

Perché abbandonare i formati proprietari (e i software che li producono)

(Specialmente quelli di Microsoft)

Luca Lorenzetto – l.lorenzetto@dancetj.net – gio 5 febbraio 2009



Note sull'autore:

Luca Lorenzetto, si occupa di informatica libera fin dal 2003 quando ha incontrato il sistema operativo Libero “Linux”. Da quel momento è un entusiasta dell'Open Source a tutto campo e ne è un convinto sostenitore. Attualmente studia “Sicurezza dei Sistemi e delle Reti Informatiche” presso il DTI dell'Università Statale di Milano. Lavora come collaboratore di una ditta privata presso una Pubblica Amministrazione e fa parte del Local Community Team Italiano di Ubuntu oltre che essere presidente del TvLUG – Treviso Linux Users Group.

Da sempre appassionato dal nuoto, ma grande adoratore della montagna :-)

Questo Documento è © 2009 Luca Lorenzetto ed è rilasciato pubblicamente secondo i termini della licenza Creative Commons BY-NC-SA 2.5, reperibile all'url <http://creativecommons.org/licenses/by-nc-sa/2.5/it/>



Introduzione

Nell'ambito delle scelte di software, al giorno d'oggi, non ci si trova più davanti a scelte obbligate come 10/12 anni fa. Attraverso la diffusione della rete e la discesa in campo del software Open Source¹, lo scenario presenta enormi possibilità di scelta sulla quale possono influire molti parametri come l'affidabilità, la sicurezza e l'aderenza agli standard, oltre che alla libertà che questo software fornisce e all'economicità. Attualmente buona parte delle scelte non sono vere e proprie scelte, ma delle “riconferme” di decisioni effettuate molto tempo addietro sulla base di criteri che non sono gli stessi di oggi.

Le funzionalità e potenzialità che otteniamo al giorno d'oggi con i software liberi ci aprono gli occhi e rendono visibile a tutti, problemi e grossi buchi del modello di sviluppo proprietario.

In questo documento verranno analizzate le problematiche e vulnerabilità l'adozione di un formato non libero come quello di Microsoft Office e del software Microsoft Office stesso, unico programma attualmente sul mercato capace di interpretare correttamente tale formato.

Il formato Microsoft Office (noto anche come doc) per anni si è dimostrato essere lo standard *de-facto* per l'automazione d'ufficio: si badi però che tale formato non è uno standard *de jure* poiché non è stato ratificato da nessun ente ufficiale né nazionale, né sovranazionale. Gli standard attuali per i documenti per l'automazione d'ufficio e i formati per l'archiviazione documentale sono rispettivamente ODF (Open Document Format for Office Applications – Standard ISO/IEC 26300:2006²) e PDF/a (Portable Document Format for Archiviatiion – Standard ISO 19005-1:2005³).

Attualmente il quadro normativo italiano invita le Pubbliche Amministrazioni a muoversi verso scelte che fanno uso di formati aperti⁴, andando in contrasto quindi con le scelte conservative rispetto al passato

Per la valutazione di vantaggi e problematiche nell'adozione del formato ODF nella PA rimando al documento redatto da Flavia Marzano “Formati aperti per la pubblica amministrazione: opportunità e rischi” reperibile online⁵.

Nelle prossime pagine verranno riportate delle citazioni da fonti reperibili online, segnalate alla fine di ogni paragrafo che consigliamo di consultare per approfondire meglio l'argomento.

1 <http://www.opensource.org/>

2 http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=43485

3 http://www.iso.org/iso/catalogue_detail?csnumber=38920

4 D.Lgs82/2005 noto come “Codice dell'Amministrazione Digitale” – Articolo 68 Comma 2

5 http://api.ning.com/files/YyiXBfR*BwBGgBlodALR*dGOyzikzwqqMOPclwExuHA_/F_Marzano_Formati_Aperti_V5.pdf

Il formato DOC

Aspetti negativi:

I file .doc contengono spesso altre informazioni oltre alla formattazione del testo, come ad esempio la lista delle operazioni annullate oppure script, che a loro volta sono un problema poiché possono rappresentare una minaccia alla sicurezza del sistema operativo Windows.

