



Office Open XML: la storia e i segreti di uno standard che non è uno standard

Il Decreto del Presidente del Consiglio dei Ministri del 13 novembre 2014 (Regole tecniche in materia di formazione, trasmissione, copia, duplicazione, riproduzione e validazione temporale dei documenti informatici, nonché di formazione e conservazione dei documenti informatici delle pubbliche amministrazioni, ai sensi degli articoli 20, 22, 23 bis, 23 ter, 40, comma 1, 41, e 71, comma 1, del Codice dell'Amministrazione Digitale, di cui al decreto legislativo n. 82 del 2005) inserisce tra i formati indicati per la conservazione (Punto 5 dell'Allegato 2) anche Office Open XML (OOXML), il formato di Microsoft Office dalla versione 2007 in avanti.

Il formato OOXML, nonostante l'approvazione da parte di ISO, non possiede le caratteristiche di un formato standard, come conferma la recente decisione del Governo del Regno Unito, che ha deciso che solo il formato Open Document Format (ODF) risponde alle esigenze di una pubblica amministrazione attenta alle esigenze dei cittadini.

La decisione del Governo del Regno Unito è avvenuta al termine di un processo che è durato due anni, e ha coinvolto esperti di formati standard di altri governi, di aziende di information technology e di aziende utenti, di enti e organizzazioni, ed esperti indipendenti, a livello soprattutto europeo.

Durante questo processo, sono emersi dei problemi significativi - in relazione allo status di standard ISO - di cui è facile comprendere la genesi analizzando il modo in cui il formato Office Open XML è nato ed è stato standardizzato.

I problemi che fanno di Office Open XML uno standard ISO che nella realtà non è uno standard afferiscono a quattro aree principali, e confermano il fatto che OOXML è un formato proprietario mascherato, e - allo stato attuale delle cose - non sarà mai un formato standard.

Il processo di standardizzazione

La prima area è quella del processo di standardizzazione. Nel caso di OOXML, è stato adottato il processo breve o "fast track", inadeguato per un formato complesso come quello dei documenti da ufficio. Secondo questo processo, che



ECMA (European Computer Manufacturer's Association) offre come servizio a pagamento ai propri associati, in quanto - come ente di standardizzazione - è in grado di eseguire la due diligence sul formato stesso prima della presentazione a ISO, la revisione tecnica delle 7.200 pagine della descrizione del formato OOXML è stata fatta in modo sintetico, in soli 838 giorni. Peccato che, nel caso di OOXML, sia poi emerso - a cose fatte - che ECMA non ha effettuato l'analisi del formato, "fidandosi" del lavoro svolto da Microsoft.

Di contro, la revisione tecnica delle 720 pagine della descrizione del formato ODF - che ha seguito la procedura normale - ha richiesto 1239 giorni, e ha prodotto uno standard che risponde a tutti i criteri.

Tra l'altro, il processo "fast track" è sfociato, al termine del Ballot Resolution Meeting di Ginevra (un meeting di cinque giorni, durante il quale sono stati affrontati - ma non risolti - i circa 3.500 problemi emersi durante l'analisi del formato OOXML da parte dei comitati nazionali di standardizzazione), in un compromesso dovuto al fatto che il gruppo di esperti si è diviso tra quelli favorevoli e quelli assolutamente contrari alla standardizzazione. La decisione, nell'impossibilità di approvare come standard quella versione del formato, è stata di approvare una versione transitoria, chiamata OOXML Transitional, nell'attesa della versione successiva del formato - definita OOXML Strict - perfettamente compatibile con la definizione dello standard.

Per quanto riguarda il processo di analisi all'interno dei comitati nazionali, c'è un'ampia letteratura che dimostra come in molti casi ci siano stati eventi poco chiari che hanno inquinato in modo irreparabile la decisione finale. In Italia, la commissione UNINFO è passata in pochi mesi da 5 a 83 membri, tra i quali una delle agenzie di relazioni pubbliche e due tra gli studi legali - uno dei quali con specializzazione in diritto del lavoro - di Microsoft, portando il voto da negativo a neutrale.

OOXML Transitional, invece di essere il formato transitorio di compromesso, è diventato il formato standard di Microsoft Office 2007, 2010 e 2013, in tre diverse versioni di volta in volta incompatibili rispetto alla definizione dello standard, per motivi ogni volta diversi. OOXML Strict è comparso con Microsoft Office 2013, come ultima tra le opzioni di salvataggio (una posizione davvero strana, per un formato standard), e con significativi problemi, che vedremo più avanti.



