

ATELIER FORMATION MCS

Vérification de la conformité sémantique des documents médicaux CDAr2

Thierry DART

21 avril 2016

ATELIER FORMATION MCS

Vérification de la conformité sémantique des documents médicaux CDAr2

21/04/2016

Sommaire

A. Pourquoi ?

B. Comment ?

C. Cas pratiques : testContenuCDA



CHAPITRE A

Pourquoi ?



Clinical Document Architecture release 2 (CDA R2)

Dialecte XML développé par l'organisation Health Level 7 International (HL7) et reconnu par l'ISO (*ISO/HL7 27932:2009 Data Exchange Standards -- HL7 Clinical Document Architecture, Release 2*) dédié aux documents de santé électroniques véhiculant les données médicales

- **Un modèle s'appuyant sur:**

- un modèle d'information du message restreint R-MIM (Restricted Message Information Model) dérivé du modèle d'information de référence RIM HL7 (Reference Information Model)

- **Un schéma xml unique: CDA.xsd**

- Validation de tout document XML conforme au standard
- base des modèles des documents de santé spécifiés dans les profils IHE et dans les spécifications d'interopérabilité du CI-SIS:
 - ✓ ensemble de contraintes appliquées au standard CDA
 - sur la **structure** et sur les **vocabulaires codés** employés
 - validées à l'aide d'un **schématron** (ISO/IEC 19757-3)

Pourquoi contrôler les contenus ?

- **S'assurer que les documents CDA:**
- **sont visualisables correctement**
 - Conformité sémantique (structure et jeux de valeurs) au volet « Structuration minimale »
- **possèdent des métadonnées cohérentes par rapport à leur contenu**
 - Conformité sémantique au volet « Structuration minimale »
- **s'ils sont structurés en niveau 3, comportent des données structurées exploitables correctement par les LPS ayant implémenté le modèle dont ils se réclament**
 - Conformité sémantique (structure et jeux de valeurs) au volet « Structuration minimale »
 - Conformité sémantique au volet de contenu spécifiant le modèle de document structuré



CHAPITRE B

Comment ?

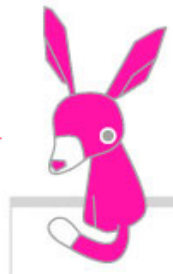


Deux outils complémentaires pour vérifier la conformité



Un seul schéma **CDA.xsd** commun à tous les documents

- ne vérifie que la conformité au standard
- analyse grammaticale : respect de la structure et de quelques jeux de valeurs embarqués



Des schématrons par modèle de document

- vérifient la présence de configurations de branches dans le document
- règles de gestion, sémantique externe, co-contraintes
- contrôles inter-branches, inter-documents
- basés sur des expressions xpath 2

« Cascade » pour vérifier la conformité

Le document est-il un document XML bien formé ?

Éditeur XML, parseur

Le document est-il valide pour le dialecte CDAR2 ?

Un schéma XML unique: « CDA.xsd »

Le document est-il conforme au « volet structuration minimale » ?

Un schématron commun à tous les volets de contenu :
« CI-SIS_StructurationCommuneCDAR2.sch »

Le document est-il conforme au volet de contenu
dont il se réclame ?

Des schématrons métier par volet de contenu : « CI-SIS_[Volet]_[Fiche].sch »

- Modèles de section et d'entrée (templated, code , modèles d'entrée utilisés par section, ...)
- Cardinalités des sections et des entrées
- Jeux de valeurs

Attention: certaines règles métier du volet de contenu ne peuvent être vérifiées par ces méthodes

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="cda.xsl" />
<ClinicalDocument xmlns="urn:hl7-org:v3"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="urn:hl7-org:v3:cda.xsd"
  <realmCode code="FR"/>
  <typeId root="2.16.840.1.113883.1.3" />
  ...
```



Cadre d'interopérabilité des SIS
Couche Contenu
Volet Compte Rendu d'Examens de Biologie Médicale

Identification du document	
Référence	CI-SIS_CONTENU_VOLET-OR-BIOLOGIE_V1.2.0.0.Docx
Date de création	07/10/2009

Exemple 1 : *date&heure de production du document*

ClinicalDocument/effectiveTime

CDA.xsd vérifie

- que l'élément obligatoire <effectiveTime> est présent juste devant l'élément <confidentialityCode>, lui aussi obligatoire ;
- est vide ou contient une chaîne de chiffres avec une éventuelle time zone, ou contient un nullFlavor.

Le schématron CI-SIS_StructurationCommune.sch vérifie

- que la date et l'heure sont cohérentes (pas de mois > 12, de jour > 31, d'heure > 24) ;
- que si une heure est présente, la timezone contient exactement 4 chiffres ;
- que l'attribut nullFlavor n'est pas utilisé.

