

In so many words

Semantic interoperability across organisations and domains: an ‘open’ perspective

Paul Oude Luttighuis & Mariette Lokin¹

1. Introduction: getting a grip on meaning

Many people consider semantic interoperability to be key for realisation of an effective and efficient electronic government. No doubt they are right: the co-operation between government agencies that is needed to effectuate adequate e-services can never be brought about if we don't understand each other. Taking this into account, the Dutch policy on e-government holds an important place for semantic interoperability. Before we go into this, it is useful to define the term semantic interoperability, as it is subject to different meanings itself. Mostly the term is used in computer science as a synonym for "Computable Semantic Interoperability". In this sense, it is the ability of computer systems to communicate information and have that information properly interpreted by the receiving system in the same sense as intended by the transmitting system.

However, we prefer to use the term in its more general sense, as the ability of any communicating entities (not only computers) to share unambiguous meaning. In this broader sense, the sender (be it a person or a machine) must be able to reliably transmit all necessary information, the receiver must be able to correctly interpret it, and both must be aware of and agree upon each others behaviour or context for given interactions.²

Now, back to the issue of semantics in Dutch e-government. The ‘spine’ of Dutch e-government is a substantive system of basic registries, e.g. on organizations (businesses), persons, land, geo-information, containing basic information needed in most e-services. It creates an important condition for operationalizing Dutch e-government principles such as ‘single delivery, multi use’, in which we see semantic interoperability clearly reflected. Still, the basic registries only represent the mere visible and tangible side of semantics in e-government. In fact, semantics are omnipresent in this field, though often not explicitly addressed or even acknowledged.

So even despite the focus on semantics in Dutch e-government, there still tends to rise confusion on what semantic interoperability entails and how it can be brought about. One of the reasons for this is that semantics is an abstract subject. However intricate technology may be, it is at least tangible. Semantics, on the other hand, is hard to define and rapidly leads to semantic (!) discussions.

In this paper, we try to lift (at least the rim of) the veil around semantics and we present an initiative on open innovation in the field of semantics, aiming at bringing semantic interoperability, and thus better e-services, closer at hand.

This article consists of the following:

- Section 2, *Semantics: much more than data* introduces semantics as an issue and addresses some important aspects.
- Section 3, *Semantic interoperability: sharing meaning* shows what is required when multiple parties have to agree on meaning.

¹ Paul Oude Luttighuis is **principal researcher** at Novay (www.novay.nl) and Mariette Lokin is legal advisor at Logius, digital government service of the Dutch Ministry of the Interior (www.logius.nl).

² http://en.wikipedia.org/wiki/Semantic_interoperability

- Section 4, *Semantic models: describing meaning* describes the implications for the way in which meaning may be described.
- Section 5, *Towards semantic interoperability*, describes lessons that may be learned for realisation of semantic interoperability in the field of Dutch electronic government and presents an initiative for a public-private partnership, aiming at turning these lessons into practicable solutions by means of open innovation.

2. Semantics: much more than data

Half a word

What is clear about semantic interoperability is that it has to do with data. And, since the objective is interoperability, discussion soon turns to standards. But semantics is much more than just data and standards. Rigid data standardisation may even diminish semantic interoperability.

Information systems are not good at understanding. They even have trouble understanding entire words. When information traffic is left to communication between information systems, we have to make sure meaning comes across correctly. When this is done successfully, the systems and organisations are said to be mutually interoperable.

The ambiguous phrase ‘come across’ may be considered in all its meanings. Of course, information sent must first of all reach its destination via the technological means available: *technical* interoperability. The information must also be presented in a form that is legible to the recipient. Many people in the Netherlands will not be able to read a text in Cyrillic characters. Similarly, information systems rely on information in a readable form. Otherwise, they will be *syntactically* non-interoperable.

But technical and syntactical interoperability are only preparations for ‘the real thing’. Electronic information still does not come across if the meaning and intention of the information are misinterpreted. In the end, *semantic* interoperability determines whether the information traffic is meaningful and useful. Lack of semantic interoperability may render such information traffic not only meaningless and useless; it might even turn out to be risky, because information that is readable but not understood may lead to incorrect conclusions.

