



DataSet-JSON – Smaller than XML, But Just as Powerful

Fabian Noelle / DH03 / PHUSE EU Connect 2022

JSON – The Basics

Key/value pairs and data types

```
{
  "Image": {
    "Width": 800,
    "Height": 600,
    "Title": "View from 15th Floor",
    "Thumbnail": {
      "Url": "http://www.example.com/image/48",
      "Height": 125,
      "Width": 100
    },
    "Animated": false,
    "IDs": [116, 943, 234, 38793]
  }
}
```

- ***null***
- Boolean
true/false
- Number
integer, float, decimal
- String
“ ... ”
- Array
[... , ... , ...]
- Object
{ “key” : value , ... }

```
{
  "FileType": "Snapshot",
  "ODMVersion": "1.3.2",
  "DatasetXMLVersion": "1.0.0",
  "FileOID": "www.cdisc.org.Studycdisc01-Define-XML_2.0.0(IG.EG)",
  "PriorFileOID": "www.cdisc.org.Studycdisc01-Define-XML_2.0.0",
  "Originator": "CDISC Dataset-XML Team",
  "CreationDateTime": "2014-04-01T09:31:03",
  "xmlns": "http://www.cdisc.org/ns/odm/v1.3",
  "xmlns:xlink": "http://www.w3.org/1999/xlink",
  "xmlns:data": "http://www.cdisc.org/ns/Dataset-XML/v1.0",
  "ClinicalData": {
    "StudyOID": "cdisc01",
    "MetaDataVersionOID": "MDV.CDISC01.SDTMIG.3.1.2.SDTM.1.2",
    "IG.EG": [
      {
        "IT.STUDYID": "CDISC01",
        "IT.EG.DOMAIN": "EG",
        "IT.USUBJID": "CDISC01.100008",
        "IT.EG.EGSEQ": "1",
        "IT.EG.EGSPID": "ECG-06",
        "IT.EG.EGTESTCD": "INTP",
        "IT.EG.EGTEST": "Interpretation",
        "IT.EG.EGPOS": "SUPINE",
        "IT.EG.EGORRES": "ABNORMAL",
        "IT.EG.EGSTRESC": "ABNORMAL",
        "IT.EG.EGBLFL": "Y",
        "IT.EG.EGEVAL": "INVESTIGATOR",
        "IT.EG.VISITNUM": "1",
        "IT.EG.VISIT": "SCREEN",
        "IT.EG.VISITDY": "-13",
        "IT.EG.EDTC": "2003-04-15",
        "IT.EG.EDDY": "-14"
      },
      {
        "IT.STUDYID": "CDISC01",
        "IT.EG.DOMAIN": "EG",
        "IT.USUBJID": "CDISC01.100008",
        "IT.EG.EGSEQ": "2",
        "IT.EG.EGSPID": "ECG-06",
        "IT.EG.EGTESTCD": "INTP",
        "IT.EG.EGTEST": "Interpretation",
        "IT.EG.EGPOS": "SUPINE",
        "IT.EG.EGORRES": "ABNORMAL",
        "IT.EG.EGSTRESC": "ABNORMAL",
        "IT.EG.EGEVAL": "INVESTIGATOR",
        "IT.EG.VISITNUM": "10",
        "IT.EG.VISIT": "WEEK 24",
        "IT.EG.VISITDY": "169".
      }
    ]
  }
}
```

FileType					
ODMVersion					
DatasetXMLVersion					
FileOID					
PriorFileOID					
Originator					
CreationDateTime					
xmlns					
xmlns:xlink					
xmlns:data					
ClinicalData					
StudyOID					
MetaDataVersionOID					
IG.EG					
IT.STUDYID	IT.EG.DOMAIN	IT.USUBJID	IT.EG.EGSEQ	IT.EG.EGSPID	IT.EG.EGTESTCD
CDISC01	EG	CDISC01.100008	1	ECG-06	INTP
CDISC01	EG	CDISC01.100008	2	ECG-06	INTP
CDISC01	EG	CDISC01.100008	3	ECG-03	PRMEAN
CDISC01	EG	CDISC01.100008	4	ECG-03	PRMEAN
CDISC01	EG	CDISC01.100008	5	ECG-02	QRSBUR
CDISC01	EG	CDISC01.100008	6	ECG-02	QRSBUR
CDISC01	EG	CDISC01.100008	7	ECG-07	QTCB
CDISC01	EG	CDISC01.100008	8	ECG-07	QTCB
CDISC01	EG	CDISC01.100008	9	ECG-08	QTCF

