

THE DANISH NATIONAL XML PROJECT

Michael Bank Kjeldgaard

**Chairman of the National XML Committee
Denmark**

Michael Bang Kjeldgaard has been chief consultant for the National IT and Telecom Agency, Ministry of Science, Technology and Innovation. He is responsible for co-ordinating general IT architecture and is Chairman of the Danish XML community. Michael previously worked in local government in Denmark and on getting local administration on the web.

Thank you for this opportunity to talk about the Danish XML project. As this headline indicates, I'll focus down and narrow on one of the issues we have been talking about yesterday and today. I think it is a core issue about using XML, and of course, it is only part of a greater picture. I could say to the last two speakers, we have just heard that the Danish XML project is within a context which is very similar to what is going on in Australia or in Sweden and in fact, in many ways, we are quite similar in our approaches and yesterday I was not here unfortunately, but, I suppose most of you were here and heard my colleague Niels Pagh-Rasmussen tell about the Danish IT architecture project and the framework for that. And I think that indicates, more or less, that we are on the same kind of track as a lot of other countries here. I also expect that what we have experienced regarding the XML project is more or less on the cutting edge of a number of issues. Of course, when we started it, we looked at experience elsewhere and we constantly do that but we also know that we are, in some respects at the tail in places, and constantly a lot of new issues are popping up in front of us that we have to deal with and that's what I am going to tell you a bit about. Sort of a snap shot of our experience with the XML project.

First, I would like to show this slide which I know you had presented yesterday and which is basically what most people will show when they talk about e-government, something like this. The point is that since we started talking about e-government like the new agenda for e-government that was mid 90's when we started getting the reception of the internet as something interesting for the government on a wider basis, and we found out that HTML was some interesting technology to put around our information. At the same time, more or less, actually I started working in this area around mid 90's, 96-97. At the same time we had the first debate about XML, the symantic web and stuff like that. Really sophisticated technology talk for specialists only. We had new debate about whether metadata would probably become very important and we can't really touch that issue because, it's about standardisation and standardisation takes years and our bosses don't want years they want months for development and you all know about the internet; that any project beyond three months is a dead project. You have only have three months. We have some problems here understanding that. But the first project on the internet were those with simple home pages and then we got more sophisticated portals, gathering content from different distributors. We had also some self-service simple forms, and search engines and stuff like that, that worked well with HTML and well without XML. Now we are on the level where we have transactions and some limited self-service solutions, where XML becomes more and more important, still fairly simple, but it

becomes more and more important, more and more used in practice. But if we want to move up to the highest level, it is our thinking at the moment where we are now in technology involvement, that XML is the strategic choice regarding the general wrapping of our content data information whatever you call it, documents. So it's a core technology to achieve the last level which may be years ahead, we will need a long term effort to develop the XML vocabularies for all our data and documents etc.

So what's the scope of the Danish XML project? Well, I think its necessary to say that in a way as we can't see the end of it, it's everything. It's not really going to be everything, but I think we need to come to terms with the fact that we are in a way reaching for the stars and we don't know how far we are going to reach but at the same time we need to have our feet on the ground, because otherwise it will stop and nobody will support this kind of effort. If we are just reaching the stars but at the end of the day, don't tell your politicians while it is just a short-term development and then we are finished. It's huge, but you have to go step-by-step and it covers services and administrative processes across all tiers of government all sectors and it concerns, both highly structured data in registries and less structured data in documents. I think we have to work a lot with documents which are traditionally not structured and develop better structure in those different types of documents in order to, for instance, support personalisation, multi-channel access to information etc. Also it covers the whole of government and you have to include the whole of government in the process, in the project of XML standardisation and you very quickly realise that you need some international angles to your project. For instance Denmark is part of the European Union and there are European Union initiatives with respect to public government and public services, talking about Pan European services and you have also the international aspect for instance, as mentioned many times today and yesterday, SARS, terrorism etc. So there is a whole lot of areas where the international angles are important. Another international angle is that our suppliers are global actors and we want to inspire our suppliers to develop good and cheap products for us so alone for that reason we have to act on a more international scale. We don't want to buy specialised products for Denmark, we are a very small country we'd rather buy standardised products which support our needs.

