



Università degli Studi di Foggia

Corso di aggiornamento INPS “Valore P.A.”

Gestione del Documento Informatico

Anno 2022

Modulo I

Formati documentali

Prof. Crescenzo Gallo

crescenzo.gallo@unifg.it

<http://www.crescenziogallo.it/unifg/>

Professore Aggregato di Sistemi di Elaborazione delle Informazioni

Università degli Studi di Foggia

Dipartimento di Medicina Clinica e Sperimentale



Acquisire familiarità con il concetto di (formato di) documento digitale



Il modulo si pone l'obiettivo di chiarire e trasmettere il concetto di **documento informatico** e dei relativi formati, proprietari e — soprattutto — *aperti*, insieme al relativo software per la produzione, gestione e interoperabilità dei documenti digitali.



- Documento analogico e digitale
- Formati aperti dei documenti digitali e loro standardizzazione
- Il software open source per la produzione e gestione nel tempo dei documenti informatici
- Lo standard XML per l'interoperabilità



Dal documento analogico al documento digitale



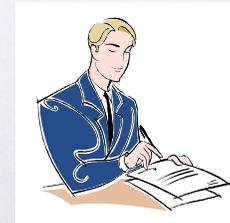
Le sfide delle nuove tecnologie costringono a ripensare il concetto di **documento** (e della sua relativa archiviazione), non più considerabile esclusivamente nella sua tradizionale forma cartacea.

L'era attuale della società dell'informazione post “gutenberghiana” impone in maniera sempre più pressante la trasformazione del supporto documentale dalla sua forma tradizionale analogica alla moderna (e ormai imprenscindibile) forma *digitale*.

Secondo il CAD il **documento analogico** è la *rappresentazione non informatica di atti, fatti o dati giuridicamente rilevanti* (art.1 lettera p-bis)

Il documento analogico pertanto:

- è un oggetto *materiale*
- è *unico* (originale diverso dalla copia)
- viene sottoscritto con *firma autografa*



Secondo il CAD il **documento informatico** è la *rappresentazione informatica di atti, fatti o dati giuridicamente rilevanti* (art.1 lettera p)

Il documento informatico (digitale) pertanto:

- è un oggetto (quasi) immateriale
- il duplicato è identico all'originale
- può essere sottoscritto con firma digitale



Tutti i documenti presenti in un archivio contengono *informazioni*.

L'informazione è caratterizzata dai seguenti aspetti:

- **emittente** (da chi/cosa origina l'informazione)
- **ricevente** (a chi/cosa è destinata l'informazione)
- **utilità** (ragione d'essere dell'informazione)
- **canale** (mezzo fisico per la trasmissione dell'informazione)
- **linguaggio** (forma condivisa in cui l'informazione è espressa)
- **supporto** (fisico di registrazione dell'informazione)

Analizzando in dettaglio la metamorfosi subita dai supporti documentali si può parlare di era *analogica, digitale e post-digitale*.

La prima è stata connotata da supporti fisici come: carta, supporti magnetici, pellicole foto/ cinematografiche, etc.



I formati digitali comprendono memorie contenenti informazioni tradotte in codice binario e conservate su supporti fisici come dischi, CD, DVD, pendrive USB.

Nell'era post-digitale il supporto fisico delle informazioni binarie diventa invisibile e intangibile, trasformandosi in una posizione (non nota, remota) non accessibile direttamente in locale (*cloud*).



Queste trasformazioni sintetizzano la transizione dalla forma materiale all'attuale forma *immateriale* di gestione e fruizione dei documenti.

L'evoluzione dei documenti e dei relativi supporti di memoria può essere infatti vista con la metamorfosi delle *transizioni di stato* della materia.

Nello stadio analogico, le informazioni erano incorporate nel supporto fisico, creando un documento “**solido**”: il documento era contemporaneamente l'informazione e il supporto fisico.

Con la diffusione di supporti digitali a basso costo (CD, DVD, HD) l'informazione divenne più indipendente dal supporto di memoria, più **liquida**, ma ancora tangibile grazie ai supporti su cui era registrata.

Ora stiamo assistendo al processo di **evaporazione** dei dati, nascosti in posizioni remote. Le registrazioni digitali fluiscono attraverso server distanti (*cloud*) e qualsiasi connessione diretta tra l'informazione e il supporto fisico è definitivamente cessata.



I formati documentali proprietary e aperti



Formati proprietari



Con *formato proprietario (chiuso)* si intende qualsiasi formato di file di cui non siano liberamente fruibili le specifiche tecniche complete, al contrario di formato aperto.

Molti formati proprietari sono nati come forma di monopolio per sfavorire la concorrenza, ad esempio impedendo il *reverse engineering* mediante la crittografia dei dati, rilasciando le specifiche in modo incompleto o coprendole con un brevetto software limitativo.



Formati proprietari

- Un formato proprietario richiede quindi l'acquisto di un programma specifico capace di leggere l'informazione contenuta nei file.
- Lo scambio di file in formato proprietario presuppone cioè che ogni destinatario del file posseda il software necessario per leggere ed eventualmente modificare questo formato.





Alcuni formati proprietari

- Documenti di testo formattato MS Word (DOC, DOCX)
- Foglio elettronico MS Excel (XLS, XLSX)
- Presentazioni MS PowerPoint (PPT, PPTX)
- Immagini bitmap ad alta qualità (BMP, TIF)
- Windows Media Audio (WMA)



Formati aperti



Un **formato aperto** è una specifica pubblica per la descrizione e l'archiviazione di dati digitali, solitamente gestita da un ente di standardizzazione non proprietario e libera da restrizioni legali per il suo utilizzo.

Un formato aperto può essere implementato da un software proprietario, open source o da software libero, ciascuno con le proprie modalità di licenza.



Alcuni Formati Aperti

- TXT (testo semplice)
- OASIS Open Document Format for Office Applications (documenti office)
- RTF (documenti di testo)
- PS, PDF (descrizione di documenti, standard ISO 32000)
- XML (un linguaggio di markup per l'interoperabilità)
- JPEG, PNG (immagini)



Differenze tra formato aperto e proprietario

- La differenza tra un formato aperto e un formato proprietario può passare inosservata finché si fa un uso locale dei propri file.
- Per uso locale intendiamo un uso personale o con altri utenti con cui esista un implicito accordo sul tipo di programmi disponibili per aprire i file.



Differenze tra formato aperto e proprietario

- Non appena l'uso non è più locale e possiamo parlare di scambio di file, la distinzione tra formato aperto e formato proprietario diventa *cruciale*.
- Nel caso di un uso non locale è preferibile utilizzare formati aperti, in quanto sono leggibili da qualsiasi tipo di editor.



Il formato aperto dei documenti (*Open Document Format*, ODF), basato su uno standard (XML) pubblicamente accessibile ed implementabile da chiunque senza restrizioni, è un formato per documenti libero, standardizzato dall'ISO/IEC nel 2005 ed attualmente gestito dall'Organizzazione per la Promozione degli Standard per le Informazioni Strutturate (*Organization for the Advancement of Structured Information Standards*, OASIS)



OpenDocument Format

standardised by OASIS



I principali documenti di questo tipo sono caratterizzati dai seguenti suffissi:

- ▶ **ODT**: Documento di testo (corrisponde ai suffissi DOC e DOCX dei documenti di MS Word).
- ▶ **ODS**: Foglio di calcolo (corrisponde ai suffissi XLS e XLSX dei documenti di MS Excel).
- ▶ **ODB**: Database (corrisponde ai suffissi MDB e ACCDB dei documenti di MS Access).
- ▶ **ODP**: Presentazione (corrisponde ai suffissi PPT e PPTX dei documenti di MS PowerPoint).



Il seguente è l'elenco completo delle **estensioni dei formati open** e relativi **tipi MIME** (*Multipurpose Internet Mail Extension*, uno standard Internet che definisce il formato degli allegati di posta elettronica):

Tipo di documento	Estensione	Tipo MIME
Testo	.odt	application/vnd.oasis.opendocument.text
Database	.odb	application/vnd.oasis.opendocument.database
Foglio di calcolo	.ods	application/vnd.oasis.opendocument.spreadsheet
Presentazione	.odp	application/vnd.oasis.opendocument.presentation
Disegno	.odg	application/vnd.oasis.opendocument.graphics
Grafico	.odc	application/vnd.oasis.opendocument.chart
Formula	.odf	application/vnd.oasis.opendocument.formula
Immagine	.odi	application/vnd.oasis.opendocument.image
Modello di documento	.odm	application/vnd.oasis.opendocument.text-master



- Ampiamente utilizzato in tutto il mondo, l'ODF è supportato da una varietà di programmi, incluse le suite open source *LibreOffice* e *OpenOffice*.
- Gli utenti possono leggere e scrivere file OpenDocument senza dover aderire a licenze proprietarie, e i programmatori sono liberi di scrivere applicazioni che supportino l'ODF senza il timore di infrangere brevetti o licenze di varia natura.





- Le **pubbliche amministrazioni**, le aziende, le biblioteche possono utilizzare l'ODF per assicurare che documenti critici possano essere letti negli anni a venire, senza essere costretti a pagare per mantenere aggiornati pacchetti applicativi proprietari di gestione dei documenti.
- L'utilizzo di formati liberi è una delle cose più semplici e più importanti che si possa fare per difendere la libertà di informazione (*freedom of information*).