L'ente che amministra lo standard OOXML

La seconda area è quella dell'ente che amministra lo standard, che nel caso di OOXML è ECMA (European Computer Manufacturer's Association). Secondo le regole ISO, questo dovrebbe essere indipendente rispetto all'azienda che propone il formato standard, che in questo caso è Microsoft.

Ebbene, un gran numero dei membri ECMA intrattiene rapporti commerciali con Microsoft, in quanto rivende la versione OEM del sistema operativo Windows, o addirittura l'applicativo Microsoft Office. Due condizioni che compromettono l'indipendenza di ECMA rispetto a Microsoft.

Quindi, il fatto che la due diligence sia stata fatta in modo approssimativo, come dimostra il numero di incompatibilità con le regole che governano gli standard, è una conseguenza della non indipendenza di ECMA rispetto a Microsoft.

A questo si aggiunge il fatto che ECMA ha dei processi interni, relativi anche all'area degli standard, che contraddicono le regole della trasparenza richieste per lo sviluppo di uno standard, in quanto i membri, le mailing list e i commenti non sono accessibili, così come le agende e le minute dei meeting.

Di contro, OASIS - il consorzio che amministra lo standard ODF - è composto da un insieme eterogeneo di software vendor (tra cui la stessa Microsoft), aziende ed enti pubblici utenti dei software, organizzazioni not for profit e singoli individui, che non hanno nessun rapporto commerciale con le fondazioni che propongono ODF, ovvero Apache Foundation e The Document Foundation.

Inoltre, OASIS pubblica in modo trasparente l'elenco dei membri, le mailing list, i commenti e le minute dei meeting, per cui il processo di sviluppo dello standard è pubblico.

Il riuso degli standard esistenti

La terza area è quella del riuso di standard esistenti, che ha un impatto diretto sulle dimensioni della descrizione dei due standard (e anche sulla loro fruibilità). OOXML riutilizza solo lo standard Dublin Core, mentre ODF riutilizza sia Dublin Core sia altri standard: XLS:FO, SVG, MathML, Xlink, SMIL e XForms. In tutti questi casi, OOXML introduce formati di provenienza



Microsoft che "gonfiano" la descrizione del formato stesso.

L'aspetto del riuso, ovviamente, ha un impatto più ampio, in quanto costringe a implementare le tecnologie Microsoft nei filtri di importazione ed esportazione, in luogo delle librerie già esistenti - e collaudate - per la gestione degli standard, con un impatto diretto sull'interoperabilità (più sono gli standard, più sono le librerie standard, maggiore è la compatibilità dei formati).

Le suite di produttività scrivono un formato in modo nativo, e gli altri formati per mezzo di filtri di esportazione e importazione che utilizzano delle librerie software per la conversione. Maggiore è il numero delle librerie che riusano uno standard, minore è il numero di librerie che devono essere scritte ex novo, e di conseguenza minore è il numero dei problemi di interoperabilità che si possono verificare nel documento finale.

Un formato standard dovrebbe essere concepito per favorire, e non per rendere più difficile l'interoperabilità da parte degli altri software, in quanto l'interoperabilità è il motivo alla base dell'esistenza degli standard. Nel caso di OOXML, sembra che il formato sia stato concepito per rendere più difficile l'interoperabilità, e questo contraddice lo spirito degli standard.

Le caratteristiche del formato OOXML

La quarta area è quella delle caratteristiche del formato, e delle modalità di gestione del formato stesso da parte dell'implementazione di riferimento, ovvero di Microsoft Office.

Prima di tutto, ogni versione di Microsoft Office, a partire da Microsoft Office 2007, gestisce come formato di default una versione diversa - e non standard - di OOXML Transitional, senza fornire all'utente gli elementi necessari per sapere di quale versione si tratta, in quanto l'estensione del file è sempre la stessa. Solo Microsoft Office 2013 è in grado di scrivere e leggere il formato OOXML Strict, che però non viene gestito nel modo tradizionale, in quanto la produzione del formato standard viene messa nelle mani dell'utente.

