

# A multidisciplinary approach to the use of technology in research: the case of interview data

**Louise Corti**  
**Christoph Draxler**

Parthenos workshop for  
CEE Countries  
7-9 October 2019,  
Sofia



# Welcome

Thank you for joining us here in Sofia!

Part of Parthenos workshop for CEE Countries

Thank to the great organising team!

# Who are we?

A multidisciplinary group of European scholars

- tools and data professionals
- fields - speech technology, social sciences, human computer interaction, oral history and linguistics

Interested in strengthening the position of **interview data** in Digital Humanities

Arjan van Hessen

Stef Scagliola

Silvia Calamai

Henk van den Heuvel

**Christoph Draxler**

**Louise Corti**

Jeannine Beeken

Norah Karrouche

# Thanks to our Oral History.eu colleagues

- Stefania Scagliola, University of Luxembourg
- Silvia Calamai, University of Siena
- Norah Karrouche, Erasmus University Rotterdam
- Jeannine Beeken, University of Essex
- Arjan van Hessen, University of Twente
- Henk van den Heuvel, Radboud University
- Maureen Haaker, University of Essex
- Max Broekhuizen, Erasmus University Rotterdam



# Aims of our multidisciplinary work

- Explore the applicability and usefulness of existing infrastructure & tools for non-digital humanities approaches
- Elucidate why **'language & linguistic' tools are not typically used** by social science or humanities scholars
- Identify the barriers to using such methods and tools
  - paradigms, methods, jargon, technical ability, tool familiarity
- Explore **how** these can be utilised by scholars using interview data

# An open mind...expanding one's toolbox

- Consider **preparing** - **reading** - **listening** - **viewing**
- Features of audiovisual and textual data that users may not have previously considered
- We believe that open-source tools can offer benefits to preparing 'data' and to interpreting them
- Workshop feedback very useful to gather uEx

# Background and user gatherings

- Stream of activity around the exploitation of techniques and tools for working with oral history (OH) data
  - [Oralhistory.eu](http://Oralhistory.eu) site: [Oxford](#), [Arezzo](#), [Munich](#) and [Utrecht](#)
- Explore the diversity of scholarly practices across disciplines who use [interview data](#) sources in their daily work: digital humanities, linguistics, oral history and traditional social science
- Previous 4 workshops - tools specialists, data stewards and scholars

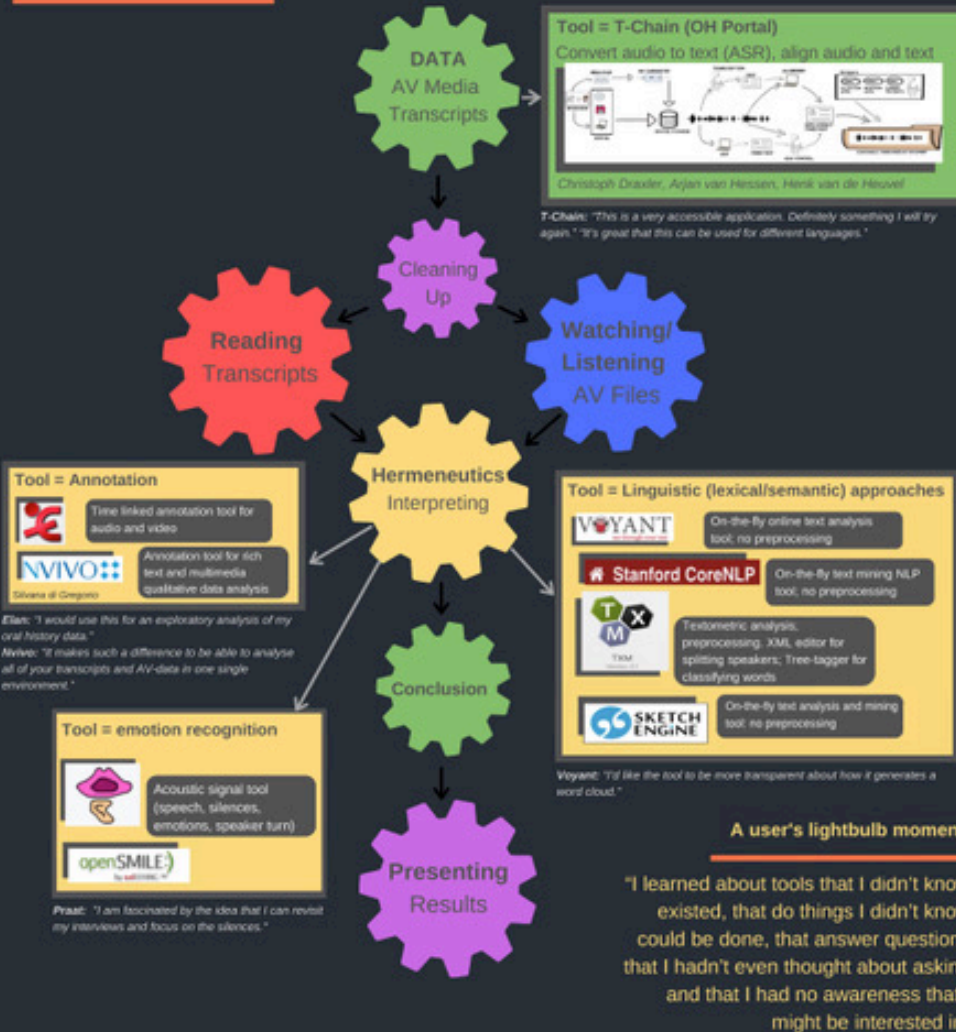
# Types of tools our workshops have explored

