E-Government strategy in Italy

F. Amato, A. Mazzeo, A. Picariello

Italian e-Government activities

E-Government, or electronic management of public services (or e-Gov), or processes of democratic governance, concerns the reorganization of the bureaucratic processes in both central and local Public Administrations. In this context, one of main goal of e-Gov is that of providing a strong computerized management of electronic documents in order to optimize the work of the governmental offices and offer the users (citizens and businesses) both faster and more effective services and new ways of accessing such services.

From a general point of view, the theme of e-Government can be traced back to the overlap between two worlds that are apparently different and distant from each other; in particular it can be considered as the application of Information and Communication Technology (ICT) to problems that are typical both of the Public Administration and the legal domain.

The use of ICT in the public administrations is not new, being introduced some decades ago with a series of specific projects, which were often the evolution of pre-existent legacy applications, conceived to automate single parts of the information and bureaucratic system and devoid of a systemic and global vision.

Many initiatives, often supported by facilitated finances, were introduced in the eighties within the Community in order to deeply introduce ICT into public administrations and realize strong and flexible information systems, flexible to changing and with the objective of supporting the principal bureaucratic processes within specific domains (Ministries, Local Bodies, Regions, etc.).

In the nineties and until the beginning of the present decade, with the spread of the Internet and the related technologies, the focus has been moved towards the opening of such systems to the web, in order to carry out initiatives of e-Gov and define a first level of interconnectivity shared among the administrations belonging to different domains, principally in the national environment, but also in an international one.

Nowadays, the process of combining the effectiveness of the services and their transparence within Public Administration context, goes through a strong automation of the internal processes and in addition through the capacity of using open systems, able to cooperate at application levels, following federate models: in this way, it is possible to ensure the observance of legal and organizational binding forces established by the autonomy of the various governmental Entities and the achievement of automatic and inter-domain bureaucratic processes.

Note that such technologies are not always directly and easily suitable to the specificities of the Italian bureaucratic applications (of e-Government) because of the binding forces of the specific regulations.

Generally speaking, the strategic plans provided for by all the actions of e-Government have the aims of establishing cooperation and coordination among the different subjects of Public Administration. In the

last decade and more, Public Administration in Italy has been changing its own organizational structure to enable the development of its own information systems with respect to the new application requirements, by opening and reorganizing itself, enacting new regulations, implementing its own standards and using European and international ones, resorting to solutions that often realize real “technological leaps” in the automation solutions applied.

Why revolutionize a bureaucratic organization existing since more than a century and based on paper documents and mechanical processes? Why change?

A first simple answer is given in the following. Looking at the Italian system, the need of change is principally due to the strong necessity of a de-bureaucratization and a simplification of the processes in order to: i) provide the public and private administrative acts with transparencies; ii) to increase in the quality of the offered services; iii) to decrease the costs of the organization, thus increasing its efficiency.

Looking, instead, from a wider point of view, we conclude that there is a great need to arrange, for a national system, convenient instruments able to ensure its growing, development and competitiveness.

The system of a Nation can’t compete in the International and Community environment without a modern and suitable bureaucratic system, based on the use of the new technologies, operating in the Internet, and able to grant to the administrative actions continuity, definite times, quality, safety and privacy.

It is necessary to pass from systems based on computerized procedures, which are often centralized and supporting organizations based on paper documents and manual processes, to information systems focused on processes, which are often so totally automated and completely based on electronic documents that are able to optimize and rationalize the use of the human resources involved.

The incentives to change are above all represented by the spread usage of electronic documents and the related processes of dematerialization, by the implementation in full cooperation and interoperability of inter-intra domain processes, by the availability of qualifying and low-cost technology; by the evolution of the communication networks both in terms of available band and capillarity, by the safety of the various levels of the system, by effective systems of access control and profiling of the users.

The main instruments achieved, but still in evolution, concern electronic signature for documents legal validity, temporal mark-up for providing temporal evidence, digital protocol, long term preservation of electronic documents according to the regulations, the service of certified electronic mails to give evidence to the posting and receipt of documents.

In Italy the CNIPA has regulated a model of reference for the interoperability and the applicatory cooperation for the Public Administration named “Architecture of the Public System of Connectivity and Cooperation (PSC)”; the Public System of Cooperation (PSCoop) is a set of technological standards and infrastructural services whose objective is enabling the interoperability and the cooperation of the information systems for the fulfillment of administrative actions; the services offered aim at creating a groundwork to which all the Regions can connect in order to use and distribute services through standard protocols, with rules of safety and access that are shared and with a prearranged and monitored quality of the service.

Many Regions and Local Bodies have been equipping themselves to take advantage of the offered services and many initiatives promoted by the Ministry of Innovation are leading to the sharing of the models and the solutions adopted in order to achieve in a short-term period a real solution of interoperability.

