



Add More MSG to Your Annotations

SI09 – PHUSE EU Connect

date: 2025-11-19



Authors

- Davy Baele – Team Leader, CDM Programming
davy.baele@iddi.com
- Nicolas Deravet – Expert CDM Programming Specialist
nicolas.deravet@iddi.com

Adding MSG Causes Headaches



Topics

- SDTM-MSG Challenges
- Opportunities
- Strategy
- Solution
- Results



SDTM-MSG V2.0

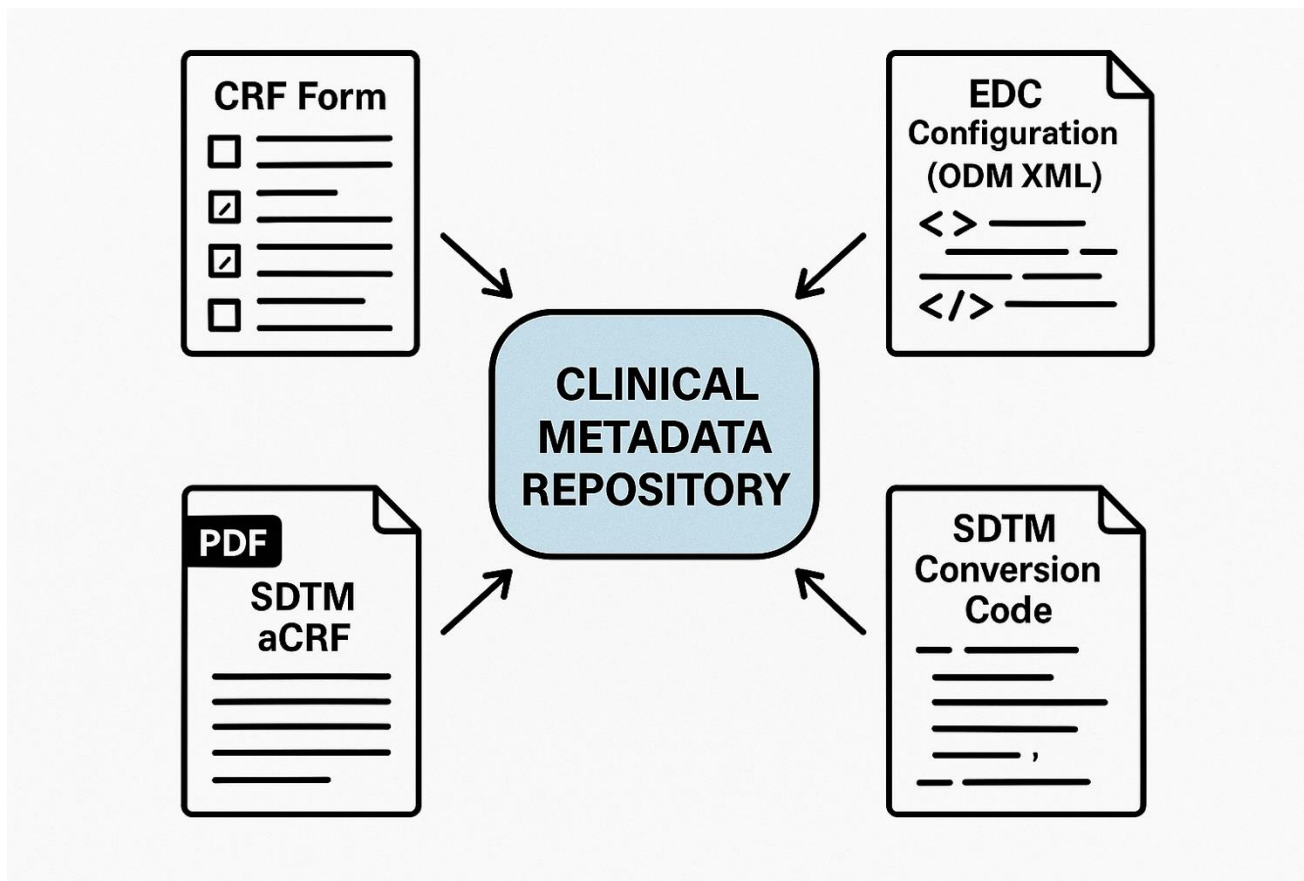
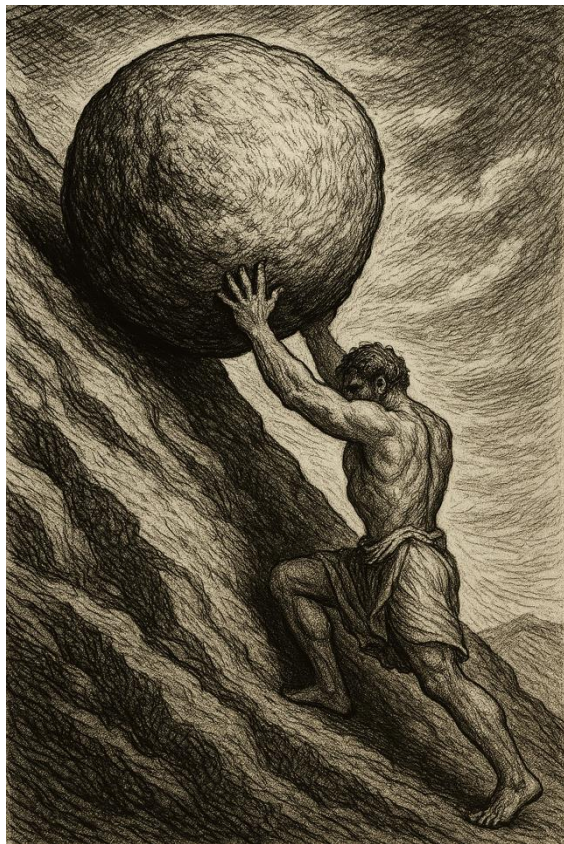


