

Core Public Service Vocabulary: The Italian Application Profile

Gabriele Ciasullo, Giorgia Lodi, Antonio Rotundo
{ciasullo, giorgia.lodi, antonio.rotundo}@agid.gov.it

Agency for Digital Italy

ABSTRACT

This paper introduces an on-going national initiative, carried out by the Agency for Digital Italy (AgID) in accordance with the relative legislation, for the definition of the Italian catalogue of public services. The catalogue has three main objectives: (i) it can be used to facilitate the discovery of public services by citizens and enterprises; (ii) it provides public administrations with a comprehensive platform through which sharing best practices on services, and building a community that discusses and potentially re-uses those best practices; and (iii) it can be used by AgID itself in order to monitor the degree of standardization and digitalization of the services of the public sector, thus reporting to the political level for strategic decision fine-tuning purposes.

The catalogue is defined in terms of metadata that contribute to the specification of the so-called Core Public Service Italian Application Profile. The metadata can be specified by public administrations (be they local or central) to represent their available, or under-development, online and physical public services. The core public service Italian application profile is defined through the use of core vocabularies, as released by the European Commission in collaboration with W3C. In particular, the paper presents the current preliminary data model of the Italian profile that is mainly based on the core public service vocabulary and its application profile, although other core vocabularies are considered (e.g., core location organization ontology and registered organization vocabulary).

1. Introduction

In the Italian digital agenda context, the role of public services, designed around end-users needs, is assuming an increasingly importance. The presence of several under-development national shared infrastructures (e.g., the Italian public system for e-identification, the Italian public system for e-payment, the Italian public platform for smart cities, to cite the most important ones) is posing the challenge for a common representation of the concept of user-centric public services.

The stratification over years of, also notable but isolated, initiatives carried out at different administrative levels led to a very heterogeneous scenario. This heterogeneity entailed the lack of (i) a clear and agreed classification of public services, (ii) a shared naming of the same public services offered by a variety of administrations (e.g., municipalities); and (iii) a standard definition of the elements that allows one to clearly represent the concept of “public service”. In this situation, it turns out that interoperability, semantic but even technical, might be very difficult to obtain.

In the light of these observations, the Agency for Digital Italy (AgID) is currently operating on a wider scale with the intent to provide public administrations with a general, common definition of interoperability profiles for data and metadata. Following the concrete initiatives promoted at the European level in working groups for semantic interoperability, AgID firstly started the definition of an Italian data catalogue application profile (DCAT-AP IT) and, very recently, the specification of the so-called *Core Public Service Vocabulary Italian Application Profile (CPSV-ITAP)*, which this paper introduces.

CPSV-ITAP has the ultimate objective to cope with the earlier mentioned heterogeneity scenario, through the specification of the minimum set of metadata required to document public services made available to citizens and enterprises by public administrations. In particular, AgID claims that defining

such a common profile can be a crucial step for achieving harmonization in the public service data representation and can ease the exchange of data on public services in national shared infrastructures, thus improving semantic interoperability.

The profile is the basis for the design and development of the Italian public service catalogue. Initially conceived for the use only in the Italian smart city platform, (since so it is prescribed by the Italian legislation), the scope of the public service catalogue has been extended so as to make it an actual authentic source for national common infrastructures like the Italian public system for e-identification and possibly the Italian public system for e-payments.

During the preliminary phases of the design of the catalogue, its main objectives emerged: (i) it can be used to facilitate the discovery of public services by citizens and enterprises; (ii) it can provide public administrations with a comprehensive platform through which sharing best practices on services, and building a community that discusses and potentially re-uses of those best practices; (iii) it can be used by AgID itself in order to monitor the degree of standardization and digitalization of the services of the public sector, thus reporting to the political level for strategic decision fine-tuning purposes.

Related works. Similar initiatives have been carried out in other countries. In particular, it is worth mentioning here the Brazilian public service catalogue [10], a Spanish ontology for public services [9] that has been used by the City of Saragossa in order to document its public services, and other Italian initiatives such as the one of the Lombardy Region, which defined an RDF ontology of its public services based on the Core Public Service Vocabulary [8].

Structure of the paper. This paper describes (i) the *methodology* adopted so far for the definition of the catalogue underlying data model (Section 2); the preliminary version of the *data model* (Section 3); (iii) the role of the core vocabularies, released by the European Commission in collaboration with W3C, in the definition of the minimum set of metadata required for describing public services and other context-based entities (e.g., organizations, locations and persons) (Section 3); and, finally, ongoing works on it (Section 4).

