Data eXchange Unit Description

IWG8 DXU – Extragalactic Analysis Pipeline

Document No.: MST-TNO-4MOST-47110-9238-0001

Issue No.: 2.0

DocuShare URL: https://ds-web.aip.de/docushare/dsweb/Services/Document-4319

Prepared by _____

Luke Davies (IWG8 Lead)

Approved by C. Clare Worley Digitally signed by C. Clare Worley Date: 2018.03.19 08:52:07 Z

ODG System Engineer

Released by Jakob Walcher Digitally signed by Jakob Walcher Date: 2018.03.17 08:19:26 +01'00'

C. Jakob Walcher (OM)

Change Record

Issue	Date	Page, Section, Paragraph Affected	Reason, Remarks
1.00	2017-02-15	all	New Document
1.01	2018-01-25	all	Updates to new template + PI comments
1.02	2018-02-13	Minor Throughout	Small changes based on Clare Worley's comments
2.0	2018-03-17	Minor Throughout	Small updates for FDR release

List of Contributors

Name	First Name	Institute	Comments
Davies	Luke	UWA	

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 IWG8 4MOST extragalactic analysis pipeline (4XP) data products to be used by 4OR and the survey teams. IWG8 is responsible for determining redshifts and measuring spectral properties for all extragalactic surveys.

Any changes to the DXU must be agreed upon by the IWG8 and DMS including 4OR and the survey teams in advance of data product submission to allow all parties to update their software to accommodate any agreed changes.

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]	4MOST ICD – General Definitions	MST-ICD-PMO-02000-0002	1.a	2014-07-21
[AD2]	4MOST Acronym List	MST-LIS-PMO-30500-9350-0001	2.00	2015-03-06
[AD3]	VLT Software Programming Standards	VLT-PRO-ESO-10000-0228	2.00	11.02.2010
[AD4]	Software Engineering Development Standard	MST-STD-PMO-40200-9420-0001	1.00	06.04.2016

AD ID	Document Title	Document Number	Issue	Date
[AD5]	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
[AD6]	ESO Science Data Products Standard	GEN-SPE-ESO-33000-5335	5.0	11/01/2013
[AD7]	IWG8 MP	MST-PLA-PSC-20308-9238	0.03	2015.14.12
[AD8]	IWG8 Requirements	??	0.01	2016.06.10
[AD9]	IWG8 4XP-Sim	??	0.01	??

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]	Data Interface Control Document (ESO FITS Header standards)	GEN-SPE-ESO-19400-0794	5.0	8/07/2011
[RD2]	DMS DRPD	VIS-DER-4MOST-47110-1410-0002	1.0	2017-02-27
[RD3]	DMS DR Archive	VIS-DER-4MOST-47110-1440-0001	1.1	2016-11-15
[RD4]	Back-End ICD	VIS-ICD-4MOST-47110-9700-0002	1.0	2017-02-27
[RD5]	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, as described in [RD2]. The L2 products are the deliverables from the L2 pipelines that are measured on the L1 products. The L1 and L2 products comprise the Phase 3 requirements for ESO.

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

4OR is the operational repository for 4MOST that will hold all 4MOST products ingested on a continuous basis as described in [RD3].

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 [RD3].

The data flow of the entire back end operations is described in [RD4].

4XP is the extragalactic analysis pipeline that will be produced by IWG8.

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 [RD1],
- The data and metadata described here must meet the requirements for ingestion into the 4MOST archives as described in [RD3] section 4.3.2.2,
- The data and metadata must be VO compliant. Unified Content Descriptors[1] must be specified per measurable as listed in [RD5].
- For the purposes of delivery as ESO Phase 3 these products must also adhere to the ESO Science Data Product Standards as described in [AD6].

5.1 Data Package Delivery to 4OR

The data products will be delivered to the 4OR via secure ftp upload as described in [RD3] section 4.3.2.2.

6 FITS File and Structure

The data package is the output of the analysis by IWG8 on the L1 reduced and calibrated spectrum. The provenance of the input products (input product filename and version number) is retained in the data products by the Fibre Information Binary Table.

