Specifications and requirements for Universal POF Template (UPT) data format for in-line inspection of pipelines

Standard Practice

POF 110

November 2020
Amended November 2021

Foreword

This document has been reviewed and approved by the Pipeline Operators Forum (POF) and is based on knowledge and experience available from POF members and others at the date of issue. It is stated however, that neither POF nor its member companies (or their representatives) can be held responsible for the fitness for purpose, completeness, accuracy and/or application of this document.

Comments on this specification and proposals for updates may be submitted to the Administrator at specifications@pipelineoperators.org with the form which is available on the POF website (www.pipelineoperators.org).

Changes November 2021

The purpose of this revision is to comply with the new POF document numbering system. Changes consist of updated references to other POF documents. In addition, editorial corrections may have been made.

Contents

1 Introduction

This document specifies a universal data format (UPT, i.e. *Universal POF Template*) to report processed data from an in-line inspection run. It is intended as an addition to POF 100 *Specifications and requirements for in-line inspection of pipelines* [1].

2 Definitions and abbreviations

All definitions and abbreviations in this document shall be understood as per POF 100, unless otherwise specified.

3 Required input data for run comparison

Input data necessary for run comparison should take into consideration anomaly growth. In the ILI industry, two types of methods are generally used, which require very different input data: segment growth estimation and local growth estimation methods. First type of methods is based on statistics extracted on a pipe joint level and is too coarse to allow an accurate run comparison. Therefore, segment growth estimation will not be further considered in the current work. Among local growth estimation methods, generally two techniques are used [2]:

- Defect matching: employs the information on the feature box list as input data;
- Signal matching: employs the raw signal measured by the ILI tool as input data.

The defect matching technique has the drawback of the low level of detail of the box list information, which, apart from very specific cases, turns to be insufficient to fulfil the run comparison requirements listed above. As opposed, the signal matching technique uses very detailed input information, however the raw data from two runs might not be directly comparable, given that: i) raw data from different ILI technologies might not be comparable; or ii) differences might exist in the data acquisition and on-tool data processing from different ILI vendors.

Therefore, a new approach employing a universal data format (UPT) needs to be defined. That processed data format differs by technology, e.g. direct as UT and indirect as MFL based remaining wall thickness measurement. Each individual technology must be assessed towards its processing methodology. In this document, the input data format for every technology is specified, as well as some input data fields always needed to identify and document a run, and perform a run comparison.

The format being proposed is generally a data grid of, for metal loss inspections, calculated depth values that will simplify defect matching by making the export agnostic of tool sampling, circumferential sensor spacing and raw signal processing. This will provide a simple means to ensure accurate data alignment and overlay. Having the data on a defined grid will also remove typical issues of analysis boxing methodology that can lead to incorrect box-to-box matching when looking purely at box listing level.

By using a grid of depth values, it also removes risk from raw signal processing presumptions particularly around comparability of raw signals between different tools. This means the indirect measurement signals are provided as data interpreted by trained and certified analysts on a specific technology, consistent with each vendor’s competency management program. This also removes the risk of incorrect data processing (including normalisation, calibration factors, models and thresholds) being applied that are only possible by the vendors who know the influence on their defined specifications.

3.1 Data file format

Required data file format is HDF5 (Hierarchical Data Format), a versatile and free to use opensource file format supported by the HDF Group: https://www.hdf-group.org.

The overall data for one run will consist of one header file and several UPT data files, all in HDF5 format. The header file will contain all general data relative to the run and the pipe tally of the run. The UPT data files will contain the recorded data of the run.

One HDF5 UPT data file is produced for every defined segment consisting of a fixed number of odometer ticks, that number being at least ten thousand (10 000) and being specified in the header file. Each file contains data in shape of numbers, vectors and matrices that are next detailed. File name should be structured as “runName_numSegment.hdf5” and follow the convention detailed below:

Data field	Description	Number of characters
runName	Identifier of the run	6
numSegment	Number of segment following an incremental counter. Number “000000” indicates the header file. From number 1 onwards, the segments of a fixed number of ticks in which the inspected pipeline is divided, padded with “0” from left.	6

Table 1: Description of the file naming convention

Files should be written in little-endian byte order. The dimensions of each data field are indicated in the tables of the forthcoming sections, as well as the data type for each field. Data types are defined as (the latter code is the HDF5 Library native datatype when applicable):

- Short: signed 16-bits integer number; H5T_NATIVE_SHORT;
- Long: signed 32-bits integer number; H5T_NATIVE_LONG;
- Float: 32 bits single-precision floating-point format; H5T_NATIVE_FLOAT;
- Double: 64-bits double-precision floating-point format; H5T_NATIVE_DOUBLE;
- String: array of UTF-8 characters (1 byte per character).

