



***Standard du gouvernement du Québec
sur les ressources informationnelles***

**Cadre commun d'interopérabilité
Volet 1 - Intégration des données
et des processus d'affaires
Version 1.3**

31 mars 2005

Avertissement

Ce document ne constitue pas pour l'instant un standard officiel du gouvernement du Québec. Il a franchi l'étape de l'enquête élargie qui précède son approbation par le Conseil du trésor. Son contenu est encore sujet à modification sans préavis.

Toute référence à ce document doit donc obligatoirement inclure la mention « ébauche de standard en cours d'officialisation », en plus du numéro et du nom du standard, ainsi que du numéro et de la date de la version.

Pour obtenir des renseignements supplémentaires à propos de ce document, veuillez vous adresser au Sous-ministériat à l'encadrement des ressources informationnelles (SMERI) du Ministère des Services gouvernementaux au (418) 528-6148.

Table des matières

SIGLES ET ACRONYMES	5
AVANT-PROPOS	7
1 MODÉLISATION ET REPRÉSENTATION DES DONNÉES	8
1.1 MÉTHODE D'ANALYSE ET DE CONCEPTION.....	8
1.2 RÉUTILISATION ET PARTAGE DES DONNÉES ET DE LEURS DESCRIPTIONS AU GOUVERNEMENT DU QUÉBEC	11
1.3 CAS DES DOMAINES ÉNUMÉRATIFS CODIFIÉS	11
1.4 LANGAGE DE DESCRIPTION STRUCTURÉE DE DONNÉES (XML)	13
1.5 FORMATS DES IDENTIFIANTS DE RESSOURCES	14
1.6 VOCABULAIRE DE DESCRIPTION UTILISÉ DANS LES MODÈLES DE DOCUMENTS ET DANS LES DOCUMENTS XML	17
1.7 DÉFINITION ET PUBLICATION DES MODÈLES DE DOCUMENTS	20
2 DOCUMENTS ET PROCESSUS D'AFFAIRES DANS L'ADMINISTRATION QUÉBÉCOISE	24
2.1 CLASSIFICATION DES DONNÉES DANS L'ADMINISTRATION QUÉBÉCOISE.....	24
2.2 ACCESSIBILITÉ ET MANIPULATION DES DOCUMENTS XML	28
2.3 UNE INFRASTRUCTURE STANDARD POUR LE COMMERCE ÉLECTRONIQUE.....	33
2.4 COMPOSANTS LOGICIELS (SERVEURS ET SCRIPTS) RÉUTILISABLES POUR LA LOGIQUE D'AFFAIRES.....	39
2.5 REGISTRE RÉFÉRENTIEL	43
3 RÉPERTOIRE.....	47
3.1 UN OBJET PARTICULIER DU RÉPERTOIRE : LE CERTIFICAT NUMÉRIQUE.....	50
4 APPLICATIONS AU DOMAINE DE L'INFORMATION GÉOGRAPHIQUE.....	53
ANNEXE : DÉFINITIONS.....	54

Table des illustrations

<i>Figure 1 : Diagrammes UML.....</i>	<i>8</i>
<i>Figure 2 : Inclusion des ensembles d'identifiants de ressources.....</i>	<i>15</i>
<i>Figure 3 : Exemple fictif de qualification du terme titre grâce aux espaces de nominatifs.....</i>	<i>19</i>
<i>Figure 4 : Interactions d'affaires entre deux compagnies, basées sur ebXML.....</i>	<i>36</i>
<i>Figure 5 : Schéma de principe du registre référentiel du gouvernement du Québec.....</i>	<i>45</i>
<i>Figure 6 : Description des champs d'un certificat X.509 (source : www.hsc.fr).....</i>	<i>51</i>

Sigles et acronymes

AEG	Architecture d'entreprise gouvernementale
AGSIN	Architecture d'entreprise de la sécurité de l'information numérique
ASCII	American Standard for Information Interchange
AVI	Audio Video Interleave (format AVI)
BIND	Berkeley Internet Name Domain
BMP	Bitmap
CGRI	Cadre de gestion des ressources informationnelles
CSS	Cascading Style Sheets (feuilles de style en cascade)
CSV	Comma Separated Values
CTP	Coût total de possession
DLAI	Direction de l'architecture et de l'infrastructure
DNS	Domain Name Service
ESMTP	Extended Simple Mail Transfer Protocol
FTP	File Transfer Protocol
FTPS	FTP/ SSL (Secure Socket Layer)
GIF	Graphics Interchange Format (format GIF)
HDML	Handheld Device Markup Language (langage HDML)
HTML	HyperText Markup Language (langage HTML)
HTTP	HyperText Transport Protocol (protocole HTTP)
HTTPS	Hypertext Transport Protocol Secure
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
IMAP	Internet Access Message Protocol
IP	Internet Protocol
IPSec	Internet Protocol Security (association de sécurité)
ISO	International Organization for Standardization
JPEG	Joint Photography Experts Group (norme JPEG)
LAN	Local Area Network
LDAP	Lightweight Directory Access Protocol
MathML	Mathematical Markup Language
MGCP	Media Gateway Control Protocol
MIME	Multipurpose Internet Mail Extensions (protocole MIME)
M/O	Ministères et organismes
MPEG	Moving Picture Experts Group (norme MPEG)
MP3	MPEG-1 Audio Layer 3
NNTP	Network News Transfer Protocol
OASIS	Organization for Structured Information Standards
PCL	Printer Command Language (langage PLC)
PDF	Portable Document Format (format PDF)
PES	Prestation électronique de services
PGRI	Plan de gestion des ressources informationnelles

PLC	Packet Loss Concealment
PIM	Protocol Independent Multicast
PNG	Portable Network Graphics (format PNG)
POP	Post Office Protocol
PS	PostScript
RADIUS	Remote Authentication Dial-in User Service
RDF	Resource Description Framework (système RDF)
RFC	Request for Comment
RTF	Rich Text Format (format RTF)
RTP	Real Time Protocol (protocole RTP)
SDC	StarOffice Spreadsheet
SFTP	Secure File Transfer Protocol
SGQRI	Standard du Gouvernement du Québec pour les Ressources Informationnelles
S/MIME	Secure-Multipurpose Internet Mail Extensions
SMTP	Simple Mail Transfer Protocol
SSH	Secure SHell
SSL	Secure Socket Layer
SWF	Shock Wave Flash
SSML	Speech Synthesis Markup Language
SVG	Scalable Vector Graphics
TACACS	The Terminal Access Controller Access Control System
TAR	Tape Archive
TCP	Transmission Control Protocol
TIFF	Tagged Image File Format (format TIFF)
TSIG	Transaction Signature
UDP	User Datagram Protocol
UIT	Union internationale des télécommunications
UTF	Unicode Transformation Format
URI	Uniform Resource Identifier
WAP	Wireless Application Protocol (protocole WAP)
W3C	Consortium W3C
WMA	Windows Media Audio
WML	Wireless Markup Language (langage WML)
XHTML	eXtensible HyperText Markup Language
XML	eXtended Markup Language (langage XML)
XSL	eXtended Stylesheet Language

Avant-propos

L'intégration des données¹, qui existe dans une très grande variété de formats et de sources, et celle des processus d'affaires exigent l'adoption de normes et de standards afin de favoriser l'interopérabilité.

Ces processus mettent au centre des technologies de l'information le document, qui, dans l'univers de la normalisation internationale, et tout autant dans la loi sur [le cadre juridique des technologies de l'information](#), devient un « objet informationnel générique », soit :

- la représentation d'une structure logique de contenu unifiée pouvant être représentée dans divers médias et formats;
- un composant qui sert d'interface pour la perception humaine;
- la délimitation d'un ensemble précis de données pour la signature et la responsabilité légale.

L'adoption d'un formalisme de *modélisation* (UML) devrait faciliter grandement la compréhension et les échanges entre les divers intervenants impliqués dans le développement de systèmes, et encourager le partage de définitions et de modèles, quel que soit leur domaine d'application.

En outre, le recours à un *langage de description de données* standard, XML, associé à la définition de *modèles de documents* (schémas XML) et à leur validation, permettra la réutilisation des documents dans l'Administration québécoise et leur classification en fonction de profils de métadonnées.

Toutefois, la classification des *données* et des *processus d'affaires* à l'échelle de l'Administration québécoise ne peut se faire qu'au moyen de la définition préalable des *profils de métadonnées* à retenir (sémantique).

De plus, il faut appliquer des règles d'affaires pour gérer la conservation des documents, les formulaires. Les documents alimentent les processus d'affaires en tant qu'objets valables juridiquement tout en alimentant les systèmes d'information en données.

La spécification et le déploiement des *solutions technologiques* doivent permettre de produire le *registre² référentiel³* qui hébergera les schémas, les documents partageables et les métadonnées ainsi que les composants partageables de la logique applicative (composants logiciels serveurs, scripts ou composants élémentaires ebXML), le tout garantissant l'interopérabilité et la réutilisation.

Dans cette toile tissée entre acteurs, processus et documents de transaction, des privilèges et droits d'accès doivent être définis respectivement pour les acteurs et les documents en fonction de leur profil, lequel sera enregistré dans un répertoire⁴.

Enfin, un domaine d'application essentiel, la géomatique, mobilise de façon horizontale, par sa richesse et sa complexité, l'ensemble de ces ressources.

¹ Voir la définition de *donnée* en annexe.

² Voir la définition de *registre* en annexe.

³ Voir la définition de *référentiel* en annexe.

⁴ Voir la définition de *répertoire* en annexe.

1 Modélisation et représentation des données

1.1 Méthode d'analyse et de conception

Contexte

UML se définit non comme une méthode, mais plutôt comme un langage de modélisation graphique et textuel destiné à comprendre et décrire des besoins, spécifier et documenter des systèmes, concevoir des solutions et communiquer des points de vue.

Il est composé de plusieurs types de diagrammes permettant de représenter différentes perspectives du système à bâtir, à partir de la capture des besoins jusqu'au modèle de déploiement. Plus précisément, UML est basé sur l'élaboration de *cas d'utilisation* et sur la conception d'un ensemble de diagrammes représentant la *structure statique* d'un système (ex. : diagramme de classes, diagramme d'objets) et la *structure dynamique* (ex. : diagramme d'états, diagramme de séquence). On compte au total neuf types de diagrammes différents.

Le schéma de la figure 1 représente les types de diagrammes qui peuvent être utilisés dans la conception d'un système avec UML.

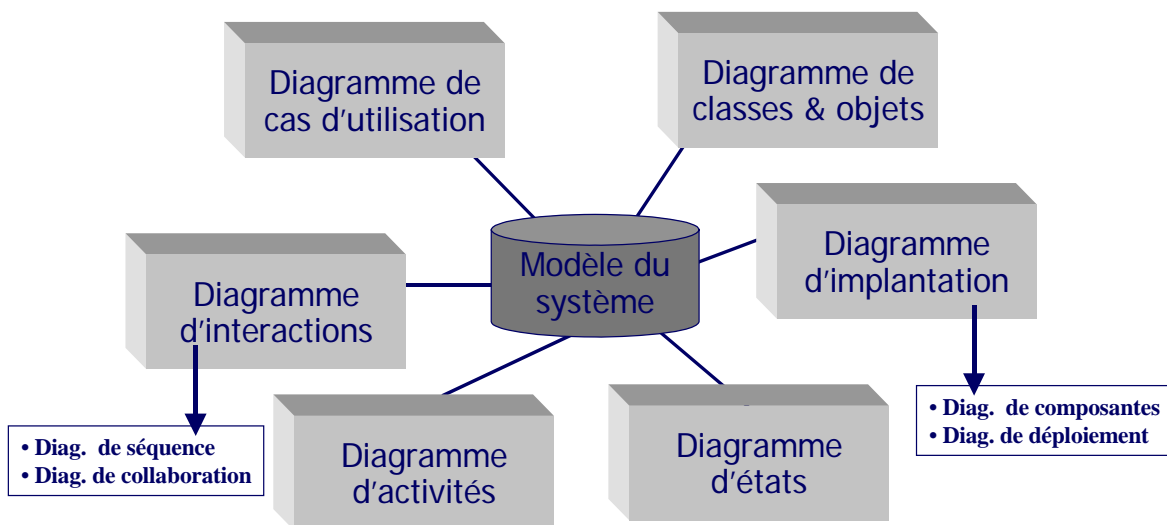


Figure 1 : Diagrammes UML

(Description du composant d20)

Les *diagrammes de cas d'utilisation* sont utilisés pour modéliser les besoins d'un point de vue fonctionnel selon la perspective de l'utilisateur.

La représentation statique du système s'appuie sur les quatre diagrammes suivants :

- le *diagramme de classes*, qui représente un des concepts les plus importants en programmation orientée objets. Il sert à illustrer la structure des classes d'objets et le *diagramme d'objets* qui sert à modéliser des instances de classes;
- les diagrammes d'implantation, qui sont constitués par : le *diagramme de déploiement* qui sert à illustrer le support informatique (serveurs, central, etc.) sur lequel le système sera installé, et le *diagramme de composantes*, qui sert à illustrer la structure des composantes physiques, comme les bibliothèques, les instances de bases de données, les exécutables, etc.

La représentation dynamique du système s'appuie pour sa part sur les quatre diagrammes suivants :

- le *diagramme d'états*, qui représente les différents états par lesquels passe un objet durant son cycle de vie;
- le *diagramme d'activités*, qui représente les enchaînements des activités dans le système;
- les diagrammes d'interactions, qui sont constitués du *diagramme de collaboration* et du *diagramme de séquence*. Tous deux servent à représenter les échanges de messages entre les objets. Le diagramme de séquence met l'accent sur l'aspect temporel, alors que le diagramme de collaboration met l'accent sur l'aspect structurel.

