Review report of the Experiment to Archive Interface Control Document (EAICD)

EAICD: XXXXX (e.g. ACP)

1.2.3 Preamble

As an independent reviewer of the Huygens archive, you are asked to read first the Data Archive Plan. This document provides you with an overview of the archiving activities within the Huygens project. It also summarizes the products that will be archived (see the appendixes C to K). Those products have been negotiated with the teams and are listed in the table:

1.2 Experiment	**Raw data**	**Calibration information**	**Reduction algorithm**	**Calibrated data**	**High level data**
HASI	X	X		X	X (TBD)
SSP	X	X	X	TBD	
ACP	X	X		X	
GCMS	X	X		X	X (TBD)
DISR	X	X	X		
DWE	X	X	X	X	
DTWG			X		X
HouseKeeping	X	X		X	

TBD: To be defined

Then, you are asked to read the individual Experiment to Archive Interface Control Documents. You are only assigned to review some of these documents, depending on your field of expertise.

The EAICD provides users of each experiment with a detailed description of the product and a description of how it was generated, including data sources and destinations. Also, it is the official interface control document between each team and the archiving authority.

As the EAICD itself will be part of the documentation of each data set, it is one of the entry points for scientists interested in the Huygens data. It is therefore very important that such a document must be clearly written.

The Huygens Data Archive Team (Olivier Witasse, Joe Zender for ESA; Lyle Huber for PDS) has been working with the teams and helped them generating this document, which contains three important parts:

Section 2: Overview of process and product generation

Section 3: Archive format and content

Section 4: Detailed interface specifications

We ask you to answer to the following questions:

Main Comments

I do not understand the many sensors, transducers, and blocks that make up PPI. Several sections concerning PWA are incomplete. ACC statistics are not defined. It is unclear whether "derived parameters" as defined by the table in section 2.5 will be archived or not. If they are not, then this archive will not be useful to the scientific community. Lots of products "are expected to be" archived. Such non-binding language should not be used in this document. Times in the archived data files appear to be native times, not UTC. If the "'derived parameters" are archived, then additional information must also be archived so that they can be reproduced by other users.

First topic: Structure and completeness of the EAICD.

Question #1: In your point of view, is each section understandable? Do you miss some important information?

Section 1.
1.2 discusses software, but no software will be archived.
RD.4 - 7 are not referenced.
The acronyms list is very long, suggesting that there are too many acronyms in the document. Some acronyms used in the text are not in this list (eg IPMO)
Mailing addresses and telephone numbers should be given for all people in 1.8.

Section 2.
Reference to AD.4. There is no AD.4. AD.4 should be available to users, ie not an internal Huygens document.
Many of the acronyms in Fig. 1 are not defined, eg Relax., RX, TX, EL, Acoustic, COG.
Figure 2 is very detailed and hard to read.
2.2 How is the probe symmetry axis related to XYZ? X-servo/Xservo, choose one only. X-servo selection - selection of what? Values are arithmetically averaged - in the DPU, the main Huygens computer, on Cassini, or in Padua?
PPI. I am very confused by the many sensors, transducers, and blocks. What do the 6 reference sensors measure? What do the 7 constant sensors measure? I do not know what the "2", "1", or "P" in 2.1P means.
What is PWA-ACDC?
2.3 I presume that Tdata = T0+10s defines Tdata, rather than being a rough guide to when Tdata will occur. If so, state this more clearly. If not, then how is Tdata (and others) defined? Is "last km" based on the altimeter-derived range to surface or something else?

Table 3 is helpful to users. What is Time (min) referenced to? What triggers Td1, Tpmed, and Tphigh?
2.3.1 What are statistics data? Means, variances, max, min, or something else?
2.3.3 PPI Calibration report. What RD is this?
2.3.4 Is AC_DC the same as ACDC? Table 3 says that PWA mode D has no RP, but 2.3.4 says otherwise.
2.3.5 What are the times next to THASI, TACC, etc, referenced to? This table is also a good summary for the user. However, I expect that its format could easily be misread by older/newer versions of Word, so a Figure might be more stable. I am surprised the ACC XYZ PIEZO will hardly operate during descent, this data might have been useful for understanding the probe attitude. I accept that (a) it's too late to change and (b) there were probably good reasons for doing so.
2.1 (after 2.3.5) Section numbers are messed up.
2.3 This references the TEM calibration report as RD.4. Earlier, RD.4 was referenced as the TEM calibration report. This is not helpful to users.
Who decides whether the “calibration derived from analysis...” is “necessary”? HASI? ESA? JPL?
2.5 This table suggests that the derived parameters listed will be part of the archived data products, yet, a few lines later, 2.8 says that HASI does not commit to archiving any derived products. A long list of (eg) resistances from TEM is of no use to users who want properties of Titan's atmosphere (eg temperatures). The subsequent sections (3 and 4) mix measured and derived parameters without saying which WILL be archived and which MAY (or MAY NOT) be archived.
Underscores and hyphens are mixed in some of the PWA names, which is very confusing. Why not just use underscores?
What does pck mean? What does ACDCACU mean?
2.7 “Documentation could be provided” Will it be provided or not?

Section 3.
PDS likes CAPITALS in file and directory names, not /calib or /data.
3.1.1 What is the difference between raw, converted, and calibrated data?
3.1.2 Level 4. There is no commitment to archive any Level 4 data.
3.1.3 This list of subdirectories appears often, sometimes it contains PROFILES and sometimes it doesn't.
3.1.4 Datatypes should all be CAPITALS. Give the full list of datatypes. I don't see PPIXY anywhere in Section 4.1.3. IMP is not listed in the modes.
3.4.2 This table is repeated from earlier.
3.4.3.1 /HASI/data is not root. /HASI/ is root.
3.4.3.3 INSTHOST.CAT, INST.CAT, PERSON.CAT, not HOST.CAT, HASI.CAT. PERSONNEL.CAT, see fig 19.4 in PDS Standards
3.4.3.9 Is a LABEL directory optional or required by the PDS? I thought they were always part of an archive.
3.4.3.10 This list of documents is vague.
3.4.3.12 “contains all the data products of HASI” - all the data products in the table from Section 2.5 (which includes derived parameters) or not? Will the PROFILES directory be included or not?

Section 4.
4.1 List of {sensor} is not the same as 3.1.4. I've no idea what “nsession” refers to.
4.1.1 “are expected to be” is nonbinding. Use binding language instead.
I don't know what STAT refers to. What kind of statistics?
The Data Archival Plan says that the time standard is UTC. Will these times (and those in the other files) be UTC?
In HASI_ACCD_X..., what is X?
Regarding HASI_ACC_4..., does HASI commit to including this derived data product or not? Altitude is not measured in units of “h”. What is this altitude relative to? A 2575 km sphere? The radar-detected surface? Something else? Will this altitude be identical to the DTWG trajectory or not? What is Ref sensor?
4.1.2 What are VF, 0VF, etc?
4.1.3 “are expected to be” is nonbinding. Use binding language instead. What does NSESSION refer to? What are Ysi1 and 2?
Level 4 products, velocity is not measured in m/s2.
4.1.4 “are expected to be” is nonbinding. Use binding language instead. Filenames here are like HASI_PWA_2_ACU... whereas filenames defined in 3.1.4 are like HASI_PWA-ACU_2... There are no descriptions of the table contents for Level 2 and not even filenames for Levels 3 and 4.
4.1.5 “are expected to be” is nonbinding. Use binding language instead. “DPU” is used in filename, 3.1.4 suggests “HK” instead. The E/D/S mode is not included in this filename, unlike 3.1.4.
4.1.6 “are expected to be” is nonbinding. Use binding language instead. Altitude is not measured in units of “h”. The PWA pressures should be in Pa for consistency with everything else.
4.2 refers to the LABEL directory. There is no LABEL directory. Table 5 - Only tables 1-3 are numbered.
Is it XServo or Xservo? Filenames for the ACCI Piezo stats (S) are not in CAPITALS.
PWA files are not defined.
HP_HASI_4_ALTITUDE_PROFILE has 3, not 4, columns.
Section 4.3 What is the difference between data product A and B? These terms are not used anywhere else.
Spacecraft clock start/stop count is given to nearest second. Should milliseconds be used?
Label A - Instrument name and type are not stated clearly. Quality.cat is mentioned, but it isn't in section 3.4.3.3. PDS standards 5.1.2 says that only CAPITALS should be used in the label, unlike the product names here.
Label B - Column 8, gain, needs a description.

Question #2: Is the EAICD itself understandable with respect to potential future users (taking into account the long-term preservation of the data - overall coherence of the document)?

Details about the PPI are not clear to me. Several sections concerning PWA are sketchy. The operational modes are discussed well. A consistent directory and filenaming structure is clearly described.

Question #3: Is the EAICD coherent with the Data Archive Plan? Check in particular the conformance to the standards (section 6.6) and to the appendixes.

Appendix G of the Data Archive Plan says that HASI will provide Level 4 products. If only the "measured parameters" of the table in section 2.5 are archived, then the archive will be useless to most users. Many products "are expected to be provided", which is not a binding commitment. These issues should be corrected before this EAICD is coherent with the Data Archive Plan.

Second topic: Scientific and technical content.

Question #1: Are the scientific objectives clearly and concisely described (in section 2)?

Yes.

Question #2: The processes involved in the data flow from the Huygens probe to the ESA Planetary Science Archive are very important, in order to understand how the data are processed and transformed.
Is it clearly described in the document (section 2)?

I don't understand the PPI overview. It is not clear to me what the various parts of PWA actually measure. ACC statistics are not defined, yet they seem to be a major data product. The operating modes are well defined.

Has the team committed to provide to providing algorithms that will allow long-term use of the data and comparison to future datasets?

Section 2.3 promises that calibration reports will be available to users and that they will be detailed. The Level 3 ACC data is not very useful, Level 4 is needed.

Question #3: Taking into account the reference papers (e.g. the space science review paper), do you think that the data products are clearly identified?
Are they clearly described?
Will these products support the scientific goals?

Yes, except for PWA. Level 4 data products for ACC are needed to do any science.

Question #4: Do you think the calibration information is carefully addressed (sections 2 and 3)?

The reader is directed to calibration reports, which were not available to me.

Question #5: Do you think that the validation of the data is carefully addressed in this document (section 3)?

Three lines of non-specific text discuss validation.

Question #6: Is the geometrical information addressed?

N/A

Question #7: Is the set of documentation (intended to be delivered with the dataset) is complete and sufficient for data calibration and processing, data visualization and analysis?

The calibration documents are not available, so it is incomplete. The derivation of densities from entry acc data cannot be reproduced without the aerodynamic database, the probe mass and area, the entry state, and more details on the ACC positions and alignments.

Third topic: Long-term access to the data.

The data will be archived under the directory /DATA. Each team is free to organize the content of this directory. For each data product (e.g. a table, an image, etc..), a label file is provided.

Question #1: Are the selected data structure clear and useful (section 3)?

Yes, except for PWA.

Question #2: The filenaming convention is explained in section 3. Please comment on the specific choices that have been made.

The choices are good, but there are a few inconsistencies that I have highlighted above.

Fourth topic: Data Product Labels

PDS data products labels are required for describing the content and format of each individual data products within a data set. Examples of label are given in section 4 of the EAICD.

Question #1: From the proposed labels (see section 4), is the list of keywords clear and understandable?

Only two example labels are given, which are not enough to describe all the labels that will be used for the various data products. I don't know what the differences are between A and B, nor which files will get A and which will get B. The impact trace doesn't seem to fit the format of either of the two labels offered. The various names and types in these labels are clear.

Question #2: In the proposed table objects, is the description of the columns clear enough? (Column name, text description, unit…)

Yes

Please list here the additional comments you may have on this document, if any.

1 XXX EAICD	[Minor] List of editorial comment

2 XXX EAICD	[Major] Comment in section xx

Please list here the additional comments you may have on the Data Archive Plan (D.A.P.), if any.

3 D.A.P.	Comment in section xx

4 D.A.P.	Comment in section yy