



**LibreOffice**  
The Document Foundation



# **OpenDocument format: from COSM to the ODF Advocacy Open Project**

**Marina Latini**

[marina.latini@libreoffice.org](mailto:marina.latini@libreoffice.org)

**Italo Vignoli**

[italo@libreoffice.org](mailto:italo@libreoffice.org)

**ALMERÍA | 11 Sept. 2018**

A 3D perspective view of a red rectangular frame on a red background. The frame is composed of two concentric rectangles, creating a sense of depth. The text 'COSM' is centered within the inner rectangle.

**COSM**



# COSM (1)

- **Community of ODF Specification Maintainers**
- Project started by TDF in early 2018, based on the need of evolving the ODF standard beyond version ODF 1.2 to reflect the work of Technical Committee members
- Independent from LibreOffice
- Project housed at Public Software CIC (a UK Community Interest Company)



# COSM (2)

- First crowdfunded open standards project: an effective way to have essential standards maintained without needing a deep-pocketed international corporation behind them
- Seeded by TDF in early 2018, and then funded by Microsoft, Collabora, CIB, the UK Government & the European Commission
- Paying for an editor to draft the ODF 1.3 specifications using the work of the great Technical Committee contributors



# ODF 1.3

- ODF 1.3 Committee Draft approved September 9, 2019, and submitted to OASIS as a standard
- We do not expect significant reactions at this point
  - We do expect some random typo catching
- ODF 1.3 Committee Draft should be ratified by OASIS by the end of 2019
- ODF 1.3 ISO Standard expected for the end of 2020



**ODF Advocacy  
Open Project @ OASIS**



# 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 OpenDocument Format (ODF) Advocacy Open Project at OASIS is to create awareness and educate the world about the benefits of using ODF
- The group will develop advocacy materials for a communication campaign about the importance of standard-based document interoperability versus proprietary formats





# Purpose and Scope

- Educate users of personal productivity applications such as office suites and the likes, 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
- **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, as they will learn about how true document interoperability can protect their data into the future and ensure freedom-of-choice in software
- Large organizations deploying personal productivity software will discover the savings made possible by ODF and the associated advantage of increased security thanks to the reduced number of vulnerabilities affecting files based on standard file formats



# Basic Concept



- 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**



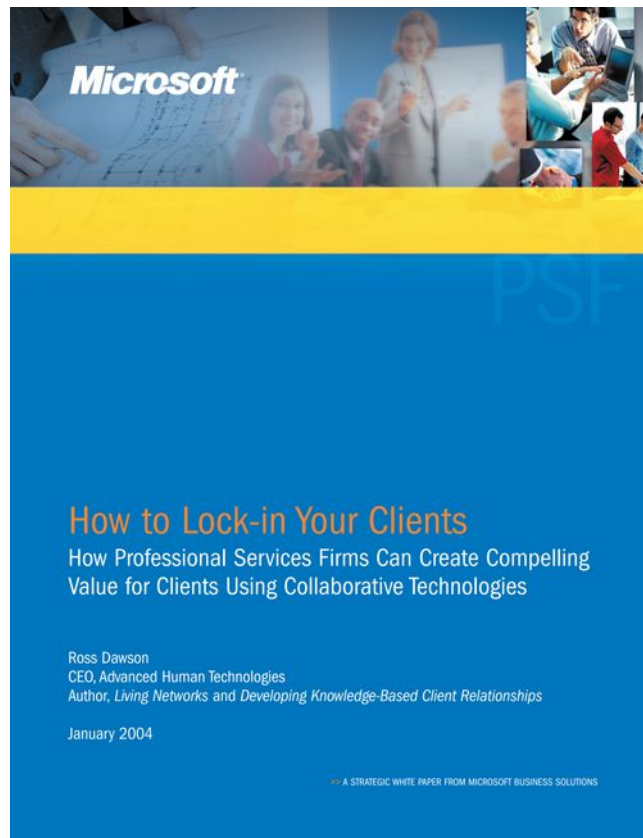
**Food for Thought**  
**ODF Advocacy**



# Lock In

# WE CANNOT READ YOUR DOCUMENTS

DOCUMENTFREEDOM.ORG





**Open Document Format**  
**the true document standard**  
**which offers freedom of choice**



# ODF is Standard

- France
- Portugal
- Sweden
- Taiwan
- UK





# Open 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 quality/price ratio
- Perennial: protects user developed contents from the “evolution” based obsolescence of technology



# Proprietary 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)



# ODF is an Open Standard

- Open and collaborative development
- Transparent access to the minutes of all meetings
- Ease of implementation in any software
- Freedom from patents and licensing restrictions
- No reliance on proprietary features or single vendor
- owned technologies
- Interoperability with any system designed to be ODF compatible



# 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)

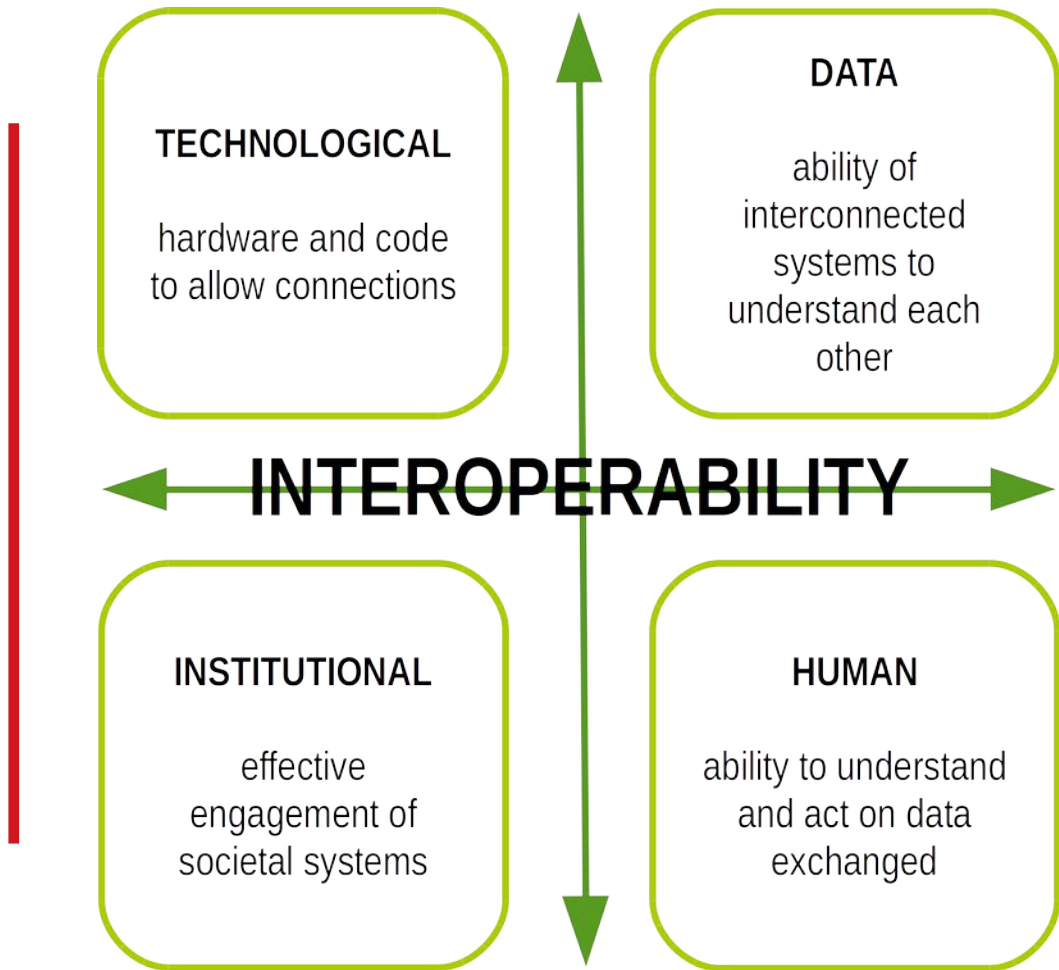
- Defined with an open and transparent process by OASIS
- Approved by the Joint Technical Committee 1 (JTC 1) of the IEC as International Standard (IS) in May 2006
- Available for deployment and use with no license, royalty payments or other restrictions

