



UNIVERSITA' DEGLI STUDI DI MACERATA

**Facoltà di Lettere e Filosofia
Facoltà di Economia
Facoltà di Beni Culturali**

*Master in
"Formazione, gestione e conservazione di archivi digitali
in ambito pubblico e privato"*

III^a edizione Anno Accademico 2009-10

L'interoperabilità dei sistemi di protocollo all'interno delle P.A.:
Quale grammatica per il formato XML?

Corsista : Dott. Daniele Russo

Relatore: Chiar.mo Prof. Stefano Allegrezza

Sommario

INTEROPERABILITÀ: COMUNICARE ATTRAVERSO IL MEDESIMO LINGUAGGIO.....	- 3 -
STATUS QUESTIONIS.....	- 3 -
QUALI METADATI?	- 5 -
<i>Informazione minima e accordi bilaterali</i>	- 5 -
<i>L'esperienza presso il Comune di Osimo</i>	- 7 -
PAROLA D'ORDINE "REENGINEERING"	- 9 -
LE CRITICITÀ DEL "VERSAMENTO" ALL'INTERNO DEL MODELLO OAIS.....	- 9 -
IL FORMATO XML.....	- 11 -
LA FUNZIONE RICOPERTA DALLA "GRAMMATICA" ALL'INTERNO DI UN DOCUMENTO XML.....	- 14 -
LIMITI E PUNTI DI FORZA DELLE GRAMMATICHE "DTD" E "XML SCHEMA"	- 16 -
CONCLUSIONI	- 19 -
BIBLIOGRAFIA	- 21 -

Interoperabilità: comunicare attraverso il medesimo linguaggio

Status Questionis

Con la delibera AIPA 7 Maggio 2001 n.28 si sono poste le basi per la creazione di uno standard gestionale che permetterà sotto il profilo operativo di mettere in comunicazione e far dialogare tra loro i software relativi ai protocolli informatici delle P.A. ovvero «*la possibilità di trattamento automatico da parte di un sistema di protocollo ricevente, delle informazioni trasmesse da un sistema di protocollo mittente*»¹ Attraverso una rigorosa disciplina tecnica², infatti, sarà possibile trasmettere e condividere non solo la documentazione registrata al protocollo, ma anche tutto un set di metadati ad essa riferita: ovvero pacchetti informativi di diversa natura contenenti dati inerenti al documento di riferimento che permettono e permetteranno di fruire anche in futuro di diverse tipologie di informazioni, categorizzate secondo schemi predefiniti e definibili in corso d'opera, che vengono raccolte di volta, in volta dagli operatori nell'atto di registrazione.

Un modello rivoluzionario che permetterà di comunicare mediante un unico linguaggio e che sicuramente riuscirà ad imporsi nel corso degli anni e a distinguersi per efficienza, velocità, risparmio economico. Allo stato attuale – purtroppo – pur trovando da più settori della P.A. un terreno comune, fatto di esigenze condivise, la fase attuativa per la realizzazione di tale progetto rimane parziale, incerta e – per molte realtà – si è del tutto arenata se non addirittura accantonata.

Abbiamo rintracciato le ragioni di questo gap in alcune situazioni sfavorevoli:

da una parte riscontriamo che la decentralizzazione e la delocalizzazione delle singole Amministrazioni italiane sfavorisce una standardizzazione delle procedure di lavoro³; l'autonomia regionale che grazie alle politiche di tipo federalista, in attuazione dell'art. 117 della Costituzione, ci mostrano un quadro normativo e burocratico piuttosto eterogeneo; infine è da ricordare come quanto detto sopra possa rappresentare un aspetto del tutto marginale rispetto al colpevole mancato stanziamento di fondi da parte dello Stato⁴, per finanziare e sostenere in modo adeguato un simile cambiamento.

A dispetto di un quadro normativo piuttosto “ambizioso”, figlio di quella *pioggia normativa*⁵ che dagli anni novanta ad oggi ha fatto da sostegno al passaggio all'era digitale, non hanno fatto seguito – “*in opere operato*” – quei provvedimenti, sotto il profilo pratico, che permetterebbero alle singole amministrazioni di uniformarsi ad un sistema unitario e standard per la P. A. Pertanto ci troviamo di fronte a situazioni completamente disomogenee tra le diverse realtà

¹ Delibera AIPA 7 Maggio 2001 n. 28

² Cfr. “Il formato XML”

³ Il consolidamento di prassi e consuetudini a valore locale a dispetto della regolare applicazione delle procedure standard è cosa nota; il fatto si amplifica quando ciò riguarda piccoli paesini di pochi abitanti con un'amministrazione che non riesce a colmare il volume di lavoro e si deve affidare a personale carente, inesperto, con poca professionalità.

⁴ L'art. 14 del D.Lgs. 82/2005, noto come Codice per Amministrazione Digitale, al comma 2 affida alla Conferenza Unificata di Stato, Regioni, e autonomie locali il compito di promuovere il «*processo di digitalizzazione dell'azione amministrativa coordinato e condiviso e per l'individuazione delle regole tecniche di cui all'articolo 71*» e ciò attuando tutta una serie di provvedimenti atti a prevenire «*il divario tecnologico tra amministrazioni di diversa dimensione e collocazione territoriale*»; tuttavia al successivo comma 3-bis viene precisato che tale politica innovativa e rivoluzionaria dovrà avvenire «*senza nuovi o maggiori oneri per la finanza pubblica*».

⁵

della nostra penisola: dal punto di vista delle conoscenze, delle professionalità, delle soluzioni adottate (soprattutto quelle in ambito informatico), delle metodologie applicate.

