Building the Business Case for Addressing Semantics in Application Integration

Eric Lambert, Electricité de France R&D Division
Greg Robinson, Xtensible Solutions

Abstract

As is the case for many large utilities, Electricité de France (EDF) is painfully aware of how its efforts to automate business processes have been hampered by incongruent data supplied by a myriad of applications and data stores. Accordingly, EDF R&D is establishing a methodology, tools, and standards for implementing a semantic-based integration infrastructure. This approach turns many data sources into one coherent body of information, making it possible to create new application functionality based on consistent data, even though this data may be coming from multiple sources.

Surprisingly, the biggest challenge is not technical, but rather getting the business participants to agree to the same overall plan. Enterprise application integration, by its very nature, involves multiple organizations, each one under pressure to meet its objectives and having limited spare capacity for other matters. Therefore a key goal of this endeavor is for these organizations to 'want' to leverage this capability for their future integration needs. To obtain the necessary support, benefits must be articulated from each participant's point of view and yet fit into one congruent business plan. This paper discusses how EDF's integration methodology is making use of industry standards, perceived and real issues regarding these standards, and then takes a look at some of the key driving and restraining forces impacting the methodology's adoption by various business participants.

EDF Group in the European Electricity Market

The implementation of the European electricity market calls for the separation of generation, transmission, supply and distribution activities. The gradual integration of the national markets into a single electricity market has created new constraints along with new opportunities. In France, 70% of the French electricity market has been open since July 2004 and it will be fully opened in July 2007.

Today EDF Group has 42.1 million customers around the world, €46.9 billion in revenues, 125,447 MW of installed capacity, and 161,000 employees worldwide. EDF Group in France is divided into several operational divisions: EDF Generation, RTE EDF Transport, Distribution, Supplier.

The Impact of Deregulation on Business Processes

Tremendous changes in business processes have occurred since deregulation took place; a new set of rules defined by the French regulator (Commission de Régulation de l'Electricité) must be satisfied. In order to be competitive and to continue increasing its number of customers, EDF adapted some of its former business processes and created new ones within this new context. EDF entities must also anticipate future changes; for instance, the impact of Distributed Energy

Resources on the MV (medium voltage) network implies new constraints on traditional business processes supported by traditional applications.

Business Processes Supporting Infrastructure Changes

In order to support the changes that occur in business processes, information systems must be able to handle these new challenges. Some parts of a formerly centralized information system management scheme have been distributed among business entities. Now each EDF Operational division (Generation, Transmission, Distribution, Supplier) has an Information Management Department which is coordinated by an Information System Division at the group level, and an Information and Telecommunication Division in charge of deployment. Some major decisions are made at the corporate level, some are made at the division level, and others at the project level. Corporate decisions are usually made after a technical and commercial analysis. Figure 1 depicts the global organization:

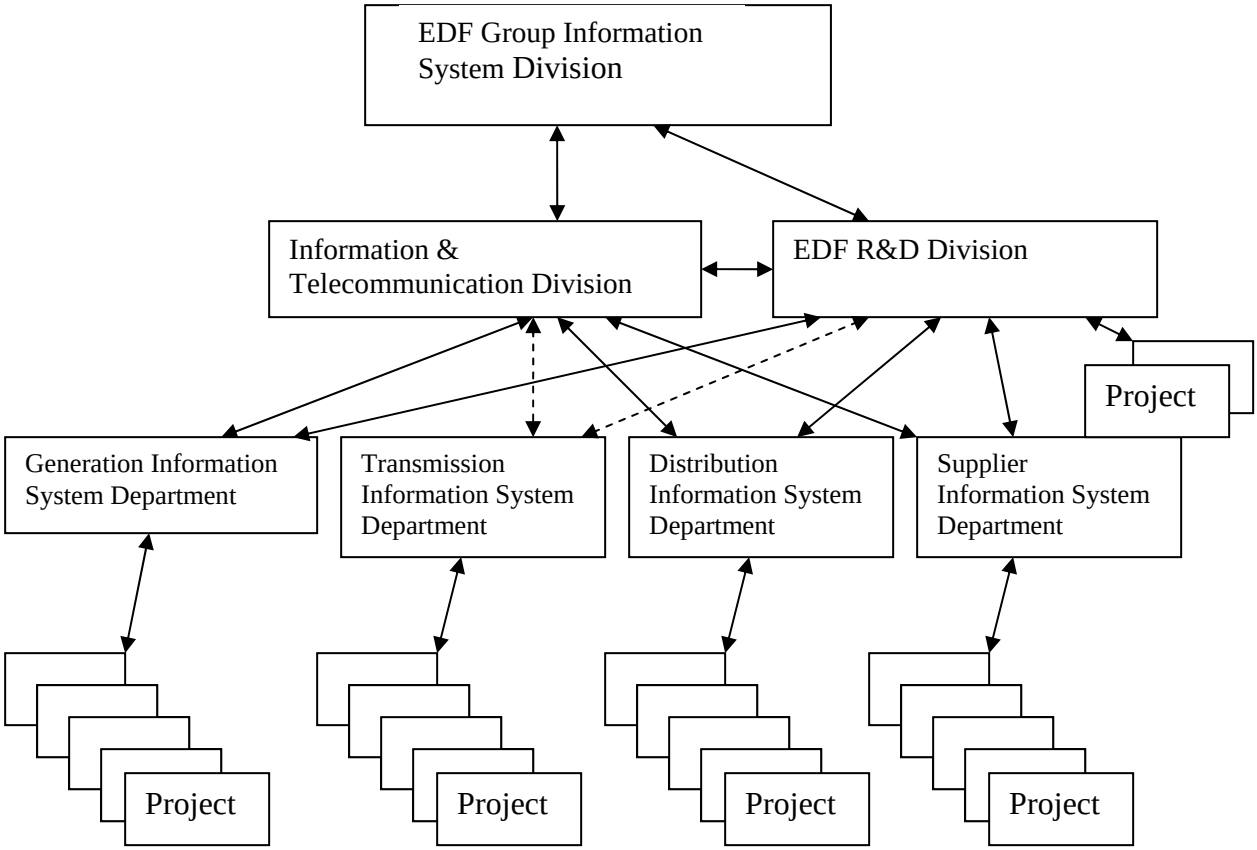


Figure 1. Information System Organisation

As a consequence of all this change, information flow has increased and new integration points have appeared. Unfortunately, so far, projects to automate business processes have produced less than optimal results because efforts have been hampered by incongruent data supplied by a

myriad of applications and data stores. EDF's experience is consistent with market analysts which state that over 50% of system integration costs are attributed to semantic issues. [1]

Facilitating Business Transformation with Standards-Based Integration

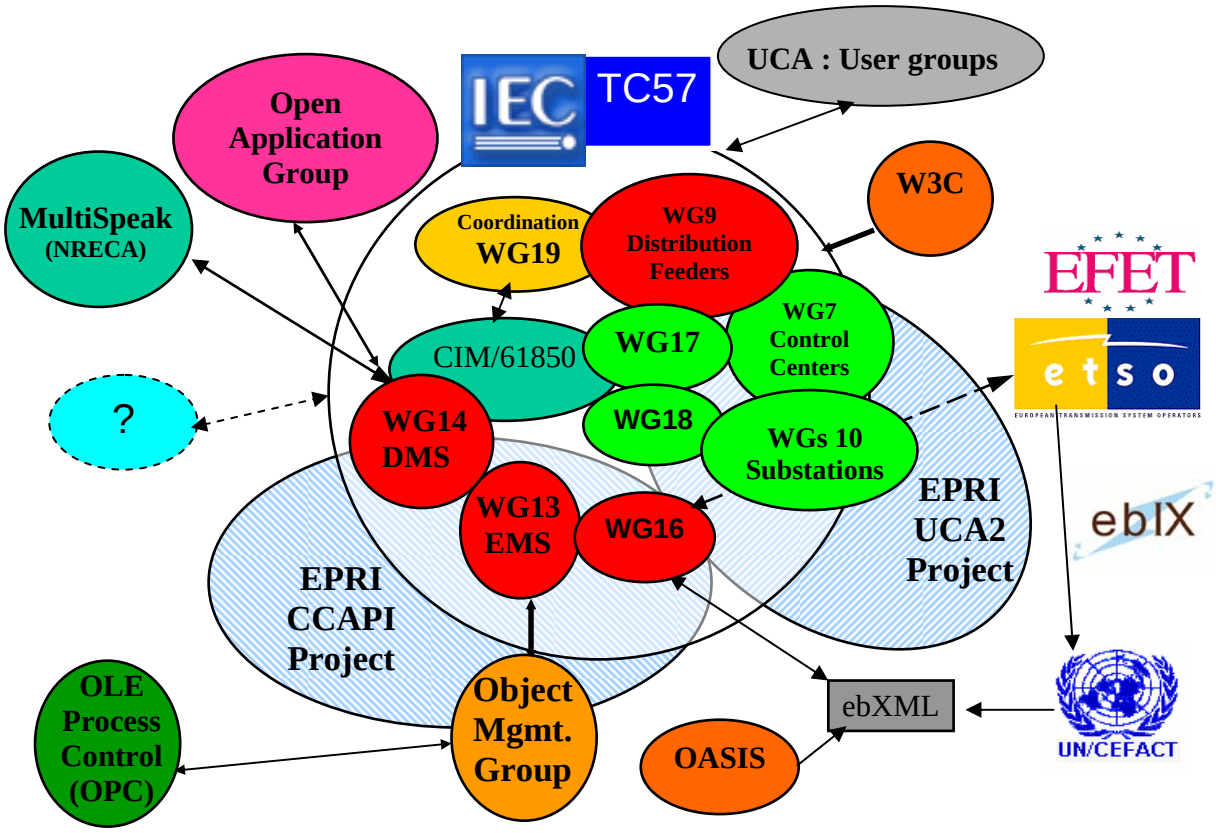
Accordingly, EDF R&D, through its Cimergy Project, is establishing a methodology and a supporting tool set that leverages standards for incrementally implementing a semantic-based integration infrastructure in step with business needs. This approach will enable the turning of many data sources into one coherent body of information that will facilitate business transformation (e.g., business process automation, business activity monitoring, decision support) and make it possible to create new application functionality based on consistent data even though this data may be coming from multiple sources.