- Non è quindi più necessario usare suite proprietarie per la gestione dei documenti, con i problemi di compatibilità dovuti ai formati proprietari: grazie a ODF non è mai stato così semplice produrre, distribuire e utilizzare documenti.
- L'Open Document Format fornisce una soluzione globale per lo scambio di documenti affidabile e semplice tra differenti prodotti e dispositivi, senza dover dipendere da un singolo prodotto o venditore.
- Esso è ampiamente supportato dai principali fornitori come *Microsoft*, *IBM*, *Google* nonché da progetti open source come *OpenOffice* (<https://www.openoffice.org/>), *LibreOffice* (<https://www.libreoffice.org/>), *Calligra* (<https://www.calligra.org/>), *GnomeOffice* (<http://live.gnome.org/GnomeOffice>).





- Molti fornitori di applicazioni per l'ufficio sono scomparsi dal mercato, e con essi sono “svaniti” i loro prodotti e la possibilità di accedere ai tantissimi documenti creati per mezzo delle loro suite applicative.



- Questo rende indispensabile l'adozione di standard internazionali per assicurare la futura leggibilità (libertà di lettura) dei documenti negli anni a venire.



- Una interessante e utile caratteristica offerta da ODF consiste nell'essere basato sugli **standard web** laddove possibile.
- Ciò consente ad esempio di aprire documenti da un normale browser come si apre una pagina web: senza plugin né altro software installato sul dispositivo, a differenza di molte suite applicative proprietarie.
- Nessun altro formato per documenti d'ufficio offre tale **flessibilità!**





Software libero e formati aperti



L'esigenza di un **software libero** emerse circa trent'anni fa quando Richard Marshall Stallman lasciò il Massachusetts Institute of Technology (MIT) e fondò nel 1985 la *Free Software Foundation* (FSF, fondazione sul software libero, <https://www.fsf.org/>).

Dietro tale termine si cela una potente idea: quella di *promuovere la libertà universale di studiare, distribuire, creare e modificare il software* secondo la cosiddetta modalità **copyleft** (in inglese letteralmente “libertà di copia”, contrapposto a copyright = “diritto d'autore”) applicata attraverso la licenza GNU General Public License della FSF.

Il termine “free”, che nella lingua inglese ha il doppio significato di *libero* e *gratuito*, è qui inteso nell’accezione di libertà e non di prezzo.

Il free software può anche essere venduto, ma permette comunque di capire le modalità di lavoro del programma e di adattarlo alle proprie esigenze.

Esso:

- ✓ offre la possibilità di **redistribuire** le copie;
- ✓ aiuta a **migliorare** il programma stesso;
- ✓ estende le **possibilità di distribuire** agli altri il contributo del proprio miglioramento;
- ✓ consente il **lavoro in collaborazione** senza vincoli.

- Alcuni importanti risultati di questa filosofia sono stati ad es. la rete Internet, il server web Apache, il pacchetto LibreOffice.



- Per la loro realizzazione è stato fondamentale GNU (<https://www.gnu.org>), un sistema operativo nato grazie a Stallman e distribuito come software libero (GNU/Linux) che rispetta la libertà degli utenti, in contrapposizione a sistemi operativi proprietari come Windows.



- Accanto al concetto di *software libero* vi è quello di *open source*.
- Il primo ha come obiettivo la libertà degli utenti (informatici) e ritiene che un programma non libero costituisca un'inutile e dannosa limitazione.
- L'*open source*, invece, non considera l'esistenza di programmi non liberi come ingiustizia verso gli utenti, credendo si tratti esclusivamente di vantaggi o svantaggi di tipo pratico.

- Mentre in inglese, come sappiamo, “free” significa sia “libero” che “gratuito”, in italiano il problema non esiste e si parla sempre di “software libero”.
- Per evitare ambiguità, da un po’ di tempo è invalso l’uso del termine francese *libre* che chiarisce (nella lingua inglese) il senso da attribuire alla parola free.
- Analogamente è stato coniato il termine “FLOSS”, acronimo di “Free/Libre and Open Source Software,” (software libero ed open source), per evitare di esprimere una preferenza tra i due movimenti.

Quando si parla di *software libero* si intende che gli utenti di uno specifico programma godono di quattro **libertà essenziali**:

- ▶ Libertà di eseguire il programma come si desidera, per qualsiasi scopo (*libertà 0*).
- ▶ Libertà di studiare come funziona il programma e adattarlo alle proprie necessità (*libertà 1*). L'accesso al codice sorgente ne è un prerequisito.
- ▶ Libertà di ridistribuire copie in modo da aiutare il prossimo (*libertà 2*).
- ▶ Libertà di migliorare il programma e distribuirne pubblicamente i miglioramenti, in modo tale che tutta la comunità ne tragga beneficio (*libertà 3*).

- Molti inoltre sostengono che il software libero sia preferibile al software proprietario dal punto di vista della **qualità**, **manutenibilità** e **sicurezza**.
- I modelli del software libero, più vicini all'artigianato che a quelli della grande industria, rendono possibile uno sviluppo più equo e armonico, con le stesse opportunità per le grandi e piccole imprese, per i Paesi ricchi ed evoluti e quelli in via di sviluppo.

- Quando Stallman sviluppò il sistema operativo GNU(*) dopo l'uscita dal MIT impose per esso un requisito fondamentale: essere open source, cioè disponibile nel formato “sorgente”, nella forma originaria nel quale il programma è stato scritto.
- Il software sorgente (*source*), prima di essere utilizzato, deve essere “compilato”, cioè tradotto nel codice della macchina (codice eseguibile, una sequenza binaria di 0 ed 1 da essa comprensibile).

(*) GNU sta per Gnu is Not Unix, una definizione ricorsiva cara alla comunità degli hacker.



- Infatti il software proprietario (ad es. Microsoft Office) viene generalmente venduto in **formato eseguibile**, per rendere praticamente impossibile la sua interpretazione e la sua modifica in funzione delle esigenze del suo utilizzatore.
- Viceversa, i programmi *open source*, proprio perché disponibili in **forma leggibile da chiunque** abbia le capacità tecniche necessarie, sono appunto “aperti” nel senso che possono essere letti, corretti e trasformati in funzione di specifiche esigenze.

- La FSF è divenuta progressivamente un **punto di riferimento** per gli ideatori di software, nonché un'istituzione di *garanzia* della qualità di un prodotto e di *protezione* dei diritti del software libero.
- Essa non si oppone al sistema legislativo del copyright, ma intende porre un limite alla tendenza giurisprudenziale della protezione di tutto ciò che viene creato nell'ambito dell'industria informatica (brevetti e diritti d'autore) e promuovere un progetto etico e un nuovo modo di lavorare.

Le ragioni e i vantaggi connessi alla libertà del software richiamano i *principi dell'etica hacker* e si possono così sintetizzare.

1. Il diritto alla libera circolazione del software e alla sua duplicazione

Questo diritto rimanda al principio generale che tutta l'informazione deve essere libera, con grande beneficio per il sistema sociale nel suo complesso.

Esso dà all'utilizzatore l'opportunità di interagire con il prodotto e di controllarlo, e aiuta a stimolare lo sviluppo della conoscenza e a diffonderla.

In senso ampio, quindi, può favorire il processo di partecipazione contribuendo a formare un membro della comunità o un cittadino più preparato.

2. La **rivoluzione digitale** e la **diffusione dei computer**, rendendo più semplice lo scambio di informazioni, possono apportare un beneficio generale.
3. Il modo migliore per favorire il libero scambio delle informazioni è quello di **promuovere un apprendimento diffuso e qualificato attraverso sistemi “aperti”**, che non pongano barriere fra il lavoro e l’informazione.
4. La **libertà di modificare il software** richiama l’imperativo del buon inventore artigiano di “metterci le mani” per capire il funzionamento delle cose e per migliorarle.

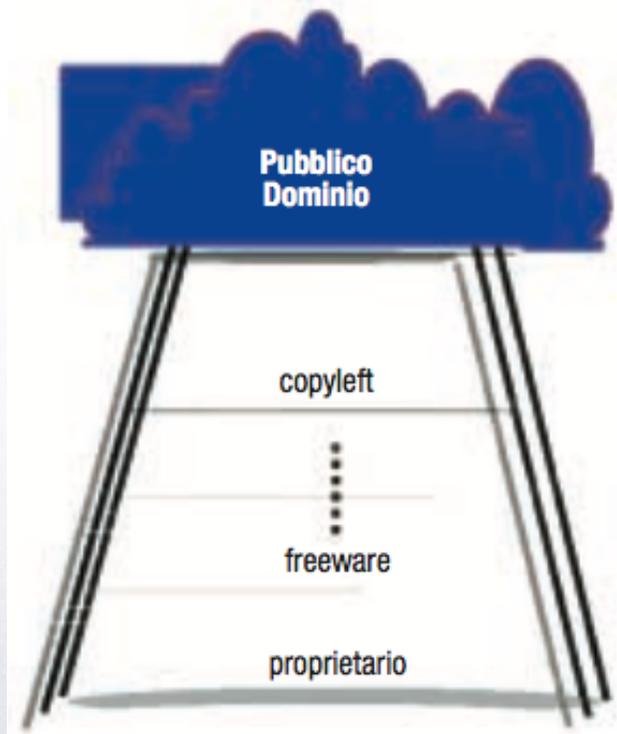
5. Il piacere e il divertimento sono un **importante incentivo alla programmazione.**
6. Le possibilità di innovazione continua offrono un **contributo alla crescita dell'intelligenza e allo sviluppo della professionalità**, a differenza di procedure routinarie che sono, invece, vere e proprie barriere.