Il formato .doc è di proprietà di Microsoft che non ne ha rilasciato le specifiche di funzionamento. Nonostante ciò, è possibile creare e modificare documenti in questo formato anche con altri programmi, come il gratuito OpenOffice.org, grazie ad un lungo lavoro di re-ingegnerizzazione effettuato da parte dei programmatori; i file creati con OpenOffice o altri programmi non vengono però visualizzati esattamente allo stesso modo se aperti con il programma di Microsoft e viceversa, ma il problema è in parte colpa di quest'ultimo: ad esempio persino i file creati con le versioni 95 e 97 del programma ufficiale non sono completamente compatibili con le versioni 2000 e successive.

Fonte: <http://it.wikipedia.org/wiki/.doc>

[...]The format specification was available from Microsoft on request until February 2008, when the .DOC format specification was made available gratis under the Microsoft Open Specification Promise patent licensing.[...]

NdR: Il formato office quindi non è libero ed è vincolato ad un accordo tra l'utente e microsoft. Open Specification Promise non è una licenza ma una serie di promesse di microsoft relative al disclosure di certe tecnologie. Non contiene note su cosa puoi fare legalmente con il materiale rilasciato, ma solo cosa Microsoft non farà nei tuoi confronti.

Fonte: <http://en.wikipedia.org/wiki/.doc>

[...] Microsoft cambia il formato dei file di Word ad ogni versione, i suoi utenti sono intrappolati in un meccanismo che li obbliga ad acquistare ogni aggiornamento, che desiderino cambiare o meno. Potrebbero anche scoprire, tra qualche anno, che i documenti Word che scrivono oggi non sono più leggibili con la versione di Word che staranno usando allora. [...]

Fonte <http://www.gnu.org/philosophy/no-word-attachments.it.html>

Il formato doc non è retro-compatibile e aprirlo con versioni più recenti potrebbe non renderizzare il documento allo stesso modo. In più vengono applicate modifiche che rendono inaccessibile all'originale.⁶

Fonte: <http://www.goldmark.org/netrants/no-word/attach.html>

Forse non ve ne siete accorti ancora, ma il recente rilascio del service pack 3 per Microsoft Office 2003 contiene una "feature" nascosta — disabilita il supporto per le vecchie versioni di Microsoft Office. Se hai un qualsiasi vecchio documento Word, Excel, 1-2-3, Quattro, o Corel Draw in giro per il tuo hard disk devi intervenire nel registro di sistema di windows per premetterne l'apertura.

⁶ Ne sono stato coinvolto personalmente. Due Colleghi con Office 2000 e 2003: il primo scrive il documento, lo spedisce al secondo che lo modifica e lo salva con office 2003. Rispedito il documento indietro, non era più leggibile su Office 2000. E non si capiva perchè.

[...]

Non c'è una via semplice per identificare i file che hanno bisogno della conversione. Microsoft non fornisce nessun strumento o flag per identificare velocemente i vecchi file, che condividono lo stesso nome file con i recenti. Eccetto per il fatto di aprirli con Office 2003 e vedere che non funziona.[...]

Naturalmente, c'è un'alternativa che è qualcosa di più semplice (e gratuito): semplicemente utilizzare OpenOffice che può gestire i vecchi formati. Una volta aperti, potrebbe essere la volta buona di convertirli in formato ODF affinché Office 2017 non decida di nuovo di disabilitare il supporto ai vecchi formati.

Fonte: <http://blog.wired.com/monkeybites/2008/01/microsoft-offic.html>

[...]

La gente crea i suoi file per un futuro recupero e reimpiego.[...] Alla fine, chi pensi che sarà capace di leggere i tuoi documenti Word97 tra dieci anni? L'architettura del computer su cui il tuo Word attuale viene eseguito sarà obsoleta, il tuo computer di adesso si romperà e tu non sarai più capace di recuperare le informazioni che si trovano nei tuoi file.

I formati proprietari distruggono le tue informazioni.

[...]

Fonte: <http://www.podval.org/~sds/data.html>

La settimana scorsa, Microsoft ha pubblicato la documentazione sui formati file binari di Office. Questi formati sembrano essere completamente impazziti [En.: almost insane]. [...]

Un programmatore normale concluderebbe che i formati file binari di Office:

- sono deliberatamente offuscati
- sono prodotti dalla mente di un Borg (NdR: cyborg dell'universo fantascientifico di Star Trek) demente
- sono stati creati da programmatori di bassa qualità impazziti
- e sono **impossibili da leggere e creare correttamente**.

