





# MPAI-MMC: Multimodal Conversations for XR

Mark Seligman

2025/01/18

# MPAI Standards and MPAI-MMC

- MPAI develops standards for:
  - **Components** (AI Modules – AIM): their function, I/O data
  - Organised in **workflows** (AI Workflows – AIW): their function, I/O data, AIM topology
  - Executed in an MPAI specified **environment** (AI Framework – AIF): metadata, API
- **Multi-modal Conversation (MMC) set of standards**

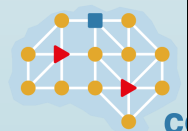
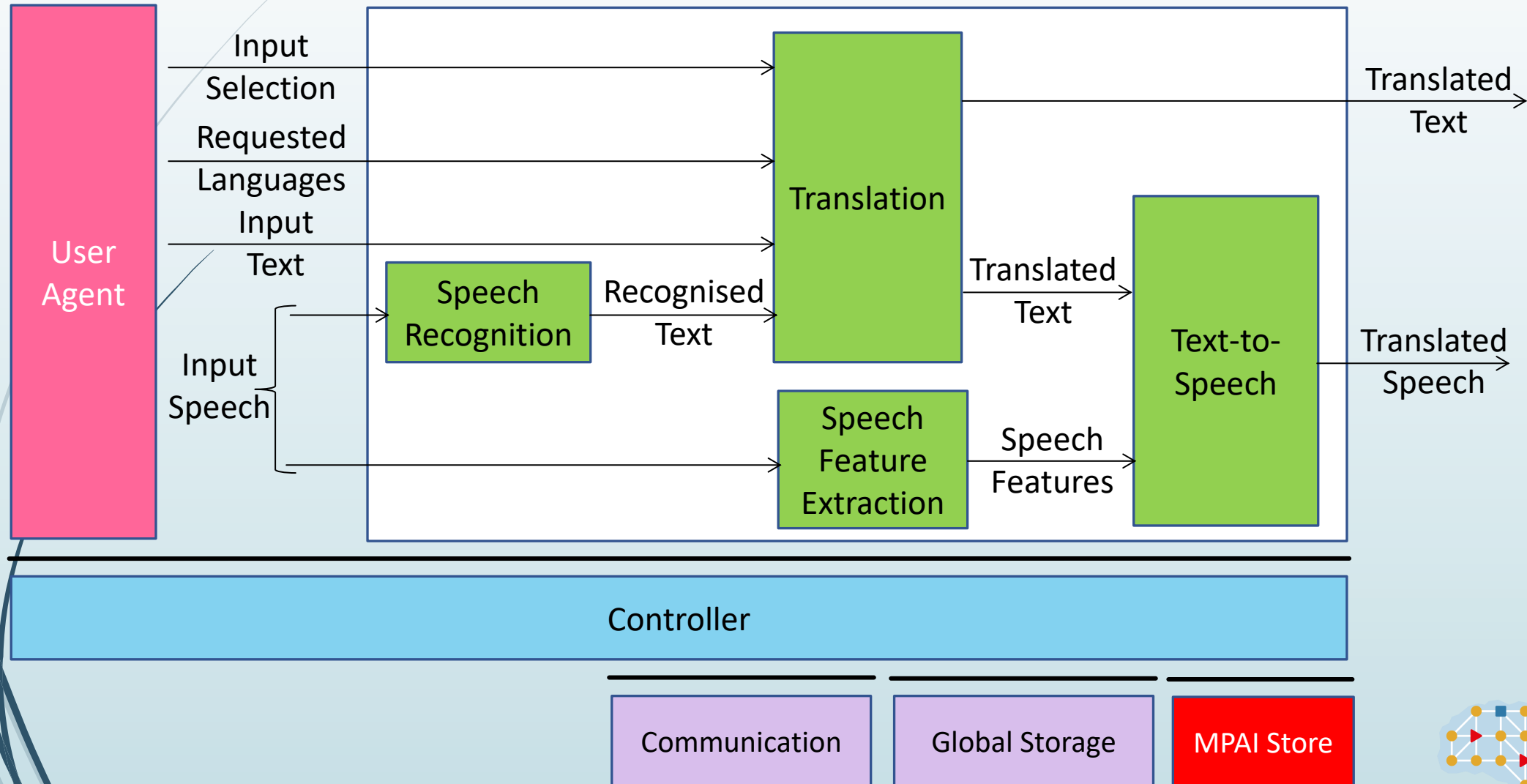
Standard Name	Acronym	V
<i>Multimodal Conversation</i>	MPAI-MMC	1.2
- Automatic Speech Translation		
- Conversation with Emotion		
- Multimodal Question Answering		

