

Formati e standard aperti

reloaded

Limana (BL), 23 ottobre 2010



Non leggo più quel file!

```
risposta = risposta libera risposta forzata
 $\hat{a}(t, t_0, x, u) = \hat{a}(t, t_0, x, 0) + \hat{a}(t, t_0, 0, u)$ 

#@CASO OMOGENEO ( u(t)=0 )## #@CONTINUO
##
La soluzione dell'equazione (vettoriale):
 $x(t) = A x(t_0) \quad x(0) = x_0 \quad x(t) \in \mathbb{R}^n$ 

del tipo:
 $x(t) = e^{At} x_0$  dove  $e = \sum_{i=0}^{\infty} \frac{A^i t^i}{i!}$ 
 $e^{At}$  chiamata esponenziale di matrice ed  $e^{At}$  non singolare per ogni t.
L'esponenziale di matrice  $e^{At}$  chiamata anche "matrice di transizione" del sistema perché governa, in assenza
```

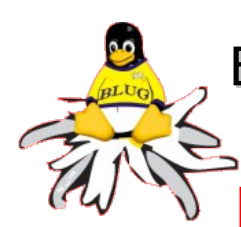


1990

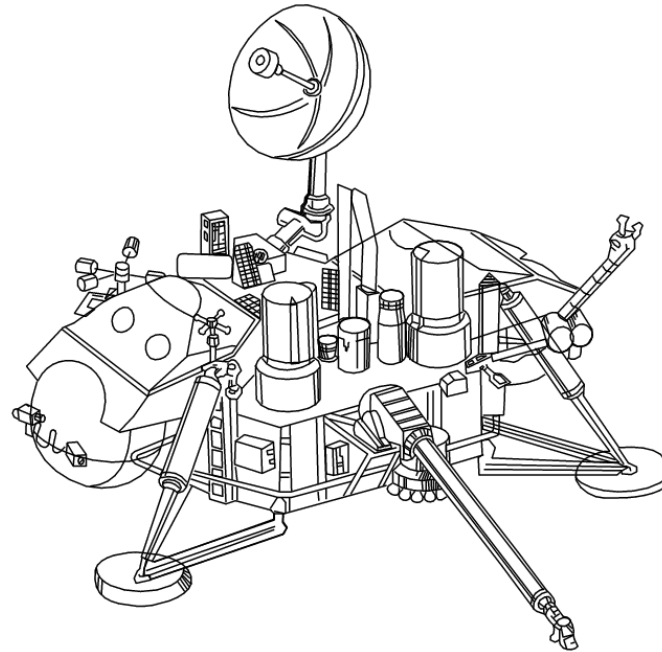
Ai tempi delle superiori (e durante i primi anni di università) ho creato centinaia di documenti con un programma di videoscrittura...

2010

"Il programma non è più disponibile"
"I formati utilizzati non sono pubblici"
"Ho perso per sempre tutti i miei documenti!"



NASA: Viking Lander data...



1975

In 1975 the NASA sent two Viking Landers to Mars planet to find out whether life forms could survive on that planet. At that time It was supposed that the datasets created by the scientists would be available for the future generations of scientific communities on magnetic tape.

1990

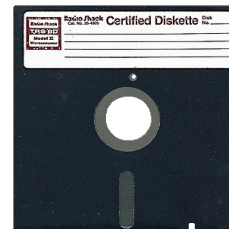
Unfortunately, only a few decades later, despite the space agency's enormous effort to keep the tapes in a climate-controlled environment, tapes cracked . **When scientists tried to re-use some of the data in the late 1990s, they realized It was not possible to decode the data formats used.** Finally they had to track down old printouts and retype everything for recovering the Information.



Obsolescenza?

"Media obsolescence"

- ➡ Vi ricordate i floppy disk?
- ➡ Ancora per quanto tempo ci saranno i CD?
- ➡ Lo sapevate che i media degradano?



"Technology obsolescence"

- ➡ Ogni quanto tempo aggiornate il PC?
- ➡ Qualcuno usa ancora Windows 3.1?



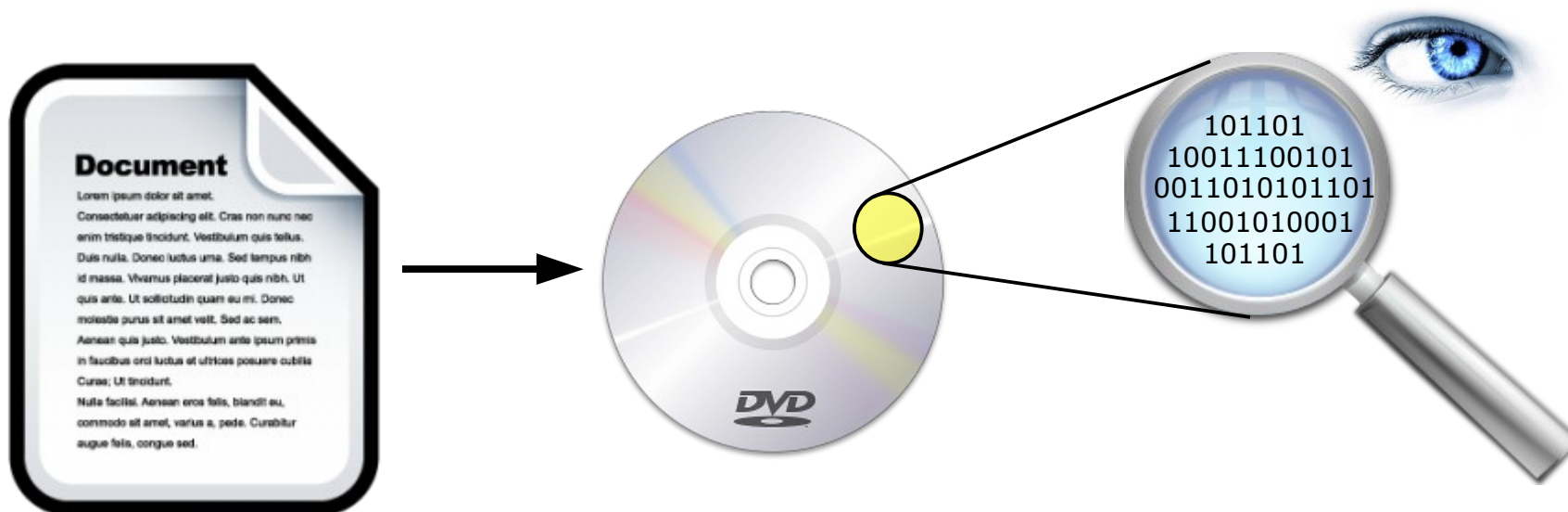
"Software/format obsolescence"

- ➡ C'è ancora qualcuno che usa Wordstar?