- Transcribing and aligning oral sources:
  - TChain
- Annotating text - using pre processing or not:
  - ELAN, NVivo and TXM
- Applying linguistic tools to oral history data
  - Voyant, SketchEngine, StanfordCoreNLP,
- Using emotion recognition tools
  - OpenSmile, Praat

# 🔊 Oral History & Technology 🗣️

Since 2016 a multidisciplinary group of speech technologists, social scientists, linguists and oral historians have come together to explore the integration of digital tools in the existing workflows of scholars who work with oral history and interview data

We are grateful to:



**Organizers:** Arjan van Hesse (Utrecht University / University of Twente), Louise Corti (UK Data Service), Stef Scagliola (University of Luxembourg), Silvia Calamai (University of Siena), Christoph Draxler (Ludwig-Maximilians-Universität München), Jeannine Beeken (UK Data Service), Norah Karrouche (Vrije Universiteit Amsterdam), Max Broekhuizen (Vrije Universiteit Amsterdam), Maureen Haaker (UK Data Service), Khiet Truong (University of Twente)  
See: <http://oralhistory.eu>

# Structure of our workshop

## Part 1: Lecture

Introducing different scholarly approaches when **working with interview data as a primary or secondary data source**: distinct traditions and differences in analytic practices and use of tools by scholars across the disciplines

**Automated speech recognition, annotation, text analysis and emotion recognition tools** are open to wider exploitation

**Presentation of TChain tool**

**Your experiences**

# Structure of our workshop


## Part II Demo and hands-on

The Transcription Chain (TChain) - a tool to convert audio-visual material into a suitable format, use automatic speech recognition (ASR), correct the ASR results, and download them

Discussion of potential of the T-Chain and whether a call for other languages to be integrated into the T-Chain

Emotion recognition video

Exercise & group discussion of your own scholarly work flows and tools when/if analysing interview text



# Information - shared drive

**The workshop materials**

**Tiny url: <http://tiny.cc/sg93dz>**



# A multidisciplinary approach to the use of technology in research: the case of interview data

**Louise Corti  
and colleagues**

Parthenos workshop for  
CEE Countries  
7-9 October 2019,  
Sofia



# The challenge - crossing disciplines

## Working with interview data as a primary or secondary data source

- Vast mounts of interview data waiting to be analysed.....
- New and often unexpected ways of confronting the raw data..'new questions of old data'
- But when it comes to techniques and tools, researchers are siloed..choosing the familiar over novel
- Observance of Tradition

# Different scholarly approaches

- Distinct traditions and differences in analytic practices and use of tools by scholars
  - Automated Speech Recognition
  - transcription
  - annotation
  - text analysis
  - emotion recognition tools
- Potential for digital humanities scholars, historians and social scientists to use speech technology, descriptive and analytical tools for language & linguistic analysis of interview data

# When, where and how?

- These approaches can support different phases of the research process, from data preparation to analysis and presentation
- But the 'whole tool set' will be unfamiliar
- And, tools are very disconnected....
- which can thwart the needs of a researcher's analytic journey
- and hinder new explorations!



# Landscape

What do social scientists do with interviews?