```
<effectiveTime value="20081988094914+0100"/>
```

```
<effectiveTime nullFlavor="MSK"/>
```

```
<effectiveTime/>
```

3 exemples conformes au schéma CDA.xsd mais non conformes au schématron CI-SIS_StructurationCommune.sch

Exemple 2 : corps non xml encapsulant une pièce encodée en base 64

nonXMLBody/text

CDA.xsd vérifie

- que l'élément <nonXMLBody> s'il est présent, possède un élément fils <text> obligatoire.

Le schématron **CI-SIS_StructurationCommuneCDAr2.sch** vérifie que <text> possède deux attributs :

- mediaType : {image/jpeg, image/tiff, text/rtf, text/plain, application/pdf}
- representation="B64"

```
<text mediaType="text/doc" representation="B64">
```

```
<text mediaType="application/pdf" representation="B6">
```

```
<text>
```

3 exemples conformes au schéma **CDA.xsd** mais non conformes au schématron **CI-SIS_StructurationCommune.sch**

Exemple 3 : valeurs conformes des éléments sources des métadonnées

```
author/assignedAuthor/code
>authorSpecialty
```

```
documentationOf/.../standardIndustryClassCode
>practiceSettingCode
```

CDA.xsd vérifie que ces éléments optionnels, s'ils existent, sont à la bonne place.

Le schématron **CI-SIS_StructurationCommuneCDAr2s.sch** vérifie qu'ils sont renseignés avec une valeur appartenant bien au jeu de valeur défini.

```
<standardIndustryClassCode code="AMBULATOYRE"
  displayName="Ambulatoire"
  codeSystem="1.2.250.1.213.1.1.4.9"/>
```