The next sections define in detail the file header and file data of every hdf5 file. It should be noted that the first file always gathers summary data from the whole pipeline, whereas the subsequent files contain the data from consecutive segments of the pipeline (defined every fixed number of odometer ticks). Figure 1 outlines the overall data file structure.

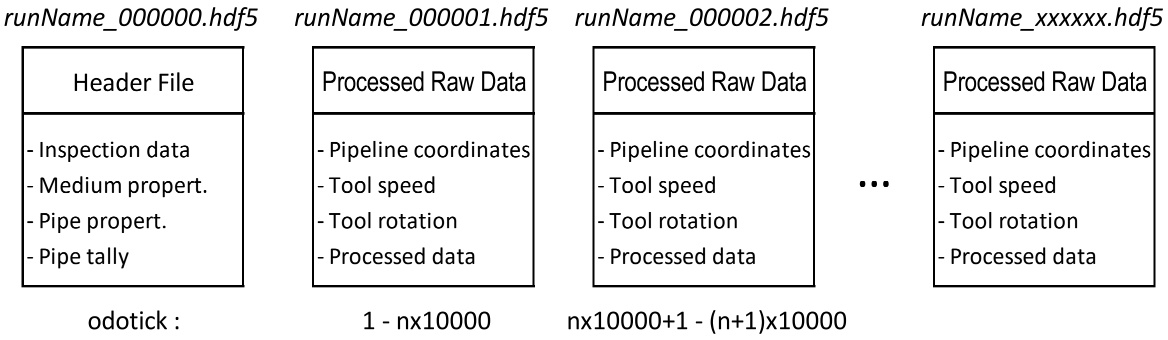

Figure 1: Data file structure

3.2 Header file

The first file should contain a set of summary data from the whole pipeline inspection, comprising the date of inspection, properties of the medium, properties of the pipe and the reported pipe tally. All those data fields are stored in a hdf5 file with the following naming convention:

<runName>_000000.hdf5

The header file should contain 3 groups: “header”, “data” and “features”.

3.2.1 The “header” group

The header group will contain the data fields listed in Table 2, stored as “attributes”. The dimensions, precision and units of every data field are also defined.

Data field	Description	Dimensions	Precision /Unit / Format / Enum
init_distance	Initial position of the given run from the origin of the pipeline.	double	0,0001 m
end_distance	Final position of the given run from the origin of the pipeline.	double	0,0001 m
data_file_size	Fixed number of ten thousands of odometer ticks use to fill data files	short	0 x 10000
date_time	Date and time (format per ISO 8601 UTC) when the inspection run started.	string	YYYY-MM-DDThh:mm:ssZ
provider	The name of the service provider who performed the inspection	string	n.a.
technology	Technology used for the pipeline inspection.	string	EMAT_METAL_LOSS, MFL, MFL-A, MFL-C, MFL-D, UT_METAL_LOSS, UT_CRACK, UT_CIRCUMFERENTIAL_CRACK, GEOMETRY, UNDEFINED
grid-or-res	Specifies if a grid type reporting will be used. In that case the two following fields define the grid mesh size.	String	GRID, RES
axial_res	Axial resolution of the inspection tool or Axial reporting grid size.	float	0.01 mm
circumf_res	Circumferential resolution of the inspection tool or Circumferential reporting grid size. Measured on outside surface of pipe wall[1].	float	0.01 mm
circumf_num	Number of reported values per circumference of pipeline.[2]	short	1
depth_res	Resolution of the depth sizing of the tool.	float	0.001 mm or 0.1% nominal wall thickness (MFL)
Specs	Any relevant auxiliary information about tool specifications as defined in chapter 4 of POF 100.	string	n.a.

Table 2: Data fields contained in the file header

3.2.2 The data group

This group will contain “attributes” and “datasets”. Content of the group is detailed below.