Dematerialization Processes

Note that all the e-Gov applications so far described have dematerialization activities as a common and fundamental factor: information, previously stored using graphic marks on material (paper) supports, is made immaterial using a codified electronic representation, and can be nowadays stored on several digital supports such as memories, magnetic or optical disks, tapes or other mature technologies nowadays in use.

Dematerialization is not only a normative and technological challenge but also an organizational matter involving various human resources. The transformation of a bureaucratic organization based on paper into one based on electronic documents is not easily achievable according to general models that are exportable among the organizations themselves.

So far, we have described the main characteristic of the e-Gov system, in particular, we note that e-Gov processes are usually characterized by a huge quantity of paper documents that need to be properly managed, stored and distributed. In order to reduce the huge amount of hard papers for optimizing information communication in terms of consumed time and resources, it is widely agreed that a semantic-based dematerialization process will greatly enhance e- Government systems and application procedures.

The dematerialization process implies the application of syntactic-semantic methodologies in order to automatically transform the unstructured or sometimes semi-structured document into a formally structured, machine readable records.

The core aspect related to a novel and efficient dematerialization process is the idea standing beyond the common document concept, that can be defined as the representation of acts, facts and figures directly made or by means of electronic processing, and stored on a intelligible support. In other words, a document consists of objects such as text, images, drawings, structured data, operational codes, programs and movies, that, according to their relative position on the support, determine the shape and, consequently the structure of the document itself through the relationships between them. During the various and different e- Government processing phases, that are really different from an application domain to another, a document is processed and eventually stored on various kinds of media, properly defined in order to archive and preserve papers, photographic films and microfilms, VHS cassettes, Magnetic Tapes, DVD disks, and more.

In the following we will provide a novel model for digital documents and we describe a system for multimedia document management, in particular for those regarding archiving and long term preservation.

A document model for e-Government

A document managed in e-Government information system is usually composed by different multimedia data types, as images, text, graphic objects, audio, video and composite multimedia. This is usually related to two main problems: a multimedia document contains heterogeneous information contents and has to manage different formats: in addition, depending on the authorities

which manages the document itself, the same information content is presented in multiple ways, using several presentation formats.

For this reason, in order to opportunely manage and preserve the real useful information contained in a certain document despite the required different presentation formats, it is necessary to provide a novel model for a multimedia document, pointing out:

- how to identify and characterize what is the minimal content of the document itself, given a certain normative context, and
- how to relate this minimal content to a presentation level, depending on different users at different times.

The proposed document model is composed by several layers, as described in the following.

- *Data Management Layer*: describes the semantic minimal content (or kernel) of a document, usually codified by different media types. This layer manages the different data types, furnishing all the necessary functionalities and facilities operating over a certain single media; for example, information extraction and indexing over texts, images, videos, audios and son on.
- *Integration layer*: provides a proper integration of the heterogeneous data sources, having the aims of regulating the coexistence of the different objects within the context of a single document.
- *Presentation layer*: this layer regulates the way in which the information has to appear to a single user within a certain context in different times.