```
<standardIndustryClassCode nullFlavor=" NAV "/>
```

```
<assignedAuthor>
  <id root="1.2.250.1.71.4.2.1" extension="801234567897"/>
  <code code="G15_10/C019"
    codeSystem="1.2.250.1.213.1.1.4.5"/>
```

3 exemples conformes au schéma **CDA.xsd** mais non conformes au schématron **CI-SIS_StructurationCommune.sch**

Qu'est-ce qu'un schématron ?

Sémantiquement :

- un ensemble de « **patterns** » à activer
- chaque pattern est une liste de **règles** dont chacune est associée à un contexte, c'est-à-dire à une configuration de branche (structurelle et sémantique) du document à contrôler
- chaque règle contextuelle contient une liste d'**assertions** dont chacune va contrôler la présence et la valeur d'éléments et d'attributs dans ou autour de la branche qui a déclenché la règle, voire par rapport à d'autres éléments d'un document extérieur.



À la différence du schéma XML, le schématron ne fait pas une analyse grammaticale, mais une succession de contrôles ciblés :

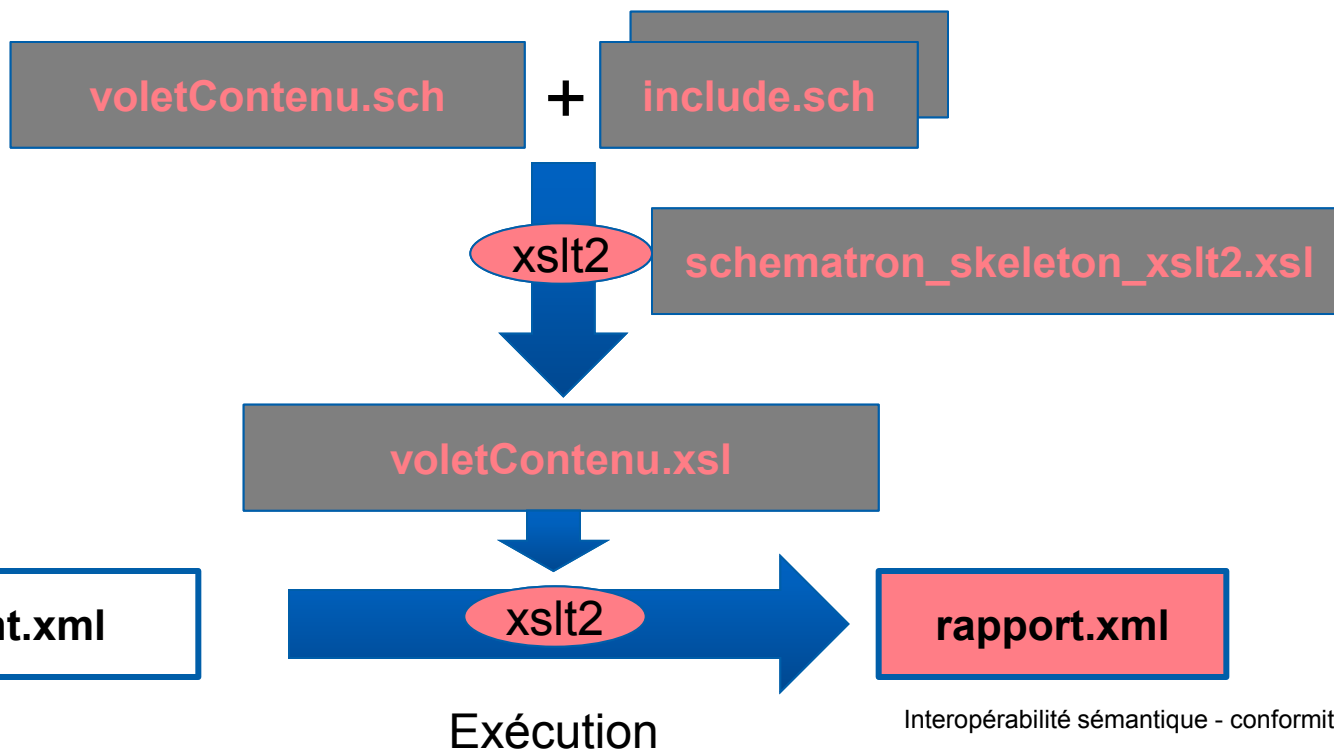
- de contenus (attributs/éléments)
- de cohérence inter-branches
- de cohérence inter-documents

Qu'est-ce qu'un schématron ?

Concrètement :



- un fichier xml avec une extension .sch, et un ensemble de sous-fichiers .sch à inclure.
- suivant la norme ISO/IEC 19757-3 (en accès libre) www.schematron.com
- compilables en une feuille de transformation XSLT 2
- dont l'application au document à vérifier produit un rapport SVRL, listant ses non-conformités.



Moteurs de schématron

Inclus dans les ateliers XML : XMLspy, oXygen ...

- exécution interactive

Autonome : SAXON

- exécutable depuis la ligne de commande, ou depuis une appli via API
- pour plateformes Java (Linux, Windows, MacOSx ...) et .NET

•composition :

- un processeur XSLT 2.0
- un processeur XPath 2.0
- un processeur Xquery 1.0
- un processeur de schéma XSD 1.0 et 1.1
- des librairies d'extensions (date time, math, ...)

•3 conditionnements disponibles :

- Saxon-HE : open source → inclus dans la diffusion du CI-SIS
- Saxon-PE (professional edition)
- Saxon-EE (enterprise edition)



Les 3 éléments fondamentaux du standard schematron

- **<pattern>** : un ensemble identifié et activable de règles contextuelles
- **<rule>** : une règle contextuelle (déclenchée par la reconnaissance d'un contexte dans le document analysé) comportant un ensemble d'assertions sur la branche sélectionnée
- **<assert>** : la vérification d'une contrainte sur la branche sélectionnée, et le message d'erreur à déclencher dans le rapport si la contrainte n'est pas vérifiée

