

Newsletter of XML4Pharma,  
Katzelbachweg 18, AT-8052 Thal, Austria  
Web : [www.XML4Pharma.com](http://www.XML4Pharma.com)  
Mail: [Info@XML4Pharma.com](mailto:Info@XML4Pharma.com)  
December 2014



**XML4Pharma is a CDISC  
Registered Solutions Provider**



### **Serious discounts on software products**

The end of the year is always a good time to buy software products that are relatively inexpensive and that make life easier. The start of a new year is always a good time to invest in software products that increase efficiency in the organization and that have a short “pay back” time.

Both is now possible with [XML4Pharma software products](#). If you order any of our software products before the end of January 2015, and pay before March 1<sup>st</sup> 2015, you get a **discount of 20%**. If you order several copies, or more than one product at the same time, the discounts can of course become considerably higher.

Many pharma companies, CROs and service providers in the clinical research space already increased their efficiency considerably by using XML4Pharma software.

For quotations for, or trial licenses of any our software products, please contact us at: [info@XML4Pharma.com](mailto:info@XML4Pharma.com).

### **FDA publishes SDTM and SEND rules**

As some of you may already have, seen, the FDA has recently published sets of rules for SDTM and SEND submissions. Each of the sets has over 300 “rules”. Unfortunately the rules come as “text descriptions” in Excel files. So they can maybe considered to be “machine readable” (as praised by some in a webinar), but what can a machine do with “rules” that are described as pure text? Even some of these “rules” cannot be considered as a real rule.

What to think of “*Variables requested by FDA in policy documents should be included in the dataset*”?

These rules will be implemented in the next version of [OpenCDISC](#), which will be released in the next few days (obviously these rules were developed in cooperation with OpenCDISC). The great disadvantage of this is that one will need to look into the source code (not an easy task) to understand how the rule has been interpreted and implemented.

In my opinion, this is not the right way to publish “rules”. Rules for standards like SDTM should be published in **machine-executable**, but still human-readable language. Therefore we have started a research project where we are re-writing these rules into XQuery, a query language (W3C standard) for querying XML documents and databases. So these rules can be tested and implemented on Dataset-XML files of SDTM submissions.

The progress of this research project can be followed on our blog “[Working with and on CDISC standards](#)”. When successful, we will probably donate these rules in machine-executable language to the FDA, and/or make them public. The great advantage will then be that **everybody is working with exactly the same rules**, without any room for different interpretations, as is unfortunately the case until now.

### **Free web services for use in clinical research**

CDISC controlled terminology (CDISC-CT) is hard to handle. CDISC-CT comes as a set of lists, with little amount of context information and not suitable for “intelligent usage” in computer systems. For example, for vital signs, if you put “mm” as unit for

systolic blood pressure in all records, your computer system will not be able to detect this error automatically because “mm” is a valid value from the “UNIT” codelist, as is “mmHg”.

Therefore, we developed a set of (free) RESTful web services that can be used from within any software program. The web services that are currently available are:

- a set of services for working with LOINC codes (--LOINC in SDTM)
- retrieve “test name” (--TEST), NCI code and CDISC definition from “test code” (--TESTCD) for SDTM and SEND test codes
- get the UCUM notation for a CDISC UNIT code (indispensable when you get information from EHRs or hospital information systems as these use UCUM)
- check whether a unit is a valid CDISC unit
- check whether a unit is a valid unit for a specific VS test code
- check whether a Vital signs "position" (VSPOS) is a correct "position" for a given (VS) test code
- check whether a Character Result/Finding in Standard Format text value (EGSTRESC) is a valid value for a given ECG Test code (EGTESTCD)

Everyone is free to call these web services from within their own software, as has already been done in some software programs. One can however also simply use them from within a web browser. The real power is of course when implemented in software.

You can find a list of the available web services with easy-to-read explanations and examples on our [application server website](#).

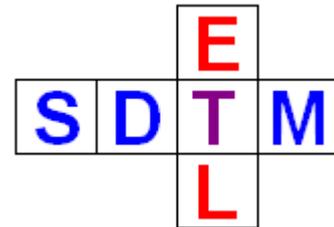
We are regularly updating these web services. In the next few days for example, we intent to add web services for obtaining information about SDTM variables, such as getting the information whether a variable is “required”, “expected” or “permissible” in context of the domain and of the SDTM version. So in future, your applications will just automatically query the service to know whether an SDTM variable is “required”, “expected” or “permissible”, instead of asking YOU to look into the SDTM-IG PDF document.

Is this the kind of services you expect from SHARE? Well, we are developing these web services in cooperation with some people at CDISC that are involved in the development of SHARE, and who closely follow what we do, as they do realize

that such services can also be implemented in SHARE in not-too-far future.

### **SDTM-ETL™ 3.1 released**

In november, we also released a new version of our popular SDTM-ETL™ software, for mapping operational data to SDTM and SEND data, with output in either SAS-XPT or Dataset-XML, and full synchronization of the underlying define.xml (1.0 or



2.0) – no need anymore to generate the define.xml seperately. Some of the new features of version 3.1 are:

- new wizards for many define.xml 2.0 elements such as the “Origin” element, the def:”WhereClause” and many other define.xml 2.0 elements
- integrated OpenCDISC 1.5 validator
- configurable “autosave” feature
- ODM: display of “repeat keys”
- extended Dataset-XML support, such as integrated “Smart Dataset-XML viewer”

But that is of course not the end of the story. We already started working on a new 3.2 version which will include:

- templates and support for SDTM 1.5
- newest version of OpenCDISC
- optional usage of web services
- and many more ...

More information is of course available at: [www.xml4pharma.com/SDTM-ETL/](http://www.xml4pharma.com/SDTM-ETL/)

### **Source code of “Smart Dataset-XML Viewer” software now freely available**

The “Smart Dataset-XML” software for inspecting, and working with CDISC Dataset-XML files (SDTM, SEND or ADaM) has now been downloaded several hundred times from the [Sourceforge website](#). The software is now stable, so I decided to also make the source code publicly available (as I promised). You can inspect and download the code when clicking the “Code” tab on the website.

As this is open-source, you can use the code as you wish, integrate it in your own code, even when it is

