Naming and Design Rules for E-Government - The Danish Approach

Mikkel Hippe **Brun** <mhb@itst.dk>
Brian **Nielsen** <bni@itst.dk>

Abstract

Denmark has adopted XML as a key to the information architecture in support of E-Government. The philosophy of the Danish E-Government strategy is based on government wide co-operation and reuse - and support also cooperation with both the private sector and with a wider international community. The Danish XML-project focuses on coordination and supports the development and standardization of XML interfaces and vocabularies. A national XML Committee is responsible for ensuring coherence and momentum in the standardization of XML-based interfaces and vocabularies. A number of handbooks provide rules and advice for developing XML vocabularies and application integration solutions "the right way". A national XML registry provides a number of tools, including a XML Schema repository, a community module which supports collaboration, shared data modeling and Schema development, re-use of data definitions, interfaces, and a UDDI registry for implemented services.

A characteristic of the Danish approach is that messages are composed from multiple namespaces. The Danish XML Committee and the related Technical Committees will only develop XML Schemas from core information objects. Private companies and public authorities are encouraged to develop and submit their own vocabularies to the XML Committee for public hearing and approval. Provided that one or more vocabularies exist in a given domain - public authorities whishing to develop XML interfaces in that domain - are required to re-use existing XML Schema fragments from these vocabularies. Reinventing XML Schemas is strictly prohibited.

Reuse from different namespaces requires that the same XML Schema Naming and Design Rules are used. XML Schema Naming and Design Rules (NDR) are key to successful application integration in E-Government. The Danish approach is to rely heavily on reuse of atomic schema fragments from a large number of namespaces related to different authorities and international standardization initiatives, such as the OASIS Universal Business Language (UBL). The Danish naming and design rules differ in a number of ways from the naming and design rules of UBL. The paper will present the naming and design rules and the rationale leading to the differences with UBL naming and design rules.

An advantage of W3C XML Schema is its introduction of strong data typing capabilities. These advantages can not be exploited in schema vocabularies such as UBL, simply because it is impossible to agree on strong data types that are relevant to all parties. The consequence is that much of the validation must be performed by back end systems. In E-Government there is a substantial benefit to be gained at a national level by expressing all elements with strong data types. The resulting messages will convey much of the integrity rules of the exchange, and the XML Schemas can be used as part of a contract between two parties. Otherwise the integrity rules have to be expressed only in the supporting documentation. This approach is off course a lot more error prone that the strong data typing approach, as the definition of the data is separated from the structure.

The Danish XML Committee has decided to embrace and re-use ebXML and UBL Core Components. This decision is a challenge since these vocabularies are composed from weakly defined elements and types. The approach taken is to define a number of national namespaces which are strongly typed versions of the ebXML and UBL namespaces. The criteria is that for each element in the national namespace - it must be possible to transform the element into the corresponding

ebXML or UBL namespace. This approach and its applicability for other national XML initiatives is discussed.

Defining a unanimous and publicly available standard for communicating, with and within the government, opens the market for vendors to create applications that comply with the standard. This should create a better competition on the marked and better interoperability between different products.

Table of Contents

1. Background

E-government is largely a matter of getting public sector IT systems geared to interoperability. The authorities must have the capability to use each other's data so that citizens, companies and case officers do not have to provide and check the same information over and over again. This requires, for example, common data definitions and coherence in the handling of security and users. And it means dispensing with 'technological islands' if we are to create a platform for new work practices.

1.1. The Danish XML-project

Past application integration efforts has evolved around the EDIFACT and X12 standards and a number of more and less proprietary methods and tools. The cost of implementing application integration has in many cases been too high and thereby out of reach especially to small and medium sized organizations. It is characteristic that application integration efforts in the public sector has been developed in an ad hoc basis and has rarely been based on established standards. Ad hoc application integration that is not based on standardized interchange formats makes it difficult and expensive to replace applications and components. Every application or component has to be tailor-made to support specific integration requirements.