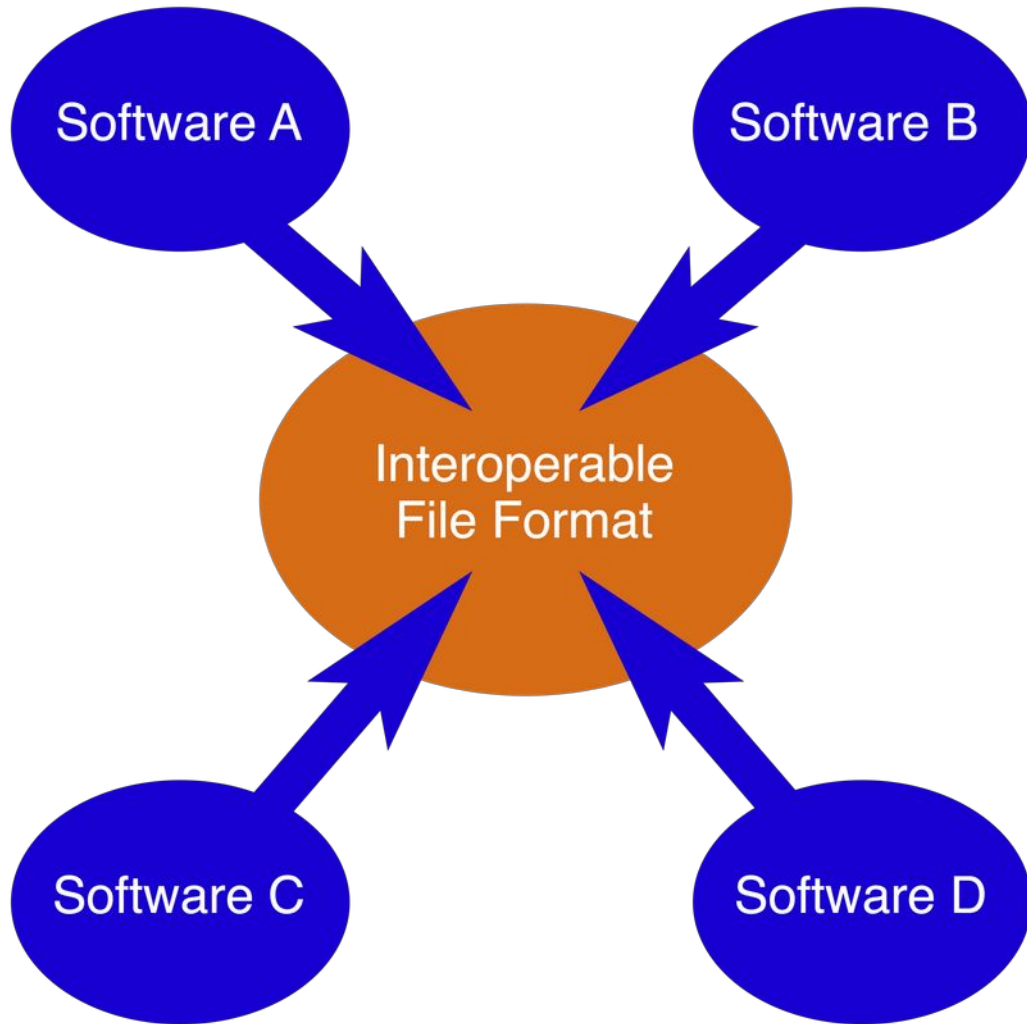


**Interoperability**



# Elements of Interoperability

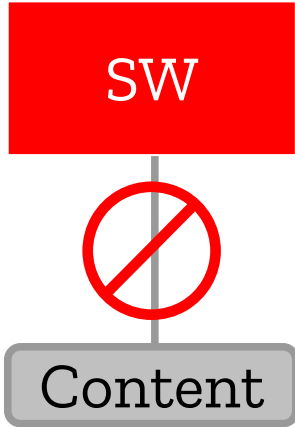








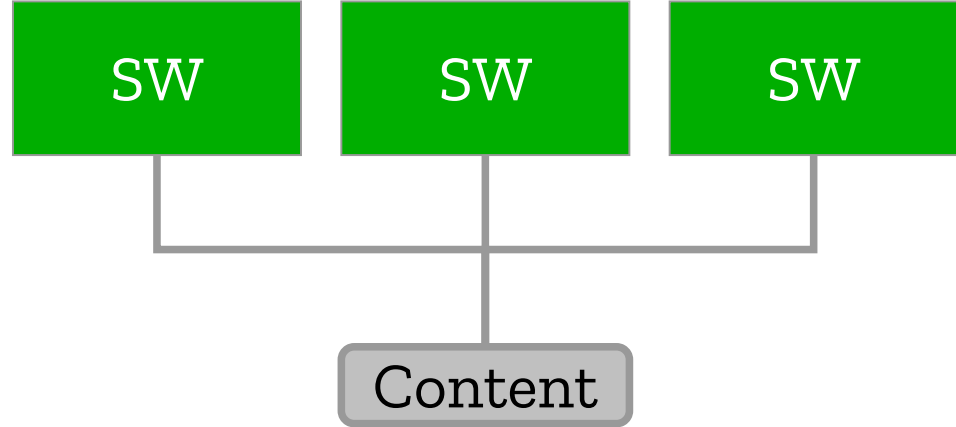
# ODF Based Interoperability



## OLD STYLE

**Content closely related to a specific application**

**Developers and not users control the application**



## NEW STYLE

**Content represented through an open standard which is not controlled by a single vendor, and supported by many applications**

