Data eXchange Unit Description

DXU

Document No.: MST-TNO-4MOST-47110-9237-0001

Issue No.: 1.00

DocuShare URL: https://ds-web.aip.de/docushare/dsweb/Services/Document-4318

Prepared by	Diane Feuillet	2018-03-15
Approved by	C. Clare Worley Digitally signed by C. Clare Worley Date: 2018.03.16 07:47:51 Z Clare Worley (ODG System Engineer)	2018-03-15
Released by	Jakob Walcher Digitally signed by Jakob Walcher Date: 2018.03.16 10:26:13 +01'00' C. Jakob Walcher (OS)	2018-03-15

Change Record

Issue	Date	Page, Section, Paragraph Affected	Reason, Remarks
0.01	2017.02.15	Full text	New Document
0.02	2017.11.20	S6.4, cross references	Semi-updated to new DXU
0.03	2017.12.07	Full text	Small edits from SF and GK, continue migration to new DXU format
0.04	2018.03.15	S6, S6.2, S6.5	Clarifying stacking and FITS structure
1.00	2018.03.15		Released

List of Contributors

Name	First Name	Institute	Comments
Feuillet	Diane	MPIA	
Lind	Karin	MPIA	
Worley	Clare	IoA, Cambridge	
Kordopatis	Georges	Observatoire de la Côte d'Azur	
IWG7 Team			

Table of contents

List of Figures

No table of figures entries found.

List of Tables

1 Objectives

This document defines the Data eXchange Unit (DXU) for the IWG7 data product to be used by 4PA. It details the input data needed, as well as the content and format of the IWG7 data products. The IWG7 data products reflect the Galactic Analysis Pipeline design. The 4MOST Galactic Analysis Pipeline, developed by IWG7, is referred to as 4GP within this document.

2 Applicable Documents (AD)

The following applicable documents (AD) of the exact issue shown form a part of this document to the extent described herein. In the event of conflict between the documents referenced herein and the contents of this document, the contents of this document are the superseding requirement.

AD ID	Document Title	Document Number	Issue	Date
[AD1]	IWG7 Management Plan	MST-PLA-PSC-20307-09237-0001	0.07	05.01.2017
[AD2]	4MOST ICD – General Definitions	MST-ICD-PMO-02000-0002	1.a	21.07.2014
[AD3]	4MOST Acronym List	MST-LIS-PMO-30500-9350-0001	2.00	06.03.2015
[AD4]	VLT Software Programming Standards	VLT-PRO-ESO-10000-0228	2.00	11.02.2010
[AD5]	Software Engineering Development Standard	MST-STD-PMO-40200-9420-0001	1.00	06.04.2016

AD ID	Document Title	Document Number	Issue	Date
[AD6]	FITS Working Group, Commission 5: Documentation and Astronomical Data, International Astronomical Union. Definition of the Flexible Image Transport System (FITS),	http://fits.gsf .nasa.gov/fits_standard.html	3.0	07.2008
[AD7]	ESO Science Data Products Standard	GEN-SPE-ESO-33000-5335	5.0	11.01.2013

3 Reference Documents (RD)

The following reference documents (RD) contain useful information relevant to the subject of the present document.

RD ID	Document Title	Document Number	Issue	Date
[RD1]	DMS-L2 Pipelines	VIS-ICD-4MOST-47110-1440-0001	1.0	18.04.2016
[RD2]	Data Interface Control Document (ESO FITS Header standards)	GEN-SPE-ESO-19400-0794	5.0	08.07.2011
[RD3]	IWG7 Software Tools and Operational Requirements	MST-TNO-PSC-20307-9237-0001	0.02	22.10.2016
[RD4]	DMS DRPD	VIS-DER-4MOST-47110-1410-0002	1.0	27.02.2017
[RD5]	DMS DR Archive	VIS-DER-4MOST-47110-1440-0001	1.1	15.11.2016
[RD6]	Back-End ICD	VIS-ICD-4MOST-47110-9700-0002	1.0	27.02.2017
[RD7]	The UCD1+ controlled vocabulary Version 1.3	http://wiki.ivoa.net/internal/IVOA/PlanetaryUCD/WD-UCDlist-1.3-20170502.pdf	1.3	2017-05-02

4 Definitions

The L1 products are the extracted and reduced science-ready spectra upon which L2 (and L3) analyses will take place. See [RD4] for more details.