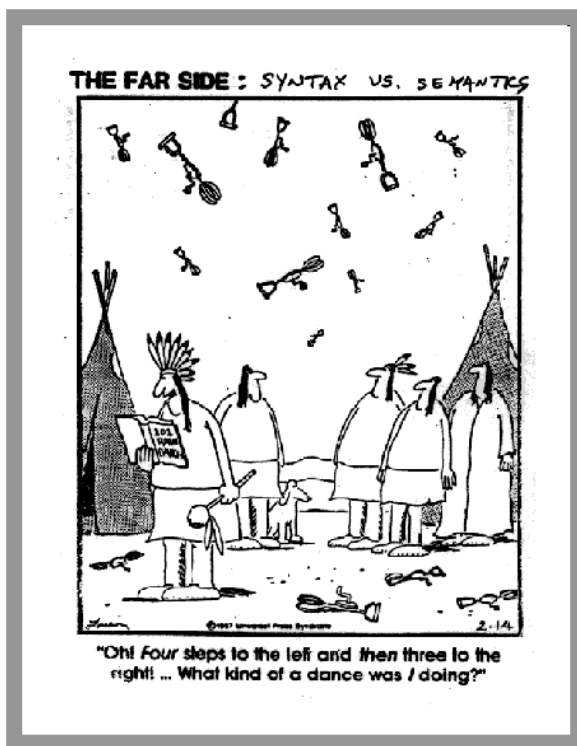
It is the ambition of the Dutch government to create an electronic government that to a large extent is based on communication between information systems. Basic registries must make electronic basic data accessible and maintain it for all government bodies, eliminating the need for citizens and businesses to provide the same information multiple times. In order to reduce the administrative burden, the volume of data and the frequency with which businesses are required to make it available to government bodies is growing. Citizens are increasingly urged to use government services via electronic portals to avoid being sent from pillar to post.

Thus, virtually every field of Dutch electronic government development is affected by the need for semantic interoperability: the basic registries, electronic communication between businesses and government institutions, service portals and all kinds of inter and intra-sector information chains. Within the scope of this article it is not possible to discuss the situations and implications in all these fields. In addition to taking up much more space, it would cause an awkward separation. After all, the term *company* used in the trade register may be related to the term *employer* used in the electronic payroll tax return, the term *business* used in the national system for environmental permits, and the *organisation* field in the electronic form used to report industrial waste. Data can be boxed, semantics cannot.

It is for these reasons that demand for semantic interoperability within the field of electronic government is growing. In its *Interoperability Agenda*³, the Dutch Standardisation Forum – an independent ‘think tank’ consisting of representatives of either government agencies and businesses in the field of IT – gives high priority to semantic interoperability. Concerning electronic information traffic between businesses and government bodies, the Wientjes Committee wrote in its report *Regeldruk Bedrijven en ICT*⁴ (*Business Regulatory Burden and IT*): “[...] government demand for information from business and the way businesses generate information are insufficiently aligned.” The recently updated *European Interoperability Framework*⁵ also reasserts the relevance of semantic interoperability.

This is not a pipe

Yet, bringing about interoperability is difficult. Reasons for this will occur in several parts of this article. First, semantic interoperability is often considered a matter of data design. This is not correct. Semantics is about meaning not only of data, but also of content, processes, services, (life) events and even (or maybe: primarily) the wording of laws.



For instance, the meaning of the term *employee* is relevant in the context of the interpretation of employee data, in income tax processes and in the event *commencement of employment*. The cartoon above shows what can happen if the meaning of business processes is not known. The unfortunate Native American duly follows the process (the dance), but it does not lead to the desired result (rain). He did not know the meaning of the process and therefore did not know what he was doing. This also is a matter of semantics.

It may only seem like a detail, but too narrow a definition of semantic interoperability will place the entire subject in the domain of system designers and IT specialists. Hence our preference for the more ‘general’ definition of the term we gave in section 1. Semantics provides the connection between data and the real world and between processes and their actual effect, regardless of whether they are automated or not. Data is only a reflection of the real world in an electronic mirror.

René Magritte has demonstrated the difference between a picture and reality in *La trahison des images*⁶. His picture states “This is not a pipe.” And rightly so: it is a picture of a pipe and not a real pipe. Data is just as treacherous as a picture of reality. Confusion of data with reality can create Kafkaesque conditions.

³ http://www.forumstandaardisatie.nl/fileadmin/OVOS/InteropAgenda-en_website.pdf

⁴ http://www.gbo.overheid.nl/fileadmin/OVOS/FS21-09-06_bijlage_-_Advies_ICT_Commissie_Wientjes.pdf. The Wientjes Committee is an advisory committee on regulatory burden, appointed by the Dutch cabinet and named after its chair, Bernard Wientjes (chairman of the largest Dutch entrepreneurs’ association, VNO-NCW).

⁵ <http://ec.europa.eu/idabc/servlets/Doc?id=31597>, page 32.

⁶ The treachery of images.



Semantics is about reality; about the domain of knowing and acting. Semantic interoperability in electronic government can therefore not be realised without the involvement of domain specialists, process specialists and legal specialists. This creates a second hurdle, as it these groups are not generally accustomed to co-operate.