Ciò che ci proponiamo di analizzare in questo breve studio è innanzitutto di verificare se esiste la possibilità di trovare possibili soluzioni concrete che permetterebbero alle P.A. di dialogare tra loro superando tutte quelle difficoltà che contraddistinguono la vita di una P.A., e che abbiamo sopra riportato; secondariamente di aprire una finestra su ambito tecnico ancora un po' oscuro e cercare di intuire se sia possibile intravedere anche come tale interoperabilità possa avvenire.

Quali Metadati?

Informazione minima e accordi bilaterali

Ogni documento inviato o conservato all'interno dell'archivio deve essere provvisto di metadati, ovvero di tutte quelle informazioni che, oltre ad identificarlo dettagliatamente (nome, autore, eventuali codici), sono in grado di garantirne il reperimento nel breve e nel lungo periodo. In funzione della tipologia documentaria, del periodo di conservazione necessaria, sarà ragionevole adottare tipologie di metadati differenti. Non ci dilungheremo ad enumerare le diverse soluzioni che l'archivistica mette a disposizione, al contrario in questo studio ci limiteremo a distinguere i metadati di base che riguardano l'informazione minima necessaria che ineriscono a tutti i documenti senza alcuna discriminazione; e i metadati che forniscono informazioni aggiuntive, che si inseriscono cioè, in un quadro di esigenze specifiche da parte dell'Ente produttore del documento o dell'Ente deputato alla gestione/conservazione dello stesso. All'interno di quest'ultima categoria si inserisce da un parte il concetto di "conservazione a lungo termine": ovvero quella gestione del documento e delle informazioni ad esso pertinenti che comprendono la conservazione degli stessi all'interno di un lasso temporale tale da interessare cambiamenti tecnologici; dall'altra parte esigenze specifiche che inducono gli Enti interessati a prendere accordi bilaterali in vista di un medesimo obiettivo.

In particolare, sarà opportuno costituire un gruppo di lavoro che si accordi – prima che le operazioni di versamento abbiano inizio – sul tipo di documentazione e soprattutto su quali metadati gli istituti interessati intendono conservare. Stipulare accordi – anche sottoforma di contratti in cui le parti creino un'intesa e decidano quale tipo di informazione possa essere privilegiata e quale omessa, in virtù degli obiettivi per i quali nasce l'esigenza della conservazione, alla luce di un futura fruizione da parte di un pubblico definito.

E' questo il caso dell'archivio OAI⁶ che stipula con gli interessanti contratti che stabiliscono minuziosamente i metadati di cui le varie parti in causa si fanno carico.

Nel caso di una P.A. oltre al ruolo di garante che l'istituto di conservazione deve svolgere, l'attenzione deve focalizzarsi sulle procedure relative al versamento della documentazione. Bisogna intendersi sul tipo di documento che dovrà essere conservato; in che formato; quale versione; bisognerà scansionare tutta la procedura relativa al procedimento amministrativo e dirigere l'attenzione su ciascuna fase ed intendersi su quali metadati ciascun ufficio, ciascun A.O.O. ha preso in carico, per poi decidere quali metadati saranno necessari per ciascun tipo di documento, tenendo conto della legislazione di riferimento e cercando di intuire quelli che potranno essere gli sviluppi futuri.

Questo è un tasto dolente: infatti non possiamo sapere cosa le generazioni future riterranno importante ricordare, o il livello di preparazione dell'utenza che si troverà a fruire di tali informazioni: ad un bambino di quattro anni – per esempio – sarà sufficiente sapere se la foto che ha di fronte è quella di un animale; ad uno di cinque dovremmo aggiungere il nome dell'animale; a quello di otto dovremmo descrivere la famiglia di appartenenza dell'animale..

Nel restituire l'informazione dobbiamo tenere conto il tipo di pubblico a cui è destinata. «...on doit connaître la Base de connaissance de cette communauté afin d'établir quelle est l'information de représentation minimale à conserver»⁷

Inoltre, quando si tratta di materiale destinato alla conservazione a lungo termine, bisogna tenere in considerazione cambiamenti relativi a quella che potremmo chiamare sensibilità epocale: le operazioni di selezione del passato in merito al materiale archivistico conservato ci la-

⁶ OAI (Open Archives Information System), <<http://public.ccsds.org/publications/archive/650x0b1.pdf>>

⁷ Modèle de référence pour un Système ouvert d'archivage d'information (OAI), Marzo 2005 CCSDS <<http://public.ccsds.org/publications/archive/650x0b1.pdf>>, ENVIRONNEMENT OAI

sciano intuire che la sensibilità umana è suscettibile di cambiamento da un'epoca all'altra. Ciò non è diverso per quanto riguarda quelle informazioni, quei metadati che gli uomini ritengono opportuno riportare alla memoria. Ad esempio fino agli anni settanta del secolo scorso – non essendo ancora diffuso l'utilizzo del PC – non avrebbe avuto molto senso prendere nota di un formato di un file e riportarlo come metadato di un documento; oggi avremmo l'imbarazzo di dover aprire lo stesso file con l'applicativo idoneo.

L'esperienza presso il Comune di Osimo

Ho voluto portare un contributo attraverso un progetto di archiviazione per ciò che concerne il versamento di alcune tipologie documentarie contenute all'interno della Regione. Onestamente l'entusiasmo è venuto meno quando mi scontrai con delle problematiche oggettive di non facile soluzione. In particolare quando venni a sapere che un Comune di trentamila abitanti, come è quello di Osimo disponeva di appena due operatori al protocollo e che, ricevendo dalle trecento alle quattrocento richieste giornaliere, in fase di registrazione, non era ritenuto possibile effettuare altra operazione se non quella di inserire metadati essenziali quali la data, l'oggetto, il mittente e il destinatario. Tutta la questione importante relativa all'archiviazione previo classificazione (definita dall'operatore "intabellinatura") è omessa dal novero delle procedure. Lo stesso dicasi per ciò che concerne l'invio di e-mail, sia quelle certificate che quelle ordinarie, destinate a rimanere nell'oblio. Di ciò ne parlo per esperienza diretta quando – al fine di inoltrare richiesta di accesso ai locali del Comune per fini di studio – inviai una e-mail attraverso il servizio CEC- PAC, pertanto certificata da Ente accreditato (Poste Italiane): tuttora attendo risposta a quella e-mail (sebbene fosse stata inviata ad un indirizzo istituzionale promosso all'interno del portale del Comune di Osimo). Successivamente vidi stampare l'e-mail e consegnarla "brevi manu" all'ufficio responsabile.

