|  |  |  |
| --- | --- | --- |
| WORLD METEOROLOGICAL ORGANIZATIONCOMMISSION FOR BASIC SYSTEMS-----------------------------THIRD MEETING OF INTER-PROGRAMME EXPERT TEAM ONDATA REPRESENTATION MAINTENANCE AND MONITORINGBEIJING, CHINA, 20 - 24 JULY 2015 |  | IPET-DRMM-III / Doc. 3.3 (1)(7. 7. 2015)-------------------------ITEM 3.3ENGLISH ONLY |

BUFR AND CREX

**On the timely addition of new data representation features and the impact of requested changes in BUFR Editions 4 and 5**

*Submitted by* *Yves Pelletier (Canada)*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Summary and Purpose of Document**

The list of proposed requirements for BUFR Edition 5 is presented anew with some additional proposals and an updated discussion

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**ACTION PROPOSED**

   The meeting is requested to consider adopting the proposals herein.

**ANNEXES:**

 1. None

**DISCUSSION**

Numerous feature requests are on the record for BUFR Edition 5. Several of these requests were initially made for the current edition, BUFR Edition 4, but, after discussion, were slated as proposals for BUFR Edition 5 instead. Typically, the argument revolved around the impact that the requested feature would have, in that it would require a change in the BUFR Edition 4 software and therefore incur software development costs for operators. It has long been a major criterion for the IPET-DRMM and its predecessor teams that within a given BUFR edition, the software specification for the encoder/decoder shall remain unchanged.

Thus, potentially useful innovations or feature requests are deferred for years even though in several cases Team members have agreed they are the most appropriate from a technical point of view.

It may not be helpful, in the bigger picture, to put further development off for years until we are ready with a new, all-encompassing, specification for the next BUFR Edition. It is doubtful that this approach meets all the adaptability requirements stemming from developments in Earth Observation data. So we will find some operators requesting new features to meet a growing set of operational needs and being stymied by a lack of agility.

We will also find a substantial population wishing for the exact opposite. Following the Migration to Table-Driven Code Forms (MTDCF) and the deployment of BUFR in the world surface and aerological observation systems, we are likely to find many operators indicating a preference that BUFR remain unchanged, in order to reduce the need for migrations and to promote operational stability. With this in mind, we believe it is likely that BUFR Edition 4 will remain in use for a significant period of time, even though it has become clear, through BUFR Edition 5 requirement gathering, that new regulatory and operational requirements will go beyond the capacities of BUFR Edition 4.

**Classifying feature requests**

Within the IPET-DRMM and the former IPET-DRC, the impact of proposed new features has often been discussed at length for specific cases, but never formally categorized in a broader context. Yet, it is also known that the impact of proposed changes can vary greatly. Now that we have a substantial list of proposed requirements for BUFR Edition 5, an analysis of the list allows us to suggest a categorization of impacts in three broad cases, by order of potential disruptiveness:

1. **Light impact: New features have limited scope. They do not affect the encoding/decoding of past or future data products that do not specifically use them.**

For example, most requests for new operators would fall into this category. New lines of software code are needed to enable the encoding/decoding of messages using the new feature, but otherwise, the change is neutral. Drawing an analogy with GRIB2, it could be noted that new Product Definition Templates also require new lines of code, yet do not require a change in edition number.
2. **Medium impact: New features result in backward incompatibility of the bit pattern within BUFR messages.**

The new features alter some underlying aspects of BUFR, but require no change to the Data Description Syntax (DDS) regulations. The existing elements and structures of Table B and D are preserved.

In this scenario, the same data, encoded with “old” and “new” BUFR, would use the same Tables B and D, and yet not produce identical binary streams. The “new” BUFR could only be decoded with a new decoder and therefore this certainly means a new Edition. This is conceptually similar with previous Edition transitions from Edition 2 to Edition 3, and from Edition 3 to Edition 4.
3. **Major impact: New features result in backward incompatibility of Table B and/or Table D.**

For example, the new features alter the DDS, such as changes that would be required to make BUFR compliant with the ISO 19000 family of standards. Table B and/or Table D would not be compatible across BUFR Editions. This requires a new Edition number. The scope of the changes induced by such new features is comparable with the migration from GRIB Edition 1 to GRIB Edition 2.

**PROPOSAL**

We propose a new change management strategy for BUFR Editions 4 and 5.

We request the Team to consider the possiblity of allowing a BUFR Edition 4 to evolve incrementally, to a limited degree.

We propose that circumstances lend themselves to a two-stream approach. In the first stream, the Team may consider whether an incremental, “point-release” approach might be applicable. This would allow for *light impact* changes, as defined above, to BUFR Edition 4 specification (thus leading to 4.1, 4.2…). This is not resource-neutral, but it allows for incremental evolution of the specification with no direct impact on existing templates or operational applications. It also provides for a suitably long transition period to let the MTDCF take hold with BUFR Edition 4.

The second stream would seek to create a BUFR Edition 5 platform that would essentially be a reengineering of the BUFR Code Form, keeping in mind the BUFR design goals[[1]](#footnote-1). The most important new design goals would be the implementation of a logical model approach for data representation and conformity with the ISO 19000 series of standards.

1. The following list of design goals was presented at the first meeting of the IPET-DRMM in 2013 and seemed uncontroversial at the time:

	1. Universality in the representation of data and metadata (achieved through the BUFR data representation syntax and the various code tables) Extensibility (meaning that new data elements or sequences can be added without changes to the encoding/decoding software; achieved by Table-driven approach)
	2. Portability across computer architectures (achieved by use of unsigned octets, scale, reference value and bit-width)
	3. Compactness (achieved by Table-Driven approach and representation of code values in binary form, using as few bits as possible).
	4. Lossless conveyance of information (no lossy compression is used. Also, while current editions do not focus on data integrity in the technical sense of the word, there is a concern for sufficient precision to preserve the accuracy of the data being conveyed.)
	5. Focus on operational data exchange between data processing centres And in light of discussions at the joint ET-MDI and IPET-DRC meeting in Exeter, we might add a new, proposed fundamental requirement:

	1. *Interoperability with ISO standards for data and metadata representation.* [↑](#footnote-ref-1)