Context: communication in perspective

The goal of interoperability is to make sure that electronic communication between multiple systems or organisations is useful and meaningful. If there is good reason to involve in (electronic) communication, there must be a common subject, problem or process that requires such communication: those involved share a context. But each person involved will also contribute a particular context – a perspective that determines the meaning someone attributes to particular terms. Context determines interpretation.

Five examples:

- A man received a summons for breaking a local regulation in the Dutch municipality of Heusden. It soon became evident that a reorganisation of local government was still in progress at the time of the supposed offence. At the time, the scene of the act was not yet part of the municipality of Heusden, and the local regulation was therefore not in force there. In this case, the point in time determines the meaning of the term *municipality of Heusden*.
- Are you a *foreigner*? For most readers, if they were asked this question at home, their answer would be negative. But the situation would be quite different if they were asked the same question by a local policeman while on a trip to the Far East.
- An example that is closer to home is the term *salary*, the meaning of which used to be quite different in tax legislation and in social security legislation in the Netherlands. In order to reduce administrative burdens, the term *salary* was harmonised in the so called ‘Walvis Act’⁷. This resulted in a common term for *salary* shared by the Tax and Customs Authority and the Employee Insurance Agency (UWV).
- Imagine you and your ailing adopted daughter go to the first-aid department of the local hospital. During the intake, you need to complete a form in which you are asked to enter the names of her *parents*. What do you enter? You may have parental authority, but if the question is asked in order to identify any hereditary correlations, this is not relevant information. Here, meaning is determined by the purpose for which the question is asked.
- The fire brigade receives a fire alarm: two *dwelling*s are on fire. When the fire brigade arrives in the street, there is only one property on fire. It turns out that the alarm was raised for two addresses in the same property. From the residential (and probably also the insurance) perspective, the property does indeed consist of two dwellings, but for the fire brigade it is only one.

⁷ Walvis is an acronym for *Wet Administratieve Lastenverlichting en Vereenvoudiging In de Sociale zekerheid* (Act on reduction of administrative burdens and simplification in social security) and the Dutch word for whale. This animal presented an adequate metaphor for the scale and impact of the project.

The list of examples could be made much longer, yet context is rarely stated explicitly. Often, one has to guess what the context is, and even more often an incorrect context is assumed, with all due consequences. People are sometimes aware of such ambiguities, but automated systems tend to just trudge on.

The reality of law

Government bodies have legal tasks. The activities of businesses are also affected by legislation, particularly when it concerns the information relationship with the government. This makes legislation an important source of meaning. This is not only true for the definitions of terms, which are often included in the first paragraph of an Act. The provisions of subsequent articles and the explanatory memorandum usually elaborate this meaning in detail.

Yet the wording of laws is not the final source of meaning for the processes and systems of government bodies. There are a number of reasons for this. Executive government bodies work with citizens, businesses and other real-world phenomena on a daily basis, and these do not readily comply with legal order. Even in the case of illegal conduct, it is not acceptable to ignore such behaviour or be unable to talk about it. Furthermore, legislation intentionally leaves room for interpretation and freedom to act, also for government bodies. Often, additional interpretation needs to take shape during the organisation of processes and systems. Finally, the wording of laws often contains the same terms for different concepts, particularly when laws are in different legal areas.

Therefore the semantic role of legislation is also contextual. The meaning of many terms – however zealous we strive for harmonization – often depends on the Act describing them. Legal context is very important, particularly for more artificial terms that stem solely from legislation. But, the legal context is not the only or final context.

3. Semantic interoperability: sharing meaning

Finding common sense

Semantic interoperability is mainly a matter of uniting and reconciling the contexts of the parties involved. This involves their history, their legal tasks, their environment and their situation. It is, of course, not necessary to add up all contexts and let everybody cope with every body else's context. It is rather a question of looking for elements shared by all contexts, common factors. Next to common factors recognised, it is still possible to acknowledge aspects that are specific to a particular context.

In the example of the term *foreigner*, this might work as follows. Two countries wish to communicate about foreigners, but attribute context-related and therefore different meanings to the term. Both countries find that the term *foreigner* is related to the term *nationality*. Based on this, they are able to establish a common meaning, for instance based on a table of valid nationalities. They then decide to communicate only about people's nationalities. They no longer discuss whether or not the person is a foreigner. The ambiguous term *foreigner* was too specific. The term *nationality* is used as the common factor.

In order to become semantically interoperable, parties involved must look for common factors. These might be outside of their own strict context, or, in terms of the illustration, they might be outside of your own box. Only if the scope is widened beyond a specific context, is there a fair chance of common factors being found. The only other options are to either force your own context upon the other party or decide not to communicate at all.