{} ClinicalData

- FileType:Snapshot
- ODMVersion:1.3.2
- DatasetXMLVersion:1.0.0
- FileOID:www.cdisc.org.Studycdisc01-Define-XML_2.0.0(IG.EG)
- PriorFileOID:www.cdisc.org.Studycdisc01-Define-XML_2.0.0
- Originator:CDISC Dataset-XML Team
- CreationDateTime:2014-04-01T09:31:03
- xmlns:http://www.cdisc.org/ns/odm/v1.3
- xmlns:xlink:http://www.w3.org/1999/xlink
- xmlns:data:http://www.cdisc.org/ns/Dataset-XML/v1.0

{} IG.EG

{} 0

{} 1

{} 2

{} 3

{} 4

{} 5

{} 6

- IT.STUDYID:CDISC01
- IT.EG.DOMAIN:EG
- IT.USUBJID:CDISC01.100008
- IT.EG.EGSEQ:7
- IT.EG.EGSPID:ECG-07
- IT.EG.EGTESTCD:QTCB
- IT.EG.EGTEST:QTcB - Bazett's Correction Formula
- IT.EG.EGPOS:SUPINE
- IT.EG.EGSTRESC:426.2
- IT.EG.EGSTRESN:426.2
- IT.EG.EGSTRESU:msec
- IT.EG.EGBLFL:Y
- IT.EG.EGDRVFL:Y
- IT.EG.VISITNUM:1
- IT.EG.VISIT:SCREEN
- IT.EG.VISITDY:-13
- IT.EG.EDTC:2003-04-15
- IT.EG.EDDY:-14

{} 7

{} 8

http://json2table.com

GSK

20 December 2022

3



Implementation

Considerations



Datetimes / dates / times

No special data format in JSON nor XPT

XPT

- Store this data as ISO 8601 strings
- Store this data as integer counting units of time from a cut-off data, time or datetime
 - The SAS cut of date is different from the more commonly used standards
 - Formats could easily make sense of these integer values

JSON

- Store this data as ISO 8601 strings
- Store this data as integer counting units of time from a cut-off data, time or datetime
 - Just keep the ISO string that the number represents, so there is no confusion
 - ADTC for ADT etc.