The Cimergy Project team members firmly believe that the pragmatic use of standards is the linchpin in the methodology for lowering integration costs. The goal of this Model-Driven Integration (MDI) approach is to enable IT staff to effectively work with business-sponsored projects to achieve the following business benefits:

1. Reduces labor to maintain overlapping data in multiple applications.
2. Systematically generated common structure and common vocabulary reduces design time effort as well as coding errors, thereby avoiding many of the semantic errors that are difficult to find and correct during implementation.
3. Reduces performance errors caused by inconsistent information. A given piece of information can be retrieved about an asset, such as a transformer, from the right source – with the confidence that the information is actually correct.
4. Provides the least cost approach for enabling timely and accurate reporting and decision support for managing Key Performance Indicators (KPIs) and handling regulatory obligations.
5. Performs faster implementation of application functionality and business processes.
6. Reduces cost to maintain and extend existing applications.
7. Reduces risk of project schedule and budget overruns by having rules of engagement articulated for every well defined component of a project.

A Quick Look at the Relevant Standards

Accomplishing this is a challenge when so many standards, most still evolving, are relevant to EDF R&D's integration strategy. The following figure depicts the global collaboration of standards bodies that the Cimergy Project deals with:



The standards organizations we are participating in, or are following include the following:

- The International Electrotechnical Commission (IEC) is the leading global organization that prepares and publishes international standards for all electrical, electronic and related technologies. These serve as a basis for national standardization and as references when drafting international tenders and contracts. EDF participates in IEC TC57, the Technical Committee in charge of Power Systems Management and information exchange. This technical committee is composed of several working groups (WGs), among them WG16 is in charge of Electricity markets, WG13 of Energy Management Systems (61970), WG14 of defining System Interfaces for Distribution Management (61968). WG13 and WG14 are responsible for collaboratively building and standardizing the industry's Common Information Model (CIM). WG10 has developed a new standard (61850) for Substations and Intelligent Electronic Devices. [2]
- Object Management Group is an open membership, not-for-profit consortium that produces and maintains computer industry specifications for interoperable enterprise applications. Among several activities it defines UML (Unified Modeling Language).
- W3C (World Wide Web Consortium) develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential. W3C is a forum for information, commerce, communication, and collective understanding. Among other specifications it defines eXtensible Markup Language.
- ETSO is the Association of European Transmission System Operators, which oversees the harmonization of the conditions governing the access and use of the networks while

ensuring both transparency and equity. ETSO covers the entire interconnected European Transmission grid.

- EFET (European Federation of Energy traders) is a group of more than 70 energy trading companies from 18 European countries dedicated to stimulate and promote energy trading throughout Europe.
- EbIx (Energy Business Information Exchange) advances, develops and standardizes the use of electronic information in the energy industry for the electricity and gas European markets. It covers the needs for the wholesale market (upstream) and the retail market (downstream).
- UN-CEFACT (United Nations Centre for Trade Facilitation and Electronic Business) provides Trade Facilitation Recommendations, Electronic Business Standards, and Technical Specifications. Among the latter, the Core Components Technical Specification, which is part of the ebXML Framework, is of particular interest for us.
- OASIS is a not-for-profit global consortium that drives the development convergence and adoption of e-business standards. They define ebXML.
- OAG (Open Applications Group) is a not-for-profit open standards group building process-based XML standards for both B2B and A2A integration. The Open Applications Group was formed in late 1994 as the first post-EDI organization focusing on improving the state of application integration
- Harmonization between UN/ECE, IEC, ISO and ITU. [3] The relations between UN/ECE (a body for UN/CEFACT), IEC, ISO and ITU are defined in a MoU (Memorandum of Understanding). The MoU defines precisely the responsibilities of each body in the domain of electronic business, and was done to minimize the risk of divergent and competitive approaches to standardization, to avoid duplication of efforts and to avoid confusion amongst users. This MoU states that UN/CEFACT is responsible for defining guidelines for electronic business and the EDIFACT syntax and messages. IEC is responsible for Business Processes, Data Definition and Semantics for its domain, the electrotechnical field.

Leveraging these standards requires EDF R&D to define a methodology and tools so that the standards may be used in a systematic and efficient manner. Unfortunately, these standard development efforts never seem to finish; they continue to progress and evolve, and each have some issues that need sorting. So while EDF can enjoy the benefits of the ‘half full glass’, we must also manage the side effects of the ‘glass being half empty’.

Perceived and Real Problems with Standards

The CIM is easily criticized; let’s look at the ‘half-empty’ part of the glass for a moment. As technologies and standards have evolved, there has necessarily been some convergence. For example, overall integration would be simplified if data from Enterprise Resource Planning (ERP) systems using OAG-based information exchange could be understood by applications using CIM-based information exchange. To accommodate this goal, a revised version will be necessary that will encompass several add-ins provided by Distribution Business Process Analysis and methodology concepts following the UN-CEFACT Core Components Technical Specification and ISO 1179 (Specification and Standardization of Data Elements).

Even though CIM is defined as an international standard, it is currently only provided in English. There are plans to provide a French definition for each class, each attribute, each relation of the UML CIM model, but in the meantime, it is awkward for non-native English speakers to use and therefore is subjected to criticism.