Utilisé par l'Object Management Group depuis 1997, UML est devenu un outil de modélisation et de communication incontournable dans les projets de développement informatique. En fait, outre l'aspect lié à la standardisation, il existe de nombreux avantages à utiliser UML pour la modélisation et la spécification des systèmes :

- UML est un langage de modélisation flexible, sémantiquement riche, pouvant s'appliquer à différents types (et tailles) de systèmes, de domaines et de méthodes;
- UML favorise le développement par composantes;
- UML solutionne les problèmes d'aujourd'hui (distribution, concurrence, capacité d'extension, etc.);
- UML est supporté par de nombreux outils de modélisation disponibles sur le marché;
- UML n'est pas une approche propriétaire et fermée, mais plutôt un langage ouvert, extensible, largement adopté par l'industrie;
- les spécifications complètes du langage et les évolutions en cours (nouvelles versions, etc.) sont accessibles sur le site de l'[Object Management Group](#);
- l'évolution et la maintenance des systèmes sont plus faciles et moins coûteuses parce que ces derniers sont mieux spécifiés et mieux documentés;
- le fait d'avoir un seul langage de modélisation facilite grandement la compréhension et les échanges entre les divers intervenants impliqués dans le développement de systèmes;
- plusieurs nouvelles approches de développement visant l'utilisation d'UML comme support de modélisation et ayant comme objectifs la réutilisation et le développement par composantes sont actuellement disponibles sur le marché.

Orientation

Le langage de modélisation UML doit être retenu par les ministères et organismes (M/O).

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d20	Langage de modélisation	UML	Retenir	http://www.omg.org/uml/	<ul style="list-style-type: none"> • UML, le langage de modélisation utilisé par l'Object Management Group (OMG) depuis 1997, est devenu un outil de modélisation et de communication incontournable pour les projets de développement informatique. • Conformément aux critères d'évaluation du groupe Gartner (voir l'article 117247, <i>Evaluating OOA&D Functionality Criteria</i>), UML demeure le standard en regard à la notation utilisée dans les diagrammes pour la modélisation. • Il y a de nombreux avantages à utiliser UML pour la modélisation et la spécification des systèmes (langage sémantiquement très riche et flexible, favorise le développement par composants, etc.). • UML n'est pas une approche propriétaire et fermée, mais plutôt un langage ouvert, extensible, largement adopté par l'industrie.

1.2 Réutilisation et partage des données et de leurs descriptions au gouvernement du Québec

Contexte

Afin de favoriser l'interopérabilité des systèmes entre les différents M/O, il est nécessaire de définir et de partager des descriptions communes pour les objets réutilisables au sein du gouvernement. Loin d'être nouvelle, cette préoccupation est mieux prise en compte par les outils de modélisation. Certains objets peuvent être réutilisables de manière dynamique (notamment les composants logiciels); ils sont dans ce cas dits partageables*.

Au préalable donc, la démarche passe, pour les objets existants, par un état des lieux des formats communs de description et d'échange de données par domaine d'application (tourisme, géomatique, etc.) et, pour les objets nouvellement créés, par la recherche d'une portée et d'une définition communes.

Concernant les données et les métadonnées, le cas particulier de l'emploi d'identifiants pour les domaines énumératifs codifiés est abordé ci-après.

Par ailleurs, une part croissante des documents seront décrits à l'aide du langage XML dans les prochaines années.

Les métadonnées ainsi que les documents réutilisables (blancs de formulaires, bandeaux, etc.) doivent également être publiés dans le registre référentiel⁵, banque de données accessible en ligne qui permet de se partager les objets et leurs définitions.

Concernant les traitements constituant la logique d'affaires, la même préoccupation doit exister chez les développeurs.

Le registre référentiel est présenté dans le chapitre 2.

1.3 Cas des domaines énumératifs codifiés

Contexte

Un domaine énumératif codifié^{6,*} a pour but de fournir une description et une structure communes partageables d'un domaine de valeurs utilisées par plus d'un M/O.

* C'est le cas par exemple des services Web. En effet, rendre un composant partageable (donc « invocable dynamiquement ») plutôt que réutilisable est un choix de l'ordre de la bonne pratique qui revient au concepteur et ne relève pas du présent cadre. Ce choix devrait prendre en compte les critères de sécurité, dont en particulier la disponibilité. De la même façon, la réutilisabilité pose le problème de la gestion des différentes versions d'un même composant. Ces pratiques sont cadrées par le document *Gestion des composantes réutilisables* dans le cadre de l'architecture d'entreprise gouvernementale.

⁵ Voir les définitions de *registre* et de *référentiel* en annexe.

⁶ Voir la définition de *domaine énumératif* et celle de *domaine énumératif codifié* en annexe.

• Au besoin, consulter le document intitulé *Métadonnées : lignes directrices pour décrire un domaine énumératif codifié de valeurs permises* sur le portail de l'infrastructure gouvernementale (http://www.gouvernement-en-ligne.qc/Normalisation_Ri/documents/Lignes-directrices-domaines-codifiés.pdf)

Un domaine énumératif codifié est constitué d'un ensemble d'identifiants⁷, et dans certains cas de codets⁸, qui désignent les valeurs permises à l'intérieur d'un domaine de valeurs de manière à rationaliser et à favoriser l'intégration des données à l'occasion d'échanges de données tant au sein de l'Administration québécoise qu'avec l'extérieur.

Les identifiants permettent de :

- favoriser l'interopérabilité par le partage des descriptions;
- assurer la permanence de l'identification;
- pour certains domaines, atteindre une reconnaissance internationale lorsque l'identification s'appuie sur des normes internationales (ex. : codes de pays selon la norme ISO 3166-1) ou des standards;
- réutiliser le contenu des domaines énumératifs codifiés en le rendant disponible et repérable pour les développeurs;
- faire évoluer la liste placée sous la responsabilité d'une autorité source;
- atteindre la consistance des résultats utilisant ces identifiants (par exemple, lors des recherches statistiques).

Orientations

Chaque domaine énumératif codifié est documenté dans un standard du gouvernement du Québec pour les ressources informationnelles (SGQRI).

L'autorité source⁹, ministère ou organisme responsable du domaine (ex. : l'Institut de la statistique du Québec a la responsabilité du domaine « Régions administratives du Québec »), a la responsabilité de déposer le contenu de ce domaine dans le registre référentiel du gouvernement du Québec.

L'autorité source a la responsabilité de l'évolution du domaine énumératif codifié et de mettre à jour le contenu disponible dans le registre référentiel.

À titre indicatif, voici une liste des standards qui s'appliquent en la matière : codification des noms de pays (SGQRI 46-1), des provinces et des territoires du Canada (SGQRI 46-2), des États des États-Unis (SGQRI 46-3), des langues (SGQRI 46-4), des monnaies et des types de fonds (SGQRI 46-5), des ministères et des organismes du gouvernement du Québec (SGQRI 46-6), des centres de responsabilité au sein du gouvernement du Québec (SGQRI 46-7), des régions administratives du Québec selon l'Institut de la statistique du Québec (SGQRI 46-8).

D'autres standards doivent encore être rédigés afin de couvrir les ressources financières, les ressources humaines, la description d'une entreprise, etc.

En fait, ce sont les projets informatiques qui ont besoin de définir des données et des descriptions, que le caractère réutilisable rend éligibles à la normalisation. Une fois normalisées, les descriptions peuvent en retour alimenter tous les projets qui en ont besoin.

⁷ Voir la définition du terme *identifiant* en annexe.

⁸ Voir la définition du terme *codet* en annexe.

⁹ Voir la définition d'*autorité source* en annexe.

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d25	Domaines énumératifs codifiés	Voir les SGQRI concernés	Retenir	Communiquer avec Yves Hudon, Yves.Hudon@msg.gouv.qc.ca	<ul style="list-style-type: none"> • Chaque domaine énumératif codifié est documenté dans un standard du gouvernement du Québec pour les ressources informationnelles (SGQRI). • L'autorité source, à savoir le ministère ou l'organisme responsable du domaine, a la charge de déposer le contenu de ce domaine dans le registre référentiel du gouvernement du Québec. • Certains standards ont été développés, alors que d'autres, notamment ceux concernant les ressources financières, les ressources humaines et la description des entreprises, doivent être rédigés.

1.4 Langage de description structurée de données (XML)

Contexte

Depuis de nombreuses années, les documents électroniques ont fait l'objet d'études qui ont conduit à la classification de leurs caractéristiques selon quatre dimensions :

- la dimension logique (par exemple : agencement structuré d'un document par paragraphes, titres, chapitres, etc.);
- la dimension spatiale (par exemple : disposition relative sur une page des différents éléments la constituant, polices, couleurs, images, etc.);
- la dimension hypertextuelle (par exemple : liens hypertextes vers lesquels pointe un document-pivot);
- la dimension temporelle (par exemple : rythme auquel se succèdent les différentes diapositives d'une même présentation).

Héritier du langage à balises SGML (*Standard Generalized Markup Language*, <http://xml.coverpages.org/sgml.html>), le langage XML (*eXtended Markup Language*) est un résultat majeur de ces études qui permet de décrire de manière structurée des documents, indépendamment de leur dimension spatiale ou temporelle.

Pour répondre aux besoins du Web que HTML ne pouvait plus soutenir, XML offre une syntaxe plus simple que SGML. Il permet d'associer au document des règles de sémantique ou de syntaxe couvrant la dimension logique, grâce aux espaces nominatifs et aux schémas XML. Les schémas ainsi que les espaces de nominatifs sont traités ci-après.

La dimension spatiale est couverte dans XML par l'association au document XML de feuilles de style et d'autres règles. Cet aspect est traité dans le volet 2 du cadre commun d'interopérabilité.

La dimension hypertextuelle est couverte grâce à Xlink, qui permet l'inclusion de liens hypertextes dans un document XML.

Enfin, la dimension temporelle est couverte par un ensemble de langages modulaires tels que VoiceXML et SMIL, traités dans le volet 2 du cadre commun d'interopérabilité.

Orientations

Le langage XML est la pierre angulaire pour les années à venir de l'interopérabilité des systèmes et de la réutilisation des documents.

Les M/O doivent retenir le standard XML. En fait, ils devraient considérer l'utilisation des modèles XML applicables à leurs domaines d'affaires. Ces modèles permettent la standardisation, la réutilisation et l'amélioration du temps de réponse. Pour plus d'information, se référer à l'article 117247, *Evaluating OOA&D Functionality Criteria*, publié par le groupe Gartner.

No	Composant	Norme ou standard	Orientations	Référence	Commentaires
d30	Langage à balise	XML	Retenir	Recommandation du W3C http://www.w3.org/TR/REC-xml/	<ul style="list-style-type: none"> • XML permet de stocker dans un document des données structurées. • Les données sont indépendantes de l'affichage. Ainsi, à partir d'un seul fichier XML, on pourra créer divers documents. • XML permet d'utiliser le potentiel de toute la famille XML : schémas, XSLT et XHTML, dont ses langages modulaires spécialisés (SVG, XML SIG, etc.). • XML permet d'améliorer l'interopérabilité et l'échange de données entre les M/O, comme avec les citoyens et les partenaires.

1.5 Formats des identifiants de ressources

Contexte

Le format utilisé pour identifier une ressource est unique dans le Web. Il s'agit de l'URI (*Uniform Resource Identifier*). Un URI peut identifier une ressource ou ses méthodes.

Les URI peuvent être des URL (*Uniform Resource Locator*), terme populaire mais sans définition formelle, ou des URN (*Uniform Resource Names*), ou un identificateur qui n'est ni un URL ni un URN, ce qu'illustre le schéma suivant.

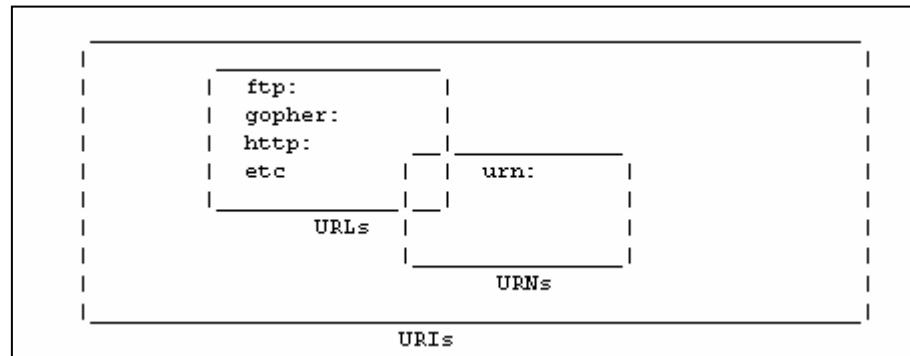


Figure 2 : Inclusion des ensembles d'identifiants de ressources

(composant d40) - Source : W3C

Pour le cas particulier des URL du gouvernement du Québec, les standards SGQRI documentent le format que les M/O doivent respecter.

Les URN forment un sous-ensemble des URI (une particularisation) et désignent les noms des ressources *persistants et indépendants de la localisation*.

Cette initiative vise à permettre la persistance d'une référence, afin de contrecarrer progressivement la mobilité imprévisible des liens sur Internet. Les URN devraient être utilisés pour identifier toute ressource permanente, comme les espaces nominatifs, par exemple.