The L2 products are the deliverables from the L2 pipelines that are measured on the L1 products and are essential for scientific studies.

The L1 and L2 products comprise the Phase 3 requirements for ESO.

L2 pipelines primarily refer to the analysis pipelines being developed by IWG7 and IWG8. Other L2 products potentially being produced via other analysis pipeline, and thus requiring ingestion into 4OR, are from IWG4.

The L3 products are added value products that are supplementary to the primary L2 products but are not included within the ESO Phase 3 product delivery. They include but are not limited to:

- Supplementary products produced by the L2 pipelines measured on the L1 products
- Measurements on the L1 products by subgroups within an IWG or DRS but exterior to the L2 pipelines
- Scientific products based on the L2 products determined within an IWG or DRS

4OR is the operational repository for 4MOST that will hold all 4MOST products ingested on a continuous basis as described in [RD5].

4PA is the public archive for 4MOST that will hold L1 and L2 products for release to the 4MOST community and the world on a regular basis as described in [RD5].

The data flow of the entire back end operations is described in [RD6].

5 General Format

FITS format is the designated format for data transport for 4MOST. Standard notation and naming conventions for the files and content are provided at the start of the pertinent sections below. There are four key standards that all DXUs must adhere to:

- The use of FITS must adhere to the ESO FITS standards outlined in [RD2].
- The data and metadata described here must meet the requirements for ingestion into the 4MOST archives as described in [RD5] section 4.3.2.2,
- The data and metadata must be VO compliant
- For the purposes of delivery as ESO Phase 3 these products must also adhere to the ESO Science Data Product Standards as described in [AD7].

5.1 Data Package Delivery to 4OR

The data products will be delivered to the 4OR via secure ftp upload as described in [RD5] section 4.3.2.2.

6 FITS File and Structure

The data package is the output of the analysis by IWG7 on the L1 data products. The provenance of the input products (input product filename and version number) is retained in the data products in a Fibre Information Binary Table. IWG7 needs to be designed to handle both singlespec and manyspec files as input. The default input for nightly observations will be manyspec files, which are per fibre per OB. However, for multiple exposures of a single star within a single OB, IWG7 will analyse the L1 stacked and unstacked products. The L1 stacked products follow the standard stacking procedure of only applying helio/barycentric corrections and sky subtract [RD4].

The L1 products could be in manyspec format if the fibre is the same for all exposures of a single star. However, if the fibres are reconfigured between exposures within an OB, then singlespec format is preferred.

For stars observed more than once in different OBs, an updated stacking will be done by IWG7 accounting for the radial velocities (RVs) of the individual observations. In this case, IWG7 will use the stacking routine that DMS intends to issue as a standalone casutool for consistency.

IWG7 will analyze both the L1 stacked and unstacked spectra of any star observed multiple times, even if observed multiple times in a single OB. If the unstacked spectra of a star have RV variations larger than the individual RV uncertainties then the stacking of the spectra will be redone assuming different RVs for each spectrum. The RVs will be determined by the GP from the individual, unstacked spectra. The DMS stacking routine will again be used with the new RVs. These new stacked spectra will be L2 data products.

6.1 FITS Filename

The filename of each delivered product should have the format defined below where the naming fields are separated by an underscore:

Table 6-1 Filename Fields

Field	Description	Example
FM	Code-usable abbreviation for 4MOST	
1	Data Release abbreviation and number to 2.d.p	DR1.00
2	Source of FITS file	IWG7
3	Batch number if required, integer to 3 s.f.	001

FM_<Field1>_<Field2>_<Field3>.fits

The fields are defined in Table 6-1.

The length of the FITS filename should not be excessive (≤30) and 3 fields should be sufficient to distinguish files for a particular data release. Further file specific information should be placed in the primary header (see Section 6.4).

