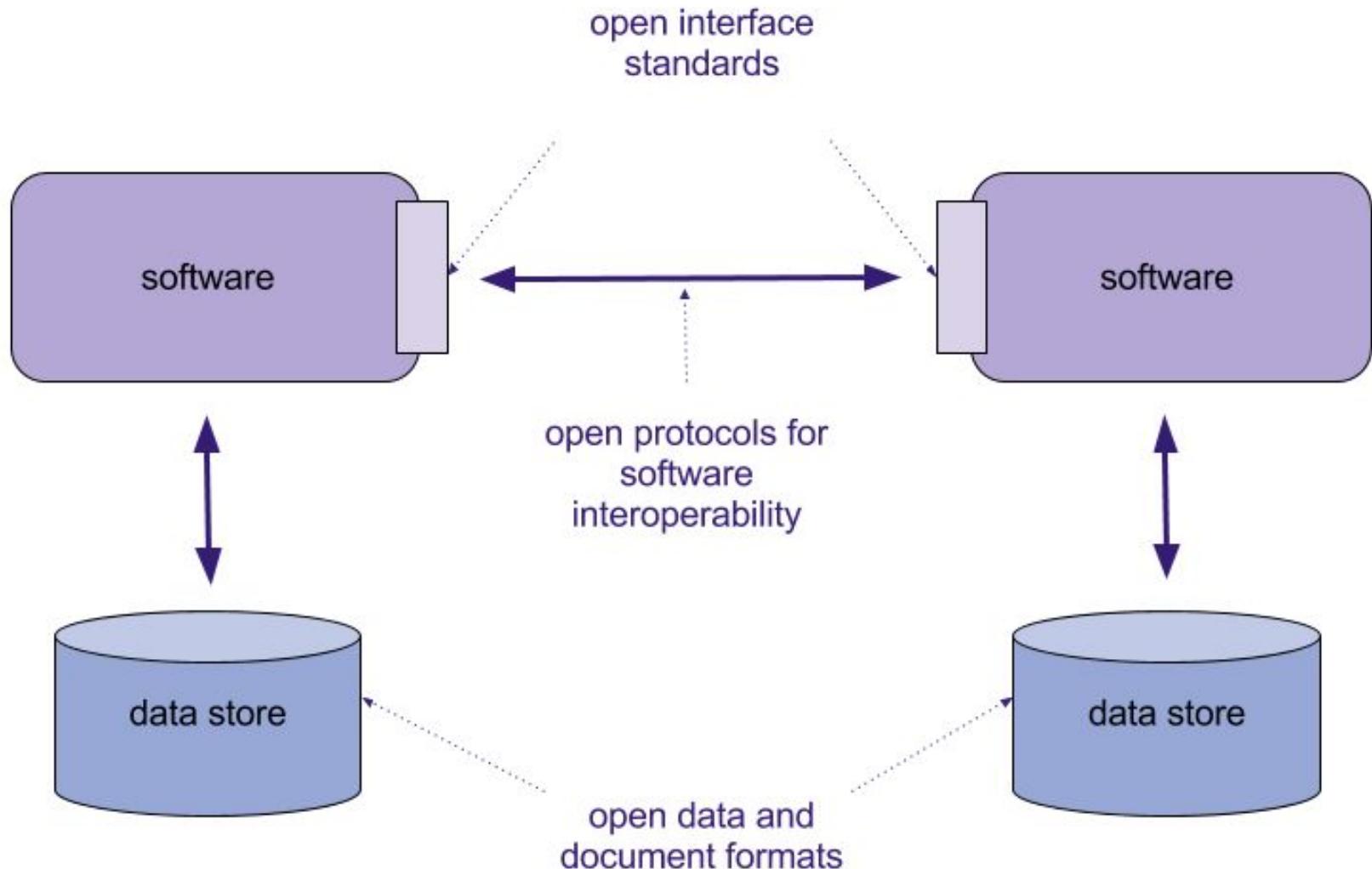


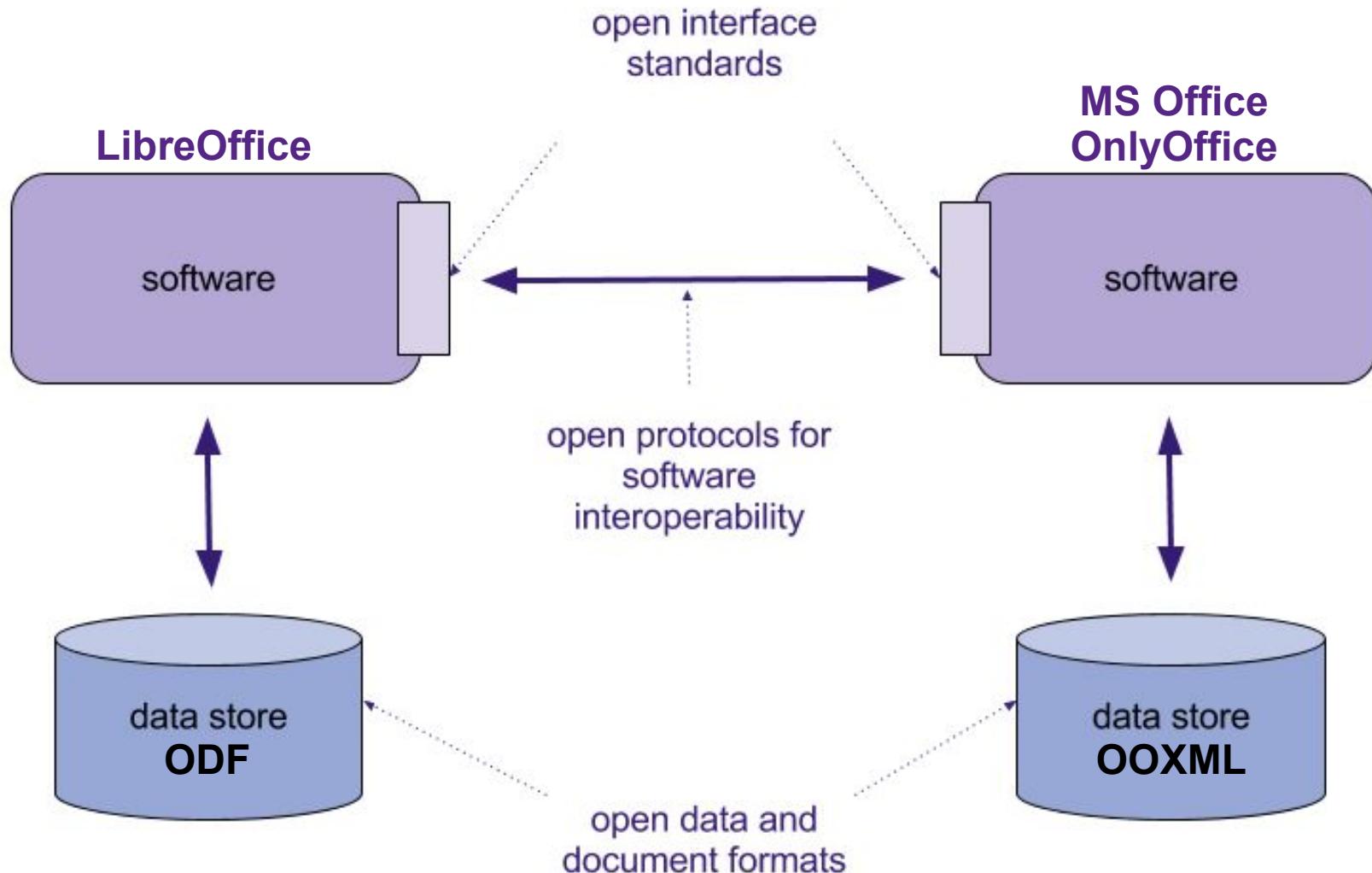
Code without Decoder

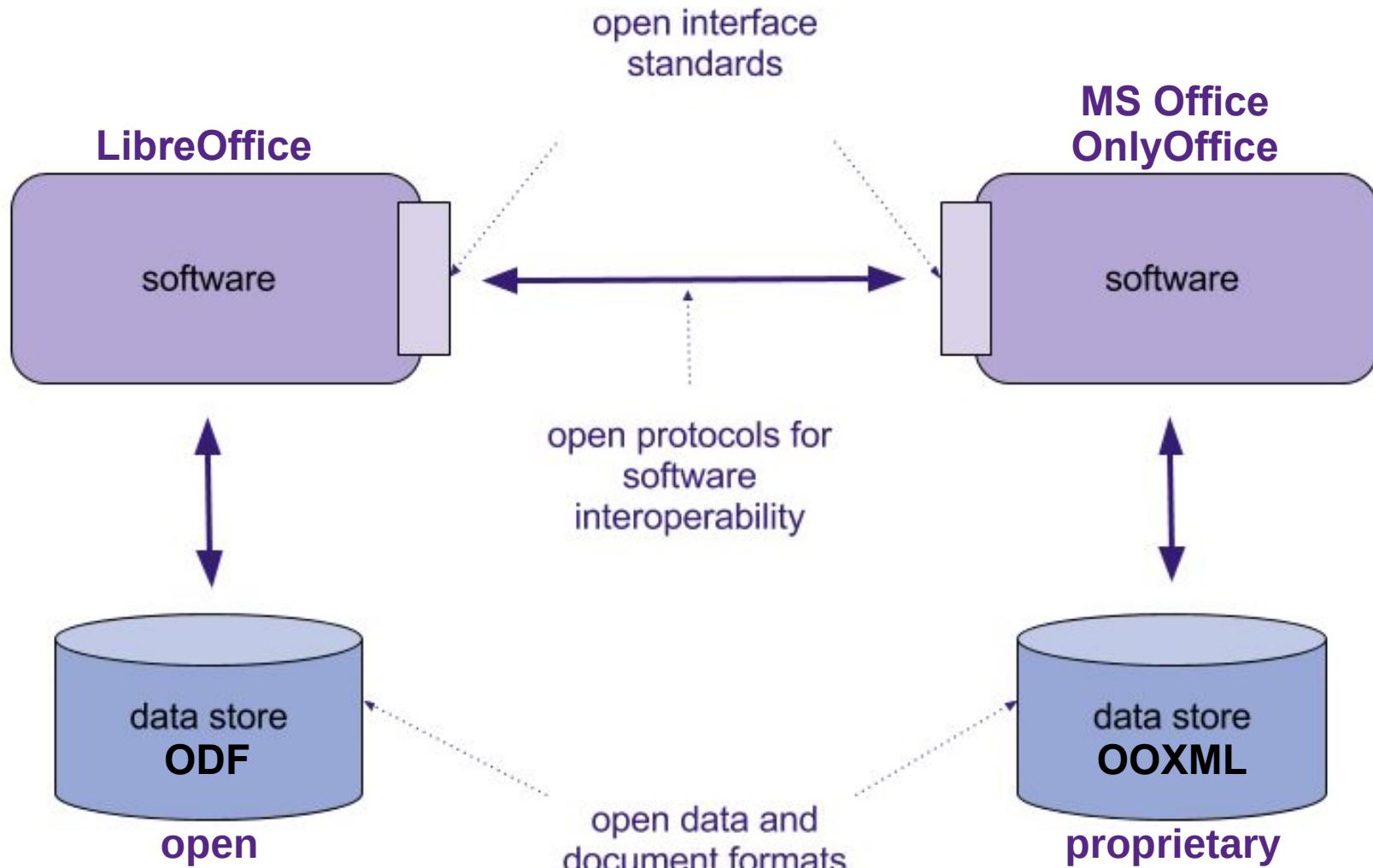
- Our ancestors inscribed on a hard stone, hoping to preserve their expression for posterity
- In fact, they managed to preserve the code, but not the decoder
- Lesson: an unused language becomes like a stone, as using a language means interpreting it over and over again











Lock In

WE CANNOT READ YOUR DOCUMENTS

DOCUMENTFREEDOM.ORG



PSF

How to Lock-in Your Clients

How Professional Services Firms Can Create Compelling Value for Clients Using Collaborative Technologies

Ross Dawson
CEO, Advanced Human Technologies
Author, *Living Networks and Developing Knowledge-Based Client Relationships*

January 2004

>> A STRATEGIC WHITE PAPER FROM MICROSOFT BUSINESS SOLUTIONS

UK Cabinet Office

- Open standards are key for software, data and document interoperability
- Needs of users should be at the heart of any choice of standards
- Users, whether they work for the Government or are just citizens, have the right to read all official documents
- They should not be expected to purchase new equipment or install new software in order to read an official document





