



DATI A CHI?

WWW.LINUXDAY.IT
23 OTTOBRE 2021



Formati standard: c'è chi è aperto e c'è chi mente

*Emiliano Vavassori
BgLUG - LibreItalia - TDF*

#LinuxDay2021 – 23 ottobre 2021



Indice

- Cos'è un formato di file
- I formati proprietari e aperti
- Cos'è uno standard
- Standard *de jure* e *de facto*
- Formati di documenti: ODF e OOXML
- Processo di standardizzazione
- La prova sul campo
- Conclusioni

Cosa è un formato di file?

Indica la *convenzione* che viene usata per leggere, scrivere e interpretare i contenuti di un file.

**Formato di file,
Wikipedia in Italiano**



Tecnologia proprietaria (non standard)

- Un monopolio naturale nella tecnologia può portare a un mercato di monopolio per prodotti e servizi basati su quella tecnologia
- La tecnologia è di fatto disponibile solo a chi ne detiene i diritti



Tecnologia aperta

- Esiste competizione completa e naturale sul mercato per prodotti e servizi basati su quella tecnologia
- È il risultato di un mercato aperto in cui tutti i soggetti hanno eguale accesso alla tecnologia, senza alcun vantaggio a priori basato sulla presenza di diritti e/o della definizione della tecnologia



Formato proprietario

- Progettato per essere manipolato da un solo software
- Evolve sulla base di precise esigenze commerciali e non di necessità dei suoi utenti
- Spesso è una rappresentazione seriale degli oggetti in memoria (contenuti *binari*)
- Il software è il formato!
- Produce *lock-in* sul software
- Gli utenti *prendono in prestito* il contenuto che scrivono dai produttori di software attraverso specifiche *End User License Agreement* (EULA)

Formato aperto

- **Indipendente** da un singolo prodotto: chiunque può scrivere un software in grado di gestire un formato aperto;
- **Interoperabile**: permette la condivisione trasparente di informazioni fra sistemi eterogenei;
- **Neutrale**: non deve forzare l'utente ad adottare (più spesso comprare) un prodotto specifico, ma lascia una ampia scelta in base al rapporto qualità/prezzo;
- **Perenne**: protegge il contenuto sviluppato dagli utenti dall'evoluzione/obsolescenza della tecnologia.

Cosa è uno standard (tecnologico)?

Un insieme di norme e requisiti prestabiliti per ottenere **un risultato tecnico ripetibile**. Solitamente si tratta di un **documento formale** che stabilisce **criteri tecnici/ingegneristici, metodi, processi e pratiche**.

Al contrario, **usi e convenzioni comuni o aziendali, prodotti di aziende** e così via possono essere **generalmente accettati e dominanti sul mercato**, tali che sono spesso chiamati *standard de facto*.

**Standard tecnologico,
Wikipedia in Inglese**



Standard *de jure* e *de facto*

Uno standard **de jure** è il prodotto di un accordo collettivo

- Maggiore interoperabilità
- Previene il *lock-in*
- Riduce i costi
- Crea un mercato trasparente
- Riduce la complessità

Uno standard **de facto** è il prodotto di una quota significativa di mercato

- Minore interoperabilità
- Aumenta il *lock-in*
- Aumenta i costi
- Crea un mercato opaco/monopolistico
- Aumenta la complessità