IWG7 is considering producing multiple tables for a given output. The primary driver of this is the availability of asteroseismic data for a subsample of stars. Multiple tables would be created to provide results derived with and without the influence of asteroseismic data. This idea is explained further below. This would be indicated by the Batch number.

6.2 FITS Structure

The structure of the IWG7 L2 data products is defined in Table 6-2 FITS Structure. IWG7 will produce two separate FITS files. The first file will contain the parameters determined by the 4GP and will be delivered as a 2D Binary Table. IWG7 would prefer to be able to use multi-element arrays as entries in the 2D Binary Table, making it more like a 3D Data Cube. However, the Binary Table would also contain single element entries. The structure of this file is given in Table 6-2 and the content of this file is given in Table 6-6.

The second FITS file will contain the L2 1D spectra products. The spectral product of IWG7 will be in singlespec format. This will require that all manyspec format files be converted to singlespec format to allow for the possibility of stacked spectra, which will be produced by IWG7. The extensions of the L2 data product are described below in Table 6-2. This file will also contain pointer information in the header to indicate which L1 data product was used so that the L1 and L2 spectra can be correctly cross-matched.

Table 6-2 FITS Structures

EXTN #	Extension Name	Description
FITS Structure 1: 2D Binary Table		
0	PHU	This is the Primary Header Unit. There will be no data in this HDU. The header will have all the general information about the IWG7 L2 data product.
1	star_parameters	Binary table of the L2 products per star per OB. 2D Binary Table with some multi-elemental array entries.
FITS Structure 2: SingleSpec L2 Spectra		
0	PHU	This is the Primary Header Unit. There will be no data in this HDU. The header will have all the general information about the IWG7 L2 data product as well as pointers as to which version of L1 data product was used.
1	normalised_spectra	Final normalised spectra based on the final stellar parameters. 1D spectrum.
2	normalised_ivar	The inverse variance of each spectrum (similar to a weight map). The variance was defined by the optimal extraction algorithm used and modified during all subsequent processing. Bad pixels will be flagged with an inverse variance of zero. 1D spectrum.
3	synthetic_spectra	Synthetic spectra generated at the respective target respective parameters. 1D spectrum.

6.3 Specific Usage Conventions for 4MOST DXUs

The following conventions are used across the 4MOST DXUs for consistency and standardisation when being ingested by the archives.

6.3.1 NULL Values

For NULL values the standard usage per data type are listed in Table 6-3.

Table 6-3 NULL Values per Data Type

Column Type	Data Type	NULL Value
AlphaNumeric (e.g. Object Name, Flags)	String	‘’
Measurements, Errors	Floating point	Nan, INDEF, NULL
Limits, Number Counts	Integer (positive)	-1

6.3.2 Concatenation symbol

When multiple strings are concatenated within a string cell (e.g. multiple flags, names etc) the delimiter for use in this DXU is the pipe symbol '|'.

6.3.3 Coordinate Units and Precision

The coordinate columns used in this DXU are in units of degree decimals and specified in DOUBLE precision.

6.3.4 Boolean Values

The use of any Booleans with data format of 'L' in this DXU adhere to the correct usage within FITS file of 'T' and 'F'. IWG7 does not know yet if we will use this format.

6.4 Primary Header Unit

The PHU contains further information regarding the source and processing of the data products that is not encoded in the filename.

Keywords added to the standard FITS primary header are listed in Table 6-4.

Table 6-4 Primary Header Keywords

Keyword	Content	Description
FMNODE	IWG7	4MOST node that has produced this data product
RELEASE	DRX.xx	Data release of this data product
PVERS	XX.xx	IWG7 Pipeline Version
PROV0001	FM_*L1product1name*.fits	1st L1 product analysed to produce data product
PROVNNNN	FM_*L1productNNNNname*.fits	NNNNth L1 product analysed to produce data product
DXUDOC	IWG7 DXU document	Data eXchange Unit document in which this product is described
MODULE	XXX	Pipeline module used (e.g. "WD", "FGK", "OBA", etc.)
LINELISTVERS	XX.xx	L2 linelist version
GAIAVERS	GDRX.xx	Gaia Data Release used for prior data
SEISMIC	XX	Flag to indicate the use of asteroseismic information in analysis

Gaia parameters will be used in the derivation of some IWG7 output parameters. In addition to preserving the exact Gaia parameter values used, the relevant Gaia Data Release will be recorded in the Primary Header.