**Users are in full control of their contents**



# Interoperability Costs (1)

NIST GCR 04-867

**NIST**

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

---

---



# Interoperability Costs (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



# Interoperability Costs (3)

<b>Construction</b>	<b>Avoidance Costs</b>	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
	<b>Mitigation Costs</b>	Interoperability research and development expenditures	0.003	0.04	3,924
		Manual reentry costs	0.024	0.26	27,750
		Design and construction information verification costs	0.006	0.07	7,377
		RFI management costs	0.05	0.53	55,656
		<b>Subtotal</b>	Avoidance costs	0.05	0.49
	Mitigation costs	0.08	0.86	90,783	
	<b>Subtotal</b>	<b>0.13</b>	<b>1.35</b>	<b>146,952</b>	
<b>Operations and Maintenance</b>	<b>Mitigation Costs</b>	Post-construction redundant information transfer costs	0.01	0.15	15,660
<b>Total Cost</b>					<b>1,169,762</b>

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



# **ODF vs OOXML**



# Battle of 2 Standards



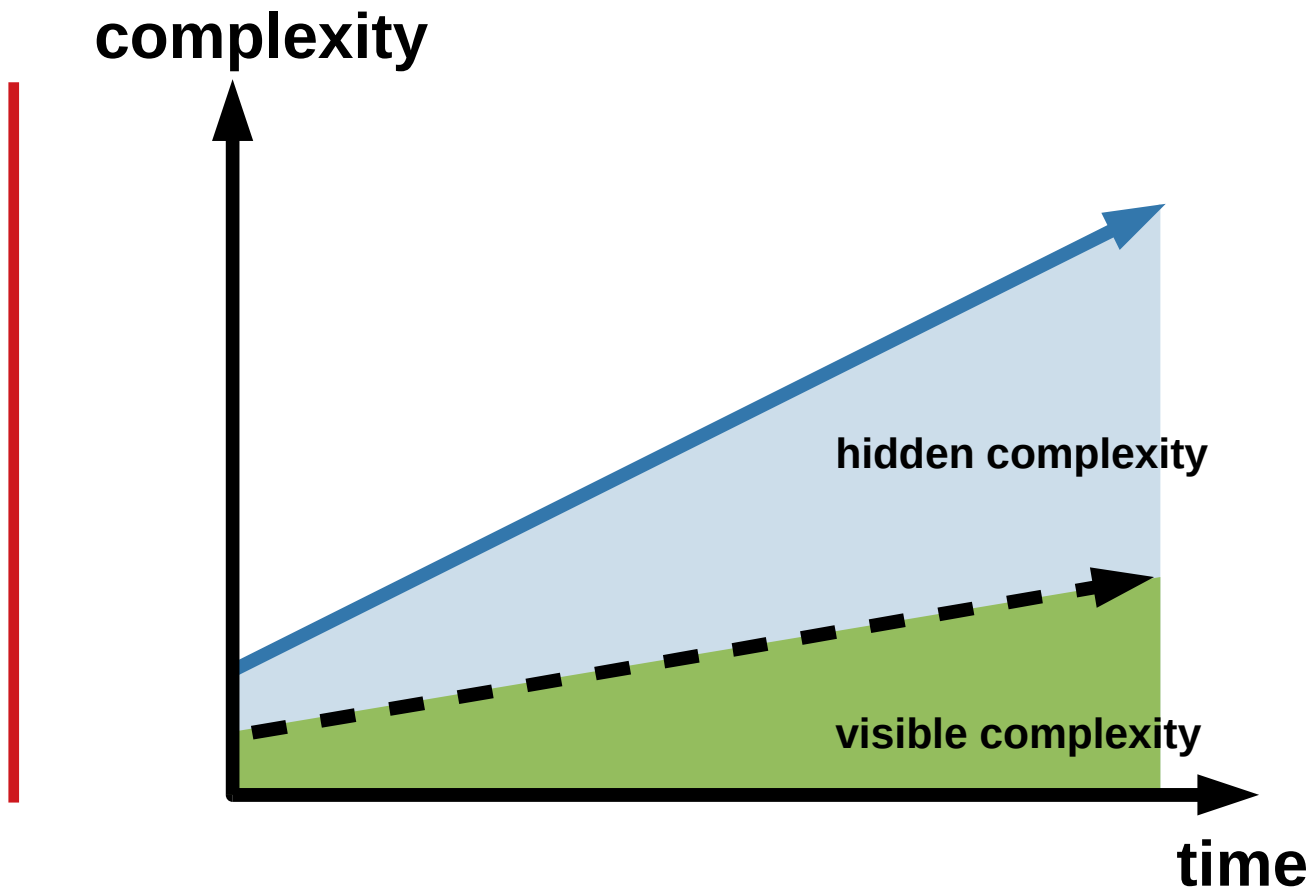


# Document Complexity

- **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 ensuring the correctness, and value of each individual document.
- **DATA** - You need to understand the structure of data, storage locations and techniques, extraction, backup, transmission methods, normalisation, consolidation, translation, manipulation, and sorting, plus security, privacy, and data governance.



# Hidden Complexity







# Brain and Computer

Brain

**Red**

Computer

**#FF0000**



# Description of Colours

## **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"



# Real Dates vs Excel Dates

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 PC	01/09/1983	30560
Italo Vignoli Wedding	08/09/1984	30933
Italo Vignoli Installs OOo	02/01/2003	37623
Italo Vignoli Launches LibreOffice	28/09/2010	40449



# Comparison ODF/OOXML

File Edit View Insert Format Styles Table Tools Window Help

Text Body Liberation Serif 12

### Lorem Ipsum

"Neque porro quisquam est qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit..."

Lorem ipsum dolor sit amet, consectetur adipiscing elit. In at sodales arcu, sollicitudin eleifend justo. Integer in ornare leo. Pellentesque ultrices sapien phanetra lectus finibus cursus. Vestibulum pulvinar, ex sed interdum phanetra, est risus efficitur arcu, ac aliquet ante diam ac eros. Sed vitae dui nec odio rutrum ornare. Cras in neque et neque cursus vehicula. Mauris posuere sed lorem at faucibus.