This thing with reaching the stars and then keeping a connection to the ground, in Denmark we do that by having the process driven mostly by actual implementation of leading projects. To some extent I'll come back to that. We have some general initiatives also on a very high level of standardisation that have a very broad perspective and not just one specific project driving it. There is another image also I saw yesterday and its just to sort of clarify the scope of what we are talking about, this is the image of some of the systems in a county in Denmark. Its not covering the health sector which is the largest task in the county but it is covering administration and environment. In order to make whatever needs business might require, in order to make these systems communicate you need some kind of integration and of course this is just one county and one tier of Government etc. its very clear it is complicated.

The traditional approach to do that is to pin-point this and this system need to talk together and then we will make these bilateral relations. In the beginning it's quite easy to do that and you can do it just on a project scale and just ask any developer to develop that integration and that's it go. It's very easy but after some time it becomes a bit complicated – you make all these lines between all these systems and this was

just you know a snapshot of a little part of what we are talking about. It becomes very simple to see that you get a lot of lines and that is what we can call integration.

Spaghetti integration cement is very expensive to develop, its very expensive to maintain so the new approach is to say let's have what we could call a common playing field where we have a common standard for integration and that was decided in Denmark a couple of years ago in relation to a cross-governmental report covering both the state and the municipalities and the counties report on e-Government and a report also specifically on access to the agency's own data and other agencies data. It was decided to or agreed to use XML as a common format for exchanging data when there was a business case for it. Of course if you have a good integration existing without XML its okay, go with it but otherwise use XML but that's not enough if you really want a common playing field XML does not do the job. What everybody would do if we just said XML would be like asking a transformation engine (*re slide*) Please take this content and make it into XML and throw it out here and then people having the other systems wanting to use the data they would go into the field and say okay there is some data and its in XML but really apart from the fact that it was XML it wouldn't necessarily be much easier for them to grab and then use it in their own systems. They would have to retransform it here and put it into their own system –a different platform -but this transformation process would be really expensive for this type, over here cheap.

Push the golden button at the other side, expensive – we have to understand what's out there. So what you have to do is to develop a common language for the common playing field – common vocabularies and this is going on in a lot of areas, different communities are working on XML vocabularies, standardisation efforts like Oasis, there is a lot of technical committees doing this especially the business communities, the international business community has been developing vocabularies for this but e-Government is really only at the starting point of this issue and I think that's my main message here today that e-Government needs a common vocabulary for this playing field.

Now I'll go back a few slides. It could just be that the Manager in this county said okay I just want a common playing field for my own business, my county; but that's not the general agenda in Denmark. We have a national e-Government project so in Denmark this playing field is at least the nation, the scope I was talking about before, but we are also part of Europe so it will increasingly also be part of the agenda I think, in Europe I hope to co-ordinate at least to some extent the development of vocabularies here and also on an international level.

In Denmark we have developed a sort of paradigm related to XML and we called it the OIOXML Paradigm and OIO stands for Open Information and I'll come back to the OIO. We have a website where you can find all the information regarding what we are talking about here but the goal for the OIOXML Paradigm is a common shared data model. I talked about reaching the stars, we could also talk about the impossible dream; we are trying it nevertheless.

The purpose is inter-operability and re-use – this will be core of not only data exchange but also of common services, we want to develop common services and common components to share in a common architectural framework. This will be very important. The concept is information object as building blocks. I'll come back

to that and the coherence that we need in order to be able to play the game together efficiently is a focus on both semantics and syntax. I'll come back to that too and the method is that's the point with the coherence common guidelines and procedures. Our approach is talk down and button-up approach and the point here is as I said before that you need to have drivers which are projects but at the same time you have to have some governance to it, some co-ordination effort to it and I'll come back to the organisational set up for this.