Formati standard di documenti d'ufficio



Microsoft OpenOfficeXML
ISO/IEC DIS 29500:2008



ODT



ODS



ODP



ODG

Open Document Format
ISO/IEC 26300:2006

OOXML

- Dublin Core

ODF

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

Il processo di standardizzazione

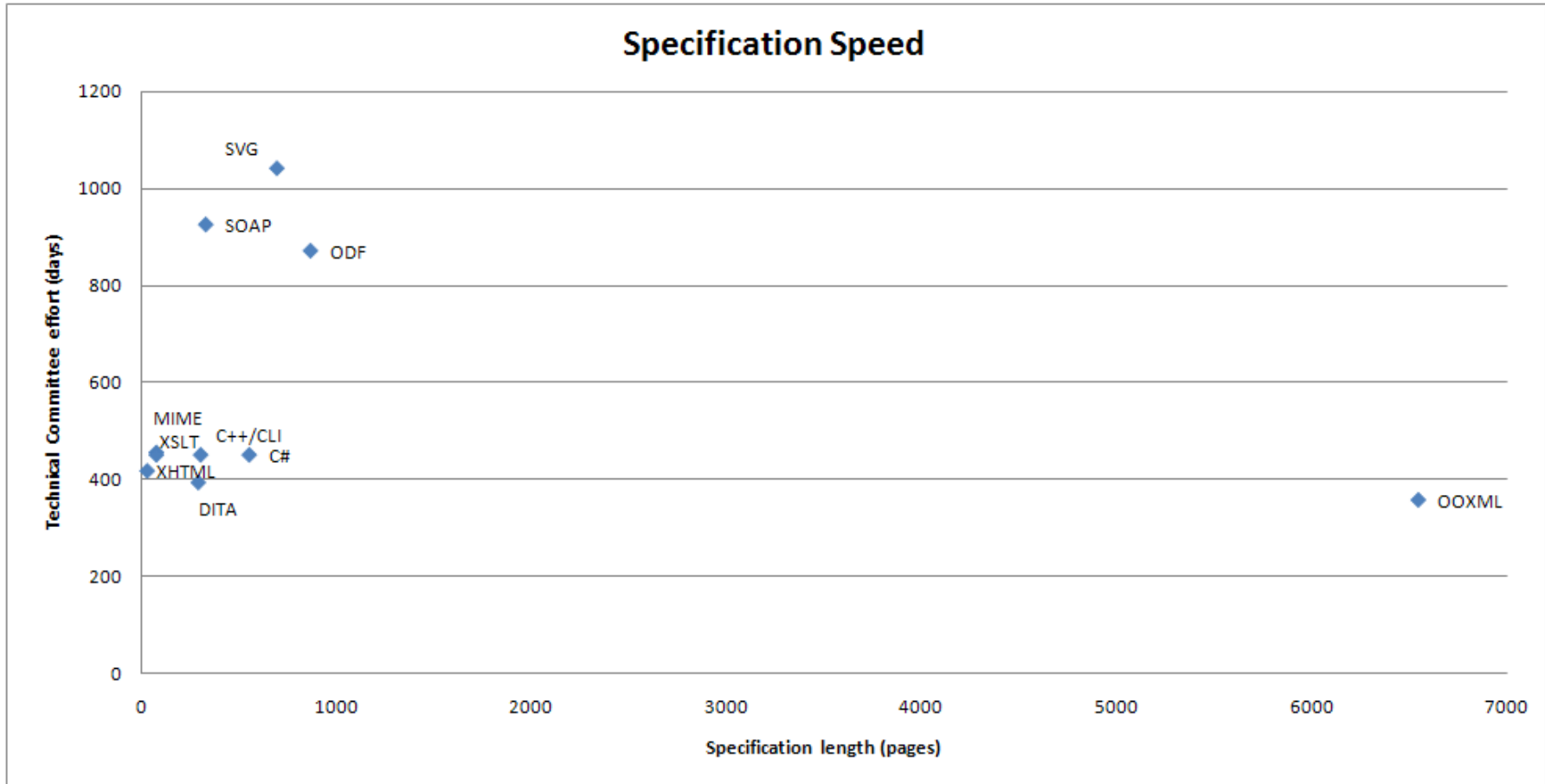
OOXML

- Basato sul formato XML di Microsoft Office 2003
- 15 dicembre 2005: il formato è presentato a ECMA
- 31 dicembre 2006: standard ECMA approvato dall'Assemblea Generale
- 31 gennaio 2007: OOXML è presentato come ISO/IEC JTC1 basato su Fast Track
- 31 marzo 2008: lo pseudo standard OOXML è approvato
- Revisione: 7200 pagine in 838 giorni

ODF

- Basato sul formato XML di OpenOffice.org
- 12 dicembre 2002: il formato è presentato ad OASIS
- 1 maggio 2005: ODF è rilasciato da OASIS
- 16 novembre 2005: ODF è presentato come ISO/IEC JTC1 basato su Specifiche Pubblicamente Disponibili (Publicly Available Specification, PAS)
- 3 maggio 2006: ODF è approvato come standard ISO/IEC IS 26300 standard
- Revisione: 720 pagine in 1239 giorni

Un grafico lo spiega meglio...