[1] This is to ensure that the grid is not altered due to wall thickness changes.
[2] See 3.3.2.6 for more details.

3.2.2.1 *Date of the inspection*

Date of the inspection should be provided, on a string data type. When available, the commissioning date and periods of non-operation of the pipeline should also be provided, using the same data type.

Data field	Description	Dimensions	Precision /Unit / Format / Enum
Inspec_date	Date of the inspection (format per ISO 8601 UTC)	string	YYYY-MM-DDThh:mm:ssZ
Comm_date	Commissioning date, when available (format per ISO 8601 UTC)	string	YYYY-MM-DDThh:mm:ssZ
Non-operate_date	Periods of non-operation of the pipeline, when available (format per ISO 8601 UTC)	string	YYYY-MM-DDThh:mm:ssZ

Table 3: Input data fields required about the date of the inspection

These data are stored as "attributes".

3.2.2.2 *Medium properties*

The following properties of the transported medium along the pipe are required.

Data field	Description	Dimensions	Precision /Unit
Medium	Name of the medium transported during operation at the time of the inspection	string	n.a.
Att	Attenuation of the UT wave in the medium	float	0.1 dB/m
Speed_med	Propagation speed of the UT wave in the medium	float	0.1 m/s
Density_med	Density of the medium	float	0.1 kg/m^3

Table 4: Input data fields required about the medium

[Att] and [Speed_med] will be set to Zero (0) value for non-relevant technologies (i.e. non-UT technologies)

These data are stored as "attributes".

3.2.2.3 *Pipe properties*

The following properties of the inspected pipe are required. The data field information should be provided for every single pipe joint. Therefore, the data fields should be arrays with length equal to the number of pipe joints within the pipeline (i.e. nPipeJoints). Missing data should be set to Null value.

Data field	Description	Dimensions	Precision /Unit
MAOP	Maximum Allowable Operating Pressure	float[nPipeJoints]	0.01 MPa
Nominal wall thickness	Nominal wall thickness of the pipe	float[nPipeJoints]	0.001 mm
Pipe type	Type of pipe according to the axial seam: long welded (LW), half-casqued (HC), Spiral weld (SPRL), Seamless (SML), Flexible (FLEX), etc.	string[nPipeJoints]	LW, HC, SPRL, SML, FLEX, UNDEFINED
Welding method	Type of welding method.	string[nPipeJoints]	ERW, DSAW, UNDEFINED

Data field	Description	Dimensions	Precision /Unit
Out-diam	Diameter of the outer wall of the pipe	float[nPipeJoints]	0.1 mm

Table 5: Input data fields required about the pipe

These data are stored as “datasets”.

3.2.3 The features group

This group is used to store the pipe tally.

Containing component features and anomaly features, following the prescribed terminology defined in POF 100. For every component and anomaly feature, the following data fields should be included, as defined in Appendix 2 of POF 100 (see Table 6). Additionally, to the POF 100 data fields, information about the relative weld position is also requested. All localisation data refer to the start point (S) of the feature as defined in POF 100 (figures 2.1, 2.9 or 2.10 for example).

Data field	Description	Dimensions	Precision /Unit or Enum
Feature id	Name given to the feature	string[nFeatures]	n.a.
Log distance	Axial starting point [see fig. 1 in POF 100]	double[nFeatures]	0.0001 m
Up weld dist.	Distance to upstream weld	float[nFeatures]	0.0001 m
L joint	Joint length to downstream weld	float[nFeatures]	0.0001 m
Feature type	Type of component or anomaly feature. See Appendix 2 in POF 100 for up to date enumeration of acceptable values.	string[nFeatures]	AGM, ADME, ANOD, ANOM, CRAB, CRAE, CASB, CASE, CHWT, CPCO, ESUP, ANCH, OFFT, OTHE, PFIX, MGNT, REPA, TEE, VALV, WELD, UNDEFINED
Feature identification	See Appendix 2 in POF 100 for up to date enumeration of acceptable values.	string[nFeatures]	DEBR, TMTM, OTHE, ARCS, ARTD, BUCK, CORR, COCL, CRAC, DENT, DEML, GOUG, GRIN, GWCR, GWAN, HIC, LAMI, LWCR, LWAN, OVAL, MIAN, MIAC, SCC, SPAL, SWCR, SWAN, WRIN, WSLB, WSLE, CSLB, CSLE, WDPB, WDPE, COTB, COTE, OTHB, BENB, BENE, CHDI, CHWT, ADTA, LOSE, SPSE, NISE, SMLS, UNDEFINED
Feature class	See Appendix 2 in POF 100 for up to date enumeration of acceptable values.	string[nFeatures]	AXGR, AXSL, CIGR, CISL, GENE, PINH, PITT, UNDEFINED
Clock position	Circumferential starting point [see fig. 1 in POF 100]	string[nFeatures]	hh:mm or degrees
Nominal t	Nominal wall thickness of every joint	float[nFeatures]	0.001 mm

