.kpn File – General Description

The **.kpn** file defines a structured container that groups images and audio recordings with session metadata. These files are designed to be processed by the **Engine**.

Each **.kpn** file includes:

- A **folder "Images"** containing the **images**

- A **folder "audio"** containing **optional audio files**

- A **JSON file "metadata.json"** containing essential **metadata** required by the Engine

---

Structure - metadata.json

The JSON file contains the following sections:

1. Metadata

The `metadata` object includes:

- **descriptor** – Identifies the file type

- **schemaVersion** – Specifies the schema version of the format

- **compatibility** – Defines compatibility information for version control

2. Session Information

- **sessionId** – The unique identifier of the session to be analyzed

- **step** – Indicates the step number within the session workflow

3. KPN File Information

The `kpnfile` object contains:

- **order** – The sequence number of this .kpn file within the step

- **status** – The current status of the file transmission

This allows proper tracking of multiple .kpn files associated with the same step.

---

4. Source Information

The `source` object contains information about how and where the .kpn file was generated.

---

5. Time Information

The `time` object includes:

- **start** – Capture start timestamp

- **end** – Capture end timestamp

---

6. Images

The `images` object contains image-related metadata:

- **width** – Image width

- **height** – Image height

- **format** – Image format (e.g., jpg, png)

- **orientation** – Image orientation metadata

Additionally, it includes a `list` array, where each element contains:

- **order** – Image sequence number

- **filename** – Image file name

- **timestamp** – Timestamp of image capture

- **eyetracking** – Eye-tracking data:

    - **x** – Horizontal gaze coordinate

    - **y** – Vertical gaze coordinate

---

7. Speech

The `speech` object contains speech capture and transcription data:

- **language** – Detected language

- **accuracy** – Confidence score of speech recognition

- **text** – Full transcribed text

It also includes a `list` array containing:

- **timestamp** – Timestamp when the word was detected

- **word** – The detected word

Example - metadata.json

An example of the JSON file can be found in this [link](#)