Regarding the guide lines we have developed so far, a number of handbooks within the greater architecture project we are developing more handbooks but so far in the XML project which has worked as a spearhead project for the IT architecture programme we have developed five handbooks, one for Project Leaders called 'The Implementation Handbook', one for domain experts called 'The Handbook for Standardisation', this is where you don't need to know a lot about XML but you do need to know about business processes, business concepts and modelling of these matters. More specifically on modelling we have developed a handbook for domain experts and system architects and XML schema developers, this is where they meet sort of, then we have an XML schema handbook that's for the experts developing the XML schema and our XML schema handbook is very inspired by the UBL naming and design methods and we have an integration handbook which is aimed at system architects and we will have much more in-depth look at that in the future. At the moment it is on a rather general level because this is really a complicated technical matter you cannot write everything into one handbook. Then we have developed some OIOXML classes and this is a very important part of our experience because for the first couple of years in our project it was either OIOXML or not but during the past few months when looking back and evaluating we have recognised that we need to have different classes, we can see from the propositions for schemas that we get in and we have to evaluate we can see that it is very different what different projects provide so we have developed these classes and you could say that the higher the better, its not that simple but adopted components and vocabularies, that actually means that if there is something out there on the international level and its good quality and its good reason to use it we'll take it. It does not specifically have to confirm with our schema rule or the things that are within our own set of rules because we may have a good business case for adopting something international even though it is deviating a bit.

Then we have the core components – these are developed both from grabbing something international for instance a lot of the core components in Denmark are taken from UBL which is based on EBXML but if we need some core components which we do not find we develop them ourselves or maybe we sophisticate that's something we can take from the international standards.

Core components are, well I have to say in Denmark we don't have a set up like in Australia and Sweden we don't have a set up where it is mandatory to use those standards; they are voluntary public standards so if you choose to use them then the core components are mandatory. Then we have the domain components; they are not mandatory for everybody; they are only mandatory for say within a domain, for instance a health domain, if the health domain decides that these are the components we use then they should use it.

Then we have lower levels of design schema, design compliance schemas and syntax compliance schemas – they are based on a lot of efforts like the adopted and core and domain schemas that we get into the project. They have been developed through a process which typically should provide good quality but some schemas that we get in have not been through this process so for instance we have a project developing schemas for 1200 forms within a rather short time span, a few months and they should try to reach the quality level of design compliant but maybe they won't so we have the level of syntax compliant and we tell that project it's good that it's syntax compliant. It's better that its just XML and not OIOXML but its not necessarily good enough quality yet to make it design compliant or good enough quality to make it core components.

So this is a very brief introduction to what we are doing but the point is you need some different classes if you want to have the approach of both central and de-central work effort, common work effort. The building blocks that we are working with consist of something, you could use the legal image here I think, the core components here are core types and elements and value lists and we can use them to build a bit more complex schemas which we call composite schemas and you can use them again to build on a higher level interfaces between actual systems so the actual building blocks are of no use its only when they are combined they are of use and this is a bit opposite to for instance the EDI approach that builds standards for messages. Here we focus on the standards as more of the finer granularity, small building blocks and the way we try to govern all the different kinds of elements that are developed and all the different schemas that are developed is by using something called Name Space Technology and to have and to decide who is having the authority over types and elements.

If we look at what has been accomplished so far and this is like a snapshot of the infrastructure base we have, I will have to go back, before introducing infrastructure base we have an infrastructure base which is an important part of the XML project which contains four tools, an information site and you can access this its www.isb.oio.dk. I'll give the address in the end.

The site contains a lot of information including the handbooks I was talking about, including the organisational set up including a tutorial to the whole infrastructure base. Then it includes some community tools to support the different projects developing XML schemas and propositions for standards like core components and domain components and then we have the repository which stores the actual schema and we have a UDDI which is like the yellow pages of actual web services, and now for the snapshot.

At the moment we have a couple of days ago I looked into it we had 361 documents most of them are XML schema and examples of the output is core components for public registries that surround fifty core component elements mostly taken from UBL and we have admitted data core for exchange between electronic records management systems and we have education and enrolment documents and schemas related to buildings and properties which is a large number of schemas here request for reimbursement on illness schemas for the structure of letters, mails, notes agendas and minutes they are not finished yet but they are under development as well as XML schemas for the generic structure of Danish legislation documents these are only also on a very basic level and will be developed further in the future. We have a common