Mais les URN nécessitent la mise en ligne de serveurs d'URN et la prise en compte des URN dans les fureteurs.

Dans l'attente de la généralisation des URN, les PURL (*Persistent Uniform Resource Locator*) peuvent être utilisés en lieu et place des URN. Le PURL pointe sur le service de redirection qui, lui, retourne l'URL correct à l'utilisateur.

L'information sera fiable si un administrateur est chargé de mettre à jour le service de redirection quand l'URL change effectivement. En d'autres termes, la persistance résulte d'un processus d'organisation et non d'une astuce technologique.

La mise en ligne d'un serveur de redirection d'URL par le gouvernement du Québec est à prendre en compte; cela permettrait la définition de PURL au gouvernement du Québec.

Les PURL représentent une transition douce vers les URN. Ainsi, la Metadata Dublin Core Initiative a fait appel à des PURL (par exemple <<http://purl.org/dc/elements/1.1/>>) pour désigner leurs espaces nominatifs par des identificateurs persistants : ces identificateurs sont utilisés en effet par tous les documents XML intégrant le profil de métadonnées du Dublin Core (voir <<http://dublincore.org/documents/2001/10/26/dcmi-namespace/>>).

Orientations

Il est recommandé aux M/O de se conformer au standard du W3C et de l'Internet Engineering Task Force (IETF) en matière d'URI et d'URL, et d'utiliser les produits qui s'y conforment. Les M/O doivent se reporter aux standards SGQRI en ce qui concerne le cas particulier des URL d'adresses de noms de domaine et de courriels.

Pour assurer ce service, la mise en ligne d'un serveur de redirection d'URL par le gouvernement du Québec est à considérer.

Remarque importante

Concernant le soutien de la langue française et de l'internationalisation en général, les M/O doivent se reporter au volet 4 du Cadre commun d'interopérabilité.

No	Composant	Norme ou standard	Orientations	Références	Commentaires
d40	Format des identifiants de ressources et de noms	URI (dont les URN)	Retenir	RFC 2483 RFC 1736, 1737 RFC 2392 RFC 2396 RFC 2255 RFC 2056 RFC 2369 RFC 2384 RFC 2717, 2718 RFC 2732 http://ftp.ics.uci.edu/pub/ietf/uri/ http://www.w3.org/Addressing/	<ul style="list-style-type: none"> Les M/O doivent se reporter aux standards SGQRI en ce qui concerne le cas particulier des adresses de noms de domaine (SGQRI 44) et de courriels (SGQRI 21). URI (<i>Uniform Resource Identifier</i>) est le format utilisé pour désigner une ressource. Un URI peut désigner une ressource ou ses méthodes. Les URN forment un sous-ensemble des URI (une particularisation) et désignent les noms des ressources <i>persistants</i> et <i>indépendants de la localisation</i>. En attendant la généralisation des URN, les M/O peuvent utiliser les redirections de PURL quand ils ont besoin de persistance. Concernant le soutien de la langue française et de l'internationalisation en général, les M/O doivent se reporter au volet 4 du Cadre commun d'interopérabilité.

No	Composant	Norme ou standard	Orientation	Références	Commentaires
d41	Redirection d'un identificateur	PURL	Retenir	http://purl.org/	<ul style="list-style-type: none"> • Dans l'attente des URN, les M/O peuvent enregistrer un PURL pour tout identificateur qui nécessite une persistance (exemple : URI d'un registre référentiel, d'un service applicatif, etc.). • Au lieu de pointer directement sur la localisation de la ressource, le PURL pointe sur le service de redirection, qui retourne l'URL correct à l'utilisateur. • Voir la foire aux questions <http://purl.oclc.org/docs/purl_faq.html>. • La mise en ligne d'un serveur de redirection d'URL par le gouvernement du Québec est à prendre en compte.

1.6 de description utilisé dans les modèles de documents et dans les documents XML

Contexte

Dans un document XML, le vocabulaire de description utilisé à l'intérieur des balises XML contient des éléments et des attributs qui sont définis pour être utilisés par des applications modulaires différentes.

Mais un même document ou ensemble de documents XML utilise fréquemment des vocabulaires issus de domaines différents et utilisés par diverses applications, ce qui peut *entraîner des collisions au moment de la validation* d'un document.

Les *espaces nominatifs* définis par le W3C permettent que des documents XML provenant de sources différentes soient combinés sans ambiguïtés.

Par exemple, l'attribut « *titre* » peut désigner le titre d'un livre et, dans le même document XML, être un titre de transport ou encore un titre d'ingénieur. Or, le schéma, lors de la validation, ne saura quel type utiliser pour valider le terme « *titre* ».

Pourtant, ce vocabulaire, situé dans un contexte, doit être commun afin d'encourager sa réutilisation.

Les espaces nominatifs permettent de lever toute ambiguïté en qualifiant chaque terme utilisé et ce, en le rapportant à un espace donné (en lui donnant un contexte, en quelque sorte).

Pour ce faire, les espaces nominatifs sont référencés dans le document XML par un URI et factorisés par un préfixe plus facile d'utilisation (dans notre exemple, <http://www.rr.gouv.qc.ca/.../CCI/volet> est factorisé par le préfixe *volet*). Ce préfixe vient qualifier chaque élément et attribut ambigu du document et lève toute ambiguïté.

Ainsi, l'élément ou l'attribut XML « *titre* », préfixé par l'espace nominatif auquel il se rapporte, répond à la définition du domaine énumératif codifié donnée plus haut et garantit l'interopérabilité.

Le schéma suivant illustre de manière fictive la collision qui pourrait survenir avec le terme «*titre*» du document XML *Cadre commun d'interopérabilité*. Ce terme *titre* désigne aussi bien le titre du document que le titre d'un volet du document. Pour lever l'ambiguïté, les deux termes sont rapportés aux espaces nominatifs <http://www.rr.gouv.qc.ca/.../CCI> et <http://www.rr.gouv.qc.ca/.../CCI/volet>.

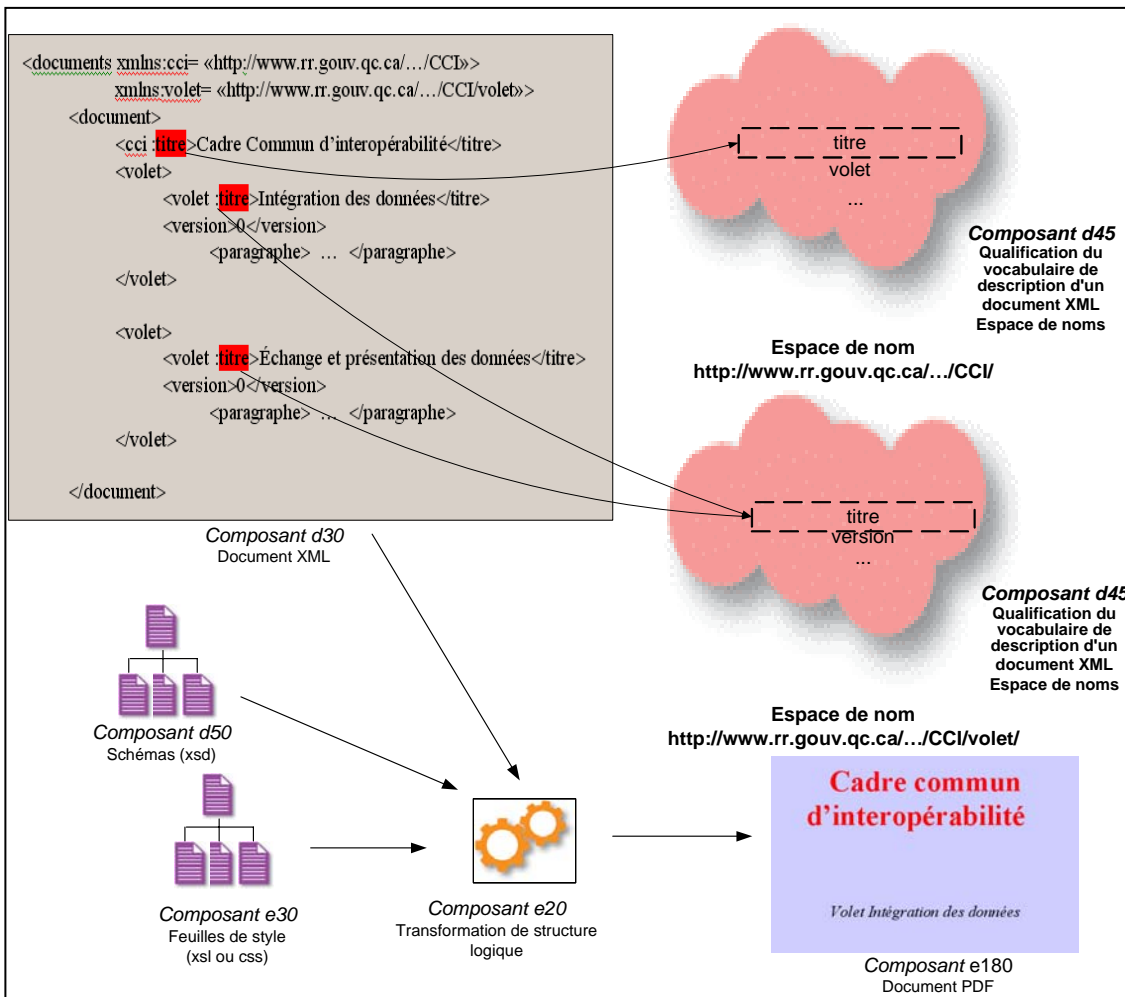


Figure 3 : Exemple fictif de qualification du terme titre grâce aux espaces nominatifs

(mise en situation du composant d45)

Orientations

Les M/O doivent utiliser les espaces nominatifs XML. Les espaces nominatifs composés par les M/O doivent référer aux espaces nominatifs du W3C.

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d45	Qualification du vocabulaire de description d'un document XML	Espaces de noms XML	Retenir	Recommandation du W3C http://www.w3.org/TR/REC-xml-names/	<ul style="list-style-type: none"> • Un même document XML utilise fréquemment des vocabulaires issus de domaines différents et utilisés par des applications différentes, ce qui peut entraîner des collisions au moment de la validation du document par le schéma. • Les espaces nominatifs définis par le W3C ont pour rôle de lever toute ambiguïté en qualifiant chaque terme utilisé (en lui donnant un contexte, en quelque sorte). • Les espaces de nominatifs sont référencés dans le document XML par un URI et factorisés par un préfixe plus facile d'utilisation. Ce préfixe vient qualifier chaque élément et attribut ambigu du document. • Les M/O doivent utiliser les espaces de nominatifs XML.

1.7 Définition et publication des modèles de documents

Contexte

La famille des langages XML s'appuie, pour définir les *règles* (modèles et restrictions) s'appliquant à un document XML, sur les schémas XML ou sur les définitions de type de document (DTD).

Cependant, il est à noter que les définitions de type de document (DTD), qui étaient utilisées antérieurement avec SGML, ont une syntaxe plus compliquée que ne l'est celle des schémas XML.

Le langage de schéma XML répond au besoin d'unifier le vocabulaire relatif aux schémas afin de mieux les définir, les décrire et les cataloguer en vue de leur réutilisation. La partie 2 du langage de schéma XML est consacrée aux *types de données*. Les influences principales sur sa conception sont ISO 11404, SQL (*Structured Query Language*) et Java, en plus du W3C.

Pour les développeurs, le langage de schéma XML est celui qui a la meilleure capacité de description et de validation de la structure et du contenu d'un document XML, tant au niveau général du document que pour son élément le plus fin.

Les schémas définissent ainsi tous les types de données à partir d'un certain nombre de types prédéfinis (types simples, types complexes, listes, restriction sur des types) qui peuvent être **étendus** par combinaison de types, héritage et modularité. Ils déterminent également le nombre d'occurrences des éléments, gèrent les contenus mixtes, les éléments nommés et les groupes d'attributs. Le type « annotation » permet l'inclusion de commentaires par le développeur.

Un document conforme à la syntaxe XML est certes déjà structuré, « bien formé », mais le ou les schémas XML qui s'y appliquent au document permettent en outre de valider la structure générale et les éléments du document en conformité avec les règles contenues dans le ou les schémas. Le document XML est lié à un ou plusieurs schémas XML en y faisant référence dans son en-tête. La validation est faite côté serveur ou côté client par un analyseur XML contenu dans l'API (*Application Program Interface*)^{*}, voire par l'application manipulant le document.

Il est nécessaire que le gouvernement du Québec publie ses schémas XML afin que les différents M/O, en tout premier lieu, mais aussi les entreprises et les citoyens, partagent les mêmes définitions d'un document quand il y a lieu de le faire.

Le schéma XML est un document d'extension .xsd à syntaxe XML qui peut être stocké dans une base de données à l'usage de tous.

Pour les utilisateurs, la validation de la conformité à un schéma ou à une définition de type de document devrait se faire de façon transparente dans les suites bureautiques et les outils d'édition.

Orientations

Les schémas XML doivent être réutilisés pour les développements ou les mises à jour d'applications par les M/O au fur et à mesure de leur publication dans le registre référentiel du gouvernement.

Si aucun schéma gouvernemental ne répond au besoin, les M/O rechercheront l'utilisation de modèles standards existant en dehors du gouvernement ou, à défaut, en définiront de nouveaux susceptibles d'être étendus et réutilisables.

^{*} Voir le chapitre « Interface standardisée pour la manipulation des documents XML par les programmes ».