Questo è un aspetto particolarmente critico, che merita un approfondimento. Infatti, il comportamento di Microsoft Office 2013 contraddice la consuetudine, secondo la quale la responsabilità del formato standard viene affidata al



software (che nella norma scrive il formato nativo standard, a meno che l'utente decida di scegliere un altro formato).

Nel caso di Microsoft Office 2013, il software scrive sempre e solamente il formato non standard a meno che l'utente faccia qualcosa che è contrario a ogni logica, ovvero salvi nel formato OOXML Strict all'apertura del nuovo file. Infatti, l'inserimento di un solo carattere prima del salvataggio crea - nella memoria di lavoro del personal computer - un file in formato OOXML Transitional, che a sua volta crea una serie di errori "random" nel caso in cui il documento venga salvato in formato OOXML Strict.

Si tratta di un comportamento talmente inusuale da essere poco comprensibile, e sicuramente difficile da spiegare in modo chiaro. Fortunatamente, è molto facile da riprodurre, per cui bastano pochi passaggi per realizzare la portata di quello che sembra un bug ma che nella realtà è un comportamento probabilmente voluto.

Sempre nell'area delle caratteristiche del formato, ci sono altri tre problemi rilevanti: l'assenza di un'implementazione libera del formato OOXML, che forza all'uso di Microsoft Office per qualsiasi tipo di attività di sviluppo (la presenza di un'implementazione libera dovrebbe essere una condizione imprescindibile per un formato standard, perché altrimenti si genera una forma di lock-in addirittura in fase di sviluppo); la presenza di elementi proprietari "incapsulati", che dovrebbero essere banditi dai formati standard; e la presenza di bug "storici" come quello che rende OOXML Transitional incompatibile con il Calendario Gregoriano.

In particolare, due dei problemi evidenziati, il bug dell'anno bisestile (che impedisce la compatibilità con il Calendario Gregoriano) e il comportamento in fase di scrittura dei file (che privilegia il formato non standard rispetto a quello standard), lasciano presupporre che il rispetto dello standard OOXML da parte di Microsoft Office, indipendentemente dalla versione dello standard stesso, possa avvenire solamente a fronte di una riscrittura di porzioni significative sicuramente di Excel e probabilmente anche di Word e PowerPoint.

Conclusioni

Riteniamo quindi sbagliato che il Decreto del Presidente del Consiglio dei

WHITE PAPER



Ministri del 13 novembre 2014 inserisca Office Open XML tra i formati per la conservazione dei documenti, in quanto - non essendo standard - perpetua il lock-in dei documenti all'applicazione Microsoft Office, e costringe le pubbliche amministrazioni all'acquisto di un software proprietario in aperta contravvenzione al dettato dell'Articolo 68 del Codice dell'Amministrazione Digitale.

Questo, ovviamente, può derivare dal fatto che le informazioni sullo standard siano state - in larga parte - fornite unilateralmente dal vendor, che ha un interesse commerciale verso la presenza di Office Open XML nell'elenco dei formati per la conservazione, e forse da una incompleta conoscenza dei principi, della storia e dei meccanismi di standardizzazione nel dominio specifico.

Eppure, i vantaggi che derivano dall'uso di un formato standard sono chiari, e sono la possibilità di interoperare - ovvero, di lavorare meglio insieme, in modo trasparente - senza pensare all'interoperabilità, perché ci pensa il software attraverso il formato standard - scambiando documenti che rispecchiano sempre l'originale (perché uno standard è uno standard).

Il Governo del Regno Unito, che ha compreso come solo un vero standard può garantire l'interoperabilità, dopo una lunga consultazione - in cui non sono mancati i tentativi di forzare la decisione, agendo direttamente sugli uomini politici - ha deciso a favore di ODF.

La presenza dello standard "non standard" OOXML all'interno del documento pubblicato dalla Presidenza del Consiglio rappresenta un enorme problema per la comunità (sistema Paese, istituzioni, pubblica amministrazione, aziende e singoli cittadini), perché un'interoperabilità limitata o addirittura assente - oltre a essere una forma estremamente sofisticata di digital divide - rappresenta un costo per tutti, e un ostacolo alla crescita della cultura digitale in Italia.

Quest'opera è soggetta alla licenza Creative Commons Attribuzione - Condividi allo stesso modo 3.0 (<http://creativecommons.org/licenses/by-sa/3.0/it/deed.it>).

