

| 1 | 2 | (3) | 4 | 5 | (6) | (7) |
|-----------------|---|--|------------------------------|--|---|--|
| MB ¹ | Clause No./ Subclause No./ Annex (e.g. 3.1) | Paragraph/ Figure/Table/ Note (e.g. Table 1) | Type of comment ² | Comment (justification for change) by the MB | Proposed change by the MB | Proposed Editors Disposition |
| VXML 1 | 7.3 | 7.3.2 | Te | Although UTF-8 is the dominant encoding that may not be the case in the future. | Make UTF-8 the default but allow UTF-16 and possibly UTF-32 | Accept |
| VXML 2 | 7.3 and 7.4 | 7.3.8 and 7.4.3 | Te | Original-channel in Session header and Channels in Turn header Recording is unlikely to change the audio much unless lossy compression is used. Even so, the changes that will occur will apply to all the audio recorded. Also, use of a wav file will always be a single channel | 1. Restore the Channels element in Turn to its prior state (per document M1 071 24) 2. Add an optional element to the Session header for the recorded utterance. Consider using a codec name or listing only the attributes that will be affected, notably audio format, compression, precision and the SNR estimate | Accept Information about the recording of the end user utterance to be moved to the Session level but need more discussion regarding prompts. Prompts be recorded, recorded + text (synthesized), both, or just text (synthesized) and this may vary from turn to turn. |
| VXML 3 | 7.3 | 7.3.8.4 | Te | Testing compliance for interoperability is difficult. Because of the rapid changes going on in telecommunications we can't pick one standard over another and mandate that. Nor can we select a small group of standards and hope that those will be sufficient. As time passes those standards will be replaced with newer ones. The suggested listing and the Note referring to the Bagwell reference is a move in the right direction because it is a dynamic, open source listing of the standards and because it is a conversion utility. If the coding is publicly available rather than it should be included. Bagwell provides that. VIEW 1: Since we are dealing with raw audio and not processed audio it may be that a generic specification of the type of processing is all that is needed. There are only three ways you can go from an analog signal to a discrete signal: pulse code, wave, and amplitude. | Proposed alternatives 1. follow the VoiceXML approach by mandating one or two formats for conformance testing and indicate a list of optional formats, perhaps using Bagwell, that may be used as well. 2. Have a required audio-format field that lists only the following low-level processing methods - family of formats: Pulse-code modulation Amplitude modulation Wave-code modulation Then you can say it is fixed point or floating point or you can say it is mu-law or Alaw. Then, to optionally specify the codec you can use the Bagwell resource. | Reject But, move list to informative or just refer to Bagwell (also informative) as a source |

¹ MB = Member of M1 - affiliation

² Type of comment: **ge** = general **te** = technical **ed** = editorial – For technical comments, please indicate whether your comment is a MAJOR or MINOR technical comment. **NOTE** Columns 1, 2, 4, 5 and 6 are compulsory.

| 1 MB ¹ | 2 Clause No./ Subclause No./ Annex (e.g. 3.1) | (3) Paragraph/ Figure/Table/ Note (e.g. Table 1) | 4 Type of com- ment ² | 5 Comment (justification for change) by the MB | (6) Proposed change by the MB | (7) Proposed Editors Disposition |
|----------------------|--|--|--|--|---|-------------------------------------|
| | | | | <p>Linear encoding is a kind of PCM and linear must be described in terms of the number of bits (8, 16, 24, 32 bit etc,) which we already do with the precision attribute.</p> <p>These are all pure, fundamental standards at the lowest level.</p> <p>You can also add mu-law and Alaw because they are saving bits and are well-understood standards. All we need is to abide by those and have codec history.</p> <p>No matter what codec you use raw data ends up being translated into a standard logarithmic coding (e.g., Alaw, mu-law) or it is a pure number of bits (8, 16,, etc.), or it is a float.</p> <p>These are the only options you have. All of the other options are really a matter of coding and compression - which goes beyond the raw data level and which ends up being decoded into one of the above values. These things are already represented in the format through sampling rate, precision, and type of sample and possibly compression. Those are the important elements of the format. Once you have those things specified it doesn't really matter what the codec is. For example, if you do basic pulse code modulation and allow for standard mappings, like Alaw and mu-law, then you focus on the bare essential of pulse-code modulation that would have either logarithmic or non-logarithmic coding. Note that, today, most people do pulse code modulation. PCM can be done linearly, in mu-law or in Alaw format. That is within a sample. They are all still pulse code modulation.</p> <p>This approach (use of the low level formats) makes the data exchange format truly independent of codecs.</p> <p>The only benefit of specifying the codec - whether for the original coding or the recording - is to determine the kinds of losses that have occurred due to the codec conversion.</p> | <p>Then you can add the historical information about the signal: where it came from, compression, and other channel information.</p> <p>3. Use the existing field to indicate specific codec. Consider using MIME</p> | |

1 MB = Member of M1 - affiliation

2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial – For technical comments, please indicate whether your comment is a MAJOR or MINOR technical comment. **NOTE** Columns 1, 2, 4, 5 and 6 are compulsory.

| 1 | 2 | (3) | 4 | 5 | (6) | (7) |
|-----------------|---|--|------------------------------|--|--|-----------------------------------|
| MB ¹ | Clause No./ Subclause No./ Annex (e.g. 3.1) | Paragraph/ Figure/Table/ Note (e.g. Table 1) | Type of comment ² | Comment (justification for change) by the MB | Proposed change by the MB | Proposed Editors Disposition |
| | | | | VIEW 2: VoiceXML has specified only two mandatory formats for compliance testing and has a number of optional formats that the platform is responsible for supporting. VIEW3: We need to consider using MIME | | |
| VXML 4 | 7.4 | | Te | Storage of binary data could explode the size of the format | Use Winzip compression of the entire format | Noted. Need further clarification |
| VXML 5 | 8 | | Te | Formatting | Use CDATA or a comparable tag for binary data | Noted. Need further consideration |
| VXML 6 | 8 | | Te | Schema is in XML 1.0. In the future, this standard will be used by companies that are using VoiceXML 3.X. VoiceXML 3.x does not yet exist so a schema cannot yet be writing in it but a schema should be made available when it is released. | Determine how to add another schema in the future. | Noted. A revision is recommended |
| VXML 7 | 8 | | Te | This is the first pass for the schema. | Review schema | Accept |

¹ MB = Member of M1 - affiliation

² **Type of comment:** ge = general te = technical ed = editorial – For technical comments, please indicate whether your comment is a MAJOR or MINOR technical comment. **NOTE** Columns 1, 2, 4, 5 and 6 are compulsory.
page 3 of 3