
Cooperazione amministrativa e semantica: Il ruolo del Documento Intelligente

Romeo Pruno

L'attuazione dell'e-Government ha reso possibile la presentazione sul mercato IT di una grande quantità di soluzioni per il document management. Un'attenta analisi delle soluzioni adottate evidenzia l'inesistenza di modelli risolutivi efficienti. In particolare non si è fatta attenzione al bisogno di un interfacciamento comune alla compilazione di modelli di documento che non vincoli la Pubblica Amministrazione (PA) all'utilizzo di soluzioni chiuse. In questo articolo verrà introdotta la "cooperazione amministrativa e semantica" come esempio di cooperazione applicativa per risolvere problemi della tracciabilità del dato all'interno delle procedure di workflow e document management e come sistema di repository di conoscenza in grado di essere consultato sia dagli attori del sistema interno, sia da un dominio esterno, utilizzando tecniche di accesso standard come ad esempio i servizi web...

Abstract

Il concetto di "cooperazione amministrativa e semantica" viene introdotto in riferimento a soluzioni per la gestione documentale con particolare attenzione al dominio della PA. Le normali procedure di cooperazione applicativa tra PA vengono prese come esempio per sottolineare i vantaggi nell'adozione del modello proposto. Vengono inoltre evidenziati i problemi della modellazione, compilazione e gestione di un documento digitale e la rintracciabilità del dato introducendo il concetto di Documento Intelligente come vera e propria metodologia di gestione dell'informazione.

1 Introduzione

L'attuazione dell'e-Government, in particolare in Italia, ha reso possibile la presentazione sul mercato IT di una grande quantità di soluzioni per il workflow e document management. Un'attenta analisi delle soluzioni adottate evidenzia l'inesistenza di modelli risolutivi efficienti. In particolare non si è fatta attenzione al bisogno di un interfacciamento comune alla compilazione di modelli di documento che non vincoli la Pubblica Amministrazione (PA) all'utilizzo di soluzioni chiuse.

In questo articolo viene introdotta la "cooperazione amministrativa e semantica" come esempio di cooperazione applicativa per risolvere problemi della tracciabilità del dato all'interno delle procedure di workflow e document management e come sistema di repository di conoscenza in grado di essere consultato sia dagli attori del sistema interno,

sia da un dominio esterno, utilizzando tecniche di accesso standard come ad esempio i servizi web. Tutto questo è reso possibile principalmente da due definizioni:

- ◆ **Il Documento Intelligente**, non come singolo componente ma come frutto dell'orchestrazione di servizi offerti dai business actors;
- ◆ **CID (Cooperative Intelligent Document)**, che permette di definire le fasi della modellazione, formazione, tracciabilità e reperimento del documento digitale all'interno del sistema;

A supporto delle definizioni di cui sopra, il sistema proposto permetterà altresì:

- ◆ **La definizione di una ontologia** che descriva il dominio degli atti amministrativi e che definisca uno schema di conoscenza ed il relativo vocabolario semantico;
- ◆ **La derivazione di un sistema formale** che consenta di strutturare il dominio di riferimento, ed in particolare i documenti digitali, implementando la tecnologia XML (livello sintattico), facendo uso di una terminologia di riferimento (livello semantico) definita e consistente che permetta, attraverso una architettura modulare la specializzazione dell'ontologia a seconda delle reali esigenze della PA di riferimento;
- ◆ **La realizzazione di un PM (Point Management)** che permetta la creazione, gestione e tracciabilità di documenti elettronici dell'ente attraverso la loro catalogazione semantica utilizzando l'ontologia ed i rispettivi file schema.

In questo articolo sarà data una definizione del Documento Intelligente come frutto di una orchestrazione di servizi, attraverso la cooperazione tra le PA intervenute alla sua creazione, inglobando nella sua formazione e/o consultazione il concetto di cooperazione applicativa descrivendo le specifiche per un sistema CID per la PA.