Qualche piccolo problemino...

- Interoperability woes with MS-OOXML
Free Software Foundation Europe
- Google: Microsoft's OOXML shouldn't be a standard
Larry Dignan, ZDNet
- Google to ISO: Say no to OOXML
Ryan Paul, Ars Technica
- ODF backers urges cooperation to fix OOXML's
deficiencies
Ryan Paul, Ars Technica

Più nella pratica:

- OOXML *Strict* è lo standard aperto approvato da ISO, ma risulta essere l'ultima opzione dell'elenco dei formati supportati da (quasi) tutte le versioni della suite MS Office nella funzionalità “Salva con nome...”
- Non essendo di default ed essendo “difficile da trovare”, possiamo supporre che il 100% dei documenti prodotti da MS Office usino il formato OOXML *Transitional*, pertanto siano **proprietary**
- Il formato OOXML *Transitional* doveva essere un formato di transizione, per l'appunto, dai vecchi formati *legacy* di Office (.doc, .xls, .ppt) e avrebbe dovuto essere abbandonato dalla versione 2010 della suite MS Office (dichiarazione Microsoft nel 2007).

Presenza del formato OOXML Strict in MS Office

- MS Office 2010: NO
- MS Office 2013: Sì, ma il default è Transitional
- MS Office 2016: Sì, ma il default è Transitional
- MS Office 2019: Sì, ma il default è Transitional
- MS Office macOS: NO
- MS Office 365: NO

La prova sul campo – struttura documentale

The screenshot shows a word processing application window with a menu bar (File, Edit, View, Insert, Format, Styles, Table, Tools, Window, Help) and a toolbar. The document content is as follows:

>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 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.

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, eu 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 veli arcu, ultricies in veli 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 a 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

volutpat. Proin sed mattis elit. Morbi et diam facilisis, laoreet quam sit amet, egestas libero.

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 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.

