

# INTEROPERABILITÀ E COOPERAZIONE NELL'E-GOVERNMENT: IL RUOLO DEI METADATI

Flavio Corradini, Francesco De Angelis, Alberto Polzonetti, Barbara Re  
Dipartimento di Matematica e Informatica  
Università di Camerino  
*nome.cognome@unicam.it*

*L'adozione di standard di metadati rappresenta un elemento fondamentale per la realizzazione della piena interoperabilità tra sistemi informativi eterogenei. La loro applicazione, infatti, permette di definire una rappresentazione efficiente ed efficace della realtà inerente il dominio di riferimento verificando, al tempo stesso, lo scambio di informazioni tra soggetti che cooperano per il raggiungimento di un obiettivo specifico. Con questo lavoro si vuole offrire una review dell'applicazione dei metadati nel dominio e-Government discutendone l'importanza. Inoltre, si propone un'infrastruttura basata su metadati per la rappresentazione di processi e documenti capace di garantire l'interoperabilità tra sistemi informativi.*

## 1. Introduzione

Il termine “metadato” rappresenta l'insieme delle informazioni relative ai dati, sui dati, ed intorno ai dati. Questa definizione informale prende spunto dall'inglese: “metadata”, che a sua volta è formato dal termine greco “meta” (attinente) e dal latino “data”, pl. di “datum” (informazione).

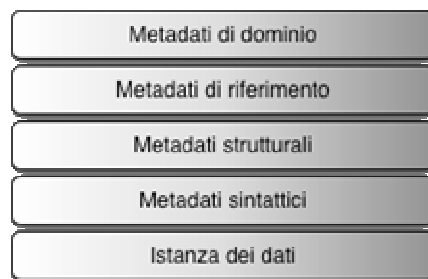
Diverse sono le definizioni presenti in letteratura ne riportiamo di seguito alcune di maggior rilievo per questo lavoro.

- “Machine-understandable information about Web resources or other things” [1].
- “Data associated with objects which relieves their potential users of having to have full advance knowledge of their existence or characteristics. A user might be a program or a person” [5].
- “Structured data about resources that can be used to help support a wide range of operations” [4].

Resta chiaro che la distinzione tra dati e metadati non è intrinseca, ma dipende dal contesto di utilizzo. Gli stessi metadati sono dati, e perciò possono essere memorizzati come dati, (nella risorsa stessa o in un'altra risorsa), e possono essere descritti da altri metadati, e così via. Il numero di metalivelli da specificare dipende quindi dalle caratteristiche delle applicazioni e dalle specifiche esigenze.

Tipicamente, parlando di metadati si distingue una classificazione su cinque livelli (Figura 1). Al livello più basso troviamo le istanze dei dati che effettivamente rappresentano elementi *value-oriented* che non dicono nulla su se stessi. I metadati sintattici, invece, sono inerenti il tipo di dato che stiamo trattando e rappresentano il vero primo metadato nello stack. Esso informa il

software che lavora sui dati relativamente a questioni inerenti, al formato del linguaggio, alla lunghezza del documento, al bit rate, ai permessi, etc. A livello superiore i metadati strutturali forniscono una struttura alle unità di dati, organizzandoli secondo una struttura gerarchica, relazionale o ad oggetti. Su tale strutturazione si basano i metadati di riferimento che forniscono link tra differenti modelli dei dati. Questo tipo di metadati può variare sensibilmente per far sì che differenti applicazioni che sfruttano modelli di dati diversi possano dialogare tra loro. Infine, al più alto livello si trovano i metadati di dominio che tipicamente sono descritti sotto forma di ontologia [10] fornendo un supporto cruciale per la condivisione di informazioni in contesti eterogenei.



**Figura 1: Stack di Metadati**

Da sottolineare è il ruolo che i metadati hanno assunto con l'avvento del World Wide Web ed in particolare in relazione alla necessità sempre più incombente di trovare informazioni utili nella massa di informazione disponibili. In questo contesto, la capacità dei metadati di essere compresi oltre che dagli uomini anche dalle macchine, risulta essere un'arma vincente per la riduzione di tempi e dei costi legati alla ricerca/comprendimento delle informazioni.