© David // Armano

darmano.typepad.com

Internet Nodes



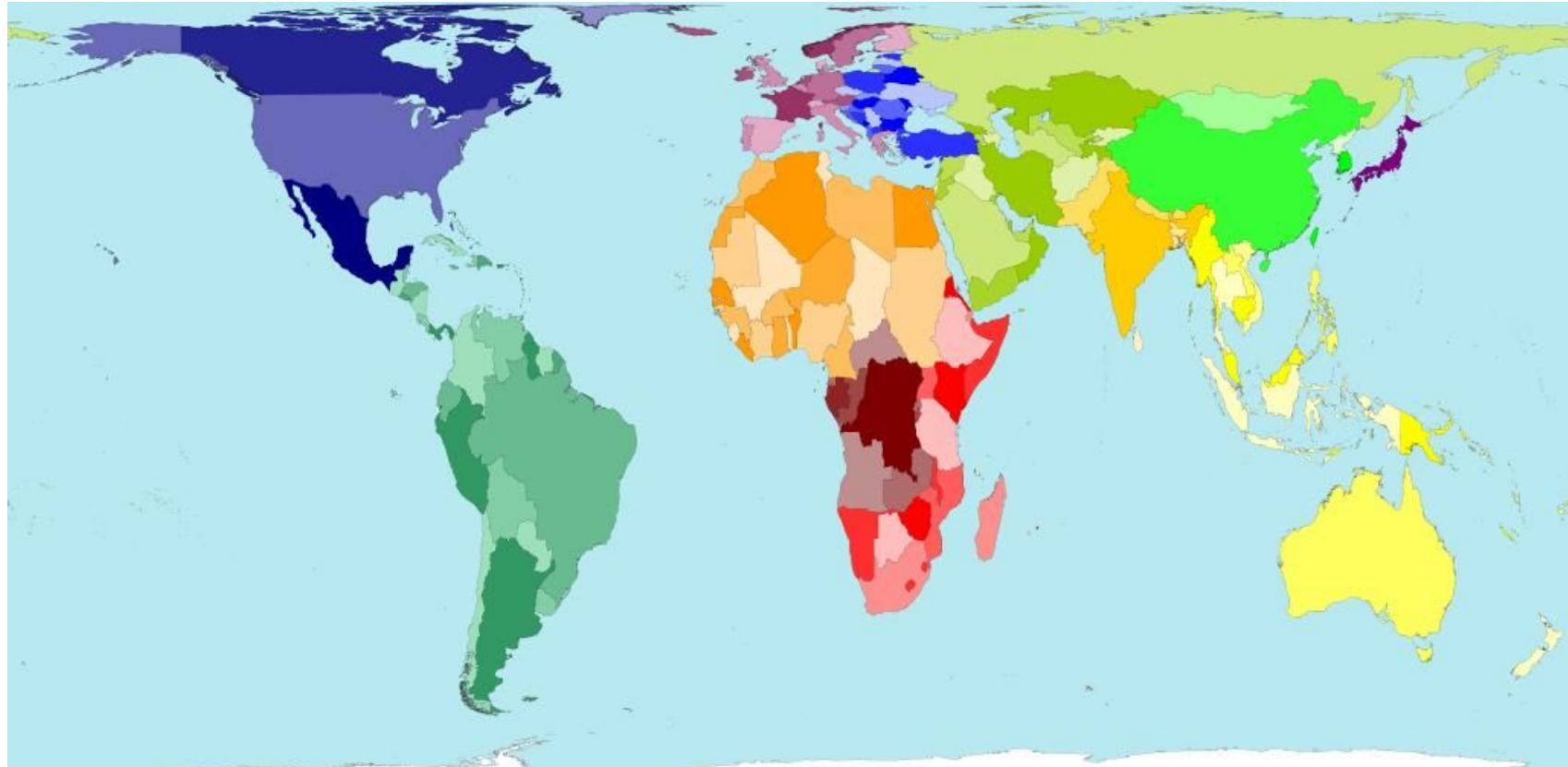
Facebook Conversations



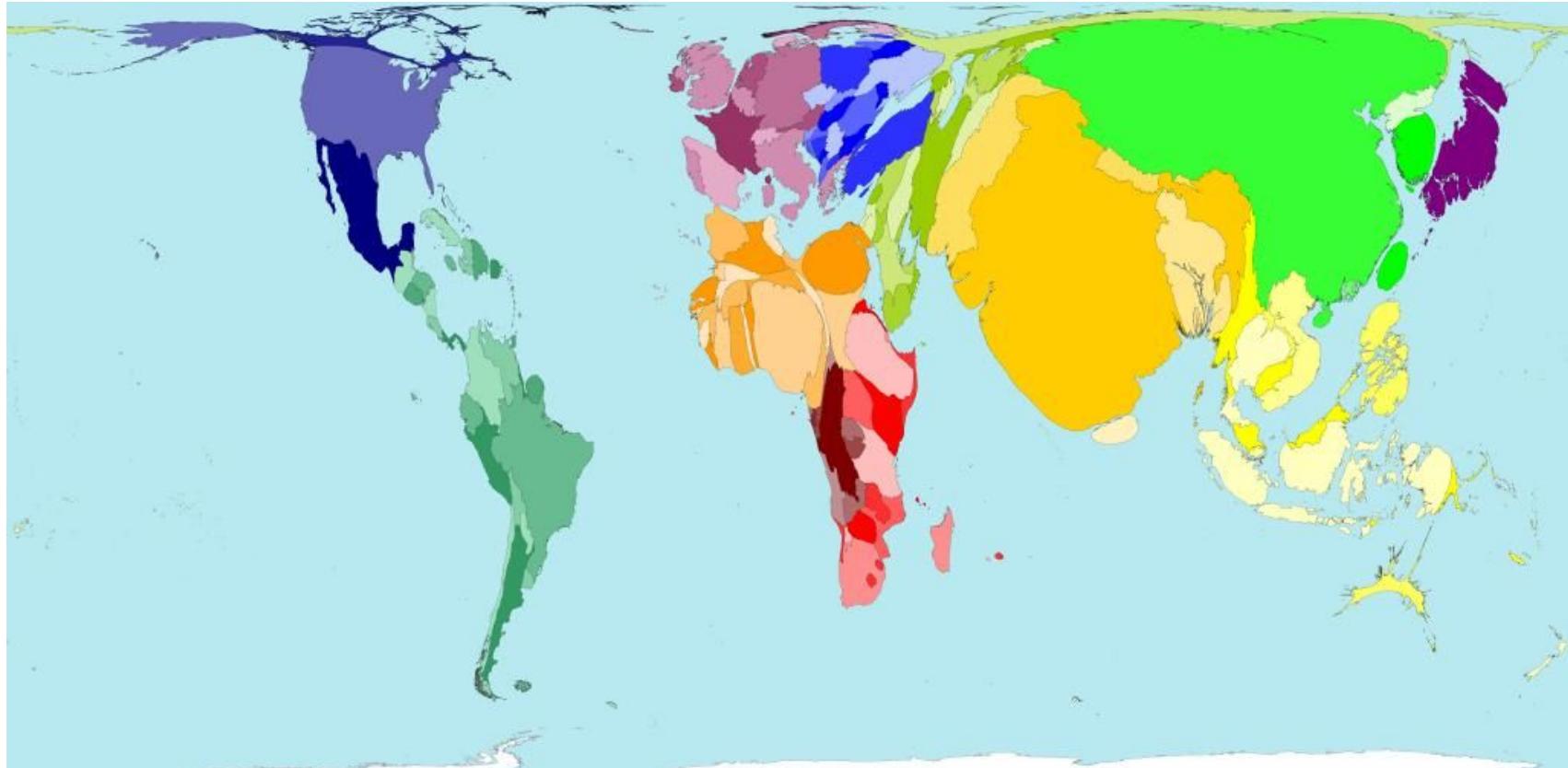
Twitter Threads



Countries by Surface



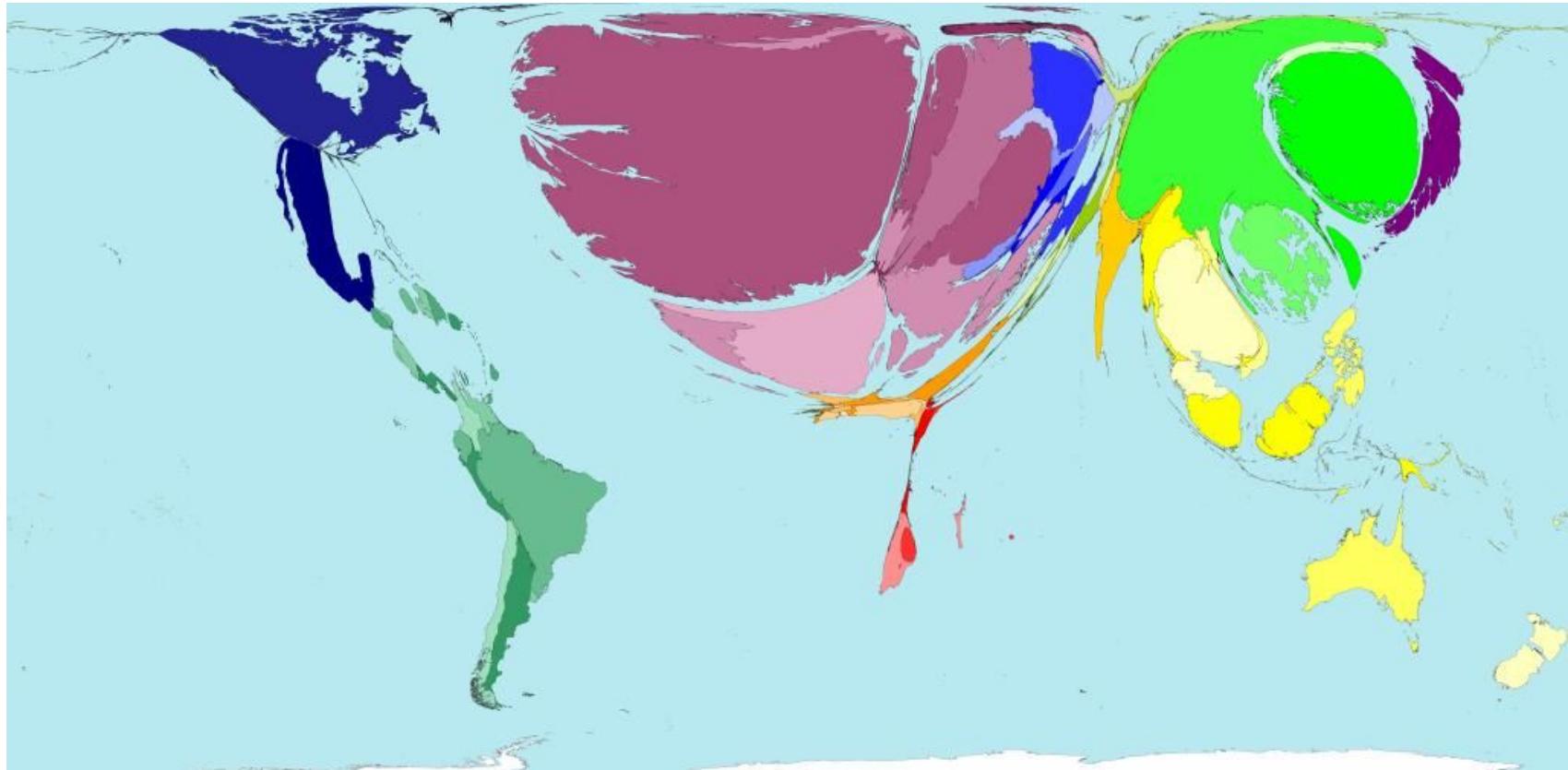
Countries by Population

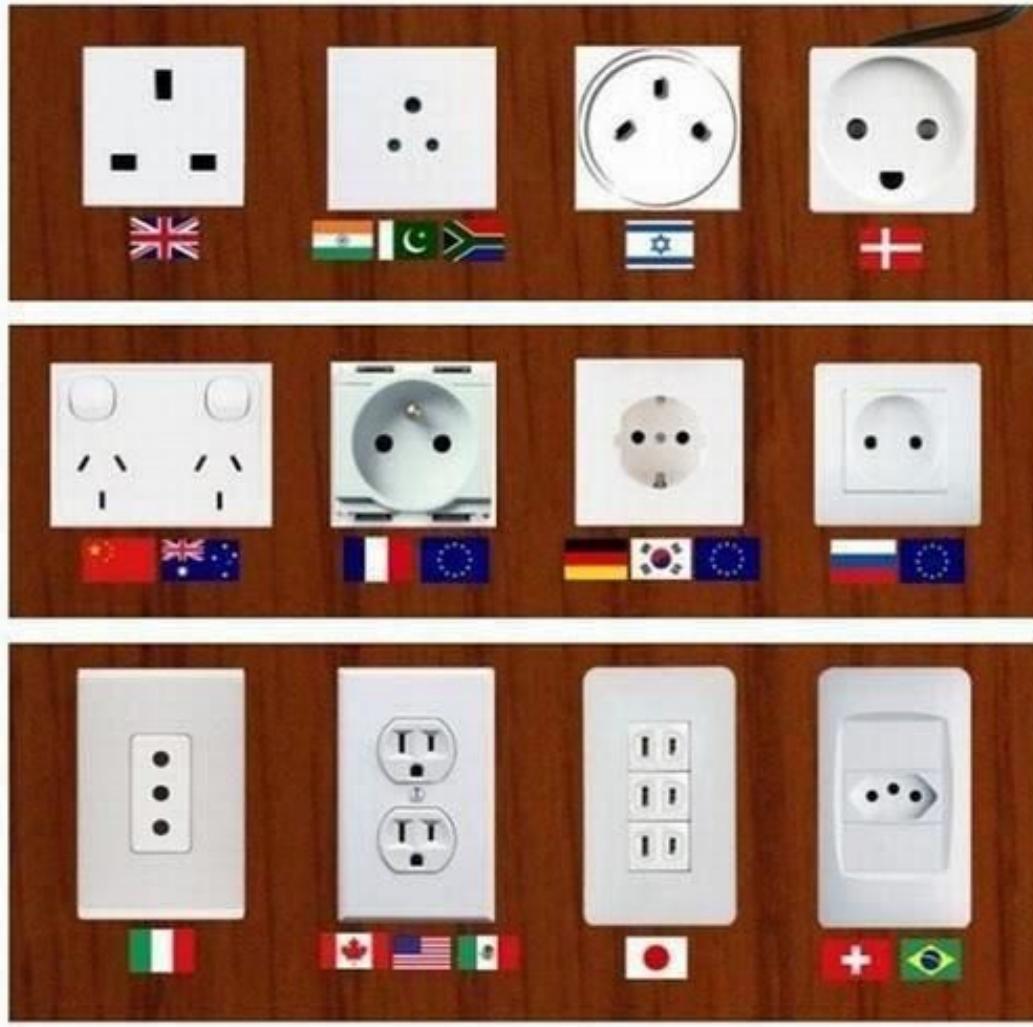


Countries by Export of SW Licences



Countries by Import of SW Licences





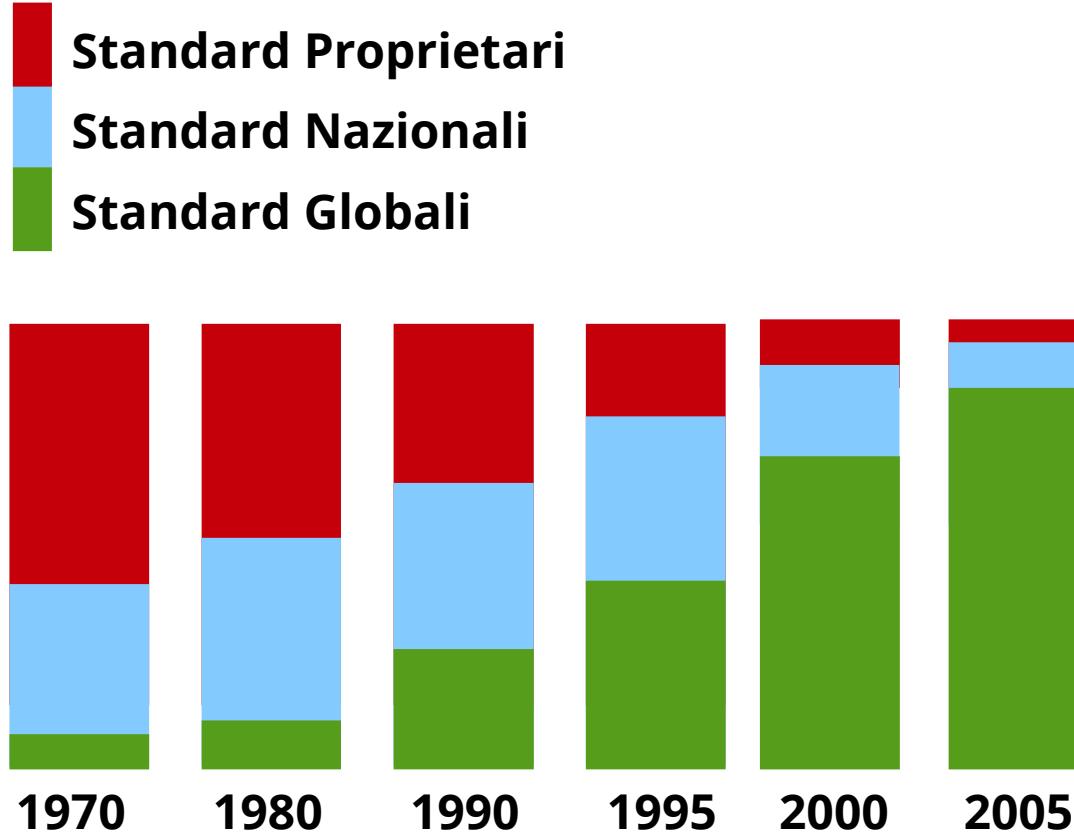
Importance of the HTML Standard

- It was the standardization of the HTML format that allowed the web to take off. And not just the fact that it's a standard, but the fact that it's open and royalty-free...
- Had HTML not been free and open, and a proprietary technology, the business of selling HTML and competing products would have been born...
- This means we need standards, because this avoids competition over technology, and fuels the value-added business built on the platform...

*Tim Berners-Lee, CERN
world wide web inventor*



Adoption of Standards



What is a Standard?

- A document, established by consensus and approved by a recognized body, that provides rules, guidelines or characteristics for activities or their results, aimed at the achievement of the optimum degree of order in a given context, for common and repeated use
- Standards should be based on the consolidated results of science, technology and experience, and aimed at the promotion of optimum community benefits

ISO/IEC Guide 2:2004 "Standardization and Related Activities - General Vocabulary"



Proprietary Standard

- Natural monopoly in technology leads to natural monopoly in market for products and services based on that technology
- Results when access to the technology is available only to the rights holders



Semi-Open Standard

- Natural monopoly in technology arises (de facto) or is defined (de jure) but some competition provided for in market for products and services
- Results when access to the technology is available to players other than the rights holders/originators, who perhaps retain some advantages



Open Standard

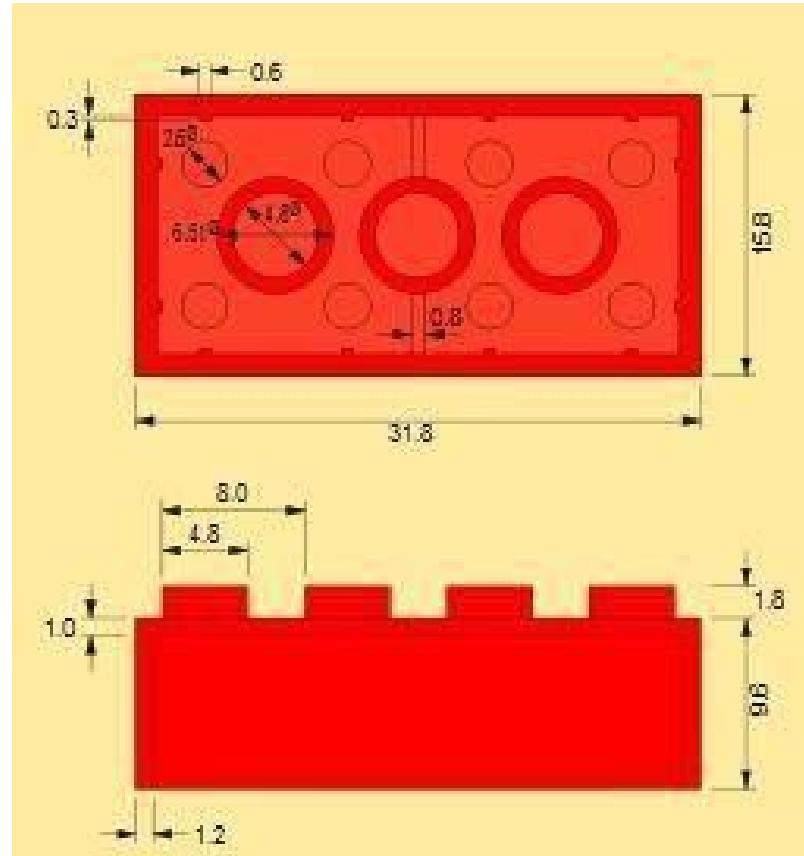
- Natural monopoly in technology arises (de facto) or is defined (de jure) but full competition ensured in market for products and services
- Results when access to the technology is available to all (potential) players on equal terms providing no a priori advantages based on ownership of rights, or definition of the technology



Standard and Interoperability

Interoperability is the ability of information and communication technology (ICT) systems, as well as of the business processes they support, to exchange data and enable the sharing of information and knowledge.