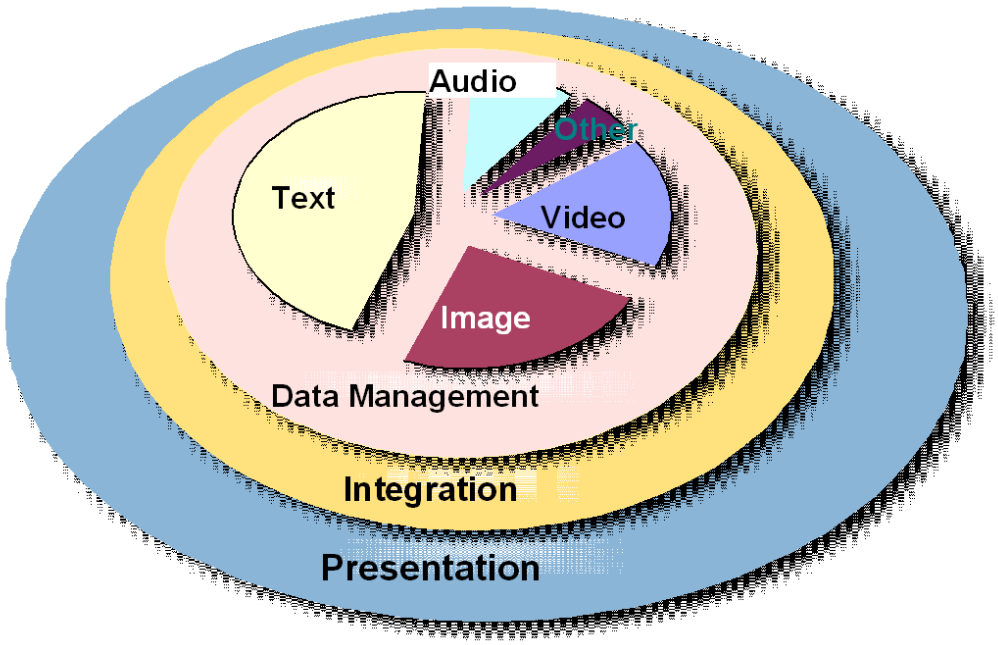

Figure 1:The Document Model

Note that usually the juridical validity of a document is nowadays provided on the whole file document, without any discrimination between the effective content and the different ways in which it is showed out. Differently, our model try to explore the possibility of giving juridical validity to the single level of the documents, thus giving the possibility of validating both the minimal content alone independently from the presentation layer and the complete presentation content.

A System Architecture for e-Government information system

Considering the theoretical aspects depicted in the previous sections, we are in a position of sketching a novel architecture for supporting e-Government activities: in particular, we propose a multimedia information system that integrates and processes different multimedia data types (as images, text, graphic objects, audio, video, composite multimedia, etc.) and provides facilities for indexing, storage, retrieval, control of multimedia data together with long term preservation strategies [1,2,3].

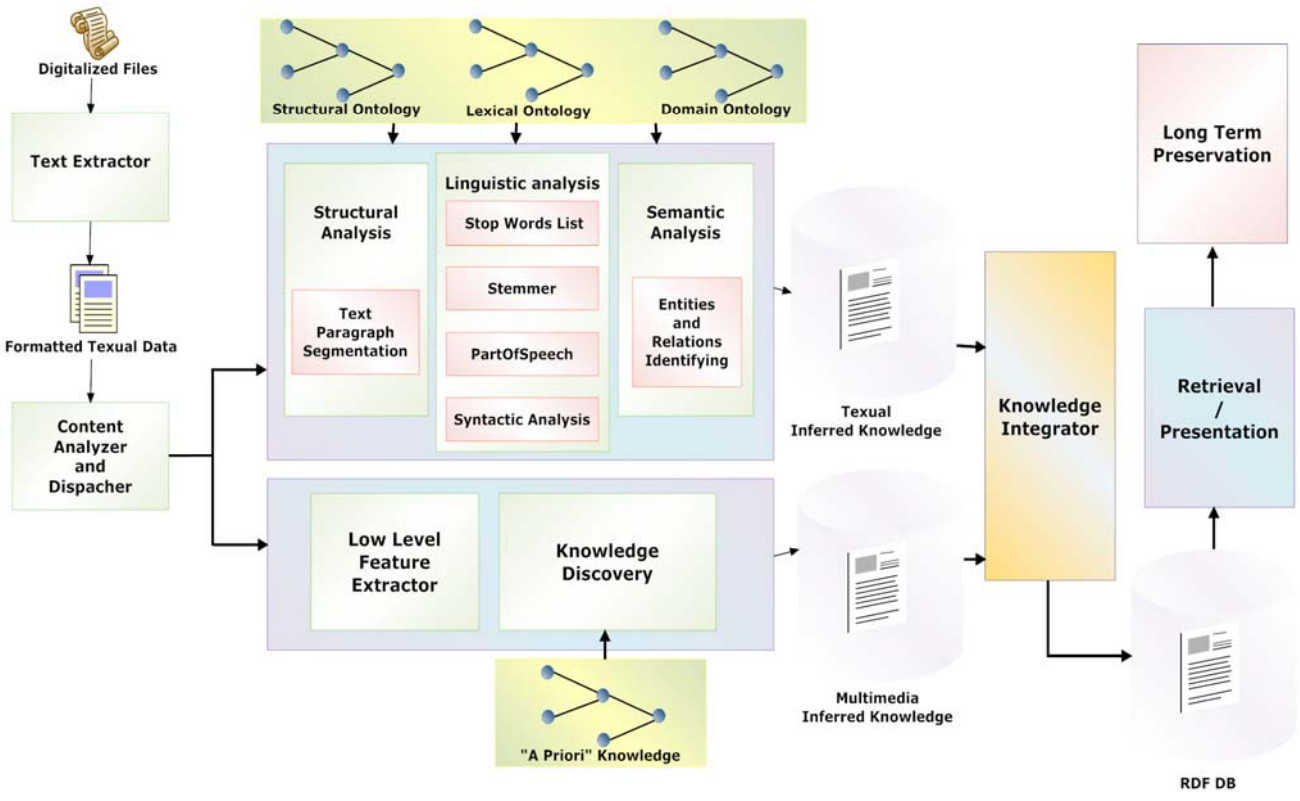

Fig. 2. The proposed architecture

The architecture of the proposed system, as shown in figure 1, is constituted by modules delegated to manage the *Information Extraction and Indexing* process and those related to *Retrieval and Presentation* applications. The knowledge associated to E-Government document activity is codified using appropriate *ontology repositories*.