No.	Composant	Norme ou standard	Orientation	Références	Commentaires
d50	Définition et validation de structures logiques	Langage de schéma XML	Retenir	http://www.w3.org/TR/xmlschema-0/ http://www.w3.org/TR/xmlschema-1/ http://www.w3.org/TR/xmlschema-2/	<p><u>Pour les développeurs</u></p> <ul style="list-style-type: none"> Les schémas définissent tous les types de données à partir d'un certain nombre de types prédéfinis (types simples, types complexes, listes, restriction sur des types), lesquels peuvent être étendus par combinaison de types, héritage et modularité. Un document conforme à la syntaxe XML est « bien formé », mais le schéma XML permet en outre de valider les éléments du document en conformité avec les règles définies par le ou les schémas qui s'y appliquent. La validation est faite par un analyseur XML contenu dans l'API de l'application manipulant le document, côté serveur ou côté client. Il est nécessaire que le gouvernement du Québec publie ses schémas XML afin que les différents M/O, en tout premier lieu, mais aussi les entreprises et les citoyens, partagent les mêmes définitions de documents et d'éléments de données. Le schéma XML est un document d'extension .xsd à syntaxe XML qui peut être stocké dans une base de données en vue d'être utilisé. Cette base de données, à l'usage de tous, est le registre référentiel du gouvernement du Québec. Le schéma XML est à préférer aux descriptions de type de document (DTD). <p><u>Pour les utilisateurs</u></p> <ul style="list-style-type: none"> La validation de la conformité à un schéma ou à une définition de type de document devrait se faire de façon transparente dans les suites bureautiques et les outils d'édition.

No.	Composant	Norme ou standard	Orientation	Références	Commentaires
		DTD Définition de type de document	Retenir		<ul style="list-style-type: none"> • La définition de type de document (DTD) n'a pas une syntaxe XML, mais peut être convertie automatiquement en schéma XML. • Les schémas XML vont progressivement remplacer les DTD.
		Relax NG	Éligible dans une version ultérieure du cadre	http://www.oasis-open.org/home/index.php	<ul style="list-style-type: none"> • Relax NG vise à simplifier le schéma XML. Il est plus efficace dans le traitement des messages XML (pas de perte d'information).

2 Documents et processus d'affaires dans l'Administration québécoise

Le formalisme de modélisation UML et la famille XML, les documents et processus d'affaires du gouvernement du Québec peuvent être progressivement informatisés. Pour cela, il convient de *classifier* les *données* et les *processus d'affaires* selon les *métadonnées* prédéfinies (la *sémantique*), de spécifier et de déployer les *solutions technologiques* permettant de réaliser le registre référentiel (bases de données) et la logique applicative (scripts, composants serveurs, composants élémentaires), le tout en garantissant l'interopérabilité.

2.1 Classification des données dans l'Administration québécoise

Définition des métadonnées

Contexte

Les affaires électroniques se font avec des objets devant être conformes aux structures logiques définies, qui sont principalement des documents.

L'enregistrement des données et des documents dans le registre référentiel doit être automatisé et doit comporter une inscription de renseignements conformes à un profil approprié de métadonnées.

Ce profil doit répondre aux finalités du repérage de l'information ainsi qu'à celles de la gestion et de la tenue de dossiers, ainsi qu'à la conservation à des fins légales. Le Cadre de référence gouvernemental en gestion intégrée des documents (CRGGID) est au centre des pratiques recommandées, en particulier le document *Profils de métadonnées pour les documents gouvernementaux* dans sa version du 10 mars 2003. Le CRGGID vise à définir et à rendre disponibles des modèles concernant des points essentiels pour la mise en œuvre d'une meilleure gestion intégrée des documents dans les ministères et organismes.

Le contenu du standard Dublin Core (voir

<<http://www-rocq.inria.fr/~vercoust/METADATA/DC-fr.1.1.html>> pour la traduction française) consiste en un ensemble de quinze métadonnées : « titre », « créateur », « sujet », « description », « éditeur », « contributeur », « date », « type », « format », « identifiant », « source », « langue », « relation », « couverture » et « droits ».

Du point de vue du langage XML, le Dublin Core n'est ni plus ni moins qu'un espace nominatif particulier contenant les quinze éléments définis par des schémas.

Pour sa part, le document *Profils de métadonnées pour les documents gouvernementaux* dans sa version du 10 mars 2003 définit pour le gouvernement du Québec un profil de 25 éléments.

Orientations

Les M/O doivent retenir pour la gestion documentaire la norme ISO 15489. Ils doivent également définir leurs profils ministériels de métadonnées en se basant sur les profils publiés dans le document *Profils de métadonnées pour les documents gouvernementaux* dans sa version du 10 mars 2003.

No.	Composant	Norme ou standard	Orientations	Références	Commentaires
d60	Principe de gestion documentaire	ISO 15489 CRGGID	Retenir	Norme ISO 15489	<ul style="list-style-type: none"> Normalisé
d70	Profil de métadonnées utilisés pour la description des objets	Profils de métadonnées du gouvernement du Québec (conforme au Dublin Core) CRGGID	Retenir	<p>Document <i>Profils de métadonnées pour les documents gouvernementaux</i> dans sa version du 10 mars 2003</p> <p>Contact : Richard Parent, richard.parent@msg.gouv.qc.ca</p> <p>Dublin Core : http://dublincore.org/documents/</p>	<ul style="list-style-type: none"> Ensemble de 25 éléments définis pour les documents de transaction et les documents de référence,. Les schémas et espaces nominatifs du registre référentiel du gouvernement du Québec instrumentent les profils de métadonnées. Les espaces nomsinatif du Dublin Core peuvent être référencés dans tout document grâce à l'URN persistant http://dublincore.org/documents/1999/07/02/dces/. Une description de type de document (DTD) ou, de préférence, des schémas XML définissent les profils de métadonnées du Dublin Core (voir http://dublincore.org/documents/2002/07/31/dcmes-xml/) Un modèle de métadonnées est essentiel dans les systèmes de gestion documentaire (voir l'étude 113711, <i>Records Management Needs Metadata and XML</i>, du groupe Gartner).

Sémantique gouvernementale de classification

Contexte

Le profil de métadonnées (titre, créateur, date, etc.) visé par le composant d70 est destiné à une utilisation générale au gouvernement du Québec pour identifier et décrire les documents ou d'autres ressources informationnelles. Le profil est constitué d'une liste d'attributs obligatoires et facultatifs parmi lesquels se trouvent des schèmes de classification selon quatre facettes génériques. Le but visé est que ce classement sémantique soit organisé sur un double plan : celui des perceptions sociales et celui des exigences fonctionnelles :

- La classification par domaine/objet, formant une arborescence dont le plus haut niveau correspond aux champs d'intervention essentiels du gouvernement (santé et services sociaux, économie et environnement, gouvernance et justice, etc.);
- La classification par processus (déclarer conforme, doter, etc.), décomposables eux-mêmes en activités;
- La classification par activités comprises dans la séquence d'un processus;
- La classification par types de documents, comprenant une appartenance de type général (exemple : formulaire) et l'identifiant de structure logique du document (exemple : formulaire numéro 2).

L'objectif important est ici d'élaborer une sémantique générale de classification dans l'Administration québécoise et en appuyer la documentation dans un thésaurus de l'activité gouvernementale et une sélection de schèmes de classification. Pour plus d'explications, voir les documents à l'adresse <http://www.banq.qc.ca/portal/dt/services/archivistique_ged/crggid/crggid_outils/profil/profil.jsp>.

Orientations

Les schèmes sont en cours d'élaboration dans le cadre du développement du Cadre de référence gouvernemental en gestion intégrée des documents. Une préoccupation particulière est d'en arriver à établir des règles de sélection et d'extension pour que chaque organisation procède aux adaptations dont elle a besoin.

Par exemple, l'équipe de gestion documentaire pourra adapter ces schèmes aux besoins ministériels de classification, y compris les schèmes associés à des vocabulaires spécialisés par domaine ou par discipline. Un ministère pourra choisir les parties qui le concernent et les adapter à sa propre réalité. La simplification pour l'utilisateur passera donc par une sélection des éléments formant la liste d'options d'un schème, de façon à en réduire le nombre à ce qui est le plus pertinent dans une subdivision organisationnelle, pour un rôle, etc. Le schème des processus peut devenir important dans une architecture d'entreprise gouvernementale (AEG) : si un ministère veut modéliser un de ses processus particuliers et s'il réutilise pour cela des modèles génériques en les adaptant de façon assistée à son contexte. Dans un cadre comme ebXML, c'est quand un processus générique est adapté pour un contexte de domaine/objet particulier qu'il complète sa définition et prend tout son sens.

No.	Composant	Norme ou standard	Orientation	Référence	Commentaires
d80	Schèmes de classification pour la sémantique gouvernementale du Québec	Domaines /objets d'affaires, processus d'affaires, activités, types de document	Retenir	Les schèmes, en cours d'élaboration, peuvent être consultés à http://www.anq.gouv.qc.ca/thesaurus/index.htm	<ul style="list-style-type: none"> Chaque élément des schèmes de classification est documenté dans le thésaurus de l'activité gouvernementale accessible au même endroit.

Standards de description pour les métadonnées

Contexte

La technologie RDF (*Resource Description Framework*) permet de mettre en œuvre technologiquement la sémantique gouvernementale. RDF, en standardisant le format des métadonnées, permet leur échange aisé avec d'autres moteurs.

Outre la recherche d'information et de ressources par les employés de l'Administration ou les citoyens et partenaires (au travers, par exemple, d'un portail gouvernemental pour ces derniers), les applications de RDF peuvent être l'intégration de sources d'information hétérogènes, la gestion de connaissances, la formation électronique, etc.

RDF spécifie un modèle permettant de décrire les métadonnées rattachées à un document et par conséquent d'échanger ce type d'information avec d'autres parties (M/O, partenaires, clients, etc.). Il permet donc de maximiser la gestion des métadonnées dans un contexte gouvernemental.

RDF utilise la syntaxe XML (espaces nominatifs, schémas, etc.) et est applicable à toute page et partie de page Web ainsi qu'à tout objet répondant au schéma d'un URI. Il permet de relier les descriptions à leurs objets par les triplets sujet-propriété-objet ou ressource-propriété-valeur.

Ainsi, RDF peut intégrer les profils de métadonnées définis dans le recueil des éléments normatifs (à venir) du Secrétariat du Conseil du trésor au moyen d'espaces nominatifs définis en réponse aux profils. Il peut en particulier intégrer les métadonnées du Dublin Core en s'appuyant sur l'espace nominatif du Dublin Core (<http://www-rocq.inria.fr/~vercoust/METADATA/DC-fr.1.1.html>) et en utilisant les métadonnées associées (« *title* », « *description* », « *publisher* », « *date* », etc.).

RDF n'inclut pas de capacité de raisonnement à partir des relations construites.

Orientation

Cette technologie est actuellement très peu déployée, et il existe très peu d'expertise sur le marché. Le langage RDF est donc à prendre en compte au gouvernement du Québec pour la description de métadonnées.

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d90	Standard de description des métadonnées	RDF	Prendre en compte	Recommandation du W3C http://www.w3.org/RDF/	<ul style="list-style-type: none"> • RDF (<i>Resource Description Framework</i>) est un standard de description de métadonnées qui utilise la syntaxe XML (espaces nominatifs, schémas). • Il est applicable à toute page et partie de page Web, à tout objet répondant au schéma d'un URI. • Il relie les descriptions aux objets par les triplets sujet-propriété-objet ou ressource-propriété-valeur. • Il peut intégrer les métadonnées du Dublin Core (par exemple) en s'appuyant sur l'espace nominatif. • Cette technologie est actuellement très peu déployée, et il existe très peu d'expertise sur le marché.

2.2 Accessibilité et manipulation des documents XML

Accès aux données et aux métadonnées

Contexte

Dans les systèmes actuels, les données constituant le patrimoine informationnel du gouvernement du Québec et alimentant les processus d'affaires sont stockées dans des bases de données.

Les données alimentant les processus sont dans leur grande majorité des données à fine granularité (correspondant à des types élémentaires), et les bases de données les stockant sont à ce titre des bases « orientées données ».

Avec la dématérialisation des procédures et des documents les alimentant, le besoin des concepteurs est de manipuler également des documents et de les stocker. De plus, ces documents peuvent revêtir un caractère légal, auquel cas il est nécessaire de les archiver dans un format conforme à l'original. Ces documents utiliseront de plus en plus la syntaxe XML.

Par ailleurs, pour les systèmes à venir, le choix d'un type de base de données (orientée données ou orientée documents ?) doit être fait par le concepteur en fonction des besoins. Pour les déterminer, il doit notamment se poser les questions suivantes :

- à quels usages est destinée l'information (données ou documents) stockée ?
- avec quelle granularité, avec quelles contraintes légales de format, de conservation et de sécurité la stocke-t-on ?
- Selon quels critères de performance doivent se faire les accès ?

Selon les réponses à ces questions, la solution pour le concepteur peut être d'utiliser un produit mettant en œuvre une base de données orientée données ou documents ou de recourir à un produit mettant en œuvre les deux solutions.

Ce choix relève de la conception et ne remet pas en question l'exigence d'interopérabilité. On ne trouvera donc pas, dans le cadre commun d'interopérabilité, de standards retenus pour la base de données elle-même ni pour l'interface de programmation (API, Application Program Interface) qui permet la manipulation des données dans la base, d'autant plus que l'API est très souvent en partie propriétaire et étroitement liée au schéma de la base de données.

