Dutch eGovernment: Past and Present Semantic Initiatives

Emile Van der Maas, NL

An example of semantic in-operability

After World War II the Dutch government encountered problems with Semantic Interoperability during the first ICT automation wave and rise of the internet in the 1980's. On basis of the age-old political autonomy of government organisations, over 1600 registries existed, all based and built on domain specific laws, having all their own definitions, interpretations and meaning of what was in reality the same object, making exchange of data via Internet difficult if not impossible.

As an example there was the case for an average pig. On basis of domain specific laws and regulations from the ministry of Agriculture, Transport, Food, Public Health and in the domain of consumer goods, more than 40 definitions of 'pig' co-existed in systems and processes. With Internet and digital exchange of data, it became apparent that the old way of working within autonomous organisations and systems, encountered serious interoperability problems.

The need was felt to re-organise government to maximise the benefits of ICT

Goals of Other Government

In order to deal with the negative side effects of the old way of working a program was initiated in the 1990's, called the Other Government. Main goals were to achieve a government that:

- Does not ask for information it already has.
- Is focused on better service towards citizens and businesses.
- Will not allow its facilities to be misused.
- Is well-informed.

- Is efficiently organised and 'in control' of its internal affairs.

To realise these goals, an ICT working-organisation was formed, called stichting ICTU. ICTU still is the central organisation responsible for designing, developing, realisation and implementation of ICT components or assets. www.ictu.nl

Work began at ICTU in 2001 within the program 'streamlining base-data' on the restructuring of the 1600 + decentralised registrations. The idea was to create national centralised registrations, containing all possible data about persons, addresses, cars etc. As a start, definitions on generic data to uniquely identify persons, cars, buildings etc. was standardised and use of these data in base-registrations was made mandatory by law.

Though the need for streamlining of data and definitions was not questioned, there was little support however for the idea of a strongly centralised and rigorous standardisation. The rigorous approach hardly fitted the decentralised centuries-old way of working at the Dutch government. Moreover maintenance was considered a major issue as was the cost of implementing.

Nonetheless at the end of 2003 a stable semantic core was laid down with the appointment by law of the first 6 base-registrations. A further 9 followed soon. Its content eventually consisted of the generic data to uniquely identify and authenticate objects throughout all government processes.

It was recognised however that in this way with just base-registries, Semantic Interoperability could not be realised. The connection still had to be made with the context of use and semantic meaning in everyday processes. It also was clear that for automated systems standardisation of definitions is a great thing, but you cannot force civil servants or citizens to adjust in everyday speech and use of information to this one formal definition. It is a more realistic approach to recognise that there always will be variation in meaning, depending on the context of use.

So how to realise coherence between semantics within a more decentralised approach?

The need for coherent architecture: NORA

Thus work was started at the end of 2004 on an information architecture- as part of an ICT architecture for the Dutch government as a whole. This Dutch Government Reference Architecture or NORA, was based on the European Interoperability Framework (E.I.F.) 1.0 and adopted its main principles. The challenge behind NORA being to achieve interoperability, better chain management, function clarification, the by law mandatory re-use of data, and the strive for more uniform definitions and better harmonized laws.

The aim was and still is to work toward a service-oriented architecture, in which major components like base registries delivering single-point access to nationwide basic services for all government organisations. This certainly also stresses the need to find a solution for Semantic Interoperability: how to exchange data based on domain specific law-defined definitions and values and still being able to know how to interpret these data, whilst re-using them in other domains?

Outlines for a federated (semantic) architecture

The E.I.F. principle of subsidiarity outlined the organisational model of the data- and information architecture for Dutch eGovernment. Thus also providing a better alignment with the traditional decentralised way of working. Within this federated information architecture the connection can be made with domain-specific meanings and values in working processes and life-events. The authentic data in base-registrations fulfilling their linking-pin role in this to uniquely identify and authenticate objects in all of these processes and events.

From the political goals from the Other Government programme, it was stressed that the main focus shouldn't however be on the internal organisation of government and the formal, technical semantics, but on giving the citizen and society better access to government information and products. Web-metadata, clear description of products and, for example a good search engine and and personalised information page, also have to support life-event driven semantics.