The CIM has evolved in a way that has made configuration management across projects a complex endeavor for companies. Fortunately, this problem has been recognized by the CIM community and is being addressed by two groups working in concert: the IEC and the CIM Users Group, the later having been formed in 2005. Two developments will help: (1) a more formalized versioning and release process and (2) a goal for the CIM editors of the IEC working groups to make future extensions be just that, extensions. In other words, increased effort will be made to add new functionality into the CIM in a way that has minimal impact on existing classes.

Some people think that the general idea of "establishing a common language" is an unreachable utopia, with people trying to make the CIM become the "Esperanto" of Electric Utilities. But this is a misunderstanding for two reasons: (1) the intended use of the CIM is for inter-application integration, not intra-application integration. Therefore the scope is limited to only the information that needs to be exchanged among 'black box' application systems. It is not intended to provide the language internal to application systems – these systems still use their native language for internal communication. (2) The CIM is aimed at being a technical integration language. Using it as a basis for common business terms is unnatural and will create frustration in business-oriented people. Rather, business process modeling should be performed in natural business terms that are mapped, and in some cases are identical, to CIM terms. For example, Fact Modeling is used by some companies for establishing their business terms. It keeps the scope of this exercise at the right level and avoids the opposite problem - trying to use business terms as a basis for technical models.

Others argue that a methodology deriving XML message types from a UML model is not necessary, and that when you have an interface file, you only need to translate it in XML. However, this is merely a disguise for a traditional point-to-point interface implemented with the latest 'hype' technologies. This approach results in brittle integration solutions that do not effectively decouple data elements from their sources so that they can be used for scalable business solutions (such as business process automation, business activity monitoring and decision support). The model is requisite for ensuring that every element of every message is used consistently across messages on an enterprise-wide basis, independent of which system produced data elements and which projects developed system interfaces consuming disparate data elements.

Some do not believe in the value of standards and prefer to just use the interfaces supplied with off-the-shelf products. Application project managers are particularly susceptible to this way of thinking because it appears to be the easiest solution for their project and all of the risk is apparently born by the vendor. Unfortunately, this approach to enterprise application integration adds yet another set of technologies and methods to a convoluted web of disparate technologies and methods that the enterprise must maintain, modify and extend. With every project supplying a different semantic solution, any type of enterprise-wide business functions, such as business process automation, business activity monitoring, or decision support that draws on data from more that one system, becomes complex to implement and very hard to maintain. Furthermore,

this type of ad-hoc integration infrastructure debilitates enterprise-wide business functions and prevents them from being scaled-up as needed to transform business. Without realizing it, the enterprise becomes less flexible with each project and costs associated with business change is increased far beyond what would have been necessary.

Others argue also that products compliant to the standards are not offered in the marketplace and therefore the real benefits of the standards do not really exist. For example, in the distribution area, there are not many planning tools or GIS systems offered with off-the-shelf CIM APIs. This is certainly a reasonable perspective and steps to resolve this problem are being taken by the CIM community. Several rounds of interoperability tests have been performed around the part of the CIM (IEC 61970) dealing with the traditional Energy Management Systems (EMS). However, the good news is that the first round of interoperability tests are scheduled to be performed with a number of these types of vendors during the first half of 2006. And now that a CIM Users Group has been established, it is likely that many more tests will follow under its guidance.

Another argument is often made regarding a questionable return on investment. What is the competitive advantage the utility will have if all business processes are standardized? Since 1998, there have been many papers and presentations about how large utilities successfully put into production numerous applications using CIM-based messaging solutions. Unfortunately, maintaining and extending these solutions has been a challenge because of the complexity of evolving esoteric methodologies, architectures, and standards that are used in their solutions. Projects had to coble together a vast array of products requiring significant and often redundant data entry to support the full software lifecycle. After projects were completed, many of the experts required to complete them were lost to other projects or companies, resulting in significant knowledge gaps for follow-on projects. Fortunately, with the advancements of W3C standards, along with their use the utility industry IEC standards, the complexity and amount of expertise required by project teams has been reduced with each passing year. Furthermore, vendors are now offering products that greatly facilitate project teams' use of these standards and industry models. Consequently, it can now be shown that it is cost competitive to implement standards-based solutions when compared with traditional vendor proprietary-driven interface solutions. In business cases nowadays, standards-based solutions can show reasonable pay-back on a tactical basis while simultaneously offering major strategic benefits.

How EDF is Overcoming these Problems

Usefulness of Interoperability Tests

Interoperability tests are essential to stabilize the CIM and expand its use among market participants, as well as to obtain concrete feedback on how it is being implemented by utilities and vendors. It is for that reason that EDF R&D, through its Cimergy Project, has been participating in these tests since 2004. Organized by EPRI since 1999, interoperability tests have been very useful for EDF R&D to understand what CIM interoperability really means. Participating in these tests are time consuming and costly, but they are worthwhile because they help to consolidate the standards (CIM model, profiles based on CIM, common interfaces, and common services allowing the exchange of data).