```
<pattern xmlns="http://purl.oclc.org/dsdl/schematron" id="patientId">
  <p>
    Vérification de la conformité au CI-SIS :
    l'INS-C doit être une chaîne de 22 chiffres
  </p>
  <rule context="cda:ClinicalDocument/cda:recordTarget/cda:patientRole/cda:id">
    <assert test="
      (@root = $OIDINS-c and string-length(@extension) = 22 and number(@extension) > 1)
      or (@root != $OIDINS-c)">
      Erreur de conformité CI-SIS : L'INS-c doit contenir une chaîne de 22 chiffres
      (valeur trouvée dans le document : <value-of select="./cda:id/@extension"/>)
    </assert>
  </rule>
</pattern>
```

L'élément `<phase>` sélectionne les `<pattern>` à activer

```

<schema xmlns="http://purl.oclc.org/dsdl/schematron"
  defaultPhase="latotale"
  xmlns:cda="urn:hl7-org:v3" queryBinding="xslt2"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  schemaVersion="CI-SIS_StructurationCommuneCDAr2.sch">

  <phase id="latotale">
    <active pattern="variables" />
    <p>Vérification complète de la conformité au CI-SIS</p>
    <active pattern="addr" />
    <active pattern="administrativeGenderCode" />
    <active pattern="assignedEntity" />
    <active pattern="authenticatorName" />
    <active pattern="authorPersonName" />
    <active pattern="authorSpecialty" />
    <active pattern="authorTime" />
    <active pattern="documentCode" />
    <active pattern="documentEffectiveTime" />
    <active pattern="healthcareFacilityTypeCode" />
    <active pattern="informantAssignedPersonName" />
  
```



Jeux de valeurs (SVS.xsd)

● CI-SIS_jdv_practiceSettingCode.xml ×

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <RetrieveValueSetResponse xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3   xsi:schemaLocation="urn:ihe:iti:svs:2008 SVS.xsd  xmlns="urn:ihe:iti:svs:2008">
4   <ValueSet id="1.2.250.1.213.1.1.5.4" displayName="practiceSettingCode" version="20101116">
5     <ConceptList>
6       <Concept code="AMBULATOIRE" displayName="Ambulatoire" codeSystem="1.2.250.1.213.1.1.4.9" />
7       <Concept code="DEPISTAGE" displayName="Dépistage" codeSystem="1.2.250.1.213.1.1.4.9" />
8       <Concept code="MAD" displayName="Maintien à domicile" codeSystem="1.2.250.1.213.1.1.4.9" />
9       <Concept code="SAD" displayName="Soins à domicile" codeSystem="1.2.250.1.213.1.1.4.9" />
10      <Concept code="HAD" displayName="Hospitalisation à domicile" codeSystem="1.2.250.1.213.1.1.4.9" />
11      <Concept code="ETABLISSEMENT" displayName="Etablissement de santé" codeSystem="1.2.250.1.213.1.1.4.9" />
12      <Concept code="PALLIATIF" displayName="Soins palliatifs" codeSystem="1.2.250.1.213.1.1.4.9" />
13      <Concept code="SAMU_SMUR" displayName="SAMU/SMUR" codeSystem="1.2.250.1.213.1.1.4.9" />
14      <Concept code="EXP_PATIENT" displayName="Expression personnelle du patient" codeSystem="1.2.250.1.213.1.1.4.6" />
15    </ConceptList>
16  </ValueSet>
17 </RetrieveValueSetResponse>
```