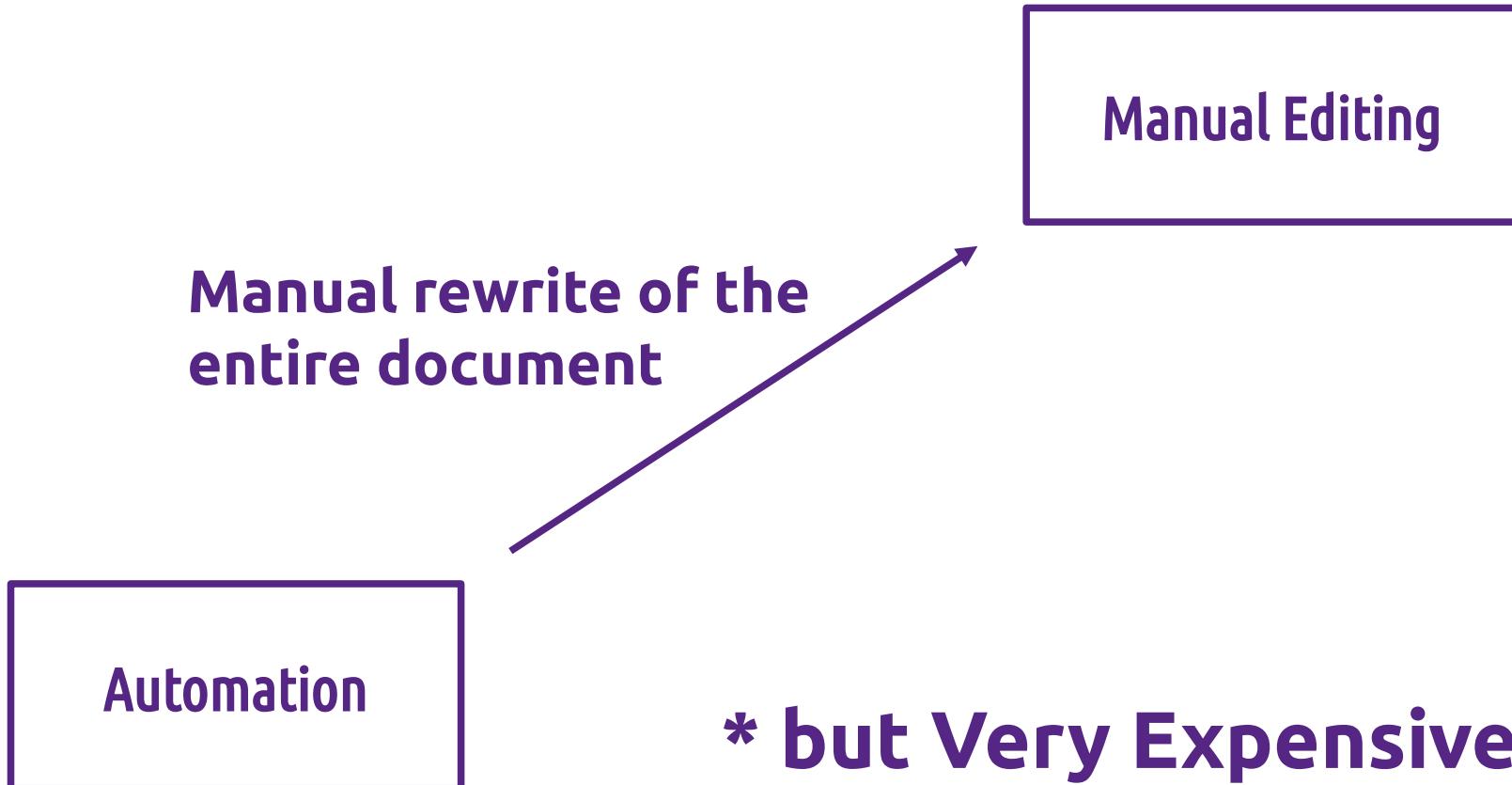
*European Interoperability
Framework, IDABC*

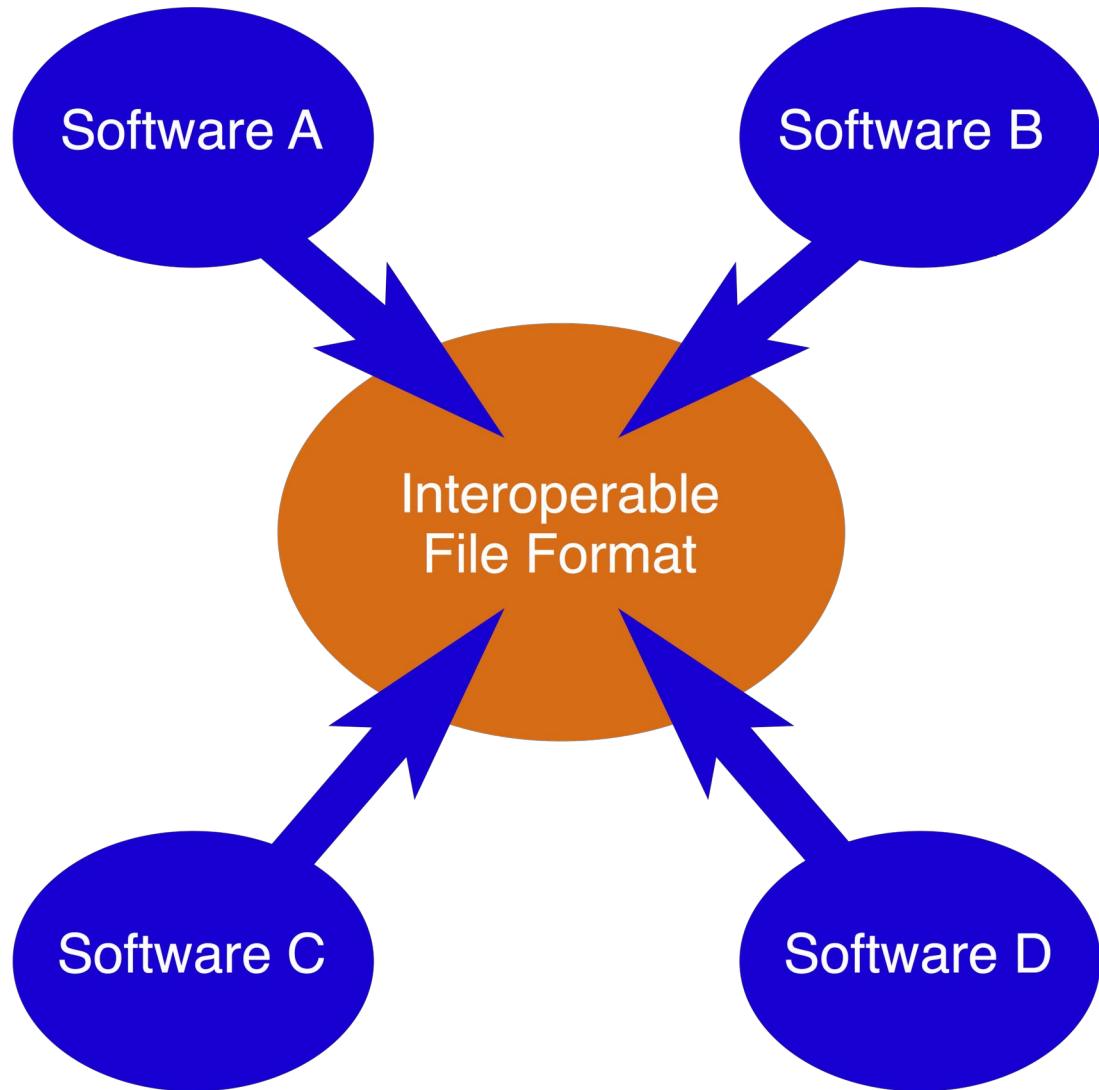


Benefits of Interoperability

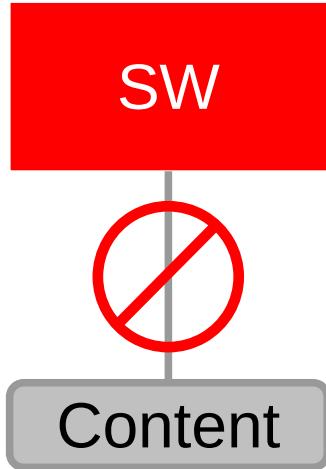
- **FINANCIAL**
 - Cost savings for users
 - Reduced operation costs for owners
 - Vendor lock-in avoidance for owners
 - Facilitate reuse, sharing & adoption
- **QUALITY**
 - High service satisfaction for users
 - Improved compliance for owners
 - Better data quality for owners
 - Better data availability for users
 - Improved security for owners
- **TIME**
 - Owner time savings
 - User time savings
- **OTHER**
 - Foster innovation
 - Increase transparency
 - Protection of user's rights
 - Furthering public policy goals
 - Reduced CO2 emissions

Perfect Interoperability is Easy *





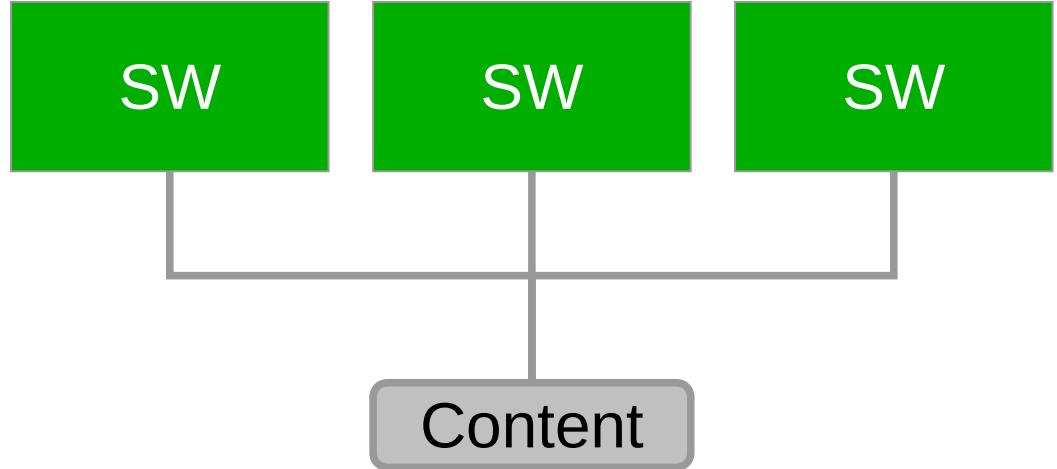
ODF Based Interoperability



Old Style

Content closely related to the application used to create it

Controlled by the application developer and not by the user



New Style

Content represented through an open standard which is not controlled by a single vendor, so many applications can create and modify it

Controlled by the user and not by the software vendor

Interoperability without Standards



Cost of Inadequate Interoperability (1)

NIST GCR 04-867



U.S. Department of Commerce
Technology Administration
National Institute of Standards and Technology

Advanced Technology Program
Information Technology and Electronics Office
Gaithersburg, Maryland 20899

Cost Analysis of Inadequate Interoperability in the U.S. Capital Facilities Industry

Michael P. Gallaher, Alan C. O'Connor, John L. Dettbarn, Jr., and Linda T. Gilday



Cost of Inadequate Interoperability (2)

Table 6-5. Costs of Inadequate Interoperability for Architects and Engineers

Life-Cycle Phase	Cost Category	Cost Component	Average Cost per Square Foot	Average Cost per Square Meter	Inadequate Interoperability Cost Estimate (\$Thousands)
Planning, Engineering, and Design		Inefficient business process management costs	0.31	3.37	356,126
		Redundant CAx systems costs	0.0001	0.001	158
		Productivity losses and training costs for redundant CAx systems	0.04	0.45	47,947
		Redundant IT support staffing for CAx systems	0.0004	0.005	501
		Data translation costs	0.002	0.02	2,139
Avoidance Costs		Interoperability research and development expenditures	0.02	0.21	22,234
		Manual reentry costs	0.41	4.38	462,734
Mitigation Costs		Design and construction information verification costs	0.10	1.08	114,342
		Reworking design files costs	0.0009	0.009	968
		Avoidance costs	0.38	3.85	429,106
		Mitigation costs	0.51	5.47	578,044
Subtotal		Subtotal	0.89	9.32	1,007,150



Cost of Inadequate Interoperability (3)

Construction	Avoidance Costs	Inefficient business process management costs	0.04	0.41	43,290	
		Redundant CAx systems costs	0.00003	0.0003	28	
		Productivity losses and training costs for redundant CAx systems	0.007	0.08	8,461	
		Redundant IT support staffing for CAx systems	0.00008	0.0008	88	
		Data translation costs	0.0003	0.004	378	
		Interoperability research and development expenditures	0.003	0.04	3,924	
		Manual reentry costs	0.024	0.26	27,750	
Mitigation Costs		Design and construction information verification costs	0.006	0.07	7,377	
		RFI management costs	0.05	0.53	55,656	
Subtotal		Avoidance costs	0.05	0.49	56,169	
		Mitigation costs	0.08	0.86	90,783	
		Subtotal	0.13	1.35	146,952	
Operations and Maintenance	Mitigation Costs	Post-construction redundant information transfer costs	0.01	0.15	15,660	
Total Cost					1,169,762	

Source: RTI estimates; totals may not sum correctly due to rounding.



Document Formats

- **Pseudo standard for document formats**
 - Microsoft Office document format (OOXML)
 - DOC, DOCX, XLS, XLSX, PPT, PPTX
- **True standard for document formats**
 - Open Document Format, native document format of LibreOffice and many other software (including Microsoft Office)
 - ODT (text), ODS (spreadsheet), ODP (presentation)



Main Characteristics of ODF

- Allows access to documents regardless of the application with which they were created (and its version)
- Shifts the focus from an application-focused model to a document-centered one
- Helps users think in terms of interoperability



ODF is a Standard

- France
- Netherlands
- Portugal
- United Kingdom
- Sweden
- Taiwan



United Kingdom

- Decision to switch to ODF after a two year process, with extensive research into the two standards
- Different government bodies have different agendas, so the migration to ODF is slower than expected by the Cabinet Office (which was rather optimistic)
- Brexit process has unfortunately changed the priority for several government bodies
- Decision to move to ODF has never been questioned



ODF is Standard in UK



Press release

Open document formats selected to meet user needs

From: Cabinet Office, The Rt Hon Francis Maude and Government Digital Service

First published: 22 July 2014

Part of: Government transparency and accountability and Government efficiency, transparency and accountability



France

- Move to ODF based on the following document:
 - Direction Interministérielle du Numérique et du Système d'Information et de Communication de l'Etat, Référentiel Général d'Interopérabilité [standardiser, s'aligner et se focaliser pour échanger efficacement]"
- The document mentions ODF and only OOXML Strict, which is not possible to detect unless you open the underlying XML text to find the "<w:conformance="strict">" variable
 - As such, public administrations are only allowed to use ODF (unless they put in place a mechanism to recognize OOXML Strict)

The Netherlands

- ODF is recognized by the Dutch law as the only document standard format allowed in public administrations
- Unfortunately, the law does not sanction use of OOXML, so many public administrations are still using Microsoft Office document formats
- Dutch government has hired ODF specialists to help public administrations



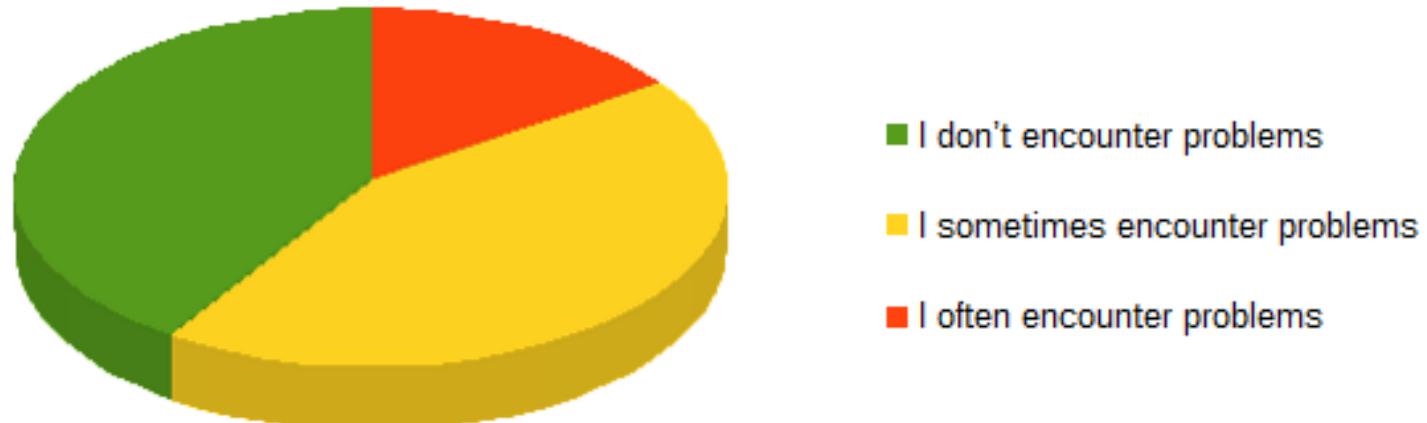
Sweden

- Move to ODF based on the list of software supporting open standards which can be purchased by public admins: “KammarKollegiet, State Purchasing Center, Open Standards, Software and Services 2014”
- Swedish public administrations are requested to purchase only applications listed in the document, to comply with the government open standard policy
- ODF is listed while OOXML is not listed, which means that public administrations must switch to ODF