Sudette criticità non vengono riportate al fine di denunciare un disservizio, o per polemizzare con un Amministrazione che difficilmente potrebbe avere colpe di una situazione di ristrettezza, anche economica, nella quale si trova in qualche modo... a dover improvvisare; diversamente ho voluto descrivere in modo concreto e per auto testimonianza quanto sostenevamo sopra⁸ riguardo alla disomogeneità della situazione italiana in fatto adeguamento al digitale e alle ICT e, soprattutto alla mancanza di reale sostegno da parte dell'Amministrazione Centrale dello Stato.

A tale proposito, sono stati gli stessi operatori del Comune a denunciare le loro difficoltà in tal senso, sostenendo di avere attrezzature adeguate, ma che riescono ad utilizzarle al dieci per cento delle loro funzionalità. E' questo il caso del "Paleo 2": un software per la protocollazione informatica che il Comune di Osimo ha la fortuna di condividere con la Regione Marche e presso la quale dovrà, poi, versare alcune tipologie documentarie, ma per le quali oggi non dispone delle competenze e delle professionalità idonee per predisporre il sistema ad una corretta e completa funzionalità.

Mentre mi concentravo sui punti focali del mio progetto, un aspetto apparentemente aleatorio e marginale attirò la mia attenzione e mi pungolò a tal punto, che mi avrebbe condotto ad approfondire la tematica dei linguaggi software nelle P.A. In effetti notavo che la registrazione a protocollo prevedeva il riempimento non già di campi predefiniti che potessero aumentare la velocità delle registrazioni, grazie ad una puntuale automazione di alcuni processi di base, e impedire errori accidentali relativi alle operazioni di data-entry; ma di campi a stringa di testo libero. Tale sistema risulta essere piuttosto precario ed è consigliabile per quei campi discorsivi (ad esempio le note)⁹ che non devono essere soggetti a filtri di alcun tipo. Nella stringa a testo libero la precisione e correttezza delle registrazioni è vincolata all'attenzione e alla preparazione del responsabile del protocollo. L'errore di battitura – sempre in agguato e fortemente condizionato dal copioso numero di registrazioni giornaliere cui è soggetto un Comune che deve rispondere per più di trenta mila abitanti – preclude di fatto l'accesso a quei documenti che non corrispondono a criteri selezionati. Possiamo solo immaginare a che tipo di fatica dovrebbe sottoporsi uno storico che dovesse valutare dati statistici o – peggio – un giuri-

⁸ Cfr. "Status Questionis"

⁹ Cfr. quanto riportiamo in merito alla grammatica DTD ("Limiti e punti di forza delle Grammatiche "DTD" e "XML Schema").

sta che cercasse – privo di riferimenti temporali precisi – elementi comprovanti atti o fatti di qualunque genere.

Osservando analiticamente il campo “oggetto” constatiamo che esso, consistendo in una sintesi sommaria del documento che andiamo a registrare, dovrebbe permettere con uno sguardo sinottico, di intuire il contenuto del documento allorché ci prepariamo ad una ricerca. In assenza di riferimenti temporali certi saremo costretti leggere tutte le registrazioni relative al campo “oggetto” poste all'interno del lasso temporale interessato dalla nostra ricerca. Tale operazione può essere anche corretta all'interno di un numero ragionevole di documenti ma – nel nostro caso – trattandosi di trecento-quattrocento occorrenze giornaliere, rischia di diventare un lavoro dispendioso, in termini di tempo e pazienza se non agevolato da riferimenti standard e opzionali che permettano di adottare un filtro che restringa il campo di ricerca.. Non sottovalutiamo da questo punto di vista l'arbitrio del responsabile del protocollo. Questi, infatti, oltre agli errori di data-entry, è anche chiamato a decidere a quali informazioni, poste tra le righe del documento, dare la preminenza e quali omettere. Ponendo l'esempio di una richiesta di avanzamento lavori relativo ad un progetto edilizio: cosa andremo a sottolineare primariamente, il fatto che sia una richiesta o il fatto che riguardi un progetto andando – di conseguenza – ad evidenziare quelli che sono i caratteri distintivi di detto progetto?

Quello che abbiamo riscontrato è che lo stesso operatore in circostanze simili ha proposto soluzioni differenti, spiazzando il lettore attento.

Le criticità presentate sopra sono solo una parte di quei elementi che posti all'interno di un sistema chiuso come quello di un ufficio, di un Comune, o comunque di un'unica Area Organizzativa Omogenea possono essere definiti semplici cavilli le cui conseguenze si traducono in “accettabili” disservizi per l'utenza. Differente è la questione se tali criticità sono poste, o meglio proposte, per essere considerate all'interno di un sistema aperto come quello che prevede il versamento di documentazione di un Ente (Comune) all'interno del sistema di un altro Ente (Regione). Le problematiche che potrebbero essere facilmente risolte attraverso il ricorso alla “relazione informale” presso l'ambito convenzionale e ristretto del Comune, non sono altrettanto applicabili in modo agevole all'interno di un sistema aperto dove non conosco abitudini, consuetudini, modus operandi, degli operatori.