Data field	Description	Dimensions	Precision /Unit or Enum
Reference t	The actual not diminished wall thickness surrounding a feature	float[nFeatures]	0.001 mm
Length	Anomaly length in axial direction	float[nFeatures]	0.01 mm
Width	Anomaly width in circumferential direction	float[nFeatures]	0.01 mm
d (peak)	Peak depth of the anomaly. Given by the % of remaining wall thickness, relative to Reference t. If that is not available, given in absolute terms in mm.	float[nFeatures]	0.001 mm or 1 %
d (mean)	Mean depth of the anomaly. Given by the % of remaining wall thickness, relative to Reference t. If that is not available, given in absolute terms in mm.	float[nFeatures]	0.001 mm or 1 %
Surface location	Location relative to the surface.	string[nFeatures]	INT, EXT, MID, UNDEFINED
Relative weld position	In weld, at weld, base material …	string[nFeatures]	n.a.
GPS coordinates	Northing, Easting and altitude coordinates	double[3][nFeatures]	$10^{-7°}$ / 0.01 m
ERF	Estimated Repair Factor per POF 100	float[nFeatures]	0.001
Comments	Additional comments to the feature	string[nFeatures]	n.a.

Table 6: Input data fields required about the pipe tally

3.3 Processed data

This section defines the specific processed data required depending on the technology used for ILI. In all cases, not only the data within the boxes but also the data from the complete line is requested. As defined above, one file is created every segment of a fixed number of odometer ticks. The files containing processed data should be named as:

<runName_00000X.hdf5>

where X indicates the number of segment (of a fixed number of odometer ticks). It should be noted that other data fields used to handle the processed data (i.e. pipeline coordinates, tool speed and tool rotation) are requested.

3.3.1 Data storage

The hdf5 files, one every a fixed number of odometer ticks, will be arranged in two groups:

- Group “properties” will store information on tool position and movement.
- Group “data” will store the UPT data measured by the ILI tool.

For computing reasons, processing speed and storage size, all values in the "data" group will be stored as "short", i.e. 16 bits signed integers, with values ranging from -0 to 32767.

Due to this storage specification, some of the measured values must be multiplied by a factor (10, 1000, etc.) before storage. The multiplication factor is indicated in the tables describing each set of data (column "Factor").

HDF5 is a binary format that can manage data in a compressed GZIP format. Tools using HDF5 format should be designed to be able to interpret datasets in either compressed or uncompressed format. To optimise file sizes, it is advised to store the UPT data in GZIP compressed datasets.

3.3.2 Specific data

This is the data enabling a detailed comparison of the ILI runs. The input data fields requested are dependent on the technology used to inspect the line. In the next sub-sections, the data files requested for every type of technology are defined, together with the dimensions and units requested.

3.3.2.1 UT Pipeline coordinates, tool speed and tool rotation

Circumferential and axial positions on the pipeline measured by the ILI tool along the pipe are requested. It is important to underline that those coordinates refer to location on the pipeline and not on the tool itself (i.e. the pos_circumferential field already takes into account the rotation of the tool). The number of measured axial positions within one sensor is given by nOdoticks, whereas the number circumferential positions measured in a unique axial position is given by nSensors. In addition, also the tool speed and the tool rotation are required.

