E-GOVERNMENT ADMINISTRATIVE AND SEMANTIC COOPERATION: THE ROLE OF “INTELLIGENT DOCUMENTS” 1

Abstract: In this paper we introduce and discuss the concept of "administrative and semantic cooperation" in the domain of Public Administrations as an example of applicative cooperation. The approach allows to overcome the problem of “data tracking” through the procedures of workflow management and to manage a knowledge repository system that can be queried by both internal and external actors of the system.

1. Introduction

The realization of e-Government projects made it possible the development of a large number of IT solutions for the workflow management. A suitable analysis of the adopted solutions shows the lack of proper and adequate workflow management models. In particular, interoperability and standard compilation interfaces of document models have to be properly taken into account in order to not to constraint a Public Administration (PA) to necessarily use close systems [1]. In this paper we introduce and discuss the concept of "administrative and semantic cooperation" in the domain of PA as an example of applicative cooperation. The approach allows to overcome the problem of “data tracking” through the procedures of workflow management and to manage a knowledge repository system that can be queried by both internal and external actors of the system also through the web (e.g. web services). This is possible mainly by the definition of the following concepts: the intelligent document and the cooperative intelligent document (CID). The intelligent document is a document that results from the orchestration of services offered by possibly different actors. The cooperative intelligent document defines the architecture that supports the document life cycle. In particular, it supports the modeling, creation, tracking and retrieval of the intelligent document. Two significant aspects of our administrative and semantic cooperation are:

1. An ontology that describes the (public administrative) documents models, their semantics information and a semantic dictionary;

2. Notations and languages to provide a structure to e-Documents by means of XML technology at the syntactic level, and a formal meaning at the semantic one. These two aspects allow to specialize the respective public administrations ontologies via a modular architecture.

1 Work funded by Halley Informatica
In the following sections we provide a detailed description of the concept of intelligent
document, as the result of services orchestration and cooperation among PAs, and a detailed
specification of the CID.

2. Administrative and Semantic Cooperation

Ontology is a set of the categories of a certain entity that exists or that it could exist in a
determined domain. It is also defined “a catalog of the types of an entity of which its existence
is assumed in a domain D in perspective of its access from an actor that a language uses L to
enter or to communicate with D” [2]. Ontology is a synonymous of communication!

It, in fact, allows defining the meaning of the concepts that belong to a determined domain; it
establishes the relationships between terms, it defines the semantics that are the core of the
system and the guarantee of information persistence. Particularly it allows the formalization
of a model of world (or a part of it) that includes concepts and relationships in groups of
simple actors (they suffer the system) and/or experienced actors (they manage the system) [3].

Now it is also important to define:

Informal Ontology, it describes, given a catalog of types, what are the defined and indefinite
entities, using only a natural language description;

Formal Ontology, specified by an arrangement of names in concepts and types of
relationships, organized in classes and sub-classes of types;

An approach like this for the workflow management inside a public administration allows the
achievement of the followings objective:

Semantic management of documents using e-Documents models based on standards;
Support for administrative processes;
Information retrieval and knowledge management specialized according to the demands of the
single public administration;

The first footstep toward the ontology construction is the creation of the reference dictionary.
(E.g. a list of keywords that describes, given the reference model [Fig. 1], the content of the
document) (Informal Ontology) [13].

After the creation of this catalog we have to compose a standard representation of the
information. In these years The International Organization for Standard [4], has providently
define a standard notation for data representation and interchange. The basic notation for the
representation is SGML [8] (Standard Generalized Mark-Up Language), particularly we will
use XML which derives from this standard.

The topic maps compiled by this organization offer the tools for the creation of a
“Knowledge” separated from the area of the document management (Formal Ontology).
Analyzing other e-Government solutions [9] that have been implemented in the last years the
use of the ontology is only a part of the general infrastructure able to cover various aspects of
the system:
Administrative Cooperation, that allows the interoperability between public administrations sharing the same ontology;

Scalability, dynamic specialization according to the real necessities of the public administration, without a redefinition of the components;

Knowledge Access and information sharing;

Integration of the existing technology;

![Ontology structure](image)

Figure 1: An example of Ontology structure

Figure 1 show a possible example of document ontology for "Stato Civile" belonging to the Italian legislation. Once defined, every public administration, is able to share or to access the common ontology and to specialize it [Fig. 2], therefore, to use all the various possible document models of "Stato Civile" that are described. This is the better way to share the ontology.