subject scheme called the OIO subject scheme which is important because that's the way of sort of system enticing the services in the public sector and we have forty five corporation and institution names. We will be developing more in coding schemas like the last two ones in the future. Also here is a brief overview of the standardisation process more or less idealised. Someone, again typically project driven, will identify a need for a standard. Maybe they do not need as such a standard but if they think within the paradigm they have the same mind set they will typically think our project will be better and it will be secure for the future if we think in terms of standardisation so lets contact the XML Secretary and have a dialogue with them, if there is something we can reuse or if they have a recommendation on whom we should approach on our project regarding the exchange of data and use of XML. If there is a good reason for it they will establish a working group and then the next step should be that they should search international and national schema repositories to see if there a similar effort going on in other areas. Then if not they should adapt something they should start developing schemas and documentation for their need and for their proposition for national standards. They should register that in the repository in the infrastructure base and submit proposed standards to a Hearing Committee and we have the national XML Committee and we have some working groups and other committees. Then there will be a quality check in related committees within the relevant domains there will be a 30 day hearing period collection of hearing responses and approval or rejection, in case of rejection they may give up or they may go back with few steps in the process and redo things. This is a simplified image of the e-Government set up in Denmark again with the approach of the XML project, what is most relevant for the XML project and everything here is across state municipalities and counties, except the ministry of course. The Joint Board of Project E-Government consists of tough administrative CEOs from five ministries and from the associations of the municipalities and the counties, together with the ministry they are being served by what we call the digital task force or the e-Government task force which is also set up as a joint secretariat from those parties. Then we have the Coordinating Information Committee that's one or two steps below the CEO level. Its not exactly the CIO, we do not have CIOs in Denmark as such, but they have the task of looking more closely into the technology and information issues as the Joint Board of Project E-Governments looks more on the business level the Information Committee looks more on the technical issues and we have an XML Committee for about two years, but in June it was decided as you heard yesterday I guess, to form also an IT Architecture Committee so under the Information Committee we now have the Coordinating Information Committee, we now have an IT Architecture Committee and an XML Committee, and the XML Committee works within the frame work of the general IT architecture framework but if you look at it on this perspective there are more or less horizontal but of course we try to co-ordinate very closely. Niels Pagh-Rasmussen is Chairman of the IT Architecture Committee and I am Chairman of the XML Committee then we have a number of working groups or what we also call subcommittees the Danish Information and Documentation TC, the Danish Core Components TC, the Danish e-Business TC and we have new groups developing. We have debate with domains like the health domain to develop a health domain subcommittee and we expect there will be number of domain subcommittees developed within the next year and further on. Then we will establish soon the programme for OIOXML because most of the committees are for government only but it is very important to have a good dialogue with the private sector.

OK to sum up some of our practical experiences and plans; the balance between formalism and pragmatism is very important if we are to do what we are trying to do - and that goes back to what I started talking about in the very beginning; standardisation 10 years ago, no not ten years but in 1996-7 we thought standardisation takes such a long time, its impossible and government is not really that involved traditionally in standardisation, let the private sector do that, but we have to do it somehow, but we cannot really do it in the very former context of traditional government and traditional standardisation also we have to use the different dynamics like projects. We have to combine the top down and bottom up approaches. We have to facilitate very complex processes so that is quite a task not least for us in the XML secretariat to actually facilitate what's going on in a lot of different projects and we have to remove barriers and solve problems for all the projects because otherwise they wont use OIOXML they will not use our paradigm. They will just push the golden button; that's much cheaper, much easier and much faster.

We are beyond the first phase now so what we are really getting to now is beyond the pilot phase, beyond the first inner circle of actors we have to reach out much much more to a lot of institutions, agencies, etc but also we have to deal with larger domains and try to help them organise the standardisation effort and we have to look into the international scope of this much more also. My last thing here is to say that integration without standardisation wont work. Standardisation with the wrong approach wont work and low co-ordination of standardisation initiatives, however local however international, if we do not co-ordinate it will be very costly. We will be going a lot of different directions and re doing a lot of things and doing a lot of things wrong.

I will stop now my time is up there are some links you will be provided with, a copy of this later on there are some links here especially the first one could be like an entrance to some of the content we have. Thank you. Now for a short commercial - the 'www.Oasis.open.org' has an e-Government technical committee you should look into