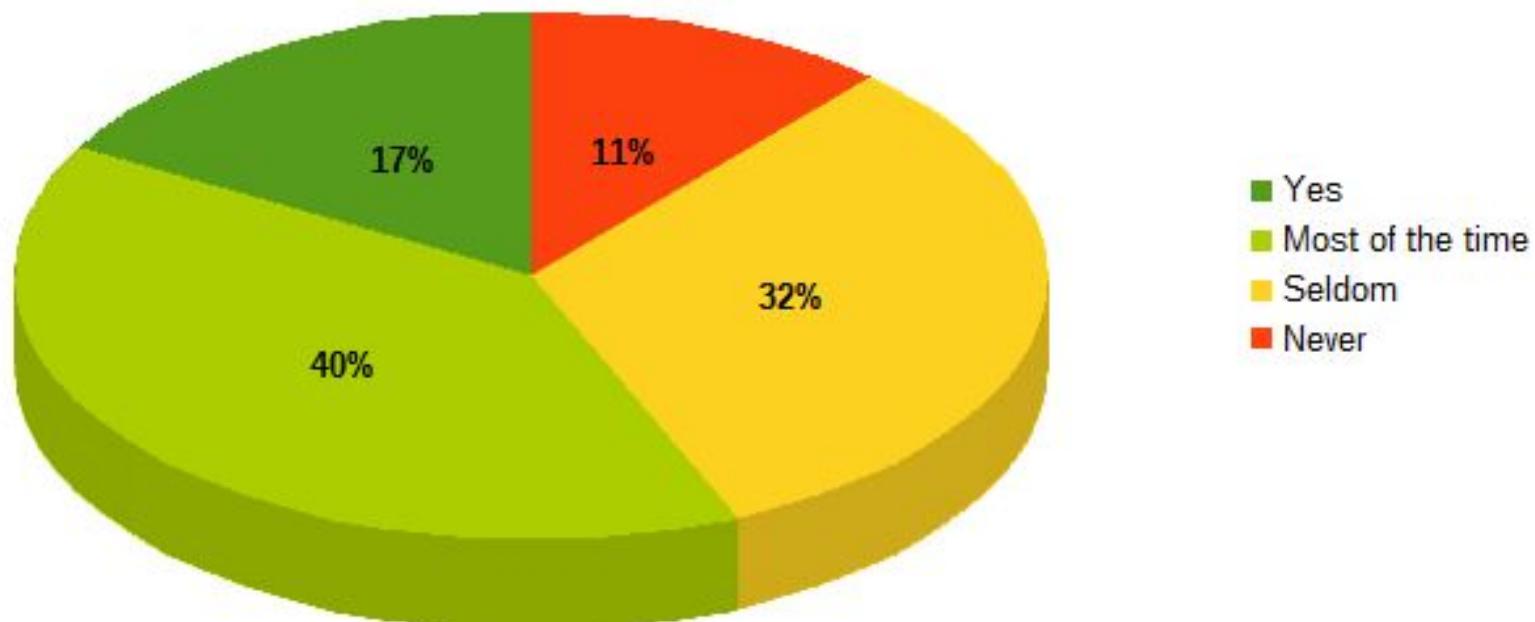
Research from Nantes Metropole

What happens when sharing documents outside the organization?



Research from Nantes Metropole

Do staff from other companies accept to use LibreOffice when you ask them?



Nantes: Dealing with Others

- SUPPLIERS & SATELLITES
 - Wherever we have control we require ODF usage
- OTHER ADMINISTRATIONS, PARTNERS
 - Where we don't have control we recommend ODF



Migration Experiences (1)

Project	Desktops
French Government	500.000
Region of Valencia (Spain)	120.000
Italian Ministry of Defence	100.000
French Gendarmerie	72.000
Hospitals in Copenhagen Region	25.000
Taiwan Ministry of Finance	24.000
City of Toulouse (France)	9.000

Migration Experiences (2)

Project	Desktops
Lithuania Police Department	8.000
Central Region of Denmark	8.000
Taiwan Yi-Lan County	8.000
Schools in Tallinn (Estonia)	4.000
City of Bari (Italy)	1.700
Healthcare 5 Veneto (Italy)	1.500
City of Las Tirana (Albania)	1.000

Migration Experiences (3)

Project	Desktops
Healthcare System	45.000
Federal Bailiffs Service	35.000
Several Regional Administrations	unknown
Russian Helicopters	unknown
VimpelCom	unknown
Gazprom Teploenergo Yaroslavl	unknown
Citygroup	31.000

Digging into Document Formats



***The Banana standard for your
worst Office documents***

De Jure vs De Facto Standards

- A **de facto** standard refers to a significant market share
- A **de jure** standard is based on a collective agreement
- As such they are innately different, as are their value and effect on the market
- **De jure standards for document formats**
 - Foster interoperability, create network externalities, prevent lock-in, cut transaction costs, create a transparent market and reduce variety
- **De facto standards for document formats**
 - Tend to be the exact opposite, to increase supplier-dependence and create an obfuscated market

One vs Two De Jure Standards

	Effect on the Market	
	One Standard	Two Standards
Information	<ul style="list-style-type: none">• Increases market transparency• Reduces transaction costs (and information asymmetry)• Correct wrong selection of tool• Facilitates exchange	<ul style="list-style-type: none">• Reduce market transparency• Increase transaction costs (need of converters and converting)• Make comparison of product quality more difficult• Hinder trade with higher information transaction costs
Compatibility	<ul style="list-style-type: none">• Creates network externalities and increases competition• Decreases vendor lock-in	<ul style="list-style-type: none">• Reduce interoperability and adds switching costs• Reduce network externalities and decrease competition• Increase likelihood of standard based vendor lock-in
Variety Reduction	<ul style="list-style-type: none">• Allows economies of scale• Facilitates building a critical mass	<ul style="list-style-type: none">• Reduced economies of scale• Reduced chances of building a critical mass



Standardization Process

ODF

- Based on OOo XML format
- Dec 12, 2002: document format presented to OASIS
- May 1, 2005: ODF released by OASIS
- Nov 16, 2005: ODF presented to ISO/IEC JTC1 based on Publicly Available Specification (PAS)
- May 3, 2006: ODF approved as ISO/IEC IS 26300 standard
- Review: 720 pages in 1239 days

OOXML

- Based on Microsoft Office 2003 XML format
- Dec 15, 2005: document format presented to ECMA
- Dec 31, 2006: ECMA standard approved by General Assembly
- Jan 31, 2007: OOXML presented to ISO/IEC JTC1 based on Fast Track
- Mar 31, 2008: OOXML pseudo standard approved
- Review: 7200 pages in 838 days



Reuse of Existing Standards

ODF

- Dublin Core
- XLS:FO
- SVG
- MathML
- XLink
- SMIL
- XForms

OOXML

- Dublin Core

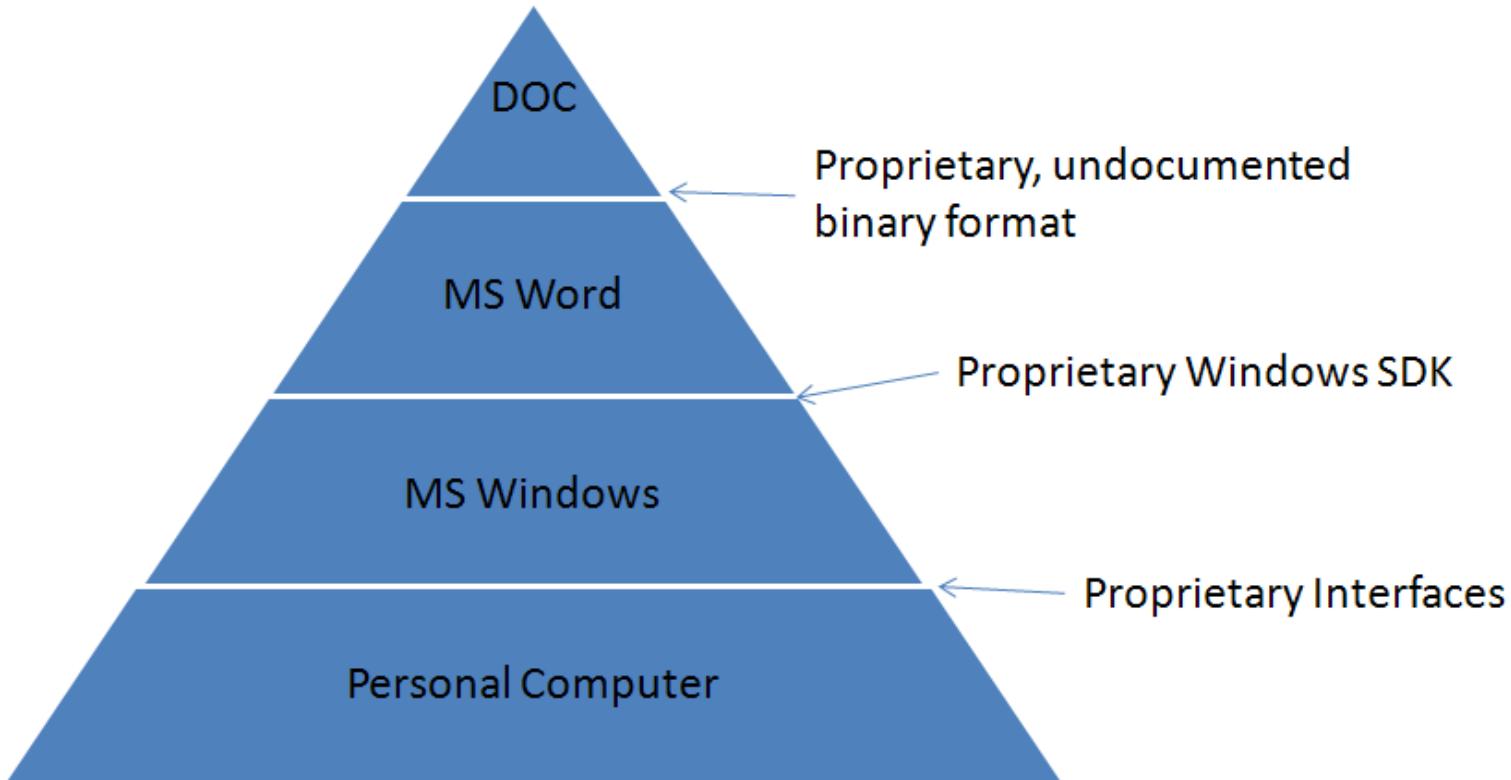


Proprietary Document Format

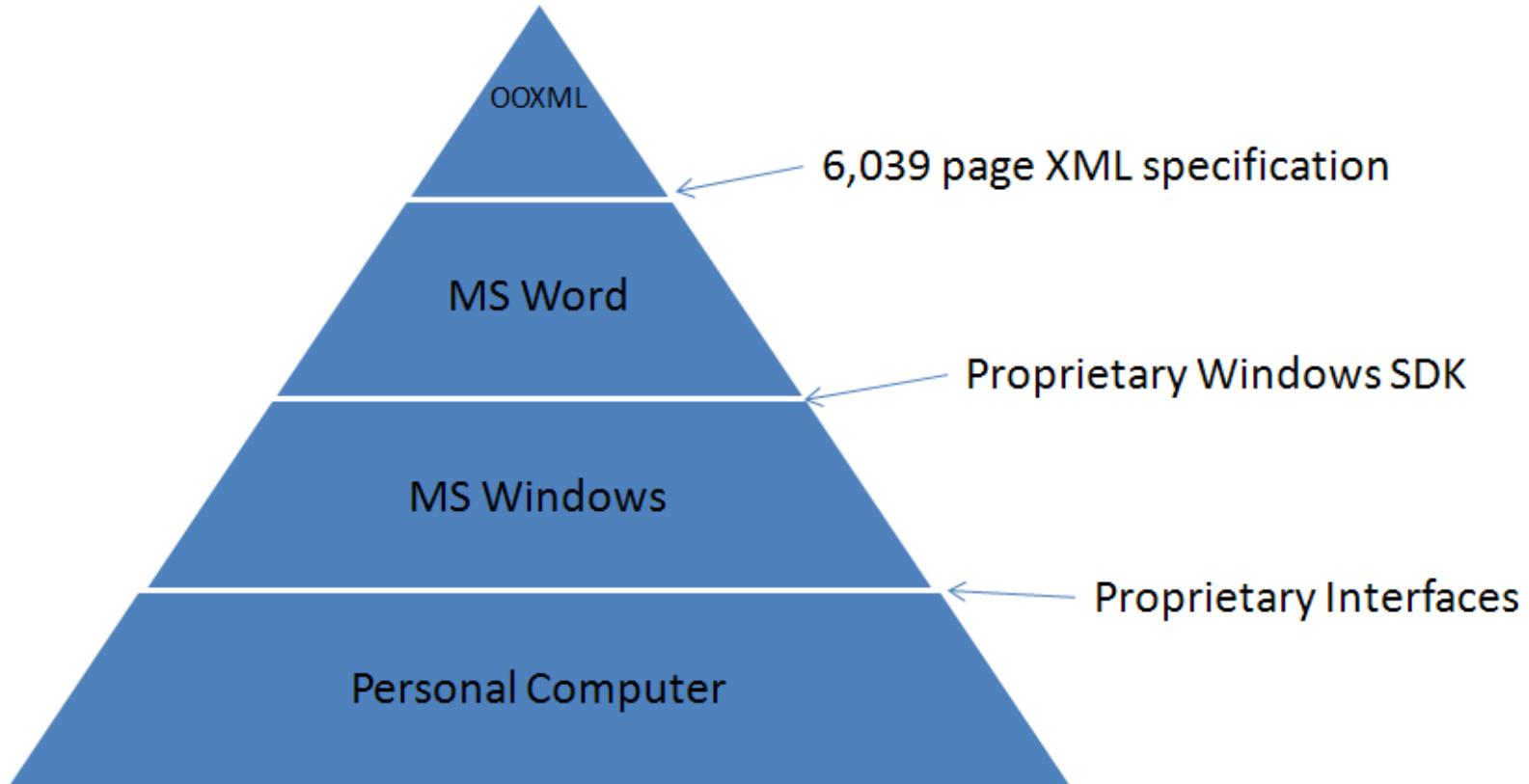
- Designed to be manipulated by a single software
- Evolves over the years based on commercial strategies and not on user needs
- Often, a direct serialization of data structures in memory
- The software is the format!
- Users borrow content from vendors through the format
End User License Agreement (EULA)



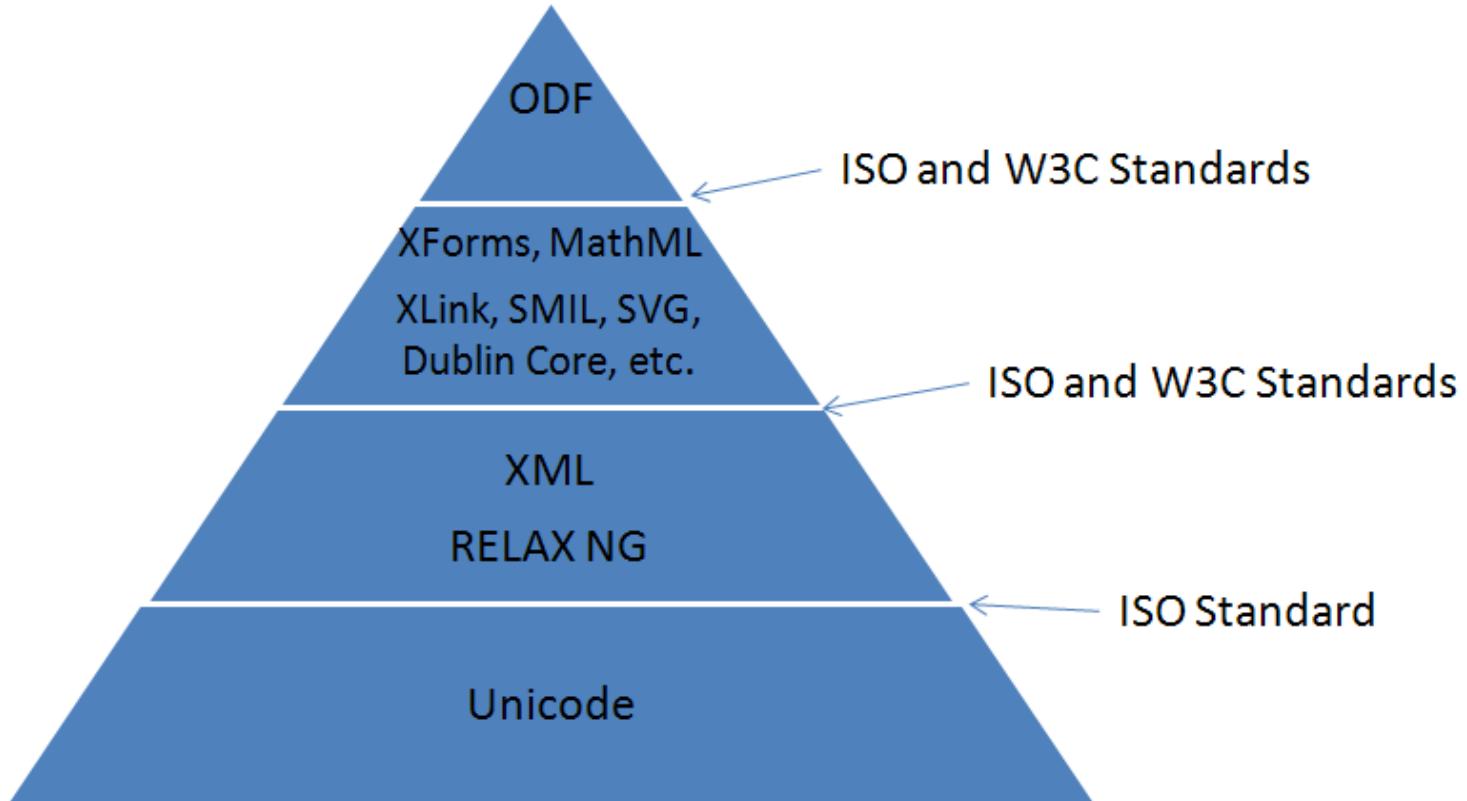
Proprietary File Stack



Pseudo-Standard File Stack



Standard File Stack



Document Complexity

- **OBJECTS** - Typography, bitmap and outline images, colour, business rules, text, steganography, and much more are used to make up documents, however the management and approval process of these objects are also key to ensure the correctness, and value of each individual document
- **DATA** - Structure, storage locations and techniques, extraction, backup, transmission methods, normalisation, consolidation, translation, manipulation and sorting, plus security, privacy, and data governance