Nell'ambito della pubblica amministrazione si sta osservando un processo di decentramento amministrativo delle competenze. Si osserva quindi una trasformazione dei procedimenti da *uni-ente* a *multi-ente*. I primi rappresentano attività indipendenti in grado di assolvere a compiti specifici, mentre quelli multi-ente rappresentano il risultato di una cooperazione delle varie amministrazioni le quali concorrono, ognuna per la sua parte di competenza, all'erogazione del servizio richiesto dall'utente finale. L'introduzione di più enti nel processo di erogazione dei servizi fa riferimento ad un modello che valorizzi sia l'uniformità delle interazioni, ma anche la specificità di ogni erogatore nonché la chiara identificazione sia dell'erogatore che del fruitore. Non pochi sono gli ostacoli da superare per la piena realizzazione del decentramento amministrativo. Innanzitutto, la repulsione da parte dell'apparato dirigente a seguito di una distribuzione e riduzione del potere decisionale, ma anche problematiche tecniche dovute alla molteplicità delle soluzioni architettoniche. In questo contesto i metadati assumono un ruolo fondamentale in grado di individuare una soluzione che da un lato tenda a preservare l'autonomia amministrativa e che, allo stesso tempo, consenta ai diversi sistemi di interoperare. L'introduzione di standard condivisi risulta essere in questo contesto un veicolo per la piena realizzazione dell'interoperabilità.

In questo lavoro si vuole evidenziare il ruolo che i metadati giocano nella cooperazione applicativa nel dominio dell'e-government presentando un'architettura in grado di garantire l'interoperabilità tra sistemi informativi favorendo il processo di ammodernamento della PA e riducendo tempi e costi.

Il resto del lavoro è così organizzato. Nella Sezione 2 si discute su soluzioni a supporto della cooperazione applicativa SPCoop, nella Sezione 3 si discute sull'importanza dei metadati per l'interoperabilità, mentre nella Sezione 4 si discute un'infrastruttura con lo scopo di garantire l'interoperabilità tra sistemi informativi. Infine, la Sezione 5 conclude il lavoro.

## **2. Importanza dei metadati per l'interoperabilità**

L'interoperabilità semantica è definita [11] come una capacità dinamica derivata dall'applicazione di tecnologie software speciali come reasoner, inference engines, ontologie e modelli, che utilizzano meccanismi di inferenza, mettono in relazione, interpretano, e classificano il significato implicito dei contenuti digitali. Ciò abilita la definizione di processi di business adattativi, di processi di acquisizione della conoscenza, di regole di business, nonché l'interoperabilità tra software applicativi eterogenei.

I metadati giocano quindi un ruolo cruciale fornendo il substrato necessario per il funzionamento di questo genere di software. Sulla base delle funzionalità che i metadati sono in grado di gestire, essi possono quindi essere distinti principalmente in tre categorie funzionali di seguito proposte:

- **Metadati Descrittivi:** per l'identificazione e il recupero degli oggetti digitali; sono costituiti da descrizioni normalizzate dei documenti digitali nativi, risiedono generalmente nei sistemi di information retrieval.
- **Metadati Amministrativi e gestionali:** forniscono dati utili per la gestione della risorsa descritta, per le operazioni di gestione degli oggetti digitali all'interno dell'archivio.
- **Metadati Strutturali:** descrivono la struttura interna fisica o logica dei documenti (es. introduzione, capitoli, sezioni, indice di un libro) e le loro relazioni con le varie parti degli oggetti digitali.

I metadati possono poi essere distinti anche sulla base della loro localizzazione rispetto alla risorsa che descrivono in embedded o esterni. Nel primo caso essi sono inclusi nell'oggetto, mentre nel secondo caso essi sono esterni all'oggetto ed archiviati a parte come link all'oggetto descritto.

Le principali funzioni che essi si prefiggono di realizzare sono quelle di: (i) gestire le risorse fornendo la loro identificazione a favore della loro presentazione e recupero; (ii) consentire l'organizzazione, la localizzazione, la gestione e le statistiche; (iii) favorire l'interoperabilità e l'integrazione di risorse simili; (iv) facilitare l'archiviazione e la conservazione delle risorse.

### *2.1. Stato dell'arte*

Molti dei progetti internazionali in ambito e-government sono orientati alla fornitura di servizi. In tali progetti, numerose iniziative sono rivolte alla standardizzazione dei metadati, tra queste particolare rilevanza hanno quelle che si prefiggono di introdurre un livello semantico nei dati descrivendoli in maniera tale da poter essere utilizzati con successo per il raggiungimento dell'interoperabilità semantica.