With this in mind, Denmark has adopted XML as a key to the information architecture in support of E-Government. The philosophy of the Danish E-Government strategy is based on government wide co-operation and reuse - and support also cooperation with both the private sector and with a wider international community. The Danish XML-project focuses on coordination and supports the development and standardization of XML interfaces and vocabularies. A national XML Committee is responsible for ensuring coherence and momentum in the standardization of XML-based interfaces and vocabularies. A number of handbooks provide rules and advice for developing XML vocabularies and application integration solutions "the right way". A national XML registry provides a number of tools, including a XML Schema repository, a community module which supports collaboration, shared data modeling and Schema development, re-use of data definitions, interfaces, and a UDDI registry for implemented services.

The underlying vision is that the public sector must act as an enterprise with coordinated service development and re-use in a Service Oriented Architecture. The benefit is that ultimately citizens and companies will not have to supply the public sector with the same information twice.

2. A Service Oriented Architecture

2.1. A goal in the Danish E-Government strategy

The Danish IT Architecture Committee has written a White Paper on Enterprise Architecture which as a general principle, recommends a service oriented architecture model in which IT solutions are designed in a modular fashion, are divided into services with well-defined mutual interfaces and, as far as possible, interfaces to existing public sector IT systems. [ITA 2003]

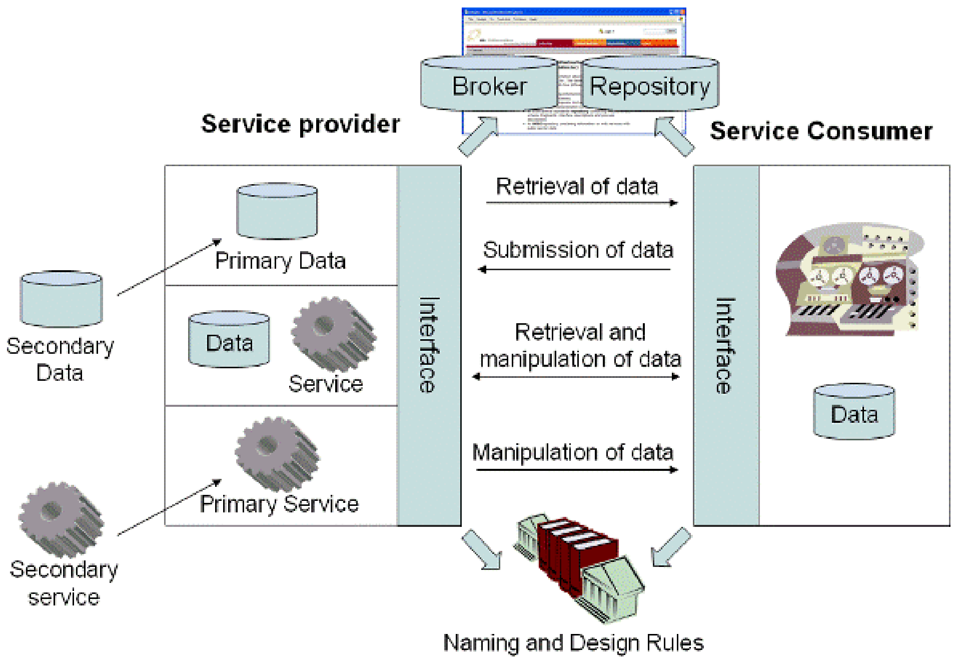
Part of the common principles is a standardization that can ensure data exchange in the public sector without technical barriers. Standardisation must focus on supporting interoperability, security and openness, and should include both the technical standards that make interconnection possible and the information structure that defines the common conception of data. One example is the choice of the XML exchange format in Denmark. Technical standardization should take its starting point in international, open standards or, where these do not exist, de facto standards.

> "Service-oriented architectures (SOAs) are a particular kind of software architecture that is designed to create a dynamically organized environment of networked services that are composable and interoperable. The fundamental building block of an SOA is a service and services are composed in specific ways to create applications. SOAs separate the services from their implementation using the notion of an interface. This interface creates a contract about how the interaction between the parties will proceed. SOAs provide a number of benefits for creating federations of services among disparate and loosely connected organizations while allowing each organization to maintain its autonomy in terms of how it builds and designs services as well as their ownership."[WP 2003]

2.2. Challenges to the implementation of a Service Oriented Architecture

A Service Oriented Architecture in the public sector is a challenging goal. SOA introduces a whole new paradigm for exposing data and functionality from one IT-system to another. Realizing a Service Oriented Architecture is not only a matter of technology and architecture. The SOA paradigm requires authorities and companies to take an entirely new approach to data integration.