For a subset of the sample, asteroseismic data will be available, which can provide more precise parameters and improve the derivation of some of the IWG7 output parameters. A possible way to deal with using this information for only a subsample of data will be to produce multiple output tables, with and without asteroseismic data included as a prior when deriving other parameters. The SEISMIC keyword in the Primary Header will indicate whether the seismic data has been included.

Each of the Extensions listed in Table 6-2 FITS Structure are described further in the following subsections.

6.5 FITS Extensions

6.5.1 Extension 1 of 2D Binary Table: Stellar Parameters Table

The IWG7 parameters measured on the L1 products, as well as any subsequently stacked spectra, are provided in a 2D Binary Table in Extension 1. The Extension name is Stellar Parameters Table.

Table 6-5 lists specific keywords used in the Extension 1 header that are not referencing the columns in the binary table.

IWG7 does not currently have entries for this table. These may arise later.

Table 6-5 Extension 1 Header Keywords

Keyword	Content	Description
EXTNAME	PARAM_TAB	The table of parameters derived by 4GP by IWG7.

The IWG7 parameters provided in the Stellar Parameters Table are listed in Table 6-6. The columns reflect the information needed to construct the FITS binary table, where TTYPE is the name of the column, TFORM is the data format of the column, TUNIT is the associated unit, TNULL is the value defined as NULL in the binary structure, TDISP[1] is the display precision for the measurement, TDMIN is the minimum expected value for the measurement if numeric, TDMAX is the maximum expected value for the measurement if numeric, and TUCD is the Unified Content Descriptor for the measurement taken from [RD7].

[1] https://heasarc.gsfc.nasa.gov/ftools/fv/doc/displayFormat.html

Table 6-6 Parameter Binary Table Columns

TTYPE	TTYPE comment	TFORM	TFORM comment	TUNIT	TUNIT comment	TNULL	Data Range	TDISP	TDMIN	TDMAX	TUCD
CNAME	4MOST object name from coordinates	A	data format of field: ASCII Character	none	none	empty string	Alpha-numeric	A16	none	none	meta.id
OB	Observation Block	A	data format of field: ASCII Character	none	none	empty string	Alpha-numeric	A	none	none	meta.id
RA	Catalogue RA of object in decimal degrees	D	data format of field: 8-byte DOUBLE	Degree decimals	physical unit of field	NULL	Positive-real	D9.2	0	360	pos.eq.ra
DEC	Catalogue Declination of object in decimal degrees	D	data format of field: 8-byte DOUBLE	Degree decimals	physical unit of field	NULL	Real	D9.1	-90	90	po.eq.dec
SNR	Signal-to-noise ratio	E, vector [3]	data format of field: 4-byte REAL	none	none	NULL	Positive-real	F9.2	0		meta.snr
N_OBS	Number of spectra taken	I	data format of field: 2-byte INTEGER	none	none	NULL	Positive-integer	I8.5	0		meta.id
OBJ_TYPE	Classification of object	A	data format of field: ASCII Character	none	none	NULL	Alpha-numeric	A50	none	none	meta.id
4GP_MOD	4MOST Galactic Pipeline module used for analysis	A	data format of field: ASCII Character	none	none	empty string	Alpha-numeric	A10	none	none	meta.id
RV	Radial Velocity	E	data format of field: 4-byte REAL	km / s	physical unit of field	NULL	Real	F10.3	-inf	inf	spec.velocity
E_RV	Error on RV	E, vector [2]	data format of field: 4-byte REAL	km / s	physical unit of field	NULL	Positive-real	F10.3	0	inf	stat.error;spec.velocity
RV_FLAG	Flag on RV	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.velocity
V_SINI	Rotational velocity	E	data format of field: 4-byte REAL	km / s	physical unit of field	NULL	Real	F10.3	-inf	inf	spec.velocity
E_V_SINI	Error on V_SINI	E, vector [2]	data format of field: 4-byte REAL	km / s	physical unit of field	NULL	Positive-real	F10.3	0	inf	stat.error:spec.velocity
V_SINI_FLAG	Flag on V_SINI	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.velocity