- SAGA - Standards und Architekturen in eGovernment Anwendungen in Germany [14].
- e-GIF - eGovernment Interoperability Framework in United Kingdom [19].
- ADEA - l'Agence pour le développement de l'administration électronique in France [9].
- EIF - European Interoperability Framework for European solutions [7].
- FEAF - Federal Enterprise Architecture Framework in USA [8].

Analogamente si segnalano progetti sviluppati da entità non nazionali solo per menzionarne alcuni citiamo OntoGov [12], Terregov [17], EPRI [7], QUALEG [13], SmartGov [16], etc.

### **3. Soluzioni a supporto della cooperazione applicativa SPCoop**

I modelli di interoperabilità semantica [11] nella cooperazione applicativa si caratterizzano usualmente per: (i) l'uso dei vocabolari (condivisi o separati), e (ii) la distribuzione della logica di integrazione (centralizzata o decentralizzata).

Nella pubblica amministrazione italiana la cooperazione applicativa è decentralizzata in quanto le amministrazioni stringono accordi per regolare la cooperazione su base prioritaria (accordi di servizio), mentre, la condivisione dei metadati si realizza nell'ambito di accordi multilaterali. Da notare come l'implementazione degli accordi di servizio è demandata ai singoli sistemi informativi così come la logica di integrazione.

Dal punto di vista dell'uso dei vocabolari si possono individuare entrambe le soluzioni: questi possono essere *separati* o *condivisi*. Nella prima, l'erogatore del servizio pubblica il proprio modello concettuale, gli utilizzatori traducono poi il modello di servizio nei termini della propria applicazione, e la semantica del sistema di cooperazione è data dall'utilizzatore sotto forma di regole di traduzione del modello dell'erogatore. Nei modelli decentralizzati con vocabolario condiviso, invece, erogatori e utilizzatori dei servizi condividono un vocabolario di metadati dove ciascuno interpreta il vocabolario condiviso rispetto alla propria concettualizzazione, in questo caso la semantica del sistema di cooperazione è data da ognuno sotto forma di regole di interpretazione del vocabolario condiviso.

Si considera quindi che avere un insieme di termini condiviso permetta di evitare la costruzione di una matrice di traduzione tra modelli concettuali

eterogenei, per gestire la semantica della cooperazione tra  $N$  sistemi, la cui dimensione può collocarsi nell'ordine di  $N^2$ . Al tempo stesso non ci si affida alla correttezza delle ipotesi semantiche del traduttore per gestire l'accuratezza delle traduzioni. La gestione della cooperazione applicativa con metadati standard prevede quindi di gestire la semantica della cooperazione tra  $N$  sistemi attraverso un insieme di interpretazioni la cui dimensione è  $N$ . L'accuratezza di una interpretazione è legata alla comprensione della semantica condivisa da parte dell'interprete per la quale è possibile ipotizzare una più agevole verifica.

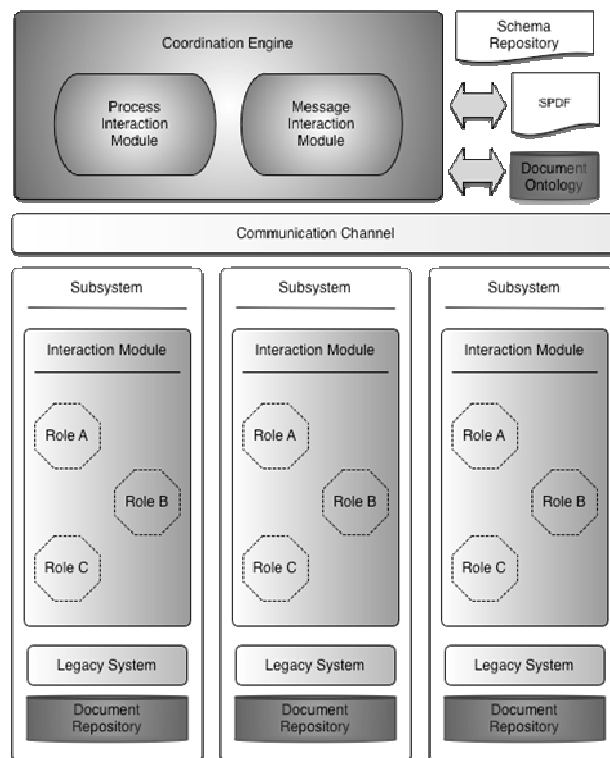
Quindi a fronte di un adeguato investimento, la standardizzazione dei metadati può portare alla cooperazione applicativa notevoli benefici tra cui una maggiore accuratezza nella gestione degli aspetti semantici: flessibilità, semplicità e robustezza. Si possono poi trovare una serie di vantaggi quantitativi a proposito della reattività al cambiamento (in termini di riduzione dei tempi per la ricezione di leggi e normative), alla riduzione dei costi di integrazione ed all'incremento della riusabilità delle soluzioni in quanto le applicazioni potrebbero essere basate su modelli concettuali omogenei.