Ce qui relève de l'interopérabilité est par contre la faculté qu'ont les applications de différents M/O, utilisant des bases de données souvent différentes, à pouvoir collaborer entre elles, avec les applications des entreprises et avec le citoyen. Il s'agit donc de partager la définition et la description des données, ce à quoi s'attache la normalisation des domaines énumératifs codifiés (voir la section 1.2.1). Il faut également normaliser le format d'insertion et d'extraction en entrée /sortie de l'interface de programmation d'une base de données.

Orientations

Si le choix se porte sur une base de données orientée documents, le concepteur devra s'assurer que le format retenu pour l'extraction et l'insertion à l'entrée et à la sortie de l'API est XML, et ce, pour limiter les transformations successives faites par les applications sur les données et les métadonnées.

Le format de stockage des documents dans la base peut être différent, et la base peut être construite sur le modèle natif XML ou relationnel. Les deux modèles sont utilisés dans les produits de bases de données disponibles sur le marché.

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d100	Format d'insertion et d'extraction des documents d'une base de données orientée documents	XML	Retenir	Recommandation du W3C http://www.w3.org/TR/REC-xml/	<ul style="list-style-type: none"> Le concepteur doit s'assurer que le format retenu pour l'extraction et l'insertion à l'entrée et à la sortie de l'API est XML, et ce, pour limiter les transformations successives faites par les applications sur les données et les métadonnées. Les bases de données natives XML répondent entre autres à ce format : elles utilisent pour l'insertion, la mise à jour et l'extraction le modèle logique XML et stockent les documents au format XML.

Interface normalisée pour la manipulation des documents XML par les programmes

Contexte

Ressources normalisées pour les développeurs, les interfaces de programmation SAX (*Simple API for XML*) et DOM (*Document Object Model*) permettent aux applications et aux scripts de modifier de manière dynamique le contenu, le style ou la structure des documents XML, évitant ainsi aux développeurs de réécrire cette partie commune.

Plus précisément, SAX et DOM effectuent *l'analyse (parsing)* du document XML, c'est-à-dire sa lecture en vue de construire une représentation arborescente en mémoire, ainsi que l'opération inverse.

Pour cela, les deux API ont recours à des méthodes différentes.

SAX découpe le flux d'entrée en jetons et retourne des événements à chaque élément, attribut ou texte rencontré. SAX est donc efficace, car il ne peut renvoyer à l'application s'appuyant sur lui que les nœuds nécessaires; il permet donc de traiter un volume de données important avec peu de mémoire, côté client ou côté serveur. SAX laisse ainsi aux développeurs la responsabilité de créer leur propre structure objets et de gérer les événements qui sont retournés.

La version 2 de SAX prend en compte les espaces nominatifs XML.

DOM, standard du W3C, construit pour sa part une représentation arborescente complète du document et permet de la manipuler. Cet arbre, nommé le DOM, est constitué de nœuds de différents types : document, élément, attribut, commentaires, etc. Ce fonctionnement impose une mémoire vive côté serveur ou client adaptée aux volumes de données à traiter.

Le DOM de niveau 2 permet de prendre en compte les espaces nominatifs, et la spécification est découpée en plusieurs parties :

- *DOM views* permet de gérer les différentes vues d'un document;
- *DOM events* permet de gérer les événements sur les nœuds de l'arbre (clic de souris, etc.);

- *DOM style* permet de manipuler les arbres de style;
- *DOM traversal & range* permet de parcourir l'arbre.

Cette spécification est ici détaillée afin de mettre en lumière toute l'importance qu'a cette API pour les documents XML, XSL et les schémas. C'est elle qui, à partir de données en entrée (arbres de style, arbres de document, modèles et restrictions d'un ou de plusieurs schémas) ou d'événements (passage, clic de la souris, etc.), accomplit diverses opérations sur l'arborescence d'un document XML au travers de ses méthodes.

DOM est utilisé notamment par XSLT et Xpath (voir volet 2).

Il est à noter que les éditeurs étendent souvent les API DOM et SAX par des méthodes propriétaires pour les rendre plus efficace l'application qui les utilise. Cela ne les rend compatibles qu'entre langages et plates-formes du même éditeur.

Orientation

Les M/O doivent acquérir des outils de traitement de documents XML recourant aux API DOM de niveau 2 et SAX v2. Les développeurs des différents M/O doivent avoir recours à ces API pour développer leurs applications.

No.	Composant	Norme ou standard	Orientation	Références	Commentaires
d110	Interface de programmation pour manipuler les documents XML (sérialisation d'un document XML)	DOM niveau 2	Retenir	Recommandation du W3C http://www.w3.org/DOM/DOMTR	<ul style="list-style-type: none"> • DOM (<i>Document Object Model</i>), standard du W3C (http://www.w3.org/DOM/DOMTR) construit un arbre complet du document constitué de nœuds de différents types : document, élément, attribut, commentaires, etc. • À partir de données en entrée (arbres de style, arbres de document, modèles et restrictions d'un ou de plusieurs schémas) ou d'événements (passage, clic de la souris, etc.), DOM accomplit diverses opérations sur l'arborescence d'un document XML au travers de ses méthodes. • Ce fonctionnement impose une mémoire vive côté serveur ou côté client adaptée aux volumes de données à traiter. • Le DOM de niveau 2 permet de prendre en compte les espaces nominatifs; la version 3 (document de travail) permet à l'API de charger et de sauvegarder les arbres manipulés. • DOM est utilisé notamment par XSLT et Xpath (voir volet 2). • À noter que, dans les applications commerciales, l'API DOM est souvent étendue par des méthodes propriétaires. • Voir l'implémentation la version JDOM et une implémentation version récente, XOM, qui mélange les approches. • Les M/O doivent baser leurs composants et scripts sur DOM niveau 2 ou sur SAX v2.

No.	Composant	Norme ou standard	Orientation	Références	Commentaires
		SAX v2	Retenir	http://www.saxproject.org/	<ul style="list-style-type: none"> • API événementielle, SAX (<i>Simple API for XML</i>) découpe le flux d'entrée en jetons et retourne des événements à chaque élément, attribut ou texte rencontré. • SAX est efficace, car il ne peut renvoyer à l'application s'appuyant sur lui que les nœuds nécessaires; il permet donc de traiter un volume de données important avec peu de mémoire. • À l'origine, cette API n'était disponible qu'en JAVA. Actuellement, elle est disponible dans d'autres langages (voir <http://www.saxproject.org/?selecte d=lang>). • La version 2 de SAX permet de prendre en compte les espaces de noms XML. • Les M/O doivent baser leurs composants et scripts sur DOM niveau 2 ou sur SAX v2.

2.3 Une infrastructure standard pour le commerce électronique

Après une vingtaine d'années de faible croissance de l'échange de données informatisé (EDI-*Electronic Data Interchange*) en raison de sa complexité et des coûts élevés de sa mise en place, l'organisme des Nations Unies qui s'occupe des affaires, du commerce et de l'administration (UN/CEFACT (*United Nations Center for Trade Facilitation and Electronic Business*)) a lancé en 1999 un projet visant à fonder un nouvel échange de données informatisé à faible coût et à haute sécurité dans Internet et en recourant à des normes ouvertes uniquement. Le projet ebXML (*eb* pour *electronic business*) a été mené en commun avec l'OASIS (*Organization for the Advancement of Structured Information Standards*) et consiste à développer une architecture standard de normes dans le domaine du commerce électronique. Parce qu'il peut donner à chaque citoyen ou entreprise des mécanismes d'échange électronique sûrs, il peut s'avérer un puissant moyen de facilitation de la croissance de la prestation électronique de services (PES) grâce à la confiance qu'il permettra.

ebXML fournit un ensemble de spécifications, de pratiques et une bibliothèque de *composants en libre service* permettant de dématérialiser un nombre croissant de transactions commerciales entre organisations de façon standard, et pour cela d'échanger des messages, de définir et d'enregistrer des processus d'affaires ainsi que les documents de transaction qui les alimentent.

ebXML compte parmi ses réalisations concrètes la définition en composants élémentaires ebXML des entités d'information d'affaires UBL (*Universal Business Language*), autre initiative très importante du comité technique d'OASIS.

Description de l'infrastructure

Contexte

L'architecture technologique de ebXML est basée sur les concepts suivants:

- La messagerie sûre : l'approche générale de ebXML consiste à mettre sous enveloppe le contenu à transmettre d'une manière qui met à profit une couche XML capable de fonctionner avec tout protocole de communication, par exemple HTTP ou SMTP. Un message ebXML utilise l'habit d'une enveloppe de message MIME, et ce, d'une façon conforme à la spécification d'un message SOAP 1.1 avec des pièces jointes (SOAPAttach), le tout constituant un colis-message. Ce colis-message est donc au croisement des standards SOAP et MIME, ce qui lui permet de miser en même temps sur la structure XML de l'enveloppe SOAP, sur la souplesse et les mécanismes de sécurité de MIME ainsi que sur le large déploiement de ces protocoles dans les réseaux existants. Le protocole fournit une fonction d'opérateur de service de messagerie. Si l'on fait une analogie avec le service postal (enveloppes, timbres, adresses, classes de courrier, courrier recommandé, etc.), on peut dire que cet opérateur place tout objet de contenu dans une enveloppe avec une étiquette d'acheminement visible pour le transport et retire les objets de contenu des enveloppes reçues. Pour plus d'explications, voir

<<http://www.autoroute.gouv.qc.ca/publica/normes/norme124.htm>>.

- Le registre référentiel : ce standard ebXML décrit l'organisation du lieu de conservation pour tout ce qui sert aux échanges de commerce électronique. Il peut s'agir de schémas et de documents XML, de descriptions de processus, de composants élémentaires, de descriptions de contextes, de modèles UML, d'information sur les parties, de composants logiciels. L'hébergement de cette diversité d'information se fait grâce à un modèle qui aide à établir quelles métadonnées doivent être enregistrées, pour quels types d'objets. Ce mécanisme est conçu conformément à la norme ISO 11179 sur les registres de métadonnées. Pour plus d'explications, voir

<<http://www.autoroute.gouv.qc.ca/publica/normes/norme106.htm>> et
<<http://www.autoroute.gouv.qc.ca/publica/normes/norme105.htm>>.

- Le profil et l'entente de protocole de collaboration : le profil CPP (*Collaboration Protocol Profile*) définit comment décrire les capacités d'échange de messages de chaque partie dans un échange d'affaires. L'intersection de deux profils permet de déterminer comment deux parties peuvent échanger des messages entre elles en étant assurées de l'interopérabilité. L'idée de base est d'alléger les conditions techniques nécessaires pour amorcer des échanges d'affaires, et d'offrir des procédures bien contrôlées d'échange. Profil et entente précisent les détails du transport et de la messagerie et permettent de simplifier la négociation en vue d'une entente CPA (*Collaboration Protocol Agreement*) en ajoutant les contraintes de sécurité ainsi que la signification dans les affaires des parties par le rattachement à un document de spécification du processus d'affaires. Ce genre d'entente de collaboration est compatible avec les concepts d'entente de sécurité entre domaines de confiance de l'architecture gouvernementale de sécurité de l'information numérique (AGSIN). Pour plus d'explications, voir

<<http://www.autoroute.gouv.qc.ca/publica/normes/norme123.htm>>.

La spécification du processus d'affaires définit la collaboration en termes d'interactions entre les parties. L'entente fait ainsi se rejoindre l'aspect affaires et l'aspect fonctionnement technique. Par la suite, chaque partie configure ses systèmes selon l'entente. C'est la spécification de processus qui, par l'intermédiaire de l'entente hébergée par chaque partie, contrôle le flux des échanges de messages faits dans le cadre d'une collaboration. Globalement, une entente de collaboration comprend une partie branchée techniquement fonctionnant en conformité avec ce qui en constitue l'autre partie, soit la définition de tout ce qui est possible dans une collaboration d'affaires : le but est que rien d'imprévisible ne puisse se produire sans rejet immédiat. Il s'agit d'une condition de confiance pour les affaires électroniques. UML (composant d20) est utilisé pour capturer et modéliser les échanges entre partenaires commerciaux. Cette partie d'ebXML reste inachevée en avril 2003. Pour plus d'explications, voir

<<http://www.autoroute.gouv.qc.ca/publica/normes/norme118.htm>>.

Une correction visée par rapport à l'échange électronique informatisé (EDI) traditionnel a conduit à s'intéresser aux composants élémentaires, en réduisant à la taille de l'élément de données le plus atomique possible le référencement supporté dans les conventions d'échange. L'interopérabilité est fixée comme le but essentiel à atteindre. La stratégie globale comporte l'institutionnalisation de bibliothèques de composants « élémentaires » ou assemblés et l'adoption d'une méthodologie de développement d'applications qui intègre bien les étapes de repérage dans ces bibliothèques aux étapes d'analyse et de conception. Le but est une interopérabilité dans laquelle la sémantique d'affaires est garantie sur le plan conceptuel malgré les différences d'appellations et de représentations : cela suppose que la méthodologie exige de déclarer ces points de détails de sorte que les automates puissent y puiser l'information pour garantir la mise en équivalence conceptuelle. Pour plus d'explications, voir

<<http://www.autoroute.gouv.qc.ca/publica/normes/norme111.htm>>.

Ces points de détails, ces variations locales, cette fine granularité, ce sont les composants élémentaires qui en forment le pivot descriptif. Les composants élémentaires sont des pièces d'assemblage qui ont chacune une définition unique en sémantique d'affaires. L'assemblage de ces pièces se fait dans des contextes précis (différents processus d'affaires) où elles acquièrent une signification particulière : par exemple, le composant élémentaire « date » peut être employé dans l'entité d'information d'affaires « date de publication ». Quand un composant élémentaire est ainsi placé dans un contexte d'affaires, il est introduit dans la définition d'une entité d'information d'affaires.