In sostanza:

- ▶ ad un'organizzazione del lavoro burocratica si vuole contrapporre un'*organizzazione sociale interattiva e creativa*;
- ▶ a un'organizzazione sociale basata su status e ruoli definiti dal reddito o dalla collocazione sociale si contrappone una *comunità di eguali basata sul merito*.

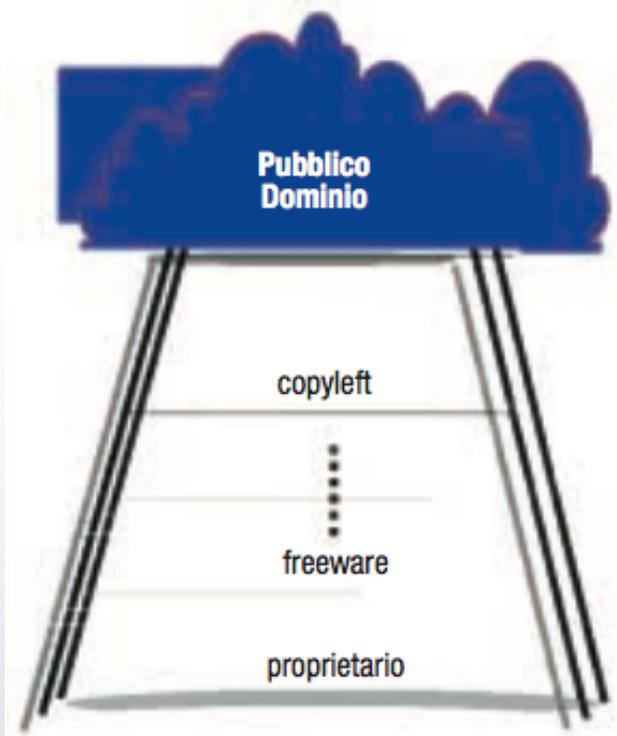


La figura rappresenta schematicamente la complessa *piramide delle libertà del software*.



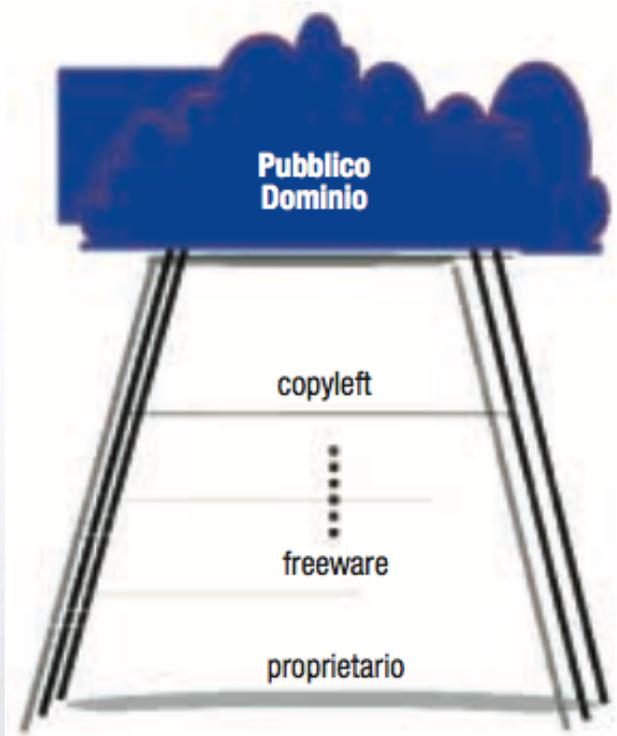
- Al gradino più basso, ossia al minimo livello della libertà, si trova naturalmente il software **proprietario**.
- Il **freeware** è il software distribuito gratuitamente (qui “free” sta per “gratuito”) in forma di codice eseguibile, liberamente utilizzabile ma non modificabile dall’utente finale, che non potrà vantare alcun diritto su di esso.

La figura rappresenta schematicamente la complessa *piramide delle libertà del software*.



- Il livello massimo è quello di “pubblico dominio” (**public domain**), totalmente svincolato da qualsivoglia limite di utilizzo e disponibile in formato sorgente — al limite *anche utilizzabile nello sviluppo di software proprietario*.
- Per questo Stallman non ritenne opportuno rendere GNU di pubblico dominio poiché la catena virtuosa del lavoro e dell’uso cooperativo si sarebbe spezzata e avrebbe interrotto anche il valore insito nel prodotto iniziale.

La figura rappresenta schematicamente la complessa *piramide delle libertà del software*.

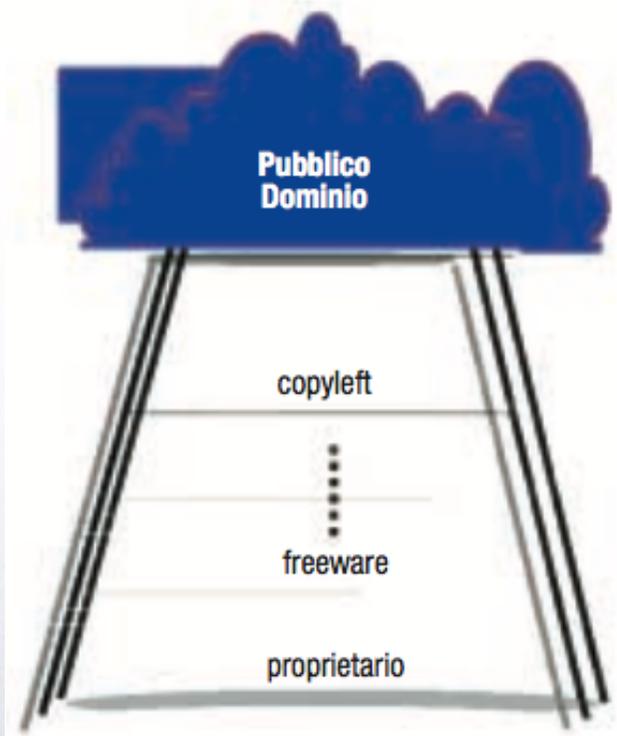


- La scelta fu quindi quella di proteggere il prodotto con un nuovo tipo di licenza, formalmente denominata G.P.L. (*General Public License*) ma scherzosamente chiamata “**copyleft**” in contrapposizione al più noto “copyright”.
- In realtà il gioco di parole è su copyleft inteso come “diritto di copia di sinistra” e copyright inteso come “diritto di copia di destra” nel senso politico dei termini (*absit iniuria verbis!*).



La figura rappresenta schematicamente la complessa *piramide delle libertà del software*.

- Il “copyleft”, che Stallman chiama anche *permesso d'autore*, consente a chi acquista un programma di utilizzarlo in un numero indefinito di copie, di modificarlo a suo piacimento, di distribuirlo nella forma originale o modificata, gratuitamente o a pagamento, alle sole condizioni di distribuirlo in formato sorgente e di imporre a chiunque acquisisca il prodotto di firmare lo stesso tipo di impegno.



- L'utilizzo dei pacchetti software proprietari per la produttività individuale (ad es. MS Office) impatta notevolmente sui **costi di licenza** a carico degli utenti, e in special modo della Pubblica Amministrazione, che in Italia si stima ammontino a circa 1 miliardo di Euro l'anno.
- I formati aperti, intesi sia per i documenti che come software rilasciato completo dei codici sorgenti e delle informazioni necessarie per la compilazione, rappresentano oggi una opportunità preziosa per garantire **grandi economie di scala**.
- In un'ottica di lungo periodo, l'importanza dell'utilizzo di formati aperti assume particolare rilevanza anche a fronte del processo di *dematerializzazione* attualmente in atto.



L'interoperabilità



Obiettivi

*Conoscere le tecnologie per
l'interscambio dei documenti tra
sistemi/formati differenti*



È opportuna una corretta adozione di formati aperti per i documenti informatici che siano coerenti con le tecnologie e normative vigenti, nonché idonei per l'uso nei *sistemi di conservazione documentale*, oltreché per lo *scambio di dati tra sistemi informativi*.

L'evoluzione delle tecnologie, unita alla crescente disponibilità e complessità dell'informazione digitale, ha infatti introdotto la necessità di:

- **gestire** sempre maggiori forme di informazioni digitali (es. testo, immagini, filmati);
- **disporre** di funzionalità specifiche per agevolare la creazione, la modifica e la manipolazione dei documenti.

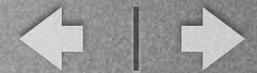


Questo fenomeno ha portato all'aumento del numero dei formati disponibili e dei software per gestirli, nonché delle piattaforme su cui questi operano.

A seconda del loro specifico utilizzo, i **formati** possono essere distinti in:

- ▶ *testi/ documenti* (ad es. doc, html, pdf, odt);
- ▶ *calcolo* (ad es. xls, ods);
- ▶ *immagini* (ad es. gif, jpg, bmp, tif, eps, svg);
- ▶ *suoni* (ad es. mp3, wav);
- ▶ *video* (ad es. mpg, mpeg, mp4, avi, wmv);
- ▶ *codici eseguibili* (ad es. exe, bat, sh);
- ▶ *archiviazione e compressione* (ad es. zip, rar);
- ▶ *email* (ad es. smtp, pop3, mime).