Parola d'ordine "reengineering"

Le criticità del "versamento" all'interno del modello OAIS

Abbiamo già ricordato le difficoltà riguardo la reperibilità di un documento e della necessità di provvedere all'inserimento di tutti quei riferimenti che abbiamo definito metadati. Ora immaginiamo di dover provvedere al versamento di tutta una tipologia documentaria all'interno di un archivio appartenente ad un altro Ente. Il modello OAIS a questo proposito costituisce un riferimento per assicurare l'attendibilità e l'affidabilità delle operazioni di versamento, conservazione e futura fruizione del materiale documentario. L'OAIS è un archivio costituito da un insieme di rapporti e di responsabilità che sono gestiti da organi che assicurano il rispetto degli accordi presi a monte delle operazioni di versamento. Questa fase in effetti è molto delicata poiché rappresenta la costituzione di quell'insieme di regole e di informazioni (metadati) che Ente produttore ed Ente gestore dovranno redigere alla luce di che tipo di documento dovrà essere conservato «*Le contenu de l'Opération de versement est basé sur un modèle de données négocié entre l'OAIS et le Producteur dans le Protocole de versement*»¹⁰ E' sempre in questa fase che vengono stabilite tutte quelle procedure che entrambe le parti devono rispettare nell'operazione di versamento e che confluiscono nel "Protocollo" «*Le Producteur établit un Protocole de versement avec l'OAIS*»¹¹.

In tal senso sebbene ciascun organo è incaricato di responsabilità precise, esiste un interesse condiviso affinché il documento in questione possieda tutto un set di metadati e di disposizioni che permetta la sua conservazione e la sua fruizione nel tempo «*le rôle du Producteur et le rôle de l'Archive sont de la responsabilité d'une même entité*»¹².

Come abbiamo anticipato sopra parlando di metadati¹³, sebbene ciascun documento abbia una natura ed un'intelligibilità che gli è propria, esistono delle informazioni che possiamo definire basilari:

Provenienza, contesto, identificazione, integrità.

Provenienza:

sono quei metadati relativi all'autore del documento e tutta la storia che lo stesso ha subito nel corso degli anni (modifiche, versamenti sostitutivi, ecc.).

Contesto:

Spiegano le relazioni esistenti tra il documento e altri documenti simili, o l'ambito in cui tale documento è inserito, le motivazioni che hanno portato alla sua creazione, la pratica, il fascicolo all'interno del quale è inserito o il procedimento amministrativo a cui attiene.

Identificazione:

sono quelle informazioni che permettono ad un documento di essere identificato univocamente: un esempio è rappresentato dal codice identificativo di un prodotto in un negozio o al codice ISBN dietro i libri.

Integrità:

¹⁰ Modèle de référence...(op. citata), INTERACTION AVEC LE PRODUCTEUR

¹¹ Modèle de référence...(op. citata), INTERACTION AVEC LE PRODUCTEUR

¹² Modèle de référence...(op. citata), CONCEPT OAIS

¹³ Cfr. "Informazione minima e accordi bilaterali"

è rappresentata da tutti quei meccanismi atti ad impedire alterazioni non documentate all'interno dei SIP¹⁴; spesso sono costituiti da dispositivi in grado di effettuare il checksum sulle evidenze informatiche dei documenti in modo tale da poter segnalare eventuali alterazioni.

Trattando nello specifico quella documentazione destinata al protocollo, e di interazione tra sistemi di protocollo, avranno per noi un interesse particolare quelle informazioni in grado di identificare quelle specifiche tecniche che fanno parte del sistema informatico adottato dall'A.O.O. di riferimento. Quei metadati che riescono a descrivere «...*la nature des techniques utilisées pour représenter l'information (par exemple, le support physique, le support numérique, le format)*»¹⁵.

Tale cura sarà tanto più urgente quanto più eterogenei sono i sistemi che devono entrare in comunicazione: in particolare – come nel caso di documentazione scambiata tra due P.A., dove ci troviamo di fronte ad una conservazione a lungo termine, tale “scrupolo” si renderà indispensabile per poter accedere ai dati in modo intellegibile. Tale concetto sarà chiarito ed approfondito meglio in seguito¹⁶.

L'OAIS inoltre ha il compito di rendere disponibili ad una comunità di possibili fruitori i documenti conservati nel proprio archivio. Non ci dilungheremo a descrivere in questa sede i procedimenti caratterizzante questa funzione: ci limitiamo a sottolineare come tale operazione sia atta a fornire quelle tipologie di metadati che possano essere comprensibili al pubblico a cui sono destinate. «*L'Archive, [...] s'assure qu'ils sont compréhensibles pour les communautés susceptibles d'y accéder*»¹⁷

Ente produttore, Archivio di gestione OAIS, e pubblico di riferimento nonché futuro fruitore del documento: questi – semplificando molto - insieme al documento e ai suoi metadati, costituiscono gli attori principali di un modello OAIS, dove non si tratta semplicemente di spostare un documento da un folder ad un altro (fisico o logico che sia), ma che consta di tutta una serie di procedure e di relazioni che tutelano la possibilità di conservare un documento, o meglio di una serie di documenti, nel lungo periodo e di poter fruire degli stessi da parte di un pubblico prestabilito.

Insomma, appare chiaro come una singola amministrazione che voglia inserirsi all'interno di tale contesto, debba operare una totale rivoluzione del proprio *modus operandi* e attuare una reingegnerizzazione dei processi di lavoro, stabilendo e concordando con le amministrazioni coinvolte, quanto necessario ad intraprendere un cammino unitario e, per quanto possibile...standard. Solo in questo modo sarà possibile lenire ed attenuare quelle differenze che attualmente contraddistinguono il nostro Paese.