TEFF	Effective temperature	E	data format of field: 4-byte REAL	K	physical unit of field	NULL	Positive-real	F10.3	0	inf	spec.temperature.effective
E_TEFF	Error on TEFF	E, vector [2]	data format of field: 4-byte REAL	K	physical unit of field	NULL	Positive-real	F10.3	0	inf	stat.error:spec.temperature.effective
TEFF_FLAG	Flag on TEFF	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code; spec.temperature.effective
LOGG	Surface gravity	E	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.gravity
E_LOGG	Error on LOGG	E, vector [2]	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error:spec.gravity
LOGG_FLAG	Flag on LOGG	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.gravity
M_H	Overall metallicity	E	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.abundance
E_M_H	Error on FE_H	E, vector [2]	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error:spec.abundance
M_H_FLAG	Flag on FE_H	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.abundance
ALPHA_FE	Abundance of alpha elements	E	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.abundance
E_ALPHA_FE	Error on ALPHA_FE	E, vector [2]	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error: spec.abundance
ALPHA_FE_FLAG	Flag on ALPHA_FE	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code; spec.abundance
V_MICRO	Microturbulent velocity	E	data format of field: 4-byte REAL	km / s	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.velocity
E_V_MICRO	Error on V_MICRO	E, vector [2]	data format of field: 4-byte REAL	km / s	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error:spec.velocity
V_MICRO_FLAG	Flag on V_MICRO	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.velocity
V_MACRO	Macroturbulent velocity	E	data format of field: 4-byte REAL	km / s	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.velocity
E_V_MACRO	Error on V_MACRO	E, vector [2]	data format of field: 4-byte REAL	km / s	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error:spec.velocity

V_MACRO_FLAG	Flag on V_MACRO	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.velocity
L_BOL	Bolometric luminosity	E	data format of field: 4-byte REAL	solar luminosities	physical unit of field	NULL	Real	F10.5	-inf	inf	phys.luminosity
E_L_BOL	Error on L_BOL	E, vector [2]	data format of field: 4-byte REAL	solar luminosities	physical unit of field	NULL	Positive-real	F10.4	0	inf	stat.error:phys.luminosity
L_BOL_FLAG	Flag on L_BOL	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;phys.luminosity
X_H	Abundance of individual elements [X/H]	E, vector [N]	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.abundance
E_X_H_UP	Upper error on X_H	E, vector [N]	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error:spec.abundance
E_X_H_LO	Lower error on X_H	E, vector [N]	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error:spec.abundance
X_H_FLAG	Flag on X_H	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.abundance
X_FE	Abundance of individual elements [X/Fe]	E, vector [N]	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.abundance
E_X_FE_UP	Upper error on X_FE	E, vector [N]	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error: spec.abundance
E_X_FE_LO	Upper error on X_FE	E, vector [N]	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error: spec.abundance
X_FE_FLAG	Flag on X_FE	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code; spec.abundance
NLTE_CORR	Flag to indicate NLTE correction applied	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.id
HE1	Abundance of neutral helium	E	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.abundance
E_HE1	Error on HE1	E, vector [2]	data format of field: 4-byte REAL	dex	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error: spec.abundance
HE1_FLAG	Flag on HE1	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code; spec.abundance
EW_HA	EW Halpha	E	data format of field: 4-byte REAL	Angstroms	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.line.width