Scale matters

The greater the number of different contexts to be combined or reconciled, the more important it is to recognize and acknowledge these contexts. Many of the ambitions of the Dutch electronic government are indeed of a large or very large semantic scale.

First of all, the *basic registries*. Their objective is to satisfy – as much as possible – the need for information of all bodies with legal tasks. This implies that the basic registries must be able to serve the contexts of many different legal tasks. In this case, semantics are therefore even inter-sectoral. On the scale of individual sectors, there are ambitions to establish *integral client files*, for instance for patients or students.

On an inter-sectoral scale, the objective is to offer *integral service portals* to citizens and businesses, as a way to implement a no-wrong-door policy. This does not only involve the contexts of a wide variety of government bodies that might provide services, but also the contexts of users. A recent development is that attempts are made to grasp these user contexts in terms of so-called *life events*, a user-centric approach which takes us even further away from the traditional government ‘stove pipes’.

Some individual information chains also operate across sector borders, and therefore often also across different laws. For instance, consider the information chains between the Tax and Customs Authority (acting in a tax context) and the Employee Insurance Agency (acting in the context of social security). Other information chains will operate within one sector, but across multiple layers of government. For instance, the Wabo Act⁸, which regulates environmental and spatial permits (e.g. for building, disposal and recycling of waste) issued by either central government, provinces and municipalities.

All these situations require a perspective that transcends the specific business processes of the parties mentioned. In other words, this calls for an increase in the semantic scale.

Diversity unites

Attentive readers will have observed that the word *standardisation* has not been used since the introduction.

This may seem strange. Isn't standardisation the perfect means of achieving interoperability?

Above, we described semantic interoperability as a problem of reconciling contexts. This may have given rise to the idea that all the various contexts must be respected. You might think this will never work: if every party is allowed to retain their peculiarities, interoperability will never be achieved.

It is correct that, at least with *technical* and *syntactical* interoperability, there are few reasons to allow deviating choices. Data traffic will be more efficient, transparent and error-free if a standardised process is used. In many cases this is possible, since technology by now has become a flexible commodity that may be tailored to specific processes and environments at will.

But an attempt to standardise semantics would be an attempt to standardise reality. This is not just a technical choice; it is a step with far-reaching consequences. The aforementioned standardisation of the term *salary* had real financial consequences for businesses, citizens and the treasury.

Legal tasks are about the real world. The real world and the legal tasks are not uniform. They are characterised by variety. Often, such variety cannot be avoided and may even turn out to be desirable. And again: the larger the scale, the greater the variety. Being able to handle such variety is a condition for successful organisation of semantic interoperability in the field of electronic government.

Consider, for instance, the term *employer*. The Dutch Tax and Customs Authority and the Employee Insurance Agency (UWV) share data on employees, provided by employers. The Tax and Customs Authority uses this data for payroll tax, while the Employee Insurance Agency uses this data to execute social security

⁸ Wabo is short for [Wet algemene bepalingen omgevingsrecht](#) (Act on general provisions for environmental permits).

schemes. These differences in use of the data lead to considerable differences in the definition (or interpretation) of the term *employer*, dependent of the context of the user. A party considered an employer by one body may not yet be treated as an employer by the other, and a party treated as a single employer by one body may be multiple employers to the other. This does not mean that there is no common factor. There is variety, but at the same time a connection.

Eff-icient or -ective?

Semantic interoperability is not a matter of choosing between either standardising or not standardising a term. It is rather a matter of identifying the differences as well as the similarities. Usually, these both exist at the same time. Similarities may offer possibilities for standardisation. Differences first of all need to be acknowledged. Next, there can be a discussion as to whether it is worth maintaining the differences. Eliminating them often makes processes simpler and more efficient. But there is also a risk that the processes will become less adequate and effective. Finding a balance between efficiency and effectiveness is a matter of tailoring and cannot be handled in one go.

With respect to efficiency, an important objective is reduction of the administrative burden for businesses resulting from their obligations to provide information to the government. In this context, standardisation of electronic information traffic is obviously an important subject. However, the semantic variation to be overcome is enormous. This not only concerns variations resulting from legal government tasks; variations are at least as much a result of the differences in businesses' operations.

Also, introduction of a standard does not automatically result in efficiency. Variation is hard to get rid of. In cases where variety is maintained, irrespective of the reasons why, intermediary processes might be implemented to connect the particular process and the standard, as 'semantic adaptors'. However standardised a situation may look from the outside, standardisation does not actually save costs; costs are only moved elsewhere⁹. The *polluter pays* principle makes sense. But, what is considered pollution depends on what is seen as the standard.