Data field	Description	Dimensions	Factor	Precision /Unit
Pos_axial	Measured axial positions along the pipeline.	double[nOdoticks]	1	0.0001 m
Pos_circumferential	Measured circumferential positions. 0° is at the top with the angle information (12 o'clock). Counting is clockwise viewing the flow direction of the medium (downstream).	short[nSensors]	100	0.01 degree
Tool speed	Speed of the tool along the inspection	short[nOdoticks]	100	0.01 m/s
Tool rotation	Rotation of the tool in the circumferential axis along the inspection	short[nOdoticks]	100	0.01 degree

Table 7: Pipeline coordinates, tool speed and tool rotation

3.3.2.2 UT Metal Loss detection

Input data fields requested for UT Metal Loss detection are defined in Table 8. The required dimensions, units and precision of the data fields are also described. Not only the data within the feature boxes but also the data from the complete in-line inspection is required in this case.

Data field	Description	Dimensions	Factor	Precision /Unit
WT CScan	Measured wall thickness for every sensor and odometer tick.	short [nOdoticks, nSensors]	100	0.01 mm

SO CScan	Measured distance of the sensor to the inner wall of the pipe for every sensor and odometer tick.	short [nOdoticks, nSensors]	100	0.01 mm

Table 8: Input data fields required for UT Metal Loss detection

3.3.2.3 UT Longitudinal Crack detection

Input data fields requested for UT Longitudinal Crack detection are defined in Table 9. The required dimensions, units and precision of the data fields are also described.

Data field	Description	Dimensions	Factor	Precision /Unit
max CScan	Maximum amplitude of the reflected signals for every axial and circumferential position in the pipeline grid.	short [nOdoticks, nSensors]	10	0.1 dB
depth CScan	Depth for every axial and circumferential position inside a box. Outside the boxed area, no depth data needs to be provided.	short [nOdoticks, nSensors]	1000	0.001 mm
Reference WT	Reference wall thickness for every odometer tick.	short [nOdoticks, nSensors]	100	0.01 mm

Table 9: Input data fields required for UT Longitudinal Crack detection

3.3.2.4 UT Circumferential Crack detection

Input data fields requested for UT Circumferential Crack detection are defined in Table 10. The required dimensions, units and precision of the data fields are also described.

Data field	Description	Dimensions	Factor	Precision /Unit
max CScan	Maximum amplitude of the reflected signals for every axial and circumferential position in the pipeline grid.	short [nOdoticks, nSensors]	10	0.1 dB
depth CScan	Depth for every axial and circumferential position inside a box. Outside the boxed area, no depth data needs to be provided.	short [nOdoticks, nSensors]	1000	0.001 mm
Reference WT	Reference wall thickness for every odometer tick.	short [nOdoticks, nSensors]	100	0.01 mm

Table 10: Input data fields required for UT Circumferential Crack detection

3.3.2.5 EMAT Metal Loss detection

Input data fields requested for EMAT Metal Loss detection are defined in Table 11. The required dimensions, units and precision of the data fields are also described. Not only the data within the feature boxes but also the data from the complete in-line inspection is required in this case.

Data field	Description	Dimensions	Factor	Precision /Unit
WT CScan	Measured wall thickness for every sensor and odometer tick.	short [nOdoticks, nSensors]	100	0.01 mm
SO CScan	Measured distance from the sensor to the inner wall of the pipe for every sensor and odometer tick.	short [nOdoticks, nSensors]	100	0.01 mm
Reference WT	Reference wall thickness for every odometer tick.	short [nOdoticks, nSensors]	100	0.01 mm

Table 11: Input data fields required for EMAT Metal Loss detection

3.3.2.6 EMAT Longitudinal Crack detection

Input data fields requested for EMAT Metal Loss detection are defined in Table 12. The required dimensions, units and precision of the data fields are also described. Not only the data within the feature boxes but also the data from the complete in-line inspection is required in this case.

Data field	Description	Dimensions	Factor	Precision /Unit
max CScan	Maximum amplitude of the transmission signal for every axial and circumferential position in the pipeline grid.	short [nOdoticks, nSensors]	1000	1 dB
depth CScan	Depth for every axial and circumferential position inside a box. Outside the boxed area, no depth data needs to be provided.	short [nOdoticks, nSensors]	1000	0.001 %t
Reference t	Reference wall thickness for every odometer tick.	short [nOdoticks, nSensors]	1000	0.001 mm

Table 12: Input data fields required for EMAT Crack detection

3.3.2.7 MFL position

The individual data readings will be provided as an export to 2D grid of locations along the entire length of the pipeline. The grid dimensions are specified in the header file n=by the fields [axial_res] and [circumf_res].

