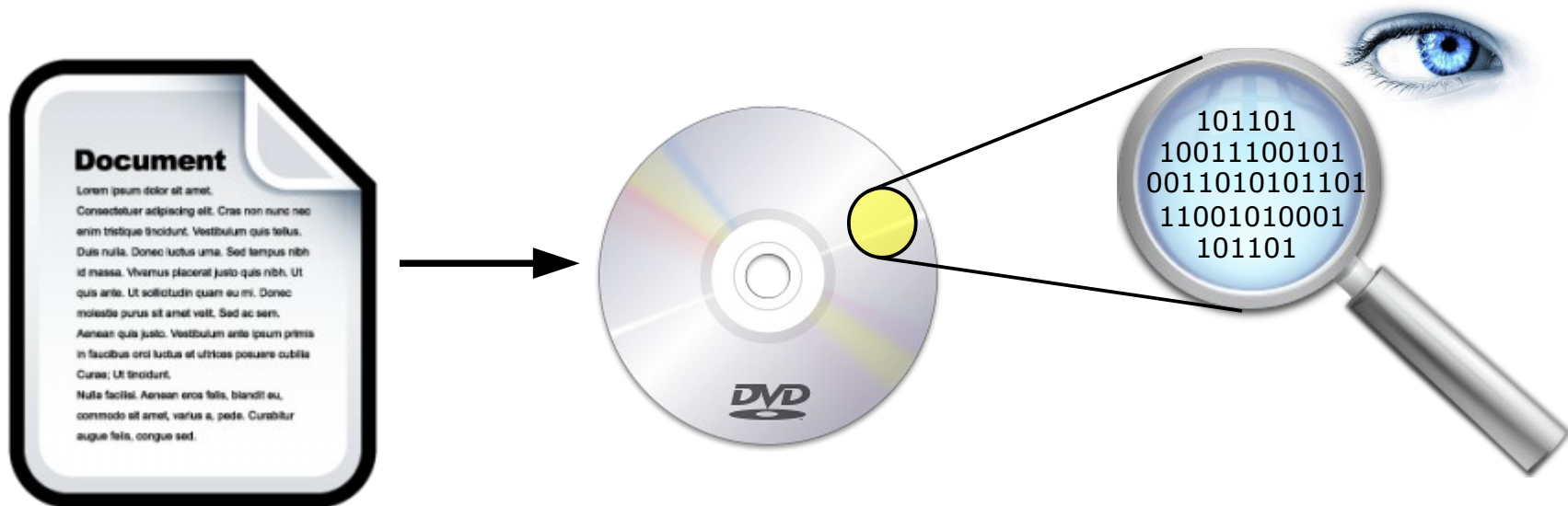
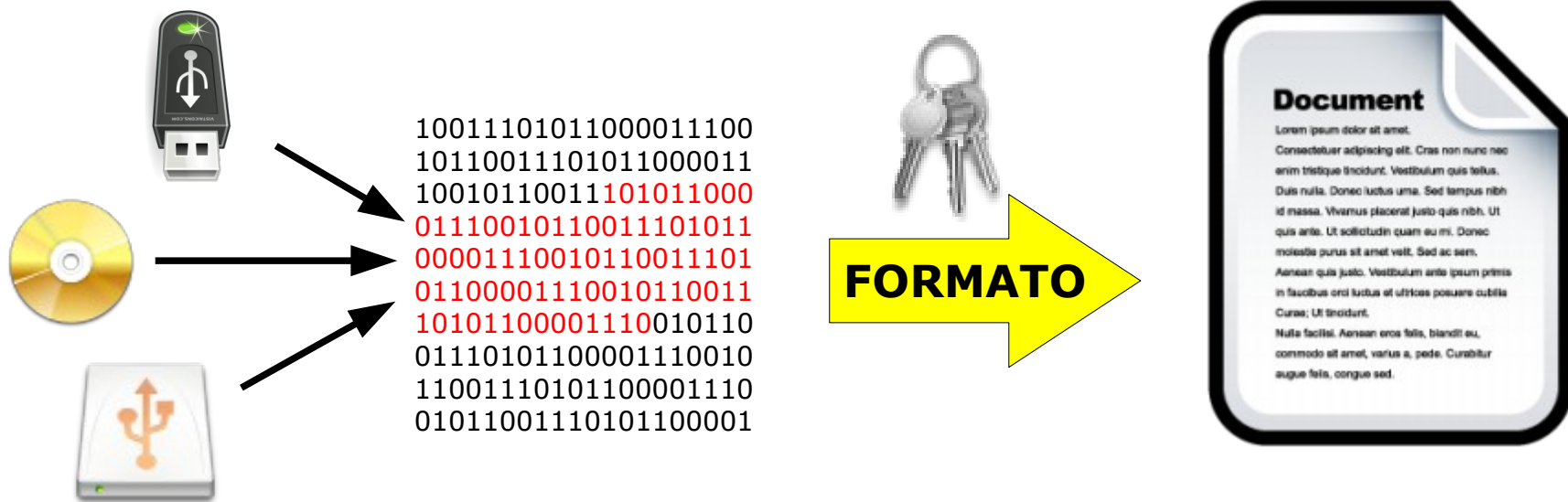