Morbi sagittis dignissim libero, at porta lectus ultricies in. Integer dui erat, mollis malesuada nisi nec, malesuada porttitor metus. Donec congue pretium mollis. Nulla hendrerit sit amet felis sit amet euismod. Ut quis sem sem. Nulla volutpat justo neque, eget sagittis orci faucibus eget. Vivamus vestibulum nulla et metus cursus iaculis. Proin nec iaculis sapien. Donec pulvinar erat id tellus rhoncus aliquam. Nunc et nisl sagittis, congue sapien in, egestas risus. Mauris efficitur volutpat dictum.

- Praesent tempus sodales lacus at accumsan.
- Pellentesque vel neque rutrum, posuere sapien eget, eleifend leo.
- Proin ut ultricies magna.
- Nulla facilisi.
- Donec sed risus quis tellus facilisis elementum.
- Vivamus sit amet semper nisl, vel tristique mi.
- Aenean lobortis neque in malesuada ornare.

**Sed condimentum sit amet lectus vel cursus.**

Mauris vel ornare nibh, ex gravida lectus. Etiam elementum nunc at tortor cursus, in bibendum dui finibus. Ut non lacus pretium, dictum dui et, fringilla leo. Nulla quis porta mauris, eget aliquet diam. Maecenas laoreet sem erat, eu mollis ex blandit eu. Nam vel fringilla magna, vel convallis massa.

Vestibulum tristique urna sit amet molestie volutpat. Sed in dapibus est, in tristique risus. Vivamus consequat odio nec consectetur eleifend. Ut tempus felis finibus congue porta. Aliquam ante tellus, faucibus nec ultrices hendrerit, ultrices vitae urna. Phasellus hendrerit eros eget sapien imperdiet semper. Nullam vestibulum mauris ut vehicula mollis. Ut id tincidunt diam, varius placerat lorem. Sed vitae accumsan orci. Quisque id mollis elit. Integer consectetur tellus eget sem porta sagittis. Suspendisse hendrerit tristique eleifend. Ut elementum eu nisl mollis elementum. Nunc at est nec tortor varius malesuada. Integer eget odio ac mi sodales rutrum.

**Nulla facilisi.**

Quisque pretium urna sit amet urna posuere posuere. Ut velit arcu, ultricies in velit sed, posuere laoreet mauris. Cras sed ante vitae urna ornare congue sed id nisl. Ut vel mollis sem. Aliquam vulputate finibus quam sed sagittis. Cras sed lacus eget mauris hendrerit suscipit ac nunc. Donec at nibh et nulla vestibulum aliquam. Sed erat dolor, finibus vel sapien et, lobortis hendrerit elit. Nullam dictum ante vitae tempor suscipit. Aliquam erat

Aenean ultricies turpis eu erat euismod porta.	Maecenas nec ullamcorper odio, ac tincidunt metus.
Duis volutpat sapien at interdum consequat.	Donec vel semper nisl.
Sed scelerisque odio non justo suscipit congue.	Maecenas convallis quis neque at elementum.
In fringilla tortor id orci porttitor, a euismod velit molestie.	

**Sed vel enim molestie, viverra sapien in, lobortis erat.**

Suspendisse at mauris eget mi tempus dapibus vel in ante. Vivamus euismod, nisl commodo phanetra 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.

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.

**Aliquam lobortis rutrum massa.**

Aenean iaculis 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 ultricies 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 punis, 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.

Page 2 of 2 811 words, 5,315 characters Default Style Italian (Italy) 95%



# Length of Content XML

Version	XML Lines
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 Content XML