2 Cooperazione amministrativa e semantica

Lo studio delle categorie di una certa entità che esiste o che potrebbe esistere in un determinato dominio, viene chiamato ontologia, ovvero "un catalogo dei tipi di una entità di cui viene assunta la sua esistenza in un dominio D in prospettiva del suo accesso da parte di un attore che utilizzi un linguaggio L per accedere o comunicare con D".

A seguito di questa affermazione l'ontologia è sinonimo di comunicazione. Essa, infatti, permette di definire il significato dei concetti che fanno parte di un determinato dominio, stabilisce le relazioni intercorrenti tra i termini stessi, definisce quindi la semantica che costituisce l'"humus" del sistema e la garanzia della persistenza delle informazioni trattate.

E' importante inoltre definire:

- ◆ **Information Ontology**, specifica, dato un catalogo di tipi, quali sono le entità definite ed indefinite utilizzando solo una descrizione in linguaggio naturale.
- ◆ **Formal Ontology**, specificata data una collocazione di nomi per concetti e tipi di relazioni organizzata in ordine di classi e sotto classi di tipi.

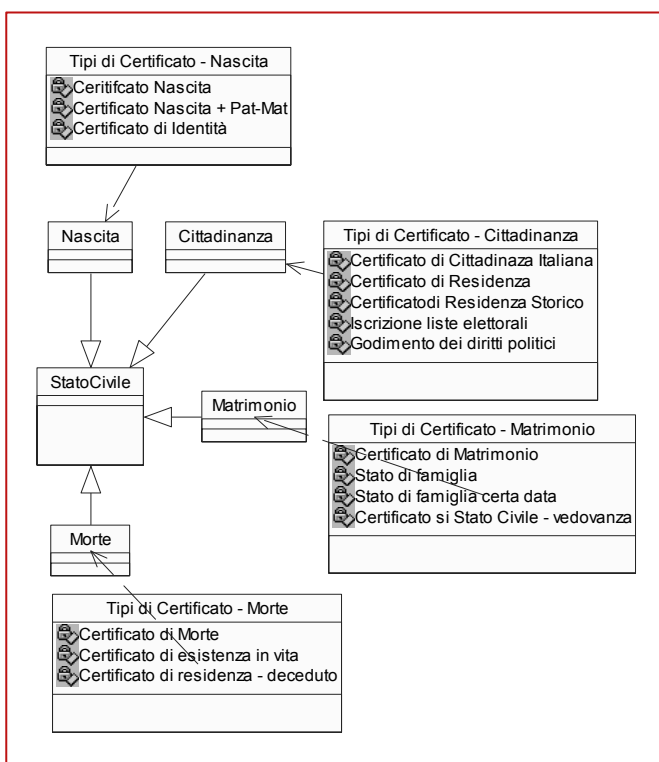
Un approccio di questo tipo per il workflow management all'interno di una PA permette il raggiungimento dei seguenti obiettivi:

- ◆ **Semantic management** di flussi documentali utilizzando documenti elettronici basati su standard;
- ◆ **Supporto ai processi amministrativi**;
- ◆ **Information retrieval e knowledge management** specializzati a seconda delle singole esigenze dell'ente;

Il primo passo verso la costruzione di una ontologia per l'e-document management è la formazione del vocabolario di riferimento. Per esempio si può creare una lista di keywords che descrivano, dato il modello di riferimento [Fig. 1], il contenuto del documento, Information ontology.

La figura sotto rappresenta un possibile esempio di ontologia per i documenti di "Stato Civile" appartenenti alla legislazione dello stato italiano.

Figura 1 Ontologia di un documento elettronico



A seguito della formazione di questo catalogo si deve costituire una rappresentazione standard delle informazioni. La International Organization for Standard, in questi anni ha provveduto a definire una notazione standard per la rappresentazione e l'interscambio dei dati.