Numbers

Possible representations of numbers

- 12.3
- 12.30
- 0.123E2
- 1.230e+1
- 123E-1
- 12300e-3

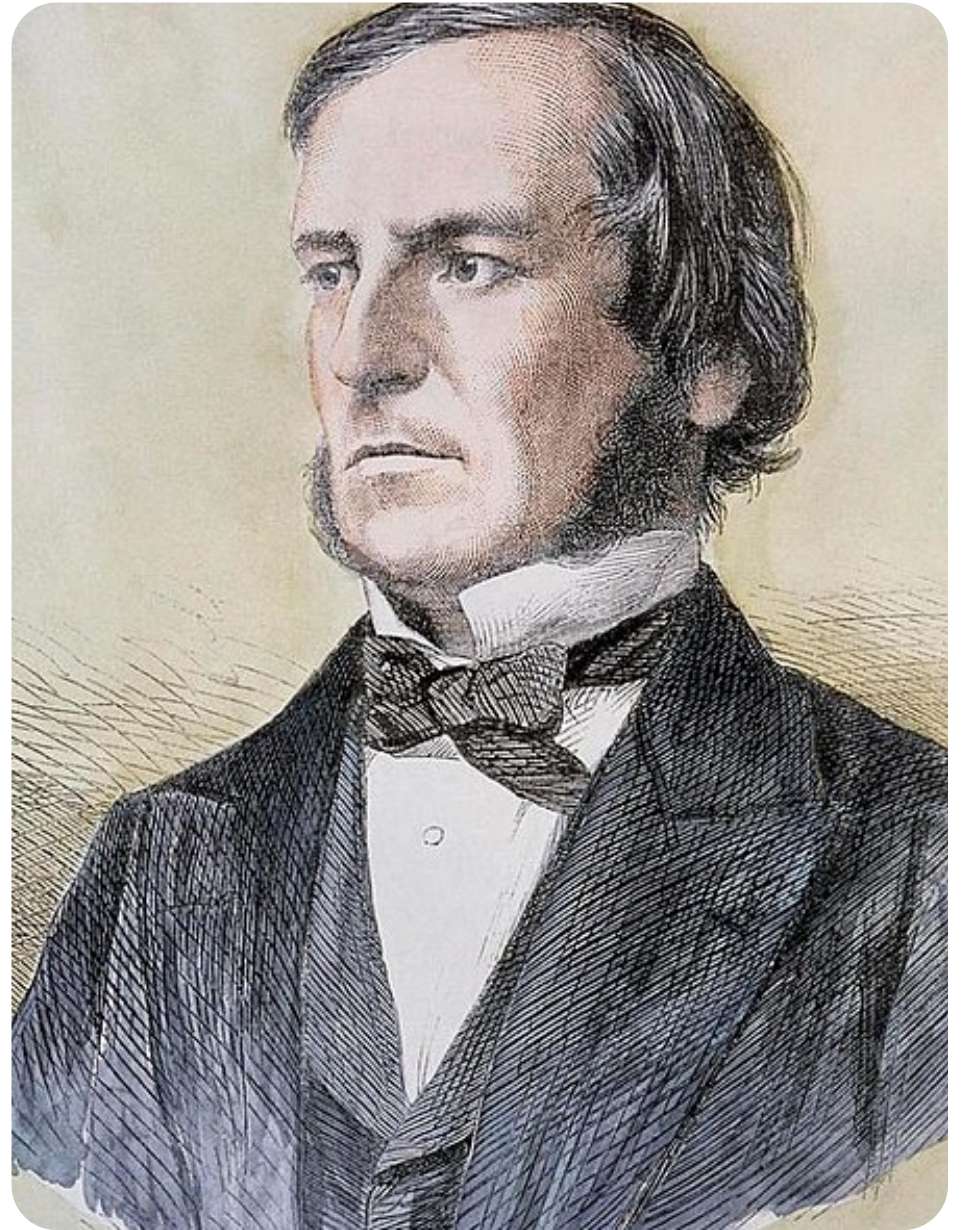
- These are all the same number.
- One might reasonably assume these two values have different precision:
 - 12.3
 - 12.30
- Different implementations might handle these as different value some as the same.
- Storing precision is currently done via metadata
 - Could also be stored in data as measurement uncertainty
 - Decimal vs float

Flags

Boolean vs No Yes Response

- Data Collection
 - Was a pregnancy test performed?
No, Yes, Not Applicable
 - Is there a family history of autoimmune diseases?
No, Yes, Unknown
- Analysis & Population Flags
 - This subject part of the safety population.
true, false

Picture of George Boole (public domain)



Whitespace

Entirely Optional

```
{ "Image": { "Width": 800, "Height": 600, "Title": "View from 15th Floor",  
  "Thumbnail": { "Url": "http://www.example.com/image/48", "Height": 125,  
    "Width": 100 }, "Animated": false, "IDs": [116, 943, 234, 38793] } }
```

DataSet-JSON

```
{
  "clinicalData": {
    "StudyOID": "cdisc01",
    "MetaDataVersionOID": "MDV.CDISC01.SDTMIG.3.1.2.SDTM.1.2",
    "itemGroupData": {
      "IG.DM": {
        "records": 5, "name": "DM", "label": "Demographics",
        "items": [
          {"OID": "IT.STUDYID", "type": "string", "name": "STUDYID", "label": "Study Identifier"},
          {"OID": "IT.USUBJID", "type": "string", "name": "USUBJID", "label": "Unique Subject Identifier"},
          {"OID": "IT.DM.AGE", "type": "integer", "name": "AGE", "label": "Age"},
          {"OID": "IT.DM.AGEU", "type": "string", "name": "AGEU", "label": "Age Unit"},
        ],
        "itemData": [
          ["CDISC01", "CDISC01.100008", 72, "YEARS"], ["CDISC01", "CDISC01.100014", 66, "YEARS"],
          ["CDISC01", "CDISC01.200001", 80, "YEARS"], ["CDISC01", "CDISC01.200002", 70, "YEARS"],
          ["CDISC01", "CDISC01.200005", 66, "YEARS"]
        ]
      }
    }
  }
}
```

A large, flowing orange shape that starts wide on the left and tapers towards the right, creating a sense of movement and depth.

Thank You

Q & A

GSK