Prépare l'utilisation du profil IHE **SVS** pour diffuser les jeux de valeurs

Vérification d'un élément codé par rapport à un jeu de valeurs

```

<pattern xmlns="http://purl.oclc.org/dsdl/schematron" id="dansJeuDeValeurs" abstract="true"
  xmlns:svs="urn:ihe:iti:svs:2008">
  <p>Conformité d'un élément codé obligatoire par rapport à un jeu de valeurs du CI-SIS</p>
  <rule context="$xpath_elt">
    <let name="att_code" value="@code"/>
    <let name="att_codeSystem" value="@codeSystem"/>
    <let name="att_displayName" value="@displayName"/>
    <assert test="(
      (@code and @codeSystem and @displayName) or
      ($nullFlavor and
      (@nullFlavor='UNK' or @nullFlavor='NASK' or @nullFlavor='ASKU' or @nullFlavor='NAV' or @nullFlavor='MSK'))
    )">
      Erreur de conformité CI-SIS : L'élément "<value-of select="$vue_elt"/>" doit avoir ses attributs
      @code, @codeSystem et @displayName renseignés, ou un nullFlavor autorisé.
    </assert>
    <assert test="@nullFlavor or
      (document($path_jdv)//svs:Concept[@code=$att_code and @codeSystem=$att_codeSystem])">
      Erreur de conformité CI-SIS : L'élément <value-of select="$vue_elt"/>
      [<value-of select="$att_code"/> : <value-of select="$att_displayName"/> :
      <value-of select="$att_codeSystem"/>]
      doit faire partie du jeu de valeurs <value-of select="$path_jdv"/>.
    </assert>
  </rule>
</pattern>