¹⁴ SIP (Submission Information Package) costituisce un pacchetto di versamento provvisto di metadati e specifiche concordate in fase di progettazione, che l'Ente Produttore ha il compito di inoltrare all'Archivio OAIS che provvederà alla sua gestione.

¹⁵ Modèle de référence...(op. citata), NEGOCIER ET ACCEPTER LES INFORMATIONS

¹⁶ Cfr. “IL formato XML”.

¹⁷ Modèle de référence...(op. citata), CONCEPT OAIS

Il formato XML

Il concetto di “interoperabilità” include dentro di sé il concetto di “sistema” e quello di “comunicabilità”: sistema poiché come ci ricorda Bertalanffy¹⁸ al modificarsi di un solo elemento di un sistema viene a modificarsi l'intero sistema; d'altra parte la comunicazione qui gioca un ruolo essenziale in quanto due o più Enti sono tenuti a “parlare” lo stesso linguaggio e – trattandosi di linguaggio informatico – la soluzione non è sempre immediata. Soprattutto per l'intervento dal punto di vista hardware di due colossi quali sono “Microsoft” e “Apple”, entrambi interessati a conservare un'esclusiva delle loro strutture, tendono a rendere incompatibili le loro piattaforme nei confronti di quelle concorrenti. Medesimo discorso può essere fatto da un punto di vista software dove la concorrenza delle software house è ancora più accanita e dove esiste un forte interesse a criptare o rendere illeggibili dati che provengono da applicazioni differenti. Pensiamo soltanto che tipo di business possa rappresentare la P.A. per gli operatori commerciali, i quali tutti interessati a mantenere l'esclusiva della commessa, creano tutta una serie di legami di dipendenza tra l'azienda fornitrice e il cliente (nel nostro caso la P.A.).

In un contesto tale, come posso proporre un ambiente di interoperabilità?

Difficile per non dire improponibile sarebbe l'opzione di vincolare la P.A. a sistemi hardware e software uniformi: costi troppo elevati e poteri troppo forti da contrastare.

Un aiuto in tal senso ci viene dal formato XML (l'extensible markup language) il quale «*serve esclusivamente per definire altri linguaggi [...] non è altro che un insieme standard di regole sintattiche per modellare la struttura di documenti e dati*»¹⁹. Attraverso questo formato è possibile strutturare, grazie ad un set di istruzioni, in modo dettagliato, le specifiche di un documento. La macchina non fa altro che eseguire le istruzioni sintattiche e il documento verrà visualizzato secondo le modalità prestabilite. «*Concretamente, un documento XML è un file di testo che contiene una serie di tag*²⁰, attributi e testo secondo regole sintattiche ben definite»²¹

Entrando più nel dettaglio possiamo pensare ad un documento XML come ad un albero gerarchico all'interno del quale ciascuna componente è considerata “elemento” dell'intero albero che costituisce, così, l'intera struttura del documento (“document tree”). Ciascun elemento, aggrappato all'albero principale (“root element”), può contenere a sua volta dei “sottoelementi”. Insieme elementi e sottoelementi concorrono alla valenza logica del documento. Ogni elemento può essere legato a delle proprietà che abbiamo deciso di conferirgli, allo stesso modo di come decidiamo di conferire un valore (ad esempio valuta) ad una casella nel noto applicativo “Microsoft Excel”. Tali proprietà sono chiamati “attributi”.

E' questo carattere di libertà nel conferire attributi specifici agli elementi che rende possibile la personalizzazione e la progettazione del documento che, grazie a una vasta gamma di scelta, può rispecchiare al massimo obiettivi e specifiche determinate.

¹⁸ Ludwig Bertalanffy, Teoria generale dei sistemi, 1968, Oscar Saggi Mondadori, 2004

¹⁹ <<http://xml.html.it/guide/lezione/1839/introduzione/>>

²⁰ Differentemente dai tag che siamo abituati a selezionare nel formato html, quest'ultimi sono personalizzabili, posso cioè decidere che tipo di attributo dare ad un certo elemento piuttosto che ad un altro, in funzione degli obiettivi che intendo perseguire.

²¹ <<http://xml.html...>(op. citata), Struttura dei documenti XML

Fattore non trascurabile è il fatto che all'interno del documento posso inserire elementi che abbiano semplice valore di testo²², in modo tale che posso portare all'attenzione di altri operatori elementi che, pur disancorandosi da una valenza prettamente logica o di visualizzazione, consentiranno l'archiviazione di metadati importanti intrinseci al documento (tipo di documento, versione, operatore, ecc).

Possiamo riassumere che all'interno di un documento XML sono io che decido primariamente cosa dovrà riportare il documento: nella fase di progettazione in cui – grazie ad «*un insieme standard di regole sintattiche per modellare la struttura di documenti e dati*»²³, potrò avvalermi di uno strumento che rimane sempre uguale a se stesso indipendentemente della piattaforma hardware o software che utilizzo per la visualizzazione. Infatti «*Le regole in questione sono standard e questo garantisce l'indipendenza da una specifica piattaforma hardware e software o da uno specifico produttore*»²⁴. Tale caratteristica risolve anche tutte quelle problematiche derivanti dall'adozione di documenti in formati o versioni differenti: l'informazione binaria, memorizzata all'interno di un file può assumere significati e – di conseguenza - visualizzazioni differenti in funzione di quale “logica” viene utilizzata per tradurre o decodificare tale informazione.

Nell'ultimo periodo, con il moltiplicarsi in breve tempo di sistemi operativi, applicativi d'ambiente, finanche software periferici, si è resa più insistente la problematica della compatibilità e dell'interoperabilità dei diversi sistemi. Ad aggravare la situazione e a rendere più urgente soluzioni come quella che abbiamo sopra presentato, c'è la crescita esponenziale di documenti che Enti e aziende come le banche o gli istituti governativi mettono a disposizione on-line sul proprio portale e la necessità – pertanto – da parte dell'utenza di poterli condividere e visualizzare in modo adeguato.