## Dimensions of an interview

<b>Focus</b>	<b>Research question topics</b>
<b>The Interview</b>	Identity (e.g. race, ethnicity, gender, age, etc.), experience and "truth"
<b>Recording of the interview</b>	Power, performance
<b>Transcription of the interview</b>	Narratives, memory
<b>Interpretation of the interview</b>	Subjectivity and intersubjectivity; emotion; cultural representation

Abrams (2010)

	Research questions	Approach
<b>Content analysis</b>	What is the amount and nature of attention/reporting on a subject in a specific media?	Largely quantitative
<b>Conversation analysis</b>	How is social order accomplished in and through interaction in everyday life?	Ethnography and conversation analysis
<b>Discourse analysis</b>	What is the discourse doing? Why some meanings become privilege or taken for granted and others become marginalized?	Study of well-established meanings/ideas around a topic which shape how we talk about it

	Research questions	Approach
<b>Narrative analysis</b>	How do people make sense of what happened and to what effect?	Interpretation not dependent on veracity of the account <ul style="list-style-type: none"><li>• narrative thematic analysis - focus on content and sequence - what is said/not said</li><li>• structural analysis - focus on language and linguistic practices - how is the story told</li><li>• Interactional context - focus of co-construction of narratives - why, when, to whom, in what context</li></ul>
<b>Thematic analysis</b>	What are recurrent ideas and topics in the data?	Look at patterns across interviews. Veracity of the account important. Can be combined with narrative analysis



# Annotation

Texts, images, audio, video - approaches to annotation differ across disciplines

- Linguistics
- Behavioral sciences, communication studies
- Media studies (film, performance, dance, television)
- Social sciences
- Digital humanities

Different tools allow for these perspectives



# Landscape

What do oral historians do with interviews?

Slides by **Norah Karrouche** and **Max Broekhuizen**



## Oral history as interviewing practice

- Oral history methods mostly deal with oral history as a **practice** and as an **approach** to history (minority groups, voice)
- Focus on data gathering and publishing the oral histories
- Weak tradition of reflecting on analytical frames

# Key questions **Oral History**

What can this person tell me about the past that I cannot find in the archive? (narrative is treated as a factual source)

## **After the linguistic turn (mid 1980s)**

Narrative is treated as an interpretation of a memorized personal past

How does this person perceive themselves looking back on the past?

How do they attribute meaning to their experiences in the past?

# Oral history as testimony and as identity

Stories of events, people (and places)

- Interviews **complement** archival sources
- Life stories
  - Interviews gathered in a particular style
  - Historians (oral historians, biographic approaches, ethnographic approaches)
  - Construction of **identity**, intersubjectivity
- Testimonies
  - Historians (oral history, social history, cultural history)
  - Mediation of memory of **events**

# Ideal-typical analytical approaches

## Holistic

- One or more interviews
- Interpretation of a completed sequence of events or a **whole** life story
- *Life story of an immigrant*

## Categorical

- Parts of one or more interviews
- Identifying and interpreting **specific parts** of one or more interview
- *Leaving home in life story of an immigrant*