Study Data Tabulation Model Metadata Submission Guidelines (SDTM-MSG): Human Clinical Trials

Version 2.0 (Final)

Developed by the
CDISC SDS MSG Team

Traditional Approach



Outdated CRF Example

SDTM-MSG V1.0

MH=Medical History	
CDISC	Screening
Study CDISC01	MHDTC
Assessment Date: ____/____/____	
PSYCHIATRIC HISTORY MHCAT	
MHTERM	MHPRESP=Y
MHSTDTC MHOCCUR=Y	
1. Date of onset of probable Alzheimer's Disease?	
2. Date of onset of depression of Alzheimer's Disease?	

Annotation Guidelines

MH=Medical History SDTM-MSG V1.0

CDISC	Screening
Study CDISC01	MHDTCT
Assessment Date: ____/____/____	
PSYCHIATRIC HISTORY	
MHCACT	
MHTERM	MHPRESP=Y
MHSTDTC	MHOCCUR=Y
1. Date of onset of probable Alzheimer's Disease? □□ □□ □□□□	
2. Date of onset of depression of Alzheimer's Disease? □□ □□ □□□□	

BLUE	DM (Demographics)	191, 255, 255
YELLOW	DS (Disposition)	255, 255, 150
GREEN	SC (Subject Characteristics)	150, 255, 150
ORANGE	VS (Vital Signs)	255, 190, 155

SDTM-MSG V2.0

- Font should be in black
- Font size should be 12
- No italic
- Domain should be formatted as MH (Medical History)
- Variable annotations should not use bold font weight
- Dashed border should be used
- Background color OK !
- ...

Opportunities

- Most formatting rules are unambiguously defined
- Variability due non-standardized defaults
 - Font, border, borderstyles, ...
- CRO should quickly toggle between different sponsor preferences
- This enables compliance checking and autocorrection

**Problem
Solved !?**

Strategy

Itex?

Java?

FDF
with
custom
Python
code

Python?

Apache
PDFBox?

pyMUpdf?