E_EW_HA	Error on EW_HA	E, vector [2]	data format of field: 4-byte REAL	Angstroms	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error:spec.line.width
EW_HA_FLAG	Flag on EW_HA	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.line.width
EW_HB	EW Hbeta	E	data format of field: 4-byte REAL	Angstroms	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.line.width
E_ EW_HB	Error on EW_HB	E, vector [2]	data format of field: 4-byte REAL	Angstroms	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error:spec.line.width
EW_HB_FLAG	Flag on EW_HB	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.line.width
AGE	Most probable age	E	data format of field: 4-byte REAL	Gyr	physical unit of field	NULL	Positive-real	F10.4	0	inf	time.age
E_AGE	Error on AGE	E, vector [2]	data format of field: 4-byte REAL	Gyr	physical unit of field	NULL	Positive-real	F10.4	0	inf	stat.error:time.age
AGE_FLAG	Flag on AGE	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;time.age
AGE_PDF_X	X-axis array for age probability distribution function, age values	E, vector [N]	data format of field: 4-byte REAL	Gyr	physical unit of field	NULL	Positive-real	F10.4	0	inf	time.age
AGE_PDF_Y	Y-axis array for age probability distribution function, probability values	E, vector [N]	data format of field: 4-byte REAL	none	none	NULL	Positive-real	F10.6	0	1	stat.likelihood
MASS	Most probable mass	E	data format of field: 4-byte REAL	Solar masses	physical unit of field	NULL	Positive-real	F10.4	0	inf	phys.mass
E_MASS	Error on MASS	E, vector [2]	data format of field: 4-byte REAL	Solar masses	physical unit of field	NULL	Positive-real	F10.4	0	inf	stat.error:phys.mass
MASS_FLAG	Flag on MASS	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;phys.mass
MASS_PDF_X	X-axis array for mass probability distribution function, mass values	E, vector [N]	data format of field: 4-byte REAL	Solar masses	physical unit of field	NULL	Positive-real	F10.4	0	inf	phys.mass
MASS_PDF_Y	Y-axis array for mass probability distribution function, probability values	E, vector [N]	data format of field: 4-byte REAL	none	none	NULL	Positive-real	F10.6	0	1	stat.likelihood
DISTANCE	Most probable distance	E	data format of field: 4-byte REAL	kiloparsecs	physical unit of field	NULL	Positive-real	F12.4	0	inf	phys.distance

E_DISTANCE	Error on DISTANCE	E, vector [2]	data format of field: 4-byte REAL	kiloparsecs	physical unit of field	NULL	Positive-real	F12.4	0	inf	stat.error:phys.distance
DISTANCE_FLAG	Flag on DISTANCE	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;phys.distance
DISTANCE_PDF_X	X-axis array for distance probability distribution function, distance values	E, vector [N]	data format of field: 4-byte REAL	kiloparsecs	physical unit of field	NULL	Positive-real	F12.4	0	inf	phys.distance
DISTANCE_PDF_Y	Y-axis array for distance probability distribution function, probability values	E, vector [N]	data format of field: 4-byte REAL	kiloparsecs	physical unit of field	NULL	Positive-real	F12.4	0	1	stat.likelihood
CHI_2	Chi2 flag	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer (0:5)	I5.1	0	5	stat.fit.chi2
VEIL	Veiling	E	data format of field: 4-byte REAL	none	none	NULL	Real	F12.5	-inf	inf	spec.line
E_VEIL	Error on VEIL	E, vector [2]	data format of field: 4-byte REAL	none	none	NULL	Positive-real	F12.5	0	inf	stat.error:spec.line
VEIL_FLAG	Flag on VEIL	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.line
MASS_LOSS	mass loss	E	data format of field: 4-byte REAL	solar masses / year	physical unit of field	NULL	Real	F10.5	-inf	inf	spec.line
E_MASS_LOSS	Error on MASS_LOSS	E, vector [2]	data format of field: 4-byte REAL	solar masses / year	physical unit of field	NULL	Positive-real	F10.5	0	inf	stat.error:spec.line
MASS_LOSS_FLAG	Flag on MASS_LOSS	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.line
V_FILL	clumping or volume filling factor	E	data format of field: 4-byte REAL	none	none	NULL	Real	F12.5	-inf	inf	phys.spec.line
E_V_FILL	Error on V_FILL	E, vector [2]	data format of field: 4-byte REAL	none	none	NULL	Positive-real	F12.5	0	inf	stat.error:spec.line
V_FILL_FLAG	Flag on V_FILL	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.line
BETA	beta-type velocity law	E	data format of field: 4-byte REAL	none	none	NULL	Real	F12.5	-inf	inf	spec.line
E_BETA	Error on BETA	E, vector [2]	data format of field: 4-byte REAL	none	none	NULL	Positive-real	F12.5	0	inf	stat.error:spec.line