Affinché sia possibile lo scambio di documenti e informazioni tra sistemi differenti è quindi necessario raggiungere l'**interoperabilità** tra gli stessi, che dipende in gran parte dalle scelte tecnologiche dei formati di produzione e scambio documentale.

- *Il concetto di interoperabilità è essenzialmente associato all'interscambio di informazioni tra diversi soggetti, pubblici o privati.*
- *Ciò fornisce le basi per una gestione dematerializzata (anche con valore legale) dei documenti, garantendo l'interscambio tra i diversi sistemi indipendentemente dagli strumenti tecnologicamente disponibili.*





- Uno dei principali ambiti dell'interoperabilità è costituito dall'**adozione di formati aperti per la produttività individuale**, che *non abbiano misure di restrizioni legali* per il loro utilizzo e siano *gestiti da enti di standardizzazione*, i quali ne rilasciano le specifiche alla comunità di sviluppo che a sua volta contribuisce all'evoluzione degli standard.
- I formati aperti hanno il vantaggio di poter essere gestiti sia da software proprietari sia da software libero, *indipendentemente dalle modalità di gestione della licenza d'uso*, lasciando all'utente la **possibilità di scegliere quale piattaforma utilizzare** a salvaguardia degli investimenti effettuati.



Caratteristiche tecniche del documento informatico

Il formato di un documento rappresenta la *convenzione* usata per redigere, interpretare, leggere e modificarne il contenuto.

Ai fini della formazione, gestione e conservazione dei documenti informatici, è necessario *definire i formati* che possano garantire la **leggibilità**, intesa come la capacità di interpretare ed elaborare correttamente il codice binario che lo costituisce secondo delle regole stabilite dal formato con cui esso è stato rappresentato, e la reperibilità del documento informatico nel suo ciclo di vita.

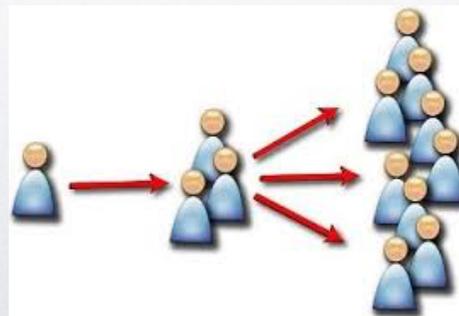
La disponibilità di formati diversi dei documenti nasce dall'**esigenza di rappresentare** in maniera differente le informazioni e dall'evoluzione tecnologica degli strumenti, dei software e delle piattaforme.



Caratteristiche tecniche del documento informatico

Nella **scelta dei formati** vanno tenute in considerazione alcune fondamentali caratteristiche:

- *apertura;*
- *sicurezza;*
- *portabilità;*
- *funzionalità;*
- *supporto allo sviluppo;*
- *diffusione.*





Caratteristiche tecniche del documento informatico

Oltre a ciò, si può anche considerare:

- la **capacità di occupare il minor spazio possibile** in fase di memorizzazione valutando, in funzione delle proprie esigenze, gli eventuali *livelli di compressione* utilizzabili;
- la **possibilità di gestire** il maggior numero di *metadati* (dati che descrivono e definiscono altri dati in un determinato contesto), compresi i riferimenti su chi ha eseguito le modifiche o le aggiunte.



Caratteristiche tecniche del documento informatico

Come già osservato, un formato si dice *aperto* quando è conforme a specifiche pubbliche che rendono sempre possibili la decodifica dei documenti rappresentati ed è disponibile a chiunque abbia interesse ad utilizzarlo.

Il concetto di apertura è collegato anche al *formato* del file che deve poter essere aperto con i diversi strumenti informatici disponibili, anche in assenza di prodotti che effettuino tale operazione automaticamente.



Caratteristiche tecniche del documento informatico

Questa condizione di apertura si verifica quando:

- il formato è documentato e pubblicato al fine di promuoverne l'adozione;
- il documento è conforme a formati definiti da organismi di standardizzazione riconosciuti.

In quest'ultimo caso è necessario che gli organismi interessati garantiscano sull'adeguatezza e sulla completezza delle specifiche stesse.



Caratteristiche tecniche del documento informatico

La **sicurezza** di un formato dipende, principalmente, da due elementi:

- il *grado di modificabilità* del contenuto del file;
- la *capacità di essere immune* all'inserimento di codice maligno (malware).

La **portabilità** di un documento informatico è rappresentata dalla facilità con cui i formati possono essere usati su piattaforme, sia hardware sia software, diverse.

Per **funzionalità** si intende la possibilità — per un determinato formato — di essere gestito con prodotti informatici che prevedono una varietà di funzioni messe a disposizione dell'utente per la formazione e gestione del documento informatico.



Caratteristiche tecniche del documento informatico

Il **supporto allo sviluppo** rappresenta la modalità con cui si mettono a disposizione le risorse necessarie alla manutenzione e allo sviluppo del formato ed ai prodotti informatici che lo gestiscono (ad es. organismi preposti alla definizione di specifiche tecniche e standard, società, comunità di sviluppatori).

La **diffusione** è l'estensione dell'impiego di uno specifico formato per la formazione e la gestione dei documenti informatici. Questo elemento influisce sulla probabilità che esso venga supportato nel tempo, attraverso la disponibilità di più prodotti informatici idonei alla sua gestione e visualizzazione.



Caratteristiche tecniche del documento informatico

È altresì importante suddividere la tipologia di documenti in base all'**utilizzo** che ne sarà fatto:

- documenti che *non necessitano di essere modificati* dall'utente (ad es. dispense, leggi, etc.), dove è necessario preservare anche l'aspetto grafico e l'impaginazione. In tale caso è opportuno l'impiego del formato PDF/A (cfr. seguito), idoneo anche all'archiviazione a lungo termine;
- documenti che l'utente deve *poter compilare* ovvero *modificare*, per i quali è invece opportuno adottare il formato aperto ODF.



Il formato PDF/A (PDF for Archiving)



Il formato PDF/A è uno standard internazionale (ISO19005), sottoinsieme dello standard PDF, appositamente pensato per l'archiviazione nel lungo periodo di documenti elettronici.

Tale standard garantisce che il documento sia *visualizzabile* sempre allo stesso modo, anche a distanza di tempo e con programmi software diversi.

Lo standard PDF/A è suddiviso in due parti. Solo il PDF/A-1 è approvato, ed è suddiviso in due livelli:

1. PDF/A-1a, livello massimo richiesto dallo standard;
2. PDF/A-1b, livello minimo richiesto dallo standard.



Il formato PDF/A (PDF for Archiving)

PDF/A contiene le sole informazioni necessarie per visualizzare il documento come nel momento in cui è stato formato.

Tutto il contenuto visibile del documento deve essere incluso, come il testo, le immagini, vettori grafici, font, colori e altro.

Un documento PDF/A non può contenere macro-istruzioni o riferimenti ad elementi od informazioni (come i font) non contenuti nel file stesso.

Sono commercialmente disponibili diversi prodotti e tools per la creazione, elaborazione e verifica di documenti in formato PDF/A.



Il formato PDF/A (PDF for Archiving)

In particolare si segnalano i seguenti **strumenti** open source:

PDFCreator. Installando la *stampante virtuale* PDFCreator (<https://www.pdfforge.org/pdfcreator>) si può creare un PDF da qualsiasi programma lanciando il documento su questa stampante virtuale. Perché il file sia generato in PDF/A, è necessario selezionare la voce PDF/A-1b nella casella “Formato del documento predefinito” nella finestra delle “Opzioni della stampante virtuale”, sezione “Salvataggio”.



LibreOffice. Mediante il prodotto LibreOffice (cfr. seguito), una volta aperto il documento, dalla voce menu “File” basta scegliere l’opzione “Esporta nel formato PDF...”; quindi, dal menu “Opzioni PDF” nella sezione “Generale” impostare la casella “PDF/A-1”, premere il tasto “Esporta” e salvare il documento, che sarà in formato PDF/A.





Problematiche nell'interscambio di documenti informatici

- La sola *definizione dei formati di scambio* dei documenti non è sufficiente a garantirne la completa corrispondenza visiva e la totale interoperabilità tra i diversi sistemi.
- Uno dei principali problemi è rappresentato dalla *difformità di rappresentazione* grafica di un medesimo documento, al variare del software utilizzato e/o della periferica impiegata.
- Problemi di questo genere non riguardano i documenti informatici nel *formato PDF*, in quanto il documento *contiene*, al proprio interno, *tutte le caratteristiche* tipografiche che verranno interpretate e riprodotte dal programma di visualizzazione.
- Di seguito si esaminano alcuni aspetti per minimizzare tali incompatibilità.



Problematiche nell'interscambio di documenti informatici Font.

- ▶ I programmi per la preparazione di un documento consentono la scelta tra una nutrita serie di *font di caratteri*.
- ▶ L'utilizzo di un font non standard ha come conseguenza una visualizzazione difforme del documento su postazioni di lavoro che ne sono prive.
- ▶ Per ridurre tale inconveniente è necessario prevedere l'utilizzo esclusivo di font standard o, in virtù dell'ampio utilizzo, divenuti uno standard de facto (come ad es. TrueType e OpenType, <https://editoria.ikonos.tv/blog/differenza-font-truetype-opentype/>).



Problematiche nell'interscambio di documenti informatici

Codifica dei caratteri.