The wonders of the Service Oriented Architecture as presented by Windley are a few years away. E.g. the Service Oriented Architecture depends on dynamic binding between services in order to be self healing. Many of the benefits of the SOA can only be realized when a critical mass of services have been exposed. For a while brokers like a UDDI registry will primarily be used for design time binding as opposed to run time binding. Furthermore some of the benefits of the SOA may be hard to harvest in the public sector where most services are only provided by one service provider. The SOA also challenges our traditional business models for ad hoc data integration, where the cost of integrating two IT-systems in most cases was paid by part(ies) that benefited from the integration.



An illustration of the SOA Service triangle.

Figure 1.

It is very important that a service interface acts as a facade. The data model exposed in the interface should not be a replica of the original data model of the service. The original data model may very well have been developed with a number of application specific constraints taken into account and these constraints must not be visible to the clients accessing the facades interface.

An interface should be composes of well modeled business information entities. The definition of cross domain business information entities should furthermore be standardized and reused. This will allow for re-use and sharing of handlers for the business information entities. E.g. an address handler may recognize the standardized address structure as being part or the message and transform the address to and from an internal data representation.

Re-use of business information entities has two advantages:

1. Code re-use: Handlers implemented to recognize and address specific data structures may be reused within an institution and possibly among institutions.

2. Less transformation hassles: If standardized business information entities are reused, less transformation is needed when exchanging data.

The consequence of following the SOA approach to implementing E-Government is that Schemas from many different namespaces with XML vocabularies and messages will be developed.

3. The importance of Naming and Design Rules

Many of the advantages of XML can only be realized if common Naming and Design Rules are followed. This is not only a regional or national issue since services on the internet are available to anyone. It is aparent that Naming and Design Rules with an international scope are important.

We now propose that coordinated work on NDR is put on the agenda in an international standardization organization.

3.1. International work on Naming and Design Rules

The Danish work on Naming and Design Rules is inspired by early work in the U.K. and the U.S. In the U.K. Guidelines for XML Schema development[SP 2002] are being developed at the Office of the [SP 2002]e-envoy under the Cabinet Office (UK-NDR). In the U.S. the Federal CIO Council Architecture and Infrastructure Committee has chartered an XML Working Group. This group has produced a "Draft Federal XML Developer's Guide"[US 2002] (US-NDR).

There are many similarities but also differences in the work from the two parties. The basic element and attribute naming rules require the uses of camel case. The US-NDR recomends the use of the ebXML modified ISO 11179 data element naming convention as solid basis for XML component creation. Furthermore the US-NDR suggests that Business Processes models and document models may use Unified Modeling Language. And that Schema development should take place as a team effort, involving functional data experts, business experts, program managers, and IT specialists. The reason for this being that "the single most critical factor in creating logical, reusable schemas for information exchange in XML is the separation of the information modeling process from the schema creation process."

The OASIS Universal Business Language TC (UBL TC) has a SC (UBL NDR SC) focused entirely on Naming and Design Rules to by used by UBL customizers and internal TC members. [UBL 2003] Some of the rules are incompatible with an E-Government scope with messages and vocabularies dependent on many namespaces. But the majority of the rules are applicable to any Schema developer. Because UBL is the first standard designed to work within an ebXML framework - the design of UBL will be trendsetting for messages in other domains.

4. The contradiction: Strong data types and international re-use

An advantage of W3C XML Schema is its introduction of strong data typing capabilities. These advantages can not be exploited in schema vocabularies such as UBL, simply because it is impossible to agree on strong data types that are relevant to all parties in cross border exchanges. The consequence is that much of the validation must be performed by back end systems. In E-Government there is a substantial benefit to be gained at a national level by expressing all elements with strong data types. The resulting messages will convey much of the integrity rules of the exchange, and the XML Schemas can be used as part of a contract between two parties. Otherwise the integrity rules have to be expressed only in the supporting documentation. This approach is off course a lot more error prone that the strong data typing approach, as the definition of the data is separated from the structure.