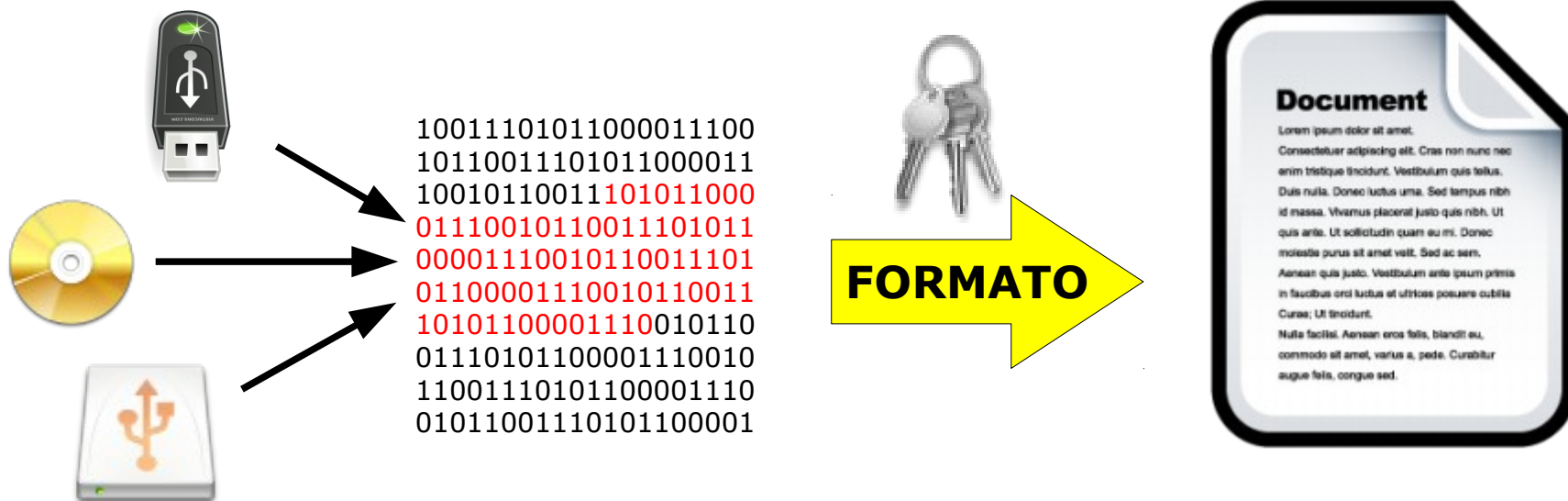
Come sono fatti i files?



➡ Un qualsiasi "contenuto digitale" (documento, immagine, musica, video, email, database...) viene memorizzato come file, ovvero come una **sequenza di bit** (0 ed 1), fissati con una certa organizzazione fisica su una memoria (es. DVD).



Cosa sono i formati?



- ➡ Perché i documenti possano essere aperti o modificati tramite i vostri programmi, è necessario che abbiano un **formato**.
- ➡ Un formato è l'insieme di **regole** che, a partire dalla sequenza di bit che costituisce un oggetto digitale, permettono di riprodurre mediante un software il relativo documento, con lo stesso contenuto e nella stessa forma che gli sono stati conferiti originariamente dall'autore.

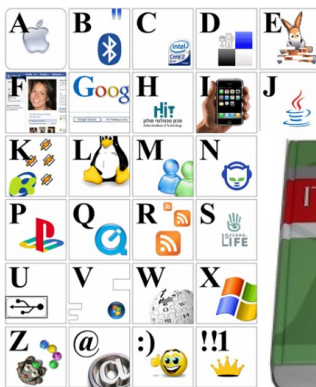


Penne e alfabeti...

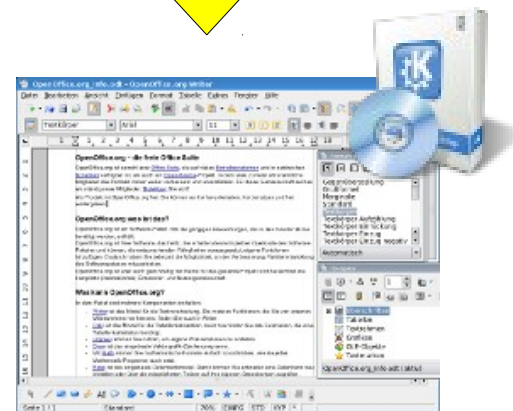
SUPPORTO FISICO



LETTERE, PAROLE, LINGUA, GRAMMATICA (regole ortografiche, lessicali e sintattiche)



"INTERFACCIA UTENTE"



I programmi software sono penne, i formati alfabeti!



Formati & estensioni



Relazione.**doc**



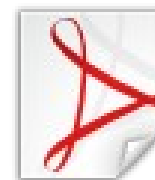
InnoDiMameli.**mp3**



SpeseAcqua.**xls**



Vacanze2010.**avi**



Modulo.**pdf**

- ➡ I formati dei documenti sono molto spesso indicati tramite la loro **estensione** ovvero il suffisso (tipicamente composto di 3 lettere) con cui termina il nome di un file.
- ➡ Quando salviamo i **nostri dati** su un supporto elettronico dovremmo sempre conoscere le **caratteristiche del formato**.