How is a Standard File

- ZIP file (regardless of extension)
- Set of XML files describing the content of the file and the presentation (the platform displays what is described by the XML file)
- XML, a standard language, simplifies both the description and the access to the contents of the file
- Same set of XML files for all applications (text, spreadsheet, presentation, etc.)
- Binary files are used only for images and multimedia



LibreOffice as Shakespeare (ODT)

2017

<text:p text:style-name="P1">To be, or not to be,
that is the question</text:p>

2018

<text:p text:style-name="P1">To be, or not to be,
that is the question</text:p>

2019

<text:p text:style-name="P1">To be, or not to be,
that is the question</text:p>

MS Office as Shakespeare (DOCX)

2017

<w:t>To be</w:t>

<w:t>,</w:t>

<w:t xml:space="preserve"> or not to be</w:t>

<w:t>,</w:t>

<w:t xml:space="preserve"> that </w:t>

<w:t>is the question</w:t>

MS Office as Shakespeare (DOCX)

2018

<w:t>To be</w:t>

<w:t>, or</w:t>

<w:t xml:space="preserve"> not to be</w:t>

<w:t>,</w:t>

<w:t xml:space="preserve"> that is the</w:t>

<w:t xml:space="preserve"> question</w:t>



MS Office as Shakespeare (DOCX)

2019

<w:t>To be</w:t>

<w:t>,</w:t>

<w:t xml:space="preserve"> or not to be, that </w:t>

<w:t>is the question</w:t>



Brain & Computer

Brain
Red

Computer
#FF0000



Brain & Computer

ODF (LibreOffice)

- Writer
fo:color="#FF0000"
- Calc
fo:color="#FF0000"
- Impress
fo:color="#FF0000"

OOXML (MS Office)

- Word
w:color w:val="FF0000"
- Excel
color rgb="FFFF0000"
- PowerPoint
a:srgbClr val="FF0000"



A screenshot of Microsoft Excel showing a table of dates from February 1, 1900, to February 29, 1900. The table has columns A through F and rows 1 through 30. The header row contains the date 01/02/1900. The data starts at cell A1 with the value 01/02/1900 and continues sequentially down column A. The text "OOXML Microsoft Office" is overlaid in red on the left side of the table.

	A	B	C	D	E	F
1	01/02/1900					
2	02/02/1900					
3	03/02/1900					
4	04/02/1900					
5	05/02/1900					
6	06/02/1900					
7	07/02/1900					
8	08/02/1900					
9	09/02/1900					
10	10/02/1900					
11	11/02/1900					
12	12/02/1900					
13	13/02/1900					
14	14/02/1900					
15	15/02/1900					
16	16/02/1900					
17	17/02/1900					
18	18/02/1900					
19	19/02/1900					
20	20/02/1900					
21	21/02/1900					
22	22/02/1900					
23	23/02/1900					
24	24/02/1900					
25	25/02/1900					
26	26/02/1900					
27	27/02/1900					
28	28/02/1900					
29	29/02/1900					
30						



A screenshot of LibreOffice Calc showing a table of dates from February 1, 1900, to March 1, 1900. The table has columns A through D and rows 1 through 30. The header row contains the date 01/02/1900. The data starts at cell G10 with the value 01/02/1900 and continues sequentially down column G. The text "ODF LibreOffice" is overlaid in green on the right side of the table.

	A	B	C	D
1	01/02/1900			
2	02/02/1900			
3	03/02/1900			
4	04/02/1900			
5	05/02/1900			
6	06/02/1900			
7	07/02/1900			
8	08/02/1900			
9	09/02/1900			
10	10/02/1900			
11	11/02/1900			
12	12/02/1900			
13	13/02/1900			
14	14/02/1900			
15	15/02/1900			
16	16/02/1900			
17	17/02/1900			
18	18/02/1900			
19	19/02/1900			
20	20/02/1900			
21	21/02/1900			
22	22/02/1900			
23	23/02/1900			
24	24/02/1900			
25	25/02/1900			
26	26/02/1900			
27	27/02/1900			
28	28/02/1900			
29	29/02/1900			
30				

Writing Dates the Excel Way

Event	Calc	Excel
Italo Vignoli Birthday		19948
Italo Vignoli Graduation		28813
Italo Vignoli First Job		29860
Italo Vignoli First Computer		30560
Italo Vignoli Wedding		30933
Italo Vignoli Installs OOo		37623
LibreOffice Announcement		40449



Writing Dates the Excel Way

Event	Calc	Excel
Italo Vignoli Birthday	12/08/1954	19948
Italo Vignoli Graduation	19/11/1978	28813
Italo Vignoli First Job	01/10/1981	29860
Italo Vignoli First Computer	01/09/1983	30560
Italo Vignoli Wedding	08/09/1984	30933
Italo Vignoli Installs OOo	02/01/2003	37623
LibreOffice Announcement	28/09/2010	40449



Issues with Language Codes

- ISO 639 is the global standard to define short codes for language names (English = EN)
- OOXML uses its own proprietary list of numbers
- This creates interoperability issues with software ISO standard compliant



Issues with Graphics

- ISO/IEC 8632 is the global ISO standard for computer graphics metafile: 2D graphical information based on "vector graphics", "raster graphics" and "text"
- OOXML recommends Windows Metafiles or Enhanced Metafiles instead of using ISO/IEC 8632 or W3C SVG
- WMF is a proprietary Windows format which has never been approved by ISO

Conflict with SVG

- SVG is W3C "standard for describing two-dimensional vector graphics and mixed vector/raster graphic with XML"
- "DrawingML" defines the VML vector format, which conflicts with W3C SVG standard
- VML was proposed to W3C in 1998, but was never approved (so, the development was ceased in 1998)

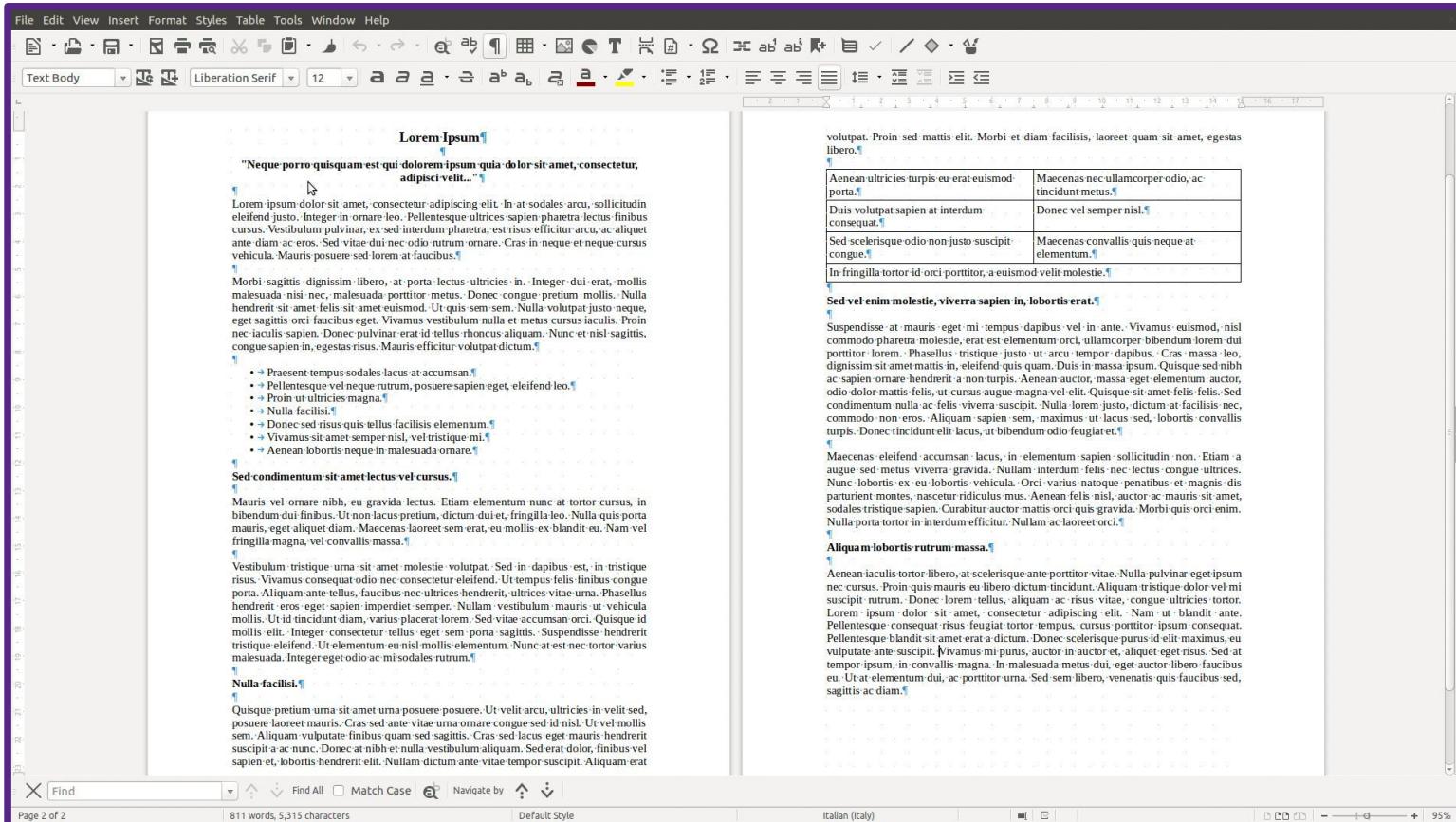


Conflict with MathML

- MathML is W3C standard for "describing mathematical notation and capturing both structure and contents"
- "Math" manages mathematical expressions, with a format in conflict and incompatible with MathML, recommended by the W3C
- MathML is part of the ISO / IEC 26300 (Open Document Format) standard, so OOXML intentionally creates a conflict with an existing standard



Comparison ODF/ooXML



Length in Lines of the XML File

Version	Lines of XML
ODF 1.2 (any version of) LibreOffice	222
OOXML 2010 Transitional (MS Office Windows)	1040
OOXML 2011 Transitional (MS Office MacOS)	12854
OOXML 2013 Transitional (MS Office Windows)	1590
OOXML 2016 Transitional (MS Office Windows)	11667
OOXML 2016 Transitional (MS Office MacOS)	11646
OOXML 2019 Transitional (MS Office Windows)	7085

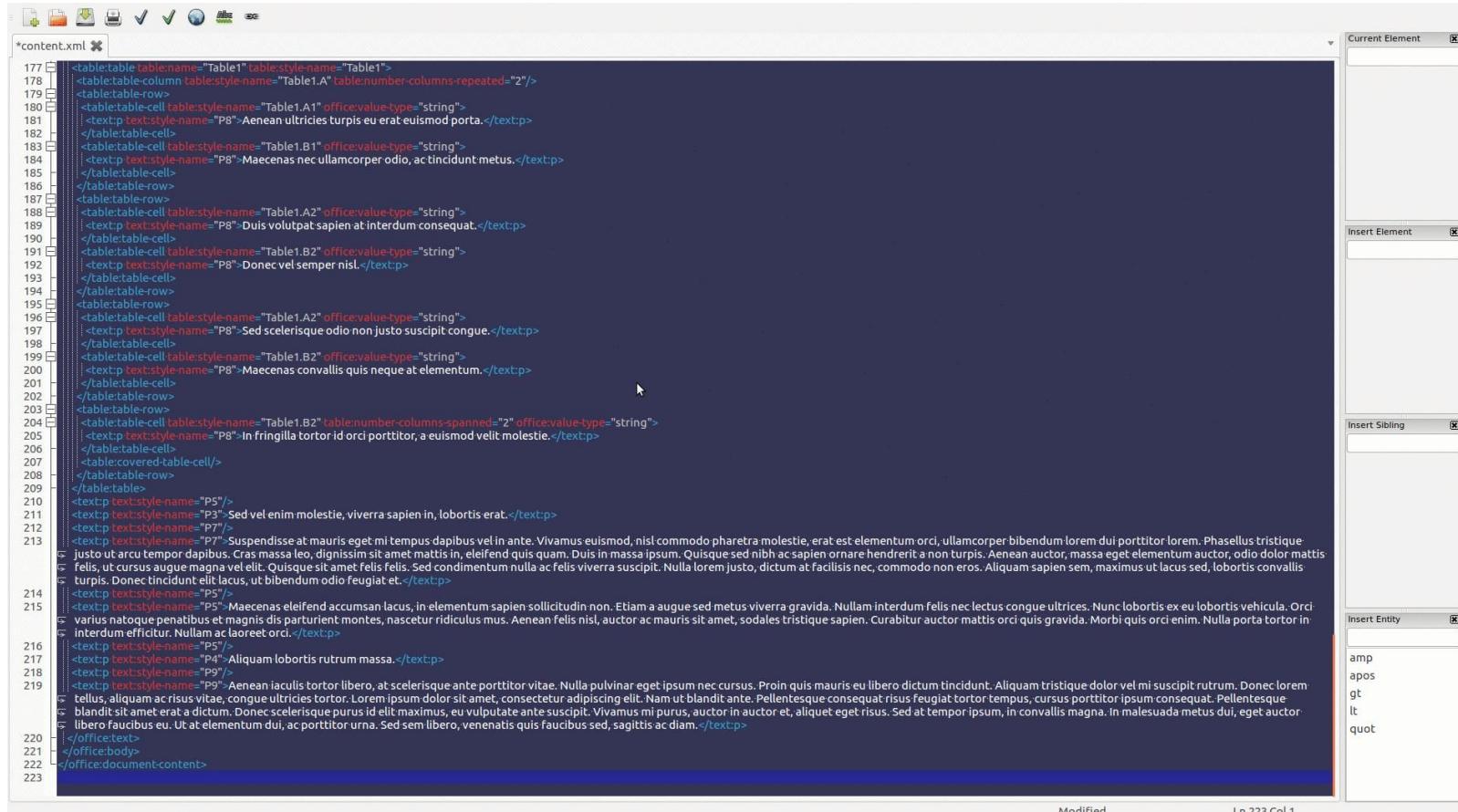


“Seasonality” of MS Office XML Files

Versione	Lines of XML
Windows OOXML 2013 Transitional Summer 2017	1590
Windows OOXML 2013 Transitional Winter 2018	13515
Windows OOXML 2016 Transitional Summer 2017	11667
Windows OOXML 2016 Transitional Winter 2018	969
Windows OOXML 2016 Transitional Fall 2018	11288
Windows OOXML 2016 Transitional Spring 2019	7085
MacOS OOXML 2016 Transitional Summer 2017	11646
MacOS OOXML 2016 Transitional Fall 2018	854
MacOS OOXML 2016 Transitional Spring 2019	7731



XML of ODT (LibreOffice)