Version	XML Lines
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 201	7731



`<start text>`**To be, or not to be,  
this is the question.**`<end text>`



# DOCX by Microsoft Office

<start text>**To**<end text>

<keep together>

<start text>**[space]**<end text>

<start text>**be**<end text>

<start text>,<end text>

<keep together>

<start text>**[space]**<end text>

<start text>**or**<end text>

<keep together>

<start text>**[space]**<end text>

<start text>**not**<end text>

<keep together>

<start text>**[space]**<end text>

<start text>**to**<end text>

<keep together>

<start text>**[space]**<end text>

<start text>**be**<end text>

<keep together>

<start text>,<end text>

<keep together>

<start text>**[space]**<end text>

<start text>**this**<end text>

<keep together>

<start text>**[space]**<end text>

<start text>**is**<end text>

<keep together>

<start text>**[space]**<end text>

<start text>**the**<end text>

<keep together>

<start text>**[space]**<end text>

<start text>**problem**<end text>

<start text>.<end text>



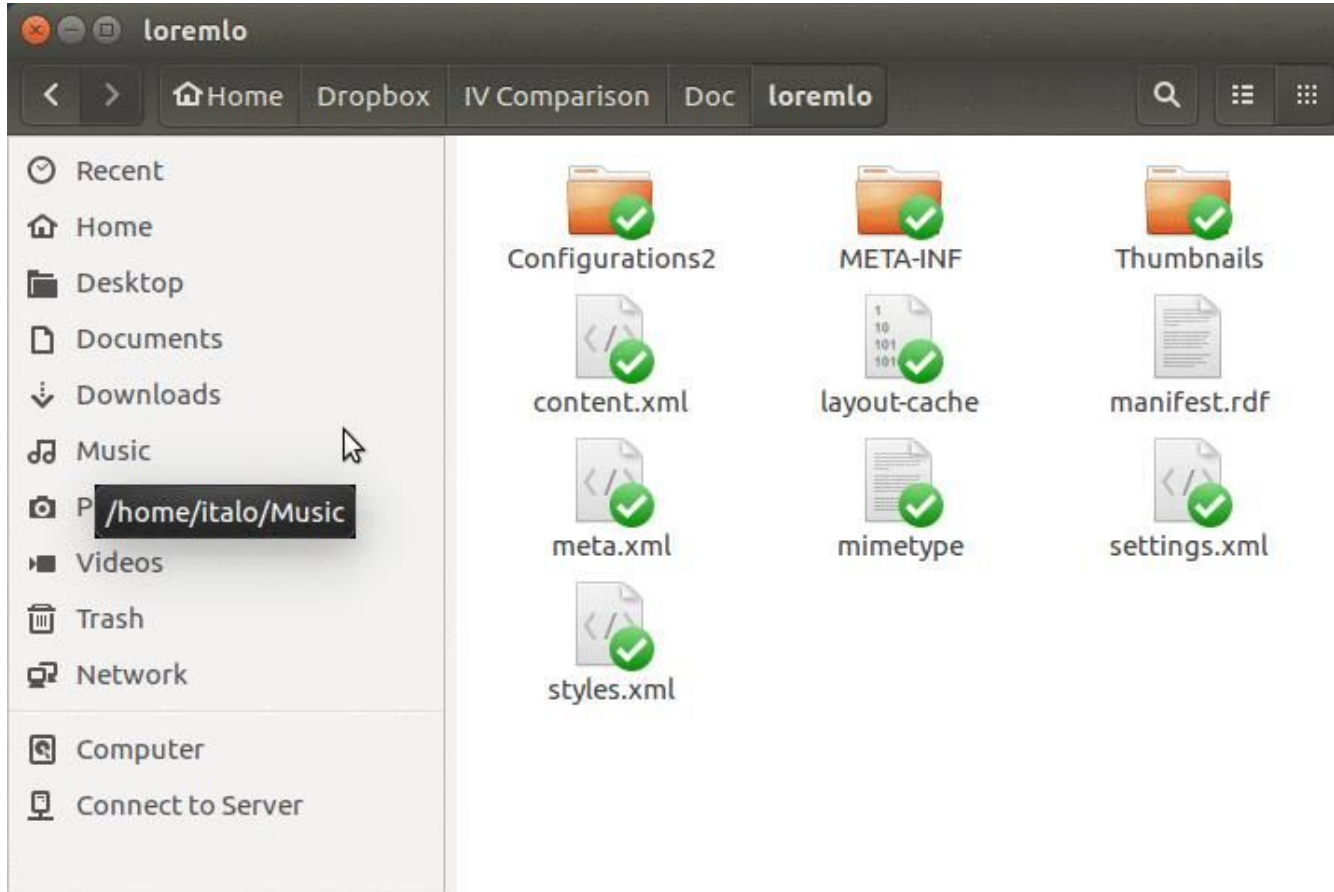


# Simplicity vs Hidden Complexity

- ODT by LibreOffice
  - Low or no hidden complexity
  - Same approach when writing OOXML
  - Files are human readable (security)
- OOXML by Microsoft Office
  - Highest option of hidden complexity
  - Same approach when writing ODT
  - Files are not human readable

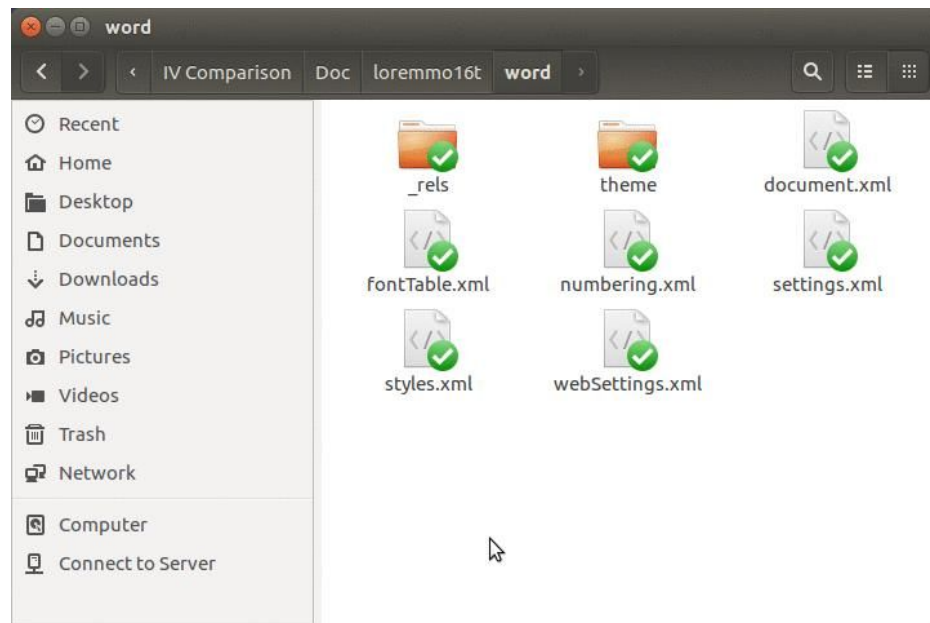
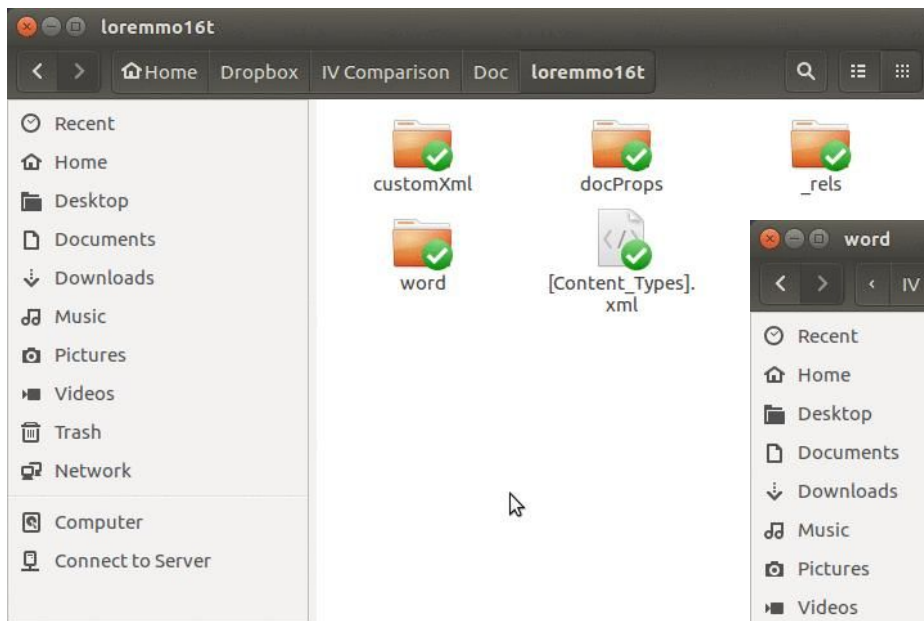