Voi lo fate?



La “non proprietà”

- Un formato è **non proprietario** (oppure **libero**) quando la gestione delle sue specifiche è affidata a una comunità di sviluppatori che cooperano per la gestione condivisa delle stesse (con processi decisionali aperti), o ad un **organismo di standardizzazione**



Il formato per applicazioni da ufficio **ODF** è **libero** (standard ISO 26300)

- Un formato è **proprietario** quando è stato creato e controllato da una organizzazione privata (ad es. un'azienda) che ne detiene i diritti di **proprietà intellettuale** (copyright, brevetti o marchi registrati). Le sue specifiche vengono gestite esclusivamente da tale organizzazione



Il formato **DOC** (così come l'XLS, il PPT e l'RTF) è **proprietario** (Microsoft);

La "apertura"

- Un formato è **aperto** (o **pubblico**) quando le sue specifiche (tutte!) sono pubbliche, liberamente accessibili (ad esempio perché sono state pubblicate sul web) e completamente documentate

➡ I formati **ODF** e **PDF** sono non proprietari ed **aperti**

➡ Il formato **RTF** è proprietario (Microsoft), ma **aperto**

- Un formato è **chiuso** (o **segreto**) quando le sue specifiche non sono pubbliche

➡ Il formato **MDB** è proprietario (Microsoft) e **chiuso**;

Tratto da: <http://www.microsoft.com/interop/docs/officebinaryformats.mspix>



Microsoft Office Binary (doc, xls, ppt) File Formats

Published: February 15, 2008 | Updated: March 12, 2008

On June 30, 2008, Microsoft published additional specifications for the Word, Excel and PowerPoint binary file formats. |

La "standardizzazione"

- **Standard *de jure***: le specifiche sono definite o approvate da un organismo di standardizzazione (ISO, ANSI, ECMA, W3C, etc.) e quindi ha ottenuto un "riconoscimento ufficiale"

➡ Esempi di standard de jure sono l'**HTML** (W3C Recommendation e ISO 15445), l'**XML** (W3C Recommendation), l'**ODF** (ISO 26300), il **JPG** (ISO 10918-1), il **PNG** (ISO 15948), il **PDF** (ISO 32000) e alcuni suoi subset quali il PDF/A, il PDF/E e il PDF/X

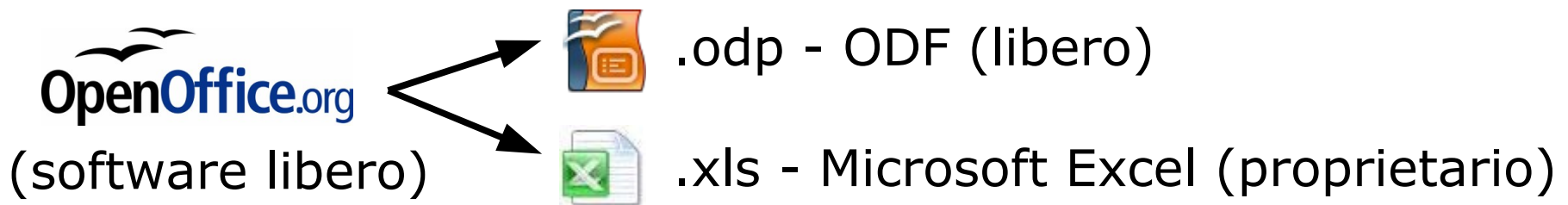
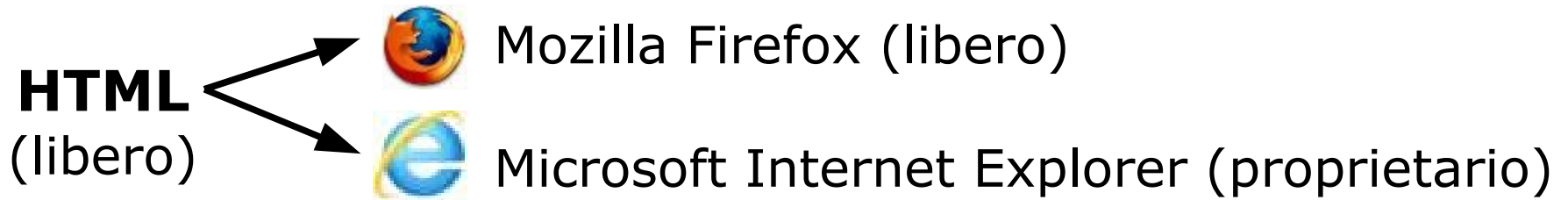
- **Standard *de facto***: le specifiche non sono state ratificate da nessun organismo di normazione ma è diventato, di fatto, uno standard grazie alla sua ampia diffusione. Può non essere libero da **restrizioni legali** per il suo utilizzo.

➡ Esempi di standard de facto sono il **DOC**, l'**XLS** e il **DWG**

Qual è il formato ideale?