# Speech-to-Speech Translation

- **Multilingual Conferencing!**
- Spoken translation in the **original source-language voice**
- Many untouched **vertical markets**
- Multilingual **social networking**



# Multilingual Conferencing



# S2S translation in original source-language voice and for vertical markets

## SPEECHMORPHING: Cross-Lingual Dialogue

- **ASR**  
(English voice-to-text)
- **Translation**  
(English text to Spanish text)
- **Speechmorphing TTS**  
(Spanish text-to-voice)



Agent asks question in English

Agent hears the response in English

- **ASR**  
(Spanish voice-to-text)
- **Translation**  
(Spanish text to English text)
- **Speechmorphing TTS**  
(English text-to-voice)

Customer hears agent's question in Spanish – in the agent's voice



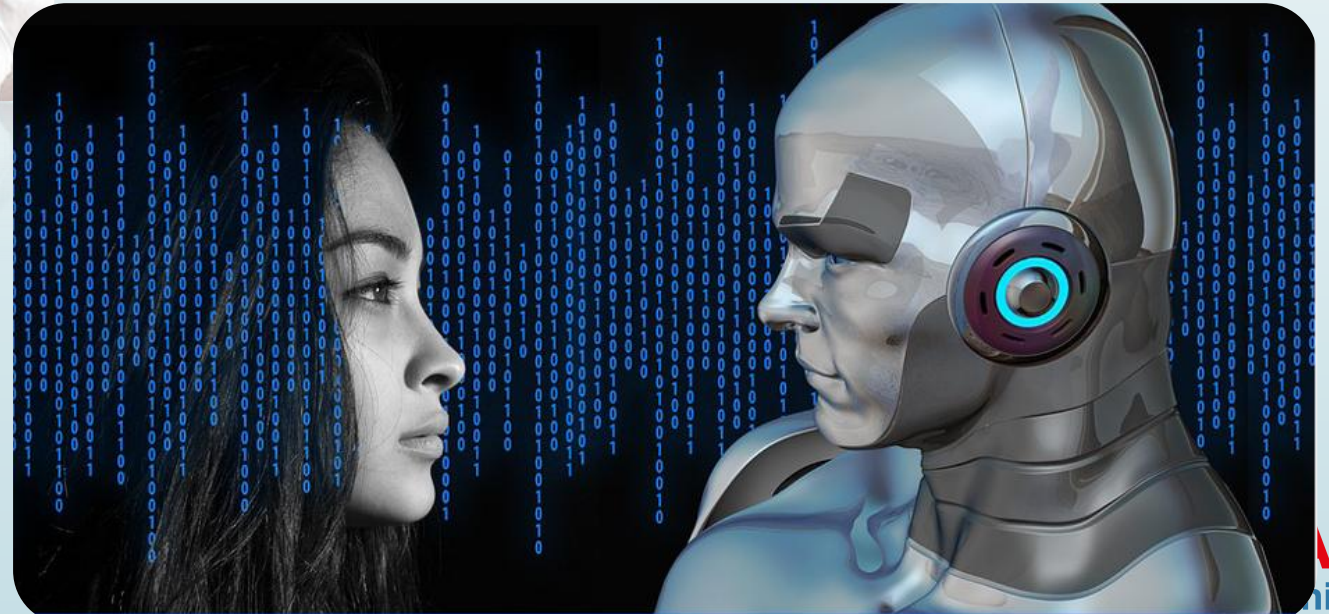
Customer responds in Spanish

# Conversation with Emotion



Conversation with Human