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:

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

The fields are defined in **Table 6-1**.

[1] http://www.ivoa.net/documents/REC/UCD/UCD-20050812.html

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	IWG8
3	Batch number if required, integer to 3 s.f.	001

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).

6.2 FITS Structure

The structure of the IWG8 data product is defined in **Table 6-2**. The types of data delivered are 1D Spectra, Functions and 2D Binary Tables

Table 6-2 FITS Structure		
EXTN #	**Extension Name**	**Description**
1	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 IWG8 L2 data, the version of 4XP that was run, the templates used in the fitting, the date of running the pipeline, etc
2	SOURCE_PARA	Binary table of the L2 products per L1 spectrum. Each spectrum will be flagged with a unique spectrum ID, source ID and survey ID.
3	CROSS_CORR	The resultant Fourier cross-correlation spectra produced in determining redshifts in 4XP-Z. This will contain one spectrum per fitted template, per source spectrum. Ultimately, these may not be retained, but may be useful in investigating spurious spectra, or searching for occulting/eclipsing sources.
4	FIT_RESIDUALS	Spectra containing the residuals of the best fit stellar continuum + emission line fits produced in 4XP-Fit. These will be useful in determining the accuracy of any derived parameters.

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 6-1** .

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

Use of any Booleans in this DXU adhere to the correct usage within FITS file of 'T' and 'F'.

6.4 Primary Header

The primary header 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	<SOURCE>	4MOST node that has produced this data product
RELEASE	DRX.xx	Data release of this data product
PVERS	XX.xx	4XP Pipeline Version
PROV0001	FM_*L1product1name*.fits	1st L1 product analysed to produce data product
DXUDOC	**MST-TNO-4MOST-47110-9238-0001**	DXU document in which this product is described.
DATE	xx.xx.xx	Date that the pipeline was run
TIME	xx.xx	Time that the pipeline was run
USER	XXXXX	User responsible for running pipeline
CONTACT	XXXXX	Who to contact regarding this product
ZTEMP	XX.xx	Version of template files used in 4XP-Z
FGTEMP	XX.xx	Version of galaxy template files used in 4XP-Fit
FAGNTEMP	XX.xx	Version of AGN template files used in 4XP-fit
LINESLIST	XX.xx	Version of spectral line list used in fitting

Table 6-5.1 Extension 1 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
SPECTRUM_ID	Unique spectrum identifier	I	data format of field: 2-byte INTEGER	none	none	NULL	Positive-integer	I5.1	1	1e7	meta.id
RA	Catalogue RA of object in decimal degrees (J2000)	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 (J2000)	D	data format of field: 8-byte DOUBLE	Degree decimals	physical unit of field	NULL	Real	D9.1	-90	90	pos.eq.dec
zBest	Best fit heliocentric redshift	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	10	src.redshift
zBestErr	Best fit heliocentric redshift error	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	10	src.redshift;stat;error
zBestCCSig	Best fit heliocentric redshift cross-correlation peak value divided by RMS CC value	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	100	stat.correlation
zBestFOM	Figure of merit of best fit redshift	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	10	stat.fit.goodness
zBestProb	Estimated probability of redshift success	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	1	stat.fit.goodness
zBestT_ID	Best fit heliocentric redshift template identifier	I	data format of field: 2-byte INTEGER	-	none	NULL	Positive-integer	I5.1	0	1000	meta.id
zBestType	Best fit heliocentric redshift template type (star, SF galaxy, passive galaxy, AGN, QSO)	A	data format of field: ASCII Character	-	none	empty string	Alpha-numeric	A16	none	none	meta.id
z2	2nd best heliocentric redshift	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	10	src.redshift

z2Err	2nd best heliocentric redshift error	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	10	src.redshift;stat;error
z2CCSig	2nd best heliocentric redshift cross-correlation peak value divided by RMS CC value	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	100	stat.correlation
z2T_ID	2nd best heliocentric redshift template identifier	I	data format of field: 2-byte INTEGER	-	none	NULL	Positive-integer	I5.1	0	1000	meta.id
z2Type	2nd best heliocentric redshift template type (star, SF galaxy, passive galaxy, AGN, QSO)	A	data format of field: ASCII Character	-	none	empty string	Alpha-numeric	A16	none	none	meta.id
z3	3rd best heliocentric redshift	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	10	src.redshift
z3Err	3rd best heliocentric redshift error	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	10	src.redshift;stat;error
z3CCSig	3rd best heliocentric redshift cross-correlation peak value divided by RMS CC value	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	100	stat.correlation
z3T_ID	3rd best heliocentric redshift template identifier	I	data format of field: 2-byte INTEGER	-	none	NULL	Positive-integer	I5.1	0	1000	meta.id
z3Type	3rd best heliocentric redshift template type (star, SF galaxy, passive galaxy, AGN, QSO)	A	data format of field: ASCII Character	-	none	empty string	Alpha-numeric	A16	none	none	meta.id
z4	4th best heliocentric redshift	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	10	src.redshift
z4Err	4th best heliocentric redshift error	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	10	src.redshift;stat;error
z4CCSig	4th best heliocentric redshift cross-correlation peak value divided by RMS CC value	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	100	stat.correlation
z4T_ID	4th best heliocentric redshift template	I	data format of field: 2-byte INTEGER	-	none	NULL	Positive-integer	I5.1	0	1000	meta.id

	identifier										
z4Type	4th best heliocentric redshift template type (star, SF galaxy, passive galaxy, AGN, QSO)	A	data format of field: ASCII Character	-	none	empty string	Alpha-numeric	A16	none	none	meta.id
SC_ID	ID of best fit SPM continuum model	I	data format of field: 2-byte INTEGER	-	none	NULL	Positive-integer	I5.1	0	1000	meta.id
SC_AGE	Age of best fit SPM continuum model	D	data format of field: 8-byte DOUBLE	Gyr	physical unit of field	NULL	Positive-rea1	D9.2	0	20	time.age
SC_Z	Metallicity of best fit SPM continuum model	D	data format of field: 8-byte DOUBLE	Zsun	physical unit of field	NULL	Positive-rea1	D9.2	0	02	phys.abund.Z
SC_eFOLD	E-folding time of best fit SPM continuum model	D	data format of field: 8-byte DOUBLE	Gyr	physical unit of field	NULL	Positive-rea1	D9.2	0	20	time.interval
SC_CHI2	Reduced Chi^2 of best fit SPM model	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-rea1	D9.2	0	100	stat.fit.chi2
MEAN_CONT	Sigma clipped mean continuum level of spectrum	D	data format of field: 8-byte DOUBLE	ergs/s/cm^2/ang	physical unit of field	NULL	Positive-rea1	D9.2	0	1	spect.continuum; stat.mean
SIGMA_CONT	Scatter in continuum level of spectrum	D	data format of field: 8-byte DOUBLE	ergs/s/cm^2/ang	physical unit of field	NULL	Positive-rea1	D9.2	0	1	spect.continuum; stat.stdev
D4000	Strength of 4000A break	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-rea1	D9.2	0	100	spect.continuum; arith.ratio
D4000Err	Error on strength of 4000A break	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-rea1	D9.2	0	100	spect.continuum; arith.ratio; stat.error
<Line>_CONT	Spectral continuum at line position	D	data format of field: 8-byte DOUBLE	ergs/s/cm^2/ang	physical unit of field	NULL	Positive-rea1	D9.2	0	1	spect.continuum
<Line>_CONTErr	Error on spectral continuum at line position	D	data format of field: 8-byte DOUBLE	ergs/s/cm^2/ang	physical unit of field	NULL	Positive-rea1	D9.2	0	1	spect.continuum; stat.error
<Line>_CEN	Observed-frame predicted line center position	D	data format of field: 8-byte DOUBLE	Ang	physical unit of field	NULL	Positive-rea1	D9.2	3000	11000	em.wl.central
<Line>_POS	Fitted line center position	D	data format of field: 8-byte DOUBLE	Ang	physical unit of field	NULL	Positive-rea1	D9.2	3000	11000	em.wl.central
<Line>_POSErr	Error on fitted line center position	D	data format of field: 8-byte DOUBLE	Ang	physical unit of field	NULL	Positive-rea1	D9.2	0	100	em.wl.central ;stat.error
<Line>_WIDTH	Fitted line FWHM	D	data format of field: 8-byte DOUBLE	Ang	physical unit of field	NULL	Positive-rea1	D9.2	0	1000	spect.line.width
<Line>_WIDTHErr	Error on fitted line FWHM	D	data format of field: 8-byte DOUBLE	Ang	physical unit of field	NULL	Positive-rea1	D9.2	0	100	spect.line.width; stat.error

<Line>_FLUX	Fitted line flux	D	data format of field: 8-byte DOUBLE	ergs/s/cm^2	physical unit of field	NULL	Real	D9.2	-1	1	spect.line.strength
<Line>_FLUXErr	Error on fitted line flux	D	data format of field: 8-byte DOUBLE	ergs/s/cm^2	physical unit of field	NULL	Positive-real	D9.2	0	1	spect.line.strength; stat.error
<Line>_EW	Fitted line equivalent width	D	data format of field: 8-byte DOUBLE	Ang	physical unit of field	NULL	Real	D9.2	-100	100	spect.line.eqWidth
<Line>_EWErr	Error on fitted line equivalent width	D	data format of field: 8-byte DOUBLE	Ang	physical unit of field	NULL	Positive-real	D9.2	0	100	spect.line.eqWidth; stat.error
<Line>_FLAG	Flag for possible errors in line fitting and derived properties.	I	data format of field: 2-byte INTEGER	-	none	-1	Positive-integer	I5.1	0	100	meta.code
BD	Balmer decrement	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real	D9.2	0	100	spect.line.strength; arith.ratio
SFR_Ha	H-alpha derived SFR	D	data format of field: 8-byte DOUBLE	Msun/yr	physical unit of field	NULL	Real	D9.1	-10000	10000	phys.SFR;em.line.Halpha
SFR_HaErr	Error on H-alpha derived SFR	D	data format of field: 8-byte DOUBLE	Msun/yr	physical unit of field	NULL	Positive-real	D9.2	0	10000	phys.SFR;em.line.Halpha;stat.error
SFR_OII	OII derived SFR	D	data format of field: 8-byte DOUBLE	Msun/yr	physical unit of field	NULL	Real	D9.1	-10000	10000	phys.SFR;em.line.OII
SFR_OIIErr	Error on OII derived SFR	D	data format of field: 8-byte DOUBLE	Msun/yr	physical unit of field	NULL	Positive-real	D9.2	0	10000	phys.SFR;em.line.OII;stat.error
METAL_R23	R23-derived metallicty (Kewley & Kobulnicky, 2004)	D	data format of field: 8-byte DOUBLE	Zsun	physical unit of field	NULL	Positive-real	D9.2	0	100	phys.abund.Z
METAL_NO_R23	NII/OII + R23 derived metallcity (Kewley & Dopita, 2002)	D	data format of field: 8-byte DOUBLE	Zsun	physical unit of field	NULL	Positive-real	D9.2	0	100	phys.abund.Z
METAL_NH	NII/Ha-derived metallicty (Pettini & Pagel, 2004)	D	data format of field: 8-byte DOUBLE	Zsun	physical unit of field	NULL	Positive-real	D9.2	0	100	phys.abund.Z
METAL_NH_OH	NII/Ha + OIII/Hb-derived metallicty (Pettini & Pagel, 2004)	D	data format of field: 8-byte DOUBLE	Zsun	physical unit of field	NULL	Positive-real	D9.2	0	100	phys.abund.Z
STELLAR_MASS	Spectral stellar mass estimate	D	data format of field: 8-byte DOUBLE	Msun	physical unit of field	NULL	Positive-real	D9.2	4	20	phys.mass
BPT_TYPE	Identifier for BPT classification (SF, AGN or composite)	A	data format of field: ASCII Character	-	none	empty string	Alpha-numeric	A16	none	none	meta.code
BPT_FLAG_Conf	Confidence of BPT classification.	I	data format of field: 2-byte INTEGER	-	none	-1	Positive-integer	I5.1	0	100	stat.probability

6.5 FITS Extensions

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

6.5.1 Extension 1: SOURCE_PARA (Binary Table)

The 4XP parameters measured on the L1 products are provided in a Binary Table in Extension 1. The Extension name is SOURCE_PARA.
Table 6-5.1 lists specific keywords used in the Extension 2 header that are not referencing the columns in the binary table.

Table 6-5.1 **Extension 1 Header Keywords**

Keyword	Content	Description
EXTNAME	SOURCE_PARA	Name of extension

The 4XP parameters provided in the Binary Table are listed in Table 6-5.1.1. 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[2] 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 [RD5].

6.5.2 Extension 2: CROSS_CORR (Spectral Product)

Extension 2 contains the set of cross-correlation spectra that are output from the 4XP analysis pipeline.

This extension will contain the Fourier cross correlation fits between the target spectrum and best-fit templates (probably those delineated by zBestT_ID, z2T_ID, z3T_ID and z4T_ID in the table above, potentially extending to lower number of fit; i.e. z5, z6….).
As such, the extension will contain NSpec X 4 1D vectors containing the cross-correlation values, and a FITS header containing CRVAL, CRPIX and CDELT values for the pixel position, with units in terms of redshift. This will allow the user to reconstruct the cross-correlation strength between the target spectrum and best-fit templates as a function of redshift. Note description column is given with a long form and FITS header from truncated to <49 characters in.

Table 6 -5.2 Extension 2 Header Keywords

Keyword	Content	Description
EXTNAME	CROSS_CORR	Extension name

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

CRVAL	0	Starting redshift in correlation
CRPIX	1	Starting pixel position in x-direction
CDELT	0.00001	redshift pixel interval
SPEC_ID_ R1_4	1039475	Unique L1 spectrum identifier linked to SPEC_ID Extension 2 Binary table for cross-correlation spectrum in row1-4 of this extension. [Unique L1 spectrum identifier R1_4]
T_ID_R1	3	Unique template spectrum identifier linked to zBestT_ID for previous SPEC_ID in Extension 2 Binary table for cross-correlation spectrum in row1 of this extension. [Unique template identifier for zBestT_ID R1]
T_ID_R2	4	Unique template spectrum identifier linked to z2T_ID for previous SPEC_ID in Extension 2 Binary table for cross-correlation spectrum in row2 of this extension. [Unique template identifier for z2T_ID R1]
T_ID_R3	1	Unique template spectrum identifier linked to z3T_ID for previous SPEC_ID in Extension 2 Binary table for cross-correlation spectrum in row3 of this extension. [Unique template identifier for z3T_ID R1]
T_ID_R4	35	Unique template spectrum identifier linked to z4T_ID for previous SPEC_ID in Extension 2 Binary table for cross-correlation spectrum in row4 of this extension. [Unique template identifier for z4T_ID R1]
SPEC_ID_R5_8	1039476	Unique L1 spectrum identifier linked to SPEC_ID Extension 2 Binary table for cross-correlation spectrum in row5-8 of this extension. [Unique L1 spectrum identifier R5_8]
T_ID_R5	5	Unique template spectrum identifier linked to zBestT_ID for previous SPEC_ID in Extension 2 Binary table for cross-correlation spectrum in row5 of this extension. [Unique template identifier for zBestT_ID R5]
…..		
SPEC_ID_RN1_N2	4759602	Unique L1 spectrum identifier linked to SPEC_ID Extension 2 Binary table for cross-correlation spectrum in rowN1-N2 of this extension. [Unique L1 spectrum identifier RN1_N2]
T_ID_RN	67	Unique template spectrum identifier linked to zBestT_ID for previous SPEC_ID in Extension 2 Binary table for cross-correlation spectrum in rowN of this extension. [Unique template identifier for zBestT_ID RN]

6.5.3 Extension 3: FIT_RESIDUALS (Spectral Product)

Extension 3 contains the set of fit residual spectra that are output from the 4XP-Fit analysis pipeline.

This extension will contain the residuals of the best-fit SPM model + emission line fits produced in 4XP-Fit (one line for each L1 spectrum). This will essentially be the SC_ID SPM template plus all Gaussian emission lines defined by [<Line>_POS, <Line>_WIDTH, <Line>_FLUX, <Line>_CONT] subtracted from the observed L1 spectrum. This will allow

the user to asses the reliability of the fit parameters derived by 4XP-Fit. Potentially this extension could also contain the 2nd, 3rd …… best fits.
This extension will contain a vector of residual values for each spectrum and will use the same CRVAL, CRPIX and CDELT values as the observed L1 spectrum.

Table 6-5.3 Extension 3 Header Keywords

Keyword	Content	Description
EXTNAME	CROSS_CORR	Extension name
CRVAL	3900	Starting wavelength in angstrom
CRPIX	1	Starting pixel position in x-direction
CDELT	0.02	Wavelength pixel interval in angstrom
SPEC_ID1	1039475	Unique L1 spectrum identifier linked to SPEC_ID Extension 2 Binary table for row 1 . [Unique L1 spectrum identifier R1]
SC_ID1	45	Unique best fit template spectrum identifier linked to SC_ID for previous SPEC_ID in Extension 2 Binary table for row1 of this extension. [Unique best fit template identifier R1]
<Line>_POS1	6568.56	Spectral line <line> wavelength position in angstrom for best fit model (one for each line) [<line> wavelength1, ang]
<Line>_WIDTH1	5.45	Spectral line <line> width in angstrom for best fit model (one for each line) [<line> width1, ang]
<Line>_FLUX1	1.234e-18	Spectral line <line> flux in ergs/sec/cm^2 for best fit model (one for each line) [<line> flux1, ergs/sec/cm^2]
<Line>_CONT1	2.362e-20	Spectral line <line> position continuum flux in ergs/sec/cm^2 for best fit model (one for each line) [<line> cont1, ergs/sec/cm^2]
…..		
SPEC_IDN	4759602	Unique L1 spectrum identifier linked to SPEC_ID Extension 2 Binary table for row N [Unique L1 spectrum identifier RN]
SC_IDN	34	Unique best fit template spectrum identifier linked to SC_ID for previous SPEC_ID in Extension 2 Binary table for rowN of this extension. [Unique best fit template identifier RN]
<Line>_POSN	3727.67	Spectral line <line> wavelength position in angstrom for best fit model (one for each line) [<line> wavelengthN, ang]
<Line>_WIDTHN	4.5634	Spectral line <line> width in angstrom for best fit model (one for each line) [<line> widthN, ang]
<Line>_FLUXN	0.453e-18	Spectral line <line> flux in ergs/sec/cm^2 for best fit model (one for each line) [<line> fluxN, ergs/sec/cm^2]

<Line>_CONTN	1.209e-20	Spectral line <line> position continuum flux in ergs/sec/cm^2 for best fit model (one for each line) [<line> contN, ergs/sec/cm^2]

7 Further Information on Provenance

Within 4XP there are a number of input data tables and spectral products that are used by our fitting processes, which will be useful to consortium members to reconstruct information from the 4XP level 2 data products. These tables consist of a set of templates used for redshift fitting, a set of SPM templates for stellar continuum fitting and a line list used for emission/absorption line fitting. Below are detailed a preliminary description of the form these datasets may take.

7.1 Input data used

Within 4XP there are a number of input data tables and spectral products that are used by our fitting processes, which will be useful to consortium members to reconstruct information from the 4XP level 2 data products. These tables consist of a set of templates used for redshift fitting, a set of SPM templates for stellar continuum fitting and a line list used for emission/absorption line fitting. Below are detailed a preliminary description of the form these datasets may take.

7.1.1 4XP-Z Template Set (Binary Table)

4XP-Z will use a sample of spectral templates provided by the survey teams and perform a Fourier cross-correlation to the reduced L1 spectrum to determine best-fit redshift. These spectral templates will be provided as a binary table of template spectrum parameters given in Table 7-1.1.

7.1.2 4XP-Fit Template Set (Binary Table)

4XP-Fit will use spectral continuum templates to fit galaxy properties using chi^2 minimization techniques. These spectral templates will be provided as a binary table of template spectrum parameters given in Table 7-1.2.

7.1.3 4XP-Fit Line List (Binary Table)

4XP-Fit will use also use a line list to fit Gaussian models to specific lines at the galaxies previously derived redshift. These line lists will be provided as a binary table given in Table 7-1.3a.

Current lines to be included in this table are given in Table 7-1.3b (note that these are an example set of lines that maybe included and this line-list is subject to change based on survey team requirements and pipeline development). Current line list is taken from SDSS - http://classic.sdss.org/dr7/algorithms/speclinefits.html#linelist.

Table 7-1.1 4XP-Z template list Binary Table Columns

TTYPE	TTYPE comment	TF	TFORM comment	TUNIT	TUNIT comment	TNULL	Data Range	TDISP	TDMIN	TDMAX	TUCD
ID	Unique template identifier	I	data format of field: 2-byte INTEGER	-	none	NULL	Positive-integer	I5.1	0	1000	meta.id
SURVEY	ID Survey team that provided this template	I	data format of field: 2-byte INTEGER	-	none	NULL	Positive-integer	I5.1	0	50	meta.id
TYPE	Type of template (star, SF galaxy, passive Galaxy, AGN, QSO)	A	data format of field: ASCII Character	-	none	empty string	Alpha-numeric	A16	none	none	meta.id
z	Intrinsic redshift of template	D	data format of field: 8-byte DOUBLE	-	physical unit of field	NULL	Positive-real	D9.2	0	10	src.redshift
MAG	Intrinsic magnitude of template	D	data format of field: 8-byte DOUBLE	AB mag	physical unit of field	NULL	Real	D9.1	-30	30	phot.mag
BAND	Band of intrinsic magnitude of template	A	data format of field: ASCII Character	-	none	empty string	Alpha-numeric	A16	none	none	meta.id
WAVE	Vector of wavelengths	D	data format of field: 8-byte DOUBLE vector	Angstrom	physical unit of field	NULL	Positive-real	D9.2	3000	15000	em.wl
FLUX	Vector of flux values	D	data format of field: 8-byte DOUBLE vector	Ergs/sec/cm^2/ Ang	physical unit of field	NULL	Real	D9.1	-1	1	phot.flux

Table 7-1.2 4XP-Fit template list Binary Table Columns

TTYPE	TTYPE comment	TF	TFORM comment	TUNIT	TUNIT comment	TNULL	Data Range	TDISP	TDMIN	TDMAX	TUCD
ID	Unique template ID	I	data format of field: 2-byte INTEGER	-	none	NULL	Positive-integer	I5.1	0	1e7	meta.id
SPM_TYPE	Origin of SPM (BC03, M05)	A	data format of field: ASCII Character	-	none	empty string	Alpha-numeric	A16	none	none	meta.id
AGE	Age of SPM	D	data format of field: 8-byte DOUBLE	Gyr	physical unit of field	NULL	Positive-real	D9.2	0	20	time.age
METAL	Metallicity of SPM	D	data format of field: 8-byte DOUBLE	Zsun	physical unit of field	NULL	Positive-real	D9.2	0	02	phys.abund.Z
eFOLD	E-folding time of SPM	D	data format of field: 8-byte DOUBLE	Gyr	physical unit of field	NULL	Positive-real	D9.2	0	20	time.interval
z	Intrinsic redshift of SPM	D	data format of field: 8-byte DOUBLE	-	physical unit of field	NULL	Positive-real	D9.2	0	10	src.redshift
MAG	Intrinsic magnitude of SPM	D	data format of field: 8-byte DOUBLE vector	AB mag	physical unit of field	NULL	Real	D9.1	-30	30	phot.mag
BAND	Band of intrinsic magnitude of SPM	A	data format of field: ASCII Character	-	none	empty string	Alpha-numeric	A16	none	none	meta.id
WAVE	Vector of wavelengths	D	data format of field: 8-byte DOUBLE Vector	Angstrom	physical unit of field	NULL	Positive-real	D9.2	3000	15000	em.wl
FLUX	Vector of flux values	D	data format of field: 8-byte DOUBLE Vector	Ergs/sec/cm^2/ Ang	physical unit of field	NULL	Real	D9.1	-1	1	phot.flux

Table 7-1.3a 4XP-Fit line list Binary Table Columns

TTYPE	TTYPE comment	TF	TFORM comment	TUNIT	TUNIT comment	TNULL	Data Range	TDISP	TDMIN	TDMAX	TUCD
ID	Unique line ID	I	data format of field: 2-byte INTEGER	-	none	NULL	Positive-integer	I5.1	0	1e7	meta.id
NAME	Unique line name	A	data format of field: ASCII Character	-	none	empty string	Alpha-numeric	A16	none	none	em.line
WAVE_ANG	Line Vaccum Wavelength in Angstrom	D	data format of field: 8-byte DOUBLE	Ang	physical unit of field	NULL	Positive-real	D9.2	1000	15000	em.wl
WEIGHT_G	Line weight for galaxy (0-10)	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real (0:10)	D9.2	0	10	spect.strength
WEIGHT_A	Line weight for AGN (0-10)	D	data format of field: 8-byte DOUBLE	-	none	NULL	Positive-real (0:10)	D9.2	0	10	spect.strength
STELLAR	Is line from stellar absorption?	L	Logical TRUE/FALSE	-	none	NULL	TRUE/FALSE	L1	none	none	meta.code
AGN	Is line from AGN/QSO?	L	Logical TRUE/FALSE	=	none	NULL	TRUE/FALSE	L1	none	none	meta.code

Table 7-1.3b 4XP-Fit line lists

ID	NAME	WAVE_ANG	WEIGHT_G	WEIGHT_A	STELLAR	AGN
1	Lya	1215.67	0	9	FALSE	TRUE
2	OIIB	3726.09	5	1	FALSE	FALSE
3	OIIR	3729.88	0	0	FALSE	FALSE
4	HB	4862.68	2	4	FALSE	TRUE
5	OIIIB	4960.30	2	0	FALSE	FALSE
6	OIIIR	5008.24	3	2	FALSE	FALSE
7	HA	6564.61	8	8	FALSE	TRUE
8	NIIR	6549.86	3	0	FALSE	FALSE
9	NIIR	6585.27	3	0	FALSE	FALSE
10	SIIR	6718.29	3	0	FALSE	FALSE
11	K	3934.78	-1	0	TRUE	FALSE
12	H	3969.59	-1	0	TRUE	FALSE
13	GBand	4305.61	-1	0	TRUE	FALSE
14	Mg	5176.70	-1	0	TRUE	FALSE
15	Na	5895.60	-1	0	TRUE	FALSE
16	NV	1240.81	0	3	FALSE	TRUE
17	CIV	1549.48	0	8	FALSE	TRUE
18	CIII	1908.73	0	7	FALSE	TRUE
19	MgII	2799.11	1	8	FALSE	TRUE

7.2 Algorithms and software

For descriptions of how these products are derived, see the IWG8 management plan and requirements/software description document.

7.3 Physical meaning of parameters

Please see descriptions in Table 5. These will be expanded upon when parameters are finalized.

7.4 Flag Definition Table

No flags require definition here.

Appendix A Interface Control Drawings

None

Appendix B List of Acronyms

List of Acronyms	
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 Surveys
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
PS	Project Scientist
QC	Quality Control
RD	Reference Document
TBD	To be defined
4XP	4MOST extragalactic analysis pipeline
4XP-Z	Redshifting branch of 4XP
4XP-Sim	Simulation branch of 4XP
4XP-Fit	Spectral fitting branch of 4XP
SPM	Stellar Population Model