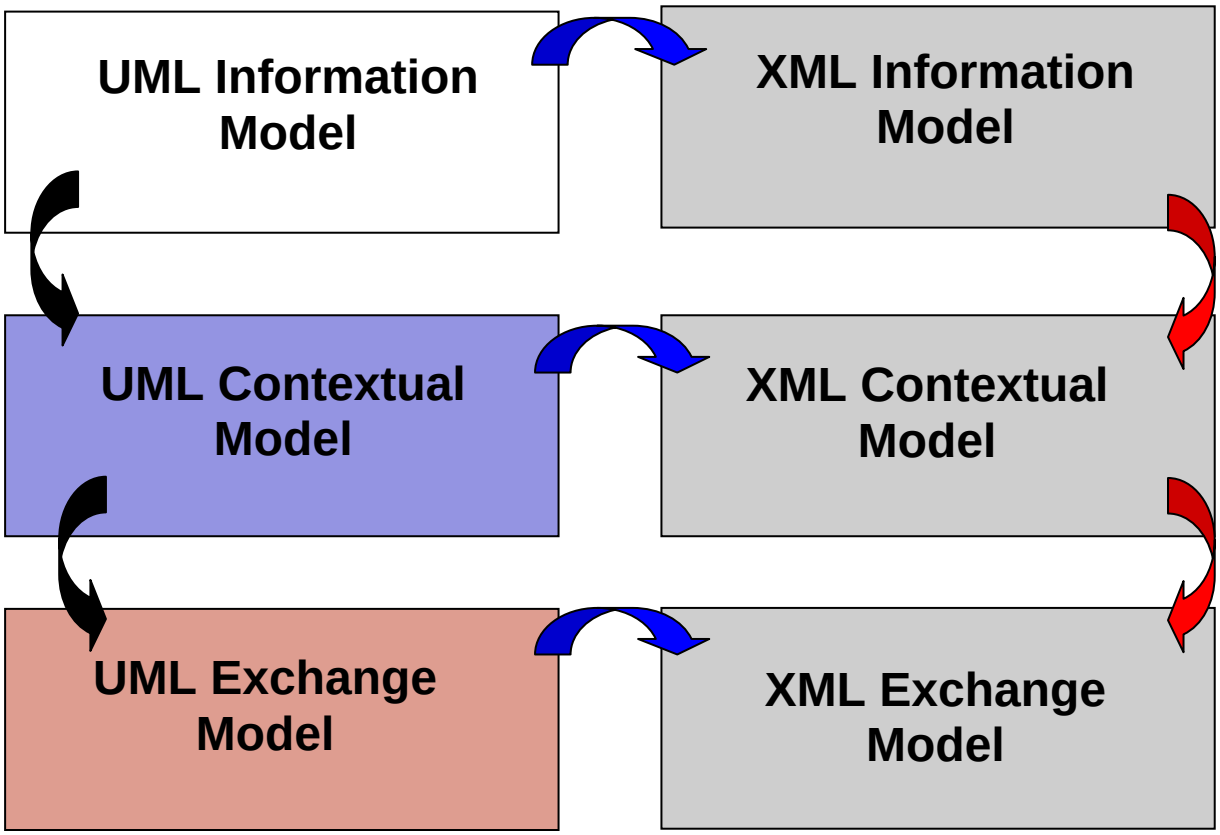
Internally, the Cimergy Project is establishing a methodology and tools to support CIM usage:

1. The methodology is based on the UN-CEFACT Core Component Technical Specification. This methodology has been retained by ETSO which has defined several message schemas. The key idea is that the CIM could play the role of the Core Component model of the Electrical Industry.

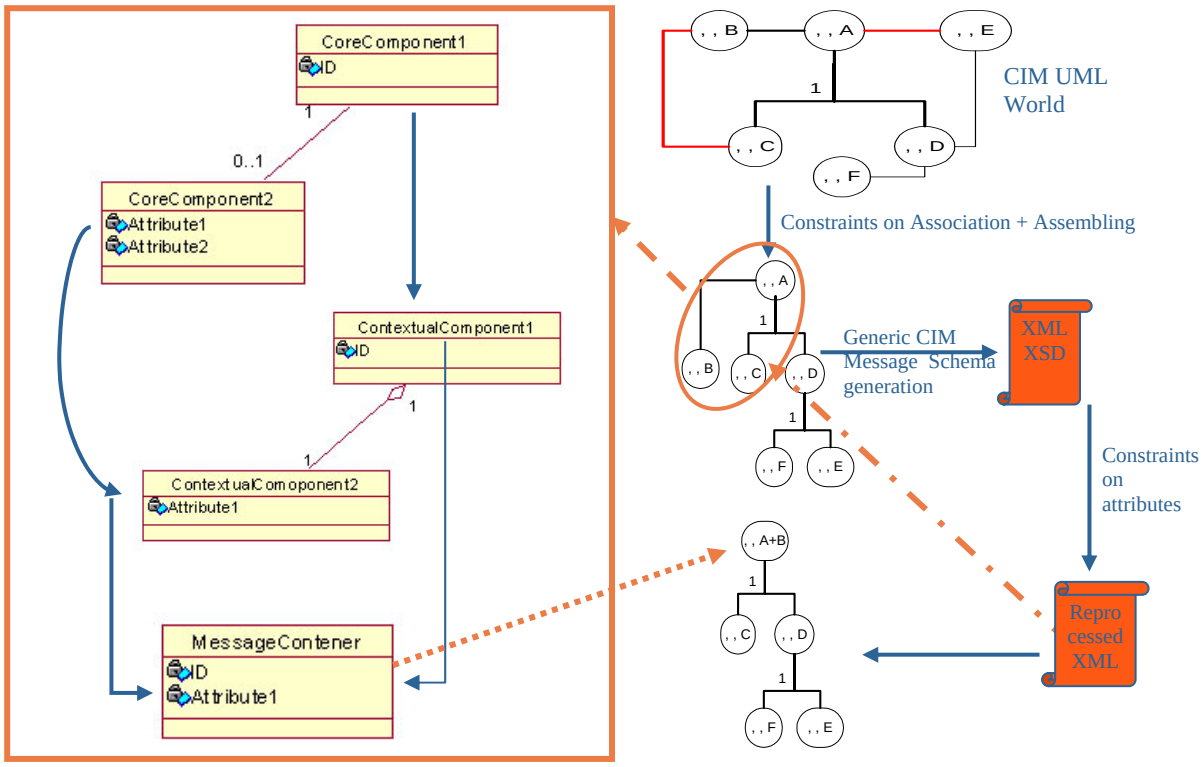
2. The tools include ones EDF R&D has either developed or selected. Among these tools, Xtensible Solutions' MDI Framework is used, which provides good support for following the methodology. We are looking forward to tools that are compliant with standards UML 2 and XMI. Our goal is to have a standardized methodology that will be implemented using several tools.

