

FOSS and Open Standards for Digital Sovereignty

Italo Vignoli



What is Digital Sovereignty?

Digital sovereignty is the ability of a state or a federation of states to provide the digital technologies it deems critical for its welfare, competitiveness, and ability to act, and to be able to develop these or source them from other economic areas without one-sided structural dependency















Of the **41** high-level EU Commission lobby meetings in 2016 that discussed **ePrivacy...**





E were with corporate interests were with civil society only

Subservience to Microsoft

Is that true? Is that real? Unbelievable, Microsoft rules

Human Genes Renamed To Please Excel

Written by Janet Swift

Friday, 07 August 2020

More than two dozen human genes have been renamed so that they can be typed into a spreadsheet without being formatted as dates. New guidelines for standardized gene naming explicitly allow for renaming genes to avoid problems with data handling.

EU Digital Sovereignty: State of Play

The influence of non-EU tech companies is a concern for EU policy-makers, especially with regard to their impact on the EU's data economy and innovation potential, on EU privacy and data protection and on the establishment of a secure and safe digital environment



Evolution of Data

Mainframe Computing	Enterprise Systems	Electronic Business	Data Economy
Information systems have been used since the 1960s and 1970s to support enterprise functions, but data wasn't shared between functions, let alone enterprises.	With the proliferation of Manufacturing Resource Planning (MRP) and Enterprise Resource Planning (ERP) in the 1980s and 1990s data enabled end-to-end business processes such as order-to- cash, procure-to-pay, make-to-stock etc.	Since the millennium change, data has increasingly become an enabler of innovative product-service- systems and integrated solutions.	Recently, data marketplaces emerged offering data APIs at a volume or frequency based fee. Data has become a product in its own right.
Data as a Process Result	Data as a Process Enabler	Data as a Product Enabler	Data

Current vs Intended Paradigm

CURRENT

- Data is a fictitious commodity, that can be sold and traded in markets
- Data handling & monetization is opaque
- Even if there is regulation, there is no possible enforcement
- We produce the data, they own it
- Benefits of data sharing are privatized (surveillance capitalism)

INTENDED

- Data is a common resource & infrastructure where to build upon (new services)
- Data handling and usage is transparent & privacy-friendly
- Data is shared according to rules set by common (enforceable) governance
- Shared benefits of data sharing
- New political, economic, and legal regime that recognize social and communal rights to data

Protecting versus Sharing

Interoperability Data Exchange "Sharing Economy" Data Centred Services

Proprietary Data Data Protection Data Value



Enabling the Data Commons

- Can we regain control of our data?
- We want to decide who to share it with, under which rules, when and for what purpose
- One can decide in a democratic way based on trust but if data is a monopoly of a few, collective intelligence is lost...



Proprietary versus FOSS

- Proprietary software protects the user by obfuscating algorithms and information, but in this way they also obfuscate the way they handle end user data
- FOSS protects the user with transparency, by sharing source code and all information about methodologies used by projects to manage end user data



Situation with Proprietary SW

- A large percentage of governments in Europe
 - at every level rely on proprietary software for desktop productivity and cloud storage of data, independently from the level of confidentiality
- This puts citizen's personal data, including several extremely confidential information (health) at risk
- This is confirmed by the recent Schrems II sentence from the Court of Justice of the European Union



Awareness of SCC* (Schrems II)



Usage of SCC* (Schrems II)





Source: DIGITALEUROPE | Base: estimated SCC users (n = 249)

* **SCC** = Standard Contractual Clauses (for data protection) when transferring data to the US

Reliance on SCC* (Schrems II)



* **SCC** = Standard Contractual Clauses (for data protection) when transferring data to the US





Source: DIGITALEUROPE I Base: SCC users that are aware of which geography they transfer data to (n = 172)⁸

Cost of Reassessing SCC*





Source: DIGITALEUROPE I Base: respondents that reassessed their use of SCCs (n = 129)

* SCC = Standard Contractual Clauses

What Would Change with FOSS

- By switching to FOSS for desktop productivity and cloud storage, European governments would regain control of citizen's personal data and manage them according to their confidentiality
- In addition, switching to FOSS would include moving from proprietary to standard document formats, with a significant advantage in term of interoperability



Apparently a No Brainer, but...