Ogni carattere corrisponde ad un codice definito con uno standard. Quello maggiormente utilizzato è lo standard ASCII, che prevede la codifica a 8 bit di 255 caratteri, opportunamente differenziati:

- i caratteri con codice da 0 a 31 non sono visualizzabili (codici di controllo);
- i caratteri con codice compreso tra 32 e 127 sono internazionalmente riconosciuti e univocamente interpretati;
- i caratteri con codifica superiore a 127 non sono standardizzati in modo univoco, ma sono personalizzati in base ad esigenze geografiche. Ciò significa che in paesi diversi, lo stesso codice corrisponde a caratteri diversi.



Problematiche nell'interscambio di documenti informatici

Codifica dei caratteri.

La codifica del carattere non pone problemi nel formato PDF, mentre ha rilevanza per i documenti in formato RTF⁽¹⁾ o HTML⁽²⁾.

Per avere la certezza che il documento realizzato sia ugualmente riprodotto su diverse postazioni di lavoro, occorrerà limitarsi all'utilizzo dei caratteri compresi tra il codice 32 ed il 127.

(1) Rich Text Format, un formato di documento di testo utilizzato da semplici programmi di videoscrittura come WordPad.

(2) HyperText Markup Language, il linguaggio con cui sono scritte le pagine Web.



Problematiche nell'interscambio di documenti informatici

Trasmissione dei documenti.

La trasmissione dei documenti digitali (in special modo mediante posta elettronica) potrebbe portare ad una possibile diversità nella visualizzazione degli stessi (anche se in generale gli applicativi sono in grado di mitigare o annullare i problemi derivanti da una diversa codifica dei caratteri).

La *posta elettronica* è sottoposta a standardizzazione da parte delle autorità preposte a tale funzione per Internet (World Wide Web Consortium, W3C), attraverso delle apposite direttive denominate Request for Comments (RFC).



Problematiche nell'interscambio di documenti informatici

Trasmissione dei documenti.

Attualmente gli standard che seguono la trasmissione di documenti informatici via posta elettronica sono le RFC (Request For Comments):

- 821 [*Simple Mail Transfer Protocol*, <https://tools.ietf.org/html/rfc821>]
- RFC 822 [*Standard for the Format of ARPA Internet Text Messages*, <https://tools.ietf.org/html/rfc822>]
- 2046 [*Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types*, <https://tools.ietf.org/html/rfc2046>]
- 2049 [*Multipurpose Internet Mail Extensions (MIME) Part Five: Conformance Criteria and Examples*, <https://tools.ietf.org/html/rfc2049>]

e successive modifiche.

Si osservi che i documenti denominati RFC sono uno dei primi esempi di testo elettronico ad ampia diffusione, che utilizza esclusivamente il set dei 96 caratteri ASCII base con codici compresi tra 32 e 127.



Lo Standard XML



XML (eXtensible Markup Language) è un *linguaggio marcatore* basato su un meccanismo sintattico che:

- consente di definire e controllare il significato degli elementi contenuti in un documento o in un testo;
- consente di definire nuovi linguaggi di marcatura.

Il nome indica che si tratta di un linguaggio marcatore (*Markup Language*) estensibile (*eXtensible*) in quanto permette di creare tag (etichette) personalizzati.



Lo Standard XML

Il World Wide Web Consortium (W3C), per uniformare il funzionamento dei browser e standardizzare le specifiche ufficiali dell'HTML, definì nel 1998 XML come un linguaggio di markup che desse maggiore *libertà nella definizione dei tag*, pur rimanendo in uno standard.



Ben presto ci si accorse che XML non era limitato al solo contesto web ma era qualcosa di più: uno strumento che permetteva di essere utilizzato nei più diversi contesti, dalla definizione della struttura di documenti, allo scambio delle informazioni tra sistemi diversi, dalla rappresentazione di immagini alla definizione di formati di dati.

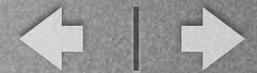


Lo Standard XML

- Mentre HTML definisce una grammatica per la descrizione e la formattazione di pagine web (layout), XML è un *metalinguaggio* utilizzato per creare nuovi linguaggi, atti a descrivere documenti strutturati.
- HTML ha perciò un insieme ben definito e ristretto di tag; XML rende invece possibile definire nuovi tag a seconda delle esigenze.
- L'XML è oggi molto utilizzato anche come mezzo per l'esportazione di dati tra diversi DBMS.

```
<?xml version="1.0"?>
<quiz>
  <qanda seq="1">
    <question>
      Who was the forty-second
      president of the U.S.A.?
    </question>
    <answer>
      William Jefferson Clinton
    </answer>
  </qanda>
  <!-- Note: We need to add
  more questions later.-->
</quiz>
```





Lo Standard XML — Sintassi

Vediamo un esempio tipico di file XML, visualizzabile mediante un qualsiasi browser.

```
<?xml version="1.0" encoding="UTF-8"?>
<corso>
  <utente anni="30">
    <nome>Mario</nome>
    <cognome>Rossi</cognome>
    <indirizzo>Foggia</indirizzo>
  </utente>
  <utente anni="52">
    <nome>Anna</nome>
    <cognome>Bianchi</cognome>
    <indirizzo>Bari</indirizzo>
  </utente>
</corso>
```

- La prima riga indica la versione di XML in uso e specifica la codifica UTF-8 per la corretta interpretazione dei dati.
- I caratteri speciali che renderebbero il documento mal formato vanno sostituiti con le rispettive entità XML:

Carattere	Entità
&	&
<	<
>	>
"	"
'	'



Lo Standard XML — Tag

I marcatori, chiamati **tag** (etichette), assegnano un *significato* al testo.

I tag possono contenere informazioni in due modi: attraverso dei **parametri** oppure racchiudendo del testo o altri tipi di informazioni.

Vi sono **tag di apertura**, necessariamente seguiti da **tag di chiusura** (tra i quali si può avere un contenuto) oppure **tag autonomi**, che possono fornire informazioni solo attraverso i loro parametri.

Ogni tag inizia e finisce con delle parentesi angolari **<>**, mentre la chiusura è rappresentata dalla barra **/**.



Lo Standard XML — Tag

Ecco degli esempi di tag:

```
<tagDiEsempio parametro="valore" caratteristiche="solo il tag di apertura ha dei parametri">  
  contenuto  
</tagDiEsempio>  
<tagAutonomo parametro="tag autonomo" caratteristiche="non ha tag di chiusura" />
```

I commenti vanno racchiusi da `<!-- -->`. Ad esempio: `<!-- Questo è un commento -->`

In XML è necessario rispettare alcune regole:

- i tag *non possono iniziare con numeri o caratteri speciali e non possono contenere spazi*;
- i tag *devono essere bilanciati*, ovvero non sono consentiti errori di annidamento.



Lo Standard XML — Tag

Esempi.

```
<contatti>
  <nome>Mario</nome>
  <cognome>Rossi
</contatti>
```

Il tag "cognome" non è stato chiuso, l'XML risulta quindi mal formato.

Il tag "cognome" è stato chiuso dopo il tag "contatti", anche in questo caso l'XML non è ben formato.

```
<contatti>
  <nome>Mario</nome>
  <cognome>Rossi
</contatti>
</cognome>
```



Lo Standard XML — Tag

Esempi.

```
<contatti>
  <nome>Mario</nome>
  <cognome>Rossi</COGNOME>
</contatti>
```

XML è case sensitive, quindi il tag "cognome" e il tag "COGNOME" sono considerati due tag diversi. Di conseguenza l'XML non è ben formato.

È possibile definire tag vuoti, che vengono aperti e chiusi immediatamente:

```
<contatti></contatti>
```

Oppure, in maniera sintetica:

```
<contatti/>
```



Lo Standard XML

Dagli esempi abbiamo visto che un documento XML deve essere *ben formattato*, deve cioè possedere le seguenti caratteristiche:

- un **prologo**, la prima istruzione che appare nel documento. Ad esempio:
`<?xml version="1.0" encoding="UTF-8"?>`
- un unico **elemento radice** (ovvero il nodo principale, chiamato *root element*) che contiene tutti gli altri nodi del documento. Ad esempio:
`<contatti>`
- all'interno del documento tutti i **tag** devono essere **bilanciati**.

Se il documento XML non contiene errori si dice *well formed* (ben formato).

Se il documento è *well formed* e in più rispetta i requisiti strutturali definiti nel file DTD o schema XML associato viene chiamato *valid* (valido).



Lo Standard XML — Schemi

Un documento XML si uniforma alle specifiche definite in uno **schema**; i più importanti sono:

- ▶ **DTD** (acronimo di *Document Type Definition*)

Specifica le caratteristiche strutturali di un documento XML attraverso una serie di "regole grammaticali". In particolare definisce l'insieme degli elementi del documento XML, le relazioni gerarchiche tra gli elementi, l'ordine di apparizione nel documento XML e quali elementi e quali attributi sono opzionali o meno.

- ▶ **XML Schema**

Come la DTD, serve a definire la struttura di un documento XML. Il W3C consiglia di adottarlo al posto della DTD, essendo una tecnica più recente ed avanzata. La sua sigla è XSD, acronimo di *XML Schema Definition*.

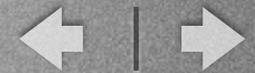


Il pacchetto open source LibreOffice



Il progetto **LibreOffice**, sviluppato da *The Document Foundation* in prosecuzione del precedente progetto *OpenOffice*, è costituito da molte comunità che operano su differenti aspetti del software: codice sorgente, documentazione, controllo di qualità, progettazione ed altro.