On the other hand, not every variation is effective in itself. Sometimes, 'being different' is a strategy in its own right. With businesses, this is often part of their competitive strategy, but such strategies are also found in government organisations, often related to task boundaries and allocation of tasks.

Therefore, there is no fixed answer to the question whether to vary or standardise. Virtually every communication is characterised by a combination of both. What is crucial is that both aspects are acknowledged and open to discussion. A taboo on either standardisation or variation is counterproductive.

A taboo on standardisation will rob organisations of the opportunity to look for a common factor and cooperate effectively and efficiently. If variation is rejected beforehand, standards may not catch on, and if standardisation is forced upon an organisation, it may even lead to additional costs and problems. It is only after variation is acknowledged that common aspects suitable for standardisation can be identified.

So, semantic interoperability involves standardisation and variety. In reality, however, the discussion almost exclusively focuses on standards.

⁹ See for instance: Rex Arendsen, *Geen bericht, goed bericht* Ph.D. thesis, University of Amsterdam, 2008.

4. Semantic models: describing meaning

Relations make sense

Meaning is relative to context. Context can comprise anything: time, a role, an event, a law, a task, a goal. In fact, what cannot function as a context? There are few things that cannot act as the context of a term. As a result, the distinction between a context and the term that derives its meaning from that context is blurred. Each term can both be inside the context of other terms, as well as provide the context of other terms. Because context determines meaning, we can also state that the meaning of terms lies within the relationship with other terms. This does not come as a surprise. Dictionaries always express the meaning of a term by using other terms.¹⁰ These other terms can also be looked up, etc. Eventually, we will return to terms already encountered, or terms that are so general that they are almost impossible to describe clearly. That way, a web of related terms unfolds. This type of terms network is probably about as close as we can get to the meaning of terms, both on paper and electronically. Let's go back to the term *employer*. This term owes its meaning to its relationship to terms such as *employee*, *work*, and *employment* – and vice versa. If you start with one term, you will inevitably start discussing the next one. There is therefore no point in addressing individual terms solely. Semantics is like spaghetti: if you pick up one strand, the rest will follow.

To implement or to complement; that's the question

Meaning lies in the relationships between terms and is therefore often represented in models as nodes (the terms), which are interconnected by lines (the relationships). In such models, two different approaches can be recognised. Let us call them the *implementary* and *complementary* approach.

The *implementary* approach takes a main term and analyses it step by step into its components. These components are then arranged hierarchically below the main term. In very general terms, this approach would analyse a car as follows: a car consists of a chassis, a body, the wheels, the engine, the drive mechanism, the interior, a year of construction, and an owner. Each wheel consists of a hub, a rim, and a tyre. The owner consists of a name, an address, and a driving licence. A driving licence consists of an issuing body and an issuing date. An issuing body consists of ...

The *complementary* approach places all terms next to each other and then focuses primarily on the relationships between the terms. In very general terms, this approach would analyse a car as follows: there are cars, there are chassis, there are bodies, there are ..., etc. A chassis, a body, a number of wheels, an engine, a drive mechanism, and an interior can have a relationship called car. A person can have a relationship with a car, called owner, etc.

Both approaches are important and are co-existent. However, the distinction between the two is crucial for semantic interoperability. The implementary approach has a closed nature because the meaning of all terms is enclosed within the meaning of the main term. The complementary approach is open, as all terms are first-class citizens. These meanings are interconnected, but could also be connected to other terms. However, the implementary approach is important as well, in particular in the realisation of information systems. At some point, a decision must be made as to which data group will be combined to form a database, message, table or system card.

These choices should not be made too early, though. A dominant implementary style will turn a model into a data model instead of a semantic information model, which happens all too often. A model that is supposed to be a semantic model is then in fact a data model, which means that various implementation choices (including

¹⁰ Or by providing examples of the context in which a term is used, thus expressing the existing variety in meaning.

storage and administrative decisions) have already been made. That does not meet the requirements of a semantic model, particularly not in large-scale environments.

One possible explanation may be that these models are created by system designers instead of by domain specialists. But there are also more fundamental causes: the development of the electronic government in the Netherlands is largely executed per facility or system. The facility concerned is identified, defined and allocated before the semantic interoperability with users and other facilities is secured. It is then developed as part of a project or programme, which must implement (!) it. The context and scope, ownership, and project or programme agreement defined in this way then obstruct a complementary approach. The facility solidified too early, or does not solidify at all. Parcelling has taken place too early.

Roles, identities and authorisation

One of the crucial issues where context plays an important role is authorisation: assigning rights to use functions or to participate in interactions. The assignment of these rights is based on a person's role within an organisation, event or situation. This role is in fact a relationship of a person to that organisation, event or situation. Or, the role is the person within the context of that organisation, event or situation. Persons almost always fulfil roles temporarily.

