

MOBS

Sistema di interscambio dati Mobilità Sanitaria

Guida al linguaggio XSD



Torino, Gennaio 2010

INDICE

1 CHE COS'È XSD.....	3
2 COME SI UTILIZZA XSD.....	3
3 CHE COSA OTTENGO VALIDANDO UN FILE XML CON UN FILE XSD?.....	3
4 A QUALE SCHEDA XML SI RIFERISCE L'ERRORE IN QUESTIONE?.....	3
5 CON CHE ORDINE VENGONO EFFETTUATI I CONTROLLI XSD?.....	4
5 QUALI CONTROLLI PERMETTE IL LINGUAGGIO XSD?.....	4
0.1 VALORE OBBLIGATORIO/FACOLTATIVO.....	4
0.2 LUNGHEZZA DEL VALORE	4
0.3 CONTROLLI FORMALI.....	4
0.3.1 <i>Lista predefinita di valori</i>	4
0.3.2 <i>Verifica formale</i>	5
6 ESEMPI DI LAYOUT DEI VARI PROGRAMMI IN COMMERCIO.....	6

1 Che cos'è XSD

XSD è un linguaggio standard generalizzato per l'implementazione di controlli formali da effettuare su un messaggio/file XML per verificarne la correttezza formale.

2 Come si utilizza XSD

I controlli XSD vengono scritti e salvati all'interno di un file con estensione “.xsd” e devono seguire le direttive dettate dal W3C (ente che definisce gli standard internazionali usati per i linguaggi mirati al WEB). I programmi in commercio in grado di verificare la correttezza di un file XML tramite un file XSD sono svariati (Microsoft XMLNOTEPAD, LIQUID XML Studio, XMLPAD Professional, OXYGEN XML Editor, ecc...)

La verifica XML/XSD può anche essere effettuata all'interno dei più comuni linguaggi di programmazione (C++, Java, VisualBasic, ecc..) utilizzando le classi apposite messe a disposizione dei singoli linguaggi.

3 Che cosa ottengo validando un file XML con un file XSD?

I dati restituiti tramite validazione XSD sono standard e generalizzati indifferentemente dal programma utilizzato.

Ad ogni occorrenza di errore riscontrato il linguaggio XSD restituisce sempre e solo i seguenti valori:

- numero riga
- numero colonna
- descrizione del messaggio di errore

4 A quale scheda XML si riferisce l'errore in questione?

Il metodo tradizionale utilizzato per l'individuazione dell'errore all'interno di un file XML è quello di:

- aprire il file XML con un editor di testo
- posizionarsi alla riga e colonna descritte nel messaggio di errore

Al fine di facilitare l'operazione di individuazione dell'errore all'interno del file XML, la maggior parte dei programmi in commercio permettono di visualizzare l'elemento XML dove è avvenuto l'errore, semplicemente cliccando con il mouse sul messaggio di errore.

Una volta trovata la Scheda XML in questione come faccio a rintracciare il dato originale all'interno del mio sistema informatico?

Ogni tracciato dei Flussi Sanitari Regionali ha un concetto di CHIAVE LOGICA, ovvero un insieme di campi che identifica la scheda

SDO -> numero cartella + anno di riferimento

C = numero ricetta

C2 -> passaggio al pronto soccorso

Per poter risalire al dato originale sul proprio sistema informatico bisogna ricercare manualmente i campi CHIAVE all'interno della Scheda XML in questione. Una volta trovati si potrà definire con certezza la fonte del dato originale sul proprio sistema informatico.

E' possibile fare in modo che un programma di validazione XML/XSD mi restituisca direttamente ogni errore associato ai campi chiave in questione?

Essendo XML/XSD un linguaggio standard generalizzato, non definisce un concetto personalizzato di CHIAVE

Per fare ciò bisognerebbe scrivere un programma complesso e personalizzato tramite un linguaggio di programmazione passibile di continue modifiche ed aggiornamenti che seguano le modifiche fatte ai flussi nel tempo.

5 Con che ordine vengono effettuati i controlli XSD?

I parsificatori XML/XSD validano prima la struttura del file XML (corretta presenza, apertura e chiusura dei TAG), solo successivamente validano il contenuto dei TAG seguendo le regole descritte all'interno del file XSD.

