# WP4 D2

# Software application based on ScriBe to automatically transcribe, partition and label speech.

Version 1.0

Date: 1. July 2026

Authors: Jan-Willem van ‘t Klooster, Teodora Spirova, Andre Bester, Elvis Vrolijk

Reviewed by: Friedrich Pawelka, Bas de Leng


Co-funded by
the European Union

# Table of Contents


*1. Introduction* ..... *3*

*2. What is ScriBe?* ..... *4*

ScriBe working mechanisms ..... 5

Analysis pipeline ..... 5

Upload Audio and GSR Files ..... 5

Working with the Analysis Table ..... 6

Codebooks ..... 6

Manuals and Tutorials ..... 7

*3. ACTIVATE use case* ..... *7*

Usage of ScriBe within ACTIVATE ..... 7

Future steps ..... 8

*4. Scribe API* ..... *9*

*References* ..... *9*

# 1. Introduction

This deliverable is the introduction to ScriBe software for transcribing multimedia files and conducting consequent qualitative data analysis. ScriBe is designed to support researchers by providing tools to manage, analyze, and visualize audio, narrative and biometric data.

Nowadays, in educational and research settings, large amounts of (multi)media data are collected and needed to be analysed. This is a challenging and timecomsuming task requiring specific expertise. Therefore, a tool called ScriBe is developed to streamline media analysis by integrating video, audio, and even sensor data, and media analysis, coding, processing and understanding. This allows the analysis of behaviour and phenomena based on video, audio and other data input.

Especially in contexts like ACTIVATE this is relevant, as complex educational settings are taking place, being observed and then analysed. These settings, including dynamic case-based learning situations, can be observed from multiple angles and captured in various ways (eg. audio, video). This is for example relevant when studying or evaluating new didactial concepts. Yet, the captures need to be managed, analysed and synthesised to become of effective value. ScriBe provides an unique way to cover these needs in a SaaS[1]—based approach.

This document covers:

- Introduction to what ScriBe is and how it can be used in research, including
    - Setting up analysis, uploading and analysing data, creating codebooks with categories and codes, creating sessions and inviting collaborators
    - Uploading multimedia and GSR (Galvanic Skin Response) files
- Creating and making use of analysis tables
- Coding data and exporting your results
- An ACTIVATE-related typical Usage scenario
- API connectivity

[1] Software as a Service

# 2. What is ScriBe?

ScriBe is a multimedia SaaS based software platform developed by BMS Lab, University of Twente, that allows learners, teachers and researchers to:

- Upload and transcribe audio/multimedia files (.mp3, .mp4) to text;
- Identify speakers using diarisation ;
- Combine audio with physiological data (e.g., GSR from Empatica or EmbracePlus devices);
- Code and categorize qualitative data;
- Visualize and export findings for usage in reports, scientific publications, theses, etc.

Figure 1 shows the main working screen of Scribe, in which multimedia controls, transcript, transcript and diarisation (speaker identification) visualization, and coding options are present.