Per quanto riguarda la P.A. abbiamo già mostrato come lo sviluppo decentrato, favorito negli ultimi anni da misure di tipo federalista, e il particolarismo della nostra penisola, nonché la disomogeneità delle città italiane in merito alla dimensione, numero di abitanti, organizzazione, professionalità e livello economico, hanno portato alla creazione di reti e sistemi chiusi in se stessi a cui – ora – si chiede di dover dialogare tra loro. Un'impresa apparentemente titanica ma che, attraverso l'adozione del formato XML troverebbe un'adeguata soluzione: l'Ente mittente, infatti, invierebbe documenti contenenti informazioni condivise tali da poter essere visualizzate tanto dall'Ente mittente che dal/dagli Ente/i destinatario/i, indipendentemente dalle piattaforme hardware e software adottate, o dai contratti stipulati con le varie software house.

Abbiamo posto l'accento sulla possibilità di personalizzare il contesto logico e valoriale delle informazioni di un documento XML, di prescindere dall'utilizzo di piattaforma hardware e software, e ciò è corretto: ma dobbiamo descrivere alcuni limiti che caratterizzano il formato XML e, contemporaneamente, ne specificano la logica interna del *document tree*.

Pensiamo alla logica di un PC che deve riuscire a “comprendere” le istruzioni che io, programmatore intendo impartirgli: il grande sforzo che dovrò fare consiste nel consegnare alla logica del PC una chiave interpretativa che gli consenta di distinguere quelle che sono istruzioni da eseguire e da una semplice stringa di testo. Problematiche di questo genere si potrebbero moltiplicare se non ricorressi ad un insieme di regole standard che permettono al PC di comprendere le istruzioni che intendiamo impartirgli. Un documento che segue queste regole è definito: ben formato (well formed).

²² sezione CDATA (Character DATA).

²³ <http://xml.html... (op. citata)

²⁴ <http://xml.html... (op. citata)

Ma quali sono queste semplici regole?

- l'obbligo a che in un documento esista uno ed un solo elemento principale o root element che contenga tutti gli altri elementi e sottoelementi del document tree;
- Ogni elemento deve essere provvisto di tag di apertura e chiusura.
- I tag di chiusura devono rispettare l'ordine dei tag di apertura, allo stesso modo con cui in matematica rispettiamo l'ordine di apertura delle parentesi prima di chiuderle.
- Riportare correttamente i nomi dei tag, tenendo conto del fatto che XML discrimina tra lettere maiuscole e minuscole.
- I valori degli attributi devono essere racchiusi tra apici singoli o doppi (<>).

«La violazione di una qualsiasi di queste regole fa in modo che il documento risultante non venga considerato ben formato»²⁵.

²⁵ <<http://xml.html...> (op. citata), Documenti ben formati

La funzione ricoperta dalla “grammatica” all'interno di un documento XML

Abbiamo visto nel precedente capitolo come un documento XML offra la possibilità – all'interno del vincolo posto da cinque semplici regole – di definire tag in maniera del tutto libera ed autonoma in funzione dell'esito che si vuole raggiungere. Ora però si potrebbe sentire l'esigenza opposta ovvero dare vita all'interno del documento di una logica predefinita che non permetta di scorrazzare in modo arbitrario tra elementi eterogenei ma, al contrario, all'interno di una struttura ben definita, in cui i diversi tag possono essere personalizzati soltanto all'interno di un protocollo, una “grammatica” che io stabilisco e personalizzo in vista di un obiettivo specifico. di un meccanismo che ne vincoli l'utilizzo all'interno dei documenti. Si può stabilire quali tag possono essere utilizzati e come per rispecchiare una struttura logica predefinita.

E' questo il significato con cui l'XML viene definito «*meta-linguaggio [...] affidabile e flessibile per gli usi più disparati*»²⁶.

In altre parole abbiamo bisogno di definire una grammatica per definire il linguaggio che abbiamo ideato. Una grammatica è un insieme di regole che indica quali elementi possono essere utilizzati e con che struttura è possibile comporre il documento.

Sintetizzando quanto detto, non è più sufficiente che il documento sia ben formato e – pertanto – rispetti i cinque dettami di cui sopra – ma è necessario che rispetti i vincoli imposti dalla grammatica adottata. Solo in questo caso il documento potrà definirsi “valido”. Ci potremmo infatti trovare nella possibilità di avere un documento ben formato ma non valido, in quanto pur rispettando le regole XML non segue quelle della grammatica. Diversamente un documento valido sarà necessariamente ben formato.

Vero è che – rispettando le regole XML può capitare che un documento non rispetti le regole di una grammatica ma segua quelle di un'altra: nel qual caso dovremmo dire che si tratta di un documento valido all'interno di una specifica grammatica; mentre sarà non valido all'interno di un'altra grammatica.

E' importante, giunti a questo punto, di domandarci perché ci siamo addentrati così tanto all'interno di linguaggi e meta-linguaggi; noi, in fondo dobbiamo semplicemente trovare una soluzione che renda possibile l'interoperabilità dei sistemi: che pertinenza hanno document tree, grammatiche e, più in generale, l'XML con l'interoperabilità?

Possiamo rispondere a questi interrogativi con un'altra domanda:
per chi progettiamo un documento?

La risposta è semplice per un operatore della P.A. che da nord a sud dell'Italia sia in grado di leggerlo, comprenderlo e – all'occorrenza – apportarvi modifiche. E' chiaro da tale considerazione che la questione si stia spostando dalle caratteristiche del documento in sé e dei suoi metadati, a quello dei software deputati alla codifica/decodifica, o meglio ai «*tipi di elaborazione*»²⁷ delle istruzioni provenienti dalle grammatiche XML.