# Come sono fatti i files?



- ➔ 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?



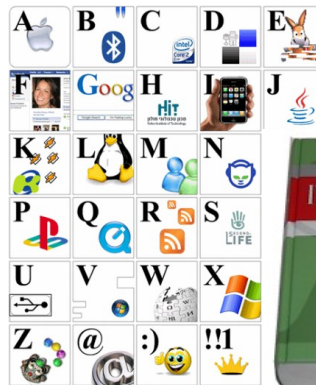
- ➔ Perché i documenti possano essere aperti o modificati tramite i vostri programmi, è necessario che abbiano un **formato**.
- ➔ Un formato è l'insieme di **regole** che, a partire dalla sequenza di bit che costituisce un oggetto digitale, permettono di riprodurre mediante un software il relativo documento,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

# Penne e alfabeti...

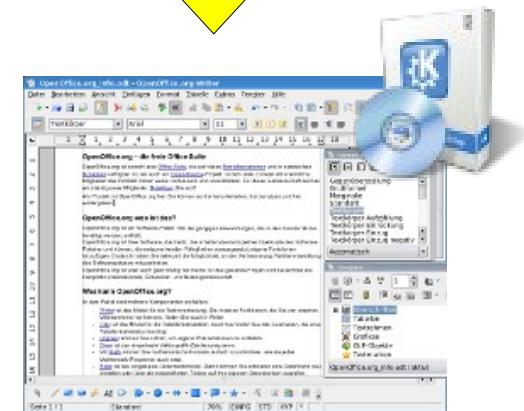
**SUPPORTO FISICO**



**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.mspx>



## 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.

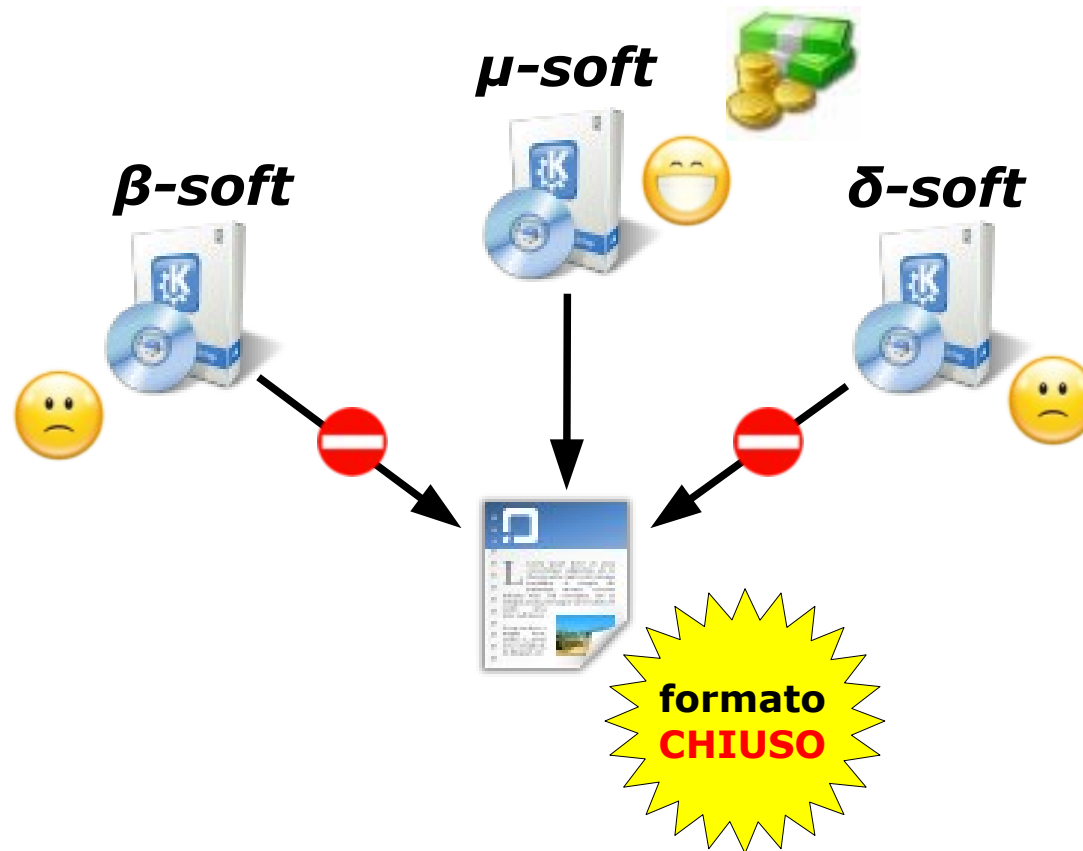
➡ Esempi di standard de facto sono il **DOC**, l'**XLS** e il **DWG**