# Inside ODT





# Inside DOCX





# Comparison Spreadsheet

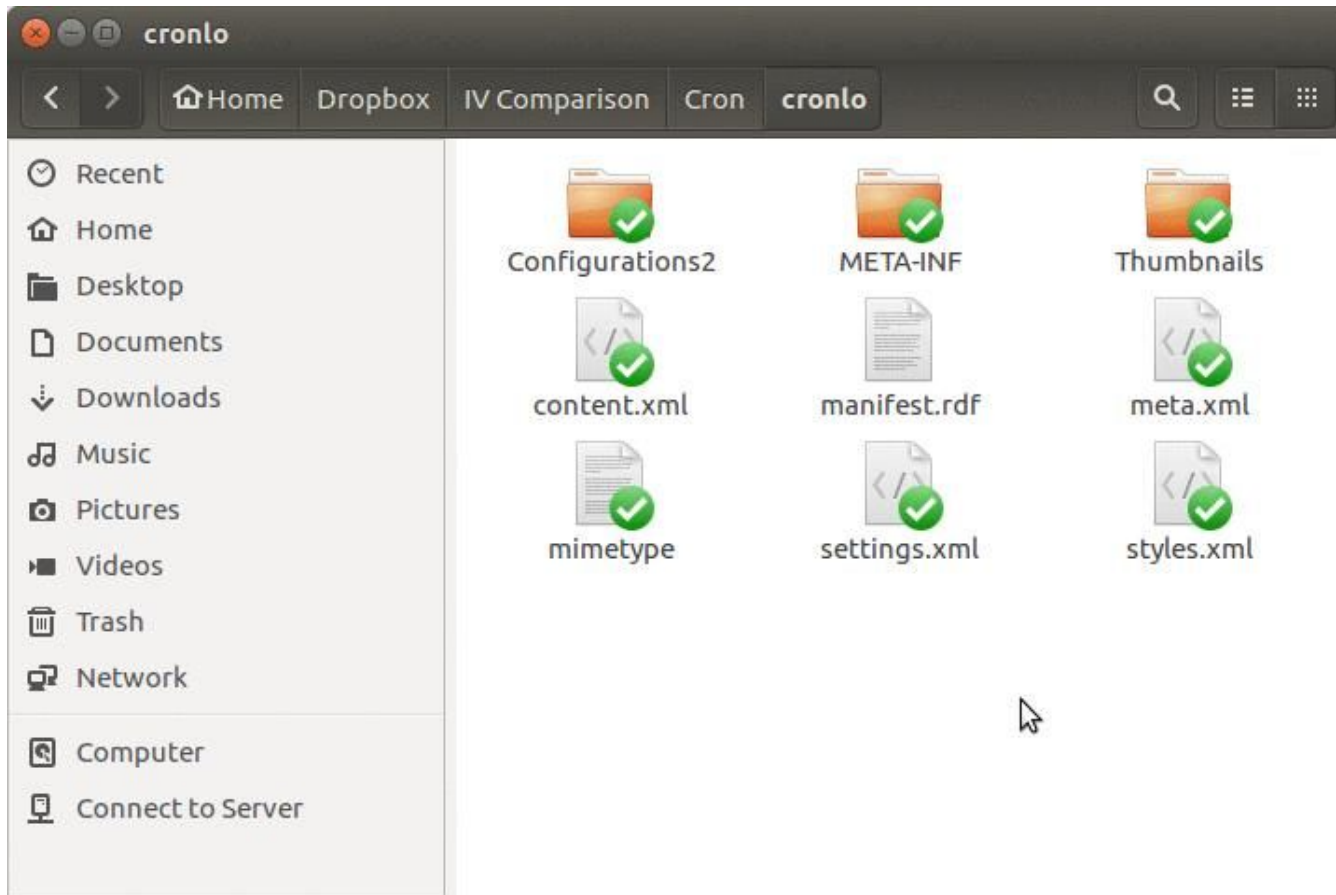
The screenshot displays a spreadsheet application interface. The main area shows a grid with columns numbered 1 to 11 and rows numbered 42767 to 42795. The first column contains dates from 2017-02-02 to 2017-02-28. The spreadsheet is titled 'R42795C10'. The Properties panel on the right is open, showing the following settings:

- Styles:** Default
- Character:** Liberation Sans, 10
- Number Format:** General, Decimal places: 1, Leading zeroes: 1
- Alignment:** Text orientation: 0°, Vertically stacked: unchecked
- Cell Appearance:** Background: #FFFFFF

The status bar at the bottom shows 'Sheet 1 of 1', 'Default', and 'Sum: 0'. A red vertical bar is positioned on the left side of the spreadsheet area.

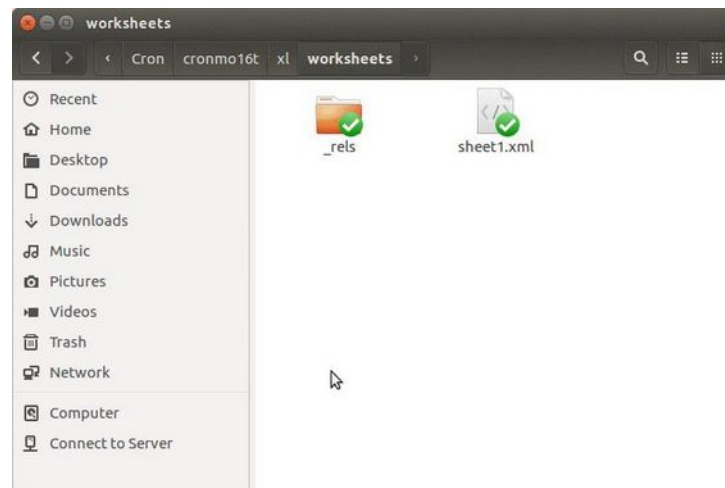
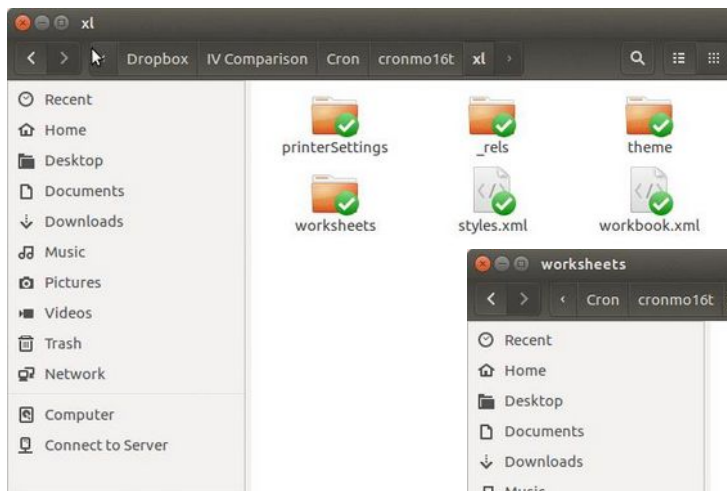
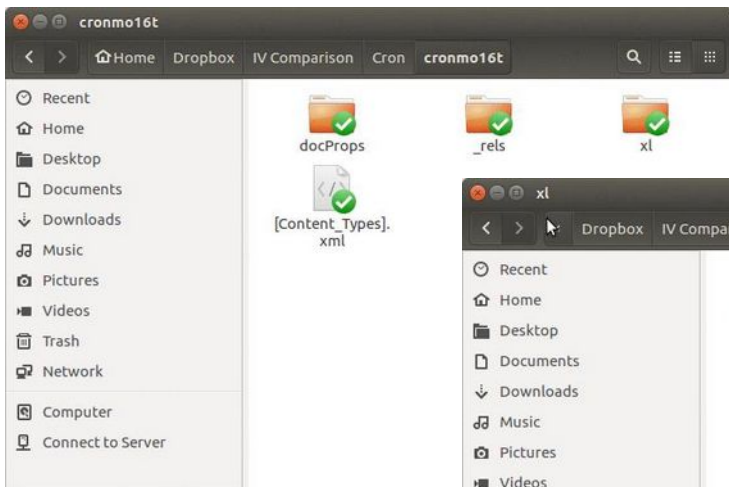


# Inside ODS





# Inside XLSX





# Comparison Presentation

File Edit View Insert Format Slide Slide Show Tools Window Help

Normal Outline Notes Handout Slide Sorter

1 **1** Lorem Ipsum

Neque porro quisquam est qui dolorem ipsum quia dolor sit amet, consectetur, adipisci velit...