![Ontology sharing](image)

Figure 2: An example of ontology sharing

3. **Online forms for standard cooperation**

Currently idea of the formation and acquisition of online data has to answer to these types of questions:

> How will the forms be visualized on different browser? How to implement a solution of data acquisition that offers the same mechanism of validation on more and/or different devices, or, How to manage the disconnected environment through an online forms managed procedure?

We answer analyzing the available technology for the creation of online forms, through an examination of the concepts of standard cooperation for the treatment of the contents among more actors and finding a scalable system for online forms creation and management.
Interoperability of the data for a common information interchange (application layer) doesn’t depend on the moment of the data request. We introduce the technology XForms [10] as the standard for web forms. XForms are chosen as the standard mainly due to the fact that they use XML schema as part of the component in the model element. The schema plays a critical role, as the schema provides information about the domain of the form which is then used to represent the semantics of the form. This is important because Web developers do not have to learn a new language, but they can simply use XForms for integration with the existing mark up language such as XHTML, Scalable Vector Graphics (SVG) and XSL. The main advantages of XForms are as follows [5]:

1. **Powerful actions, event model and validation rules:** XForms provide a wide range of client-side processing and reduce the number of round-trips to the server without the need for scripting languages;
2. **Clean separation of data, logic and presentation:** This implies ease generating data-bound controls;
3. **Highly regular XML structure:** The regular XML structure makes it possible to build WYSIWYG user interface (UI) development environments;
4. **Abstract controls:** This type of controls enable abstract application design that gets translated to device specific rendering;

XForm, using the XML syntax, allows to specify inside the form [Fig. 3]:

- file XSLT for the form presentation;
- file XSD for the definition of the schema and therefore the ties of the form;
- Personal document information, (e.g. the author, the typology of form, the document version, etc.)

This last point is very important to track the different document versions this is an important characteristic for the Intelligent Document definition.

![Figure 3: Online form structure](image)

This structure [Fig. 3] so defined can be used through a common interface as the Web. The model above described it has a need of a managerial structure that allows its creation and/or distribution (internet/intranet) additionally to the procedures of authentication and authorization for its consultation, change and forwarding toward another public administration or another office of the same public administration. Then, always through the browser, all the operations of management [11] can be allowed using a MC (Management Component) [Fig. 2]. The system base is a point-to-point architecture and every administration has a peer level as the others. This entity (Management Component) is needed to be able to check the Intelligent Document procedures.
Figure 5 shows as the personal document information are filed in the Data Repository through the RSS encoding [12]. This operation permits principal parts extraction through the association of the same file schema for the semantic use. At the same time the data of the model are inserted in the repository. At the same time document tracking is guaranteed using the information in the XSD partnership schema file. In this section we don't define the most proper technology to implement this repository we just define this entity as a structure that has to guarantee the storage of the data and their persistence. The standard cooperation, and therefore the ability to extend the processes of workflow management toward external actors to the domain, can be implemented using the online forms above described.

The solutions based on this model will be able:

- To use many information coming from many public administrations, without the necessity to modify or to program the application legacies of each public administration;
- To use a common Data Repository, that files the data of the documents, in which other public administration can draw to acquire any version of the document.

### 4. Intelligent Document

A lot of organizations have begun for a long time to use the electronic documents for processes management. The market analyses have underlined that at times the costs of the ineffectiveness of the document management can be reduced and even entirely eliminated by using the electronic documents to pick up in more efficient way the information from the subjects that participate in the processes. For the public administration the adoption of electronic documents it’s now often required by the law: in the United States, for instance, the federal administrations must conform themselves to the GPEA, a law of October 21st 2003 on the elimination of the papery activities document, that obliges them to give the possibility to citizens to send electronic formed information, encouraging the use of the electronic signature as integration of the safety. In general the use of the electronic documents [6] implies some meaningful advantages for:

*The public administration,*

- Diminution of the costs of press, elaboration, distribution, delivery;
- Diminution of compilation errors;
- Possibility to share and/or to access the shared information;
And for the user:

- Diminution of the time of the manual compilation;
- Electronic dispatch of the documents;
- Diminution of the postal, fax, etc costs;

From this analysis came out the necessity to create a process of workflow management that contemplates the fruition of the electronic document inside the application domain. The approach that has been used till now is not enough to answer to our questions, we don't have to think to an Intelligent Document as a simple form but as result of an orchestration of services offered by the community that participated in its formation. Following this affirmation we describe the principal characteristics of an Intelligent Document mandatory ("e-Document for the public administration") [Fig. 6]: Multi composed, Multi version, Multi dependence, Multi data type, Multi platform;