La constitution de bibliothèques d'entités d'information d'affaires est le but d'une initiative distincte mais coordonnée, *Universal Business Language* (UBL). Le projet UBL fait appel à tous les services offerts par ebXML, mais tout particulièrement à l'aspect composants élémentaires (*Core Components*) dont relève une partie majeure de l'interopérabilité, et ce, grâce à un ensemble de définitions de types de données primaires et de types dérivés créés par l'imposition de contraintes en fonction de contextes. Chaque entité d'information d'affaires définie en UBL est basée sur un composant élémentaire défini dans ebXML. Quant aux entités d'information d'affaires, ce peut être les blocs adresse, prix ou date qu'on utilise pour constituer les formulaires, par exemple. Les composants définis dans ebXML le sont comme des objets de données abstraits, et UBL prolonge ce travail en fournissant un encodage standard de ces définitions d'ebXML. Pour plus d'explications, voir <<http://www.autoroute.gouv.qc.ca/publica/normes/norme109.htm>>.

Une initiative de l'OASIS regroupant dans un comité technique différents gouvernements a, parmi ses mandats, celui d'alimenter ebXML en besoins et exigences communs au secteur public (voir son programme de travail :

<http://www.oasis-open.org/committees/egov/docs/work-programme-egov-TC.rtf>).

Le scénario suivant, basé sur une infrastructure ebXML, présente deux compagnies qui font des affaires. La compagnie A est un fournisseur de services, alors que la compagnie B en est un utilisateur.

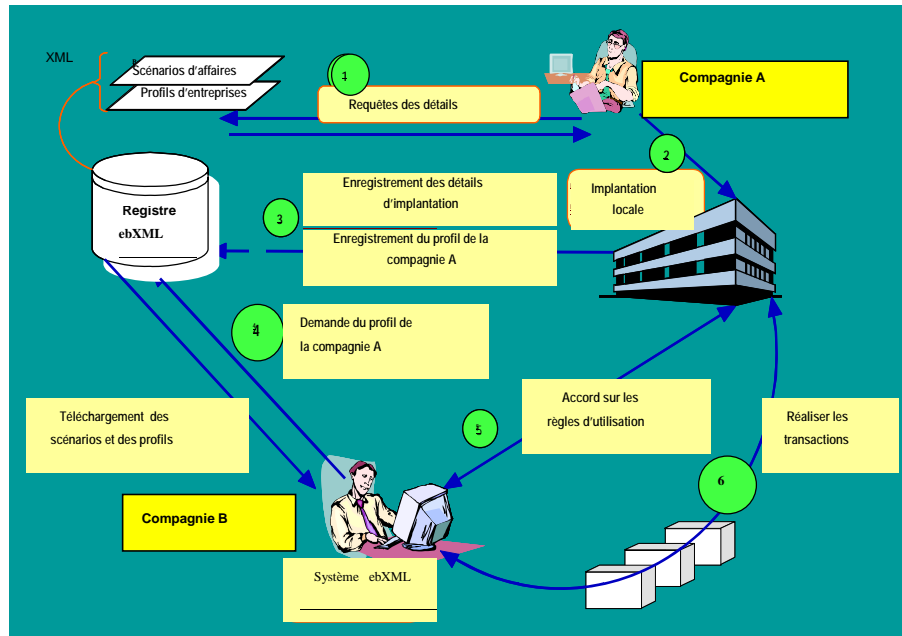


Figure 4 : Interactions d'affaires entre deux compagnies, basées sur ebXML

(mise en situation du composant d130)

1. Le scénario débute alors que la compagnie A consulte les spécifications dans le répertoire ebXML dans le but de prendre une décision sur les services qu'elle veut rendre conformes aux normes ebXML afin de les rendre disponible publiquement.
2. En se basant sur les spécifications qu'elle trouve dans le répertoire, la compagnie A modifie les applications ciblées pour les rendre conformes.
3. Par la suite, elle enregistre les informations sur la compagnie et les services d'affaires qu'elle veut publier dans le répertoire ebXML (*protocole CPP*).
4. La compagnie B cherche dans le répertoire ebXML un service qui répond à ses besoins et découvre ainsi les services de la compagnie A.
5. La compagnie B prend entente avec la compagnie A pour l'utilisation de ses services d'affaires (*protocole CPA*).
6. À l'étape finale, la compagnie B effectue des transactions commerciales via l'échange de messages SOAP, et ce, de manière sécuritaire et fiable (*ebXML Messaging Service*).

Depuis les débuts de l'initiative, un nombre important d'entreprises, de fournisseurs et un consortium de plus de 400 compagnies se sont joints au projet ebXML. Un scénario tel que celui illustré ci-dessus correspondra bientôt de plus en plus à la façon de faire des affaires par l'intermédiaire d'Internet.

Orientations

Loin d'être un compétiteur des services Web (voir l'étude 116989, *Different Goals Mean ebXML Is Not a Rival to Web Services*, du groupe Gartner), ebXML peut les compléter pour mettre en œuvre les processus d'affaires entre organisations, auxquels il apporte le socle de modélisation et d'architecture fonctionnelle nécessaire et des solutions mieux adaptées à une collaboration entre rôles.

L'utilisation d'ebXML est à prendre en compte au gouvernement du Québec pour ses caractéristiques de normes ouvertes et son orientation de service au public, de solution de qualité pour le *monde ordinaire*.

Il est du ressort du concepteur de décider, selon ses contraintes, comment utiliser les composants génériques d'affaires UBL et comment les étendre. Les extensions devraient du point de vue de la méthode faire le tri entre les besoins propres à un M/O, les besoins communs au gouvernement du Québec et les besoins communs à tout service offert par un gouvernement.

Il est de la responsabilité du chargé de projet de réutiliser aussi largement que possible les composants élémentaires publiés dans le registre référentiel du Québec et de publier dans le registre *ad hoc* tout composant réutilisable, en enrichissant son profil.

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d130	Infrastructure électronique d'affaires	ebXML	Prendre en compte	http://www.ebxml.org/	<ul style="list-style-type: none"> • ebXML fournit un ensemble de spécifications, de pratiques et une bibliothèque de <i>composants en libre service</i> permettant de dématérialiser un nombre croissant de transactions commerciales entre organisations de façon standard, et pour cela d'échanger des messages, de définir et d'enregistrer des processus d'affaires ainsi que les documents de transaction qui les alimentent. • ebXML définit également un protocole d'entente entre participants (<i>Collaboration Protocol Agreement</i>) et un protocole de configuration et d'échange de messages (<i>ebXML Messaging Service</i>). • Le cadre concret applicable au secteur public est encore peu étendu et devrait s'enrichir progressivement, puisqu'il sera alimenté par le <i>eGov TC - Electronic Government Technical Committee</i> de l'OASIS (http://www.oasis-open.org/committees/egov/).
d135	Protocole de messagerie sûre	ebXML Messaging	Prendre en compte	http://www.w3.org/TR/SOAP/ RFC2616 http://www.ietf.org/rfc/rfc2616.txt RFC2774 http://www.ietf.org/rfc/rfc2774.txt RFC2965 http://www.ietf.org/rfc/rfc2965.txt	<ul style="list-style-type: none"> • Il est encapsulé dans SOAP qui lui-même s'appuie sur HTTP. • Se référer au volet 3 du cadre commun d'interopérabilité, notamment les composants <i>i250</i> et <i>i70</i>. • Actuellement, le comité sur la messagerie ebXML s'attèle à examiner les spécifications et propositions relatives à la sécurité des services Web.

No	Composant	Norme ou standard	Orientation	Références	Commentaires
d140	Composant élémentaire alimentant un processus d'affaires	UBL	Prendre en compte	http://www.ebxml.org/	<ul style="list-style-type: none"> • Les développements en cours des composants élémentaires ebXML (<i>Electronic Business using XML</i>) et UBL (<i>Universal Business Language</i>) doivent être réutilisés à mesure qu'ils seront rendus disponibles. • Les premiers composants alimentant le cycle d'exploitation d'une entreprise (lequel fait intervenir un acheteur, un vendeur et le transporteur du fret) sont accessibles sur http://oasis-open.org/committees/ubl/lcsc/Op70/ > et font l'objet d'un appel public à commentaires. • Voir http://www.ebxml.org/implementations/index.htm pour des implémentations.

2.4 Composants logiciels (serveurs et scripts) réutilisables pour la logique d'affaires

Contexte

De nouveaux besoins de dématérialisation de traitements naissent constamment au sein du gouvernement du Québec. Pour chaque besoin qui émerge, il est important dans sa collecte et son analyse de faire le tri entre les besoins propres à un M/O, les besoins communs au gouvernement du Québec et les besoins communs à tout service offert par un gouvernement. Il en va de même des documents alimentant les processus.

La réponse au besoin impose souvent de nouveaux développements au-delà de la réutilisation des composants déjà développés, qu'il s'agisse de composants standards de l'OASIS (voir composant d140) ou de composants génériques du gouvernement du Québec. Par ailleurs, tout nouveau développement doit prendre en compte une exigence supplémentaire : l'interopérabilité.

Dans les différentes solutions d'architecture pour la création de nouvelles applications, les développements doivent prendre en compte un facteur favorisant l'interopérabilité, à savoir la structuration de type MVC (modèle - vue - contrôleur).

Le modèle de conception MVC consiste en effet à organiser l'application à développer en trois composants majeurs : un modèle, qui correspond à la logique applicative, une vue, qui correspond à la présentation visuelle de l'application, et un contrôleur, qui définit l'état de la vue en fonction des données gérées par le modèle.

Les composants serveurs des plates-formes orientées objets devraient être appelés à modéliser de plus en plus la logique d'affaires *interne* à chaque M/O en manipulant les données. L'interopérabilité « orientée documents » entre les différentes plates-formes déployées au sein de chaque ministère ou organisme devrait être garantie de plus en plus par le recours à la famille XML, notamment ebXML, et au déploiement de services Web respectant les normes et standards présentés dans le présent cadre.

Quand les logiques d'affaires (procédures) diffèrent peu d'un M/O à l'autre, les investissements consentis par un M/O dans la modélisation et la technologie peuvent être mutualisés; les bibliothèques de composants serveurs et de scripts (Java script et ECMA 262; voir volet 2) devraient être le plus générique possible afin d'être *réutilisables* et/ou *partageables* (tout en tenant compte des contraintes de gestion des versions, de sécurité et de qualité de service, dont la disponibilité) dans le temps et entre M/O, et indépendants des données qu'ils traitent.

Il est notable que le langage de modélisation UML favorise la réutilisabilité des composants.

La DGSIG (Direction générale des systèmes d'information de gestion) rend ainsi disponibles des composants génériques dans le référentiel (<http://www.referentiel.qc/>).

En outre, toute la couche intermédiaire d'infrastructure, décrite dans le volet 2, est en train de se standardiser autour d'XML et de ses composants modulaires : services Web et messagerie ebXML pour les échanges entre applications, VoiceXML pour le traitement de la voix, SMIL pour le traitement de la vidéo, SVG pour le traitement des images vectorielles. Cette couche d'abstraction obéit à une tendance générale qui vise l'affranchissement par rapport à l'hétérogénéité des technologies d'interconnexion décrites dans le volet 3.

Elle permet également de standardiser les services rendus par la couche intermédiaire et permet donc au final *le développement de composants génériques* pouvant fournir le service. Un tel exemple est donné par VoiceXML et les bibliothèques de composants Java disponibles destinées aux développeurs.

Une fois développés, ces composants doivent être enregistrés et publiés dans le registre référentiel du gouvernement du Québec, traité ci-après.

Orientations

Il est suggéré aux M/O de retenir l'approche modèle - vue - contrôleur et, pour les réalisations, de privilégier le développement de composants serveurs *génériques* et *réutilisables* en utilisant le langage de modélisation UML.

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d150	Modèle d'architecture	MVC	Retenir		<ul style="list-style-type: none">• Organise l'application à développer en trois composants majeurs :<ul style="list-style-type: none">❑ un modèle, qui correspond à la logique applicative;❑ une vue, qui correspond à la présentation visuelle de l'application;❑ un contrôleur, qui définit l'état de la vue en fonction des données gérées par le modèle.

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d160	Composants logiciels serveurs et scripts	Généricité, réutilisabilité	Prendre en compte		<ul style="list-style-type: none"> • Les composants sont appelés à modéliser de plus en plus la logique d'affaires. • Il faut pouvoir réutiliser les composants logiciels publiés dans des bibliothèques. • Toute la couche intermédiaire d'infrastructure, décrite dans le volet 2, est en train de se standardiser autour d'XML et de ses composants modulaires : VoiceXML pour le traitement de la voix, SMIL pour le traitement de la vidéo, SVG pour le traitement des images vectorielles. Cette couche d'abstraction obéit à une tendance générale qui vise l'affranchissement par rapport à l'hétérogénéité des technologies d'interconnexion décrites dans le volet 3. • Elle permet également de standardiser les entrées et sorties de chaque composant de la couche intermédiaire et permet donc au final le développement de composants génériques pouvant fournir le service. Voir par exemple VoiceXML et les bibliothèques de composants Java disponibles en ligne. • Les scripts sont traités dans le volet 2; respectant le standard ECMA 262, une partie importante des scripts devrait être remplacée par le recours à Xforms, mais la logique non prise en compte par Xforms est développée en scripts qu'il faut rendre les plus génériques possible et publier. • Une fois développés, ces composants doivent être enregistrés et publiés dans le registre référentiel du gouvernement du Québec.