FDF File Structure

```
%FDF-1.2
%âãïó
1 0 obj
<</FDF<</Annots[2 0 R 3 0 R 4 0 R 5 0 R 6 0 R 7 0 R 8 0 R]/F(/C/Users/xxx/Dow
endobj
2 0 obj
<</C[0.75 1 1]/Contents(MHDTC)/CreationDate(D:0000000000000000Z)/DA(0 0 0 rg /A
"><p dir="ltr">MHDTC</p></body>)/RD[0.50015 0.50015 0.50015 0.50015]/Rect[481
endobj
3 0 obj
<</C[0.75 1 1]/Contents(MHCAT)/CreationDate(D:0000000000000000Z)/DA(0 0 0 rg /A
"><p dir="ltr">MHCAT</p></body>)/RD[0.50015 0.50015 0.50015 0.50015]/Rect[372
endobj
4 0 obj
<</C[0.75 1 1]/Contents(MHSTDTC)/CreationDate(D:0000000000000000Z)/DA(0 0 0 rg
"><p dir="ltr">MHSTDTC</p></body>)/RD[0.50015 0.50015 0.50015 0.50015]/Rect[3
endobj
5 0 obj
<</C[0.75 1 1]/Contents(MHTERM)/CreationDate(D:0000000000000000Z)/DA(0 0 0 rg /
```

Python's fdf_annotations class

```
fdf_annotations.py > fdf_annotations > getrccontent
1  import re
2  import io
3
4  class fdf_annotations:
5      root_key=[]                #list containing all objects referenced from the root catalog object
6      ordered_fdf_key=[]
7      fdf_dict={}
8      bs_subobject_dict={}      #border style subobject references - referenced objects not included in root_key
9      popup_subobject_dict={}   #popup style subobject references - referenced objects included in root_key
10     parent_subobject_dict={}  #parent style subobject references - referenced objects included in root_key
11     interobjectcounter=0
12
13     def __init__(self, inputfdfpath: str) -> None:
14         #load input fdf
15         with open(inputfdfpath, "r", encoding='windows-1252') as file:
16             input_fdf_content=file.read()
17
18         #obtain header from input fdf [header: everything until first object identifier (in lookahead)]
19         headerendtag=r"\n(?:\d+ \d+ obj\n)"
20         headerend=re.search(headerendtag, input_fdf_content)
21         headertext=input_fdf_content[0:headerend.span()[0]]
22         self.fdf_dict["header"]=headertext
23         self.ordered_fdf_key.append("header")
```