2 **2** Lorem ipsum dolor sit amet, consectetur adipiscing elit. In at sodales arcu, sollicitudin eleifend justo. Integer in ornare leo. Pellentesque ultricies sapien pharetra lectus finibus cursus. Vestibulum pulvinar, ex sed interdum pharetra, est risus efficitur arcu, ac aliquet ante diam ac eros. Sed vitae dui nec odio rutrum ornare. Cras in neque et neque cursus vehicula. Mauris posuere sed lorem at faucibus.


3 **3** Mauris efficitur volutpat dictum

- Praesent tempus sodales lacus at accumsan.
- Pellentesque vel neque rutrum, posuere sapien eget, eleifend leo.
- Proin ut ultricies magna.
- Nulla facilisi.
- Donec sed risus quis tellus facilisis elementum.
- Vivamus sit amet semper nisl, vel tristique mi.
- Aenean lobortis neque in malesuada ornare.


4 **4** Proin sed mattis elit

Aenean ultricies turpis eu erat euismod porta.	Maecenas nec ullamcorper odio, ac trincidunt metus.
Duis volutpat sapien at interdum congue.	Donec vel semper nisl.
Sed scelerisque odio non justo suscipit congue.	Maecenas convallis quis neque at elementum.
In fringilla tortor id orci porttitor, a euismod velit molestie.	










5 **5** Aliquam lobortis rutrum massa




6 **6**



7 **7** Nullam ac laoreet orci

			<input type="checkbox"/>
			<input type="checkbox"/>
			<input type="checkbox"/>

8 **8** It could work



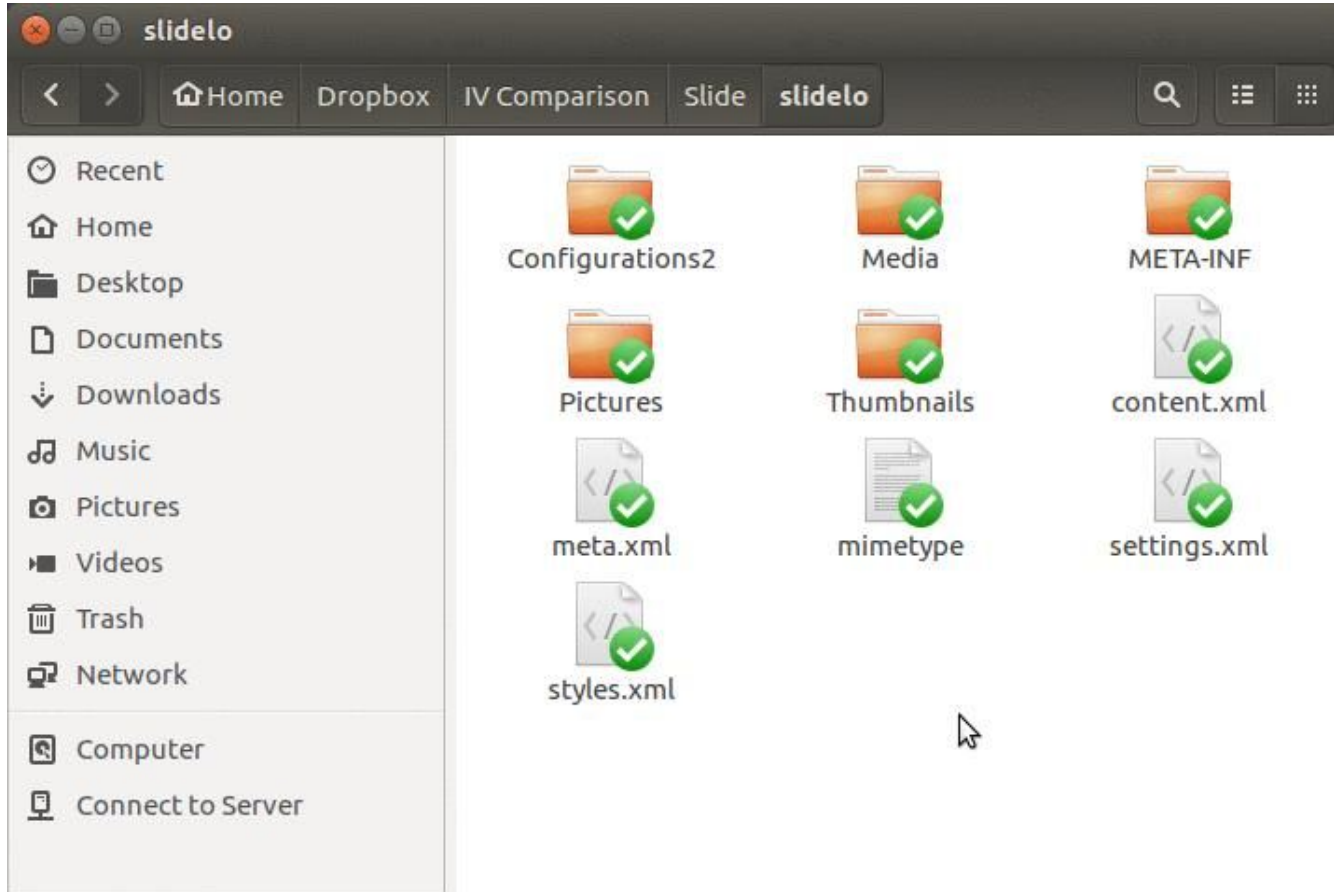
4 Slides

Find Find All Match Case

Slide 1 of 8 Default



# Inside ODP







# Inside PPTX

The image displays three overlapping Windows Explorer windows illustrating the internal structure of a PPTX file:

- Top Window (slidemo16t):** Shows the root directory of the presentation. It contains folders for `docProps`, `ppt`, and `_rels`, along with a file `[Content_Types].xml`.
- Middle Window (ppt):** Shows the contents of the `ppt` folder. It includes sub-folders `media`, `_rels`, and `slideLayouts`, and files `slideMasters`, `slides`, `theme`, `presentation.xml`, `presProps.xml`, and `tableStyles.xml`.
- Bottom Window (slides):** Shows the contents of the `slides` folder, which contains individual slide files: `slide1.xml`, `slide2.xml`, `slide3.xml`, `slide4.xml`, `slide5.xml`, `slide6.xml`, `slide7.xml`, and `slide8.xml`.



