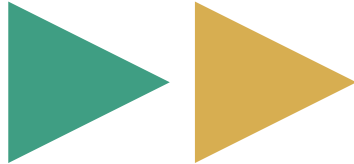


FMD K&L



Data Exchange in Clinical Trial

Chao Wang

2020-03-10

Bio & QA

- Chao Wang is the programmer manager from FMD K&L. He has about 10 years' experience on CDISC implementation. He is specialized in Define-XML, eDATA package preparation and standard development. Before FMD K&L, he worked for Pfizer and MSD about 5 years.
- Sample Questions
 - Where can I get the FDF files? Is there any specification about it?
 - What's the practice to choose the encoding when your company submit the data?

Data Exchange is Everywhere!

- Transfer study data between organizations and authorities
- Transmit data between servers and computer programs
- Translate documents between different languages
- Sample
 - CDISC (Clinical **Data Interchange** Standards Consortium)
 - Sftp
 - eCTD system

What is Data Exchange?

- Exchange information
- Three criteria
 - Structured data
 - Accurate representation
 - Shared and compatible

Annotated CRF

DM=Demographics		DS=Disposition	
Version			
Project Name:			
Form: Informed Consent			
Generated On: 11 Jun 2018 19:52:25			
DSTERM, DSDECOD = INFORMED CONSENT OBTAINED DSCAT=PROTOCOL MILESTONE			
Date Informed Consent Was Signed	DM.RFICDTC	DSSTDTC	
Protocol Version	SUPPDM.QVAL where QNAM=PROTVER		

```
67 18 0 obj
68 <</C[0.75 1 1]/Contents(DM=Demographics)/C:
69 mal;font-family:Helvetica;font-stretch:nor
70 endobj
71 19 0 obj
72 <</C[1 0.75 0.666656]/Contents(DS=Disposit:
73 l;font-family:Helvetica,sans-serif;font-st:
74 endobj
75 20 0 obj
76 <</C[0.75 1 1]/Contents(DSTERM, DSDECOD =
77 l;font-family:Helvetica,sans-serif;font-st:
78 ="text-align:center;color:#FF0000;font-fam:
79 endobj
80 21 0 obj
81 <</C[0.75 1 1]/Contents(DSSTDTC)/CreationD:
82 mal;font-family:Helvetica;font-stretch:nor
83 endobj
84 22 0 obj
85 <</C[0.75 1 1]/Contents(DS=Disposition)/Cr:
86 mal;font-family:Helvetica;font-stretch:nor
87 endobj
88 23 0 obj
89 <</C[0.75 1 1]/Contents(DSCAT=PROTOCOL MILI
90 mal;font-family:Helvetica;font-stretch:nor
91 endobj
```

- How to maintain annotation across files?
- Contents or Information?
- Annotation is reusable

Forms Data Format (FDF)

```
18 0 obj
<</C[0.75 1 1]/Contents(DM=Demographics)/CreationDate(D:20100901111330+08'00')/DA(0 0 0 rg /Helv 14
Tf)/DS(font: Helvetica 14.0pt; text-align:center; color:#FF0000 )/F
128/M(D:20191112210353-05'00')/NM(76e59e47-ff65-4315-af75-5d08bcf014cd)/Page 3/RC(<?xml
version="1.0"?><body xmlns="http://www.w3.org/1999/xhtml" xmlns:xfa="
http://www.xfa.org/schema/xfa-data/1.0/" xfa:APIVersion="Acrobat:10.0.0" xfa:spec="2.0.2"
style="font-size:14.0pt;text-align:center;color:#FF0000;font-weight:normal;font-style:nor\
mal;font-family:Helvetica;font-stretch:normal"><p
dir="ltr">DM=Demographics</p></body>)/Rect[89.8696 708.945 221.721
726.955]/Subtype/FreeText/Type/Annot>>
```

- Annotation by program
- Batch update: contents, font, font size, color, alignment, field lock
- Beyond aCRF
 - Amendment 8

SAS Transport File

- XPT v5 is a required clinical data format by regulatory
- Size limitation and uncertainty
- Limited metadata to describe the data

Clinical Data Exchange Solution

	XPT V5	XPT V8	Dataset-XML
Length of Variable Name	8	32 (stored in their original case (upper or lower))	Support long names
Length of Char Variable	200	Up to 32767	Support long text fields
Length of Var Label	40	Up to 256	Support long text
Character in Variable Name	alphanumeric and underscore	any characters other than null ('00'x).	Per CDISC Rules
Character Support	ASCII	ASCII	All Characters
Encoding Requirements	No	No	Requires stricter encoding

- Dataset-XML requires strict consistency between datasets and Define-XML.
- JSON (JavaScript Object Notation) may be another solution

Data Exchange Challenge

- Share consistent information between human and human, entities and entities, human and machine
- Interoperability: accurate, effective, security and consistent
- Three type:
 - Technical
 - Semantic
 - Process

Encoding

- How to handle °C or μ in your data package?
- &&&& or &&? (Single Byte Character Set vs. Multiple Byte Character Set)
- Transcoding message in SAS log
- Encoding translates the character sets to the codes that can be recognized by computer system
- UTF-8 is one of the most popular methods for data encoding

Translation

- How to design a database to support multiple language input?
- Controlled Terminology and Dictionary
 - CDISC CT
 - MedDRA/WHODD/SNOMED
 - LOINC
 - UNII, NDF-RT

Streaming Data Exchange

- Frequently data transfer between different entities
- SDTM is treated as a structured raw data
- Automatic mapping
- EDC API port
- Cloud computation

Dynamic Reports

- How to review a report with ten thousands pages?
- Data visualization, not only graph, but also tables
- A universal framework
- Share the source script, not the final output
- R+Shiny may be a solution

Data Exchange through blockchain

- Exchange the information, reserve data
- A new framework for modern data exchange
- Industry collaboration to discover knowledge

Security

- Data Integrity
 - HTTPS
 - SSL
 - MD5
 - Antivirus
- Information Safety

Q&A

- Thanks for your attention!
- Question and discussion

