WORLD METEOROLOGICAL ORGANIZATION
_____

COMMISSION FOR BASIC SYSTEMS
OPEN PROGRAMME AREA GROUP ON INTEGRATED OBSERVING SYSTEMS

EXPERT TEAM ON SATELLITE UTILIZATION AND PRODUCTS

EIGHTH SESSION

GENEVA, SWITZERLAND, 14-17 APRIL 2014

ET-SUP-8/Doc. 7.1
(13.IV.2014)
_____

ITEM: 7

Original: ENGLISH

SATELLITE DATA ACCESS – USER PERSPECTIVE FROM RA VI

Report of the Data Format Task Group

(Submitted by Simon Keogh (UKMO), Enrico Fucile (ECMWF), Simon Elliott (EUMETSAT))

Summary and Purpose of Document

The purpose of this document is to propose some actions and recommendations to address the concerns of satellite data users related to the perceived difficulty in utilizing data in various WMO standard formats. This work stems from an initial impression from the WMO survey 2012 user survey that users find data formats to be challenging to use, particularly if they are not working for a National Meteorological or Hydrological Service (NHMS). This paper aims to state what requirements users have for various "high level" use cases for satellite data.

ACTION PROPOSED

- The eighth session is invited to consider the following Actions and Recommendations:

Action 1: To expand and agree on the high level use cases outlined in this paper.
Action 2: ECMWF to fully publicise its "ecCodes" tools, documentation and training via the WMO web site along with a short-medium term roadmap of how these tools will evolve.
Action 3: Secretariat to make other relevant WMO groups aware of this paper.

Recommendation 1: Secretariat to engage with Logical Data Modeling teams coordinated by WMO to ask them to provide a long term roadmap on how and when users can expect format conversion tools and training to be available. This roadmap should be a "laymans guide" and assume no prior knowledge of data modelling techniques or language.
Recommendation 2: Secretariat to consider reinstating ET-ADRS to address specific issues on data representation and to update their paper
Recommendation 3: ECMWF should propose a plan (including resources required) for running a community based development of their software to provide unified access to data in several formats.

_____

DISCUSSION

Satellite Data Access – A user perspective for region VI.
DRAFT – For Review by ET-SUP

Simon J. Keogh[1], Enrico Fucile[2], Richard Weedon[1], Simon Elliott[3]
[1]Met Office, [2]ECMWF, [3]EUMETSAT.

Summary
The purpose of this document is to propose some actions and recommendations to address the concerns of satellite data users related to the perceived difficulty in utilizing data in various WMO standard formats. This work stems from an initial impression from the WMO survey 2012 user survey [1] that users find data formats to be challenging to use, particularly if they are not working for a National Meteorological or Hydrological Service (NHMS). This paper aims to state what requirements users have for various "high level" use cases for satellite data.

1. Levels of satellite data.

It is important to note that satellite data can be made available to users with several "levels" of processing applied to the raw satellite data. Each level of processing takes the user further away from raw instrument output and closer towards geophysical products that are integrated in both space and time, sometimes with other geophysical information. Some users prefer access to raw data (e.g. instrument scientists) and others prefer access to time averaged data (e.g. for climate monitoring applications).

Table 1. Data levels (from NASA web page [2]).

Data level	Description
Level 0	Reconstructed, unprocessed instrument and payload data at full resolution, with any and all communications artifacts (e.g., synchronization frames, communications headers, duplicate data) removed. (In most cases, the EOS Data and Operations System (EDOS) provides these data to the data centers as production data sets for processing by the Science Data Processing Segment (SDPS) or by a SIPS to produce higher-level products.)
Level 1A	Reconstructed, unprocessed instrument data at full resolution, time-referenced, and annotated with ancillary information, including radiometric and geometric calibration coefficients and georeferencing parameters (e.g., platform ephemeris) computed and appended but not applied to Level 0 data.
Level 1B	Level 1A data that have been processed to sensor units (not all instruments have Level 1B source data).
Level 2	Derived geophysical variables at the same resolution and location as Level 1 source data.
Level 3	Variables mapped on uniform space-time grid scales, usually with some completeness and consistency.
Level 4	Model output or results from analyses of lower-level data (e.g., variables derived from multiple measurements).

2. Data Formats.

WMO Expert Team on the Assessment of Data Representation Systems (ET-ADRS) has evaluated various ways of representing data and made available a comprehensive comparison of the data representation systems currently used [3]. There are two agreed WMO standard formats for satellite data dissemination to users. These are HDF5 (for imager instrument data) and BUFR

(for sounder instrument data). Although these formats are agreed, satellite providers sometimes fail to use the correct format for the correct purpose.
Main NWP centres also use GRIB, CREX. For WIS a much broader range of formats is envisaged so work will be needed to develop a) a standard data model for satellite data and b) a set of tools to covert between formats for different user communities (interoperability). However, it is likely to be many years before this solution is realised so until then it will be necessary for users to make the best of what tools have already been developed to help them access satellite data. It will be a challenge for WMO to make these tools more visible to users to ensure that they do not waste their resources developing tools that already exist and are widely used in other communities. In addition, it will be a challenge for WMO to ensure that users are kept abreast of developments with WIS, the Open Data Model and the development of tools to exploit products described in the WIS.

3. Data Access Tools.

Regardless as to which format data is supplied in, users are still faced with the challenge of accessing the data. Very often, expert users within NHMSs develop their own tools (either locally or in collaboration with other NMHSs in the same user community) to facilitate access to standard satellite products. However, for users outside these expert communities, there is a lack of awareness as to what tools have already been developed that could help them with the application. This can lead to user frustration and even to users abandoning using the satellite data altogether.
WMO maintains a “living” web page [4] that lists all the commonly available software packages for accessing commonly available satellite data. Some packages are open source and some are provided by commercial vendors. We do not discriminate between these tools here or advocate any in particular. It is a challenge for ET-SUP to keep this web page up to date.

The accessibility of data from the scientific user community is an area in which several organisations are working with different approaches. ECMWF has developed a GRIB decoder/encoder (grib_api) with the purpose of giving a simple application programming interface and a set of tools for the user who is not an expert of data format and needs to access the data in GRIB. The success of this simple model of accessing the data has suggested to extend it to other formats and ECMWF is at the moment developing a general purpose decoder (ecCodes) which will be able with the same function calls and therefore the same programming syntax to decode BUFR, GRIB, some Traditional Alphanumeric Codes, GTS headers and likely many other formats. The aim is to provide a single library and therefore a unified syntax for accessing the data. This will improve interoperability between different data formats and at the same time will provide a simple access to this complex data formats for scientific users. The principle of this library is very simple. It provides some “get” and “set” functions which allow the access to elements of the data format hiding the details of the binary format which can be very complex and not needed for the use and interpretation of the data themselves.

An application programming interface (API) accessing different data formats can provide a unified syntax to access the data, but the difference in semantics will still be present and the user will be forced to ask the producer or search for a source of semantics provided by WMO or some other Production Centre for the interpretation of the data. The API should be capable of providing this semantics information or of pointing to the web resource providing it.

It has to be considered that all the data formats currently used are based on a binary representation and a given semantics. The focus it has always been to develop the best format to satisfy some requirements from a specialised user community. All the formats produced are rich of useful features and affected by problems and they are all accessed with different software providing different syntax that the user has to master, sometimes spending great effort. Developing the best format to satisfy the needs of a community is a very complex and expensive task.

It is recommended to focus on the development of community software able to provide unified access to several data formats to make easier for the user to access the information in a data set.

The software under development at ECMWF can be a good starting point that can be further developed for the needs of the scientific community.

4. High level use cases for satellite data

Table 2. High level use cases and high level user requirements.

Use case	Hardware	Software	Network	Skills needed
Download	Terminal	Data transfer tools	High bandwidth connection to data providers via GTS, Internet, Direct Broadcast or EUMETCast	Ability to write scripts for data reception and routing.
Pre-process	Server	Pre-processing tools	LAN	Good knowledge of where to get software. Good scripting skills. Ability to understand specialist documentation.
Archive	Server	Archiving tools	LAN	Database use, maintenance and management.
Discover	Terminal	Catalogue tools	LAN	Catalogue use and management.
Convert	Terminal	Format Conversion tools	LAN	Good scripting skills.
Transform	Terminal	Data analysis and processing tools	LAN	Data analysis. Subject matter expertise.
Visualise	Terminal	Visualisation tools	LAN	Tools expertise. Subject matter expertise.
Disseminate	Server	Data transfer tools. Service Management tools.	High bandwidth connection to downstream users/systems	Ability to write scripts for data routing. Service Management.

5. Recommendations and Actions

Action 1: ET-SUP to expand and agree on the high level use cases outlined in this paper.

Action 2: ECMWF to fully publicise its “ecCodes” tools, documentation and training via the WMO web site [4] along with a short-medium term roadmap of how these tools will evolve.

Action 3: Secretariat to make other relevant WMO groups aware of this paper.

Recommendation 1: Secretariat to engage with Logical Data Modeling teams coordinated by WMO to ask them to provide a long term roadmap on how and when users can expect format conversion tools and training to be available. This roadmap should be a “laymans guide” and assume no prior knowledge of data modelling techniques or language.

Recommendation 2: Secretariat to consider reinstating ET-ADRS to address specific issues on data representation and to update their paper [3]

Recommendation 3: ECMWF should propose a plan (including resources required) for running a community based development of their software to provide unified access to data in several formats.

6. References.

1. WMO User Survey 2012 (http://www.wmo.int/pages/prog/sat/documents/SAT-PUB_SP-9-Survey-Report-2012.pdf)
2. NASA data levels web page (http://science.nasa.gov/earth-science/earth-science-data/data-processing-levels-for-eosdis-data-products/)
3. ET-ADRS 1st Meeting Report Version 7, 23-25 April, Washington, 2008 (http://www.wmo.int/pages/prog/www/WDM/ET-ADRS-1/ET-ADRS-1_draft_report_v7.doc)
WMO Data access page (http://www.wmo.int/pages/prog/sat/accessandtools_en.php)