The screenshot shows the LibreOffice XML Editor interface. The main window displays the XML code for a document named "content.xml". The code includes several tables and their rows, with specific cells containing text styled as "P8". The XML editor features a toolbar at the top, a status bar at the bottom, and several side panes: "Current Element", "Insert Element", "Insert Sibling", and "Insert Entity". The "Insert Entity" pane lists entities such as amp, apos, gt, lt, and quot. The status bar at the bottom right indicates "Modified" and "Ln 223 Col 1".

```
<?xml version="1.0" encoding="UTF-8"?>
<office:document-content>
  <office:body>
    <table:table table:name="Table1" table:style-name="Table1">
      <table:table-column table:style-name="Table1.A" table:number-columns-repeated="2"/>
      <table:table-row>
        <table:table-cell table:style-name="Table1.A1" office:value-type="string">
          <text:p text:style-name="P8">Aenean ultricies turpis eu erat euismod porta.</text:p>
        </table:table-cell>
        <table:table-cell table:style-name="Table1.B1" office:value-type="string">
          <text:p text:style-name="P8">Maecenas nec ullamcorper odio, ac tincidunt metus.</text:p>
        </table:table-cell>
      </table:table-row>
      <table:table-row>
        <table:table-cell table:style-name="Table1.A2" office:value-type="string">
          <text:p text:style-name="P8">Duis volutpat sapien at interdum consequat.</text:p>
        </table:table-cell>
        <table:table-cell table:style-name="Table1.B2" office:value-type="string">
          <text:p text:style-name="P8">Donec vel semper nisl.</text:p>
        </table:table-cell>
      </table:table-row>
      <table:table-row>
        <table:table-cell table:style-name="Table1.A2" office:value-type="string">
          <text:p text:style-name="P8">Sed scelerisque odio non justo suscipit congue.</text:p>
        </table:table-cell>
        <table:table-cell table:style-name="Table1.B2" office:value-type="string">
          <text:p text:style-name="P8">Maecenas convallis quis neque at elementum.</text:p>
        </table:table-cell>
      </table:table-row>
      <table:table-row>
        <table:table-cell table:style-name="Table1.B2" table:number-columns-spanned="2" office:value-type="string">
          <text:p text:style-name="P8">In fringilla tortor id orci porttitor, a euismod velit molestie.</text:p>
        </table:table-cell>
        <table:covered-table-cell>
        </table:covered-table-cell>
      </table:table-row>
      <table:table-row>
        <text:p text:style-name="P5">
          <text:p text:style-name="P3">Sed vel enim molestie, viverra sapien in, lobortis erat.</text:p>
          <text:p text:style-name="P7">
            <text:p text:style-name="P7">Suspendisse at mauris eget mi tempus dapibus vel in ante. Vivamus euismod, nisl commodo pharetra molestie, erat est elementum orci, ullamcorper bibendum lorem dui porttitor lorem. Phasellus tristique justo ut arcu tempor dapibus. Cras massa leo, dignissim sit amet mattis in, eleifend quis quam. Duis in massa ipsum. Quisque sed nibh ac sapien ornare hendrerit a non turpis. Aenean auctor, massa eget elementum auctor, odio dolor mattis felis, ut cursus augue magna vel elit. Quisque sit amet felis felis. Sed condimentum nulla ac felis viverra suscipit. Nulla lorem justo, dictum at facilisis nec, commodo non eros. Aliquam sapien sem, maximus ut lacus sed, lobortis convallis turpis. Donec tincidunt elit lacus, ut bibendum odio feugiat et.</text:p>
        </text:p>
        <text:p text:style-name="P5">
          <text:p text:style-name="P5">Maecenas eleifend accumsan lacus, in elementum sapien sollicitudin non. Etiam a augue sed metus viverra gravida. Nullam interdum felis nec lectus congue ultrices. Nunc lobortis ex eu lobortis vehicula. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aenean felis nisl, auctor ac mauris sit amet, sodales tristique sapien. Curabitur auctor mattis orci quis gravida. Morbi quis orci enim. Nulla porta tortor in interdum efficitur. Nullam ac laoreet orci.</text:p>
        </text:p>
        <text:p text:style-name="P5">
          <text:p text:style-name="P4">Aliquam lobortis rutrum massa.</text:p>
          <text:p text:style-name="P9">
            <text:p text:style-name="P9">Aenean facilius tortor libero, at scelerisque ante porttitor vitae. Nulla pulvinar eget ipsum nec cursus. Proin quis mauris eu libero dictum tincidunt. Aliquam tristique dolor vel mi suscipit rutrum. Donec lorem tellus, aliquam ac risus vitae, congue ultrices tortor. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nam ut blandit ante. Pellentesque consequat risus feugiat tortor tempus, cursus porttitor ipsum consequat. Pellentesque blandit sit amet erat a dictum. Donec scelerisque purus id elit maximus, eu vulputate ante suscipit. Vivamus mi purus, auctor in auctor et, aliquet eget risus. Sed at tempor ipsum, in convallis magna. In malesuada metus dui, eget auctor libero faucibus eu. Ut at elementum dui, ac porttitor urna. Sed sem libero, venenatis quis faucibus sed, sagittis ac diam.</text:p>
        </text:p>
      </table:table-cell>
    </table:body>
  </office:document-content>
```



XML of DOCX (MS Office 2013)

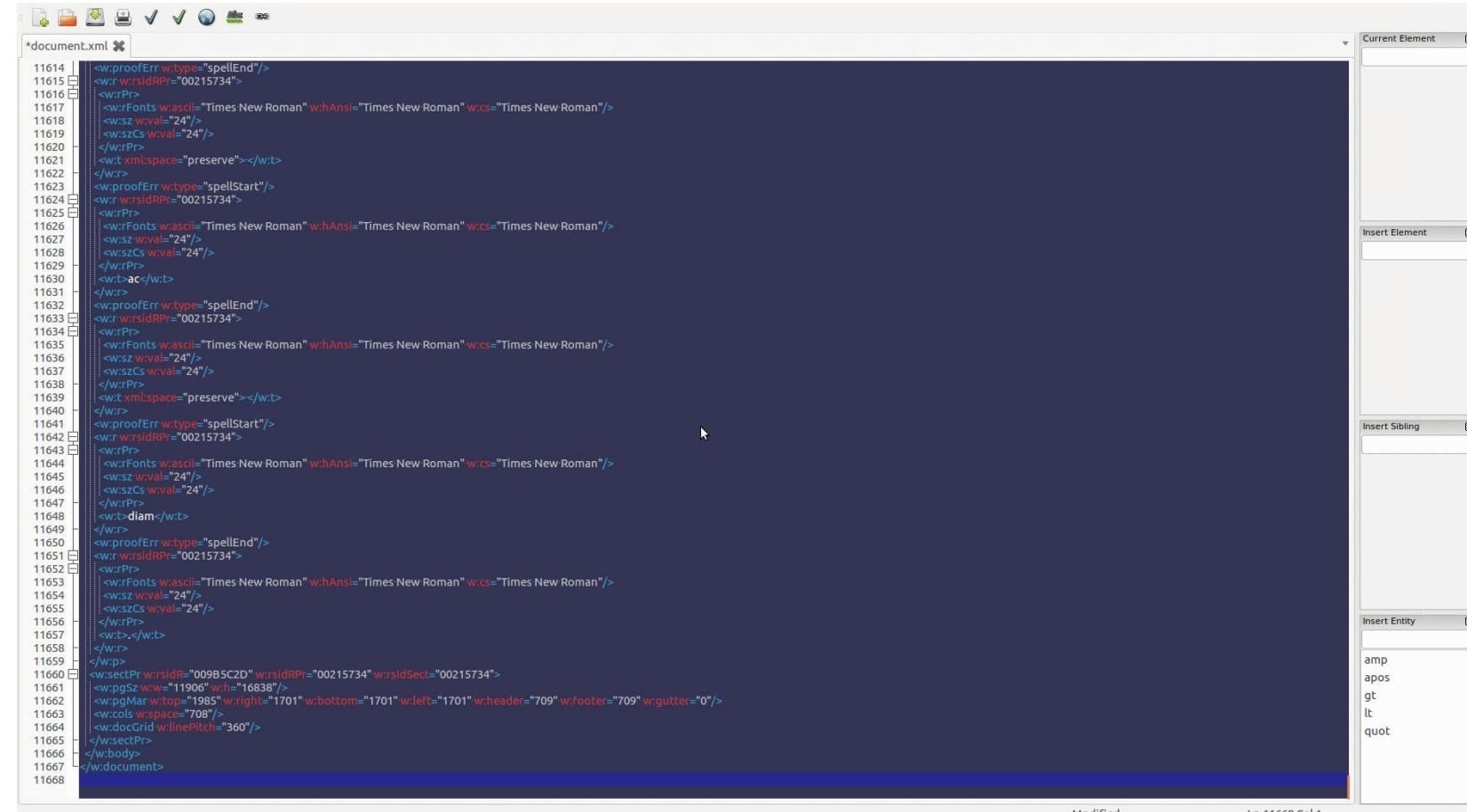


The screenshot shows a XML editor window titled "document.xml". The left pane displays the XML code with line numbers from 1537 to 1591. The right pane contains four toolbars: "Current Element", "Insert Element", "Insert Sibling", and "Insert Entity". The "Insert Entity" toolbar includes buttons for "amp", "apos", "gt", "lt", and "quot". The XML code includes various paragraph and run elements with attributes like w:font, w:sz, w:szCs, w:lang, and w:pr. It also contains sections for proofing errors and a footer section with a page number of 1.

```
<w:rFonts w:ascii="Times New Roman" w:hAnsi="Times New Roman" w:cs="Times New Roman"/>
<w:sz w:val="24"/>
<w:szCs w:val="24"/>
<w:lang w:val="en-US"/>
</w:rPr>
<w:t>purus</w:t>
</w:r>
<w:proofErr w:type="gramEnd"/>
<w:rPr>
<w:rFonts w:ascii="Times New Roman" w:hAnsi="Times New Roman" w:cs="Times New Roman"/>
<w:sz w:val="24"/>
<w:szCs w:val="24"/>
<w:lang w:val="en-US"/>
</w:rPr>
<w:t xml:space="preserve">id elit maximus, eu vulputate ante suscipit. Vivamus mi-</w:t>
</w:r>
<w:proofErr w:type="gramStart"/>
<w:rPr w:rsidRPr="00215734">
<w:rPr>
<w:rFonts w:ascii="Times New Roman" w:hAnsi="Times New Roman" w:cs="Times New Roman"/>
<w:sz w:val="24"/>
<w:szCs w:val="24"/>
<w:lang w:val="en-US"/>
</w:rPr>
<w:t>purus</w:t>
</w:r>
<w:proofErr w:type="gramEnd"/>
<w:rPr w:rsidRPr="00215734">
<w:rPr>
<w:rFonts w:ascii="Times New Roman" w:hAnsi="Times New Roman" w:cs="Times New Roman"/>
<w:sz w:val="24"/>
<w:szCs w:val="24"/>
<w:lang w:val="en-US"/>
</w:rPr>
<w:t xml:space="preserve">auctor in auctor et, aliquet eget risus. Sed at tempor ipsum, in convallis magna. In malesuada metus dui, eget auctor libero faucibus eu. Ut at elementum dui, ac porttitor urna.</w:t>
</w:r>
<w:rPr w:rsidRPr="00215734">
<w:rPr>
<w:rFonts w:ascii="Times New Roman" w:hAnsi="Times New Roman" w:cs="Times New Roman"/>
<w:sz w:val="24"/>
<w:szCs w:val="24"/>
<w:lang w:val="en-US"/>
</w:rPr>
<w:t>Sed sem libero, venenatis quis faucibus sed, sagittis ac diam.</w:t>
</w:r>
</w:p>
<w:sectPr w:rsidR="009B5C2D" w:rsidRPr="00215734" w:rsidSect="00215734">
<w:pgSz w:w="11906" w:h="16836"/>
<w:pgMar w:top="1985" w:right="1701" w:bottom="1701" w:left="1701" w:header="709" w:footer="709" w:gutter="0"/>
<w:cols w:space="708"/>
<w:docGrid w:linePitch="360"/>
</w:sectPr>
</w:body>
</w:document>
```

Modified Ln 1 Col 1

XML of DOCX (MS Office 2016)



“Naive” Deductions

- All LibreOffice developers are genius
- All Microsoft Office are just i***s



“Real” Deductions

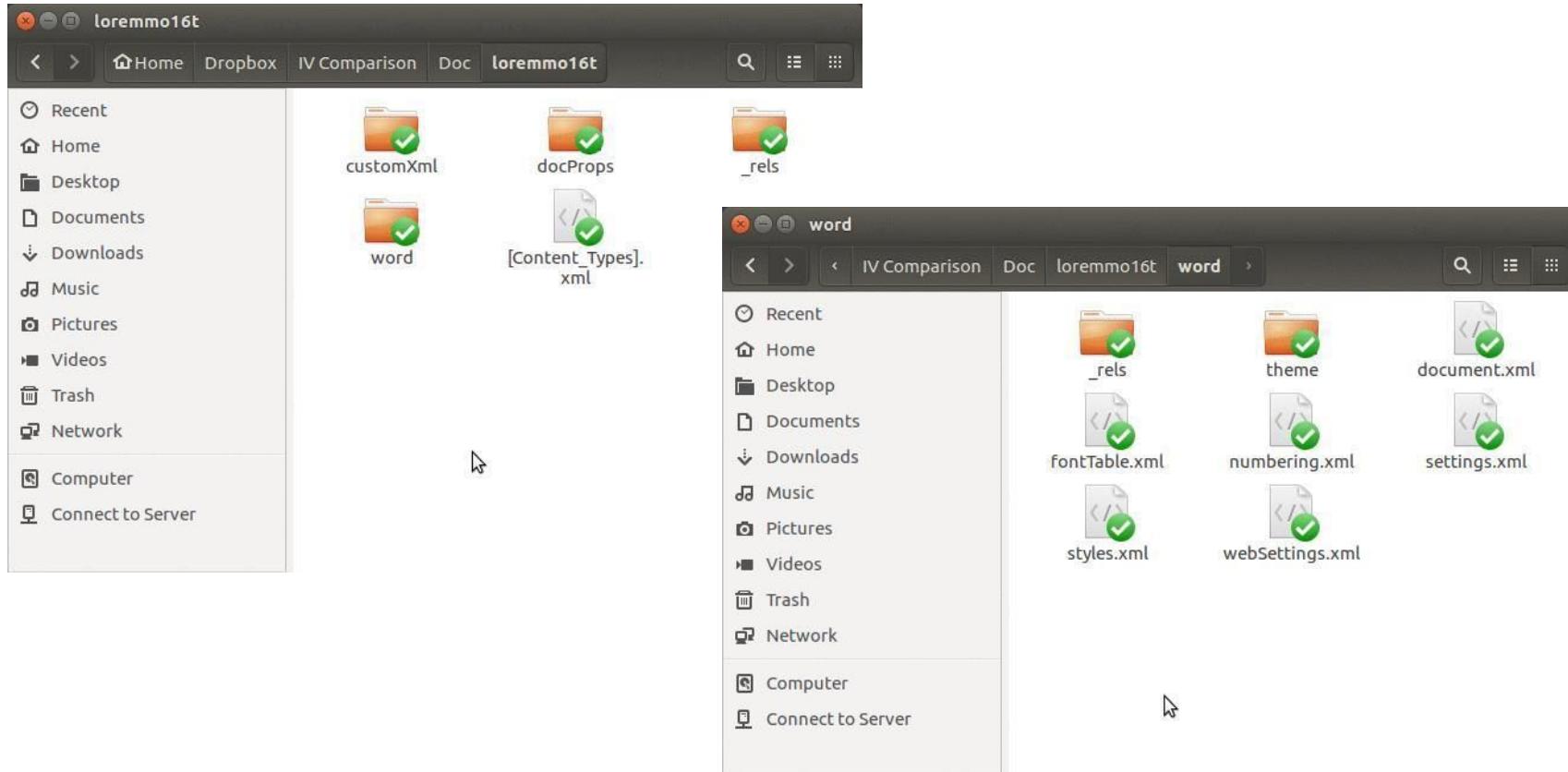
- Microsoft Office XML files are artificially filled with unnecessary content to reduce the chances that software other than Microsoft Office can open them correctly
- Microsoft has a clear commercial interest in opposing interoperability based on standard and open formats, to protect a market that is still worth more than \$25 billion
- So, documents created with Microsoft Office are standard on paper, but in reality they are built to fool users (and convince them that interoperability cannot exist)