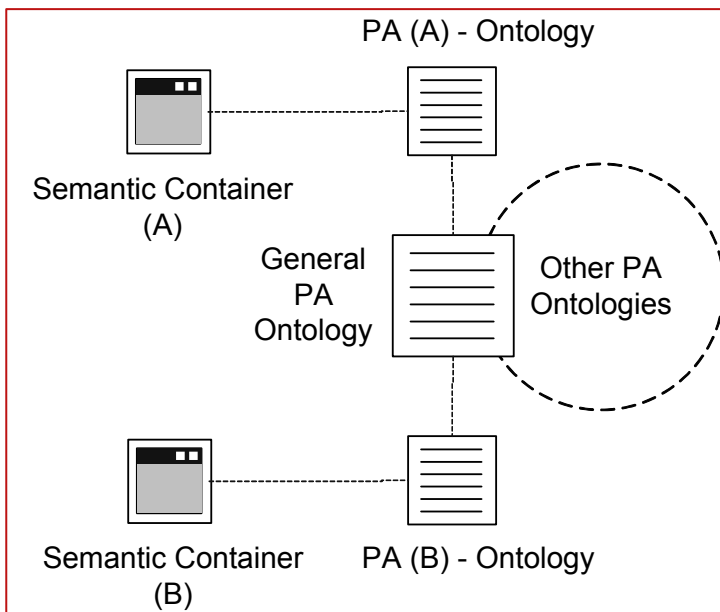
La notazione base per la rappresentazione è SGML, Standard Generalized Mark-Up Language, in particolare utilizzeremo XML, che deriva da questo standard. La topic maps redatta da questa organizzazione offre gli strumenti per la creazione della "Knowledge" separata dall'area del document management, Formal Ontology. Analizzando altre soluzioni di e-Government che sono state implementate negli ultimi anni l'utilizzo dell'ontologia è solo una parte dell'infrastruttura

complessiva, in grado di ricoprire vari aspetti del sistema:

- ◆ **Cooperazione amministrativa**, che permette l'interoperabilità tra le PA attraverso la condivisione della stessa ontologia;
- ◆ **Scalabilità del sistema**, specializzazione dinamica a seconda delle reali necessità della PA, senza che ci sia una ridefinizione delle componenti;
- ◆ **Accesso al knowledge** e condivisione delle informazioni;

◆ Integrazione della tecnologia esistente;

Figura 2 Ontologia condivisa



Una volta definita, ogni PA, è in grado di condividere o accedere all'ontologia comune e specializzarla [Fig. 2], quindi, estrarre i vari modelli possibili di stato civile che vengono descritti. Questo, è il modo migliore di condividere l'ontologia.

3 Forms on line e cooperazione applicativa

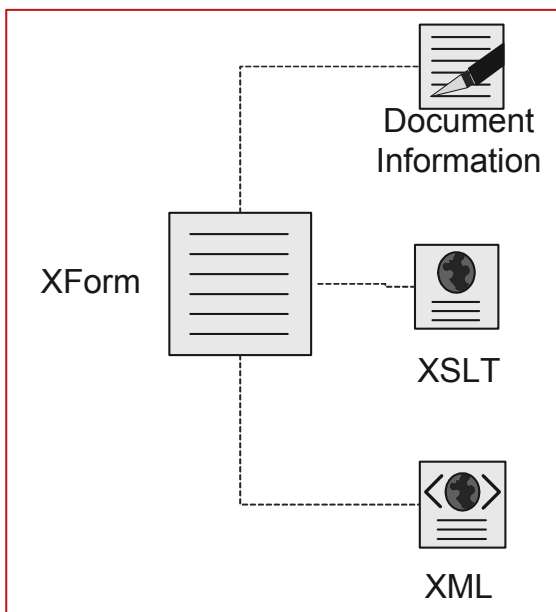
Attualmente l'idea della formazione e acquisizione di dati online deve poter rispondere a domande del tipo:

- ◆ Come saranno visualizzati i forms su browser differenti?
- ◆ Come implementare una soluzione di acquisizione dati che offra lo stesso meccanismo di convalida su più e/o diversi dispositivi, oppure, come gestire l'ambiente disconnesso attraverso una politica di gestione di forms online?