We propose that vocabularies like ebXML Core Components are developed in accordance with Naming and Design Rules based on the same principles as UBL. But the Naming and Design Rules should take into account that it can be desirable to enforce stronger data types at a regional or national level. The condition for regional or national restriction of the data types should be that the new namespace is a true subset of the original namespace.

The ultimate test of this requirement is that it must be possible to develop a XSLT stylesheet that can transform an XML-instance from a national namespace to the international namespace for that particular vocabulary.

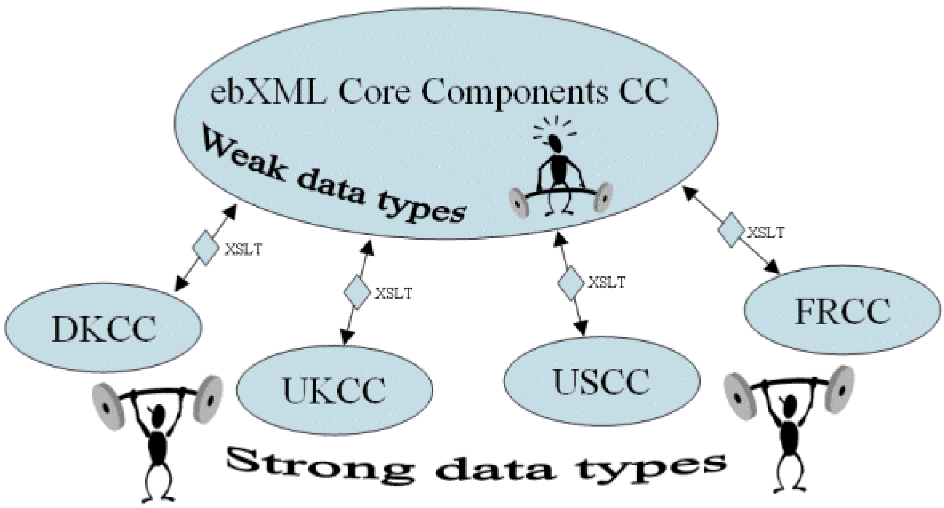


Illustration of the weak namespace of ebXML Core Components and possible strongly typed regional or national counterparts

Figure 2.

The idea is that messages composed in a regional or national context may be composed of strongly typed components. In many uses of XML for data interchange this scenario is likely to be most dominant (e.g. the health industry where patients are rarely moved across borders). In the event that a message is to be exchanged with a partner, that does not support the strongly typed namespace, at simple transformation can convert the message into the international namespace that may be accepted by the recipient.

5. The vision: International Naming and Design Rules for E-Government

International collaboration on the development of Naming and Design Rules is very important. We are still at an early stage of XML Schema adoption. There is much to be gained by not letting this opportunity for corporation slip our hands.

We propose that a coordinated effort is initiated in the OASIS eGovernment TC and that the UBL NDR work is used as a foundation.