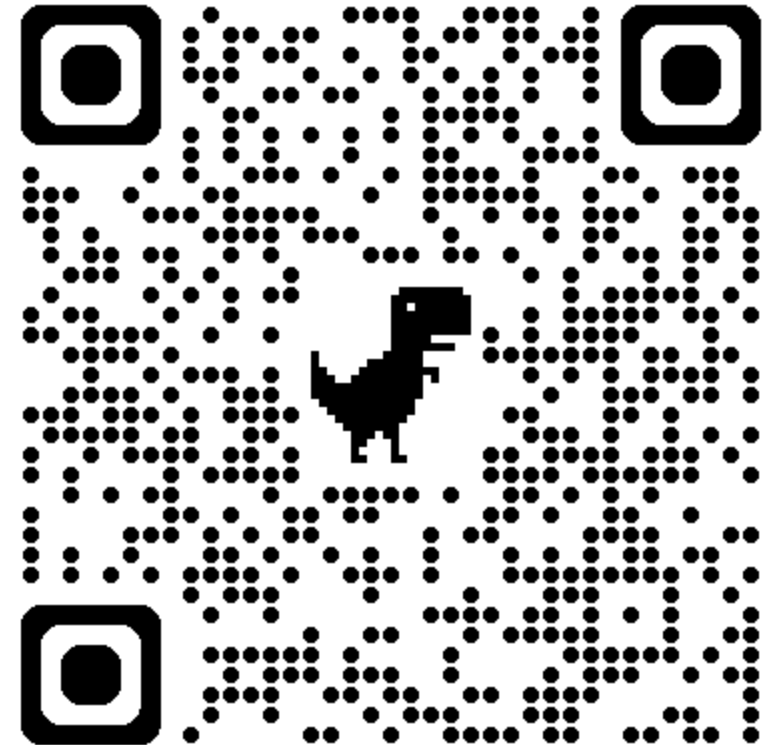
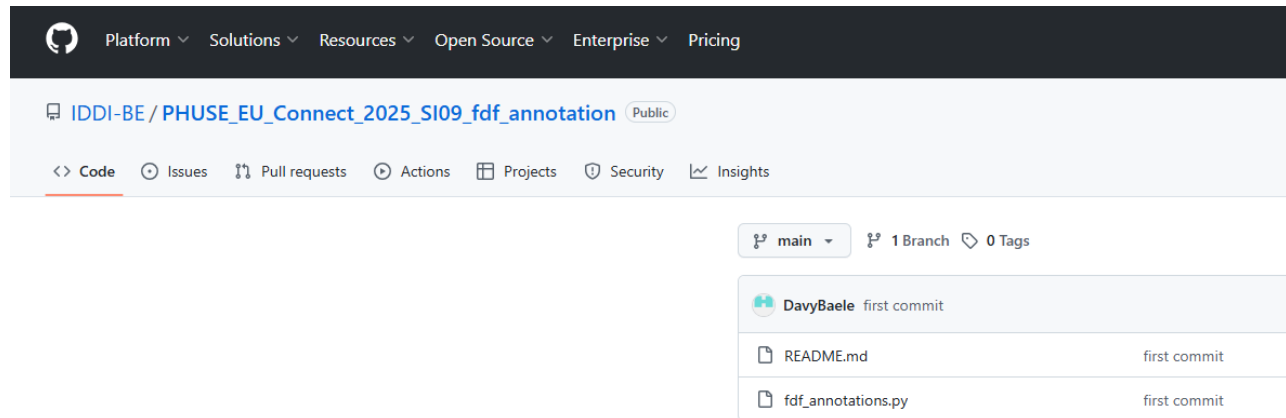
Accessing Attributes

```
def getc(self, objectid: str) -> str:
    """
    Method that returns the /C value as a string for the provided annotation id.
    This value represents the background fill within an annotation box within Text Box Comments.
    If the provided object id value is not existing in the fdf_dict it will return None.
    If the provided object id has no /C attribute it will return None.

    Input:
    |   objectid: string: object id
    Return: String containing the /C value enclosed in square brackets
    """

    if objectid not in self.fdf_dict:
        print(f"The provided object (ID= {objectid}) could not be found within the fdf.")
        return None
    ctag=r"(?<!\)/C(\[.*\])"
    cmatch=re.search(ctag, self.fdf_dict[objectid])
    if cmatch:
        return str(cmatch.group(1))
    print(f"The provided object (ID= {objectid}) has no /C attribute.")
    return None
```

Accessing the Code



https://github.com/IDDI-BE/PHUSE_EU_Connect_2025_SI09_fdf_annotation

Flexibility

```
fdf_annotations.py > ...
4 class fdf_annotations:
827     def rgb_inttohex(rgbcolorstring: str) -> str:
829         Method that translates the provided rgb colorstring integer values (0-255) to a hexad
830
831         Input: rgbcolorstr: string contaning the 3 int values (0-255) separated by a space. T
832         Output: hexadecimal string starting with a hashtag followed by the rgb components acc
833         E.g. "0 0 255" or "[0 0 255]" --> "#0000FF"
834         In case the provided rgbcolorstr does not meet the expected formatting or is emp
835
836         ""
837
838         if rgbcolorstring=="":
839             return ""
840         intrgbtag=r"^\[?\s*(\d{1,3})\s+(\d{1,3})\s+(\d{1,3})\s*\]?$" #optional square br
841         intrgbstrmatch=re.search(intrgbtag, rgbcolorstring)
842         if intrgbstrmatch:
843             return "#" + format(int(intrgbstrmatch.group(1)), "02X") + format(int(intrgbstrma
844
845
846
847     #additional methods to be added here
```