2. Definitions and methodology

In the context of this paper, with the term *public service* we mean “a set of mandatory and discretionary acts that are carried out by a public administration (or on behalf of a public administration) for the advantage of citizens and enterprises”. Public services can be online and physically available, active and under-development services. Different use cases have been considered. For the sake of brevity, we report just one of these use cases as an example.

Use case. A citizen wishes to change his/her residence from a municipality of origin to a new municipality. The citizen then presents a request to change to the new municipality. The request can be issued through different channels: via e-mail (traditional or certified electronic mail), via fax machine, and via help desk of the municipality. When a request is started, the citizen is required to provide a number of input documents following specific rules that are published by the municipality; namely, a request form filled in with all his/her personal data and a copy of his/her ID card. The output produced by the service is the change of residence according to the new information provided by the citizen.

Methodology. The methodology we followed in the definition of the data model for the public service catalogue (the CPSV-ITAP) consists of different steps depicted in Figure 1.

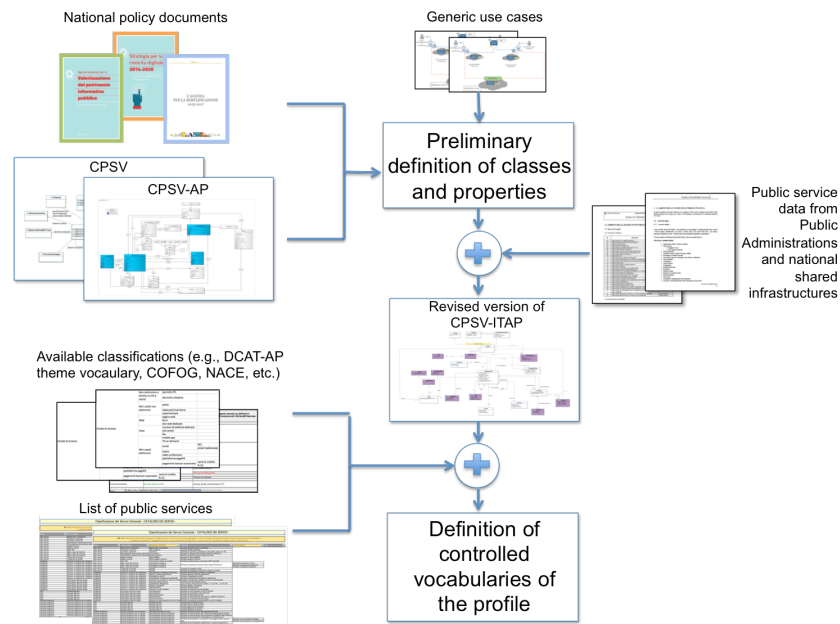


Figure 1: The methodology adopted in the definition of the data model of the Italian public service catalogue

The work started by assessing general uses cases and a number of national policy documents that illustrate the action plan for the Italian digital agenda. In these documents, key performance indicators (KPIs) and components of national shared infrastructures and public services are detailed. These have been used as primary sources of requirements necessary to envisage the definition of the core elements of the public service catalogue.

At the same time, we re-used the results obtained in the context of the European working groups for the definition of the core public service vocabulary [1] and its application profile [2]. Assessing requirements and already available vocabularies brought us to a preliminary identification of the classes and properties of the data model (top-down approach). As shown in Figure 1, a successive fine tuning of the model has been undertaken in order to reflect the actual characteristics of real public services available from a variety of administrations (bottom-up approach). In particular, we considered public services offered by municipalities, regions and central administrations as published in their institutional web sites. We also took into account the list of services envisaged in the experimental phase that is currently activated for the development of the Italian public system of e-identification, recently launched by the government. This bottom up approach led to changes of the preliminary data model and to the definition of a second version of it that we called Core Public Service Vocabulary - Italian Application Profile (CPSV-ITAP). Please note that the discussion on the final version of the profile is still open in our country; therefore, the model we introduce in this paper can be subject to modifications in the upcoming months.

When defining classes and properties, classifications have been identified; e.g., the main *theme* of a service, the channels through which services can be provided to end-users, the types of input documents, etc.

Common and shared classifications are key elements for enabling semantic interoperability and linking among different types of data [2]. To this end, we started to define controlled vocabularies using already available classifications (even official ones as released by the Italian Institute of Statistics). Examples include COFOG (published as linked open dataset in our country [4]), NACE [5], the theme vocabulary defined in the context of DCAT-AP [6], etc. At the time of this writing, the consensus has been achieved for both the main theme classification and the possible types of channels of public services. As far as the main theme classification is concerned, we decided to re-use the theme vocabulary of DCAT-AP with some small variants in the theme grouping [6]. This approach allows us to provide harmonization between services and data upon which services can be built. As for the types of channels, we re-used the classification available from UK local service catalogue [7], personalized to the Italian context

3. Core public service vocabulary – Italian application profile

Figure 2 illustrates the UML diagram of the Core Public Service Vocabulary - ITalian Application Profile (CPSV-ITAP).