Il compito di questi software elaboratori (“Parser”) è semplice:

1. verificare che il documento sia ben formato;
2. verificare che il documento sia valido.

In questo senso avremo *parser validanti* e *parser non validanti*.

²⁶ <<http://xml.html...>> (op. citata), Il mondo di XML: XHTML e oltre

²⁷ <<http://xml.html...>> (op. citata), Documenti validi.

Ora riusciamo a comprendere l'importanza di questi software elaboratori, che permettono di decodificare in modo corretto documenti e rispettivi metadati indipendentemente da piattaforme hardware o software utilizzate, ponendo una grossa ipoteca intorno alla possibilità di trovare una soluzione affidabile per quel che concerne lo scambio di documentazione tra sistemi di protocollo diversi.

Limiti e punti di forza delle Grammatiche “DTD” e “XML Schema”

il “Catalogo Schemi e Ontologie” *«che consente la condivisione e il riuso di modelli concettuali (ontologie) e di schemi dati (XML Schema) a supporto della cooperazione tra servizi.»*²⁸ E' inoltre possibile effettuare la ricerca di servizi applicativi facenti riferimento a modelli concettuali condivisi. Nel fare ciò abbiamo riconosciuto l'importanza di far seguire ai documenti che si desidera condividere tutto un set di informazioni relativi al documento e al suo contesto e che saranno da questi inseparabili. che devono seguire metadati

Posto che non è questa la sede per decidere il quadro tecnico della P.A. del futuro – tanto meno del suo sistema di protocollo – troviamo opportuno affacciarci timidamente su quella finestra che si apre verso mondi fatti di codici ed istruzioni, di automazioni e di vincoli quale è quello delle grammatiche, e scrutare se riusciamo a cogliere almeno punti di forza, criticità, o possibili aporie.

*«il primo approccio per la definizione di grammatiche per documenti XML è rappresentato dai Document Type Definition (DTD)»*²⁹

La grammatica usata dal DTD individua due funzioni precise:

- trovare gli elementi utilizzabili nel documento nonché la sua struttura;
- definire gli attributi di ciascun elemento.

In generale un Dtd ha la capacità di orientare il documento intorno ad alcune caratteristiche generali quali il numero di occorrenze di un dato elemento (+); piuttosto che la possibilità di introdurre stringhe di testo di carattere generale e prive di qualunque carattere di “business logic” o automazione predefinita (<!ELEMENT testo (#PCDATA)>). Quest'ultima funzione, tra l'altro, risulta ottimale per l'inserimento di elementi quali i campi relativi alle note, dove difficilmente si avverte la necessità di dare vita a scelte obbligate preimpostate.

Per quanto riguarda la definizione degli attributi nel Dtd esistono alcune istruzioni che indicano come l'inserimento di un dato attributo sia obbligatorio (#REQUIRED), piuttosto che opzionale (#IMPLIED), o a valore fisso non modificabile (#FIXED valore).

Una volta posto in relazione il documento XML con il suo Dtd, il parser di riferimento verificherà la correttezza del documento non solo in merito alla sua validità rispetto alla grammatica utilizzata.

Utilizzando i Dtd si riscontra un maggior controllo sulla struttura e sull'uso dei tag in un documento XML, evitando che la libertà nella definizione dei tag possa far perdere il controllo sui contenuti.

Tuttavia l'uso dei Dtd porta in seno alcune criticità che possiamo sintetizzare nel modo seguente:

- la sintassi relativa alla grammatica si discosta troppo dalle regole di definizione di un documento XML;
- non consentono di specificare il “range” all'interno del quale può ripetersi un dato tag
- non consente di specificare i tipi di dato che il documento potrebbe richiedere.

²⁸ <[²⁹ <<http://xml.html...> \(op. citata\), Dtd: Document Type Definition.](http://www.cnipa.gov.it/site/it-it/Attivit%C3%A0/Sistema_Pubblico_di_Connettivit%C3%A0_(SPC)/Servizi_infrastrutturali_di_interoperabilit%C3%A0_cooperazione_ed_accesso_(SICA)/></p></div><div data-bbox=)

E' per tali motivi che qualora avessimo bisogno di un maggiore controllo sugli elementi che possono trovarsi all'interno di uno specifico tipo di documenti XML, i Dtd non risultano più sufficienti. Ciò ha indotto alla definizione di logiche alternative per definire grammatiche per documenti XML: il più noto è XML Schema.

A differenza del Dtd, infatti, l'XML Schema offre la possibilità di vincolare maggiormente il documento.

L'aspetto rivoluzionario consiste nel fatto che la sua sintassi è basata sul meta-linguaggio XML: ovvero è un documento XML che descrive la grammatica di un linguaggio XML.

Un punto di forza di questa grammatica è la possibilità di definire la tipologia di dato a cui mi riferisco. Oltre alla definizione degli attributi, allo stesso modo della grammatica Dtd, XLS Schema definisce anche se ci stiamo riferendo ad una stringa, piuttosto che ad un numero intero, ad una data, ecc. secondo lo schema seguente:

Tipo di dato	Descrizione
xs:string	Stringa di caratteri
xs:integer	Numero intero
xs:decimal	Numero decimale
xs:boolean	Valore booleano
xs:date	Data
xs:time	Ora
xs:uriReference	URL

I tipi sopra riportati sono standard ma XLS Schema consente di personalizzare immettendo un range definito ad una quantità o creare e personalizzare altre limitazioni.

Potremmo dilungarci a descrivere molte altre caratteristiche interessanti di questa grammatica quali le funzionalità dei "tipi complessi" definizione di "namespace" o la creazione di parser; tutto ciò, sebbene interessante dal punto di vista tecnico, devia l'attenzione da quelli che rimangono i nostri obiettivi.