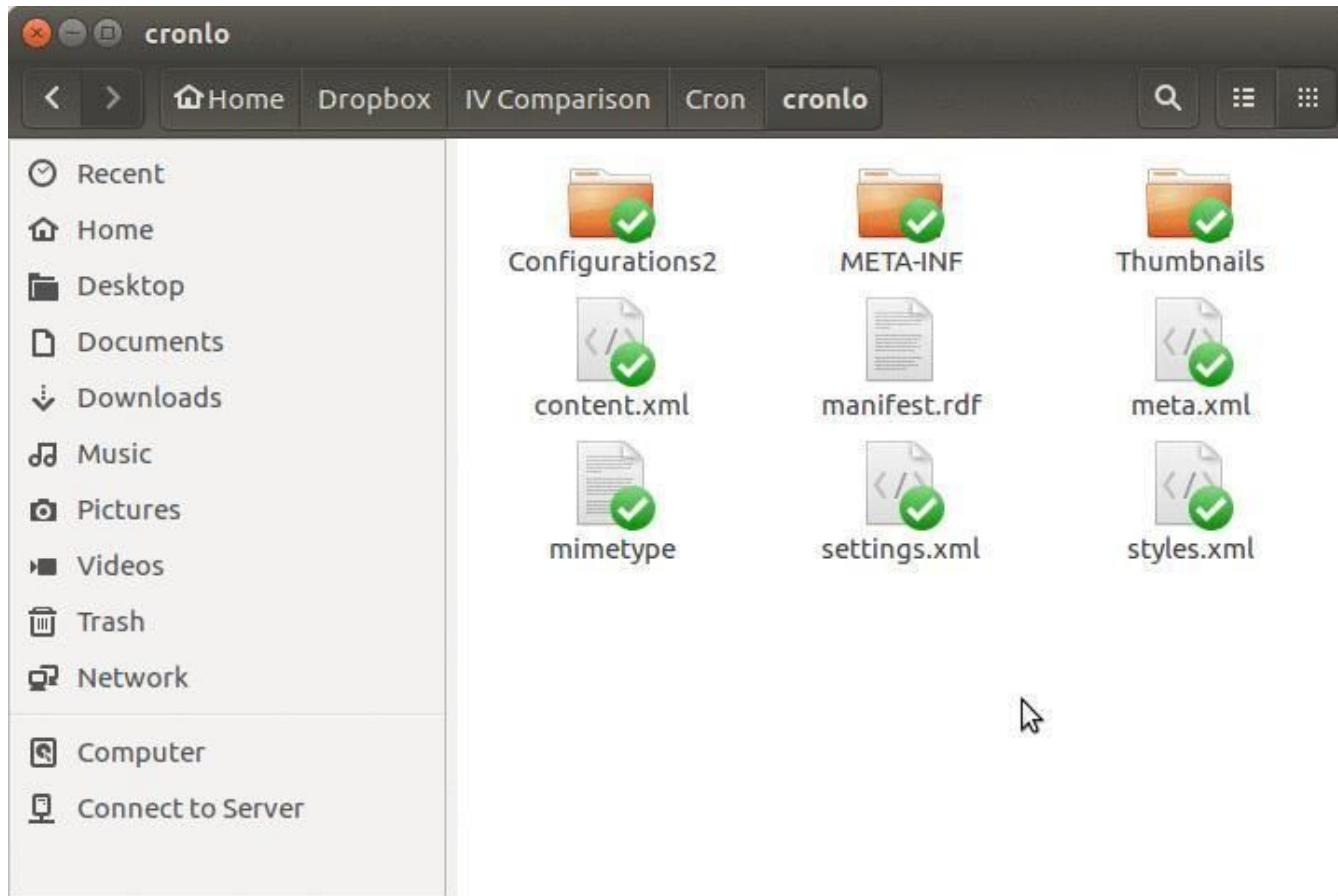
Inside ODT



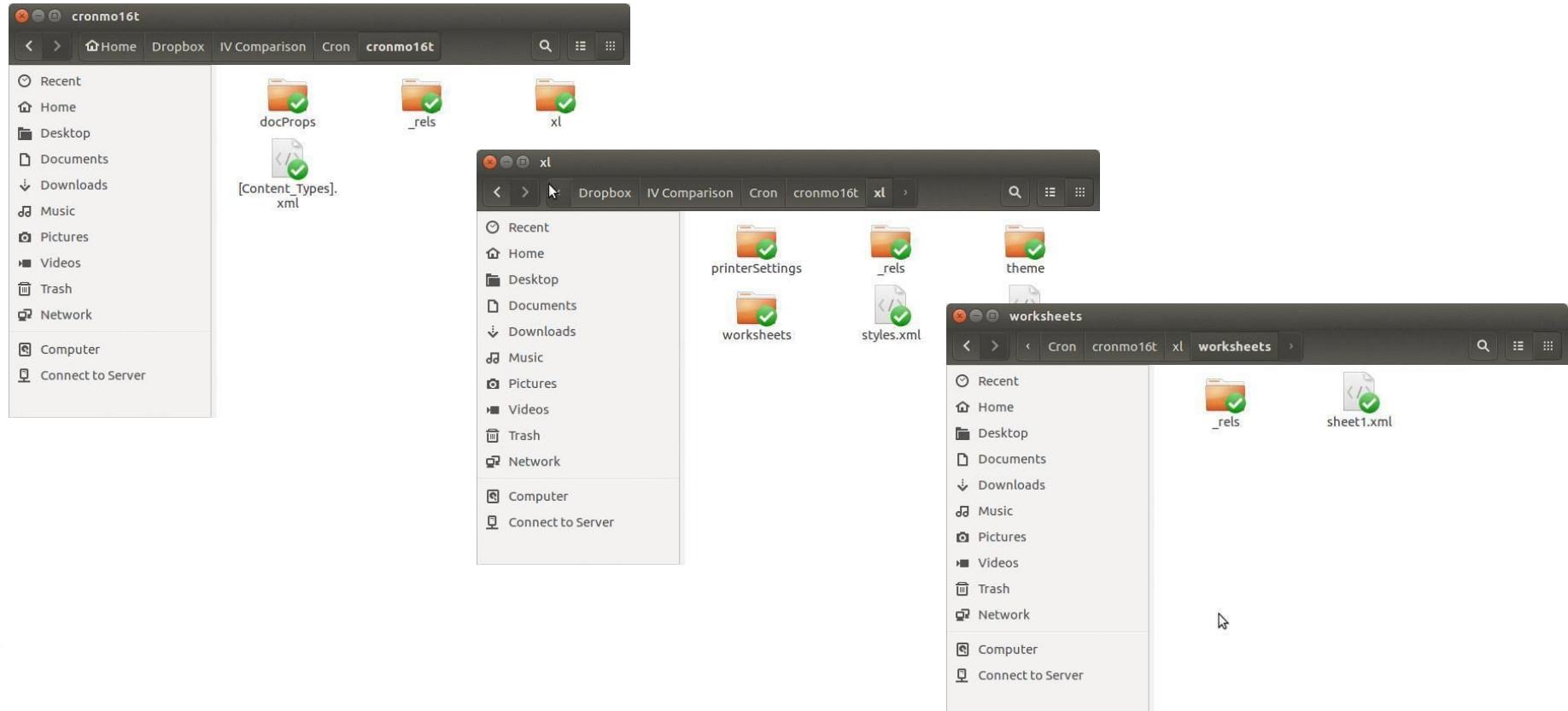
Inside DOCX



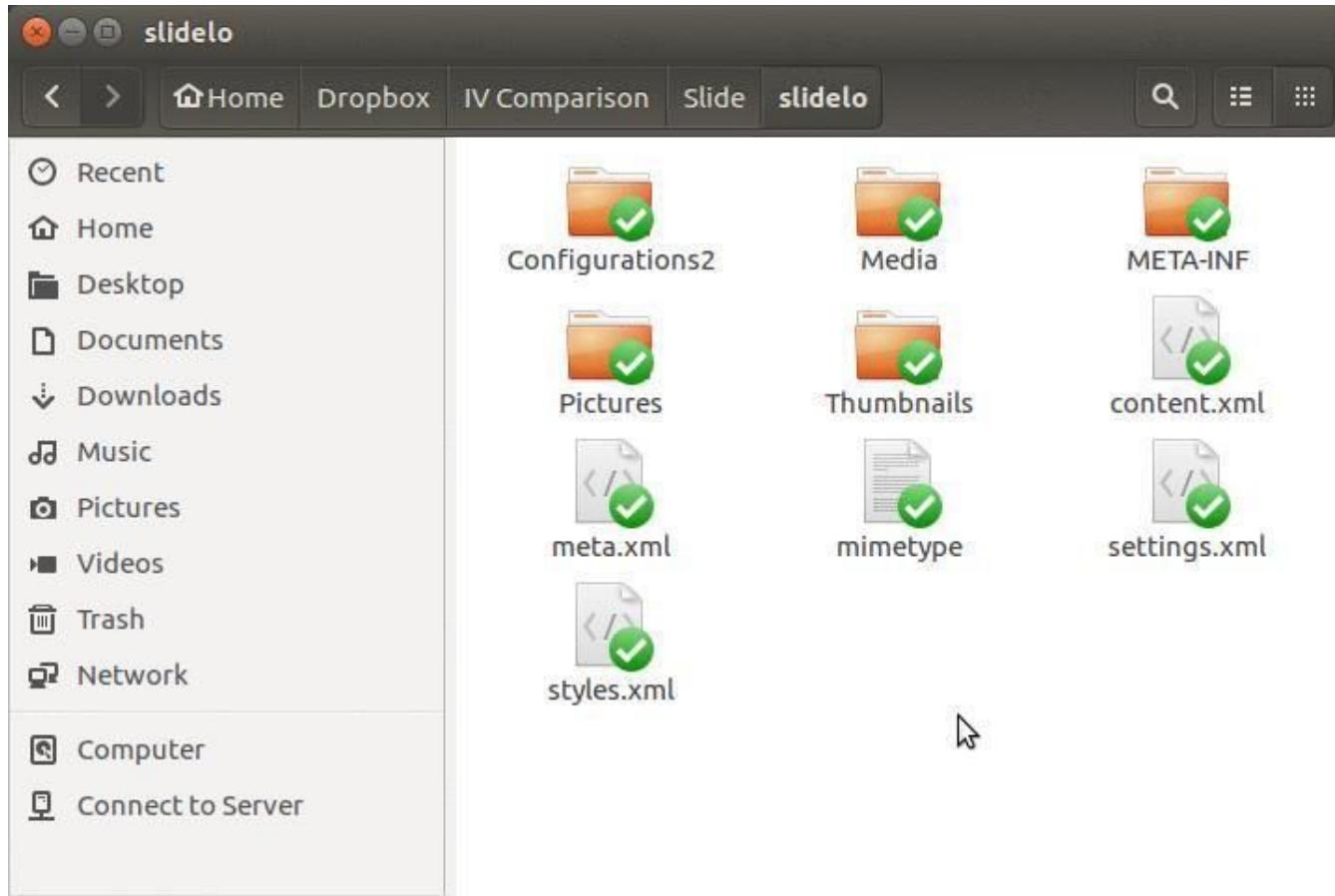
Inside ODS



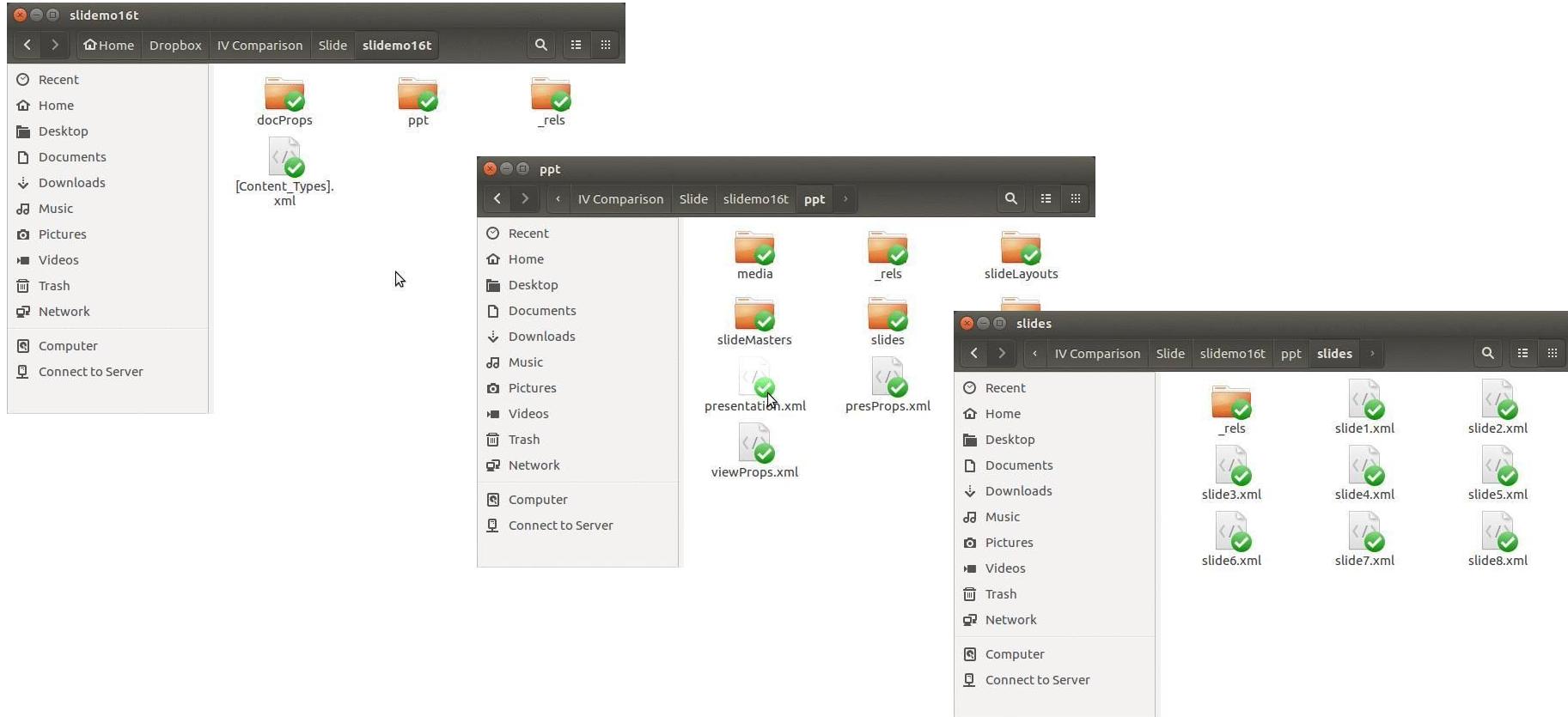
Inside XLSX



Inside ODP



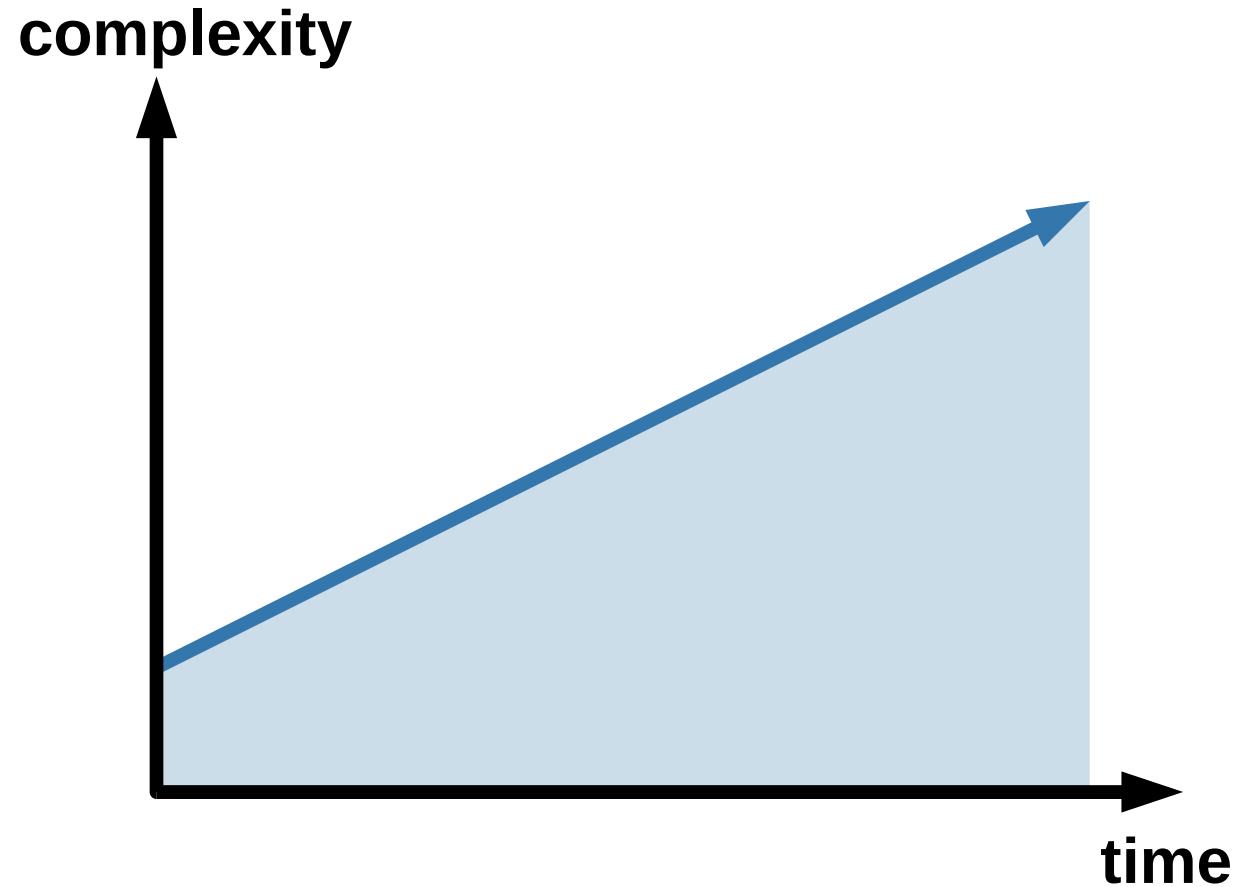
Inside PPTX



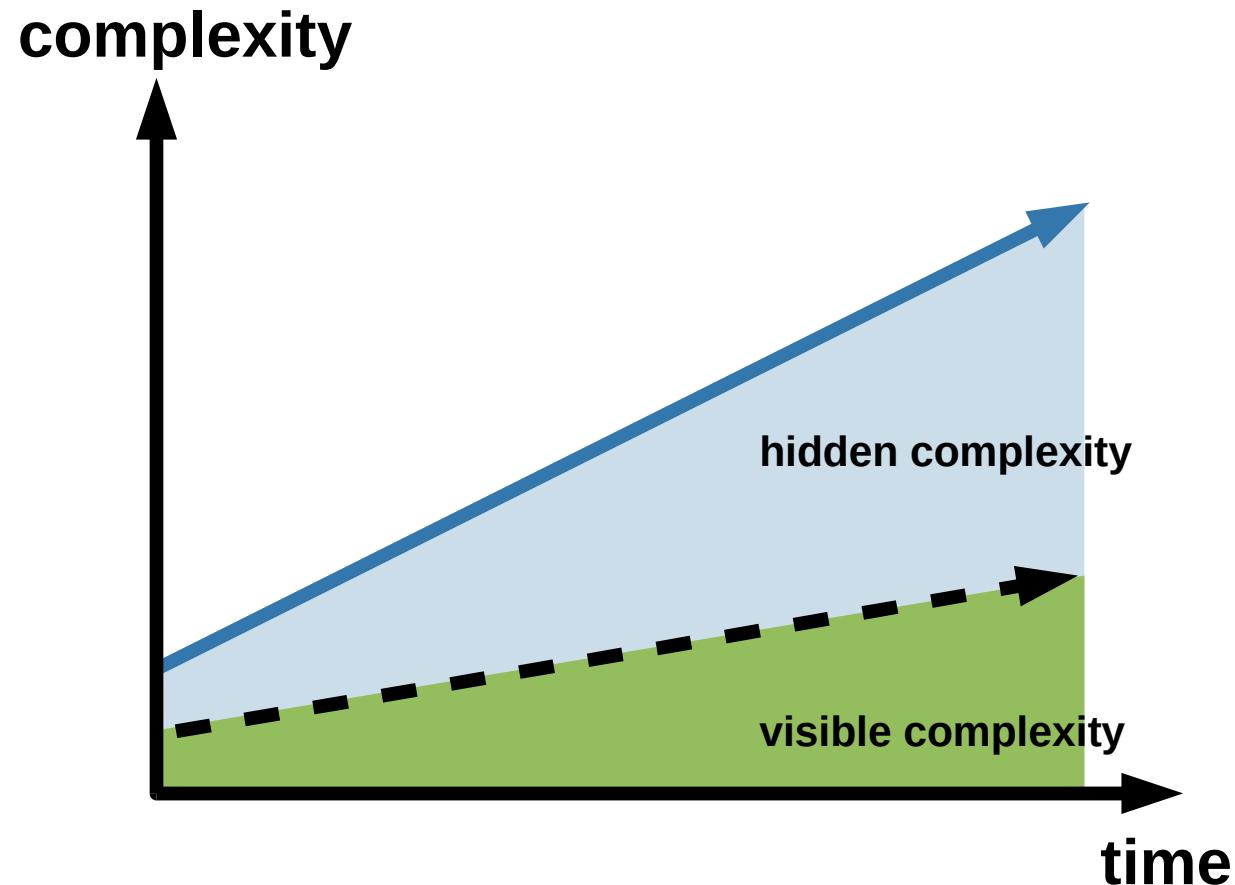
Simplicity vs Hidden Complexity

- **ODT / LibreOffice**
 - Reduced, very low or non existing complexity
 - XML files are human readable (as they should be)
- **OOXML / Microsoft Office**
 - Highest possible complexity vs technology
 - XML files are not human readable (contrary to what the XML standard language mandates)

Complexity of IT Systems



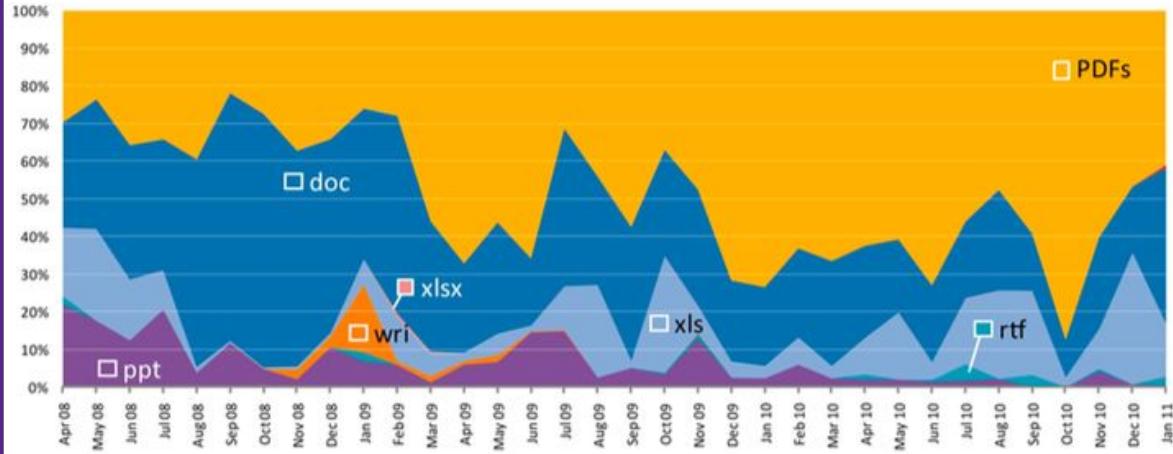
Hidden Complexity



Document Vulnerabilities in 2011



Document File Types Used in Targeted Attacks

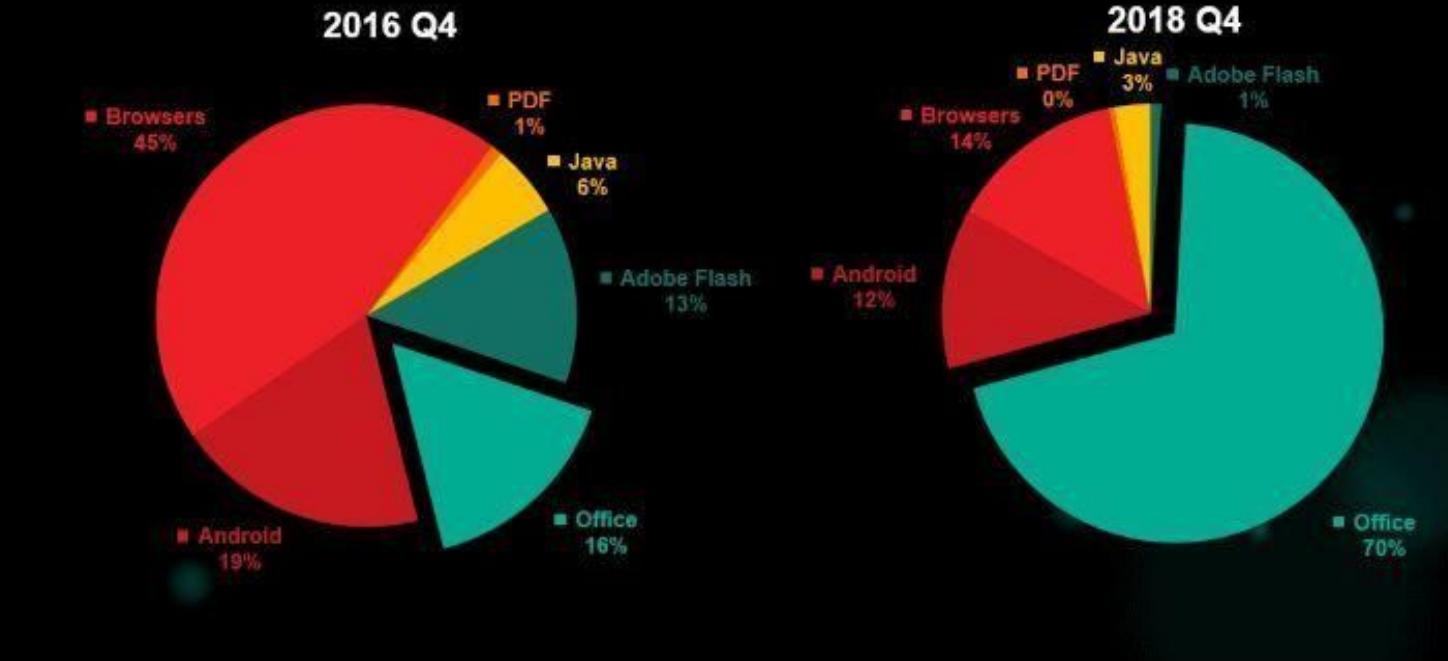


Source: Symantec MessageLabs Intelligence, February 2011 Intelligence Report



Document Vulnerabilities in 2018