In the complementary approach, a role itself has its own *raison d'être*, and therefore its own instances and identities. A student is a role within the context of education, a director is a role within an organisation, a biological parent is a role within the context of an adoption, a surviving relative is a role within the context of death, and a witness is a role within the context of an (alleged) crime. These roles have their own identity in addition to the identity of the person performing the role. Authorisations are assigned to the role, not to the person performing the role. This means that any reasonably sized authorisation system cannot do without a proper model of these roles. This type of model is nothing less than a semantic model of the domain.

Variation-proof

The Standardisation Forum has prioritised semantic interoperability in its Interoperability Agenda¹¹. It has aimed at testing an approach that does justice to the enormous semantic scale of the ambitions of the Dutch electronic government. This approach must make context explicit, allow for a complementary style, and leave room for meanings that transcend the scope of individual facilities.

In 2009, the Forum tested such an approach – the Metapattern approach¹² of Pieter Wisse. The most important, and in fact only, construction in this approach is called contextual specification. The approach provides such a rigorous account of context-dependence, as well as a complementary modelling style, that it formed a good basis for testing. The same qualities can also be realised with other complementary modelling languages — including UML, OWL and ERD. However, these languages grant the modeller more flexibility compared to Metapattern, requiring additional modelling agreements to be made in order to realise the same quality. The rigorous nature of Metapattern suited the tests very well.

In the spring of 2009, the approach itself was assessed for its ability to improve the semantic interoperability within the field of the Dutch electronic government. Three cases¹³ and their semantic relationships to four

¹¹ See note 2.

¹² P.E. Wisse, *Metapattern: context and time in information models*. Addison-Wesley, 2001.

¹³ The three cases were: applying for a disabled parking space, the assessment of property, and the transfer file of students.

basic registries were analysed. The conclusions of the analysis were favourable¹⁴. By rigorously addressing context, the approach becomes a leverage to shift focus from ‘owners’ to users of basic registries. Furthermore, the approach commands a complementary, open style that simplifies the expansion, modification and combination of models.

Two more extensive tests were performed in the autumn of 2009. One of these tests¹⁵ involved the *related persons* (children, partners) in the basic registry for non-residents (RNI). First of all, due to the shifting context, the RNI cannot simply copy the concepts of *child* and *partner* from the basic registry for residents (GBA). Contrary to the population of the RNI, the population of the GBA is restricted to Dutch residents, which provides a particular context. The test also showed, partly based on a legal analysis, that an alternative interpretation of the term *partner* that could be used by bodies obtaining information from the RNI, would comprise such a wide variety of meanings, that inclusion in the current setup of the RNI is inadvisable. This variation would also largely apply to partner data of residents.

The second test involved the term *employer* in the payroll tax return chain¹⁶, in which – as stated above – several agencies¹⁷ were supposed to share data in carrying out their legal tasks, for their own efficiency as well as to achieve reduction of administrative burdens for employers. As part of this test, the above-mentioned approach was used to expose the variety of meanings assigned to the term *employer* by the chain partners involved from the perspective of the requirements of their own legal tasks. This variety was one of the reasons why the payroll tax return chain did not work properly.

Yet another method?

The analysis and tests provide a favourable picture of the power and value of a contextual approach.¹⁸ In its present form though, it presents us with another hurdle to take, because it appears to be positioned outside of current design practice. But this is actually not true, because of the fact that it is ‘just’ an approach, a way of thinking, or looking at things, if you like. This means it can be applied using existing modelling languages and that we can also benefit from the tools already available for these languages. And it will also allow designers¹⁹ within electronic government to more easily incorporate the approach in their own processes.

But attention must not be diverted from the main issue. The heart of the approach is in the complementary aspect, not in the language. This must not be impaired, even when languages and tools provide more freedom.

5. Towards semantic interoperability

A marathon across hurdles

In this article we have identified a number of issues that need to be addressed before semantic interoperability in the Dutch electronic government will become reality. Below, we briefly list them as hurdles, suggesting how they may be overcome.

¹⁴ http://www.open-standaarden.nl/fileadmin/os/documenten/Pub_Semantiek_rapport-semantiek-op-stelselschaal.pdf.

¹⁵ <http://www.open-standaarden.nl/fileadmin/os/documenten/Rapport%20Analyse%20en%20harmonisatie%20gerelateerdegegevens%20RNI%20%201.0.pdf>.

¹⁶ <http://www.open-standaarden.nl/fileadmin/os/documenten/FS23-12-02%203%20praktijkmodellering%20werkgever.pdf>.