- Politicians who are not technology experts see GAFAMs as part of the global system, and therefore consider their issues as blockers for the entire digital transformation process (and try to help them)
- On the contrary, politicians because of their limited understanding of technology – do not see FLOSS as part of the global system, and as a consequence do ignore FLOSS as a potential solution



Time to FOSS Vulnerability Fixes

Expectation for open source vulnerability fixes







Document Vulnerabilities in 2011



Document Vulnerabilities in 2018



FSFE Project



Public Money

Public Code

publiccode.eu



Enabling the Data Commons

- Can we regain control of our data?
- We want to decide who to share it with, under which rules, when and for what purpose
- One can decide in a democratic way based on trust but if data is a monopoly of a few, collective intelligence is lost...
- We definitely need true interoperability of contents



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





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



Document Format as a Hindrance?

- Documents are one of the most important objects that move from: (1) citizen to government, (2) government to government and (3) government to citizen
- Production, updating and reproduction of documents is extremely important
- A common problem is that documents (governed by a pseudo standard) can lock users into a particular platform (proprietary operating system & application)



Document Format as a Hindrance?

- Government should be platform independent and allow only true document standards, as pseudo standards can be tweaked in a way not visible to users to prevent document interoperability
- In fact, tweaked standards force citizens to pay a fee to create documents (purchase of a proprietary license), or to accept the intrusive license / spying conditions of a cloud based platform
- Only standards associated to FOSS can solve this problem



Open Document Format

the true document standard which offers freedom of choice



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



Basic Concepts

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



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



Lock In

WE CANNET READ YELR **DICUMENT5** DECUMENT FREEDEM. ERG



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



Digging into Document Formats





The Banana standard for your worst Office documents



OOXML Transitional and Strict

- As of 2020, the Office default for .docx, .xlsx and .pptx is Transitional OOXML, a proprietary document format which was created as a bridge from legacy MS Office formats and the approved ISO Standard.
- OOXML Strict is the ISO approved open standard, but being the non publicized last option on MS Office "file, save as..." menu has not been adopted, so 100% of existing OOXML files we are referring to are proprietary (non standard).



OOXML Strict Standard Support

- MS Office 2010: NO
- MS Office 2013: YES, but default is Transitional
- MS Office 2016: YES, but default is Transitional
- MS Office 2019: YES, but default is Transitional
- MS Office macOS: NO
- MS Office 365: NO
- According to Microsoft statements in 2007, OOXML Strict should have been the default since Office 2010



ODF Philosophy

- The philosophy behind the ODF standard document format was to design a mechanism in a "vendor neutral" manner from the ground up using existing standards wherever possible
- Although this means that software vendors would need to tweak their individual packages more than if they continued down their original routes the benefits for interoperability were important enough to justify the move



OOXML Philosophy

- The OOXML pseudo-standard document format appears to be designed by Microsoft for Microsoft products, and to inter-operate with the Microsoft environment
- Little thought appears to have been exercised for interoperability with non-Microsoft environments or compliance with established vendor-neutral standards



ODF vs OOXML Strategic Difference

- ODF has been designed as a document standard for the next 20-50 years, to liberate users from the lock-in strategy built into yesterday's and today's proprietary formats, and foster interoperability
- OOXML has been designed as a pseudo-standard document format to propagate yesterday's document issues and lock-in strategy for the next 20-50 years, to the detriment of users and interoperability



XML Design Advantages



OOXML Poor XML

- Poor names and inconsistent naming conventions for elements and attributes
- Ecma 376 contradicts the goals of XML which are
 - XML documents should be human-legible and reasonably clear
 - Terseness in XML markup is of minimal importance
- Instead, Ecma 376 often uses unclear names and inconsistent naming conventions
 - These include unnecessary vowel removals, name truncations, and unusual abbreviations, as described in next slide



Differences in Tags and Tag Naming

- The OOXML has shorter tag names, which save file space and facilitates an increase in the speed used of "parsing" the data to convert it to the internal structures the application needs, but increases the number of tags needed in that format.
- The ODF naming is longer and more wordy as it follows the XML convention for naming tags, to ease interoperability when implementing the standard, while file space and slower parsing are offset by the fact there are fewer tags required in this format.



"Naive" Deductions

- All LibreOffice developers are genius
- All Microsoft Office developers 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)



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)



Less Visible More Granted





Thanks

Italo Vignoli The Document Foundation italo@libreoffice.org italo@vignoli.org

