

## Description form - PROVIDER

### Language Resources: Multi-modal

#### M.1. GENERAL INFORMATION

**Data included:**

Audio (see section M.6.)       Image (see section M.7.)       Video (see section M.8.)

**Language(s):**

Language dependent       Language independent

Language(s):

#### M.2. RECORDING INFORMATION – HUMANS

**Sex and number of humans:**

Male      Number:

Female      Number:

Imposters      Number:

Synthetic      Number:

Total number:

Number of humans visible in the same frame:

**Age class:** (please indicate number of speakers and age)

Children I ( \_\_ to \_\_ )       Adults I ( \_\_ to \_\_ )       Elderly (over 60)

Children II ( \_\_ to \_\_ )       Adults II ( \_\_ to \_\_ )       Age unknown

Teenagers I ( \_\_ to \_\_ )       Adults III ( \_\_ to \_\_ )       Other distribution:

Teenagers II ( \_\_ to \_\_ )       Adults IV ( \_\_ to \_\_ )

**Origin:**

Native       Non native       Unknown

Comments:

**Geographic distribution:**

Total number of regions:

Percentage per region:

Regions included:

**Information included about:**

Place of living       Place of birth       Dialect/accent

Comments:

**Additional speaker information included:**

Speaking/hearing impairments       Height       Weight

Smoking habits       Trained speakers       Education level

Profession       Place of (secondary) education       Other:

Comments:

**M.3. RECORDING INFORMATION – RESOURCE**

**Human body parts visible in the resource:**

<input type="checkbox"/> None	<input type="checkbox"/> Whole body	<input type="checkbox"/> Head
<input type="checkbox"/> Face	<input type="checkbox"/> Mouth	<input type="checkbox"/> Arms
<input type="checkbox"/> Hands	<input type="checkbox"/> Legs	<input type="checkbox"/> Feet
<input type="checkbox"/> Other:		

**Distractors visible in the resource:**

<input type="checkbox"/> None	<input type="checkbox"/> Hat	<input type="checkbox"/> Glasses
<input type="checkbox"/> Watch	<input type="checkbox"/> Scarf	<input type="checkbox"/> Pen/Paper/Notepad
<input type="checkbox"/> Microphone	<input type="checkbox"/> Markers	<input type="checkbox"/> Mobile phone
<input type="checkbox"/> Other:		

**Interactive media visible/audible in the resource:**

<input type="checkbox"/> None	<input type="checkbox"/> Graphical screen	<input type="checkbox"/> Computer pen
<input type="checkbox"/> Tactile screen	<input type="checkbox"/> Data glove	<input type="checkbox"/> PDA
<input type="checkbox"/> Desktop PC	<input type="checkbox"/> Laptop	<input type="checkbox"/> Mouse
<input type="checkbox"/> Other:		

**Annotated modalities in the resource:** *(for details see section M.4.)*

<input type="checkbox"/> None	<input type="checkbox"/> Speech	<input type="checkbox"/> Hand/Arm gestures
<input type="checkbox"/> Gaze/Eye movements	<input type="checkbox"/> Facial expressions	<input type="checkbox"/> Lip movements
<input type="checkbox"/> Head movements	<input type="checkbox"/> Body movements	<input type="checkbox"/> Hand manipulation of objects
<input type="checkbox"/> Other:		

Total number of annotated modalities:

**Other modalities available/visible but not annotated in the resource:** *(for details see section M.4.)*

<input type="checkbox"/> None	<input type="checkbox"/> Speech	<input type="checkbox"/> Hand/Arm gestures
<input type="checkbox"/> Gaze/Eye movements	<input type="checkbox"/> Facial expressions	<input type="checkbox"/> Lip movements
<input type="checkbox"/> Head movements	<input type="checkbox"/> Body movements	<input type="checkbox"/> Hand manipulation of objects
<input type="checkbox"/> Other:		

Total number of modalities (not annotated):

**Scene – Illumination:**

<input type="checkbox"/> Daylight	<input type="checkbox"/> Single source	<input type="checkbox"/> Multiple sources
<input type="checkbox"/> Fix	<input type="checkbox"/> Variable	<input type="checkbox"/> Other:

**Scene – Backgrounds:**

<input type="checkbox"/> Plain	<input type="checkbox"/> Complex	<input type="checkbox"/> Other:
--------------------------------	----------------------------------	---------------------------------