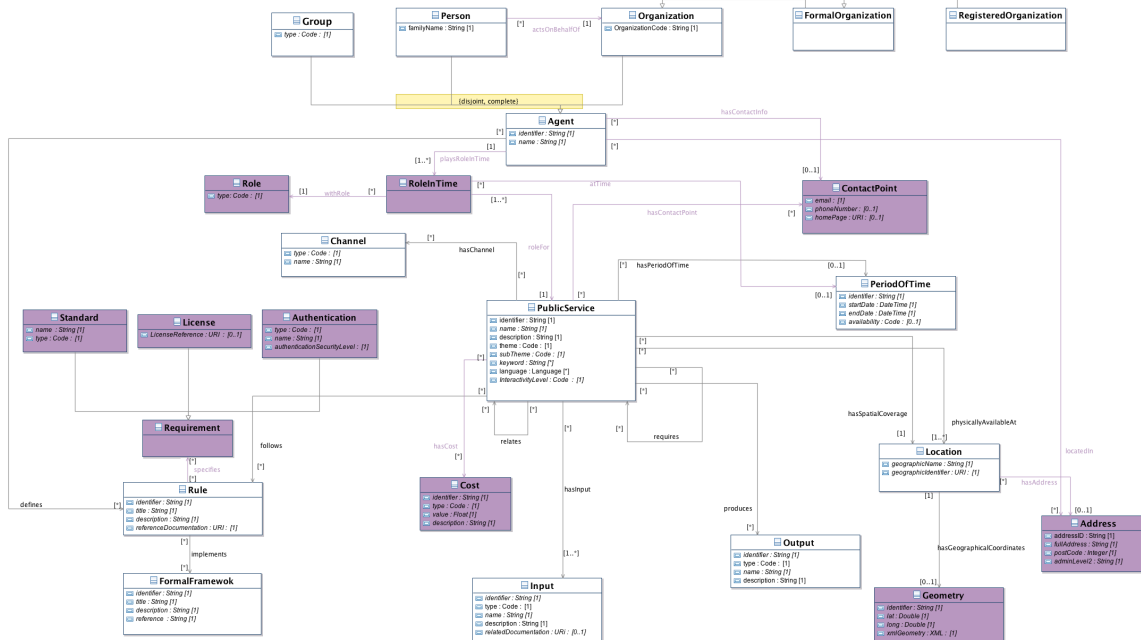


Figure 2: UML diagram of the CPSV-ITAP

The profile has been entirely defined based on the elements of the Core Public Service Vocabulary (CPSV). However, in order to cope with the specific national requirements, derived from policy documents as earlier discussed, the profile extends the CPSV with a set of additional classes and properties. In the UML diagram of Figure 2, the classes that we introduced are in purple whereas the new properties are in italic. Mandatory and optional classes and properties are shown through their cardinalities. For the sake of brevity, we do not detail all the properties and elements of the profile; rather, we highlight those that we explicitly used to extend the CPSV.

The central class is *PublicService*. Among its well-known properties as identifier, name and description, the class includes: the theme, as described in Section 2 and equivalent to the property “dterms:type” of CPSV specifications, the sub-theme, which is a more detailed item of the classification (in the CPSV application profile the term *sector* is used to indicate this property), a set of keywords that can be used to facilitate search functions in a possible catalogue, the language and finally the so-called *InteractivityLevel*. This latter property is meant to represent the fifth degrees of sophistication of online public services [11], which can be an effective indicator of the enhancements towards the provision of a comprehensive online services ecosystem.

A significant novelty we introduced with respect to CPSV is a more extended management of roles played by agents on public services. In particular, we assessed use cases where organizations of different types (public and private) exercise over time a different role on the service (e.g., competent authority, provider, managing authority). To this end, the design pattern of the time-indexed situation [13] is applied and in particular the ontology describing, among others, roles [12] is considered as reference modelling schema. As for the representation of the agents, we followed the models offered by Org [14] and RegOrg [15], both ontologies derived from the previous Core Business vocabulary.