A queste domande si può rispondere attraverso un'analisi della tecnologia disponibile per la creazione di forms online e un approfondimento dei concetti di standard cooperation per la trattazione dei contenuti tra più attori, interoperabilità dei dati per uno scambio di informazioni mutuato (a livello applicativo) che non dipenda dal tempo in cui si accede alla fonte dati, scalabilità di una eventuale politica di creazione/distribuzione di forms online.

Introduciamo la tecnologia XForm come linguaggio standard per la creazione e/o compilazione di documenti elettronici. XForm è un'applicazione XML e può essere integrata con diverse altre tecnologie come ad esempio Xlink, XForm, utilizzando la sintassi XML permette all'interno del modulo di specificare [Fig. 3]:

- ◆ file XSLT per la presentazione grafica del form;
- ◆ file XSD per la definizione dello schema e quindi i vincoli del modulo;
- ◆ informazioni, riguardanti il documento come ad esempio l'autore, la tipologia di form, informazioni riguardo la versione...;

Figura 3 Struttura di una form


Questo ultimo punto, le informazioni del documento, è molto importante per una tracciabilità del form basato sulle diverse versioni esistenti: caratteristica fondamentale per la definizione dei Documenti Intelligenti.

La struttura del form così definita potrà essere utilizzata tramite una interfaccia comune come il Web.

Il modello sopra descritto ha bisogno però di una struttura “gestionale” che permetta la sua creazione e quindi distribuzione in rete (internet/intranet) in aggiunta alle procedure di autenticazione ed autorizzazione per la sua consultazione, modifica e inoltro verso un'altra PA o un altro ufficio della medesima PA.

Quindi, sempre attraverso il browser, si potranno permettere tutte le operazioni di gestione utilizzando un PM (Point Manager). Essendo questo un sistema point-to-point, dove ogni amministrazione è al pari livello delle altre, si rende necessaria l'introduzione di una entità (Point Manager) che sia in grado di controllare l'iter procedurale dell' Intelligent Document.

L'architettura stateless su cui si basa il funzionamento di internet, non permette una facile gestione dell'ambiente disconnesso. Attualmente, la compilazione di forms online ed offline è una necessità che può essere soddisfatta utilizzando la programmazione xml-oriented.

Essa permette la persistenza delle informazioni direttamente nel file XML, dando la possibilità una volta effettuato il download del form di completare le parti mancanti in modalità disconnessa ed inviarle una volta create le condizioni per l'accesso alla rete. Questa funzionalità è utile da una parte per dare la possibilità all'utente di completare la compilazione del modulo a suo piacimento, dall'altra è parte fondamentale di una cooperazione amministrativa tra enti, i quali possono creare procedure di completamento “parziale” di quelle parti del form per cui si ha la necessità e/o autorizzazione alla compilazione.

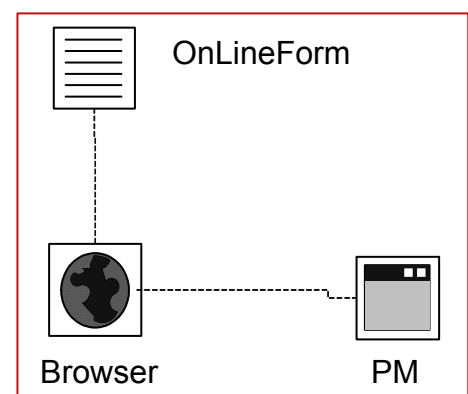
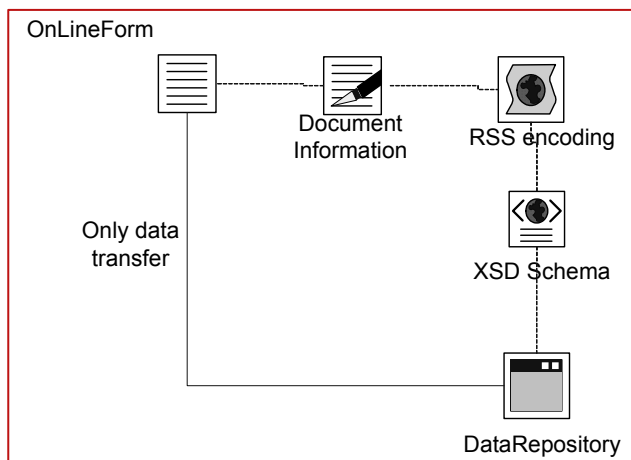