![Intelligent document life cycle](image)

*Figure 5: Intelligent document life cycle*

**Multi Composed**, an Intelligent Document has to contain information both on its state and on the application domain in which its formation and/or distribution origins. The processes or the users has to mutually compete to its creation;

**Multi Version**, the data truck has to be guaranteed from the persistence of the various versions of the file. Every office of the public administration and/or external public administrations will have only to compile the part of the document that are within their competences;

**Multi Dependence**, it consists in the creation of the electronic document from a collection of other documents, through a modular architecture that is specialized according to the application domain, in which it is produced;

**Multi Data type**, more types of data have the possibility to cohabit inside the structure of the Intelligent Document, if otherwise specified in the file schema;

**Multi Platform**, the formation and use of the Intelligent Document they have to assure the interoperability at the application level;

**Multi Dependence** automatizes the standard procedures of the documents creation in the public administration. The greater part of the documents inside the government offices derives from more than one office either inside or outside the system. For example a "Birth Certificate" has to acquire information both from the registry office of the municipality where one born, from that of residence, up to the certification given from the office of "Stato Civile". We have to consider that often the municipality where one born is different from the one where once lives. The compilation of this typology of documents requires external actors to
the public administration, where this certificate has been requested. This characteristic is supported by the concept of “Document Assembly” [7]. We redefine the concept of Document Assembly to defining automatic procedures for the creation of an Intelligent Document. Particularly in this paper we define four phases:

1. **Requirements**, user needs a type of model (Online Forms / Active Documents) in this procedure the Document Assembly function is the search in the Data Repository

2. **Performs**, the model specializes itself to display to the user enclosing in the document the proper information [Fig. 5].

3. **Repository**, the data filing is divided in two distinct parts; personal information will go to form the semantic container through the use of the RSS coding, while the data will be inserted in the Data Repository associating a file schema [Fig. 3].

4. **Content Organization**, it allows the use of semantic dictionaries for the classification of document types and their consultation through a common ontology, accessible also from external domain.

5. **CID (Cooperative Intelligent Document)**

A common vision of a system that allows the interoperability among national level public administrations has a primary importance for the definition services typologies and functionalities. Elements on which to build a shared vision are:

- The level of the involved administrations and their responsibilities in the definition of the essential aspects of the system;
- Sharing of the available data and the supplied services;
- The location of the data, of the information, of the available services;
- The formalities of data, information and local services access:

![Figure 6: CID model](image-url)
All these information compete to model of a cooperation system that is the base of our definition of Intelligent Document. The principal characteristics of the CID are [Fig. 7]:

*Point-To-Point Architecture*, through this architecture the public administration participants reside on a same level of interaction, with the advantage of having no dependences on any system.

*Orchestration*, the interoperability by more public administrations in the compilation of an Intelligent Document it’s quite frequent. Our system allows the flow of the document among more administrations using the CID orchestration services. As a practical example we examine the request of a certificate of residence; in the compilation of this document they have to participate two distinguished administrations. The first municipality (A) it has to activate the procedure and knows previously that the number of actors that they will cooperate to create the certificate is two. Therefore it will include in the document the information of the two possible versions (v1 and v2). Done this (A) begins the compilation of the online form and sends the form to (B), using the information introduced during its creation. From this point to ahead (A) becomes the “conductor”, it asks, automatically for a reply message from (B) when (B) visualizes the document, (B) on the other side will be able to compile only its parts. Once compiled (B) recover according to the operations described above and automatically sends the final document to (A). On its side (A) will always send automatically acknowledge to close the cooperation. In this last phase (A) can also affix its digital signature [15] and send the whole document to the applying office.

*Interoperability*, through the sharing of a common ontology for the classification of structures of common models, every public administration has both the possibility to specialize the existing ontology inside its application domain and to share information and/or the same ontology with the others public administrations. The interoperability in this model is also implemented from an application point of view as a data interchange using XML, that allows data access independently from the time and in a mutual way.

8. Conclusion

The model that we have proposed in this paper derives from a careful analysis to the application domain called e-Government.

We have redefined the concept of Intelligent Document using existing structures, existing technologies, watching out for the existing procedures. From this work it came out that is possible to reuse a lot of the system resources and architecture, because of the advantage to have a well defined ontology, as the law that define various types of all administrative documents. The workflow management based on our idea results efficient and advantageous both for the citizens and for the principal public administration actors.

Bibliography