```



CHAPITRE C

Cas pratiques: testContenuCDA

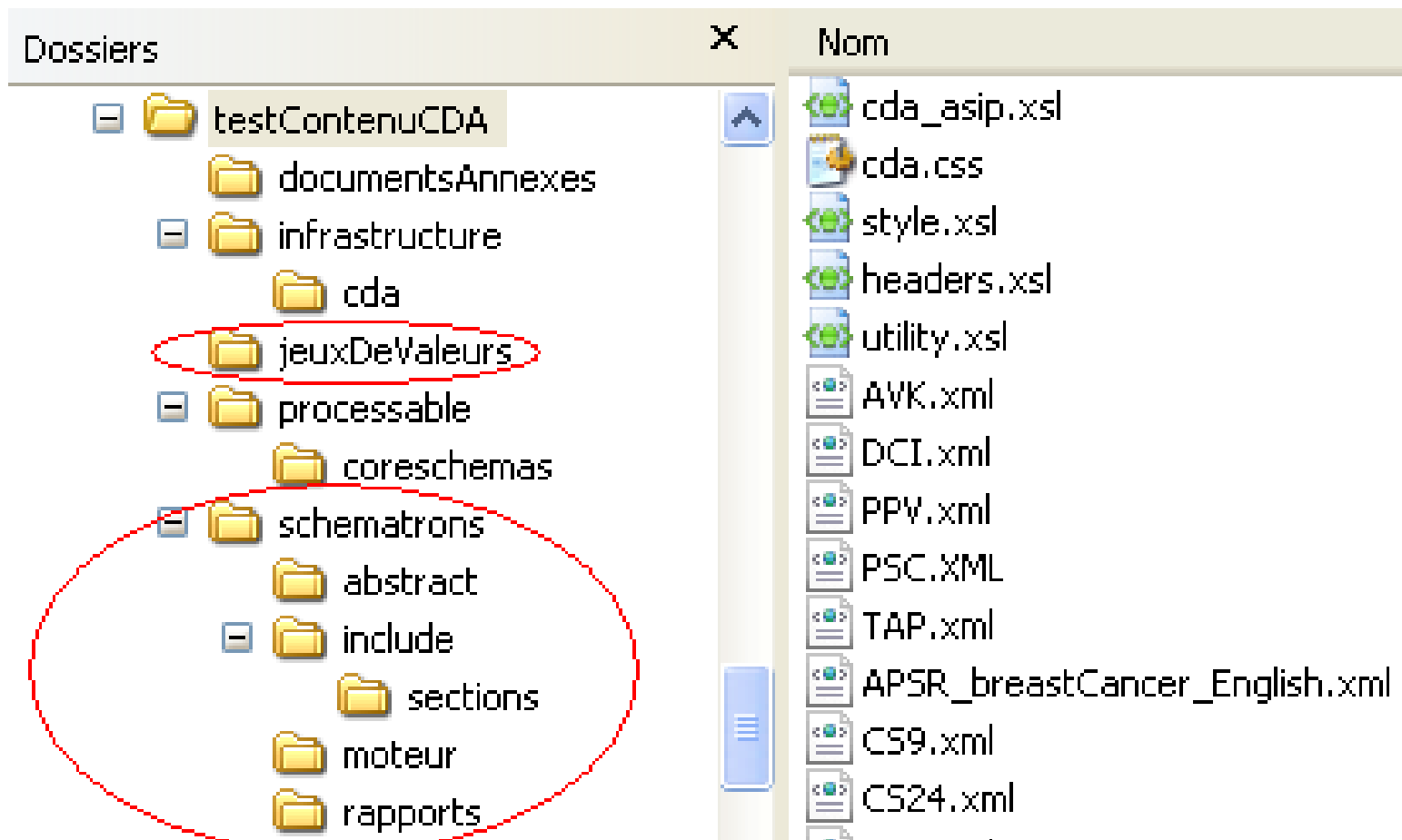


- Espace de publication CI-SIS – Couche Contenus Métiers
- <http://esante.gouv.fr/services/referentiels/ci-sis/espace-publication/contenus-metiers>
- **Annexes & outils: testcontenucda_[Date].zip**
- Exemples CDAr2
- Schéma CDA.xsd
- Schématrons
- Feuilles de style
- Jeux de valeurs

The screenshot shows the website interface for 'esante.gouv.fr Services Pro'. The main navigation bar includes 'ACTUS', 'TRIBUNES', 'LE MAG', 'PAROLE AUX RÉGIONS', 'ETUDES & RAPPORTS', and 'SERVICES'. Below this, there are links for 'Référentiels', 'Espace CPS', 'Espace DMP', 'Espace MSSanté', 'Santé connectée', 'Qualité HN', and 'Labellisation MCS'. The breadcrumb trail reads: 'Accueil > Services > Referentiels > Ci sis > Espace publication > Espace de publication CI-SIS – Couche Contenus Métiers'. The page title is 'Espace de publication CI-SIS – Couche Contenus Métiers' with a date of '06 NOV. 2015'. The main content area states: 'Cet espace présente les documents publiés pour la couche Contenus Métiers du CI-SIS organisée en :'. A list of links follows: 'Volets communs couche contenus métiers - couche services', 'Modèle de contenu CDA', 'Volets modèles de documents médicaux', 'Volets modèles de documents administratifs', 'Autres volets de contenus métiers', and 'Annexes & outils'. The 'Annexes & outils' link is enclosed in a red rectangular box. On the right side, there are social media sharing options for Facebook, Twitter, and LinkedIn, as well as 'ENGLISH VERSION', 'ENVOYER PAR MAIL', and 'IMPRIMER LA PAGE'.

- **Contenu du fichier testContenuCDA**
- **Répertoire racine**
 - « CI-SIS_TEST-CDA-LISEZ-MOI_vxxxx.pdf »
 - La totalité des exemples de documents CDA publiés dans le CI-SIS
 - La feuille de style par défaut « cda_asip.xsl » et ses fichiers annexes
- **Répertoire « infrastructure\cda »**
 - Le schéma xml « CDA.xsd » (édition normative CDA release 2, de 2005)
- **Répertoire « schematrons »**
 - Le Schématron « CI-SIS_StructurationCommuneCDAr2.sch »
 - Conformité au volet « Structuration minimale des documents médicaux »
 - La totalité des schématrons des volets de la couche Contenus Métiers
 - « CI-SIS_[Volet]_[Fiche].sch »
- **Répertoire « schematrons/moteur »**
 - Le moteur xslt2 open source écrit en Java : « saxon9he.jar »
 - Les scripts « verif.bat » et « compilverif.bat » qui doivent être lancés depuis le répertoire courant
- **Répertoire « schematrons/rapports »**
 - Les rapports de vérification des exemples de documents par rapport aux schématrons
 - Chaque rapport est au format SVRL (Schematron Validation Report Language)