The Document Foundation è un'*organizzazione senza fini di lucro*, aperta a chiunque sia d'accordo con i principi fondanti e voglia contribuire alle sue attività.



Accanto a LibreOffice, la fondazione ha anche avviato *The Document Liberation Project*, una comunità di sviluppatori uniti per liberare gli utenti dalla schiavitù del software proprietario fornendo loro potenti **strumenti per la conversione di formati** di file proprietari (come ad es. DOC, XLS, etc.) nel corrispondente formato ODF.



L'obiettivo della fondazione è infatti quello di eliminare il *digital divide* (*) e ampliare le potenzialità di tutti i cittadini, supportare la conservazione delle lingue madri, ed evitare i vincoli del software e dei formati proprietari.

(*) Il divario esistente tra chi ha accesso effettivo alle tecnologie dell'informazione e chi ne è escluso, in modo parziale o totale.



Caratteristiche principali

Informazioni su LibreOffice



LibreOffice
The Document Foundation

LibreOffice è una suite per ufficio moderna e di facile utilizzo per la redazione di testi, fogli elettronici, presentazioni e altro ancora.

Questo rilascio è stato fornito da The Document Foundation.
Copyright © 2000–2022 Contributori di LibreOffice.
LibreOffice era basato su OpenOffice.org.

[Riconoscimenti](#) [Sito web](#) [Note di rilascio](#)

Informazioni sulla versione 

Versione:	7.3.3.2 / LibreOffice Community
Build:	d1d0ea68f081ee2800a922cac...
Ambiente:	Thread CPU: 8; SO: Mac OS X 10.16
Interfaccia utente:	Resa interfaccia: predefinito; VCL: osx
Versione locale:	it-IT (it_IT.UTF-8); Interfaccia utente: it-IT
Varie:	Calc: threaded

[Chiudi](#)

Il pacchetto LibreOffice è attualmente (Giugno 2022) alla versione 7.3.3.2 ed è scaricabile dall'indirizzo: <http://it.libreoffice.org>

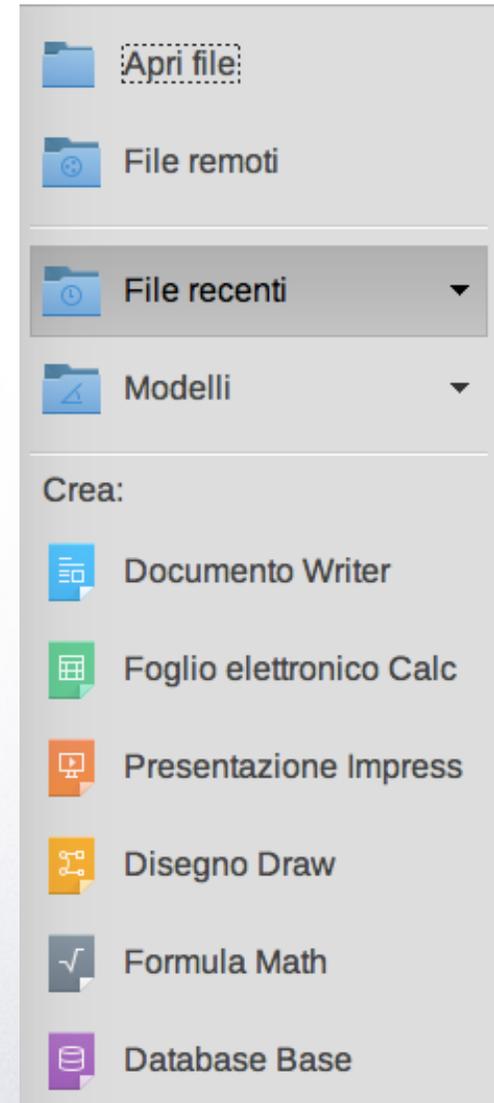


Caratteristiche principali

All'avvio presenta una schermata che consente di aprire un file già esistente o di crearne uno nuovo in base ai moduli applicativi disponibili.

Tale schermata è praticamente identica sulle varie *piattaforme* su cui il software è disponibile (Linux, Mac OS, Windows) e ciò rende molto immediato il trasferimento di documenti tra i vari sistemi nonché il “riuso” delle competenze da parte degli utenti.

Una volta imparato ad usare il software, l'utente può spostarsi da un ambiente operativo all'altro senza perdita di contenuti né di tempo per dover imparare un nuovo programma!





Caratteristiche principali

La (apparente) diffusione e successo degli ambienti proprietari come Microsoft Windows e Office non sono dovuti ad una loro (pretesa) vasta accettazione.

Semplicemente è la **pigrizia** degli utenti che, acquistando un nuovo computer e trovandovi già installati Windows e Office (non sempre con regolare licenza...), non si preoccupano di sostituirli con software libero.

Ciò è anche più grave nella P.A. dove viene speso circa 1 miliardo di Euro all'anno in licenze di software proprietario, che potrebbero tranquillamente essere risparmiati (*spending review?*) con l'adozione di software open source.



Caratteristiche principali

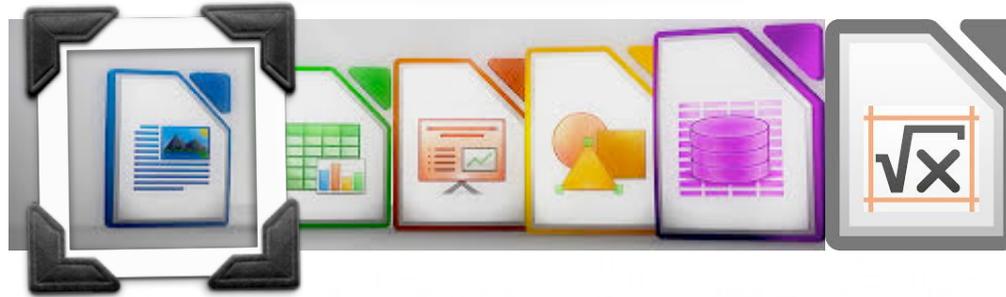
Solo quando sarà troppo tardi l'utente si renderà conto di essere divenuto “schiavo” di un sistema operativo e di una piattaforma di scrittura **complessa, costosa e non sempre facile** da usare.

Quindi, prima che sia troppo tardi, è bene **imparare l'utilizzo della suite LibreOffice** (anche e di più Linux, se possibile): vi sarà sempre il tempo per passare a software proprietari, almeno con cognizione di causa e sapendo che “esiste sempre una via d'uscita”.





Moduli applicativi



Gestione testi *Writer*.

- LibreOffice *Writer* è un ricco tool per la creazione di lettere, testi, report, newsletter, brochure e altri documenti.
- È possibile inserire nei documenti *Writer* grafici e oggetti provenienti dagli altri moduli.
- *Writer* consente di esportare file nei formati HTML, XHTML, XML, PDF nonché nelle varie versioni di Microsoft Word.
- È inoltre in grado di connettersi al programma di posta elettronica.



Moduli applicativi



Foglio elettronico *Calc*.

- Comprende tutte le caratteristiche per l'analisi avanzata, preparazione di diagrammi e strumenti di decisione che ci si aspetta da uno spreadsheet di buon livello.
- Include oltre 300 funzioni per operazioni finanziarie, statistiche e matematiche.
- Gli “Scenari” consentono la c.d. analisi *what if* (cosa succede se...).



Moduli applicativi



Foglio elettronico *Calc*.

- Genera diagrammi bi e tridimensionali, che possono essere integrati in altri documenti LibreOffice.
- È anche possibile aprire e lavorare con cartelle Microsoft Excel e salvarle nel formato nativo.
- Calc è anche in grado di esportare fogli in diversi formati, come CSV (*Comma Separated Value*) PDF e HTML.



Moduli applicativi



Presentazione *Impress*.

- Impress fornisce tutti i comuni strumenti di presentazione multimediale, come effetti speciali, animazioni e strumenti di disegno.
- È integrato con le avanzate capacità grafiche dei moduli LibreOffice *Draw* e *Math*.
- Le presentazioni possono essere ulteriormente arricchite con testo *Fontwork* (analogo a WordArt), suoni e video.
- Impress è compatibile con il formato di PowerPoint e può anche salvare il lavoro in numerosi formati grafici.



Moduli applicativi



Disegno *Draw*.

- Si tratta di uno strumento di *disegno vettoriale* che può servire per produrre qualsiasi tipo di grafica, da semplici diagrammi a artefatti tridimensionali.
- Nella *grafica vettoriale* (cfr. Modulo 2) un'immagine è descritta mediante un insieme di primitive geometriche che definiscono punti, linee, curve e poligoni ai quali possono essere attribuiti colori e anche sfumature.
- È radicalmente diversa dalla grafica tradizionale bitmap, tipica ad es. di una foto JPEG, in quanto in quest'ultima le immagini vengono descritte come una griglia di punti a video (pixel) opportunamente colorati.



Moduli applicativi



Disegno *Draw*.

- Si può impiegare per creare disegni da utilizzare in uno qualsiasi dei moduli di LibreOffice, compresa la creazione di clip da aggiungere alla Galleria.
- È anche possibile importare grafica da molti formati comuni e esportare nei formati BMP, EMF, EPS, GIF, HTML, JPEG, PDF, PNG, SVG, TIFF, WMF.



Moduli applicativi



Gestione database *Base*.

- LibreOffice *Base* fornisce degli strumenti per l'operatività quotidiana su basi di dati con un'interfaccia semplice da usare.
- Può creare e modificare maschere, report, query, tabelle, viste logiche e relazioni.
- Si può gestire una base di dati relazionale in maniera simile ad altre applicazioni per database (come ad es. MS Access).