# 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

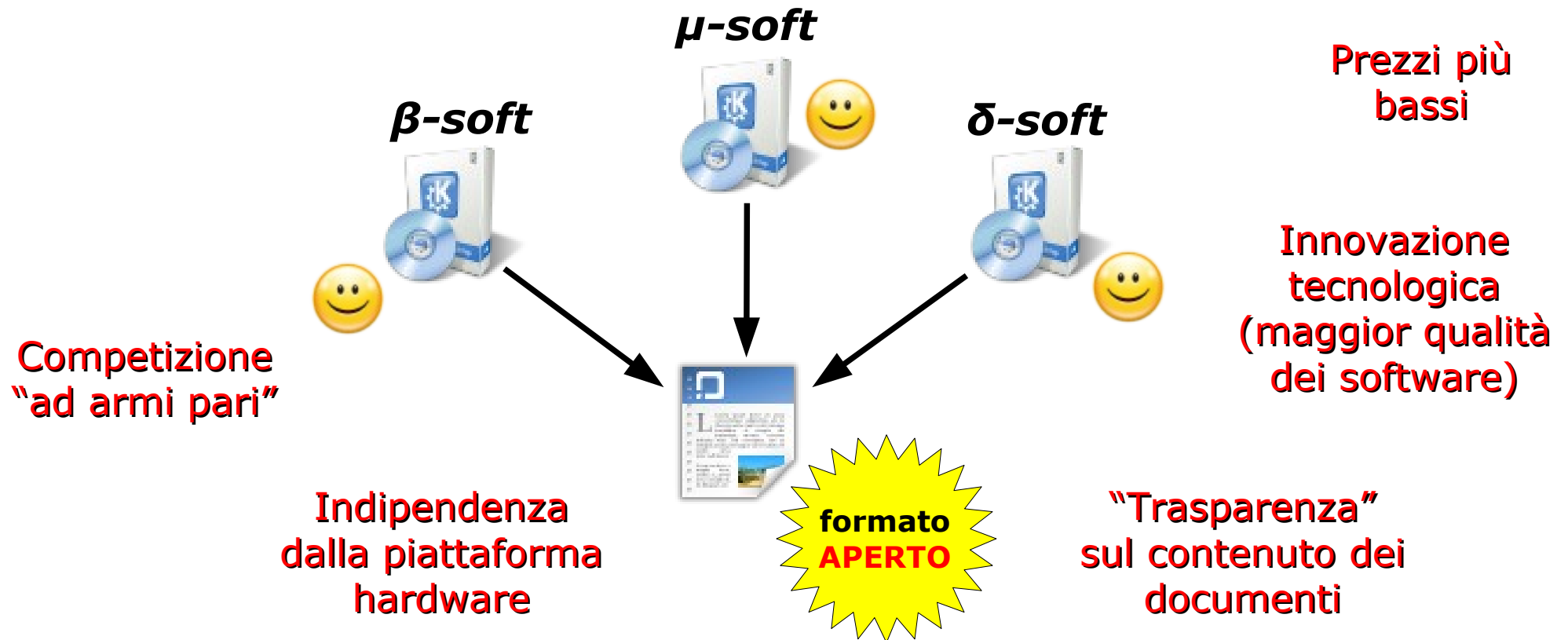


# Formati chiusi: interoperabilità?



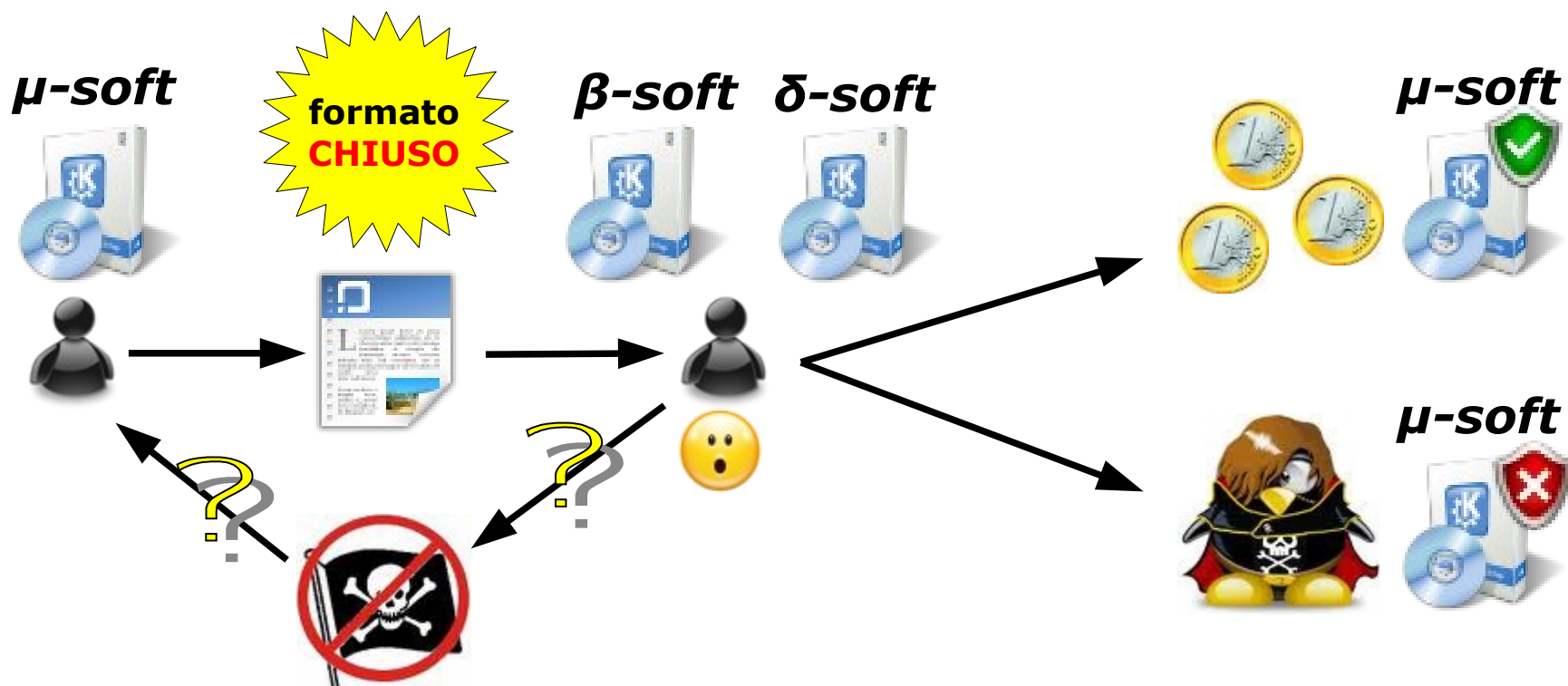
- ➔ L'informazione è intimamente legata ad 1 applicazione
- ➔ Viene inibito l'uso di prodotti concorrenti
- ➔ **Il controllo è del produttore del software e non nostro!**