Schématrons avec environnement d'exécution



testContenuCDA

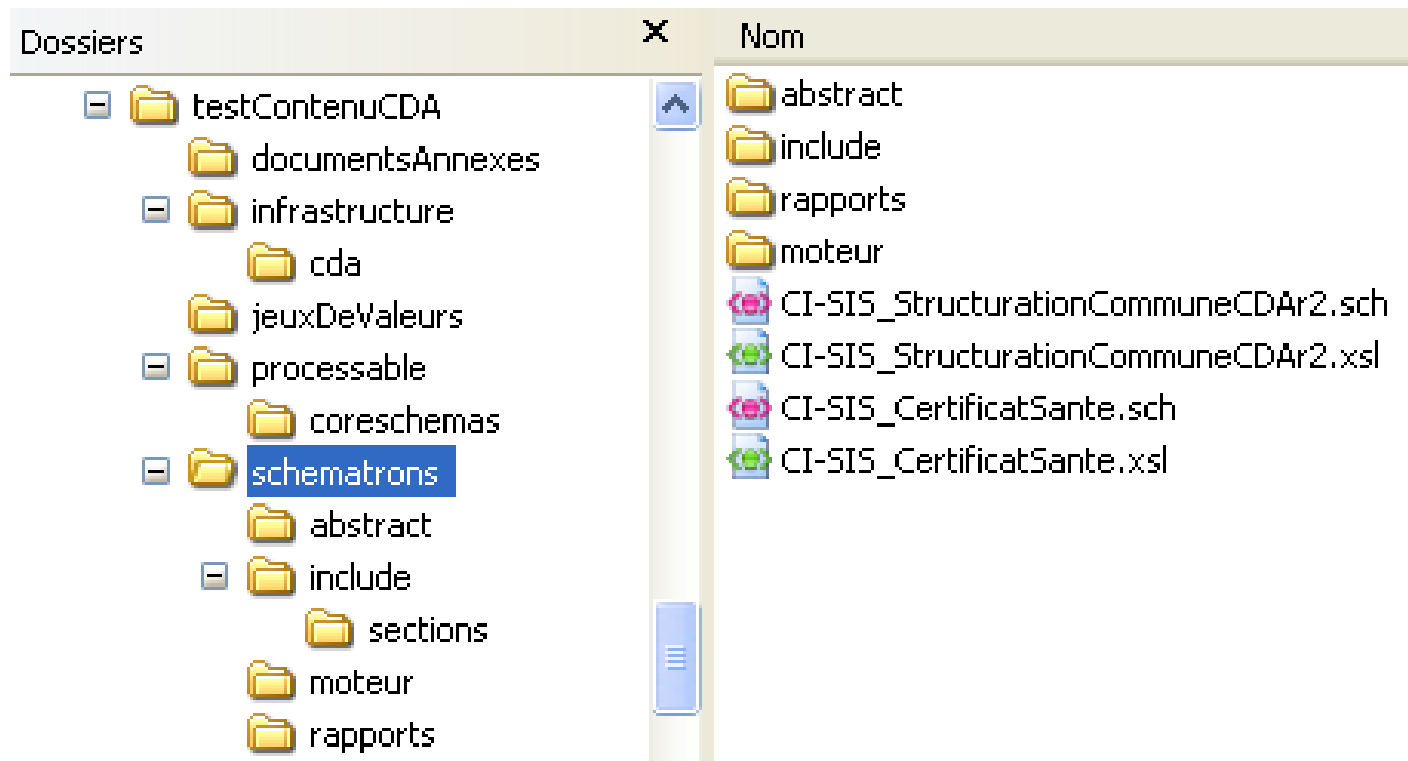
Jeux de valeurs livrés en xml (SVS.xsd)

The image shows a file explorer window with two panes. The left pane, titled 'Dossiers', displays a tree view of folders: 'testContenuCDA' (expanded), 'documentsAnnexes', 'infrastructure' (expanded), 'cda', and 'jeuxDeValeurs' (selected). The right pane, titled 'Nom', lists the files in the selected folder: 'SVS.xsd', 'CI-SIS_jdv_practiceSettingCode.xml', 'CI-SIS_jdv_healthcareFacilityTypeCode.xml', 'CI-SIS_jdv_authorSpecialty.xml', and 'CI-SIS_jdv_typeCode.xml'.