Figura 4 Point Manager

Figura 5 Procedura di “data repository”


Una risposta valida alla domanda di interoperabilità e accesso ai dati del form, si ha nella creazione di un repository [Fig. 5] che permetta il salvataggio, temporizzato ed automatico, dei dati del modulo.

Questa procedura se associata all'utilizzo di un RSS encoding per l'estrazione delle informazioni personali del form, permette da una parte lo storage dei soli dati e dall'altra l'estrazione di quelle parti del documento che una volta strutturate secondo le direttive ontologiche, andranno a formare il semantic

container della nostra ontologia.

Come illustrato nella figura, le informazioni personali del documento vengono archiviate nel DataRepository attraverso la codifica dello stream dati secondo una convalida RSS, che attraverso l'associazione dello stesso file schema, permette l'estrazione delle parti fondamentali del documento per l'utilizzo semantico. Nello stesso tempo i dati del modello vengono inserite nel repository e messi a disposizione della comunità.

La tracciabilità del documento originale viene comunque garantita, attraverso l'utilizzo delle informazioni nel file XSD schema associato.

Anche in questa parte non ci soffermiamo a definire la tecnologia più adatta nell'implementazione di questo repository, ci limitiamo a definire questa entità come una struttura che deve garantire lo storage dei dati, anche in forma non organizzata e la loro persistenza. La “standard cooperation”, e quindi la capacità di estendere i processi di workflow management verso attori esterni al dominio, può essere implementata utilizzando le componenti sopra descritte.

Le soluzioni che si basano su questo modello potranno:

- ◆ **Utilizzare informazioni** derivanti da più pubbliche amministrazioni, senza la necessità di modificare o programmare gli applicativi legacy dell'ente;
- ◆ **Utilizzare un DataRepository** comune che archivi i dati dei documenti nel quale altre PA possano attingere, per acquisire la versione del documento che interessa a loro;

4 Il documento intelligente

Molte organizzazioni hanno da tempo iniziato ad utilizzare i documenti in formato elettronico per automatizzare i processi. Le analisi di mercato hanno evidenziato che i costi associati alle inefficienze dei processi documentali possono essere ridotti e a volte addirittura completamente eliminati, utilizzando i documenti digitali per raccogliere in modo più efficiente le informazioni dai soggetti che partecipano ai processi.

Per quanto riguarda gli enti pubblici, l'adozione di documenti in formato elettronico costituisce ormai sempre più spesso un requisito legale: negli Stati Uniti, ad esempio, gli

enti federali sono tenuti a conformarsi al GPEA, una legge del 21 ottobre 2003 sull'eliminazione delle attività documentali cartacee negli enti pubblici, che impone agli enti di dare ai cittadini la possibilità di inviare informazioni in formato elettronico incoraggiando l'utilizzo della firma elettronica come integrazione della sicurezza.