# Deductions of a Stupid Me

- LibreOffice developers are a bunch of geniuses
- Microsoft Office developers are a bunch of \*\*\*\*\*  
unless
- Microsoft Office XML files are artificially stuffed with useless contents to reduce the chances that software other than Microsoft Office can open them properly
- Microsoft has a vested interest in killing standard based interoperability to protect a captive market still valued at over 25 billion dollars



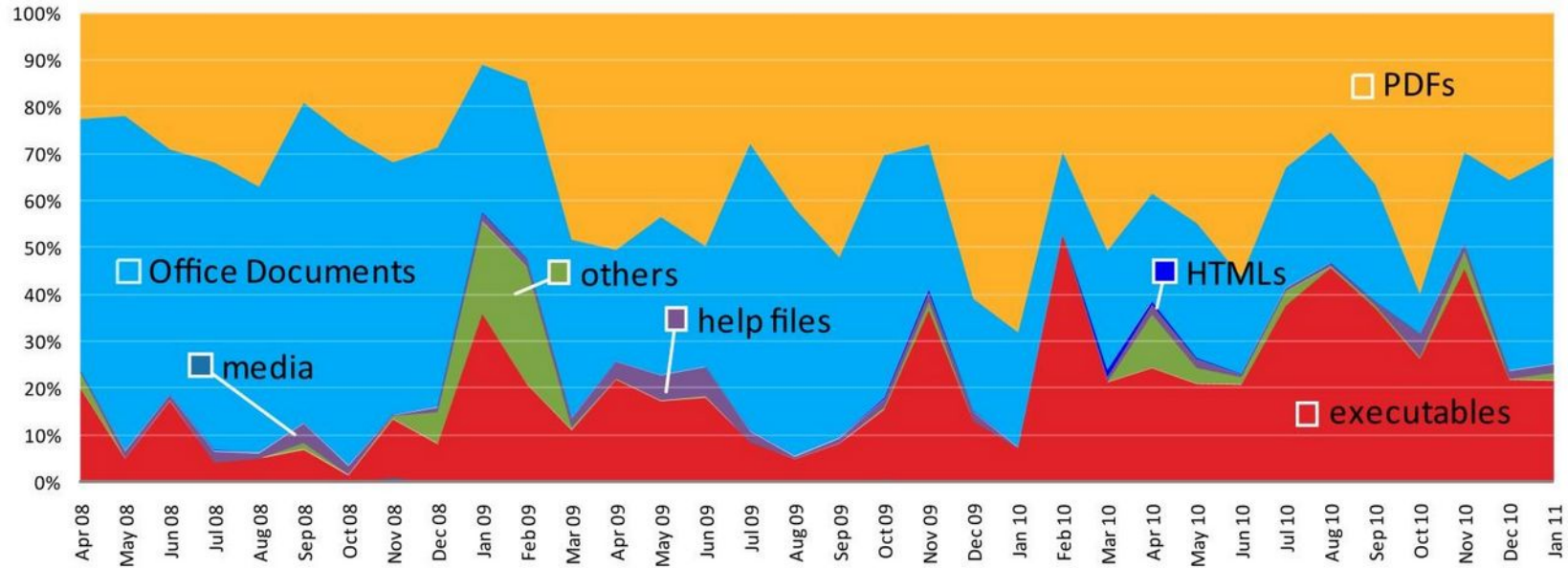
# Complexity of Document Formats

Document Format	# Elements
Office Open XML	1792
WordprocessingML	780
OASIS Open Document	530

AS OF 2006



# File Types Used in Attacks

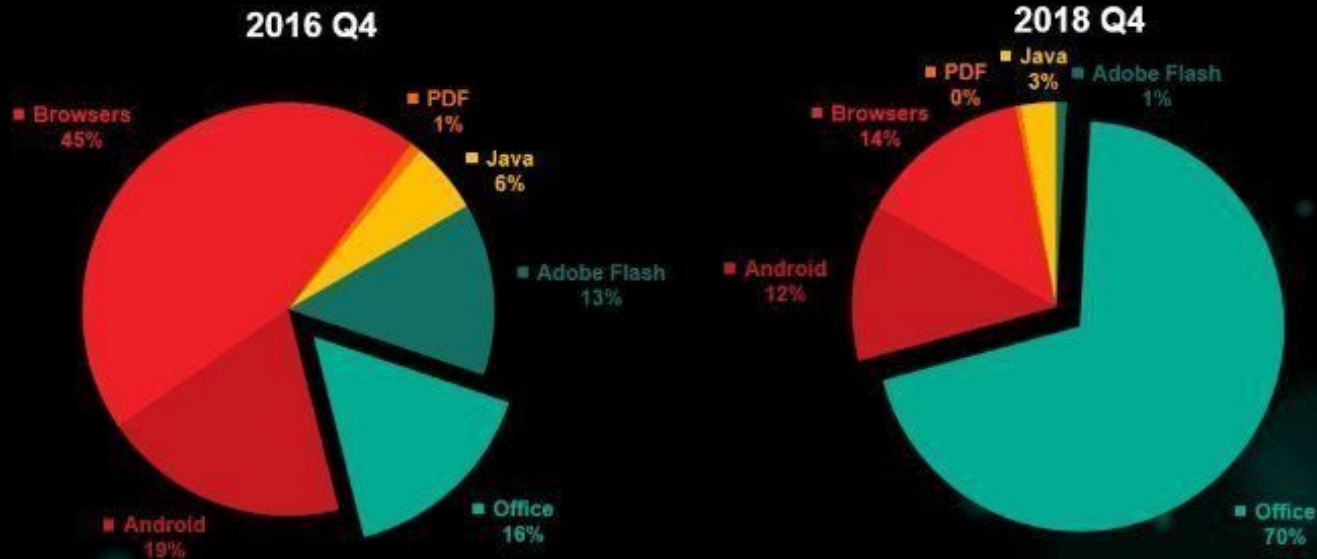


Source: Symantec MessageLabs Intelligence, February 2011 Intelligence Report



# File Types Used in Attacks

## Targeted platforms by attacked users



Source: Kaspersky Labs, Spring 2019 Worldwide Meeting

Except where otherwise noted, content on this presentation is licensed under a Creative Commons Attribution 4.0 International license



All text and image content in this document is licensed under the Creative Commons Attribution-Share Alike 4.0 License (unless otherwise specified). "LibreOffice" and "The Document Foundation" are registered trademarks. Their respective logos and icons are subject to international copyright laws. The use of these thereof is subject to trademark policy.