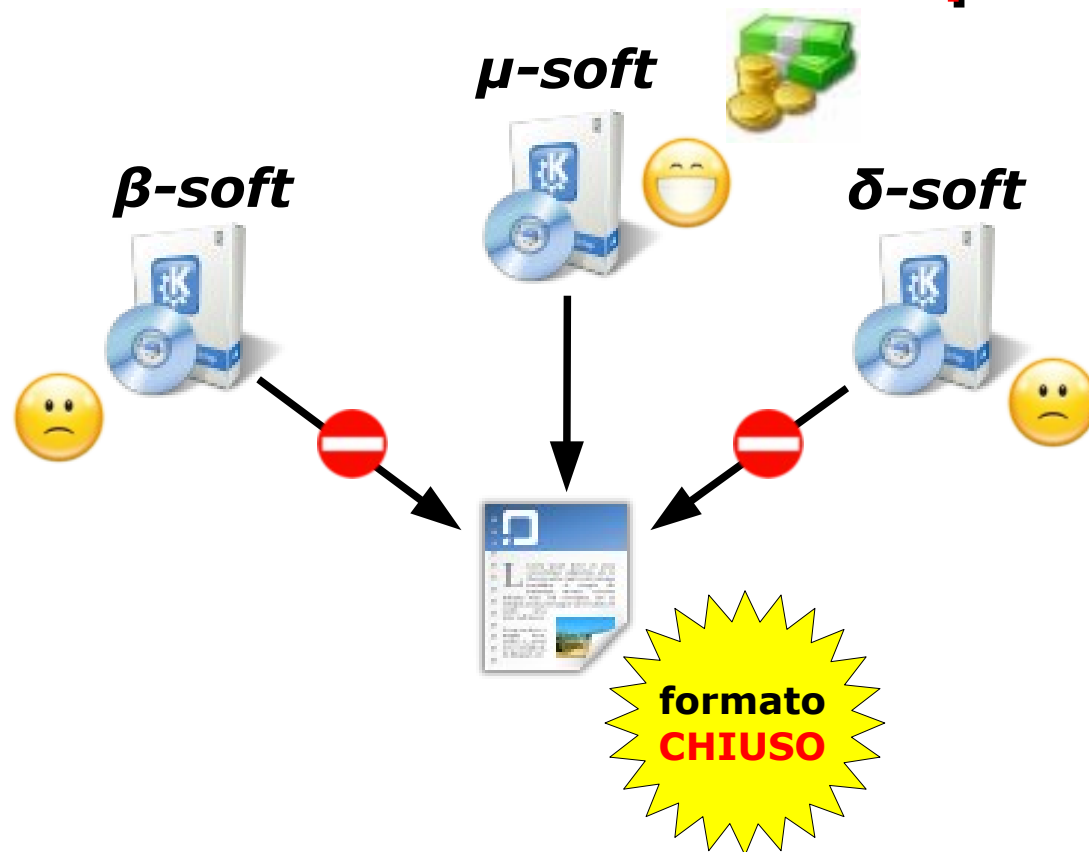
- ➔ Utilizzare formati **non proprietari** poiché non sono legati alla esistenza di una specifica azienda che ne detiene la proprietà e che potrebbe, in qualsiasi momento, modificarne le specifiche, renderle inaccessibili, o imporre restrizioni sul loro utilizzo
- ➔ Utilizzare dei formati **aperti** poiché sono i soli compatibili con un processo di conservazione digitale
- ➔ Utilizzare dei formati **standard**: gli standard *de jure* sono da preferire agli standard *de facto*, dal momento che solo il processo ufficiale di standardizzazione garantisce che non vi siano interessi di parte nella definizione ed implementazione di un formato
- ➔ Utilizzare dei formati con **ampia adozione** poiché l'ampia adozione è uno dei principali deterrenti contro i rischi legati all'obsolescenza

Centra con il software libero?



➡ Il software libero non è necessariamente collegato ai formati liberi: molte applicazioni di software proprietario fanno largo uso di formati liberi ed il software libero può a volte usare formati proprietari.

Formati chiusi: interoperabilità?



- ➡ L'informazione è intimamente legata ad 1 applicazione
- ➡ Viene inibito l'uso di prodotti concorrenti
- ➡ **Il controllo è del produttore del software e non nostro!**



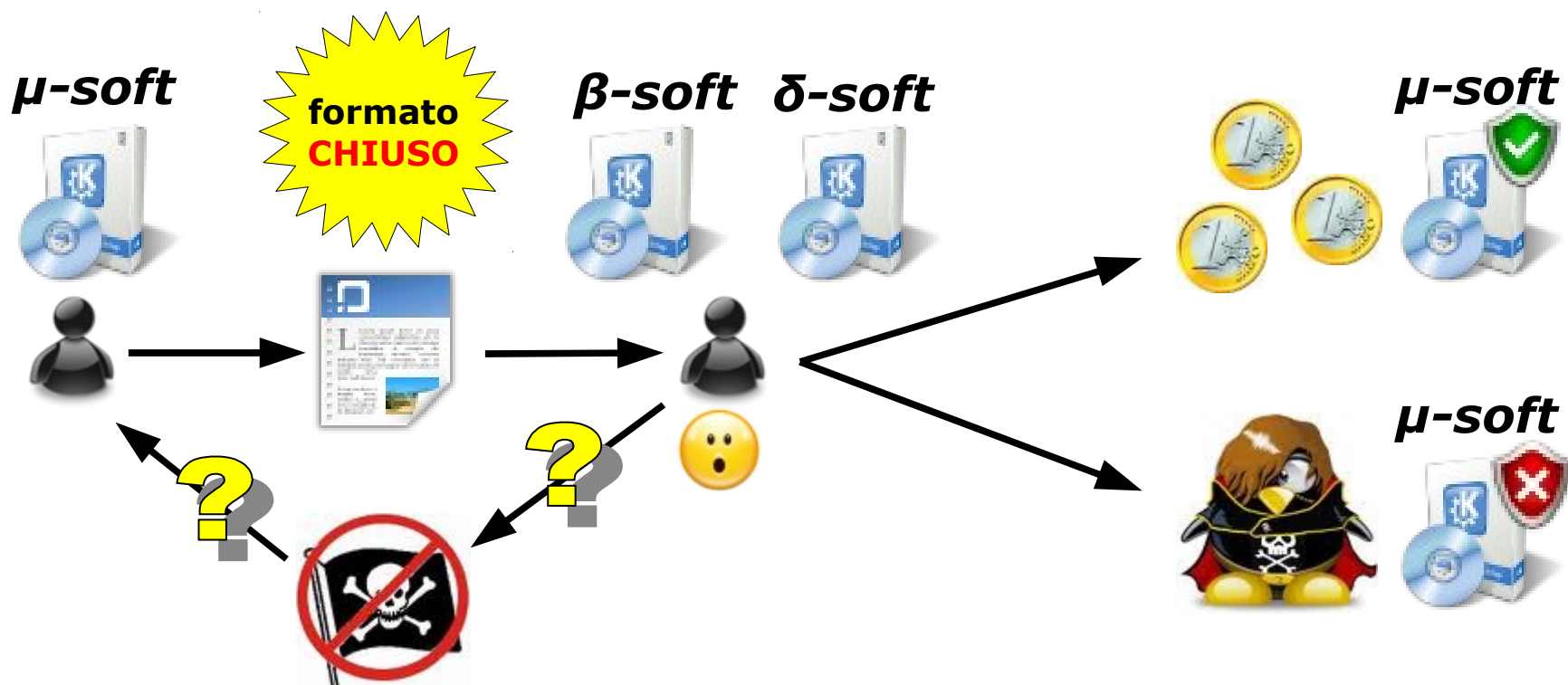
Produttori furbetti...