Conversation with Machine



# Conversation/2



Conversation with AI Speaker



Conversation with Robot

# Conversation between Human and Machine

Online meeting



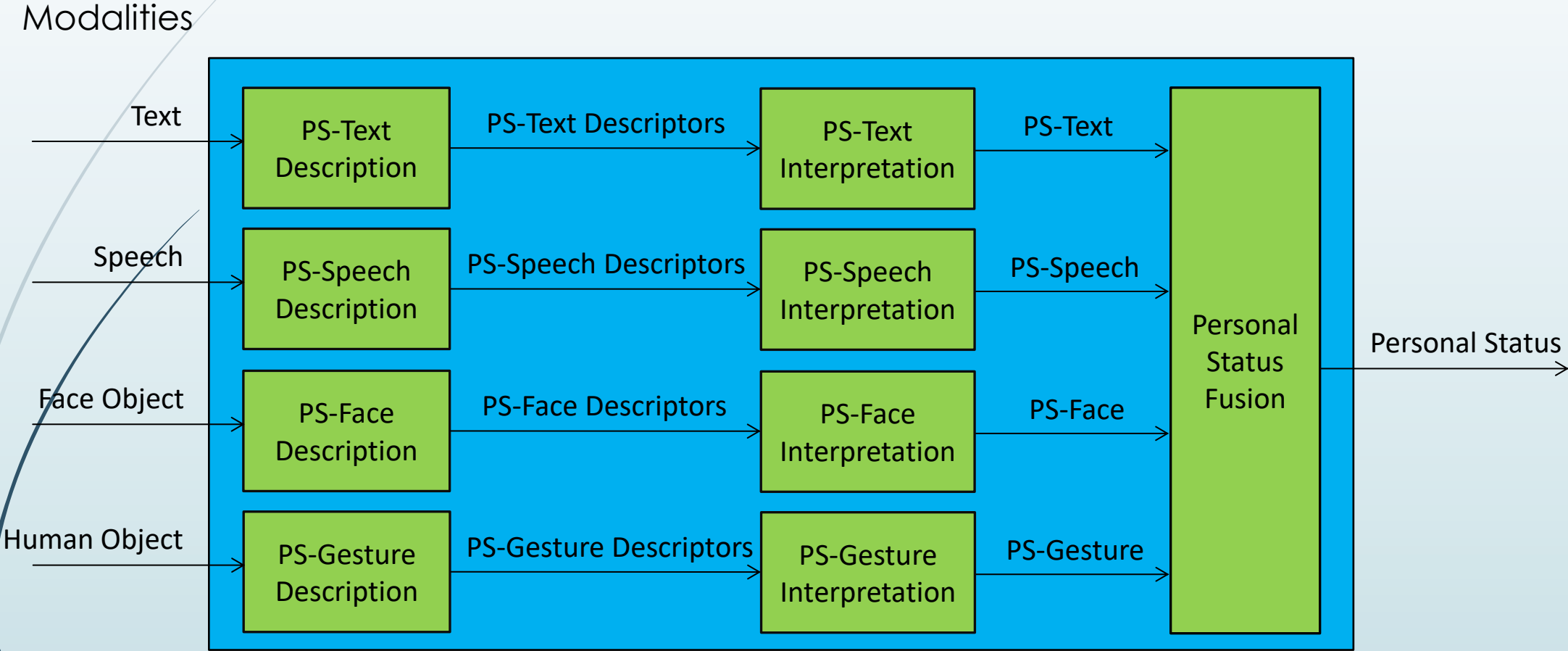
Online meeting with Avatar



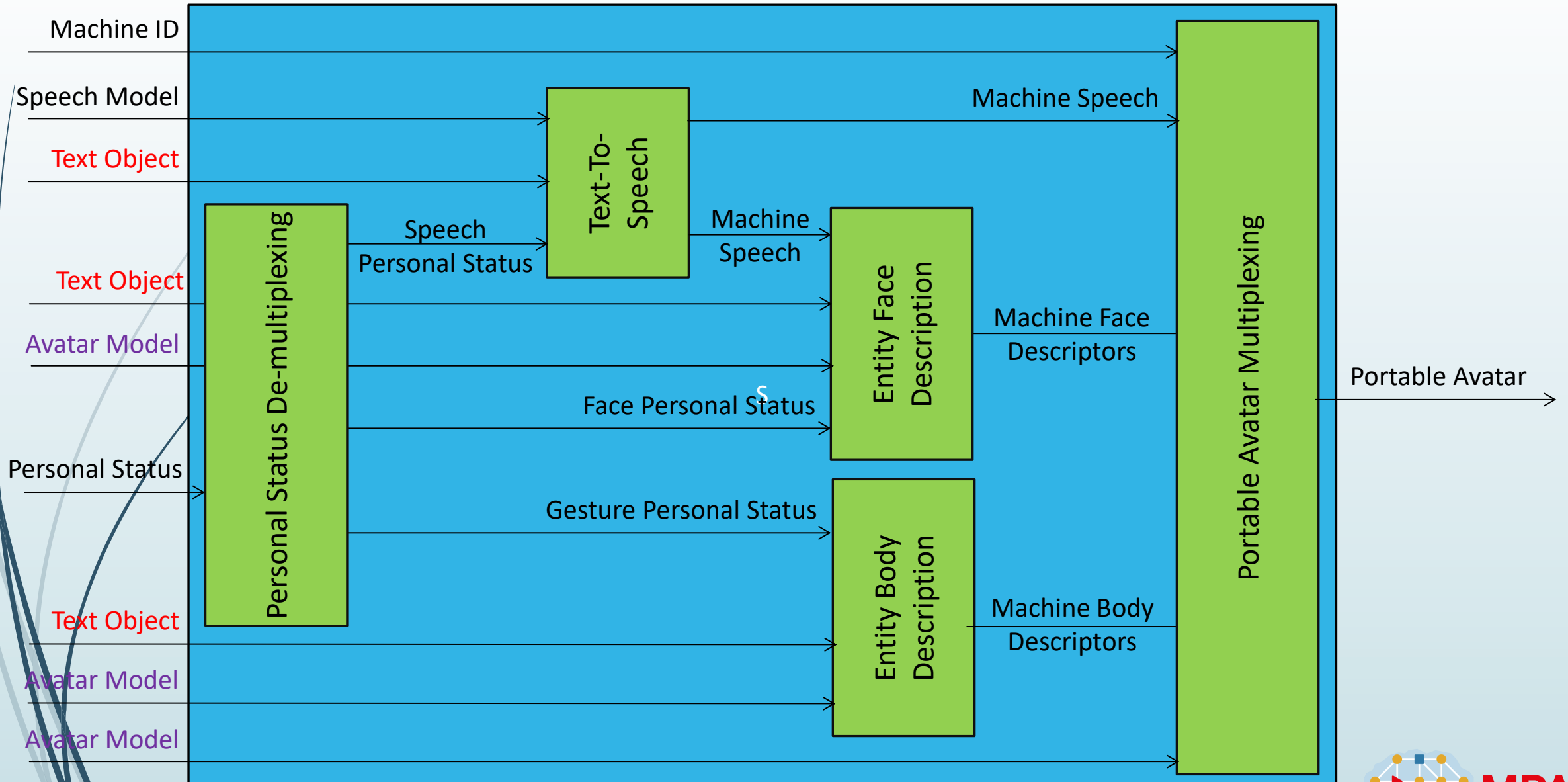
# Personal Status

- MPAI defined Personal Status (PS) as a set of **internal human characteristics**:
  - 1. Emotion and Cognitive State** result from human interaction with the Environment.
    - a. Cognitive State* is more rational (e.g., “Confused”, “Dubious”, “Convinced”).
    - b. Emotion* is less rational (e.g., “Angry”, “Sad”, “Determined”).
    - a. Attitude* is the stance a human takes when s/he has reached an Emotion and Cognitive State (e.g., “Confrontational”, “Respectful”, “Soothing”).
- PS is shown by a modality: *Text, Speech, Face and Gesture* and possibly by more than one modality simultaneously.
  - More modalities possible but currently not considered, e.g., the body as in body language, dance, etc.
- Personal Status Extraction: a composite AIM for extracting the PS of a human.

# Personal Status Extraction



# Personal Status Display



# Emotion in MPAI-MMC V2 – Basic Emotion Set

EMOTION CATEGORIES	GENERAL ADJECTIVAL	SPECIFIC ADJECTIVAL
<b>ANGER</b>	anger	furious irritated frustrated
<b>APPROVAL, DISAPPROVAL</b>	admiring/approving disapproving indifferent	awed contemptuous
<b>AROUSAL</b>	aroused/excited/energetic	cheerful playful lethargic sleepy
<b>ATTENTION</b>	attentive	expectant/anticipating thoughtful distracted/absent-minded vigilant hopeful/optimistic
<b>BELIEF</b>	credulous	sceptical
<b>CALMNESS</b>	calm	peaceful/serene resigned
<b>DISGUST</b>	disgust	loathing
<b>FEAR</b>	fearful/scared	terrified anxious/uneasy
<b>HAPPINESS</b>	happy	joyful content delighted amused