Fonte: <http://www.joelonsoftware.com/items/2008/02/19.html>

Motivi perché i documenti Word sono il male:

1. "Word" non è uno standard! I documenti word, (salvo certe eccezioni), possono solo essere letti da Microsoft Word; e talvolta nemmeno da quello. E' errato pensare che tutti abbiano Microsoft Word; non esiste per molti sistemi operativi, per tipi differenti di computer, vecchi computer, cellulari e palmari, etc.
2. I file sono dannatamente grandi. Non te ne accorgi se li trasferisci attraverso una rete locale o una linea a banda larga, ma dovresti comunque farci caso;[...]
3. Perché il formato di word è segreto, ed è difficile lavorarci. Per esempio, se volessi cercare tra le mie email quella che contiene la parola "hotel", "prenotazione" o "riferimento", probabilmente non la troverei perché non mi è possibile indicizzare i file .doc con il mio programma
4. I programmi Microsoft hanno un sacco di problemi di sicurezza che rendono quindi possibile la trasmissione di virus attraverso documenti Word.
5. Ci sono altri rischi di sicurezza: il mittente potrebbe erratamente inviare informazioni segrete che sono nascoste nel documento perché le informazioni di recupero del documento sono incluse nel documento stesso. (Non ho avuto conferma di questo ma dovrebbe essere

possibile in certi casi “recuperare” documenti totalmente non correlati dal file word che hai inviato come allegato)

[...]

Fonte: <http://www.cs.uit.no/~tobias/NEW/word-rejection.html>

Citazioni di normative e regolamenti nazionali e internazionali

Dal Quaderno n.21 del CNIPA:

2.2.2 FORMAZIONE DEI DOCUMENTI - ASPETTI DI SICUREZZA

In conformità a quanto disposto dalle norme vigenti in materia, le risorse strumentali e le procedure utilizzate per la formazione dei documenti informatici delle pubbliche amministrazioni devono garantire:

[...]

- la leggibilità dei documenti nel tempo;
- l'intercambiabilità dei documenti.

[...]

La normativa in materia prevede che i formati dei documenti della Pubblica Amministrazione “devono possedere almeno i seguenti requisiti:

[...]

- immutabilità nel tempo del contenuto e della sua struttura.

A tale fine i documenti informatici non devono contenere macroistruzioni o codice eseguibile, tali da attivare funzionalità che possano modificarne la struttura o il contenuto.

Commento personale:

Le condizioni elencate non possono essere garantite dal formato di Microsoft Office perché:

- la leggibilità nel tempo non è garantita perché l'unico titolare della tecnologia di decodifica completa e corretta di tale formato è Microsoft. Nel caso microsoft fallisca e decida di non pubblicare i formati, o semplicemente decida di chiudere definitivamente quel formato, non vi è possibilità di accedervi; come avete avuti l'opportunità di leggere in precedenza, qualcosa di simile è già stato fatto.
- il formato non è intercambiabile perché ci sono dei problemi di compatibilità tra versioni differenti dello stesso software;
- il documento se aperto con una versione più recente del software viene mutato nella sua visualizzazione rispetto all'originale;
- non è possibile eliminare la parte di macroistruzioni all'interno del documento Microsoft Office.

Fonte: http://www.cnipa.gov.it/site/_files/Quaderno%20n%2021.pdf

La Legge 4/2004, al Comma 1 dell'Art. 1 (Obiettivi e finalità) recita “La Repubblica riconosce e tutela il diritto di ogni persona ad accedere a tutte le fonti di informazione e ai relativi servizi, ivi compresi quelli che si articolano attraverso gli strumenti informatici e telematici.” [...]

La PA [...] ha la necessità che,

indipendentemente dallo strumento hardware o software utilizzato, i documenti siano:

- leggibili da tutti (cittadini, imprese, altre amministrazioni,...);
- leggibili con qualunque strumento;
- leggibili nel tempo;
- conservabili nel tempo (come anche previsto dagli Artt. 43 e 44 del Codice dell'Amministrazione Digitale¹, capo VIII, Sezione 1, Art. 72, d'ora in poi CAD).

Per poter garantire la fattibilità di quanto sopra espresso è necessario che i documenti e i

formati dei dati rispondano ad alcune caratteristiche tecniche e in particolare che siano aperti. [...]