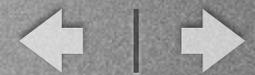


Moduli applicativi

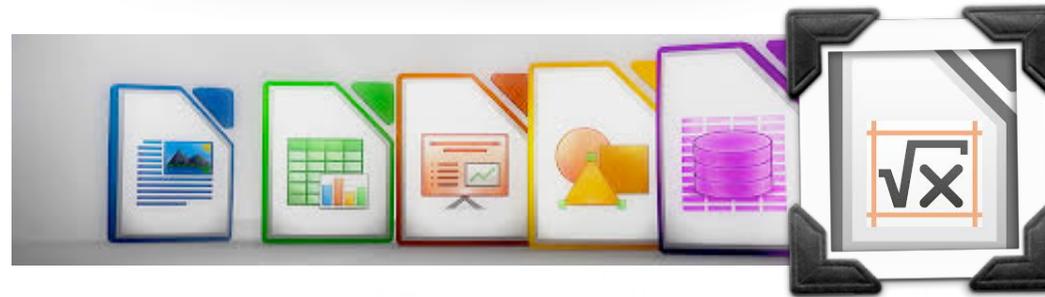


Gestione database *Base*.

- *Base* fornisce molte nuove caratteristiche, come ad es. la possibilità di analizzare e modificare relazioni in vista diagramma.
- *Base* incorpora come motore relazionale standard l'engine HSQLDB. Può anche utilizzare database dBASE, Microsoft Access, MySQL, Oracle o qualsiasi database ODBC o JDBC.
- Fornisce anche supporto per un sottoinsieme del linguaggio ANSI-92 SQL.



Moduli applicativi



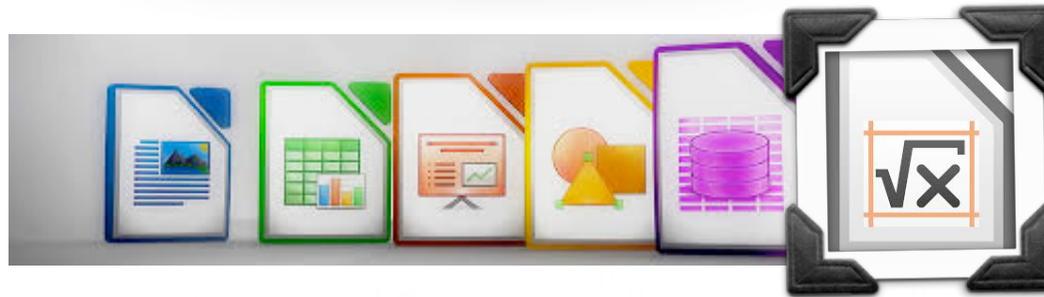
Formule matematiche *Math*.

- *Math* è l'editor di formule ed equazioni di LibreOffice.
- Si può utilizzare per creare complesse formule (vedi figura) che includono simboli o caratteri non disponibili nei font standard.

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$$



Moduli applicativi



Formule matematiche *Math*.

- Anche se è utilizzato per lo più per creare formule in altri documenti, come *Writer* e *Impress*, *Math* può anche operare come strumento a sé stante.
- Si possono salvare le formule nel formato standard MathML (Mathematical Markup Language, derivante da XML) per inserirle in pagine web e altri documenti non creati con LibreOffice.



V a n t a g g i

Elenchiamo alcuni dei principali vantaggi di LibreOffice rispetto ad altre suite applicative per l'ufficio.

1. **Nessun costo di licenza**

- chiunque può utilizzare e distribuire gratuitamente LibreOffice;
- molte caratteristiche che sono disponibili come componenti aggiuntivi a pagamento in altre suite per ufficio (come ad es. l'esportazione in formato PDF) sono gratuite con LibreOffice;
- non ci sono costi nascosti, né ora né in futuro.



V a n t a g g i

2. **Open source**

- il software può essere distribuito, copiato e modificato all'occorrenza, secondo le modalità stabilite dalle licenze cui è soggetto LibreOffice.

3. **Cross-platform**

- LibreOffice può essere eseguito su diverse architetture hardware e con diversi sistemi operativi, come Microsoft Windows, Mac OS e Linux;
- offre ampio supporto per le lingue: l'interfaccia utente è disponibile in oltre 40 lingue e il progetto LibreOffice offre ortografia, sillabazione e dizionari per oltre 70 lingue e dialetti comprese urdu, ebraico e arabo.



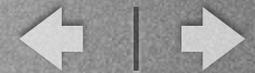
V a n t a g g i

4. Interfaccia utente

- tutti i componenti hanno un *look and feel* (aspetto) simile, che li rende facili da usare e padroneggiare.

5. Integrazione

- i componenti di LibreOffice sono ben integrati tra loro;
- tutti i componenti condividono un correttore ortografico comune e altri strumenti, che sono utilizzati in modo coerente in tutta la suite;
- non è necessario sapere quale applicazione è stata utilizzata per creare un particolare file. Ad esempio, è possibile aprire un file *Draw* da *Writer*.



V a n t a g g i

6. **Granularità**

- di solito, se si modifica un'opzione, essa riguarda tutti i componenti;
- tuttavia, le opzioni di LibreOffice possono essere impostate a livello di componente o addirittura a livello di documento.

7. **Compatibilità dei file**

- oltre ai suoi formati OpenDocument nativi, LibreOffice include la possibilità di esportazione nel formato PDF, così come il supporto per l'apertura e salvataggio dei file in molti formati comuni, tra cui Microsoft Office, HTML, XML, WordPerfect, e formati di Lotus 1-2-3;
- un'estensione (inclusa) offre la possibilità di importare e modificare alcuni file PDF.



V a n t a g g i

8. Indipendenza dal produttore

- LibreOffice utilizza OpenDocument, un formato XML di file sviluppato come standard del settore da OASIS (*Organization for the Advancement of Structured Information Standards*);
- questi file sono facilmente leggibili con qualsiasi editor di testo, e la loro struttura è aperta e pubblica.

9. Partecipazione

- i miglioramenti, le correzioni al software e le date di uscita delle versioni sono definite dalla comunità;
- è possibile far parte della comunità e contribuire allo sviluppo del prodotto che si usa.



Open source e Pubblica Amministrazione



Obiettivi

*Comprendere la problematica
dell'adozione del software
aperto nella P.A.*



I singoli cittadini, il mondo della ricerca, ma, soprattutto, le Pubbliche Amministrazioni guardano con interesse al **software open source** e al modello di sviluppo a esso collegato.

Da molte parti si ritiene che il paradigma open source possa **favorire il rilancio dell'industria informatica nazionale** ed europea e il dibattito sul software open source concerne l'impatto che esso può avere sullo **sviluppo del Sistema Paese**.



Le licenze alla base dell'open source

- Giuridicamente, *il software libero non è un software senza diritti*: esso resta governato dalle disposizioni della licenza e l'autore del software resta il titolare dell'insieme dei diritti di autore.
- Le licenze relative consistono in una *messa a disposizione del software* con l'intento di permetterne la libera evoluzione.
- *La licenza non ha per scopo il trasferimento di un diritto di proprietà* o la rinuncia al diritto di autore o di “far cadere” il software nel pubblico dominio.
- Diffondendo il proprio software libero l'autore può assicurarsi che *la libera utilizzazione del software non sia perturbata dalle azioni dei soggetti alla licenza*.



Le licenze alla base dell'open source

- Come già visto, il concetto di *copyleft* è definito in contrapposizione al tradizionale copyright: mentre quest'ultimo tende a tutelare il diritto d'autore, anche attraverso limitazioni all'accesso della conoscenza, il **copyleft intende tutelare il più generale diritto della collettività a fruire dei prodotti dell'innovazione.**
- I principi del copyleft sono formalizzati dalla FSF nella nota *General Public License* (GPL). L'utente di un software sotto licenza GPL è vincolato a utilizzare a sua volta la GPL e dovrà, quindi, fornire il codice sorgente delle estensioni realizzate.
- In altre parole, la GPL è un modello di licenza *ricorsivo*: il codice e le “libertà” a esso associate diventano così legalmente inseparabili.

Le licenze alla base dell'open source

Le licenze open source non si limitano comunque alla GPL. Ricordiamo che esistono modelli di licenza alternativi quali, ad esempio:

- LGPL (Lesser General Public License)
- BSD (Berkeley Software Distribution)
- MPL (Mozilla Public License)

che prevedono — in forme differenti — l'apertura del codice sorgente.





Le licenze alla base dell'open source

Le licenze open source non prescrivono che il software debba essere ceduto gratuitamente.

Il software open source **non è alternativo** al software commerciale: il modello open source non preclude la presenza di distribuzione commerciale, di fornitori di valore aggiunto o di servizi di supporto.

È corretto definire l'open source come alternativa al modello di licenza proprietario (*closed source*), in cui l'accesso al codice sorgente non è concesso e il fornitore del software vende all'utente una "licenza d'uso" — temporanea o illimitata — che consente la fruizione del prodotto, ma non implica in nessun modo che l'utente ne acquisisca il controllo.



L'open source nella Pubblica Amministrazione

I dati relativi all'uso di open source nella Pubblica Amministrazione sono contraddittori a causa della complessità del rilevamento delle informazioni e in quanto non esistono dati per il solo settore pubblico.