In generale l'utilizzo dei documenti elettronici porta dei vantaggi significativi dalla parte della pubblica amministrazione:

- ◆ Diminuzione dei costi di stampa, elaborazione, distribuzione, spedizione;
- ◆ Diminuzione degli errori di compilazione;
- ◆ Possibilità di condividere e/o accedere alle informazioni condivise;

e dalla parte dell'utente finale:

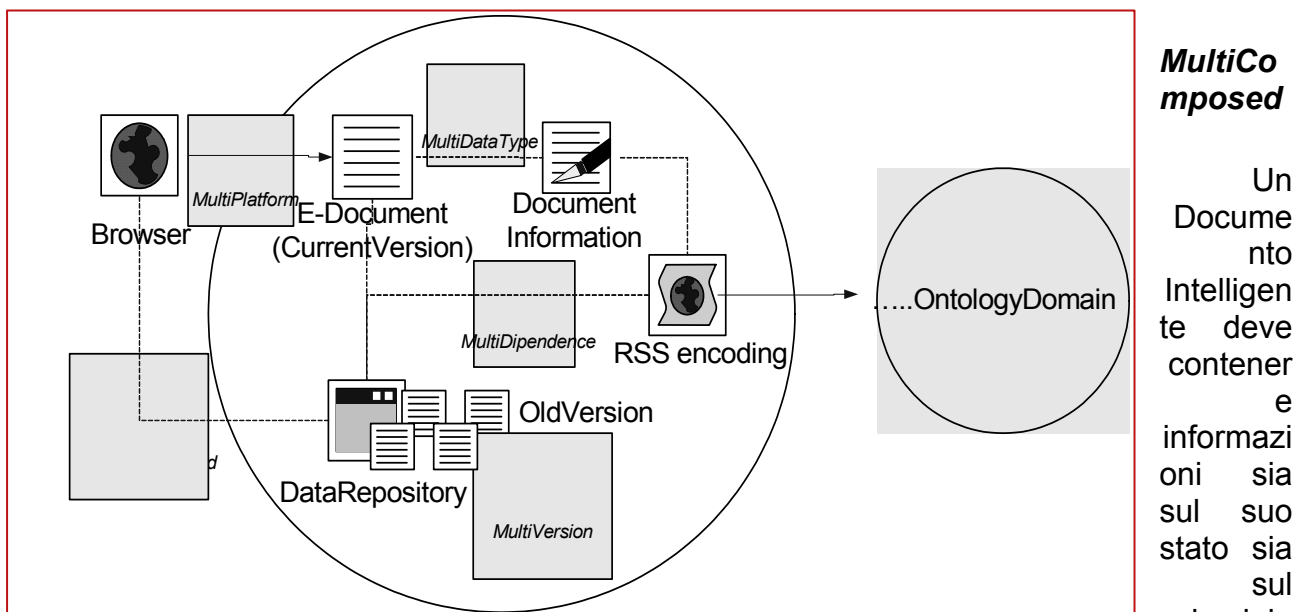
- ◆ Diminuzione del tempo della compilazione manuale;
- ◆ Invio elettronico dei documenti;
- ◆ Costi postali, fax o corriere espresso;

Da questa analisi emerge la necessità di creare un processo di workflow management che contempli la fruizione del documento elettronico all'interno del dominio applicativo.

L'approccio che è stato utilizzato fino ad oggi non basta per dare risposte alle domande sopra citate, non dobbiamo pensare ad un Documento Intelligente come semplice modulo componibile ma come frutto di una orchestrazione di servizi offerti dalla comunità che ha partecipato alla sua formazione. A supporto di questa affermazione è d'obbligo descrivere le principali funzionalità che un Documento Intelligente deve avere [Fig. 6]:

- ◆ **MultiComposed**
- ◆ **MultiVersion**
- ◆ **MultiDipendence**
- ◆ **MultidatatypeMultiplatform**

Figura 6 Le proprietà di un Documento Intelligente



applicativo in cui avviene la sua formazione e/o distribuzione. Gli utenti o i processi devono concorrere mutualmente alla sua formazione.

MultiVersion

La tracciabilità del dato deve essere garantita dalla persistenza delle varie versioni del file. Ogni ufficio dell'amministrazione e/o PA esterne riconosceranno la versione di quel determinato file che dovranno compilare.