   To complete this Top-Down approach, a bottom-up approach has been incorporated. It primarily consists of the development of several APIs based on the CIM. These APIs followed the principals existing at that time. We have defined API for several LV (Low Voltage), MV, and HV (High Voltage) applications.

Key steps of the methodology are depicted in the two following figures. The first one is a conceptual overview of the overall methodology.



The next figure depicts some of the key mechanics used to implement the conceptual view above.



This last figure shows three distinct UML levels on the left side, starting at the core component level, then contextual views are created, and finally a message view is defined. On the right side, the corresponding XML world is represented. Constraints are applied when passing from one level to another. Two key requirements are having the ability to track each step in the methodology and that each step follows a standard convention.

Validating the Methodology

In order to validate the methodology, several proofs of concept mini-projects have been implemented with the Distribution Operational Division. In accordance with the methodology, use cases were used to establish requirements for the business processes. This work has helped to promote the approach and to convince managers and other technical persons that it is valuable.

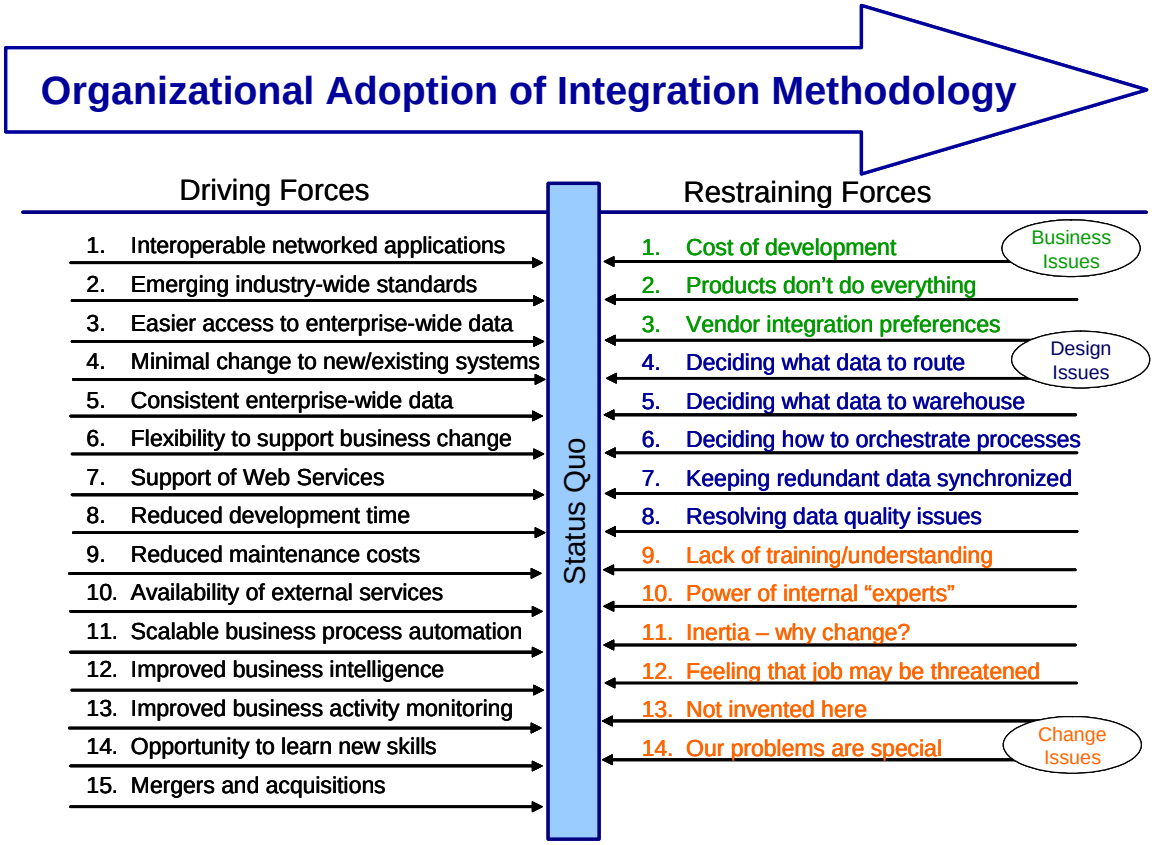
The next step is to consolidate the CIM model and the UN-CEFACT CCTS (Core Component Technical Specification). Also, the communication of this aspect is very important to succeed in this global vision - for Operational Divisions to successfully use this methodology on their concrete projects. That's the reason that since 2004, we have been offering a CIM course named "The CIM: How to link business activities around the electrical network". This course helps to promote the CIM model, the methodology, and related tools. It is dedicated to EDF employees, from both the business side as well as the information technology side.