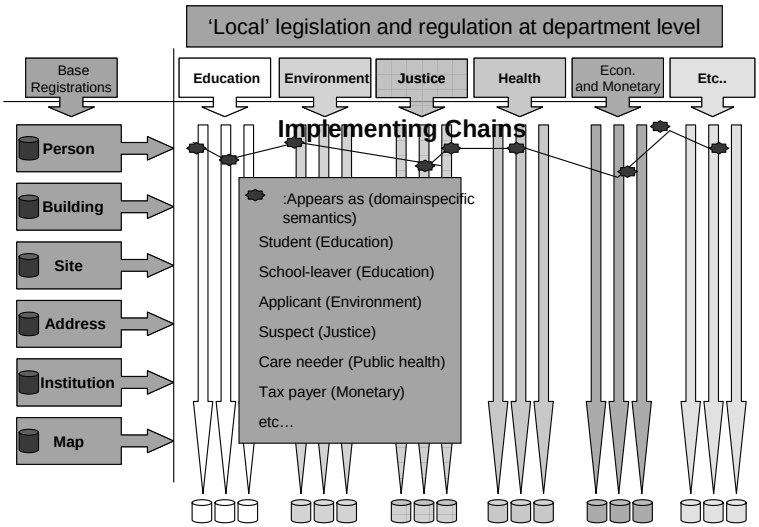
The role and function of metadata

As a solution for Semantic Interoperability on all levels between databases and documents, processes and life events, the Dutch government is now looking into the possibilities of metadata to achieve this. In NORA 2.0 an international standard for metadata was proposed. This ISO 23081 standard focuses on a standardised way to make a metadata schema and within that how to uniquely identify contexts (domains) and domain-specific meanings and values. Perhaps as a parallel with the role of the authentic data in base-registries, standardisation of metadata schema for contexts, will make it possible to uniquely identify a domain and domain-specific meanings and values.

The agreement on this kind of metadata schema for unique context identification, thus opening the way to focus on the mapping rules between contexts.

Thus in a federated semantic architecture, domain specific semantic meanings and values can co-exist, on the condition that:

- the whole of government agrees on which domains there are to uniquely identifiable and how we're going to identify them. Organisations need to decide to witch domain



their processes relates and thus to witch semantic models. The discussion on what domains there can or should be identified, recently received a new impulse with the proposal from SEMIC.EU for a generic Member State activity schema. Within the Netherlands we look to this schema as both a mapping table as a base-model for the development of domain specific semantic assets and models, like taxonomies.

- Each domain-specific semantic model, as written out in a data dictionary, UML schema or worked out taxonomy, has to connect on a national level, to at least one base registration. The base registrations thus fulfilling the role of national, cross-domain semantic assets providing the generic message core components that are mandatory to be inherited in domain-specific semantic models. Thus Semantic Interoperability within the national context can be achieved.
- The whole of Dutch government agrees on a proposed set of metadata that have to be (automatically) recorded in order to enable unique context-identification. The identifiers for context containing, amongst others, metadata about the business-context: what organisation is or was responsible for the registration or mutation of data. What is the legal framework, as part of the business-context, since domain specific laws define meanings and values at a more detailed level.

At this moment, a first try-out of the possibilities of metadata in achieving Semantic Interoperability is worked out at the level of the national ministries. In a baseline information-policy for all of the 12 ministries, metadata are the core to uniquely identify context specific semantic meaning and values. The outline for this concept base-line is expected to become available at the end of June 2008. If reactions are positive and agreement is achieved on the implementation path and time-scheme, the proposal will be rolled- out to be implemented for all of governmental organisations, both centralised as decentralised.

Other work on Semantic Interoperability and assets is being done on a national data dictionary that will provide in time the generic XML core components and access to the related domain-specific data dictionaries and taxonomies. It is estimated that this central catalogue will be operational in 2009. At this moment a national services registry, containing, amongst others, the generic data services from the base registrations, nears completion. Also a first domain-specific

taxonomy is available for the annual fiscal year reporting (xBRL). Work also has started on a web-ontology aiming at giving citizens better insight in the organisation and working of Dutch government.

This article gave but a short insight on past and current initiatives on Semantic Interoperability at the Dutch eGovernment. If you would like to know more on Dutch eGovernment initiatives and projects on Semantic Interoperability and assets under construction, you can contact us at:

Knowledge center eGovernment via www.e-overheid.nl on www.e-overheid.nl more detailed information can be found about Dutch eGovernment programmes and projects, NORA and the base registrations. You can also post an e-mail with your question.

Emile van der Maas, NL, *studied History at the University of Leiden, Netherlands, specialising in Methodology and philosophy of science. In 2004 he started as an ICT architect at ICTU for the programme streamlining of base-data. He was one of the founding architects at ICTU for the NORA, the Dutch Government Reference Architecture. Currently he is an architect at ICTU at the Dutch eGovernment Knowledge Center and works on semantics, Semantic Interoperability, base-registrations, metadata and develops NORA further.*