Pubbliche amministrazioni diverse possono avere linguaggi con una semantica definita ma con sintassi diverse. A tutt'oggi i sistemi informativi evidenziano la mancanza di una sintassi condivisa e ciò rappresenta un fattore di inibizione per lo sviluppo della cooperazione applicativa. La standardizzazione dei metadati risulta quindi essere il modo migliore per definire la sintassi in sistemi di cooperazione decentralizzati come quelli della pubblica amministrazione.

#### **4. Soluzione a supporto dell'interoperabilità**

La proposta che formalizziamo si contestualizza in un ambiente distribuito eterogeneo in cui i sistemi informativi disponibili devono essere integrati per semplificare la condivisione dei dati, informazioni e conoscenza. In particolare, l'ambiente ha il compito di veicolare i documenti tra i vari soggetti del sistema distribuito in linea con la struttura della PA italiana.

In questo modo si promuove l'evoluzione ed il riuso di soluzioni esistenti in modo da includerle in un ambiente coordinato. Ciò preserva il flusso di informazioni che venivano scambiate, tra differenti attori, in relazione ad uno specifico modello di business. Nella Figura 2 si propone l'architettura generale del sistema che di seguito verrà presentata nel dettaglio.



**Figura 2: Architettura per l'interoperabilità**

Al fine di ridurre i tempi ed i costi di ammodernamento e di gestione si propone la costruzione di un'interfaccia in grado di gestire in modo coordinato tutti gli attori del sistema distribuito (che comunque viene rappresentato da un'infrastruttura chiusa dove i partecipanti sono definiti a priori) al fine di raggiungere gli obiettivi stabiliti dalla comunità.

La soluzione proposta garantisce la distribuzione di funzionalità complesse tra differenti moduli di processo traendo vantaggio da soluzioni tecnologiche interoperabili come i Web Service. Si garantisce comunque, attraverso un repository lo storage dei documenti multiformato che tenga in considerazione attraverso l'introduzione di metadati standard la tracciatura della costruzione del documento, il versioning, etc. in linea con i principi di documento intelligente [3]

Nell'architettura proposta l'elemento fondamentale per la realizzazione della piena interoperabilità è l'*interaction module* che agisce come un mediator locale tra i ruoli (funzionalità) del sistema. Al tempo stesso esso rappresenta una parte di un mediator globale tra i sistemi legacy coinvolti nella gestione distribuita dei dati e delle informazioni. A questo livello, il modulo di interazione usa un repository di metadati condiviso e sfrutta le potenzialità di un vocabolario condiviso rispetto ad una concettualizzazione locale per fornire le funzionalità di una matrice di traduzione tra modelli concettuali.

Le comunicazioni che si vengono ad instaurare utilizzano il canale di comunicazione del sistema. Essa rappresenta un mezzo unico di interazione tra

ruoli diversi dello stesso sottosistema e trae vantaggio dall'infrastruttura di rete sottostante. Al tempo stesso il canale di comunicazione implementa le specifiche SPCoop su cui la nostra architettura si estende.

SPCoop costituisce il Sistema Pubblico di Connettività definito come “l'insieme di strutture organizzative, infrastrutture tecnologiche e regole tecniche, per lo sviluppo, la condivisione, l'integrazione e la circolarità del patrimonio informativo della Pubblica Amministrazione, necessarie per assicurare l'interoperabilità e la cooperazione applicativa dei sistemi informatici e dei flussi informativi, garantendo la sicurezza e la riservatezza delle informazioni” [15].