We also work closely with Operational Divisions, in accordance with the Information and Telecommunication Department. They are experimenting with the UN-CEFACT Core Component Technical Specification as well. One of the key issues in the near future should be how we can industrialize this process, in order to promote the methodology and tools to projects

conducted in Operational Divisions. The tools aspect is also very important, as the Information and Telecommunication Department recommends the tools and negotiates purchase conditions with editors.

Business Adoption Challenges

Surprisingly, the biggest challenge is not technical, but rather getting the business participants to agree to the same overall plan. Enterprise application integration, by its very nature, involves multiple organizations, each one under pressure to meet its objectives and having limited spare capacity for other matters. So for this methodology to be successfully used, business and IT staff must all believe that they are truly better off for having invested their time to learn and apply it. The following diagram summarizes forces driving their support as well as forces hindering it. The purpose of this type of diagram is to make driving and restraining forces visible so that decisions concerning change can be made with the best available information. Obviously, to improve chances of success, the goal is to strengthen driving forces and weaken restraining forces. [4]



Driving Forces are those which help achieve a goal, in this case organizational adoption of the methodology developed by EDF R&D:

1. *Interoperable networked applications.* Having a common integration language enables applications to be treated as interoperable black box components, in much the same manner as audiovisual electronic components (e.g., stereo receiver, speakers, CD player, DVD player, tape player, mixer, television, etc.) can be connected and swapped in and out using standard cables.

2. *Emerging industry-wide standards.* Most of the restraining forces are issues with semantics or the meaning of data. The standards provide an application-neutral common integration language that settles most of the data debates that frequently occur between project teams.
3. *Easier access to enterprise-wide data.* Data is made easy to discover and access through a consistent set of services irrespective of the sources of the data.
4. *Minimal change to new/existing systems.* EDF places high value on vendor applications that are IEC 61970 and IEC 61968 compliant. However, if not, the methodology provides the means to insulate proprietary interfaces and keep them from creeping into the rest of EDF's enterprise. Accordingly, integration is performed with the vendor off-the-shelf interfaces so that upgrades or application replacements can be put in place with minimal impact to EDF's other systems and business processes.
5. *Consistent enterprise-wide data.* Data is consistent and congruent irrespective of the sources of the data.
6. *Flexibility to support business change.* Implementing an integration architecture as described in IEC 61968-1 and based on a logical design using the IEC Interface Reference Model (IRM), helps ensure that information exchange solutions successfully transcend product variations, continuous technology changes, and business changes.
7. *Support of Web Services.* Using the methodology, semantics are consistent and independent of the middleware employed for particular application interfaces. Applications supporting interfaces based on the Web Services Description Language (WSDL) can now be accommodated and are in most cases preferred.
8. *Reduced development time.* While it does take some time to learn pre-built industry information models, canonical message models, use cases, etc., doing so results in a net productivity gain. Projects can now shift more of their time to the application at hand and less on determining how to integrate their application with other applications. Furthermore, risk of missing integration requirements is significantly reduced as these industry artifacts have been used in multiple utilities with corrections feed back to the IEC.
9. *Reduced maintenance costs.* In terms of both time and money, it is expensive to support a different set of interfaces emanating from each application system. Now vendor proprietary technologies and methods are hidden behind well-defined interfaces. This enables applications to be upgraded or replaced with significantly less effort and without the side-effects that so commonly occur when traditional integration techniques are used.
10. *Availability of external services.* While standard interfaces have proven to be desirable within the enterprise, it is imperative for business to business integration. Using the same semantics across both minimizes development costs as well as the risk of errors in transformations spanning different business worlds.
11. *Scalable business process automation.* Consistent data that is available in a semantic layer enables business process automation to be implemented independent of the number of applications and data stores involved. To do otherwise results in brittle and expensive solutions that will not scale to any significant size. A model-driven semantic layer is requisite to scalable business process automation, monitoring and management.
12. *Improved business intelligence.* (same rational as for 'business process automation')
13. *Improved business activity monitoring.* (same rational as for 'business process automation')

14. *Opportunity to learn new skills.* While it is common for many employees to fear most any kind of change, most all employees look forward to opportunities to learn new skills with "hot new" technologies and improve their value.
15. *Mergers and acquisitions.* The standards-based integration infrastructure provides a neutral vehicle for integrating disparate applications from different companies. Decisions to replace individual applications can wait until there is a business need rather than doing so merely for integration purposes.