¹⁷ Tax and Customs Authority, Employee Insurance Agency (UWV), and Statistics Netherlands (CBS)

¹⁸ This was also the conclusion of a review of the approach by Jeff Rothenberg, RAND-Europe, March 2010. His review and the English translation of the reports mentioned above are available at Standardisation Forum (forumstandaardisatie@logius.nl).

¹⁹ Business administration specialists, information specialists, legal specialists, IT specialists.

Hurdle	Approach
Semantics is considered a technical issue.	Place semantics at the beginning of the policy, legislation, process and system design process.
Domain specialists, process specialists, legal specialists and IT specialists are not yet cooperating across their professional boundaries.	Organise shared development of semantic models across these groups.
Context is rarely made explicit.	Make context visible in semantic models, for instance by means of contextual specification
In the context of interoperability, priority is given to standardisation.	Make variation in meaning visible. This will also show where standardisation is possible. Limit variation, if necessary, based on a considered effect analysis.
Semantics is handled term by term.	Develop semantic interoperability by (limited) domain. Use critical business interfaces as a starting point and keep models open for additional relationships. Make the results available in a systematic register for reuse and extension. Let models grow and adjust.
Semantic models are often closed due to an implementary approach.	Promote a complementary approach to modelling.
Semantics is secondary to a facility or an owner.	Semantics is language and ‘language that is owned’ ceases to develop. Organise development of language between parties, not within parties.
Contextual specification has not yet been incorporated in existing methods.	Point out the relationship with common modelling languages, tools and development processes.

First jumps

This article may give the impression that the Dutch electronic government still has a long way to go on the path to semantic interoperability. This however would not do justice to what has already been achieved. In several places, semantic standards have been established and applied, although often fragmented and not always geared to the scale of electronic government. There is also a model for the semantic core²⁰ of the basic registries, and recently a semantic catalogue²¹ has been compiled.

The approach to semantic interoperability described in this article is also closely related to service orientation. Service orientation promotes external orientation of service processes and information systems, just as this article promotes external orientation of semantics. Therefore the hurdles described in this article are also valid for service orientation. But service orientation is already an accepted principle for the Dutch electronic government. Some of the large executive bodies, such as the Immigration and Naturalisation Agency (IND), the Social Insurance Bank (SVB) and the Tax and Customs Authority (Belastingdienst), are actively reorganising their data systems and processes based on semantic models and service orientation. These models describe their business, using legislation as an important source. With this approach, they hope to regain the ability to quickly respond to changing legislation by amending processes and systems, as that is what a semantic approach does: decoupling business and technology.

²⁰ Vladimir Grafov and René van Hove, *Architectuur van het stelsel: architectuur en doorkijk*. ICTU/RENOIR. Version 0.6, 29 June 2006.

²¹ http://www.e-overheid.nl/sites/renoir/producten_en_diensten/stelselcatalogus/stelselcatalogus.html

This also enables them to organize their services around *life events*. These are used in an attempt to grasp the environment of citizens and businesses as a basis for organising and structuring their (shared) services in a user-centric way. Life events are, like profiles, part of the context of citizens and businesses and should therefore also be used in semantic models.

A second and bigger leap

The promising results with a semantic approach based on contextual specification in the tests we described above and the need for such an approach to further deploy electronic government in the Netherlands, call for an initiative to make this approach scalable and practicable. As interoperability is not only of interest to government, but equally important to businesses, the Standardisation Forum aims at *open innovation*. The concept of open innovation was introduced by Chesbrough²², assuming that firms can and should use external ideas as well as internal ideas, and internal and external paths to market, as they look to advance their technology. Though the concept initially focused on firms, the same goes for government. Just as the boundaries between a firm and its environment, the boundaries between government and society have become more permeable. Innovations on either sides can easily (and must) transfer inward and outward. Knowledge is no longer proprietary, nor to businesses, nor to government.

Futhermore, the need to optimize semantic interoperability not only arises with respect to the achievement of primary goals of government (such as in e-government), but also of market sectors. An example can be found in the financial sector, specifically insurance companies and pension funds. They are faced with important changes such as mergers and take overs, new products and activities (e.g. distinction between asset management and execution in the pension sector), new (electronic) services and increasing regulations, especially on compliance. This calls for substantive change in their information-household, not only within companies, but also at chain level. Without a good semantic basis, the required changes will be achieved later or less good, and companies will see their position weakened, which may cause serious negative effects for their customers as well.²³

In this respect, the Standardisation Forum launched the idea of sharing its knowledge on contextual specification to achieve a broad adoption and implementation of the approach, not only within government, but also in the commercial field. The ambition is, in short, to stimulate the market (in the broadest sense of the word) to further develop, apply and exploit the approach of contextual specification. This idea takes the shape of a precompetitive co-operation in a consortium with public and private parties, called *Essence*²⁴, working together at open results that can be used by market parties.