L'attività di coordinazione globale è realizzata da un *coordination engine* capace di gestire un processo di specifica globale. In questo contesto è previsto un linguaggio di specifica dei processi che in accordo con la specifica concettualizzazione dei documenti (SPDF) gestisca il flusso di informazioni tra sottosistemi eterogenei e distribuiti. Le funzionalità base dell'engine sono rappresentate da due moduli: il *process interaction module* e il *message interaction module*. Il primo gestisce la specifica di processo ed è capace di gestire i dati ed i documenti per soddisfare specifici obiettivi gestendo l'interazione tra sottosistemi. Il secondo gestisce gli schemi di messaggi per permettere l'interazione ed il buon fine delle operazioni. Gli schemi sono memorizzati in un appropriato repository contenente tutte le informazioni necessarie per l'identificazione permettendo di risolvere problemi di inconsistenza.

Nel dominio di riferimento un insieme di schemi evitano la costruzione di mapping di tipo statico tra modelli eterogenei. Si propone quindi un approccio ricco a livello semantico dove ontologie condivise abilitano l'interoperabilità tra sottosistemi sfruttando al tempo stesso le funzionalità di un vocabolario distribuito. Con questo tipo di soluzioni i documenti sono annotati con metadati semantici utilizzando un'ontologia di documento, permettendo il recupero dei documenti a partire da una specifica parziale contenuta nella definizione del processo. La piena interoperabilità del sistema viene raggiunta grazie ad una specifica completa delle interazioni tra i ruoli coinvolti.

L'introduzione di sistemi di interoperabilità risulta assumere un ruolo fondamentale nell'introduzione di soluzioni innovative all'interno delle pubbliche amministrazioni. Al tempo stesso risulta essere condizione indispensabile per l'offerta di servizi on-line. Inoltre è da notare come problematiche organizzative risultano essere cruciali nel momento in cui si devono affrontare questioni politiche di gestione e ci si deve confrontare con l'ambiente circostante. Questo implica la necessità di

- armonizzare differenti sistemi informativi;
- re-ingegnerizzare le organizzazioni e i processi di business;
- confrontarsi con la conoscenza necessaria a gestire la complessità dei vari problemi.

## 5. Conclusioni

Alla luce di quanto detto si evince chiaramente l'importanza della definizione di standard opportuni per la rappresentazione delle informazioni nel contesto della pubblica amministrazione.

La standardizzazione dei metadati è il modo migliore per definire la semantica in sistemi di cooperazione decentralizzati come quelli della pubblica amministrazione. A fronte di un adeguato investimento, la standardizzazione può portare alla cooperazione applicativa notevoli benefici sia quantitativi che qualitativi. Relativamente ai primi si evidenzia maggiore accuratezza nella gestione degli aspetti semantici della cooperazione, riduzione delle interdipendenze, trasparenza semantica dei dati ed equidistribuzione della logica applicativa. Dal lato qualitativo si evidenzia, invece, la nascita di applicazioni basate su modelli concettuali omogenei, riduzione dei tempi per la ricezione di leggi e normative ed una sostanziale riduzione dei costi di integrazione. Naturalmente si devono considerare una serie di costi come quelli di produzione del metamodello, di un modello generale di ontologia oltre che di modelli specifici, la gestione dei servizi di metadati e dell'autoring cooperativo, nonché la formazione, consulenza, e software specifico.

In pratica per la gestione dell'IT, ed in particolare per quel che concerne le politiche di interoperabilità all'interno delle PA, entrano in gioco una serie di fattori come: le strategie di e-Government, la domanda di servizi, l'aumento degli investimenti, i sistemi legacy preesistenti, le infrastrutture e la conoscenza disponibile.

All'interno di una singola organizzazione l'integrazione di informazioni e processi può procedere in parallelo, ma l'idea principale è raggiungere un insieme condiviso di elementi che vengano gestiti attraverso l'introduzione di una adeguata infrastruttura middleware di cooperazione. Inoltre, se sussiste un insieme ristretto di partner esiste la possibilità di instaurare relazioni continue per l'integrazione dei processi che costituiscano la parte fondamentale della cooperazione tra le diverse amministrazioni coinvolte.

Con la piena diffusione dei principi di e-Government si sta affermando una nuova dimensione con la quale le amministrazioni dovranno necessariamente confrontarsi. La nostra proposta offre una completa integrazione in rete caratterizzata da sistemi distribuiti aperti e dinamici e da una forte eterogeneità delle risorse e dei sistemi informativi. Si evince chiaramente che il dominio di e-Government deve risolvere problematiche che nascono in tale contesto.