➔ Diversamente, potendo convertire da e verso altri formati, l'utente potrebbe trasformare il file in un formato aperto editabile e riconvertirlo in quello proprietario al termine del lavoro, **senza comprare il programma della casa madre.**

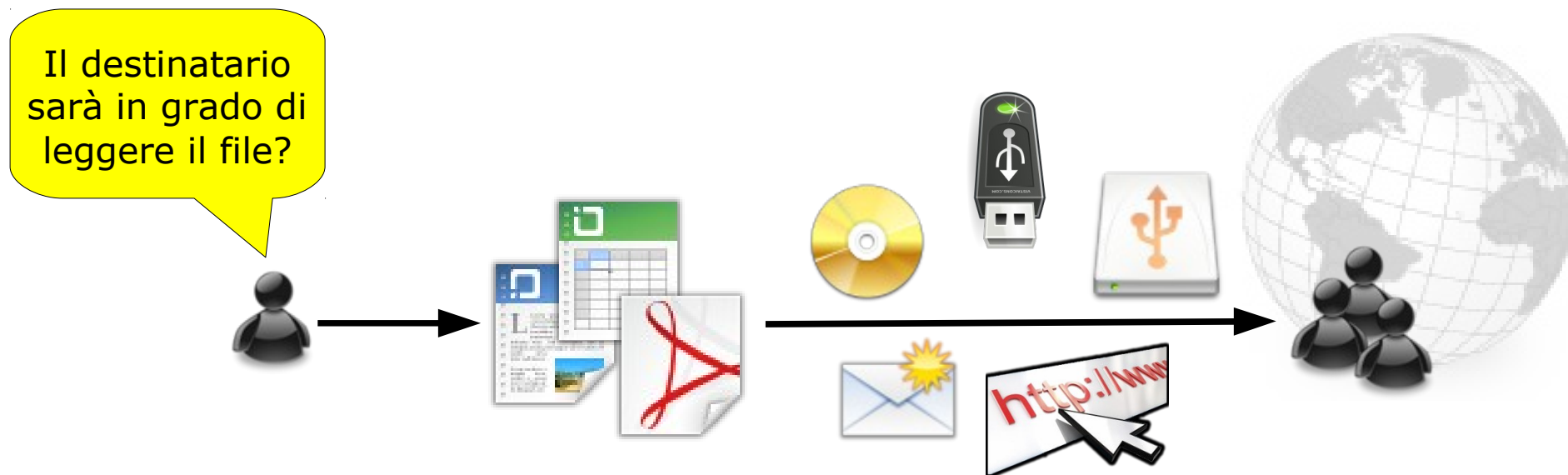


Non riesco a leggere il file!



➔ Utilizzare un formato documentale chiuso significa esercitare un'imposizione sulle persone con le quali si condividono le informazioni contenute nei file salvati in quel formato.