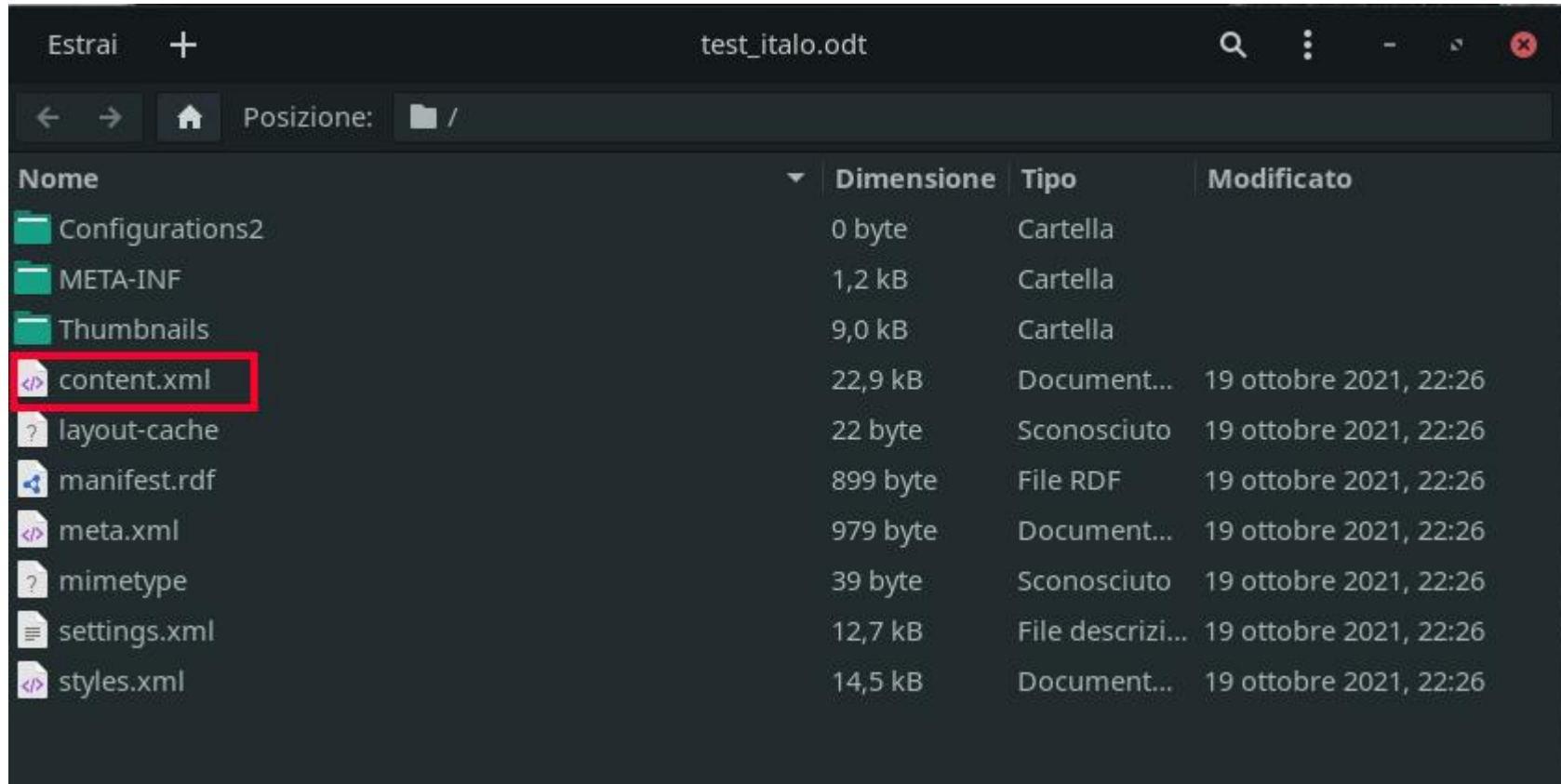
Maecenas eleifend accumsan lacus, in elementum sapien sollicitudin non. Etiam a augue sed metus viverra gravida. Nullam interdum felis nec lectus congue ultrices. Nunc lobortis ex eu lobortis vehicula. Orci varius natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aenean felis 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 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.

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

La prova sul campo – struttura del file



The screenshot shows a file manager window with the title bar 'test_italo.odt'. The address bar indicates the current position is the root directory. The file list is as follows:

Nome	Dimensione	Tipo	Modificato
Configurations2	0 byte	Cartella	
META-INF	1,2 kB	Cartella	
Thumbnails	9,0 kB	Cartella	
content.xml	22,9 kB	Document...	19 ottobre 2021, 22:26
layout-cache	22 byte	Sconosciuto	19 ottobre 2021, 22:26
manifest.rdf	899 byte	File RDF	19 ottobre 2021, 22:26
meta.xml	979 byte	Document...	19 ottobre 2021, 22:26
mimetype	39 byte	Sconosciuto	19 ottobre 2021, 22:26
settings.xml	12,7 kB	File descrizi...	19 ottobre 2021, 22:26
styles.xml	14,5 kB	Document...	19 ottobre 2021, 22:26

Risultati delle prove con formati diversi

Numero di linee di codice del file content.xml (o corrispettivo)

Versione	Righe XML
ODF 1.2 (qualsiasi versione di) 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

OOXML durante l'anno

Versione	Righe XML
Windows OOXML 2013 Transitional estate 2017	1590
Windows OOXML 2013 Transitional inverno 2018	13515
Windows OOXML 2016 Transitional estate 2017	11667
Windows OOXML 2016 Transitional inverno 2018	969
Windows OOXML 2016 Transitional autunno 2018	11288
Windows OOXML 2016 Transitional primavera 2019	7085
macOS OOXML 2016 Transitional estate 2017	11646
macOS OOXML 2016 Transitional autunno 2018	854
macOS OOXML 2016 Transitional primavera 2019	7731

Considerazioni finali 1: il *lock-in*

Il rapporto di **dipendenza** che si instaura tra un cliente ed un fornitore di beni o servizi, tale che il cliente si trova nella condizione di **non poter acquistare analoghi beni o servizi** da un fornitore differente **senza dover sostenere rilevanti costi e rischi** per effettuare questo passaggio.