Transparency

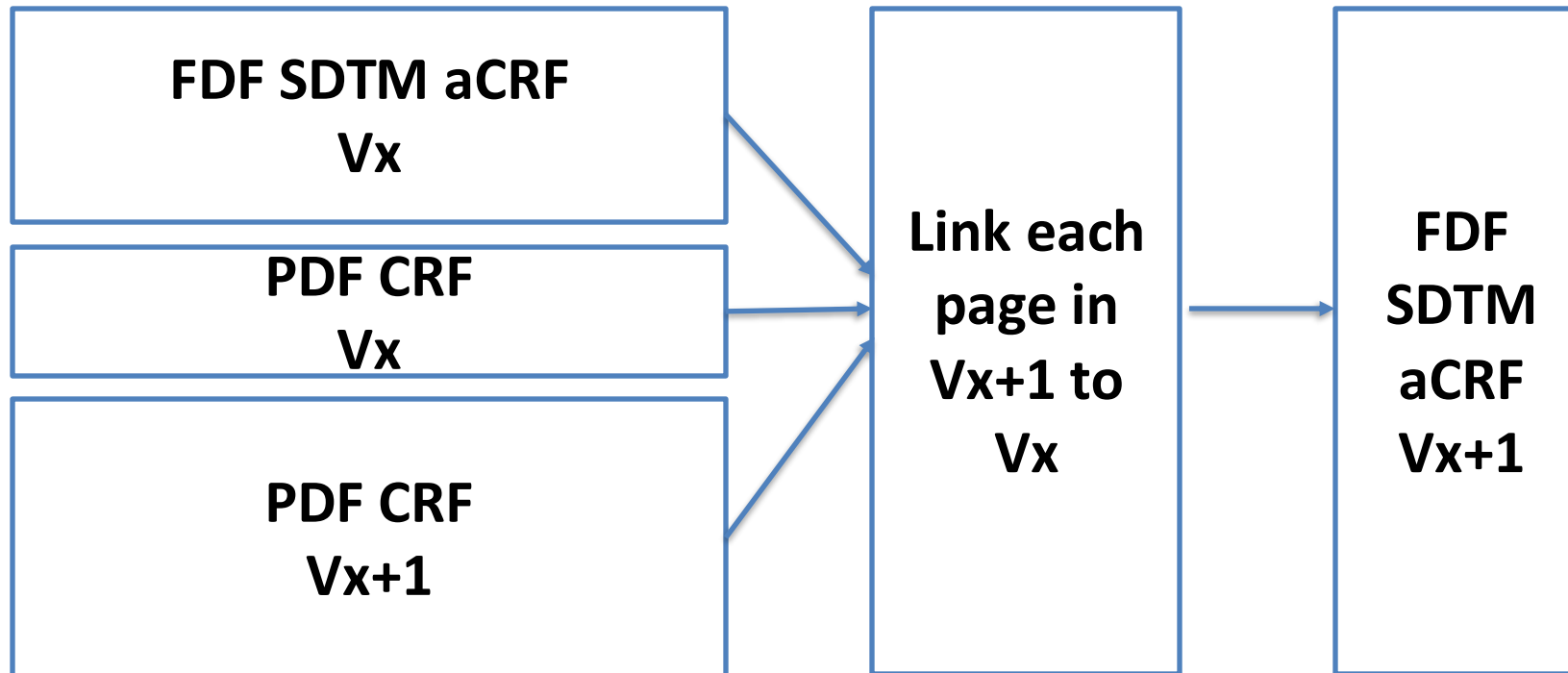
08/10/2025 08:15:15 5.247 bytes Everything Else ▾ ANSI ▾ MIX

```
%FDF-1.2
%âãĭó
1 0 obj
<</FDF<</Annots[2 0 R 3 0 R 4 0 R 5 0 R 6 0 R 7 0 R 8 0 R]/F(/C/Users/dbaele/Downlo
endobj
2 0 obj
<</C[0.75 1 1]/Contents(MHDTTC)/CreationDate(D:00000000000000Z)/DA(0 0 0 rg /Arial,B
"><p dir="ltr">MHDTTC</p></body>>)/RD[0.50015 0.50015 0.50015 0.50015]/Rect[481.187 6
endobj
3 0 obj
<</C[0.75 1 1]/Contents(MHCAT)/CreationDate(D:00000000000000Z)/DA(0 0 0 rg /Arial,B
"><p dir="ltr">MHCAT</p></body>>)/RD[0.50015 0.50015 0.50015 0.50015]/Rect[372 644.2
endobj
4 0 obj
<</C[0.75 1 1]/Contents(MHSTDTC)/CreationDate(D:00000000000000Z)/DA(0 0 0 rg /Arial
"><p dir="ltr">MHSTDTC</p></body>>)/RD[0.50015 0.50015 0.50015 0.50015]/Rect[394.937
endobj
5 0 obj
<</C[0.75 1 1]/Contents(MHTERM)/CreationDate(D:00000000000000Z)/DA(0 0 0 rg /Arial,
"><p dir="ltr">MHTERM</p></body>>)/RD[0.50015 0.50015 0.50015 0.50015]/Rect[92.1217
endobj
6 0 obj
<</C[0.75 1 1]/Contents(MH=Medical History )/CreationDate(D:00000000000000Z)/DA(0 G
p dir="ltr"><span style="text-align:center;font-family:Arial">MH=Medical History </
endobj
7 0 obj
<</C[0.75 1 1]/Contents(MHPRESP=Y)/CreationDate(D:00000000000000Z)/DA(0 0 0 rg /Ari
p dir="ltr"><span style="color:#FF0000;font-family:Arial">MHPRESP=Y</span></p></bod
```