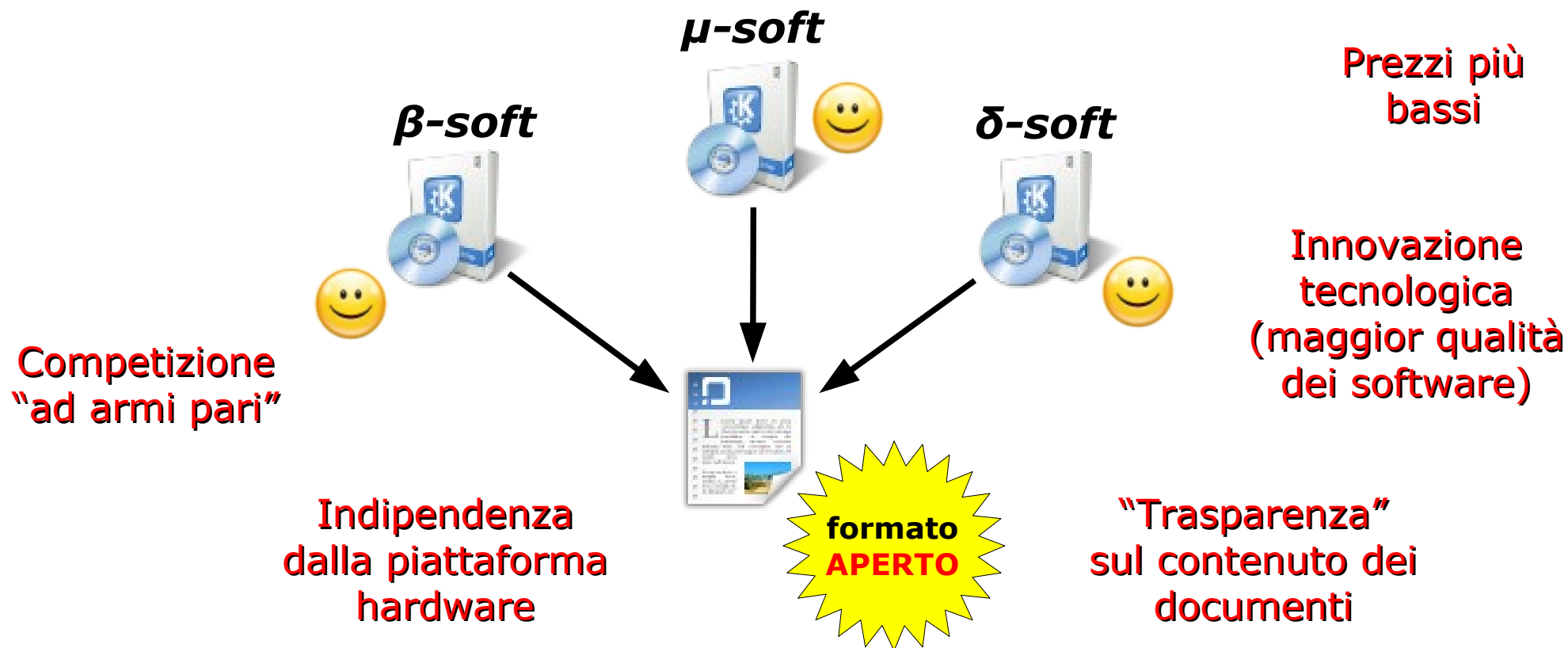
Scambio di files



- ➡ Se si fa un **uso locale** dei propri file (uso personale o con altri utenti con cui esista un implicito accordo sul tipo di programmi disponibili per aprire i file), **la differenza tra un formato aperto e un formato proprietario può passare inosservata...**
- ➡ Se l'**uso non è locale** (non sussiste un accordo con i destinatari sul tipo di programmi da utilizzare per aprire i file), **la distinzione tra formato aperto e formato proprietario è cruciale!**



Vantaggi dei formati aperti



- ➡ L'informazione può essere creata con varie applicazioni
- ➡ Viene incentivata la creazione di prodotti concorrenti
- ➡ **Il controllo è nostro e non dei produttori dei software**



Accesso ai dati

2010



*passano le ore
i giorni, gli anni,...*



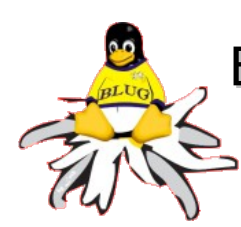
...2040...2050...

FORMATO CHIUSO



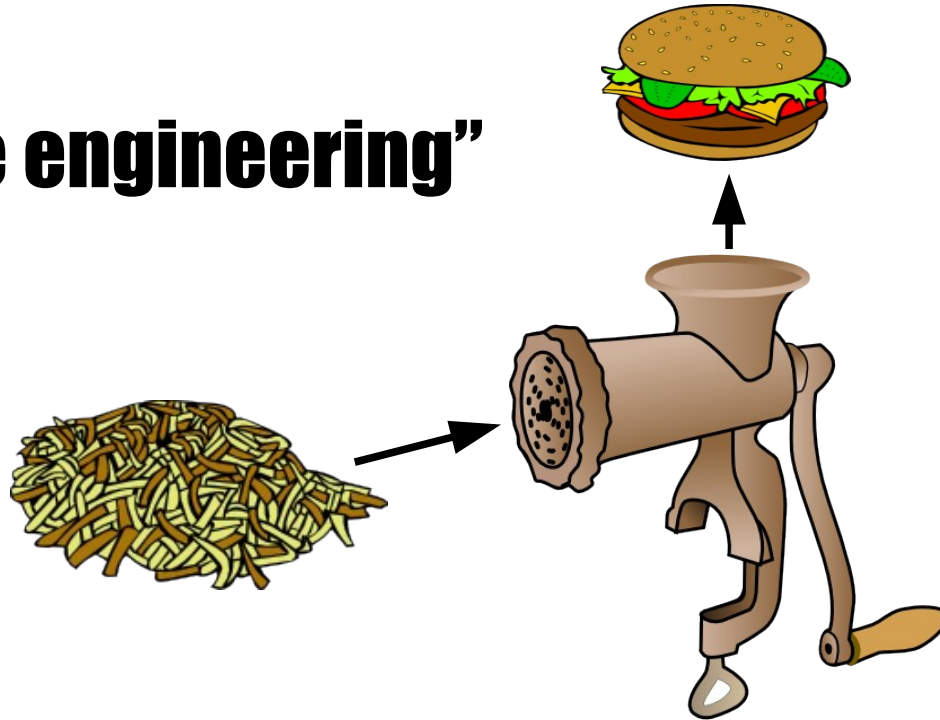
FORMATO APERTO

**È garantito l'accesso ai dati
nel lungo periodo senza
incertezza presente e futura
riguardo ai diritti legali o le
specifiche tecniche**



I filtri di conversione

“reverse engineering”



➡ I filtri di conversione per i formati proprietari **non possono mai essere perfetti**, sia per la segretezza di quei formati sia perché essi vengono **continuamente cambiati** proprio per evitare che siano disponibili filtri perfetti!



Esempi di formati

PROPRIETARI

.doc .xls .gif .bmp
.ppd .tif
.pub .dwg .psd
.mp3 .mp4 .divx
.avi .mov .wma
.flv .wmv

APERTI

.ods .odb .jpg
.odt .odp .png
.odg .ps .svg
.pdf .xml
.rtf .html .xhtml
.txt .7z .ogg
.zip .bz2



Ce ne sono molti altri!



Il formato ODF

ODT



documento
di testo

ODS



foglio
di calcolo

ODP



presentazione

ODG



grafica

ODB



database

(esistono molte altre estensioni)

- Il formato **OpenDocument (ODF)**, abbreviazione di "OASIS Open Document Format for Office Applications", è un formato aperto per file di documento per l'archiviazione e lo scambio di documenti per la produttività di ufficio come documenti di testo, fogli di calcolo, diagrammi, presentazioni etc.



Basato sul linguaggio XML



UNI CEI ISO/IEC IS 26300:2006



Esistono varie implementazioni sia open source (OpenOffice) che commerciali



E' l'opzione che al momento fornisce la migliore scelta per l'interoperabilità e la capacità di conservare nel tempo le informazioni



Il formato Office Open XML



documento
di testo



foglio
di calcolo



presentazione

- **Office Open XML (OOXML)** è un formato di file per la memorizzazione di documenti di testo, fogli di calcolo, diagrammi, presentazioni, etc. Il formato è stato sviluppato da Microsoft come formato di memorizzazione per i prodotti Office, con lo scopo di **contrastare l'adozione di ODF**. Nel 2008 è diventato (dopo non poche polemiche!) standard ISO.