2.5 Registre référentiel

Contexte

L'enregistrement dans un registre référentiel¹⁰ du gouvernement du Québec des objets communs (schémas XML, règles de présentation, domaines de valeurs, composants logiciels, etc.) permet de faciliter la réutilisation des objets et entraîne ainsi des gains d'efficacité et de qualité.

Il permet de plus l'intégration des données : il s'agit de prescrire comment identifier, nommer, définir, décrire, localiser et administrer les éléments structurants dans la gestion des règles d'affaires et des connaissances au sein de l'administration gouvernementale.

Constitué technologiquement d'une ou de plusieurs bases de données spécialisées selon l'utilisation qui en est faite, le registre référentiel doit être disponible en ligne.

Enregistrer et publier les composants ebXML

Une partie du registre du gouvernement du Québec doit être *spécialisée* pour offrir des services associés aux circuits de production ebXML dans les processus d'affaires : rôles, échange de messages, chorégraphie des parties participant à l'échange de messages, ententes de collaboration, modalités de réutilisation des composants élémentaires, multiples vocabulaires de domaines, comme XBRL (*eXtensible Business Reporting Language*) en information financière, HL7 (*Health Level Seven*) en santé, etc.

Le registre doit pouvoir être parcouru en fonction des catégories de contexte et en fonction des entités d'information d'affaires qui sont présentes dans les processus d'affaires.

Soit le développeur réutilise les composants en les dérivant selon les contraintes locales, soit il crée des entités, des processus ou des composants assemblés et les ajoute au référentiel en les inscrivant au registre. La qualité du registre doit être soutenue au cours des ans par un processus d'évaluation et d'attribution de statut.

Il faut chercher à connaître ses entités par rapport au contexte et chercher à établir ce qui est constant et ce qui varie, et ainsi ce qui est composant élémentaire (valable hors contexte) et ce qui est entité d'affaires variant en contexte.

Les catégories de contextes définies par la norme ISO/CEI 11179, autrement dit les facettes de contextes utilisées pour classifier les UBL, sont :

- le contexte des processus d'affaires, tels que ces derniers sont décrits dans la bibliothèque des processus d'affaires communs (*Catalogue of Common Business Processes*);
- le contexte de la classification des produits;
- le contexte de la classification industrielle;
- le contexte géopolitique;
- le contexte des contraintes officielles (légales, etc.);

¹⁰ Voir la définition en annexe.

- le contexte des rôles joués dans les processus d'affaires (comme vendeur ou acheteur, par exemple) décrits dans la bibliothèque des processus d'affaires communs;
- le contexte du rôle de soutien que joue tout participant à un échange de données dont il n'est ni l'émetteur ni le récepteur;
- le contexte de la capacité du système, laquelle résulte d'une contrainte du système ou de la conformité à un standard.

Les entités d'information d'affaires sont également enregistrées selon les règles de dénomination, de définition et de description conformes à la norme ISO/CEI 11179, mais une série de règles spécifiques à ebXML sont ajoutées.

Par ailleurs, un *langage de contraintes* permet d'exprimer comment la sémantique s'applique au composant élémentaire en contexte pour produire l'entité d'information d'affaires. Les contraintes d'assemblage et les enrichissements sémantiques sont exprimés séparément dans deux parties du langage de contraintes, ce qui est pratique dans le déploiement pour distinguer l'assemblage fait par les développeurs d'applications des ajustements sémantiques faits selon les besoins ou préférences des utilisateurs de l'application.

Enregistrer et publier les autres composants

Qu'il s'agisse d'un composant logiciel ou d'un document de référence, l'auteur doit au moment de son enregistrement enrichir son profil, c'est-à-dire compléter les différents champs définis par le composant d80, « Profils de métadonnées pour la sémantique gouvernementale du Québec ».

Dans cette perspective, le référentiel de composants partageables et réutilisables du gouvernement du Québec disponible à l'adresse <<http://www.referentiel.qc/>> est une initiative de la DGSIG (Direction générale des systèmes d'information de gestion) qui permet aux ministères et aux organismes du gouvernement du Québec de partager les descriptions de traitements et de données.

Cet initiative qui cumule chacune les fonctions de registre et de référentiel, devraient converger vers un seul registre référentiel, comme l'illustre le schéma qui suit.

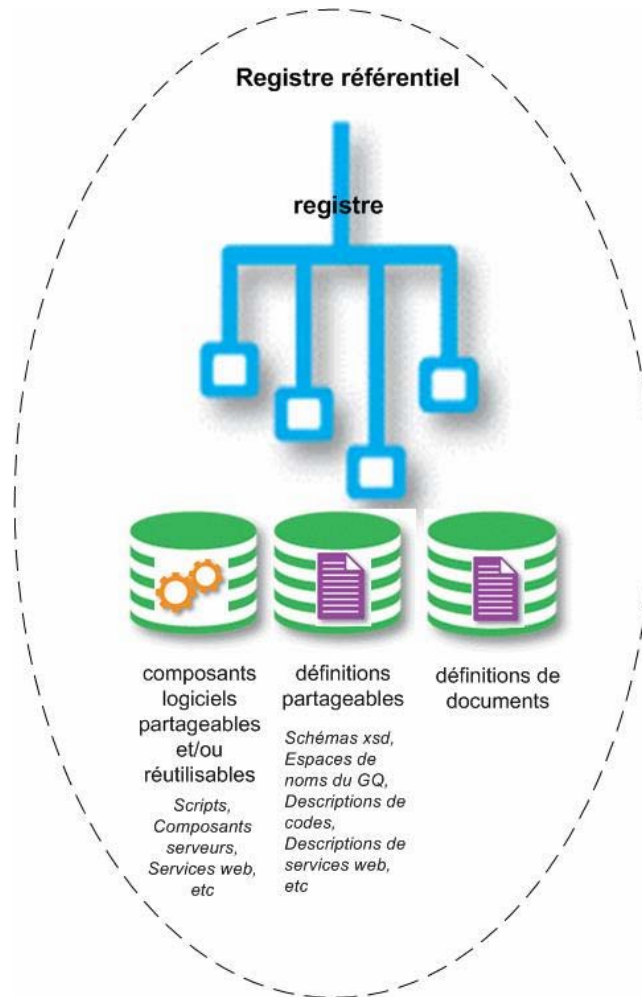


Figure 5 : Schéma de principe du registre référentiel du gouvernement du Québec

(composant d170)

Orientation

Quels que soient les objets (document, donnée, composant logiciel) à portée interministérielle, il convient d'établir à leur sujet une référence commune et de la publier dans le registre du gouvernement du Québec.

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d170	Partage et publication des définitions des objets communs	Registre référentiel du gouvernement du Québec	Retenir	<p>ISO/CEI 11179</p> <p>Registre référentiel</p> <p>ebXML/OASIS</p> <p>http://www.oasis-open.org/committees/regrep/documents/2.0/specs/ebxml.pdf</p> <p>http://www.oasis-open.org/committees/regrep/documents/2.0/specs/ebxml.pdf</p>	<ul style="list-style-type: none"> • C'est un standard en cours de normalisation. • Conceptuellement, le registre et le référentiel sont dissociés, mais technologiquement, ils peuvent cohabiter dans une base de données. • Le registre référentiel facilite la réutilisation des objets et permet l'intégration des données.

3 Répertoire

Contexte

Les répertoires informatiques sont des outils conçus pour gérer efficacement des données de référence. Ce sont des données qui sont typées, souvent structurées, hautement partagées et accessibles beaucoup plus fréquemment en lecture qu'en écriture.

Le gouvernement du Québec étudie la possibilité de mettre en place un répertoire gouvernemental pour couvrir ses besoins en matière de sécurité (identification/authentification, autorisation/habilitation) et de localisation, que ce soit pour les employés, les unités organisationnelles ou les dispositifs.

La norme X.500 approuvée en 1988 a été élaborée par l'Union internationale des télécommunications (UIT) pour la réalisation de répertoires accessibles via le protocole d'interrogation DAP (*Directory Access Protocol*). La lourdeur et la complexité de X.500 ont limité les implantations de répertoires X.500 respectant toutes les recommandations. Seules certaines ont été par la suite retenues (X.509, par exemple) ou utilisées comme base de travail (le protocole X.500 DAP a été simplifié pour devenir le protocole LDAP v3).

En 1995, l'Université du Michigan (<http://www.umich.edu/~dirsvcs/ldap/ldap.html>) a créé le premier serveur LDAP (*standalone LDAP*), qui constitue une version plus simple du répertoire en vue d'une utilisation sur Internet.

Le volet 3 du Cadre commun d'interopérabilité standardise le composant « Accès à un répertoire ». Mais le protocole d'accès, s'il est important, n'est pas tout : il s'agissait de définir la structure du répertoire elle-même.

Ainsi, l'IETF (*Internet Engineering Task Force*) a normalisé la structure du répertoire elle-même. Il existe peu de manufacturiers proposant des solutions de répertoires basées sur X.500, alors qu'il existe un choix important du côté de LDAP. La plupart des fournisseurs de logiciels de messagerie électronique, de services d'annuaires ainsi que des constructeurs informatiques ont adapté la majorité de leurs produits au protocole LDAP. De plus, la plupart des applications clients sont compatibles avec LDAP.

Le répertoire LDAP v3 s'intègre parfaitement à l'infrastructure PKIX (*Public-Key Infrastructure X.509*). Voir <<http://www.ietf.org/internet-drafts/draft-klasen-ldap-x509certificate-schema-02.txt>>.

Le RFC 2251 (*Lightweight Directory Access Protocol v3*) définit comment l'information est structurée et gérée sur les plans logique et physique (répartition entre plusieurs serveurs par exemple) ainsi que les primitives d'accès à l'information.

Le RFC 2252 (*Lightweight Directory Access Protocol (v3): Attribute Syntax Definitions*) définit un modèle d'information caractérisant les types de données *standards*, c'est-à-dire des attributs et les objets contenant ces attributs.

Les objets standards couvrent les domaines suivants : les unités organisationnelles (les directions et les services d'une organisation), les personnes, les fonctions et la sécurité, notamment le format de certificat X.509 détaillé ci-après.

Les objets utilisés dans un répertoire LDAP peuvent être soit créés, soit dérivés des objets standards. Le respect des objets standards constitue une garantie d'interopérabilité; par contre, ils ne conviennent pas toujours aux besoins spécifiques des entreprises.

Il s'agit ici de donner les orientations quant à l'organisation du répertoire et notamment aux attributs relatifs à chaque nœud, afin de favoriser l'interopérabilité à long terme du répertoire mis en ligne.

Orientations

Les orientations générales du gouvernement du Québec retiennent LDAP version 3 comme structure de répertoire informatique et interdisent l'utilisation d'attributs non standards au sens du RFC 2252.

La difficulté majeure dans la mise en œuvre d'un répertoire est de définir le schéma de répertoire, c'est-à-dire les attributs qu'il contient, les objets contenant ces attributs et la hiérarchie entre ces objets, ainsi que la multiplicité des sources de données et de formats. Les orientations fines garantissant l'interopérabilité entre répertoires et portant notamment sur les attributs ainsi que les solutions technologiques seront données une fois les livraisons du projet en cours « répertoire gouvernemental » effectuées.

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d190	Structure et modèle d'information d'un répertoire informatique	LDAP v3	Retenir	<p>IETF http://www.ietf.org/rfc/rfc3377.txt</p> <p>RFC 2251 http://www.ietf.org/rfc/rfc2251.txt</p> <p>RFC 2252 http://www.ietf.org/rfc/rfc2252.txt</p> <p>http://www.ietf.org/internet-drafts/draft-klasen-ldap-x509certificate-schema-02.txt</p>	<ul style="list-style-type: none"> • Standard de l'IETF, LDAP est le modèle d'information à retenir. • Il ne faut pas utiliser d'attributs propriétaires. • La plupart des fournisseurs de logiciels de messagerie électronique, de services d'annuaires ainsi que des constructeurs informatiques ont adapté dans leur majorité leurs produits au protocole LDAP. • Le répertoire LDAP v3 s'intègre parfaitement à l'infrastructure PKIX («Public-Key Infrastructure X.509»).
		X.500 X.501	Ne pas retenir	<p>UIT http://www.itu.int/rec/recommendation.asp?type=folders&lang=e&parent=T-REC-X.500</p> <p>http://www.itu.int/rec/recommendation.asp?type=folders&lang=e&parent=T-REC-X.501</p>	<ul style="list-style-type: none"> • Les normes ISO ont été abandonnées par plusieurs pays au profit de LDAP. • Le standard X.500 est très complexe. Il est par conséquent peu répandu.
d200	Organisation du répertoire du gouvernement du Québec (schéma d'annuaire)	En cours de réalisation	Éligible dans une version ultérieure	Livrables du projet du Secrétariat du Conseil du trésor «Répertoire gouvernemental»	<ul style="list-style-type: none"> • Communiquer avec Patrice di Marcantonio, Patrice.di-marcantonio@sct.gouv.qc.ca

3.1 Un objet particulier du répertoire : le certificat numérique

Contexte

Le certificat numérique est un document électronique qui atteste qu'une clé publique est bien liée à une personne ou à une organisation. Il permet donc de vérifier la propriété d'une clé publique pour prévenir la contrefaçon.