**Data:**

Total number of sessions:

Number of poses per subject:

**M.4. MODALITIES – DETAILED INFORMATION****Facial expressions:** (6 universal expressions of emotion)

- |                                    |                                  |                                   |
|------------------------------------|----------------------------------|-----------------------------------|
| <input type="checkbox"/> Anger     | <input type="checkbox"/> Disgust | <input type="checkbox"/> Fear     |
| <input type="checkbox"/> Happiness | <input type="checkbox"/> Sadness | <input type="checkbox"/> Surprise |
| <input type="checkbox"/> Other:    |                                  |                                   |

**Head movements:**

- |                                   |   |   |
|-----------------------------------|---|---|
| <input type="checkbox"/> Rotation | <input type="checkbox"/> Inclination forward/backward | <input type="checkbox"/> Inclination sideward |
| <input type="checkbox"/> Other:   |   |   |

**Face views:**

- |                                  |                                  |                                 |
|----------------------------------|----------------------------------|---------------------------------|
| <input type="checkbox"/> Frontal | <input type="checkbox"/> Profile | <input type="checkbox"/> Other: |
|----------------------------------|----------------------------------|---------------------------------|

Total number of face views per subject:

**Gestures:**

- |   |  |                                    |
|---|--|------------------------------------|
| <input type="checkbox"/> Interactional  | <input type="checkbox"/> Non-interactional | <input type="checkbox"/> Emotional |
| <input type="checkbox"/> Non identified |  |                                    |
| <input type="checkbox"/> Other:         |  |                                    |

**Gaze/Eye movements:**

- |                                     |   |                                      |
|-------------------------------------|---|--------------------------------------|
| <input type="checkbox"/> Saccades   | <input type="checkbox"/> Pursuit motion | <input type="checkbox"/> Convergence |
| <input type="checkbox"/> Horizontal | <input type="checkbox"/> Vertical       | <input type="checkbox"/> Other:      |

**Hand manipulation of objects:**

- |  |   |  |
|--|---|--|
| <input type="checkbox"/> Direct manipulation       | <input type="checkbox"/> Indirect manipulation    | <input type="checkbox"/> Modifying objects |
| <input type="checkbox"/> Joining/splitting objects | <input type="checkbox"/> Changing object position | <input type="checkbox"/> Other:            |

**Body movements:**

- |                                     |                                     |                                     |
|-------------------------------------|-------------------------------------|-------------------------------------|
| <input type="checkbox"/> Upper body | <input type="checkbox"/> Lower body | <input type="checkbox"/> Whole body |
| <input type="checkbox"/> Arms       | <input type="checkbox"/> Legs       | <input type="checkbox"/> Other:     |

**M.5. APPLICATION INFORMATION****Authentication:**

- |  |  |  |
|--|--|--|
| <input type="checkbox"/> Face verification | <input type="checkbox"/> Speech verification | <input type="checkbox"/> User authentication |
| <input type="checkbox"/> Other:            |  |  |

**Recognition:**

- |   |   |   |
|---|---|---|
| <input type="checkbox"/> Face recognition       | <input type="checkbox"/> Automatic speech recognition | <input type="checkbox"/> Automatic person recognition |
| <input type="checkbox"/> Expression recognition | <input type="checkbox"/> Other:                       |   |

**Analysis:**

- |                                       |  |                                 |
|---------------------------------------|--|---------------------------------|
| <input type="checkbox"/> Lip tracking | <input type="checkbox"/> Speech/lips correlation | <input type="checkbox"/> Other: |
|---------------------------------------|--|---------------------------------|

**Synthesis:**

- |   |                                  |  |
|---|----------------------------------|--|
| <input type="checkbox"/> Talking heads          | <input type="checkbox"/> Avatars | <input type="checkbox"/> Humanoid agents |
| <input type="checkbox"/> Multimedia development | <input type="checkbox"/> Other:  |  |

**Control:**

- |  |  |                                 |
|--|--|---------------------------------|
| <input type="checkbox"/> Voice control | <input type="checkbox"/> Speech assisted video | <input type="checkbox"/> Other: |
|--|--|---------------------------------|

**Miscellaneous:**

- |  |                                 |
|--|---------------------------------|
| <input type="checkbox"/> Information retrieval | <input type="checkbox"/> Other: |
|--|---------------------------------|