Basato sul linguaggio XML



UNI CEI ISO/IEC IS 29500:2008



Consente la compatibilità all'indietro con tutti i documenti MS-Office



Al momento, esiste una sola implementazione su un'unica piattaforma



I documenti nella PA

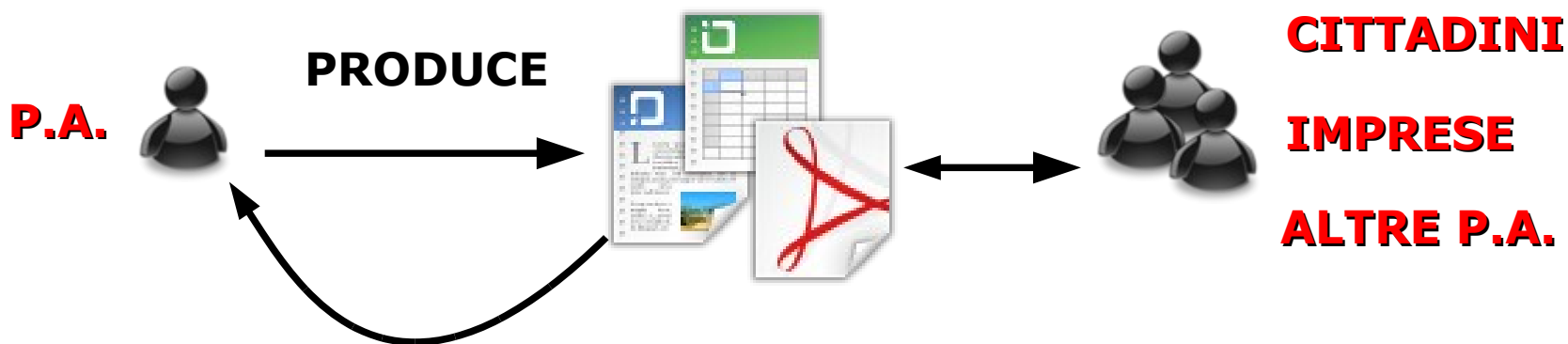
La PA nella scelta dei formati dei documenti deve:

- Assicurare interoperabilità e flessibilità
- Evitare di imporre decisioni tecnologiche ai cittadini
- Evitare vincoli nei confronti di particolari produttori
- Contenere e razionalizzare la spesa pubblica
- Assicurare l'accesso futuro ai dati



➡ Deve prediligere l'uso di Formati Standard Aperti!

DEMATERIALIZZAZIONE

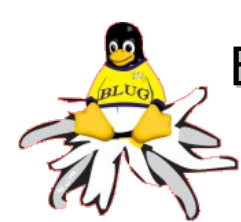


Tanto per cominciare...

- Pubblicare in rete (siti Internet) i documenti da ufficio solo utilizzando dei formati aperti
- Accettare file ODT da imprese e cittadini
- Accettare file ODT in gare, appalti, concorsi,...



Deve essere un
processo graduale!



Formati aperti nel mondo

- Le amministrazioni pubbliche ed i governi sono diventati progressivamente consapevoli sulle questioni dei formati
 - ➔ 2005: il **Massachusetts** usa ODT per i propri archivi pubblici
 - ➔ 2006: il governo del **Belgio** ha deciso di adottare il formato **ODT come standard per i documenti della Pubblica Amministrazione**
 - ➔ 2007: i governi del **Giappone** e del **Sud Africa** ha deciso di adottare il formato ODF come standard per lo scambio di documenti nella P.A.
 - ➔ 2009: Il governo della **Gran Bretagna** supporterà l'uso di ODF e garantirà l'uso per i documenti governativi
 - ➔ 2009: in **Norvegia**, a partire dal 1 gennaio 2009, **tutte le informazioni dovranno essere rese disponibili in formato ODF, PDF e HTML**
 - ➔ 2009: dal gennaio del 2010 tutte le P.A. della **Germania** dovranno essere in grado di ricevere, leggere e spedire documenti in formato ODF.
 - ➔ Argentina, Danimarca, Brasile, Perù, Spagna, India, Corea, Messico, Cina...

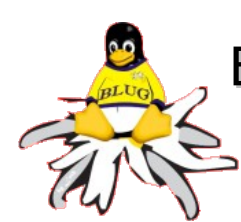
In Italia (1/4)

- Decreto legislativo 7 marzo 2005, n. 82 - "Codice dell'amministrazione digitale." (Art. 68)

➡ "acquisizione di *programmi informatici a codice sorgente aperto*"

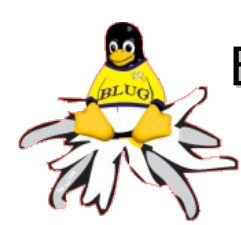
➡ "Le pubbliche amministrazioni nella predisposizione o nell'acquisizione dei programmi informatici, adottano *soluzioni informatiche che assicurino l'interoperabilità e la cooperazione applicativa*, secondo quanto previsto dal decreto legislativo 28 febbraio 2005, n. 42, e che *consentano la rappresentazione dei dati e documenti in più formati, di cui almeno uno di tipo aperto*, salvo che ricorrano peculiari ed eccezionali esigenze"

➡ "Per formato dei dati di tipo aperto si intende un formato dati reso pubblico e documentato esaustivamente"



In Italia (2/4)

- **Provincia Autonoma di Trento:** "Linee guida in materia di interoperabilità dei sistemi informativi e di software Open Source" - Delibera di Giunta Provinciale num. 1492 del 2005
 - ➡ *"nelle comunicazioni interne **privilegi l'adozione di formati di dati aperti e in quelle esterne garantisca l'adozione (pur non esclusiva) di formati dati aperti**"*
 - ➡ *"continui a promuovere e sostenga attivamente iniziative finalizzate all'alfabetizzazione, alla **diffusione e all'impiego consapevole di tecnologie informatiche che siano basate su strumenti e standard aperti**"*
 - ➡ *"progetti di formazione informatica promuovano e incentivino il completamento dell'offerta formativa anche sull'uso di **ambienti aperti ed interoperabili**"*