La Commissione Europea ha promosso alcuni studi sul tema. I più articolati e significativi sono i progetti:

- ▶ **IDABC** (*Interoperable Delivery of European eGovernment Services to public Administrations, Business and Citizens*)
- ▶ **ISA** (*Interoperability Solutions for European Public Administrations*)
- ▶ **FLOSS** (*Free/Libre and Open Source Software: Survey and Study*)



L'open source nella Pubblica Amministrazione

Non tutte le amministrazioni pubbliche dei vari stati europei hanno analoghe *policy* sul tema dell'open source, con un'attenzione comunque complessivamente crescente.

Da vari studi e report pubblici si evince che l'**utilizzo dell'open source nella Pubblica Amministrazione è concentrato nella fascia “server”**, con *Linux* quale sistema operativo preponderante e *Apache* quale server web più diffuso, insieme al middleware collegato (TOMCAT, Perl, PHP, etc.).

Sul lato “client”, invece, le suite open source per l'Office Automation OpenOffice e LibreOffice costituiscono il fenomeno più interessante, anche se con dimensioni quantitativamente ben più contenute.

L'open source nella Pubblica Amministrazione



- Nel caso particolare della *Pubblica Amministrazione Locale* italiana, secondo l'osservatorio “Software Open Source nella Pubblica Amministrazione” dell'Università di Bologna, emerge un utilizzo relativamente ridotto dei server open source Apache: **meno del 40%** nella PA italiana, contro oltre il 60% a livello mondiale.

L'open source nella Pubblica Amministrazione



- Nel 2012 il Governo Monti, con il *Decreto Sviluppo 2012*, indicava come **le pubbliche amministrazioni dovessero prendere in considerazione e soprattutto dare la preferenza a soluzioni basate su software libero** (come ad es. i sistemi operativi Debian/Linux, la suite per l'ufficio LibreOffice, etc.)



L'open source nella Pubblica Amministrazione



- La legge 134/2012 stabiliva che **solo in casi in cui non ci fossero alternative** “libere” le pubbliche amministrazioni potevano acquistare software proprietari.
- Sulla base di questi criteri l'*Agenzia per l'Italia Digitale* della Presidenza del Consiglio dei Ministri ha varato la circolare 63/2013 e la determinazione 115/2019 che incoraggiano la Pubblica Amministrazione a prendere in considerazione il software open source.



L'open source nella Pubblica Amministrazione

L'invito all'utilizzo del software open source nella P.A. è anche presente nel comma 1 dell'art. 68 del C.A.D. (*Codice dell'Amministrazione Digitale*), che così recita:

Le pubbliche amministrazioni acquisiscono programmi informatici o parti di essi a seguito di una valutazione comparativa di tipo tecnico ed economico tra le seguenti soluzioni disponibili sul mercato:

- a) **software sviluppato** per conto della pubblica amministrazione;*
- b) **riutilizzo di software** o parti di esso sviluppati per conto della pubblica amministrazione;*
- c) **software libero** o a codice sorgente aperto;*
- d) **software combinazione** delle precedenti soluzioni.*



L'open source nella Pubblica Amministrazione

- Da ciò si evince che **solo quando la valutazione comparativa** tecnico-economica **dimostri l'impossibilità** di accedere a soluzioni open source o già sviluppate all'interno della pubblica amministrazione ad un prezzo inferiore, è consentita l'acquisizione di programmi informatici di tipo proprietario.
- In particolare, non essendo il software prodotto da o per le amministrazioni pubbliche costituito da “pacchetti standard” riutilizzabili così come sviluppati, e data la necessità di fornire il software con il codice sorgente per consentire di adattarlo alle realtà locali prima di applicarlo e ridistribuirlo, è **naturale per la P.A. aderire al modello Open Source.**
- Ciò soprattutto per rispettare i requisiti fondamentali di **economicità, indipendenza** dai fornitori, **sicurezza, riusabilità e interoperabilità.**



L'open source nella Pubblica Amministrazione

Il primo effetto dell'adozione di software open source consiste in un **risparmio** in termini di costi per licenze.

Perché il confronto economico sia corretto, occorre però tener conto non solo della spesa iniziale, ma piuttosto confrontare il TCO (*Total Cost of Ownership*, costo totale di proprietà) delle soluzioni open source e di quelle proprietarie.

Oltre al costo delle licenze, nel TCO confluiscono:

- le spese dei servizi di supporto,
- le spese della formazione,
- i costi di migrazione, installazione e gestione.

L'open source nella Pubblica Amministrazione

Un elemento a favore del software open source nella P.A. è l'**indipendenza dai fornitori**, consistente nel poter affidarsi per l'assistenza di un prodotto open source ad un'azienda scelta dalla P.A., al contrario di quanto succede nel mondo del software proprietario in cui solo il produttore (o un suo partner autorizzato) può supportare il proprio software.



L'open source nella Pubblica Amministrazione

- Disporre, inoltre, del codice sorgente dei programmi utilizzati all'interno delle organizzazioni pubbliche permette (anche se non garantisce) **un grado maggiore di sicurezza**.
- Sono, infatti, più agevoli i controlli interni alla ricerca di eventuali errori (*bug*) o debolezze (*back-door*) sfruttabili da attacchi esterni.
- L'amministrazione dispone di **un miglior controllo** sulla “politica” di evoluzione del software applicativo alla base del suo sistema informativo e, in generale, **sul governo della gestione del patrimonio pubblico**.





L'open source nella Pubblica Amministrazione

- Questo argomento assume una rilevanza particolare in aspetti connessi con la sicurezza, quali l'**autenticazione** e l'**identificazione del cittadino** o quelli relativi all'**integrità, confidenzialità** e all'**accessibilità** dei dati nel corso del tempo, aspetti cruciali per l'infrastruttura pubblica di autenticazione SPID che consentirà ai cittadini di accedere ai servizi pubblici (e privati) mediante un'unica tecnologia di autenticazione (password, OTP, biometrica o firma avanzata).





L'open source nella Pubblica Amministrazione

La modernizzazione dei sistemi informatici (e quindi informativi) dello Stato passa quindi per il ricorso al *software libero*, che consente:

- ✓ di accedere a un **patrimonio considerevole di software** spesso di qualità e conforme agli standard;
- ✓ di controllare (per minimizzarlo) il **rapporto costi/benefici** della soluzione software attraverso il rafforzamento della concorrenza;
- ✓ di governare il software e di assicurarne la **disponibilità futura**.



L'open source nella Pubblica Amministrazione

Oltre al software libero in sé, l'utilizzo consequenziale di formati aperti e standard favorisce l'interscambio di dati; tale scelta assicura:

Indipendenza. La documentazione pubblica e completa del formato consente l'indipendenza da uno specifico prodotto e fornitore; tutti possono sviluppare applicazioni che gestiscono un formato aperto.

Interoperabilità. Usando formati aperti standard, sistemi eterogenei sono in grado di condividere gli stessi dati.

Neutralità. I formati aperti non obbligano a usare uno specifico prodotto, lasciando libero l'utente di scegliere sulla base del rapporto qualità/prezzo e delle competenze interne.



Limiti

Alcuni **limiti** riguardano comunque ad esempio la *bassa compatibilità* del software open source con standard commerciali, che si evidenzia in particolare nella cronica mancanza di driver software negli ambienti open.

È infatti noto che la maggioranza dei produttori di periferiche non forniscono driver per Linux, dunque la lista dell'hardware compatibile è limitata ai dispositivi cui la comunità degli sviluppatori open source ha accesso.

Quando viene lanciata sul mercato una nuova periferica, occorrono mesi prima che i driver siano disponibili; il problema è particolarmente evidente per le stampanti, schede video e modem/router.



Limiti

È d'altronde evidente che qualsiasi sviluppo e messa in produzione di software comporta dei rischi; per il software libero comporta una serie di potenziali problemi quali ad es.:

- lo sviluppo effettuato utilizzando componenti di software libero — le cui licenze possono risultare fra di loro incompatibili — può comportare il **rischio di compromettere la legalità** del prodotto finale;
- la pubblica amministrazione, in quanto responsabile del software sviluppato o modificato, è pienamente coinvolta nella problematica del **rispetto del diritto d'autore** o della **garanzia del prodotto software** rilasciato.



Limiti

Non si può quindi pensare che nella P.A. sia possibile, senza sforzo iniziale e competenze interne, beneficiare di aggiornamenti e potenziamenti gratuiti, nel quadro di un sistema a codice sorgente aperto, a meno che non siano rispettate **precise condizioni di sviluppo**, quali:

- *numero ragionevole* di operatori “con lo stesso problema”;
- *ripartizione* iniziale e flessibile *dei compiti* di gestione/direzione tra persone diverse appartenenti a organizzazioni diverse;
- *documentazione* esauriente;
- *codici non monolitici*, bensì software articolato in più segmenti di codice di dimensioni relativamente ridotte (moduli), così da agevolarne una gestione individuale;
- *chiara identificazione delle parti* mature/stabili e di quelle che vanno ottimizzate in base al principio del “versioning”.

Limiti

D'altronde molte di queste considerazioni valgono anche per lo sviluppo di software proprietari.

Il paradigma dell'open source nella Pubblica Amministrazione può però **contribuire ad applicare "best practices"** troppo spesso solo enunciate e ad **attivare un ciclo virtuoso** di sviluppo del software, in un modello originale e promettente nell'interesse collettivo.





Grazie per l'attenzione!