MultiDipendence

Consiste nella creazione del documento digitale dalla collezione di altri documenti, attraverso una architettura modulare che viene specializzata a seconda dell'Application Domain, in cui viene generato.

MultiDatatype

Più tipi di dato devono poter convivere all'interno della struttura del Documento Intelligente, se non altrimenti specificato dello schema del documento.

MultiPlatform

La formazione, utilizzo del Documento intelligente devono essere in grado di assicurare l'interoperabilità a livello applicativo, su qualsiasi piattaforma verrà implementata una simile procedura.

La MultiDipendenza automatizza le procedure standard di creazione di documenti per la PA. La maggior parte dei documenti all'interno degli uffici governativi è frutto della cooperazione di più uffici sia interni che esterni al sistema. Prendendo come esempio un "Certificato di Nascita", quest'ultimo deve acquisire informazioni sia dall'ufficio anagrafe del comune di nascita, sia da quello di residenza, fino alla certificazione di quest'ultimo posta dall'ufficio di stato civile.

Si deve considerare che spesso il comune di nascita è diverso da quello di residenza e viceversa, questo fa sì che alla compilazione di questa tipologia di documenti intervengano attori esterni alla PA, dove quest'ultimo è stato richiesto. Questa caratteristica si appoggia al concetto di "Document Assembly". Ridefiniamo il concetto di Document Assembly per permettere di definire procedure automatiche alla formazione di un Intelligent Document. In particolare definiamo quattro fasi:

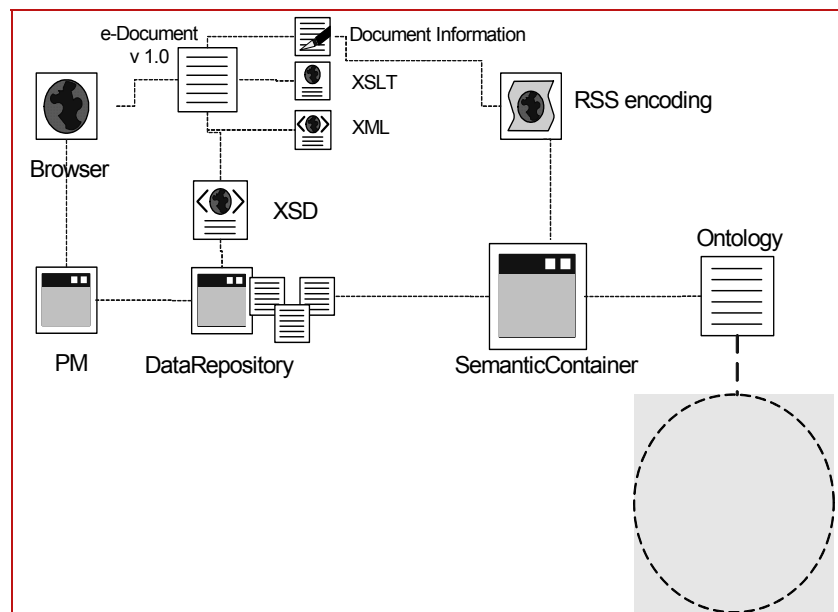
- ◆ **Requirements**, l'utente richiede il tipo di modello (OnLine Forms / Active Documents), la funzione del Document Assembly in questo specifico è la ricerca nel DataRepository [Fig. 5].
- ◆ **Performs**, il modello si specializza per la presentazione all'utente inglobando dentro di sé informazioni personali per la tracciabilità all'interno del sistema. [Fig. 4].
- ◆ **Repository**, l'archiviazione di dati si scinde in due parti distinte, le informazioni personali andranno a formare il semantic container attraverso l'utilizzo della codifica RSS, mentre i dati verranno inseriti nel DataRepository associando un file schema per la rielaborazione degli stessi [Fig. 2].
- ◆ **ContentOrganization**, permette l'utilizzo di vocabolari semantici per la classificazione di tipi di documento e la consultazione degli stessi attraverso una ontologia comune, accessibile anche da domini esterni.