In the current implementation of the system, we have realized three main separate subsystems that are responsible of information extraction and presentation tasks: one for the text processing operations, another one for processing the other kinds of multimedia information, in particular images, and the last one for

presentation aims, in according to the normative requirements of public administrations. The features of text and image management subsystems will be described in the following.

Text Processing Module: automatic extraction of RDF information triples from unstructured documents

The *Text Processing Module* extracts relevant information from the documents of the E-Government domain, starting from the analysis and the processing of the textual content of the submitted input document.

This module is based on both linguistic and statistical approaches for the early stages, and semantic function for the recognizing purpose. The semantic methods make use of a knowledge domain built over the top of an ontological system in order to control identification and extraction of relevant words in the text, representing the instances of the concept of interest. The text processing procedure is composed of several stages[6]: (i)*Text extraction*, where the plain text is extracted from the source file; (ii)*Structural analysis*, where the textual macrostructures are identified for text sections recognition; (iii)*Lexical analysis*, where each text element is associated with a grammatical category (verb, noun, adjective etc.) and a syntactic role (subject, predicate, complement, etc.); (iv) *Semantic analysis* where, proper concepts are associated with discovered entities and relations among them, by means of structural, legal domain, and lexical ontologies. The output of such procedures is a semantic annotation codified by RDF triples.

The Multimedia Processing Module: automatic annotation of images using visual information and pre-defined taxonomies

The goal of the *Multimedia Processing* subsystem is to automatically infer useful annotations for images looking at their visual content and exploiting an "a priori knowledge" (obtained in the training step of the system) in the shape of pre-defined taxonomies of image contents[4].

To such purposes, each image, belonging to a given concept (category) of the a-priori taxonomy, undergoes a particular indexing process, where in a first step a low-level description is obtained and then in a second one an apposite indexing structure is created/updated for facilitating the successive retrieval and annotation tasks. The indexing process can then support the *Knowledge Discovery* task (i.e. the "category detection" procedure presented in [5]), which automatically discovers those concepts of the a-priori taxonomy that better reflect the semantics of the input images. The obtained information can be thus used as useful annotations for each image, in order to infer knowledge about the content of the images, that is codified in a multimedia ontology (taxonomy concepts + images). The inferred knowledge is recorded using RDF triples.

The Integration and Presentation modules: merging knowledge from heterogeneous multimedia data and delivery of e-docs in different formats

The objectives of the *Integration and Presentation* modules are: from one hand, to merge in a unique "container" the heterogeneous knowledge coming from text and multimedia data, and from the other one, to delivery the content of e-docs in different formats.

In the current implementation of the system the integration module uses a human-assisted semiautomatic approach to instantiate relationships among concepts of the different ontologies. The result of such a process is an ontology that contains all the knowledge related to the e-Government documents.

The presentation module works on the top of such ontologies and exploiting the set of relations about structure of multimedia assets and e-gov documents, in order to present and delivery to final users the content of an e-Government document in different ways.

Conclusions and future directions

In this paper we presented a model for digital document suitable for e-Government activity, and a system for management of multimedia information related to e-Government procedures. At the moment we have implemented a prototypal version of the system that realizes the described information extraction and presentation tasks. Future efforts will be devoted to implement the other modules of the system and to obtain experimental results that validate the proposed approach.

References

1. F. Amato, A. Mazzeo, A. Penta, A. Picariello, "Knowledge Representation and Management for E-Government Documents", Book Chapter of E-Government Ict Professionalism and Competences Service Science, pp.31-40, Springer Boston, 2008.
2. W. I. Grosky (1997), "Managing Multimedia Information in Database Systems", Comm. Of ACM, vol. 40, n.12.
3. D. A. Adjeroh, and K. C. Nwosu (1997), "Multimedia Database Management - Requirements and Issues", IEEE Transaction Multimedia, pp. 24-33, July-September.
4. M. S. Lew, N. Sebe, D. Djeraba, and J. Rain, (2006) "Content-based multimedia information retrieval: State of the art and challenges", ACM Trans. Multimedia Comput. Commun. Appl, vol. 2, n.1, pp. 1–19.
5. G. Boccignone, A. Chianese, V. Moscato, and A. Picariello, (2008) "Context-sensitive queries for image retrieval in digital libraries.", Journal of Intelligent Information Systems, vol. 31, Issue 1, pp. 53–84.
6. F. Amato, A. Mazzeo, A. Penta, A. Picariello, "An information system for the extraction of relevant information for e-government activity", sebd, pp.366-373,2008 Sixteenth Italian Symposium on Advanced Database Systems, 2008.