A maximum of 5 mm is required for [axial_res] and [circumf_res] values. For such a value of 5 mm, the number of circumferential samples (field [circumf_num] in table 2) is set per diameter based on Table 13 below. Having a unique number of circumferential samples for a given OD pipeline ensures that the grid is evenly distributed and simplyfies alignment by keeping the samples consistent across the pipeline dataset (i.e. number of samples do not change due to wall thickness variations).

NPS	DN	OD		[circumf_num] Circumferential Samples
		(in)	(mm)	
3	80	3.5	88.90	56
4	100	4.5	114.30	72
6	150	6.625	168.28	106
8	200	8.625	219.08	138

NPS	DN	OD		[circumf_num] Circumferential Samples
		(in)	(mm)	
10	250	10.75	273.05	172
12	300	12.75	323.85	204
14	350	14	355.60	224
16	400	16	406.40	256
18	450	18	457.20	288
20	500	20	508.00	320
22	550	22	558.80	352
24	600	24	609.60	384
26	650	26	660.40	416
28	700	28	711.20	448
30	750	30	762.00	480
32	800	32	812.80	512
34	850	34	863.60	544
36	900	36	914.40	576
40	1000	40	1016.00	640
42	1050	42	1066.80	672
44	1100	44	1117.60	704
46	1150	46	1168.40	736
48	1200	48	1219.20	768
52	1300	52	1320.80	830
56	1400	56	1422.40	894

Table 13: Circumferential Grid Spacing

Data field	Description	Dimensions	Factor	Precision/Unit
Pos_axial	Measured axial position along the pipeline	Double [OdoDistance / axial_res]	1	0.0001 m
Dist_US_GW	Distance to upstream girth weld	Double [OdoDistance / axial_res]	1	0.0001 m
Pos_circumferential	Measured circumferential position clockwise (viewing downstream) from the top of pipe	Short [circumf_num]	100	0.01 degree

Table 14: Input Data fields for processed data position

3.3.2.8 MFL Metal Loss

The reported MFL anomalies' depth shall be reported to a 2D grid of locations along the entire length of the pipeline. As discussed above, this is being provided in order to simplify the alignment with other datasets while removing the risk in interpretation that raw signals would bring.

- For each grid point the peak depth between the grid point and the next grid point (axially and circumferentially) and the next grid point based on the reported individual metal loss at that location will be provided
    - Depth of single individual metal loss anomalies boxes (not the overall cluster) will be used in order to provide the full detail of the analyzed data.
    - If 2 or more individual metal losses are at the same location (overlapping) the deeper depth will be reported
    - Where no reported individual metal loss exists 0 would be reported
    - On a defined constant 5mm x 5mm grid. Circumferential spacing as defined in Table 13 (if the technology allows it a finer mesh grid can be specified in the header file).
    - Rectangular boxes will be exported with no tapering/rounding, i.e. each metal loss will only have 1 depth in the export
- All reported metal loss type anomalies will be included (corrosion, manufacturing).

Data field	Description	Dimensions	Factor	Precision/Unit
ML_Depth_Data	Depth of reported metal loss for defined reporting grid	Short [OdoDistance / axial_res, circumf_num]	10	0.1 %t

Table 15: Input Data fields for MFL

3.3.2.9 Geometry

Input data fields requested for Geometry tool are defined in Table 16. The required dimensions, units and precision of the data fields are also described. Not only the data within the feature boxes but also the data from the complete in-line inspection is required in this case.

Data field	Description	Dimensions	Factor	Precision /Unit
CScan	Measured difference to the nominal ID of the pipe along the axial direction	short [nOdoticks, nSensors]	100	0.01 mm

Table 16: Input data fields required for geometry tools

4 References

1. Pipeline Operators Forum, *Specifications and requirements for in-line inspection of pipelines*, POF 100, 2021.
2. Dawson, J. S., and Kariyawasam, S., *Understanding and accounting for pipeline corrosion growth rates*, Proc. 17th Joint Technical Meeting on Pipeline Research, EPRG – PRCI – APIA, 2009.