Les certificats X.509 v3, normalisés par l'ISO et par l'IETF, sont utilisés, entre autres, dans les applications, protocoles et standards suivants : SSL/TLS, IPSec, S/MIME, chiffrement, authentification, signature, horodatage, etc.

Les certificats X.509 v3 s'intègrent parfaitement à l'infrastructure à clé publique PKIX ([voir le RFC 2459](#)).

Selon l'utilisation qui est faite de la clé, un certificat peut être stocké dans un répertoire (cas des certificats pour le chiffrement, et ce, pour des raisons liées aux contraintes de protection des renseignements personnels) ou être conservé par l'utilisateur (par exemple, sur un support logiciel ou matériel).

Il existe plusieurs normes définissant des formats de certificats. La plus répandue est la norme X.509 version 3. De plus, au Québec, la Loi concernant le cadre juridique des technologies de l'information impose un contenu minimal aux certificats délivrés par un prestataire de services de certification.

Le certificat de clé publique doit contenir au moins les renseignements suivants :

- Le nom distinctif du prestataire de services qui délivre le certificat ainsi que sa signature;
- la référence à l'énoncé de politique du prestataire de services de certification, y compris ses pratiques, sur lesquelles s'appuient les garanties qu'offrent le certificat qu'il délivre;
- le numéro de la version de X.509 et le numéro de série du certificat;
- la période de validité (début et fin);
- le nom distinctif (ou pseudonyme) de l'entité qui est titulaire du certificat;
- s'il s'agit d'un certificat d'attribut, la désignation de l'attribut dont le certificat confirme l'existence et, au besoin, l'identification de la personne, de l'association, de la société, de l'état ou de l'objet auquel il est lié.

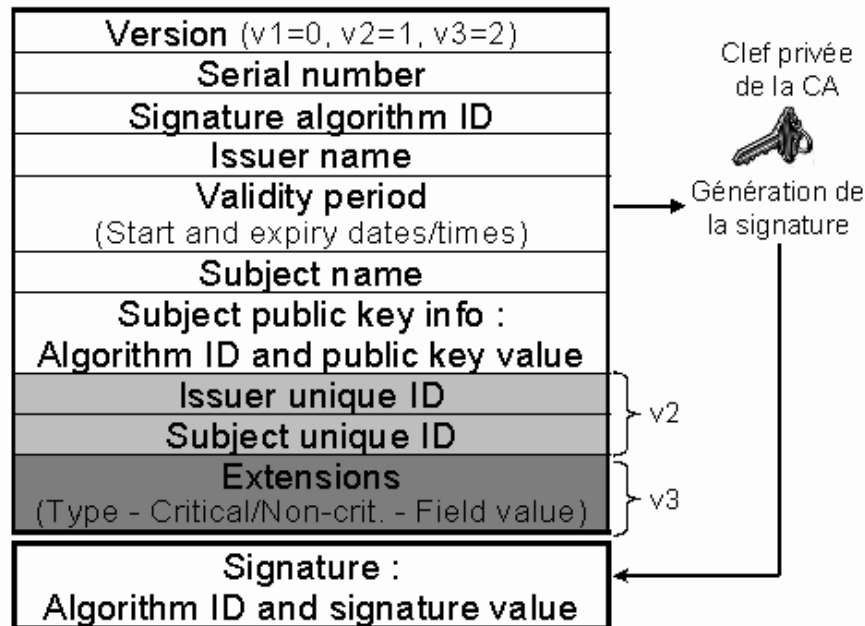


Figure 6 : Description des champs d'un certificat X.509 (source : www.hsc.fr)

- *Version* : indique à quelle version de X.509 correspond le certificat.
- *Serial number* : numéro de série du certificat (propre à chaque autorité de certification).
- *Signature algorithm ID* : identifiant du type de signature utilisée.
- *Issuer name* : nom distinctif (*distinguished name*) de l'autorité de certification qui a délivré ce certificat.
- *Validité period* : période de validité.
- *Subject name* : nom distinctif du détenteur de la clef publique.
- *Subject public key info* : information sur la clef publique de ce certificat.
- *Issuer unique ID / Subject unique ID* : extensions optionnelles introduites avec la version 2 de X.509.
- *Extensions* : extensions génériques optionnelles, introduites avec la version 3 de X.509.
- *Signature* : signature numérique de l'autorité de certification apposée sur l'ensemble des champs précédents.

La version 3 de la norme X.509 apporte par rapport aux versions précédentes, la notion d'extensions. Ces extensions, quand elles existent, sont incluses dans le certificat. La plus courante est l'extension « KeyUsage ». Cette dernière limite la clé à une utilisation précise. À titre d'exemple, l'extension du champ « KeyUsage » renseigne sur l'utilisation qui doit être faite de la clé.

Par ailleurs, aucune extension n'est obligatoire selon le standard de l'IETF, donc il convient de définir pour chaque extension si elle est considérée critique ou non dans le cadre de l'infrastructure déployée par le gouvernement du Québec. À défaut, des problèmes d'interopérabilité ou de mauvaise utilisation du certificat pourraient survenir.

Orientations

Toute l'industrie des technologies de l'information considère la recommandation X.509 version 3 comme une référence incontournable dans le domaine des infrastructures à clé publique. Il est recommandé aux M/O d'adopter cette recommandation.

Aucune extension n'est obligatoire au sens du standard de l'IETF. Cependant, les M/O doivent se référer aux spécifications de l'infrastructure à clé publique au sein du gouvernement du Québec. De même, le traitement des certificats en versions antérieures (version 1 et version 2) devrait être pris en compte.

No.	Composant	Norme ou standard	Orientations	Références	Commentaires
d210	Certificat numérique au gouvernement du Québec	X.509 v3	Retenir	<p>Norme ISO X.509</p> <p>http://www.itu.int/rec/recommendation.asp?type=folders&lang=e&parent=T-REC-X.509</p> <p>Standard IETF RFC 3280</p> <p>http://www.ietf.org/rfc/rfc3280.txt</p> <p>http://www.autoroute.gouv.qc.ca/publica/pdf/Directive_sur_les_services_de_certification.pdf</p>	<ul style="list-style-type: none"> L'utilisation de certificats X.509 v3 par l'ensemble des protocoles et des applications du monde Internet favorise l'interopérabilité entre les différents produits. X.509 v3 est à retenir par les ministères et organismes. Aucune extension n'est obligatoire selon le standard de l'IETF. Cependant, l'implémentation de l'infrastructure à clé publique au sein du gouvernement du Québec devrait définir certaines extensions comme critiques et d'autres comme ne l'étant pas. Le traitement des certificats en versions antérieures (version 1 et version 2) devrait être pris en compte. Les certificats X.509 v3 s'intègrent parfaitement dans l'infrastructure à clé publique PKIX (<i>Public-Key Infrastructure X.509</i>).
d220	Extensions des certificats X.509 au gouvernement du Québec	En cours de réalisation	Prendre en compte	<p>Communiquer avec Myriam Cyr, myriam.cyr@sct.gouv.qc.ca</p>	<ul style="list-style-type: none"> Il faut se référer aux spécifications de l'ICPG.X.509.

4 Applications au domaine de l'information géographique

Contexte

Ce chapitre, à l'état embryonnaire pour le moment, s'adresse essentiellement aux concepteurs, développeurs et responsables d'une application à base de données géographiques afin d'assurer l'interopérabilité de leur application avec celles de leurs partenaires.

La responsabilité pour élaborer le contenu de ce volet incombe au comité directeur du Plan géomatique du gouvernement du Québec.

Pour avoir un aperçu de ce que pourrait contenir ce volet, on peut consulter le projet proposé en 2002 par l'ATICA (Agence pour les technologies de l'information et de la communication dans l'Administration) et le CNIG (Conseil national de l'information géographique) pour le volet géomatique (http://www.atica.pm.gouv.fr/upload/documents/cadre_commun_IG.rtf) du cadre commun d'interopérabilité destiné à répondre aux besoins de l'Administration française.

Orientation

L'ensemble du contenu du sous-ensemble géomatique doit appliquer le cadre commun d'interopérabilité.

À titre indicatif, voici quelques éléments préliminaires :

No	Composant	Norme ou standard	Orientation	Référence	Commentaires
d230	Métadonnée	À venir	Éligible dans une version ultérieure	À venir	<ul style="list-style-type: none">• À entreprendre.• Ce composant constitue un profil de la norme canadienne CAN/CGSB-171.3-0?ISO 19115:2002. Remarque : <ul style="list-style-type: none">• Il est possible de mettre en place un profil Canada-États-Unis de la norme ISO 19115:2003.
d240	Identifiant géographique	À venir	Éligible dans une version ultérieure	À venir	<ul style="list-style-type: none">• À entreprendre. Remarque : <ul style="list-style-type: none">• Il est possible de mettre en place un profil Canada-États-Unis de la norme ISO 19112:2002.
d250	Thésaurus des entités géographiques du gouvernement du Québec	À venir	Éligible dans une version ultérieure	À venir	<ul style="list-style-type: none">• En cours d'élaboration.• Cet élément constitue une évolution du <i>Dictionnaire des entités géographiques du gouvernement du Québec</i>.

Annexe : définitions

Autorité source	Organisation qui exerce une responsabilité reconnue pour la création et l'évolution d'un objet qui peut être partagé et réutilisé par les ministères et les organismes.
Codet	Représentation d'un objet appartenant à un ensemble. Par exemple : <ul style="list-style-type: none">• les lettres « CDG » constituent le codet de « Paris-Charles-de -Gaulle » dans le jeu codé de la représentation internationale des noms d'aéroports;• dans le code Morse, les codets des lettres de l'alphabet sont des groupes de points et de traits;• les lettres « CA » constituent le codet de « Canada » dans la norme ISO 3166-1 1997, <i>Codes pour la représentation des noms de pays et de leurs subdivisions</i> – Partie 1 : <i>Codes pays</i> (définition adaptée de l'Office de la langue française, 1999).
Domaine énumératif	Domaine de valeurs qui est caractérisé par une liste de toutes les valeurs permises (traduction : ISO/CEI WD 11179-3:200x, 3.3.72).
Domaine énumératif codifié	Domaine énumératif où chacune des valeurs qui en fait partie est identifiée par un identifiant unique délivré par une autorité source (adaptation : ISO/CEI WD 11179-3:200x, 3.3.72).
Donnée	Représentation d'une information codée dans un format permettant son traitement par un programme.
Identifiant	<p>Valeur non ambiguë, unique, permanente et linguistiquement neutre, résultant de l'application d'un processus d'identification à base de règles, utilisée pour identifier un objet. Exemple : les caractères numériques « 124 » constituent l'identifiant de « Canada » dans la norme ISO 3166-1 1997, <i>Codes pour la représentation des noms de pays et de leurs subdivisions</i> – Partie 1 : <i>Codes pays</i>.</p> <p>Un identifiant est unique dans le système d'identification de l'autorité qui le délivre (définition adaptée de la norme ISO/CEI 15944-1).</p> <p>Un URI ne répond pas complètement à la définition de l'Organisation internationale de normalisation : tous les URI sont certes uniques et non ambigus et répondent à une identification à base de règles (le schéma de l'URI), mais seuls les URN, parmi les URI, sont permanents, et aucun URI n'est linguistiquement neutre. C'est que l'URI a le double rôle de permettre d'identifier et de représenter la ressource à l'utilisateur. Ainsi, l'URI est souvent exprimé dans la langue de celle ou celui qui l'a conçu.</p>
Référentiel	Le référentiel dans lequel sont conservées les composantes est un mécanisme d'entreposage, de préservation et de distribution des fichiers contenant les définitions de structures, soit les objets correspondant aux items inscrits dans un registre. Dans le modèle de l'OASIS, le référentiel est une sorte d'entrepôt spécialisé pour les objets hautement réutilisables que sont les définitions de structures.

Registre	Un registre est un mécanisme d'enregistrement d'objets; il est institué par une organisation en vue de permettre à ses acteurs légitimes l'inscription de métadonnées d'identifiant et décrivant un objet ainsi que la localisation indiquant où est simultanément placé l'objet dans un lieu distinct d'entreposage (le référentiel).
Répertoire	Un répertoire est une base de données répartie optimisée pour fournir des références, c'est-à-dire un dispositif d'emmagasinement d'information sur des « objets » d'intérêt pour une communauté d'utilisateurs, par exemple le gouvernement, un ministère, les responsables de la gestion du personnel, les clientèles du gouvernement, la population, etc.
Identifiant	<p>Valeur non ambiguë, unique, permanente et linguistiquement neutre, résultant de l'application d'un processus d'identification à base de règles, utilisée pour identifier un objet. Exemple : les caractères numériques « 124 » constituent l'identifiant de « Canada » dans la norme ISO 3166-1 1997, <i>Codes pour la représentation des noms de pays et de leurs subdivisions</i> – Partie 1 : <i>Codes pays</i>.</p> <p>Un identifiant est unique dans le système d'identification de l'autorité qui le délivre (définition adaptée de la norme ISO/CEI 15944-1).</p> <p>Un URI ne répond pas complètement à la définition de l'Organisation internationale de normalisation : tous les URI sont certes uniques et non ambigus et répondent à une identification à base de règles (le schéma de l'URI), mais seuls les URN, parmi les URI, sont permanents, et aucun URI n'est linguistiquement neutre. C'est que l'URI a le double rôle de permettre d'identifier et de représenter la ressource à l'utilisateur. Ainsi, l'URI est souvent exprimé dans la langue de celle ou celui qui l'a conçu.</p>

*Secrétariat
du Conseil du trésor*

Québec