5 CID Cooperative Intelligent Document

La visione condivisa di un sistema che permetta l'interoperabilità tra PA a livello nazionale è di primaria importanza per la definizione delle tipologie di servizi e delle funzionalità che questi devono avere. Gli elementi su cui costruire una visione condivisa sono quindi:

- ◆ Il livello dei governi coinvolti e la responsabilità nella definizione degli aspetti essenziali del sistema;
- ◆ La titolarità dei dati disponibili e dei servizi erogati;
- ◆ La localizzazione dei dati, delle informazioni, dei servizi disponibili;
- ◆ Le modalità di accesso ai dati, alle informazioni, ai servizi locali;
- ◆ Le politiche di accesso e/o condivisione dei dati;

Tutte queste informazioni concorrono alla modellazione di un sistema di cooperazione, che è alla base della nostra definizione di Documento Intelligente.

Figura 7 Architettura del CID


Le caratteristiche principali del CID sono [Fig. 7]:

Architettura P2P

attraverso questa architettura le PA che partecipano all'orchestrazione risiedono su uno stesso livello di interazione, con il vantaggio della non dipendenza, se non per l'ontologia condivisa, da qualsiasi sistema.

Orchestrazione

La cooperazione nella compilazione da più PA di un Intelligent Document è cosa molto frequente, il nostro sistema consente la circolazione del documento tra più PA utilizzando l'orchestrazione dei servizi del CID. Come esempio pratico esaminiamo la richiesta di un certificato di residenza, nella compilazione di questo documento devono partecipare due PA distinte.

La prima PA (A) deve attivare la procedura e sa a priori che il numero di attori che coopereranno alla formazione del certificato è uno.

Dunque includerà nel documento le informazioni riguardo alle sue possibili versioni (v1 e v2).

Una volta fatto ciò (A) inizia la compilazione del form online ed invia il modulo a (B), utilizzando le informazioni immesse al momento della sua creazione da parte di (A). Da questo punto in poi (A) diventa orchestratore, è lei a richiedere una ricevuta di ritorno da (B), in modo automatico, quando quest'ultima visualizza il documento, (B) dal canto suo si trova abilitata la parte che a lui interessa per la compilazione delle restanti sezioni.

Una volta compilato (B) fa il recover secondo le operazioni descritte in precedenza ed automaticamente invia il documento finale ad (A), dal canto suo (A) invierà sempre in modo automatico un acknowledge per chiudere la cooperazione. In questa ultima fase, (A) può apporre anche la sua firma digitale ed inviare l'intero documento all'ufficio richiedente.

Interoperabilità

Attraverso la condivisione di una ontologia comune per la classificazione di strutture di modelli comuni, ogni PA ha sia la possibilità di specializzare l'ontologia esistente all'interno del suo dominio applicativo e sia di condividere informazioni e la stessa ontologia con tutte le altre PA. L'interoperabilità in questo modello è vissuta anche da un punto di vista applicativo in quanto l'interscambio grazie all'utilizzo di XML permette un accesso ai dati indipendente dal tempo ed in modo mutuato.

6 Conclusioni

Il modello proposto all'interno di questo articolo è frutto di una attenta analisi al dominio applicativo chiamato e-Government. Abbiamo ridefinito il concetto di Documento Intelligente utilizzando strutture esistenti, tecnologie esistenti, facendo attenzione alle procedure esistenti, da ciò ne viene che un attento riuso sia delle risorse del sistema e sia della sua architettura si rileva vantaggioso la dove c'è a monte una ontologia ben definita (anche se nascosta), come lo è la legge che definisce i vari tipi di documenti amministrativi, la gestione del workflow management attraverso la nostra idea risulta efficiente e vantaggiosa sia per i cittadini e sia per i principali attori, le PA.