5 Quali controlli permette il linguaggio XSD?

0.1 Valore obbligatorio/facoltativo

Definisce se l'elemento deve contenere obbligatoriamente un valore oppure se la presenza di tale valore è facoltativa.

0.2 Lunghezza del valore

Il valore dell'elemento deve rispettare una delle seguenti tipologie di lunghezza definita:

- Qualsiasi lunghezza
- Lunghezza fissa (Es: 10 caratteri)
- Lunghezza variabile (Es: da 1 a 5 caratteri)

0.3 Controlli formali

Il valore dell'elemento deve appartenere ad uno dei seguenti controlli formali:

0.3.1 Lista predefinita di valori

Il valore dell'elemento deve appartenere ad uno dei valori predefiniti in una lista.

N.B. Si sconsiglia questa scelta nel caso di liste molto lunghe (oltre 400 valori) il quanto l'utilizzo di questo controllo rischia di rallentare la validazione del file.

0.3.2 Verifica formale

Le regole di validazione formale permettono di verificare l'appartenenza di ogni singolo carattere del valore dell'elemento, ad una lista di caratteri ammessi.

In questo modo è possibile personalizzare la validazione a seconda delle esigenze.

A titolo di esempio vengono elencate alcune delle possibilità di configurare le verifiche formali:

- Accettare solo caratteri numerici (0-9)
- Accettare solo caratteri Alfabetici maiuscoli (A-Z)
- Accettare solo caratteri Alfabetici minuscoli (a-z)
- Accettare solo caratteri Alfabetici (A-Z e a-z)
- Accettare valori con il primo carattere alfabetico e dal secondo in poi alfanumerici
- Accettare valori con il primo carattere numerico, secondo carattere [spazio] e dal terzo in poi alfanumerici
- Ecc...

N.B. Con questo sistema si possono costruire controlli sui campi data per verificare che tale data appartenga ad un periodo predefinito

Esempio:

- Accettare data dal "01/01/2000" al "31/12/2010"

6 Esempi di layout dei vari programmi in commercio

OXYGEN XML Editor

The screenshot displays the Oxygen XML Editor interface. The main window shows an XML document with the following structure:

```
1 <?xml version="1.0" encoding="utf-8"?>
2 <dipr_s0002_01>
3   <scheda>
4     <tipoRecord>Z</tipoRecord>
5
6     <aslTerritoriale>202</aslTerritoriale>
7
8     <istitutoDiCura>
9       <codiceStrutturaRicovero>01000300</codiceStrutturaRicovero>
10    </istitutoDiCura>
11
12    <numeroScheda>
13      <annoRiferimento>2008</annoRiferimento>
14
15      <numeroCartella>015197</numeroCartella>
16    </numeroScheda>
17
18    <nome>PAOLO</nome>
19
20    <nome>CARRARO</nome>
21
22    <codiceFiscale>CRRPL049A42C514M</codiceFiscale>
23
24    <sesto>2</sesto>
25  </scheda>
26</dipr_s0002_01>
```

The left pane shows a tree view of the document structure. The bottom pane displays a list of validation errors:

Descrizione	Risultati	Risk
[E] [xerces] cvc-enumeration-valid: Value 'Z' is not facet-valid with respect to enumeration [A]. It must be a value from the enumeration.	DIPR	
[E] [xerces] cvc-type-3.1.3: The value 'Z' of element 'tipoRecord' is not valid.	DIPR	
[E] [xerces] cvc-enumeration-valid: Value '999' is not facet-valid with respect to enumeration [201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213]. It must be a value from the enumeration.	DIPR	
[E] [xerces] cvc-type-3.1.3: The value '999' of element 'aslTerritoriale' is not valid.	DIPR	
[E] [xerces] cvc-enumeration-valid: Value '9' is not facet-valid with respect to enumeration [1, 2]. It must be a value from the enumeration.	DIPR	
[E] [xerces] cvc-type-3.1.3: The value '9' of element 'sesto' is not valid.	DIPR	
[E] [xerces] cvc-pattern-valid: Value '89998889' is not facet-valid with respect to pattern '((([0-9]{1-9}){1-2}){0-9}){3}([0-1])(01 02 03 04 05 06 07 08 09 10 11 12)(19 20){0-9}(2))?' for type '#AnonType_dataNascita'.	DIPR	
[E] [xerces] cvc-type-3.1.3: The value '89998889' of element 'dataNascita' is not valid.	DIPR	
[E] [xerces] cvc-enumeration-valid: Value '9' is not facet-valid with respect to enumeration [1, 2, 3, 4, 5, 6]. It must be a value from the enumeration.	DIPR	
[E] [xerces] cvc-type-3.1.3: The value '9' of element 'statoCivile' is not valid.	DIPR	
[E] [xerces] cvc-enumeration-valid: Value '998' is not facet-valid with respect to enumeration [100, 201, 202, 203, 206, 209, 210, 212, 214, 215, 216, 217, 219, 220, 221, 223, 224, 225, 226, 227, 227, 229, 231, 232, 233, 233, 234, 235...]. It must be a value from the enumeration.	DIPR	
[E] [xerces] cvc-type-3.1.3: The value '998' of element 'cittadinanza' is not valid.	DIPR	
[E] [xerces] cvc-enumeration-valid: Value '888' is not facet-valid with respect to enumeration [010, 020, 030, 041, 042, 050, 060, 070, 080, 090, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 201, 202, 203, 206, 209, 209, 210, 212...]. It must be a value from the enumeration.	DIPR	
[E] [xerces] cvc-type-3.1.3: The value '888' of element 'regioneResidenza' is not valid.	DIPR	
[E] [xerces] cvc-enumeration-valid: Value '9' is not facet-valid with respect to enumeration [1, 2, 3, 4, 5, 6, 7]. It must be a value from the enumeration.	DIPR	
[E] [xerces] cvc-type-3.1.3: The value '9' of element 'istruzione' is not valid.	DIPR	
[E] [xerces] cvc-enumeration-valid: Value '8' is not facet-valid with respect to enumeration [1, 2, 3, 4, 5, 6, 7, 9]. It must be a value from the enumeration.	DIPR	
[E] [xerces] cvc-type-3.1.3: The value '8' of element 'condizioneProfessionaleAttuale' is not valid.	DIPR	
[E] [xerces] cvc-pattern-valid: Value '4600 INSS VALENCIA' is not facet-valid with respect to pattern '((A-20-9\;){1,28})?' for type '#AnonType_teamCodiceIstituzioneCompetente'.	DIPR	
[E] [xerces] cvc-type-3.1.3: The value '4600 INSS VALENCIA' of element 'teamCodiceIstituzioneCompetente' is not valid.	DIPR	

Microsoft XML-NOTEPAD

The screenshot displays the Microsoft XML Notepad interface. The main window shows an XML document with a tree view on the left and a text view on the right. The tree view shows a root element 'dipr_s0002_01' containing a 'scheda' element. The 'scheda' element has several sub-elements, including 'tipoRecord', 'aslTerritoriale', 'istitutoDiCura', 'numeroScheda', 'cognome', 'nome', 'codiceFiscale', 'sesso', 'provinciaComuneNascita', 'dataNascita', 'statoCivile', 'cittadinanza', 'regioneResidenza', 'provinciaComuneResidenza', 'aslResidenza', 'istruzione', 'condizioneProfessionaleAttuale', 'ultimaProfessione', 'attivitaEconomicaPrincipale', 'posizioneNellaProfessione', 'tipoMovimento', 'teamIdCittadino', 'teamCodiceIstituzioneCompetente', 'teamNumeroIdentificazione', and 'teamDataScadenza'. The text view shows the XML content for the 'scheda' element, with values like 'Z', '202', 'PAOLO CARRARO', 'CRRPLO49A42C514M', '2', '071020', '02011949', '2', '100', '010', '001090', '203', '7', '6', 'OPERATA', '17', '6', 'S', and '4600 INSSI VALENCIA'.

The error list at the bottom contains the following entries:

Description	File	Line	Column
L'elemento 'tipoRecord' non è valido. il valore 'Z' non è valido per il tipo di dati 'String' - Vincolo Enumeration non riuscito.	DIPR_S0002_01.xml	4	6
L'elemento 'aslTerritoriale' non è valido. il valore '999' non è valido per il tipo di dati 'String' - Vincolo Enumeration non riuscito.	DIPR_S0002_01.xml	518	6
L'elemento 'sesso' non è valido. il valore '9' non è valido per il tipo di dati 'String' - Vincolo Enumeration non riuscito.	DIPR_S0002_01.xml	1304	6
L'elemento 'dataNascita' non è valido. il valore '89998889' non è valido per il tipo di dati 'String' - Vincolo Pattern non riuscito.	DIPR_S0002_01.xml	1308	6
L'elemento 'statoCivile' non è valido. il valore '9' non è valido per il tipo di dati 'String' - Vincolo Enumeration non riuscito.	DIPR_S0002_01.xml	1310	6
L'elemento 'cittadinanza' non è valido. il valore '998' non è valido per il tipo di dati 'String' - Vincolo Enumeration non riuscito.	DIPR_S0002_01.xml	1312	6
L'elemento 'regioneResidenza' non è valido. il valore '888' non è valido per il tipo di dati 'String' - Vincolo Enumeration non riuscito.	DIPR_S0002_01.xml	1314	6
L'elemento 'istruzione' non è valido. il valore '9' non è valido per il tipo di dati 'String' - Vincolo Enumeration non riuscito.	DIPR_S0002_01.xml	1320	6
L'elemento 'condizioneProfessionaleAttuale' non è valido. il valore '8' non è valido per il tipo di dati 'String' - Vincolo Enumeration non riuscito.	DIPR_S0002_01.xml	1322	6
L'elemento 'teamCodiceIstituzioneCompetente' non è valido. il valore '4600 INSSI VALENCIA' non è valido per il tipo di dati 'String' - Vincolo Pattern non riuscito.	DIPR_S0002_01.xml	80054	6
L'elemento 'codiceFiscale' non è valido. il valore 'CRRPLO1004056' non è valido per il tipo di dati 'String' - Vincolo Pattern non riuscito.	DIPR_S0002_01.xml	83223	6

XMLPAD Professional

The screenshot shows the XMLPAD Professional interface. The main window displays the XML content of 'DIPR_S0002_01.xml'. The XML structure is as follows:

```

000002 <dipr_s0002_01 xsi:noNamespaceSchemaLocation="file:///D:/MOBS/Flusso%20S0002/DIPR_S0002_01.xsd">
000003   <scheda>
000004     <tipoRecord>Z</tipoRecord>
000005   </scheda>
000006 </dipr_s0002_01>
000007
000008 <aslTerritoriale>202</aslTerritoriale>
000009
000010 <istitutoDiCura>
000011   <codiceStrutturaRicovero>01000300</codiceStrutturaRicovero>
000012 </istitutoDiCura>
000013
000014 <numeroScheda>
000015   <annoRiferimento>2008</annoRiferimento>
000016   <numeroCartella>015197</numeroCartella>
000017 </numeroScheda>
000018
000019 <cognome>PAOLO</cognome>
000020
000021 <nome>CARRARO</nome>
000022
000023 <codiceFiscale>CRPL049A42C514M</codiceFiscale>
000024
000025 <sesso>2</sesso>
    
```

The 'Task List' window at the bottom displays the following errors:

Description	File	Line
Datatype check error for element tipoRecord{}: enumeration facet violation: the text doesn't in enum.	D:\MOBS\Flusso 50002\202\245059_xs...	4
Datatype check error for element aslTerritoriale{}: enumeration facet violation: the text doesn't in enum.	D:\MOBS\Flusso 50002\202\245059_xs...	518
Datatype check error for element sesso{}: enumeration facet violation: the text doesn't in enum.	D:\MOBS\Flusso 50002\202\245059_xs...	1304
Datatype check error for element dataNascita{}: value is invalid for datatype 'string'.	D:\MOBS\Flusso 50002\202\245059_xs...	1308
Datatype check error for element statoCivile{}: enumeration facet violation: the text doesn't in enum.	D:\MOBS\Flusso 50002\202\245059_xs...	1310
Datatype check error for element cittadinanza{}: enumeration facet violation: the text doesn't in enum.	D:\MOBS\Flusso 50002\202\245059_xs...	1312
Datatype check error for element regioneResidenza{}: enumeration facet violation: the text doesn't in enum.	D:\MOBS\Flusso 50002\202\245059_xs...	1314
Datatype check error for element istruzione{}: enumeration facet violation: the text doesn't in enum.	D:\MOBS\Flusso 50002\202\245059_xs...	1320
Datatype check error for element condizioneProfessionaleAttuale{}: enumeration facet violation: the text do...	D:\MOBS\Flusso 50002\202\245059_xs...	1322
Datatype check error for element codiceFiscale{}: value is invalid for datatype 'string'.	D:\MOBS\Flusso 50002\202\245059_xs...	83223