Restraining Forces are categorized as business issues, design issues, and change issues:

Business Issues

1. *Cost of development.* Both the initial and life-cycle costs to integrate systems is arguably lower using this methodology. To avoid an unnecessary increase in development costs, when compared with traditional approaches, it is important to integrate in increments in step with actual business needs. It is not necessary to have everything updated at the same time just for the sake of using the same technology approach.
2. *Products don't do everything.* EDF R&D has worked hard to put together a methodology supported by tools to minimize the gaps between tools used in the development environment.
3. *Vendor integration preferences.* Many vendors employ strategies to achieve "customer lock-in." While they may agree to implement standard interfaces, these types of vendors will inevitably find ways to influence their customer to use their proprietary interfaces. They have many tricks and are well practiced at applying them.

Design Issues

4. *Deciding what data to route.* This is a challenge with any system integration, but is aggravated when doing so requires using a new methodology that is a paradigm shift (i.e., service oriented architecture leveraging information model-driven canonical message models). The best solution is to provide the necessary training and be sure that some of your staff (whether they be employees or contractors) have actual project experience.
5. *Deciding what data to warehouse.* (same rational as for 'deciding what data to route')
6. *Deciding how to orchestrate processes.* (same rational as for 'deciding what data to route')
7. *Keeping redundant data synchronized.* (same rational as for 'deciding what data to route')
8. *Resolving data quality issues.* (same rational as for 'deciding what data to route')

Change Issues

9. *Lack of training/understanding.* This negative should be turned into a positive developmental opportunity for employees. It is imperative that the team is competent. A common mistake projects make is to do work with available resources. While available resources are obviously the first place to look for required skills, it will be a disaster in the making if key positions are filled with people who do not have a good understand of what needs to be accomplished and how to go about it.
10. *Power of internal "experts".* Experts in old systems or technologies that are subject to being replaced by the new systems/approaches can be most innovative at coming up with reasons why the new approach will fail. Without staff capable of separating out legitimate

concerns from emotional reactions, projects can make poor decisions resulting in substantial additional costs and schedule slippages.

11. *Inertia – why change?* Many people are comfortable with the way things are and would prefer not to change them. A project challenge is to get these people to see how the change will be smoothly implemented and understand how much better it will be for them when the work is done. If they will not be part of the new way, effort should be made get those people placed in new situations that will be better suited to their needs.
12. *Feeling that job may be threatened.* (same rational as 'Inertia – why change?)
13. *Not invented here.* It is common for a person to invest tremendous amounts of energy over time into an existing solution that is optimized for that situation, at least in their minds. An off-the-shelf application or a context aware integration methodology will never stack up in that person's mind, and probably is in fact, not as good in many respects. When one's baby is at risk, it is hard for that person to think objectively for the good of the whole. The goal here is to get the person to incorporate the good from the old way into the new way, which is a people management challenge as much as it is a technical one.
14. *Our problems are special.* This is a common belief that needs to be openly worked through. Our experience is that in over 90% of the cases, industry standards can support these special needs with only a modest amount of effort.

Conclusion

EDF R&D's initial usage of EDF's model-driven integration methodology, combining CIM model and UN-CEFACT Core Component Technical Specification is a promising approach. Even though the methodology and tool set has begun to be sorted out, success hinges on business participants deciding to use it. Each of the above restraining forces along with the perceived and real standards issues will need careful attention at all impacted levels if the methodology is to become a normal part of business across EDF. While this is no simple task, the benefits are well worth it :

- The reusable methodology, with its resulting artifacts, articulate how business objectives are implemented and provide end-to-end requirements traceability.
- As more projects leverage the infrastructure, more data becomes available as part of one coherent body of information. This makes it possible to create new application functionality based on consistent data even though this data may be coming from multiple disparate sources.
- Faster integration of application systems and information should improve EDF's ability to react quickly to business changes while still providing the right information to people when they need it.
- Improved ability to integrate business processes across commercial-off-the-shelf (COTS) applications is lowering dependence on individual vendors.
- Re-use and adoption by future projects is encouraged by an extensible and adaptable design. There is one methodology and managed tool set to ensure consistency and leverage, leading to implementation of fewer, simpler, more consistent interfaces throughout the enterprise.

Acknowledgements
The authors thank the EDF R&D Cimergy Team whom are making the methodology useful for integration around EDF and are contributing many helpful ideas to the IEC standardizations processes.

References
[1] McComb, Dave, Semantics in Business Systems – The Savvy Manager's Guide, Morgan Kaufmann Publishers, 2004.
[2] Presentations and committee drafts from IEC TC57 Working Groups 13, 14, 16, and 19 (refer to http://www.cimuser.org), 2005.
[3] Memorandum of Understanding between The International Electrotechnical Commission, The International Standardization Organisation, The International Telecommunication Union and the United Nations Economic Commission for Europe concerning standardization in the field of Electronic Business. http://www.itu.int/ITU-T/e-business/mou/index.html
[4] Barry, Douglass K., Web Services and Service-Oriented Architectures – The Savy Manager's Guide, Morgan Kaufmann Publishers, 2003.