# Formati aperti: interoperabilità?



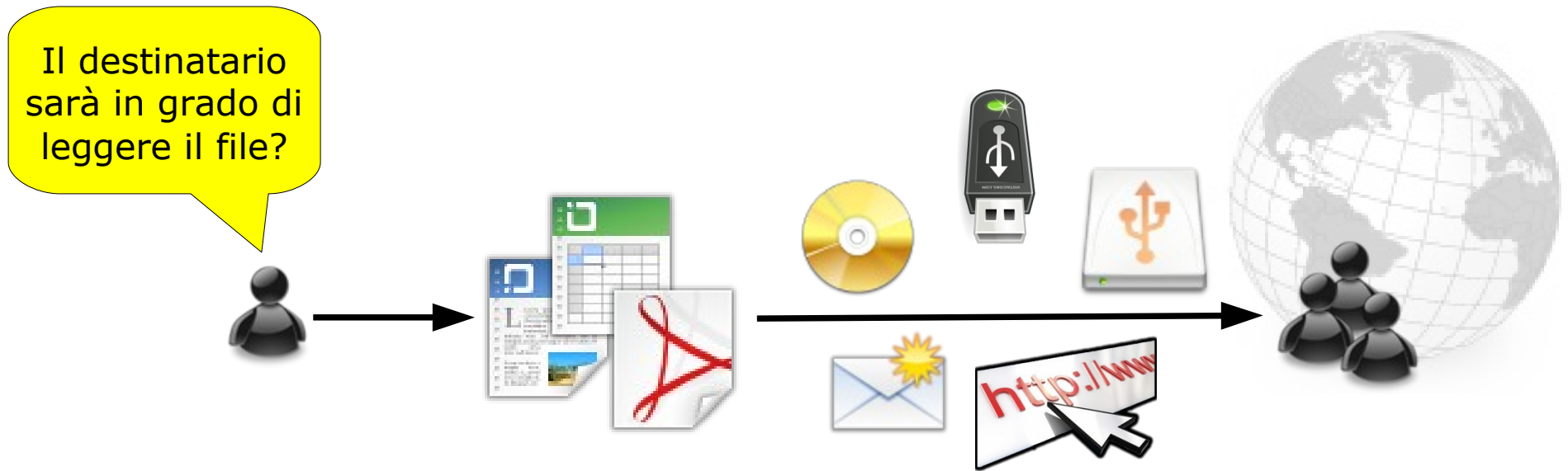
- ➔ L'informazione può essere creata con varie applicazioni
- ➔ Viene incentivata la creazione di prodotti concorrenti
- ➔ **Il controllo è nostro e non dei produttori dei software**

# 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!**

# Accesso ai dati

**2012**



*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**

# 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  
.rtf .html .xml  
.txt .7z .xhtml  
.zip .bz2 .ogg



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

# 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