The consortium welcomes participants from three main target groups:

- ICT user organizations, primarily from the government, finance and utility sectors, because they are the primary holders of the semantic interoperability issue;
- collective organisations, such as sector organizations, chain and network organizations, and standards organizations, because they often represent the semantic interoperability issue for their sector, chain, or network;
- ICT service and tool suppliers, because the market success of the approach builds on them embedding it in their commercial portfolios.

²² Chesbrough, H.W. (2003). *Open Innovation: The new imperative for creating and profiting from technology*. Boston: Harvard Business School Press, p. xxiv.

²³ Similar cases can be found in the cure/care and education sector; coping with e.g. growing demand as a result of aging, implementation of new (digital) learning methods and lack of (qualified) personnel, parties are urged to enhance their capacity to co-operate.

²⁴ Essence is an acronym for Expressing, Structuring and Sharing ENterprise ConcEpts. See also www.essence-project.nl.

Essence will work on the following results:

- a shared *language* (metamodel and notation) for applying contextual specification in semantic models;
- an approach for *establishing* such models together with users and domain experts;
- an approach for *combining and reconciling* such models from different sources;
- an approach for *implementing* such models, via their transformation to existing modelling languages;
- an approach for evolutionary *maintenance* of such models;
- *tooling* for support of these four approaches;
- real-life case *applications* of the approach for each of the consortium participants;
- valorisation of the approach by means of education, commercial exploitation models, dissemination and methodological standardization.

Participation in the consortium requires a cash investment of the participating organization. In return for this contribution, the consortium will carry out a project for the partner, based on a case from its own practice. The project can either be at the ‘front end’ (analysis of the existing practice of the partner, collecting wishes and demands for a change in this practice), or at the ‘back end’ (a pilot application). The results of the project are for the benefit of the consortium partner. A generalization and ‘anonimization’ of the results will be part of the consortium approach. Thus the consolidation, further development and open accessibility of knowledge and tooling will be guaranteed.

The approach in the case-projects will be strongly iterative (‘agile’ in terms of software and systems development), working in cycles of three months, concluding every cycle with an evaluation by the users. Every cycle will lead to a (new) version of (parts of) the results, mentioned above. Coherence will be guarded by continuous maintenance of the ‘architecture’ of the approach.

It should be very clear that all Essence’s results will be placed, from the start, in the context of the overwhelming amount of existing semantic modelling approaches. Most, not all, of these contain considerable value that is reused by Essence. Yet, Essence adds a crucial element, or even missing link, to standing practice. It is called contextual modelling and is indispensable for large-scale semantic interoperability.

Why so much hassle?

This is not an advertisement for semantic interoperability. It is not very attracting: semantics is abstract, context-specific and cannot be simply allocated to some party. Not even legislation provides a definite basis. Nevertheless, the value of information systems is in the meaning of the data stored, moved, transformed and presented. Information systems are not just gadgets that make life easier and more pleasant; they manipulate our daily reality. There's nothing that can be done about that.

Accepting this, the risks of ignoring the context-dependence of meaning are not to be ignored:

- Central electronic government facilities may not be suitable for the intended users, particularly if the group of users is large and varied.
- Misunderstandings in information chains.

If definition and meaning are placed specifically with one party, there is a risk of meaning being limited to the context of that party and therefore not suitable to other parties. What this means for the achievability of e-government principles such as ‘single delivery, multi use’ and user-centric services will be clear.

However, the solution should not be sought in standardisation alone. The risks of strictly and exclusively aiming for standardisation of meaning are:

- A common factor between parties may not be found.

- No cost reduction will be realised, but costs will be moved and increased, while effectiveness of processes will be endangered.
- An intervention in reality can have undesirable effects.

When data models are accepted as semantic models too soon, meaning may be defined in system structures rather than in reality structures. Such a reversal places the system above the context and is therefore contrary to service ambition of the Dutch electronic government and service orientation. An additional risk is poor adaptability and maintainability of the models.

If interoperability in the Dutch electronic government would involve no more than a few bilateral connections, perhaps all this intricacy would be unnecessary. But, as we have pointed out in the above, ambitions go far beyond this. Semantic interoperability is therefore a problem at e-government scale. It is this scale that inescapably makes semantic interoperability rear its tricky head and calls for scalable solutions. We hope and expect to realize these in the approach of public-private partnership and open innovation: the *Essence* consortium. The first signs are promising and we hope to be able to present even more tangible and appealing results in the near future.