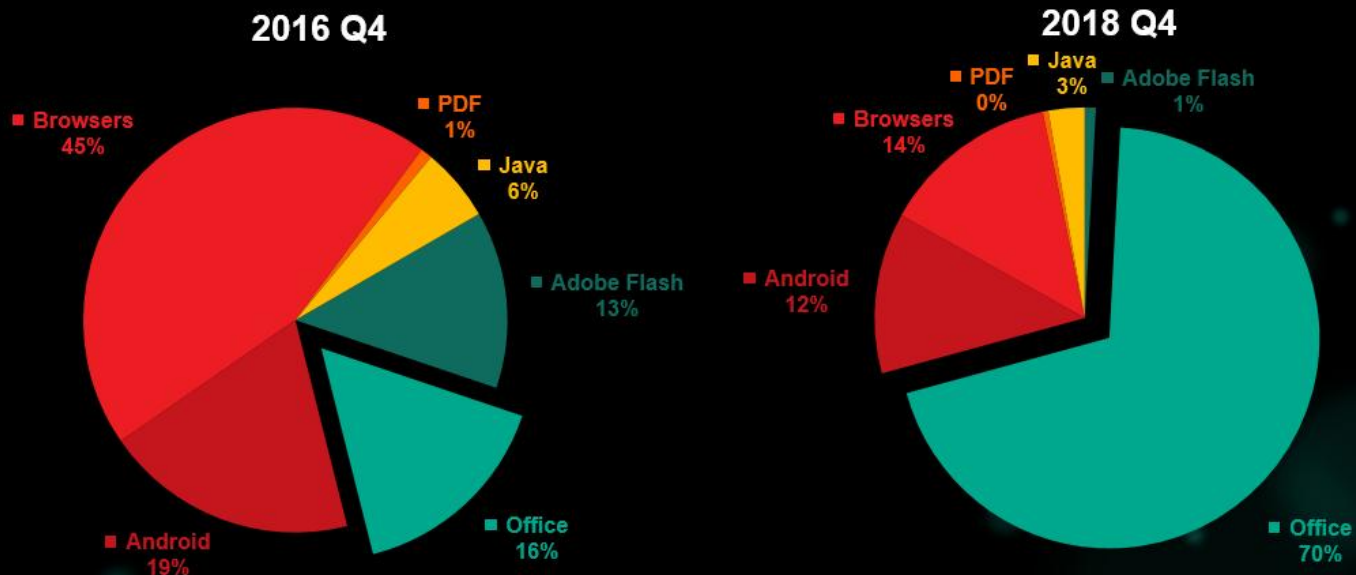
Vendor lock-in,
Wikipedia in Italiano

Considerazioni finali 2: interoperabilità, mercato e futuribilità

- I file OOXML sono riempiti artificialmente di contenuti inutili per non permettere ad altri programmi di aprirli correttamente (**non interoperabilità**)
- Microsoft ha un interesse commerciale evidente: ha un mercato che vale **oltre 25 miliardi di dollari**
- Non è possibile **avere alcuna certezza** che, in futuro, qualsiasi versione di file OOXML sia correttamente apribile e “decifrabile” da una suite di *office automation* (ma: [Document Liberation Project](#))

Considerazioni finali 3: sicurezza

Targeted platforms by attacked users



Fonte: Kaspersky Labs, 2019

Grazie!

Emiliano Vavassori

BgLUG – Bergamo Linux Users Group - bglug.it

LibreItalia - libreitalia.org

The Document Foundation - documentfoundation.org

emiliano.vavassori@libreitalia.it ✉

syntaxerrormmm@libreoffice.org ✉

[@syntaxerrormmm](https://twitter.com/syntaxerrormmm) 🐦

[@syntaxerrormmm](https://www.linkedin.com/company/syntaxerrormmm) 📌



DATI A CHI?

WWW.LINUXDAY.IT
23 OTTOBRE 2021



Quest'opera è distribuita con Licenza Creative Commons
Attribuzione/Condividi allo stesso modo 4.0 Internazionale.