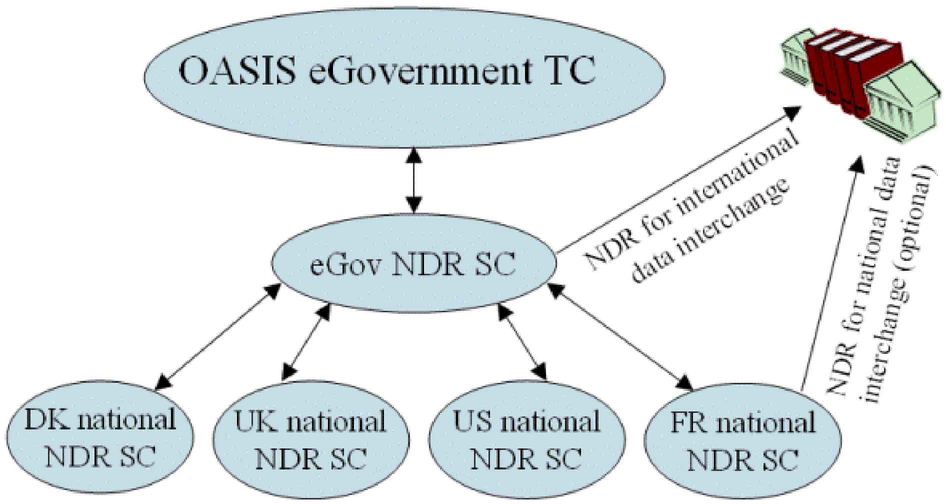


Illustration of the organization of an eGov SC's relationship with national NDR bodies.

Figure 3.

The NDR of the eGovernment TC should take into account that stronger data types may be desirable at a regional and national level and that this aspect must be addressed and allowed.

6. Summary

The requirements to Naming and Design Rules in E-Government differ from the requirements within an isolated domain like UBL. In E-Government re-use between multiple namespaces and the use of strong data types will exploit of some of the new features that XML Schema provides. The NDR of Universal Business Language is a solid foundation for future international work on an eGovernment NDR. It is proposed that such work is initiated in the OASIS eGovernemnt TC.

Acknowledgements

We would like to thank our colleagues at the National IT and Telecom Agency: Michael Bang Kjeldgaard, Palle Aagaard, Erik Hannibal Terp, Signe Wagner, Rasmus Knippel, John Gøtze and René Løhde for contributions to the development of ideas discussed in this paper.

Bibliography

[MB 2002] The Danish XML Committee, 2002: Brun, Mikkel Hippe and René Løhde 2002. *XML Schema Handbook*, The Ministry of Science, Technology and Innovation, National IT and Telecom Agency, Copenhagen, Denmark. http://purl.oclc.org/NET/oio/cookbooks/XMLschema

[SP 2002] Office of the e-Envoy, 2002: Spencer, Paul and Ann Wrightson 2002. *Schema guidelines Best Practice Advice*, Cabinet Office, London, Great Britan. http://www.govtalk.gov.uk/documents/Schema%20Guidelines%202.doc

[WP 2003] The Danish ITA Committee, 2003: Windley Phil 2003. *Service Oriented Architectures*, The Ministry of Science, Technology and Innovation, National IT and Telecom Agency, Copenhagen, Denmark. http://purl.oclc.org/NET/oio/reports/SOAexplained

[US 2002] Architecture and Infrastructure Committee, XML Working Group 2003. *Draft Federal XML Developer's Guide*, U.S. Federal CIO Council, USA. http://xml.gov/documents/in_progress/developersguide.pdf

[UBL 2003] OASIS UBL NDR SC, 2003.*Universal Business Language (UBL) Naming and Design Rules* , The Organization for the Advancement of Structured Information Standards. http://purl.oclc.org/NET/oio/reports/EnterpriseArchitecture

[ITA 2003] The Danish IT Architecture Committee, 2003.*White Paper on Enterprise Architecture* , The Ministry of Science, Technology and Innovation, National IT and Telecom Agency, Copenhagen, Denmark. http://purl.oclc.org/NET/oio/reports/EnterpriseArchitecture

Biography

Mikkel Hippe **Brun**
National IT and Telecom Agency
Copenhagen
Denmark
mhb@itst.dk

Mr. Mikkel Hippe Brun has been an active SGML/XML consultant to Danish business and government since 1995. For the past two years, Mr. Brun has been the chief technical advisor to the Danish XML Committee at the Ministry of Science, Technology and Innovation. Mr. Brun is author to a handbook on XML Schema Naming and Design Rules published by the Danish XML Committee[MB 2002]. Mr. Brun has been an active member of the Danish SGML/XML community since 1993.

Brian **Nielsen**
National IT and Telecom Agency
Denmark
bni@itst.dk

Mr. Brian Nielsen works as a consultant on the development of a common Danish data model for the public sector, documented in W3C XML Schema. Mr. Nielsen is in charge of the development and quality control of XML Schemas in various domains. Mr. Nielsen has a background in Mathematics/Physics but later focused on computer science and has worked with meteorological data for many years.