A tale proposito ci sembra di poter concludere che la grammatica XML Schema, basata su una sintassi XML, pertanto dotata anche di maggiore fluidità ed intuitività, sembra di gran lunga quella che maggiormente riesce ad interpretare le strutture di tipologie documentarie presenti e/o future, grazie soprattutto ad un ampio raggio di opzioni che consentono di determinare e vincolare un documento in modo migliore rispetto a quanto potrebbe fare la grammatica Dtd.

A onor del vero, però, non ci sentiamo di dare una risposta categorica ed univoca: né potrebbe dimostrare il contrario il presente studio che - ridimensionato sotto il profilo dello "spessore scientifico" richiede nuova linfa da opere di maggior respiro, basate su documentazione di maggior rilievo. Nel dare tale giudizio siamo consapevoli di essere, per quel che concerne questa materia, in ambito pionieristico, dove anche il piccolo contributo può essere spunto di riflessione.

In tal senso ci sentiamo di concludere che a nostro avviso i sistemi di protocollo presso la P.A. avrebbero maggiori opportunità di personalizzare i propri documenti, e prepararli per la condivisione con altre P.A., locali o nazionali, qualora avessero la possibilità di adottare un linguaggio di mark-up come quello XML e una grammatica come quella XML Schema.

Tuttavia lasciamo aperta la porta a lavori futuri, più approfonditi e meglio documentati, che stimolino la ricerca scientifica verso una direzione che conduca all'adozione da parte della P.A. di queste grammatiche, con l'auspicio che ai progressi conseguiti in ambito giuridico, facciano seguito anche quelli dal punto di vista pratico.

Conclusioni

Partendo da una situazione all'interno della nostra penisola di eterogeneità e di relativa autonomia in merito alle soluzioni archivistiche adottate dalle singole P.A., ci siamo chiesti – in virtù delle indicazioni normative esistenti – se fosse possibile avviare un dialogo tra i diversi sistemi locali, in modo da poter migliorare l'efficienza del pubblico servizio. Abbiamo richiamato l'attenzione a come la documentazione passando da un'amministrazione ad un'altra, debba essere accompagnata in modo inscindibile da tutta quella serie di informazioni che abbiamo chiamato metadati e che permettono il recupero del documento, indipendentemente dal tempo trascorso dalla sua archiviazione, o dall'operatore addetto a soddisfare la richiesta.

A questo punto è stato necessario aprire una parentesi per far comprendere come l'impianto progettuale che abbiamo appena presentato sia da considerarsi per buona parte esclusivamente teorico: come testimoniato dall'esperienza osimana, ma possiamo estendere il concetto anche a molte altre località del marchigiano o del centro-sud, la realtà organizzativa della amministrazioni locali è costretta a fare i conti con numerose difficoltà, in particolare economiche, che di fatto impediscono loro di dare sostegno alle disposizioni vigenti. Una situazione difficile che abbiamo già denunciato all'inizio del presente studio, e di cui l'imputato maggiore è il mancato stanziamento di fondi atti a rendere operative le modifiche attuate solamente “de jure”.

Dal documento e dai metadati ci siamo concentrati poi sulla fase di “versamento” della documentazione: più precisamente nel modello OAIS, attraverso l'invio di pacchetti informativi (SIP) all'interno di un archivio. Nell'OAIS abbiamo visto quale sia il gioco svolto dalle singole responsabilità e della reingegnerizzazione dei processi che coinvolgono le singole amministrazioni/A.O.O., al fine di garantire la conservazione e la futura fruizione della documentazione e dei rispettivi metadati.

Per rendere, infine, agevole la comunicazione e limitare la possibilità di errore umano tra le diverse amministrazioni – in un'epoca di chiusura e incompatibilità sia sotto il profilo hardware che software – abbiamo presentato un formato che, se correttamente adottato, è in grado di risolvere tutta quella serie di limiti, criticità e aporie che, di fatto, trasformano un documento in una vacua sequenza di bits inintelligibile.

In particolare abbiamo analizzato nel dettaglio le caratteristiche del documento XML, la sua struttura e le sue funzioni, focalizzando l'attenzione sul ruolo giocato dalle grammatiche nel vincolare scelte troppo arbitrarie che rischierebbero di porre un documento al di fuori del controllo del suo autore; e dai parser: software “indipendenti” in grado di permettere la corretta ed intellegibile visualizzazione del documento.

Quest'ultimo punto ci porta a dare una risposta affermativa al nostro interrogativo iniziale: ovvero se fosse possibile avviare un dialogo tra i diversi sistemi locali, in modo da poter migliorare l'efficienza del pubblico servizio.

Quale grammatica?

Con questa domanda esordisce il nostro lavoro che ora, giunti al suo compimento, è destinata a rimanere tale in quanto il presente studio non consente di approfondire quanto dovuto l'argomento. Vero è che scrutando timidamente dalla finestra degli addetti ai lavori sembrerebbe di notare una dinamicità maggiore nella grammatica XLS Schema che, come abbiamo

visto, oltre al fatto di essere basata sul meta-linguaggio XML, possiede una ricchezza di alternative e di opzioni che il Dtd non ha e in questo appare quella più idonea per affiancare e sostenere la P.A. nel suo sforzo di interrelazione e condivisione.

Bibliografia

Bertalanffy, Ludwig, Teoria generale dei sistemi, 1968, Oscar Saggi Mondadori, 2004
Modèle de référence pour un Système ouvert d'archivage d'information (OAIS), Marzo 2005
CCSDS <<http://public.ccsds.org/publications/archive/650x0b1.pdf>>.
<<http://xml.html.it/guide>>
Sito istituzionale Digit PA <<http://www.cnipa.gov.it>>
Delibera AIPA 7 Maggio 2001 n.28