In Italia (3/4)

- **Comune di Trento:** “Estendere ulteriormente l'utilizzo e la promozione del software libero ed open source e dei formati aperti, nonché la partecipazione a progetti di software libero ed open source da parte del comune di Trento” - Delibera del Consiglio Comunale num. 90556 del 2010

- ➡ “*proseguire ulteriormente nell'introduzione e nell'uso di software FLOSS a tutti i livelli all'interno dell'Amministrazione comunale*”
- ➡ “*adottare il formato di dati standard aperto (ad esempio ODF Open Document Format e PDF) quali formati per i documenti prodotti e gestiti all'interno e verso l'esterno dell'Amministrazione comunale [...] salvo ricorrano comprovate esigenze ad impedirlo*”
- ➡ “*a mettere a disposizione dei cittadini i moduli elettronici per accedere ai servizi dell'Amministrazione comunale in formato PDF e/o ODF*”



In Italia (4/4)

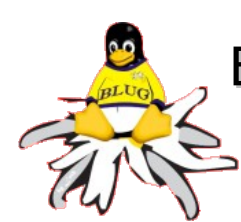
- **Comune di Ferrara:** Risoluzione collegata al Bilancio 2005 del Comune di Ferrara. Presentata da Barbara Diolaiti, Capogruppo dei Verdi del Comune di Ferrara. Approvata all'unanimità in sede di approvazione del bilancio 2005.



"promuovere l'impiego di formati di dati standard, aperti e documentati, in quanto di estrema importanza e flessibilità per mantenere la piena compatibilità con futuri cambiamenti tecnici e avendo questo notevole impatto nel caso di realizzazione di programmi ad hoc per la PA senza nessuna conseguenza sulla facilità d'uso"



"indirizzare i dipendenti all'impiego esclusivo nella spedizione di documenti di formati di salvataggio/interscambio leggibili da qualsiasi altro programma di scrittura e non semplicemente, ad esempio, documenti nell'ultima versione disponibile del programma di word processing di cui dispongono"



In Veneto

- **Regione Veneto:** "Norme in materia di pluralismo informatico, diffusione del riuso e adozioni di formati per documenti digitali aperti e standard nella società dell'informazione del Veneto" - Legge regionale 14 novembre 2008, n. 19

➡ "Al fine di garantire la più ampia libertà di accesso all'informazione pubblica attraverso il pluralismo informatico, *la Regione del Veneto promuove e incentiva l'uso di formati digitali aperti e non proprietari, per la documentazione elettronica e per le basi di dati*"

➡ "persegue, nelle forme e nei limiti previsti dalla disciplina vigente, la rimozione delle *barriere all'accesso alle informazioni*, mediante l'*adozione di formati standard per la predisposizione dei programmi* e delle piattaforme e con l'*impiego ottimale dei software a codice sorgente aperto* e chiuso"

➡ "impiega *almeno un formato di dati di tipo aperto*, ai sensi dell'articolo 68, comma 3, del decreto legislativo 7 marzo 2005, n. 82, "Codice dell'amministrazione digitale" *nelle operazioni di memorizzazione e pubblicazione dei propri documenti, al fine di garantirne la disponibilità e fruibilità.*"



In provincia di Belluno

- **Comune di Belluno:** "Linee di indirizzo su open source" - Delibera di Giunta Comunale n°2 del 13/01/2005
 - ➡ "... di avviare, compatibilmente con le esigenze e le necessità tecniche, soluzioni open source quali: uso del sistema operativo Linux nelle stazioni di lavoro degli utenti; adozione di programmi di produttività individuale open nelle stazioni utente..."
- **Comune di Feltre:** "Linee di indirizzo per l'adozione di soluzioni software open source" - Delibera di Giunta Comunale (2004)
 - ➡ "... diffondere al proprio interno l'utilizzazione di software open source [omissis] sui p.c. degli utenti, sia per quanto riguarda il Sistema Operativo (Linux) che per i programmi di produttività individuale (Open Office, etc.)..."





Conclusioni

- ➔ Per garantire la conservazione nel tempo dei contenuti digitali occorre iniziare con la corretta scelta del formato. In particolare occorre scegliere formati che siano:
 - non proprietari
 - aperti
 - standard *de jure*
- ➔ Adottare formati indipendenti da qualsiasi singolo programma software (sia esso proprietario o libero)
- ➔ Essere sempre consapevoli di quale sia il formato di file con il quale si sta lavorando
- ➔ Correre ai ripari nel caso in cui ci si accorge che il formato che si sta utilizzando è proprietario



Bibliografia

- Formato aperto (http://it.wikipedia.org/wiki/Formato_aperto)
- Openformats.org (<http://www.openformats.org>)
- OpenDocument (<http://it.wikipedia.org/wiki/OpenDocument>)
- Digital Standards Organization (<http://www.digistan.org>)
- OpenDocument Fellowship (<http://opendocumentfellowship.com>)
- Marco Fioretti - "La via maestra per il pluralismo informatico: formati aperti per innovazione, libero mercato e servizi pubblici più efficienti" - 30/03/2009
- Stefano Allegrezza - "La produzione di documenti informatici: requisiti dei formati elettronici" - Macerato 8/5/2009
- Domenico Squillace - "Formati Standard Aperti e la Pubblica Amministrazione" - UNINFO 26/10/2007
- Paolo Dongilli . "Formati aperti, il Software Libero non è solo software" - LinuxDay2008 Trento
- Open Design Alliance (<http://www.opendesign.com>)
- Associazione per il Software Libero (<http://softwarelibero.it>)
- Digital Preservation Coalition (<http://www.dpconline.org>)
- ODF Alliance (<http://www.odfalliance.org>)
- CASPAR (<http://www.casparpreserves.eu>)



Grazie per l'attenzione!

Copyright © 2010 Mauro Barattin

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation.