|  |  |
| --- | --- |
|  | Moving Picture, Audio and Data Coding by Artificial Intelligence  www.mpai.community |

|  |  |
| --- | --- |
| **N188** | 2021/03/17 |
| **Source** | CAE-DC |
| **Title** | Some issues related to MPAI-CAE Functional Requirements |
| **Target** | MPAI Members |

MPAI-5 has approved the MPAI-CAE Use Cases and Functional Requirements (N151) as attachment to the Calls for Technologies N171. However, the source CAE-DC has identified some issues that are worth a clarificaton. This is posted on the MPAI web site and will be com­mun­icated directly to those who have informed the Secretariata of their intention to respond.

**General issue**

MPAI understands that the scope of N151 is very broad. Therefore it reiterates the point made in N152 that:

*Completeness of a proposal for a Use Case is a merit because reviewers can assess that the components are integrated. However, submissions will be judged for the merit of what is proposed. A submission on a single technology that is excellent may be considered instead of a submission that is complete but has a less performing technology.*

**Emotion-Enhanced Speech** (Use case #1 in N151)

The Functional Requirements of the Use Case does not explicitly indicate the form in which speech without emotion and Emotion enter the EES system. Possible modalities are

1. A speech file and a separate Emotion file where the sequence of Emotions carries time stamps
2. An interleaved stream of speech and Emotions

MPAI welcomes proposals addressing these issues.

The assessment of submissions by Respondents who elect not to not answer this point will not influence the assessment of the rest of their submission

**Audio Recording Preservation** (Use case #2 in N151)

MPAI welcomes proposed semantics of the information conveyed in the Text output of the Musicological classifier. It is not clear what information the Musicological classifier is providing.

The assessment of submissions by Respondents who elect not to not answer this point will not influence the assessment of the rest of their submission

**Enhanced Audioconference Experience** (Use case #3 in N151)

The function of the Speech detection and separation AIM is described as “Separation of relevant speech vs non-speech signals”.

As the description is possibly misleading, we inform submitters that the correct description of the AIM should be “Separation of relevant speech vs other signals (including unwanted speech)”.

The assessment of submissions by Respondents who base their submission on the text in N151 will not be affected by this clarification.

**References**

1. MPAI-CAE Use Cases & Functional Requirements; MPAI N151; <https://mpai.community/standards/mpai-cae/#UCFR>
2. MPAI-CAE Call for Technologies, MPAI N152; <https://mpai.community/standards/mpai-cae/#Technologies>
3. MPAI-CAE Framework Licence, MPAI N171; <https://mpai.community/standards/mpai-cae/#Licence>