Per questo è necessario un confronto costruttivo tra le capacità tecniche ed organizzative di differenti amministrazioni. Non si tratta certo di un compito banale in quanto si richiede un forte livello di cooperazione per quel che riguarda la definizione di una strategia comune tra amministrazioni. La cooperazione diventa quindi una sfida sempre più complessa nel momento in cui lo scambio di informazioni viene veicolato attraverso servizi.



Concludendo, il linguaggio della Pubblica Amministrazione ha una semantica definita, ma a tutt'oggi i sistemi informativi della PA hanno in comune, per lo più, qualche sintassi. In questo modo la mancanza di una semantica condivisa è un fattore di inibizione per lo sviluppo della cooperazione applicativa nel sistema paese che deve trovare le forze per gestire il processo di standardizzazione del vocabolario della pubblica amministrazione e farne un volano per favorire l'efficienza amministrativa, i servizi ai cittadini e l'economia. Il nostro framework punta a superare questa mancanza garantendo l'interoperabilità tra sistemi informativi legacy in un ottica di specifica dei processi e standardizzazione dei metadati.

Inoltre, attraverso la specifica e la standardizzazione di concetti come quelli dei web service e del semantic web nascono nuove opportunità per il raggiungimento dell'interoperabilità. Naturalmente, queste opzioni risultano essere rilevanti per l'e-Government in quanto in futuro l'integrazione di processi ed informazioni sarà la guida per l'integrazione di applicazioni su larga scala e per la fornitura di servizi integrati, oltre che gettare le basi per una sempre più vasta adozione di architetture orientate ai servizi all'interno del dominio di e-Government.

## 6. Riferimenti bibliografici

[1] Berners-Lee, T., 1997. *Metadata Architecture*. Web available. <http://www.w3.org/DesignIssues/Metadata>

[2] Contenti, M., Termini, A., Mecella M. and Baldoni, R., 2004. *Supporting Inter-administration Cooperation: the EUPUBLI. com Approach*, EGOV'04, pp 226-233.

[3] Corradini, F., Polzonetti, A., Pruno, R., 2005. eGovernment Administrative and Semantic Cooperation: The Role of "Intelligent Documents". EGOV (Workshops and Posters), pp. 150-157

[4] Day, M. (2001). *Metadata for digital preservation: a review of recent developments*. In: P. Constantopoulos and I. T. Sølberg, (eds.), *Research and Advanced Technology for Digital Libraries: 5th European Conference, ECDL 2001, Darmstadt, Germany, September 4-9, 2001, Proceedings, Lecture Notes in Computer Science, 2163*, Berlin: Springer-Verlag, 2001, pp. 161-172. ISBN 3-540-42537-3.

[5] Dempsey, L., Rachel H, 1998. *Metadata: a current view of practice and issues*. *Journal of Documentation*, Vol. 54, No. 2, pp. 145-172.

[6] IDABC - European eGovernment services, 2005. *European Interoperability Framework for pan-European eGovernment services*. Web available. <http://europa.eu.int/idabc/en/document/2319/5644>.

[7] EPRI, 2005, Web available. <http://www.epri.org/>.

[8] FEAF - Federal Enterprise Architecture Framework, 2004, Web available. <http://www.whitehouse.gov/omb/egov/a-1-fea.html>

[9] French Government, 2004, ADEA. Web available. <http://www.adae.gouv.fr/adele/>.

[10] Gruber T. R., 1993. A translation approach to portable ontologies. *Knowledge Acquisition*, Vol. 5, Num. 2, pp. 199 – 220.

[11] Jeffrey T. Pollock and Ralph Hodgson, 2004, *Adaptive Information*, Wiley

[12] OntoGov, 2007, Web available, <http://www.ontogov.com/>.

[13] QUALEG, 2005, Web available, <http://www.qualleg.eupm.net/>.

[14] SAGA, 2005, Web available. <http://www.kbst.bund.de/-,182/SAGA.htm>.

[15] SPC, “Sistema Pubblico di Cooperazione: Architettura, Versione 1.0”, CNIPA, 25 Novembre 2004

[16] Smartgov, 2005, Web available, <http://www.smartgovproject.org/>.

[17] TERREGOV, 2005, Web available, [http://www.terregov.eupm.net/my\\_spip/index.php](http://www.terregov.eupm.net/my_spip/index.php).

[18] Tim Berners-Lee, James Hendler and Ora Lassila, 2001, *The Semantic Web*, Scientific American

[19] UK GovTalk, e-GIF, 2004, Web available, <http://www.govtalk.gov.uk/>.