Dossiers	Nom
testContenuCDA	SVS.xsd
documentsAnnexes	CI-SIS_jdv_practiceSettingCode.xml
infrastructure	CI-SIS_jdv_healthcareFacilityTypeCode.xml
cda	CI-SIS_jdv_authorSpecialty.xml
jeuxDeValeurs	CI-SIS_jdv_typeCode.xml

schématron du volet « Structuration Minimale des docs médicaux »

schématrons des volets de contenu (CDA niveau 3)



Conformité d'un document CDAr2 au volet Structuration Commune des Documents Médicaux du CI-SIS

Document analysé : file:///C:/Documents%20and%20Settings/fmacary/Mes%20documents/testContenuCDA/schematrons/moteur/././EXPPAT_bad.xml

Schematron ISO : [CI-SIS_StructurationCommuneCDAr2.sch](#)

Exécution : 19/6/2011 à 17:56:27 (temps UTC+02:00)

Phase de contrôle : "latotale" ; 26 patterns activés ; 42 règles exécutées ; **3 anomalie(s) détectée(s) détaillée(s) ci-dessous :**

- Anomalie détectée dans [/ClinicalDocument/effectiveTime](#)
 Erreur de conformité CI-SIS : ClinicalDocument/effectiveTime/@value = "20111302114207+0100" contient une date et heure invalide, différente de aaaa ou aaaamm ou aaaammjj ou aaaammjjhh[mm[ss]][+/-]zzzz en temps local du producteur.
- Anomalie détectée dans [/ClinicalDocument/component/nonXMLBody](#)
 Erreur de conformité CDAr2 : Un document avec un corps non xml doit encapsuler en format base64 son contenu dans l'élément text, avec @representation = "B64"
- Anomalie détectée dans [/ClinicalDocument/component/nonXMLBody](#)
 Erreur de conformité CI-SIS : Un document avec un corps non xml doit encapsuler en format base64 son contenu dans l'élément text, avec @mediaType devant prendre l'une de ces valeurs : {"text/plain", "application/pdf", "image/jpeg", "image/tiff", "text/rtf"}.



Conformité d'un document CDAr2 au volet Structuration Commune des Documents Médicaux du CI-SIS

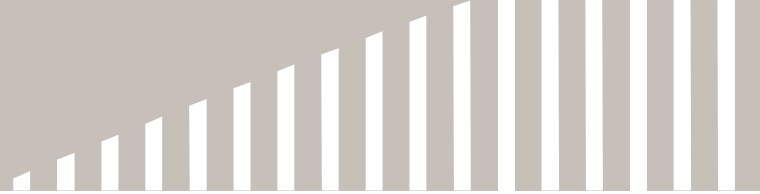
Document analysé : file:///C:/Documents%20and%20Settings/fmacary/Mes%20documents/testContenuCDA/schematrons/moteur/././EXPPAT1_exemple_20110302.xml

Schematron ISO : [CI-SIS_StructurationCommuneCDAr2.sch](#)

Exécution : 19/6/2011 à 17:59:46 (temps UTC+02:00)

Phase de contrôle : "latotale" ; 26 patterns activés ; 42 règles exécutées ; **document conforme et sans anomalie**

- **Pré-requis**
- **Dézipper le fichier testContenuCDA_[date].zip**
- **Installer un environnement Java**
- **Validation d'un CDA**
- **Ouvrir une invite de commande et se placer dans le répertoire « schematrons/moteur »**
- **Lancer la vérification de la conformité au volet de structuration minimale**
- « [compil]verif CR_BIO_Electrophorese »
 - Produit deux fichiers dans le répertoire « schematrons\rapports »
 - « CR_BIO_Electrophorese_validCDA.xml »: validation du CDA par le schéma CDA.xsd
 - « CR_BIO_Electrophorese_verif.xml »: conformité au volet (schématron)
- **Lancer la vérification de la conformité au volet métier**
- « [compil]verif CR_BIO_Electrophorese CI-SIS_CRBiologie »
 - Produit deux fichiers dans le répertoire « schematrons\rapports »
 - « CR_BIO_Electrophorese_validCDA.xml »: validation du CDA par le schéma CDA.xsd
 - « CR_BIO_Electrophorese_verif.xml »: conformité au volet (schématron)



- **Schématrons**
- **Structuration minimale**
 - CI-SIS_StructurationCommuneCDAr2.sch
- **BIO – Compte rendu d’Examens de Biologie Médicale**
 - CI-SIS_CRBiologie.sch
- **VSM – Volet de Synthèse Médicale**
 - CI-SIS_VSM_CDA_n3

CI-SIS@SANTE.GOUV.FR

Merci