BETA_FLAG	Flag on BETA	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.line
V_TERM	terminal velocity	E	data format of field: 4-byte REAL	km / s	physical unit of field	NULL	Real	F12.5	-inf	inf	spec.velocity
E_V_TERM	Error on V_TERM	E, vector [2]	data format of field: 4-byte REAL	km / s	physical unit of field	NULL	Positive-real	F12.5	0	inf	stat.error:spec.velocity
V_TERM_FLAG	Flag on V_TERM	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.velocity
F_ION	Ionizing flux	E	data format of field: 4-byte REAL	photons / s	physical unit of field	NULL	Real	F12.5	-inf	inf	spec.flux
E_F_ION	Error on F_ION	E, vector [2]	data format of field: 4-byte REAL	photons / s	physical unit of field	NULL	Positive-real	F12.5	0	inf	stat.error:spec.flux
F_ION_FLAG	Flag on F_ION	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.flux
GAMMA	Gravity sensitive spectral index	E	data format of field: 4-byte REAL	none	none	NULL	Real	F12.5	-inf	inf	spec.line
E_GAMMA	Error on GAMMA	E, vector [2]	data format of field: 4-byte REAL	none	none	NULL	Positive-real	F12.5	0	inf	stat.error:spec.line
GAMMA_FLAG	Flag on GAMMA	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.code;spec.line
SINGLE_FLAG	Flag to indicate single star systems	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.id
BINARY_FLAG	Flag to indicate binary stars	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.id
VARIABLE_FLAG	Flag to indicate variable stars	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.id
NON_STELLAR_FLAG	Flag for non-stellar objects	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.id
WD_FLAG	Flag to indicate known White Dwarfs	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.id
WD_TYPE	Flag to indicate known White Dwarf type	I	data format of field: 2-byte INTEGER	none	none	NULL	Positive-integer	I5.1	0	10	meta.id
ACTIVITY_FLAG	Flag to indicate significant atmospheric activity	I	data format of field: 2-byte INTEGER	none	none	-1	Positive-integer	I5.1	0	5	meta.id

Parameters that will be used as priors to derive other parameters are not currently included in this table. These prior parameters will need to be available and tagged with the version of the analysis they were used for, however, as they are not derived by the 4GP, they are not currently included in the output. These parameters are listed below in Section 7.1.

We would like to have some parameters reported as vectors/multi-element arrays in the binary table. This will be used for parameters such as asymmetric errors, individual element abundances, and probability distribution functions. The errors will be given as arrays with 2 elements to allow for asymmetric errors to be reported. We set this as the default for all parameters. Individual element abundances will be reported as an array with length N, equal to the number of elements we will measure abundances for. Probability distribution functions will be given for age, mass, and distance. The length of these arrays will depend on the gridding of the models used, but will likely be on the order 20-200 elements.

6.5.2 Extension 1 of the L2 spectra product: Normalised Spectra

Extension 1 of the L2 spectra product contains the final normalised spectra that are output from the IWG7 L2 analysis pipeline.

The wavelength vector (or x vector of a function f(x)) is specified using keywords in the Extension 1 of the L2 spectra product header as provided in Table 6-7.

Table 6-7 Extension 1 of L2 spectra product Header Keywords

Keyword	Content	Description
EXTNAME	NORM_SPEC	
CRVAL	3900	Starting Wavelength in Angstroms
CDELTA		Wavelength step in Angstroms

6.5.3 Extension 2 of the L2 spectra product: Normalised Inverse Variance

Extension 2 of the L2 spectra product contains the normalised inverse spectra that are output from the IWG7 L2 analysis pipeline.

The wavelength vector (or x vector of a function f(x)) is specified using keywords in the Extension 2 of the L2 spectra product header as provided in Table 6-8.