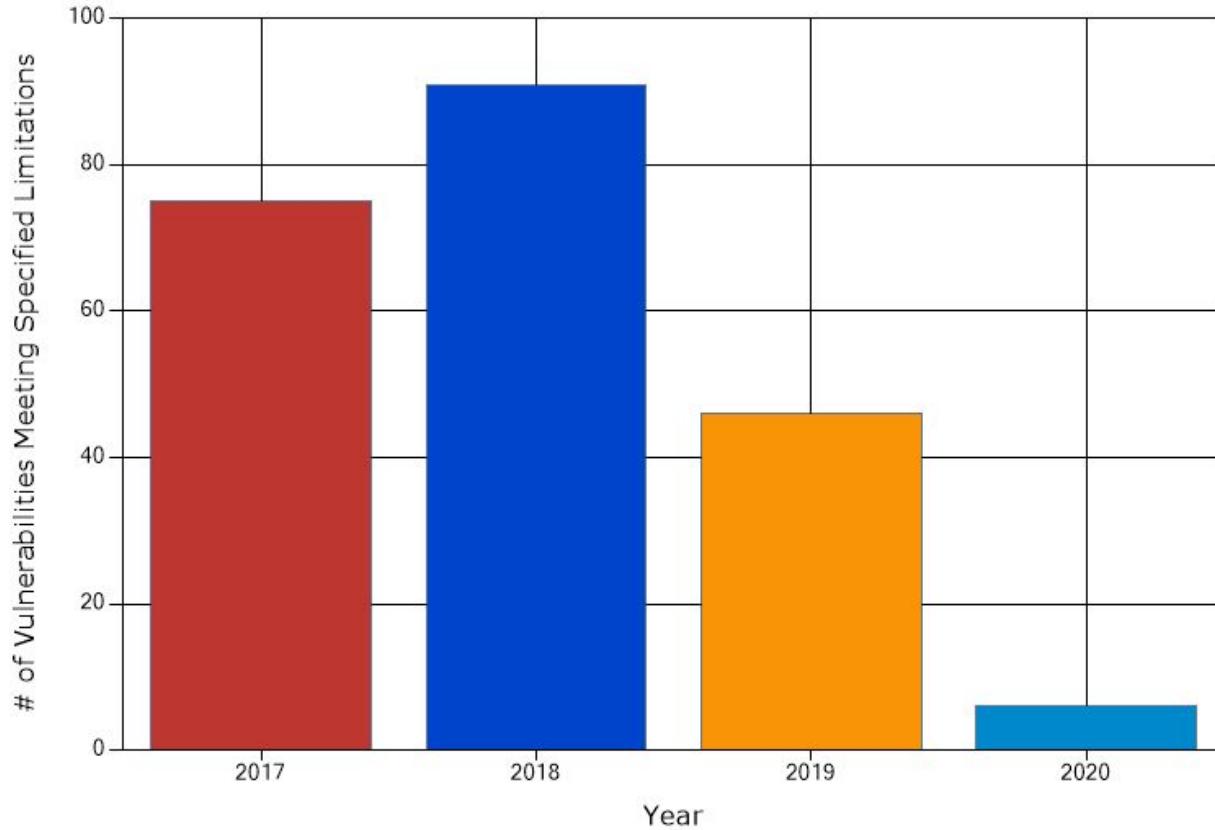
Targeted platforms by attacked users



Search Parameters:

- Results Type: Statistics
- Keyword (text search): microsoft office
- Search Type: Search Last 3 Years

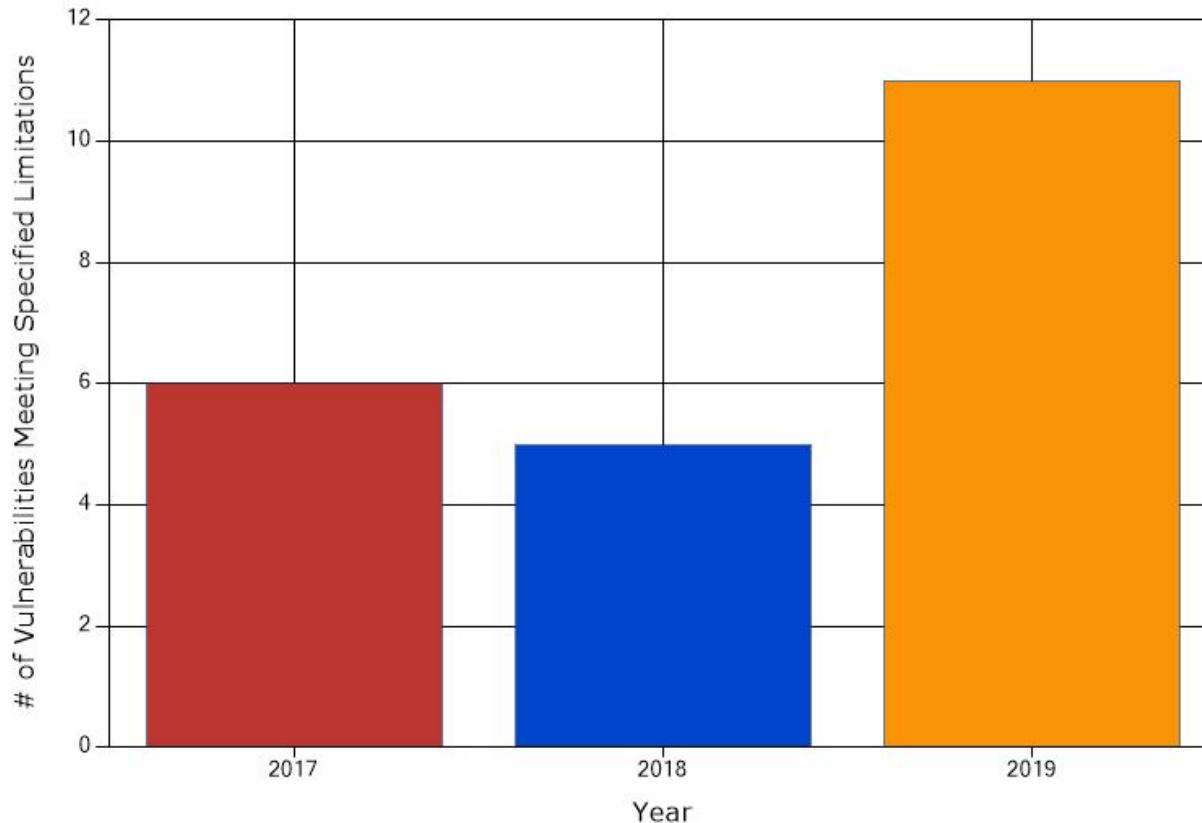
Total Matches By Year



Search Parameters:

- Results Type: Statistics
- Keyword (text search): libreoffice
- Search Type: Search Last 3 Years

Total Matches By Year





PSF

How to Lock-in Your Clients

How Professional Services Firms Can Create Compelling
Value for Clients Using Collaborative Technologies

Ross Dawson
CEO, Advanced Human Technologies
Author, *Living Networks and Developing Knowledge-Based Client Relationships*

January 2004

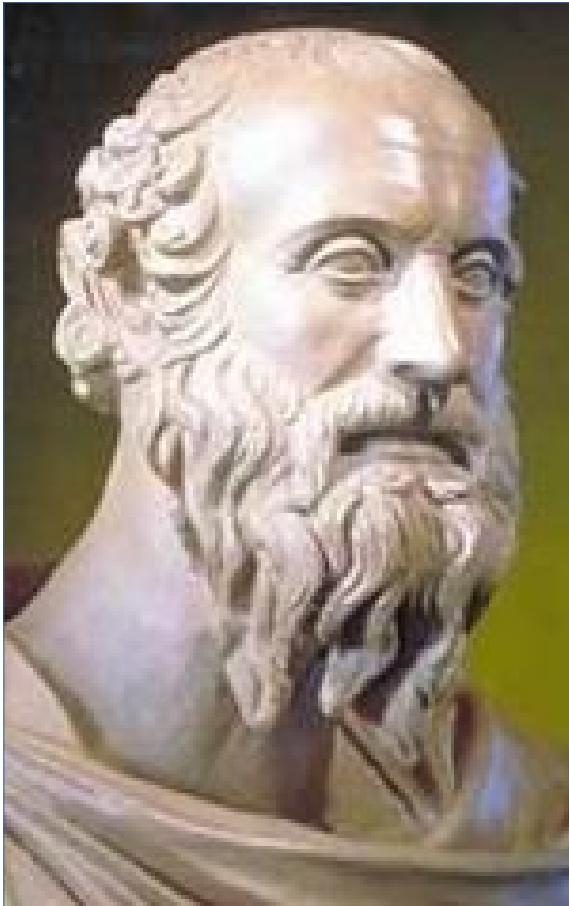
>> A STRATEGIC WHITE PAPER FROM MICROSOFT BUSINESS SOLUTIONS





Man is an
animal with
two legs and
no feathers

Plato



Diogenes the Cynic



Plato's Man

Thank You !

Italo Vignoli
Chair, ODF Advocacy
italo@libreoffice.org
+39 348 5653829