An additional novel element is the extension of the *Rule* class already included in CPSV. *Rule* specifies elements that in CPSV are just mentioned; namely, authentication, license and standards. Authentication is a key element for the Italian profile, since it allows us to model the data required in the context of the new Italian public system of e-identification such as the security level requested to authenticate to the services (in particular there are three levels: credentials, one-time password, biometrics data). License and standards are crucial concepts for promoting a possible reuse of the

service in other administrations. Finally, four new elements of the profile are described in the following. Firstly, we introduced a slightly different modelling of the location at which the service is physically available. In this case, the property *physicallyAvailableAt* links the *PublicService* class to *Location* (the property can be maintained as a sub-property of *hasChannel*). Secondly, as for the *Location*, we re-used the modelling schema offered by Core Location [16] and in particular the classes *Address* and *Geometry* used to represent the physical address at which the service can be exploited and the geographical coordinates of the area covered by the service, respectively. We believe this latter element is of utmost importance to provide end-users (citizens and enterprises) with a catalogue that is able to geo-localize services of the public sector. Thirdly, once more we adopted the specifications of the CPSV application profile when representing possible *Costs* associated with a service. We believe this entails ease integration with other procurement information of the public services that might be available. Fourthly, although it is still under discussion and thus subject to changes, we currently introduced a *ContactPoint* class that provides email (mandatory), telephone number and home page references, only, through which establishing a contact with the service. The same class can be used to indicate possible contact references of agents.

4. Concluding remarks

This paper described the Italian on-going experience on the definition of the Italian public service catalogue data model. The model is built upon core vocabularies (core public service, core location, org and regorg) released in the context of the ISA programme on semantic interoperability. Currently we are working in order to finalize both the model and all possible controlled vocabularies that are necessary to complete the profile (e.g., the input documents, the types of authentication, etc.). We are also planning to release the OWL ontology of the profile to be published as documentation of the overall public service catalogue. In parallel, we are designing the comprehensive platform of the catalogue that will be also published as (linked) open data for the re-use by anyone.

Bibliography

- [1] ISA working group, "Core Public Service Vocabulary", version 1.01, 2013.
- [2] DG DIGIT/ISA Programme, "D02.02 – Definition and development of a data model for description of the services related to key business events", version 1.0, 2014.
- [3] G. Lodi, A. Maccioni, M. Scannapieco, M. Scanu, and L. Tosco, "Publishing Official Classifications in Linked Open Data", in *the proceeding of the 2nd International Workshop SemStats, 2014, in conjunction with ISWC 2014*.
- [4] Agenzia per l'Italia Digitale, COFOG Linked Open Data, available at <http://spcdata.digitpa.gov.it/data/cofog2009.rdf>
- [5] Eurostat, NACE Classification, available at http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Statistical_classification_of_economic_activities_in_the_European_Community_%28NACE_%29, 2015
- [6] ISA working group, DCAT application profile for data portals in Europe, available at https://joinup.ec.europa.eu/asset/dcat_application_profile/description, 2015
- [7] UK local government association, ESD toolkit – standards channels list, available at <http://standards.esd.org.uk/?uri=list%2Fchannels&tab=details>, 2015.
- [8] Agenda Digitale Lombarda, Linked Open Data, versione 1.0, 2014, available at <http://www.agendadigitale.regione.lombardia.it/shared/ccurl/875/528/Agenda%20Digitale%20Lombarda%20-%20Linked%20Open%20Data.pdf>,
- [9] CTIC, "Vocabulario de trámites y servicios públicos", available at http://data.fundacionctic.org/vocab/infraestructuras/servicios_20120823.html, 2015.
- [10] Brazilian public service portal, available at <http://www.servicos.gov.br/>, 2015.
- [11] Capgemini, "The User Challenge Benchmarking The Supply Of Online Public Services", 2007, available at http://www.ut.is/media/utvefur-skjol/CapGemini_2007.pdf
- [12] Peroni, S., Shotton, D., Vitali, F. (2012). Scholarly publishing and the Linked Data: describing roles, statuses, temporal and contextual extents. In Sack, H., Pellegrini, T. (Eds.), *Proceedings of the 8th International Conference on Semantic Systems* (i-Semantics 2012): 9-16. New York, New York, USA: ACM. Open Access: <http://speroni.web.cs.unibo.it/publications/peroni-2012-scholarly-publishing-linked.pdf>, available at <http://www.sparontologies.net/ontologies/pro>, 2015.
- [13] Time indexed situation design pattern, available at <http://www.ontologydesignpatterns.org/cp/owl/timeindexedsituation.owl>, 2015.
- [14] Organizatio Ontology, available at <http://www.w3.org/TR/vocab-org/>
- [15] Registered Organization Vocabulary, available at <http://www.w3.org/TR/vocab-regorg/>,
- [16] Core Location Vocabulary, available at <http://www.w3.org/ns/locn>