**Application areas:**

- |   |                                   |  |
|---|-----------------------------------|--|
| <input type="checkbox"/> Education/Training | <input type="checkbox"/> Research | <input type="checkbox"/> Entertainment |
| <input type="checkbox"/> Banking            | <input type="checkbox"/> Tourism  | <input type="checkbox"/> Other:        |

**M.6. TECHNICAL INFORMATION – AUDIO**

**Signal encoding:**

- A-law                                        $\mu$ -law                                       Linear PCM  
 Other:

**File format:**

- AIFF                                       Wav                                       Without header  
 SAM                                       NIST/Sphere                                       Au  
 Other:

**Sampling rate:**

- 8 kHz                                       16 kHz                                       32 kHz  
 44,1 kHz                                       48 kHz                                       Other:

**Quantisation:**

- 8 bit                                       16 bit                                       32 bit  
 Other:

**Byte order:**

- Lo-hi (Intel)                                       Hi-lo (Motorola)

**Data format:**

- Signed integer                                       Unsigned integer                                       Other:

**Amount of data:**

Size (Mb, Gb, etc) or duration (minutes, hours, etc):

**Compression:**

- None                                       Shorten                                       Other:

**Number of recording channels:**

- 1 (mono)                                       2 (stereo)                                       3  
 4                                       8                                       Other:

**Annotation standard:**

- SAM                                       SGML                                       XML  
 NIST/LDC                                       Other:

**Sound quality measures included:**

- SNR                                       Cross talk                                       Clipping rate  
 Background noise                                       Other:

**Tools used for measuring sound quality:**

**Speech content:**

*No. of items:*

*No. of items:*

- |  |       |  |       |
|--|-------|--|-------|
| <input type="checkbox"/> Application words               | _____ | <input type="checkbox"/> Digit-set           | _____ |
| <input type="checkbox"/> Concatenated words              | _____ | <input type="checkbox"/> Isolated digits     | _____ |
| <input type="checkbox"/> Isolated words                  | _____ | <input type="checkbox"/> Natural numbers     | _____ |
| <input type="checkbox"/> Syllables                       | _____ | <input type="checkbox"/> Money amounts       | _____ |
| <input type="checkbox"/> VCV sequences                   | _____ | <input type="checkbox"/> Credit card numbers | _____ |
| <input type="checkbox"/> Continuous sentences            | _____ | <input type="checkbox"/> Telephone numbers   | _____ |
| <input type="checkbox"/> Phonetically rich sentences     | _____ | <input type="checkbox"/> Yes/no questions    | _____ |
| <input type="checkbox"/> Phonetically balanced sentences | _____ | <input type="checkbox"/> Other:              | _____ |

**M.7. TECHNICAL INFORMATION – IMAGE****Resolution in pixels:****Color components:**

RGB  CMYK  4:2:2  
 Other:

**Color depth:**

8 bits  16 bits  24 bits  
 Other:

**File format:**

JPG  GIF  TIFF  
 BMP  EPS  CIF  
 PPM  Other:

**Amount of data:**

Size (Mb, Gb, etc):

Duration (minutes, hours, etc):

**Compression:**

None  Zip  Other:

Compression ratio:

**M.8. TECHNICAL INFORMATION – VIDEO** **Synchronized audio** (*see section M.6.*)**Resolution in pixels:****Color components:**

RGB  CMYK  4:2:2  
 Other:

**Color depth:**

8 bits  16 bits  24 bits  
 Other:

**Frame rate:**

<25 frames/sec.  25 frames/sec.  30 frames/sec.  
 50 frames/sec.  60 frames/sec.  Other:

**File format:**

MOV  AVI  MPEG  
 Quicktime  SGI  Other:

**Amount of data:**

Size (Mb, Gb, etc):

Duration (minutes, hours, etc):

**Compression:**

None  Zip  Other:

Compression ratio:

**M.9. TECHNICAL INFORMATION – MODELING****Models:** 2D 3D Other:**File formats:** VRML Other:**Algorithms used:****M.10. FURTHER COMMENTS**

ELRA  
55-57 rue Brillat Savarin  
F-75013 PARIS, FRANCE  
Tel +33 1 43 13 33 33 - Fax +33 1 43 13 33 30  
<http://www.elra.info>