La Commissione Europea ha definito gli standard aperti nel documento “*European Interoperability Framework*” richiedendo in particolare che abbiano alcune caratteristiche comuni:

- gli standard essere adottati e mantenuti da un’organizzazione “no-profit”;
- essi devono essere sviluppati utilizzando processi decisionali aperti e accessibili a tutte le parti interessate e le cui decisioni vengono prese per consenso o a maggioranza;
- il documento di specifiche degli standard pubblicati deve essere disponibile gratuitamente o a un costo nominale, e deve essere possibile a chiunque copiarlo, riusarlo e distribuirlo liberamente senza costi o a costo nominale;
- eventuali diritti di *copyright*, brevetti o marchi registrati sono irrevocabilmente disponibili senza alcun pagamento di licenze e/o diritti;
- non ci devono essere vincoli sul riuso, alla modifica e all’estensione degli standard.

A queste caratteristiche la Commissione Europea aggiunge alcune **raccomandazioni** sulle garanzie che devono offrire i servizi di eGovernment:

- Accessibilità
- Multilinguismo
- Sicurezza
- Privacy
- Sussidiarietà
- Uso di Open Standard
- Valutare i benefici del software Open Source
- Uso di soluzioni multilaterali

Fonte:

http://api.ning.com/files/YyiXBfR*BwBGgBlodALR*dGOyzikzwqqMQPclwExuHA_/F_Marzano_Formati_Aperti_V5.pdf

Critiche relative ai formati non standard nella PA firmati digitalmente

[...] Quale nazione affiderebbe ad una azienda business oriented, per quanto potente e stabile, la conservazione della propria memoria? [...]

La firma digitale, infatti, oltre a non essere una firma - ma un marchio, un contrassegno, un sigillo informatico - non migra da una generazione di tecnologia ad un'altra. In altre parole essa è il primo strumento a soffrire dell'altra faccia del progresso: l'obsolescenza tecnologica. Sappiamo infatti che il cambiamento anche di un solo bit nella sequenza del file firmato produce come conseguenza la invalidità della firma. Pertanto, un file prodotto con MS Word 5.5 con il sistema operativo DOS 3.3 non conserva intatto il flusso originario di bit in ambiente Windows XP, ma viene modificato con le conseguenze appena descritte. [...]

Fonte: http://www.interlex.it/docdigit/cnipa_pdf.htm

[...] Per esempio se abbiamo sottoscritto, utilizzando una busta PKCS#7, un file generato con un prodotto di videoscrittura in ambiente MS-DOS di qualche anno fa, saremo ancora di verificare la firma (la rappresentazione della struttura dati firmata è binaria) ma per la lettura del file sottoscritto dovremo disporre dello specifico prodotto che la ha generato. Per quanto appena detto è

indispensabile che i formati dei documenti informatici prodotti siano "aperti" ovvero pubblicamente documentati.[...]

Fonte: http://www.interlex.it/docdigit/manca_adobe.htm

Considerazioni di sicurezza sul software Microsoft Office

In questa sezione verranno messi a confronto Microsoft Office con l'attuale suo principale concorrente opensource, Open Office il quale fa uso del formato standard ODF.

Da un ricerca nel sito “milw0rm - Exploit database separated by exploit type (local, remote, DoS, etc.)” per la voce “Microsoft Office”:

Tipi	[exploits/shellcode]
--:DATE	--:DESCRIPTION
28/11/08	Microsoft Office Communicator (SIP) Remote Denial of Service Exploit
30/03/08	Microsoft Office XP SP3 PPT File Buffer Overflow Exploit (ms08-016)
21/03/08	Microsoft Office Excel Code Execution Exploit (MS08-014)
13/02/08	Microsoft Office .WPS File Stack Overflow Exploit (MS08-011)
13/06/07	Microsoft Office MSODataSourceControl COM-object BoF PoC (0day)
23/05/07	Microsoft Office 2000 (OUACTRL.OCX v. 1.0.1.9) Remote DoS Exploit
09/04/07	Microsoft Office Word 2007 Multiple Vulnerabilities
12/10/06	Microsoft Office 2003 PPT Local Buffer Overflow PoC

Mentre per la voce “Open Office”

	[exploits/shellcode]
--:DATE	--:DESCRIPTION
10/05/08	Open Office.org 2.31 swriter Local Code Execution Exploit

Glossario:

- **DoS, Denial of Service:** Negazione di servizio; In questo tipo di attacco si cerca di portare il funzionamento di un sistema informatico che fornisce un servizio, ad esempio un sito web, al limite delle prestazioni, lavorando su uno dei parametri d'ingresso, fino a renderlo non più in grado di erogare il servizio. **Fonte:** http://it.wikipedia.org/wiki/Denial_of_Service
- **BoF, Buffer Overflow, Stack Overflow:** vulnerabilità di sicurezza che può affliggere un programma software. [...] tale programma non controlla in anticipo la lunghezza dei dati in arrivo, ma si limita a scrivere il loro valore in un buffer di lunghezza prestabilita, confidando che l'utente (o il mittente) non immetta più dati di quanti esso ne possa contenere.[...] Conoscendo molto bene il programma in questione, il sistema operativo e il tipo di computer su cui gira, si può precalcolare una serie di dati malevoli che inviata per provocare un buffer overflow consenta ad un malintenzionato di prendere il controllo del programma (e a volte, tramite questo, dell'intero computer).[...] Lo stack overflow consiste ugualmente nella

sovrascrittura dell'area dati del programma, ma questa volta la causa è l'attività del programma stesso: chiamando con dei parametri particolari una funzione ricorsiva del programma, questa accumula chiamate in sospeso sullo stack fino a riempirlo completamente e inizia a sovrascrivere la memoria vicina. **Fonte:** http://it.wikipedia.org/wiki/Buffer_overflow

- **Code Execution:** l'abilità di un attaccante di eseguire un qualsiasi comando di sua scelta nella macchina obiettivo o nel processo obiettivo **Fonte:** http://en.wikipedia.org/wiki/Arbitrary_code_execution

Fonte: <http://www.milw0rm.com/>

Secunia, provider di intelligence sulle vulnerabilità informatiche nel mondo, permette attraverso il suo sito di venire a conoscenza dei problemi di sicurezza dei prodotti software attualmente sul mercato. Da una ricerca nel loro database risulta che:

- Microsoft Office 2007 ha ricevuto segnalazioni per 74 Vulnerabilità in 19 bollettini Secunia, di cui nessuno è attualmente segnalato come “Non Corrette (Unpatched)”
- Microsoft Office 2003 Professional ha ricevuto segnalazioni per 153 Vulnerabilità in 60 bollettini Secunia dei quali il 7% sono segnalati come “Non Corrette (Unpatched)”
- Microsoft Office 2000 ha ricevuto segnalazioni per 141 Vulnerabilità in 61 bollettini Secunia dei quali il 7% sono segnalati come “Non Corrette (Unpatched)”
- OpenOffice 2.x ha ricevuto segnalazioni per 31 Vulnerabilità in 11 bollettini Secunia dei quali nessuno è attualmente segnalato come “Non Corrette (Unpatched)”
- OpenOffice 3.x non ha ricevuto segnalazioni di Vulnerabilità.

Fonti: <http://secunia.com/advisories/product/13228/> - <http://secunia.com/advisories/product/2276/> - <http://secunia.com/advisories/product/24/> - <http://secunia.com/advisories/product/6157/> - <http://secunia.com/advisories/product/20130/>

Conclusioni

Una scelta oculata sotto il profilo della sicurezza e dell'interoperabilità permetterebbe di ridurre i problemi che affliggono attualmente i formati proprietari. Nel modo specifico scegliere di abbandonare il formato doc può essere una scelta di notevole portata che porterà a lungo termine miglie e maggiore sicurezza nell'ambito delle attività di automazione d'ufficio. La diminuzione della superficie d'attacco per malintenzionati e l'indipendenza dal fornitore crea l'opportunità di sviluppare un nuovo modo di lavorare, in cui aumenta la libertà di operare dell'azienda, sia sul piano della sicurezza su cui si è maggiormente tutelati, sia sul piano della gestione degli investimenti che potranno essere mirati verso molteplici fornitori anziché verso un fornitore unico⁷.

La mancanza di libertà nel formato fa sì che si dipenda da un'azienda privata (che ha delle politiche commerciali di cui non siamo a conoscenza e alle quali non possiamo opporci) impedendoci di scegliere liberamente come investire le nostre risorse.

Siamo già vincolati abbastanza per certi aspetti nella nostra società... Perché, come dice il filosofo Silvano Agosti⁸, dobbiamo “mettere i fiorellini alla finestra della cella dove siamo prigionieri” quando avremmo l'opportunità di uscire dalla cella perché la porta è aperta?

7 In “Revolution OS II. Software libero, proprietà intellettuale, cultura e politica.” di Arturo Di Corinto, Apogeo 2006 viene affrontato questo tema dei vantaggi degli investimenti verso i prodotti Open Source

8 Silvano Agosti – Il discorso tipico dello schiavo