# Emotion in MPAI-MMC V2 – Semantics of the Basic Emotion Set

ID	Emotion	Meaning
1	Admiring/Approving	emotion due to perception that others' actions or results are valuable
2	Amused	positive emotion combined with interest (cognitive)
3	Angry	emotion due to perception of physical or emotional damage or threat
4	Anxious/Uneasy	low or medium degree of fear, often continuing rather than instant
5	Aroused/Excited/ Energetic	cognitive state of alertness and energy
6	Arrogant	emotion communicating social dominance
7	Astounded	high degree of surprised
8	Attentive	cognitive state of paying attention
9	Awed	approval combined with incomprehension or fear
10	Bewildered/Puzzled	high degree of incomprehension
11	Bored	not interested
12	Calm	relative lack of emotion

Join the fun, build the future!

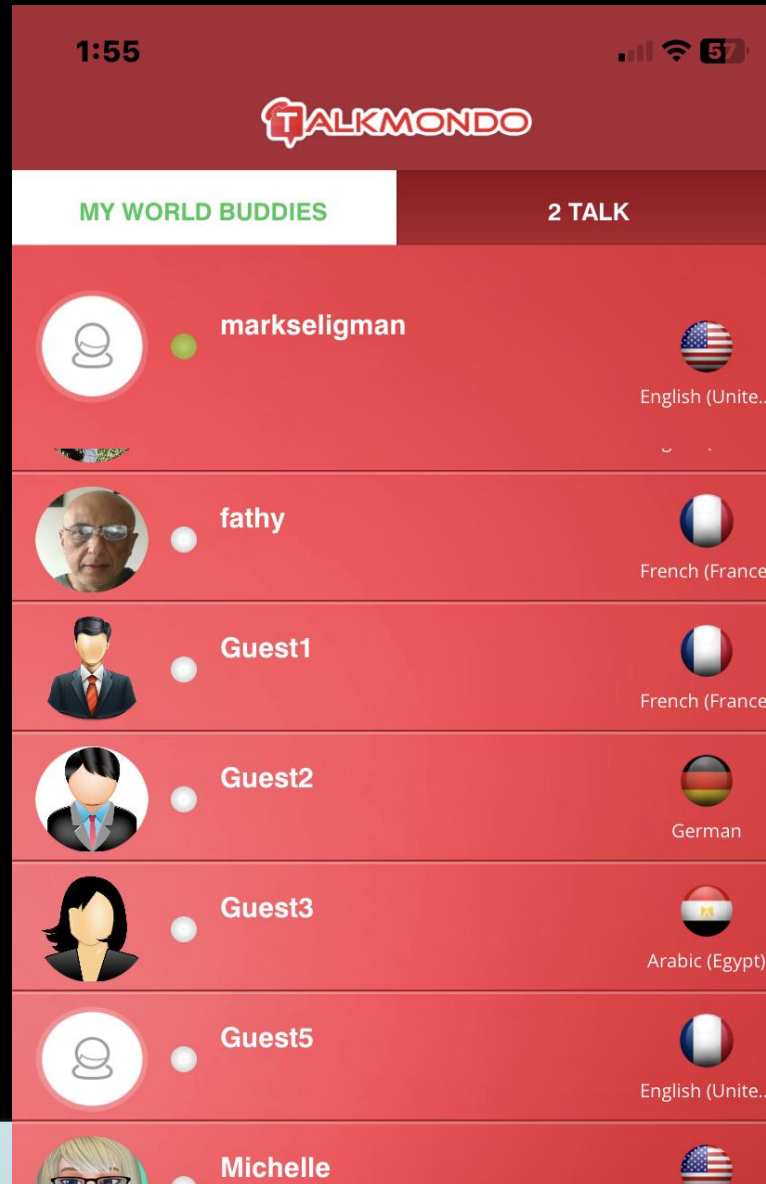


<https://www.mpai.community/>

More about this call at <http://mmc.mpai.community/>

# Multilingual Social Networking

Speech-to-speech cross-language chat with known buddies and new friends.



Each buddy has preferred language, current location, range of interests.