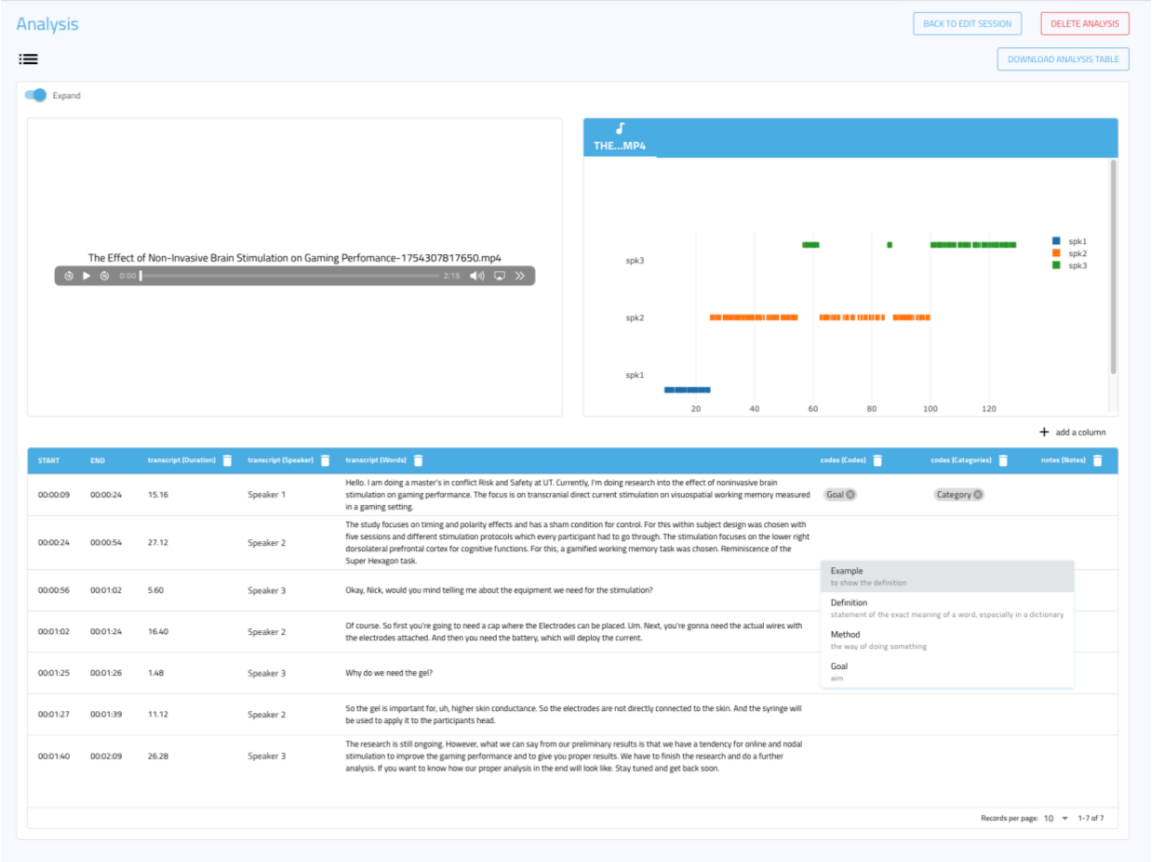


Figure 1: ScriBe software workflow showing multimedia controls (left top), diarisation visualization (right top), transcription, diarisation and coding (bottom) of a 2-speaker audio file.

## ScriBe working mechanisms

ScriBe works with *Sessions* – the main project workspace containing files, collaborators, and analysis tools; *Codebooks* – a structured collection of categories and codes used during data analysis and *Analysis* – The environment where one applies codebooks to data files, invites coders, and interprets findings.
ScriBe can be used for deductive coding, inductive coding, consensus building and testing of coding schemes, summary making, labelling of narrative data, teaching, and more. In education, it could help to teach learners structured coding and qualitative research methods, and usage of qualitative research methods tooling.

## Analysis pipeline

Setting up an analysis pipeline in ScriBe includes:

1. **Creation** of a **Codebook** with Codes and Categories,
2. Inviting and collaborating on the Codebook with other users, (optional)
3. **Inviting** and collaborating with others within a Session on the same data to be analysed, (optional)
4. **uploading files** with data to be analysed to the SaaS platform,
5. configuring **Speech to Text** and speaker **diarisation,**
6. Working with the **Analysis Table** to check, modify, and save the transcription
7. Add **codebooks**, files, and potentially other codes
8. Code the actual data
9. Export the coded result.

Details of these steps are described in the ScriBe user manual User Documentation [1].

## Upload Audio and GSR Files

Once a session has been created, users are required to upload relevant multimedia files, such as audio and video in MP3 and MP4 formats. If applicable, data files from the Embrace Plus wearable device—capturing measures such as temperature, acceleration, heart rate, electrodermal activity (GSR), blood volume pulse, inter-beat interval, and actigraphy counts—can also be uploaded for further analysis.

Following the upload, users must configure the processing settings. This includes enabling transcription and specifying the language and the number of speakers, which are required for accurate transcription. The time required to generate transcripts depends on the size of the uploaded files. Once processing is complete, users are able to review, edit, and analyze the transcript, including viewing speaker turns.

Processing of wearable data provides data, including recording duration, number of samples, and sampling rate.

## Working with the Analysis Table

To initiate the analysis, an analysis table must be created by adding a codebook, relevant files, and assign coders.

Once the codebook has been established, the structure of the analysis table should be defined by adding the needed columns. These typically include the start and end times of each speech segment, the transcript, the speaker identifier, and the codebook. Where relevant, wearable data and any additional variables useful for the analysis may also be incorporated.

The transcript can be edited in place and segmented into additional rows where necessary. Furthermore, associated multimedia files can be viewed and played directly within the same interface, enabling integrated analysis.

From within the table interface, users can access and manage codes, including their descriptions, examples, and hierarchical relationships (e.g., sub-codes), and make modifications as needed. Codes can be organized hierarchically directly within the table.

Lastly, as part of the analysis table, users are able to visualise the transcript in a graph showing segments spoken by each speaker.

## Codebooks

Codebooks consist of structured collections of codes that can be reused across multiple sessions and analyses by all users. They are dynamically updated: any new codes created directly within the analysis table—including hierarchical relationships such as subcodes—are automatically incorporated into the codebook.

The platform supports both inductive and deductive coding approaches. Users may choose to develop a comprehensive codebook prior to analysis or to begin coding directly within the table, iteratively refining and expanding the codebook throughout the analytical process.

In a deductive coding approach, users can fully configure the codebook in advance, including defining codes, subcodes, and their descriptions. As codes are subsequently applied to segments of text, illustrative examples are automatically populated within the codebook.