08/10/2025 20:28:44 5.745 bytes Everything Else ▾ ANSI ▾ MIX

```
%FDF-1.2
%âãĭó
1 0 obj
<</FDF<</Annots[2 0 R 3 0 R 4 0 R 5 0 R 6 0 R 7 0 R 8 0 R]/F(/C/Users/d
endobj
2 0 obj
<</C[0.75 1 1]/Contents(MHDTTC)/CreationDate(D:00000000000000Z)/DA(0 G 0
font-family:Arial,sans-serif;font-stretch:normal"><p dir="ltr"><span st
endobj
3 0 obj
<</BS 12 0 R/C[0.75 1 1]/Contents(MHCAT)/CreationDate(D:00000000000000Z
font-family:Arial,sans-serif;font-stretch:normal"><p dir="ltr"><span st
endobj
4 0 obj
<</C[0.75 1 1]/Contents(MHSTDTC)/CreationDate(D:00000000000000Z)/DA(0 0
1;font-family:Arial;font-stretch:normal"><p dir="ltr">MHSTDTC</p></body
endobj
5 0 obj
<</BS 11 0 R/C[0.75 1 1]/Contents(MHTERM)/CreationDate(D:00000000000000
font-family:Arial;font-stretch:normal"><p dir="ltr"><span style="font-w
endobj
6 0 obj
<</C[0.75 1 1]/Contents(MH \ (Medical History \ ) )/CreationDate(D:0000000
1;font-family:Arial;font-stretch:normal"><p dir="ltr">MH \ (Medical Hist
endobj
7 0 obj
<</BS 10 0 R/C[0.75 1 1]/Contents(MHPRESP=Y)/CreationDate(D:000000000000
font-family:Arial;font-stretch:normal"><p dir="ltr"><span style="font-w
```

Annotation Propagation



Conclusion

- fdf_annotations.py
 - Verification
 - Autocorrection
 - Easily extensible

MSG Does not Cause Headaches



⋮
thank
you

USA

✉ info@iddi.com

📍 7751 Brier Creek Parkway, Suite 204
Raleigh, NC 27617

☎ +1 984 227 8599

BELGIUM

✉ info@iddi.com

📍 Avenue Provinciale, 30
1341 Louvain-la-Neuve

☎ +32 10 61 44 44

🌐 www.IDDI.com