Table 6-8 Extension 2 of the L2 spectra product Header Keywords

Keyword	Content	Description
EXTNAME	INV_SPEC	
CRVAL	3900	Starting Wavelength in Angstroms
CDELTA		Wavelength step in Angstroms

6.5.4 Extension 3 of the L2 spectra product: Synthetic Spectra

Extension 3 contains the synthetic spectra that are output from the IWG7 L2 analysis pipeline.

The wavelength vector (or x vector of a function f(x)) is specified using keywords in the Extension 3 of the L2 spectra product header as provided in Table 6-9.

Table 6-9 Extension 3 of the L2 spectra product Header Keywords

Keyword	Content	Description
EXTNAME	SYNTH_SPEC	
CRVAL	3900	Starting Wavelength in Angstroms
CDELTA		Wavelength step in Angstroms
MODEL		Name of model grid used to create spectrum
…		Additional parameters to create synthetic spectrum

7 Further Information on Provenance

7.1 Input data used

In order to produce the output products listed above, IWG7 will need input data from other sources.

- Photometric data on all stars in IR, optical, and UV bands
- Parallax measurements from Gaia
- Asteroseismic data if available. Parameters needed are delta nu, nu max, peak separation, and evolutionary state.
- Variability classificiation

7.2 Algorithms and software

IWG7 will use the algorithms and software detailed in [RD3]. Briefly, IWG7 will develop three independent software packages: stellar spectra synthesis software, parameter determination software, and validation software. The stellar spectra synthesis software will use existent modules and pre-computed grids of stellar atmospheres to produce synthetic spectra needed for the spectra parameterization software. The parameter determination software will contain three modules: FGK stars, OBA stars, and white dwarfs (WD). The FGK and OBA modules will consist of an RV and normalisation module, both a data-driven and a model-driven pipeline, and an age-determination module. The WD module will have its own RV and normalisation, and parameter determination pipelines. The validation software will assess the quality of the derived parameters and derive corrections when needed.

7.3 Physical meaning of parameters

The physical parameters listed in Table 6-6 are standard stellar spectroscopic measurements and can be used to fully describe the physical parameters of a star. While not all parameters will be relevant to every star observed, they are all standard measurements for at least one of the three 4GP modules (WD, FGK, OBA). Some additional parameters can be provided due to the addition of photometric, astrometric, and asteroseismic data. Parameters that can be derived with external data and are relevant to the physical parameters of the star, such as mass and age, are included in Table 6-6 as products of IWG7.

7.4 Flag Definition Table

This DXU contains flags on key measurements. The flags and their specific values are defined in the following table:

We will fill in the table with the name, values, and definitions of all flags to be used.

Make sure this is consistent with the Object flags. We also will probably need to wait until we have the data to decide exactly which flags we need.

Table 7-1 Example Flag Definitions

Flag Column	**Values**	**Description per Value per Flag**
TEFF_FLAG	0	The reported TEFF is good
	1	The reported TEFF is okay
	2	The reported TEFF is marginal
	3	The reported TEFF is suspect
	4	The reported TEFF is unreliable
	5	The TEFF is not available

Appendix A Interface Control Drawings

Any relevant drawings will be placed here in the future.

Appendix B List of Acronyms

List of Acronyms	
4GP	4MOST Galactic Pipeline
4MOST	4-metre Multi-Object Spectroscopic Telescope
4OR	4MOST Operational Repository
4PA	4MOST Public Archive
AD	Applicable Document
DMS	Data Management System
DRS	Design Reference System
DXU	Data eXchange Unit
FIBINFO	Fibre Information
FITS	Flexible Image Transport System
GSF	Geometric Selection Function
ICD	Interface Control Document
IWG	Infrastructure Working Group
L1-3	Level 1 to 3 Data Products
N/A	Not Applicable
OB	Observation Block
OpSys	Operational Systems
OS	Operations Scientist
OSF	Object Selection Function

PHU	Primary Header Unit
PS	Project Scientist
QC	Quality Control
RD	Reference Document
TBD	To be defined