## Manuals and Tutorials

The latest detailed Software Manual and Usage Tutorials are published under https://www.utwente.nl/en/bmslab/infohub/scribe/.

# 3. ACTIVATE use case

## Usage of ScriBe within ACTIVATE

### *In individual skills training*

The TIIM smartphone app [2] allows for recording and secure, categorized storage of audio. Then, using ScriBe, transcripts are generated and made available again within TIIM to researchers (researcher portal) and end users (e-portfolio).
For user experience reasons, it was chosen to avoid having the transcript be viewed in the app as the transcript will be very long (as it requires a lot of scrolling). A desktop e-portfolio on larger screen is preferred.

The TIIM app was used during individual skills training to collect feedback of peers and instructor in an unobtrusive way and present it in the trainee's individual e-portfolio. For that, a new variant of the audio recording question type was implemented, which collects also the *information about the spoken language* and the *number of speakers* – information that ScriBe needs to perform a transcription. After finishing the training, the audio recordings of this new question type is transcribed via ScriBe, and then sent back to the TIIM server. The transcript is then visible within the app (next to the playable audio file). To support better readability of large transcripts and comparison of the feedback of multiple trainings an e-portfolio was created that can be used on a desktop computer (or tablet in landscape rotation) and where the transcript from ScriBe is also visible. The visualization of the different speakers is very helpful for quickly grasping the content of the transcript.

### *For inter-disciplinary feedback*

ScriBe can also be used for feedback on recorded sessions in group sessions such as inter-professional education. These audio files are then uploaded to Scribe and then transcribed automatically. Several participants from different disciplines can add codes to transcriptions and transcript events for later classification, reflection and training protocols.

## Future steps

### *Process needs*

Usage within ACTIVATE taught us a lot about the process of offering such an application in an international collaboration context. First of all, as developers and trainers sometimes speak different languages and come from different disciplines, one has to find a shared common language. Also, demonstrations are needed and have to be understood to come to see the potential of automated speech to text on one hand, but on the other the actual use case also has to be understood to make a generic but suitable system.

On the other hand, dealing with potentially sensitive recordings under GDPR, requires proper ethics consent and usage authorization. It was found during the project that different institutions have different ethical consent procedures, and that these procedures had to be aligned.

### *Content needs*

Results of usage of ScriBe within the Activate project also led to to the identification of content related needs for a future update (outside of ACTIVATE itself, but as a direct result of it). This was labelled 'Scribe 2.0'. Must haves identified include:

- **Uploading** of existing text files to further analyse (rather than audio/video files);
- increased explanatory **guidance**. To this end, a manual and Video Tutorials have been made available on https://www.utwente.nl/en/bmslab/infohub/scribe/.
- Enhanced **Collaboration** Features. Collaboration is central to ScriBe's purpose. Improvements to multi-user functionality will include:
    - Version history and change tracking for codes and annotations.
    - Action overviews for team activities.
    - Structured workgroups and simultaneous editing within shared projects
- Allow users to define **alternative units of analysis**, such as text segments instead of timestamps.
- **Technical updates.** This includes codebase optimization, possibility to be more generic in terms of TTS, and storage scalability.
- Streamline navigation and workflows (**UX**).

At University of Twente, faculty-internal social sciences & humanities research funds have been requested in 2025Q3 and allocated in 2025Q4 to accommodate these updates. These are scheduled and implemented in 2026Q2 to 2026Q4.

## 4. Scribe API

The Scribe API provides endpoints to the Scribe web application to do the following:

- Create, edit and delete sessions
- Retrieve user specific collections of sessions
- Upload and download files associated with a session
- Delete files associated with a session
- Retrieve a specific file
- Create, edit and delete analyses associated with a session
- Retrieve a specific analysis associated with a session
- Request transcription of an audio or video file into text

A separate deliverable [3] provides technical and architectural details on the API Endpoints. The ScriBe API is used within ACTIVATE to enable the TIIM ecosystem [2] audio recording and ePortfolio functionalities, described in and Activate Deliverable WP4 D4 [4] and in [5].

## References


1. T. Spirova. BMS Lab Infohub (2026). https://www.utwente.nl/en/bmslab/infohub/scribe/ University of Twente, 2026. last accessed: 26 June 2026.
2. van 't Klooster J-WJR, Rabago Mayer LM, Klaassen B and Kelders SM (2024) Challenges and opportunities in mobile e-coaching. Front. Digit. Health 5:1304089. doi: 10.3389/fdgth.2023.1304089
3. D3 Application Programming Interface for Scribe
4. D4 Usability test of Individual Deliberate Practice (IDP) support
5. Van 't Klooster, J. W. J. R., Spirova, T., Vrolijk, E., Bester, A., Pawelka, F. & de Leng, B. User-centered Design of App-assisted Feedback Portfolio Technology for Individual Deliberate Practice in Health Professional Education (2026). In: Lecture Notes in Computer Sciences (Accepted/In press).