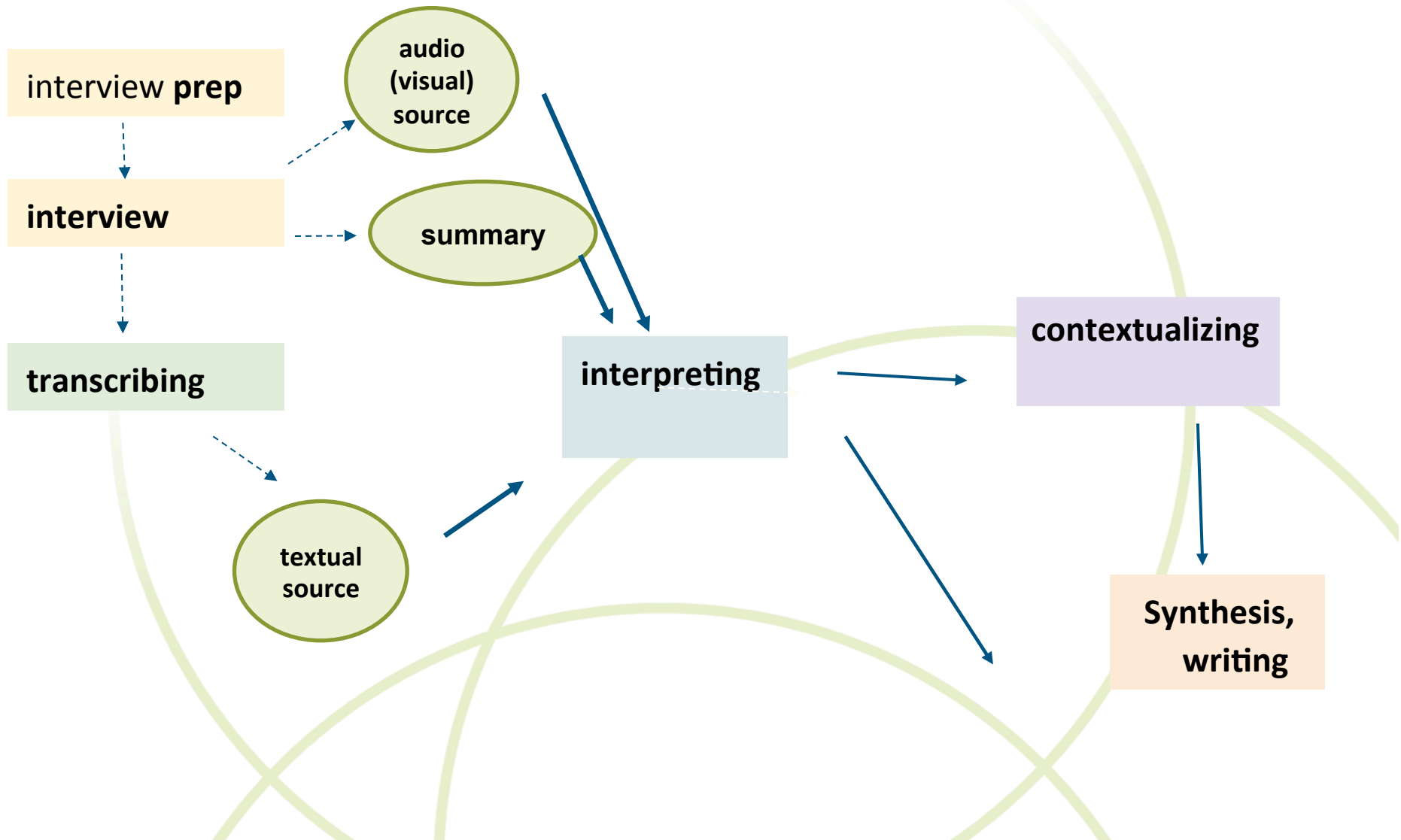
## Thematic

- (Parts of) one or more interviews
- Content: **what** is being said?
- *The experience of leaving home*

- (Parts of) one or more interviews
- Form: **how** is it being said?
- *The way in which the experience of leaving home is narrated (e.g. traumatic, heroic etc.)*

**In practice, we use combinations of these approaches ('bricolage')**

# What do oral historians do with interviews?





# Landscape

What computational linguists do with spoken corpora?

Slides by Jeannine Beeken



## What computational linguists do with spoken corpora?

- **Formulating**

- research **questions**
- **hypotheses** to be confirmed/disconfirmed

- **Exploring**

- and selecting **data** on certain criteria (representativity, language, availability, format, quality, etc.)
- Testing, selecting and evaluating **tools** for analysis

- **Pre-processing** (when needed)

- data **transformation** (e.g. speech to text, transcription), cleaning (e.g. errors correction), conversion (one format to another)
- data **enrichment** (identification of speakers, roles, characteristics - age, gender, language -, uncertain and missing elements, etc.)
- data **structuring/grouping** (chronology, language, typology of spoken data, etc.)

## What computational linguists do with spoken corpora?

- **Analysing**

- applying methods/tools for **part of speech tagging, lemmatization, frequency counts, comparing, clustering, parsing, disambiguation, computing semantic similarities, recognition of named entities, sentiment analysis, etc.**

- **Interpreting**

- analysis results taking also into account the **context** (geo-historical, social, political, linguistic, methodological - in collecting and processing the data, etc.)

- **Evaluating**

- **added value, limitations, biais, possibility of generalisation, range of errors, formulating further questions/hypotheses to be studied, etc.**

# Verbal vs. non-verbal

- verbal: spoken (sounds, speech) vs. written (text, silent)
- Non-verbal: signs, body language and symbols (silent)
- Excerpt from a silent film  
[https://www.youtube.com/watch?v=4QrR\\_1NW\\_w4](https://www.youtube.com/watch?v=4QrR_1NW_w4)  
(1.10min)

# Verbal communication: oral vs. written

- Speech vs. silence
- Spelling: punctuation marks, capital letters, apostrophe, brackets
  - Your first/family name with/without capital
  - Paris Jackson, Orlando Bloom, mother(-)of(-)P/pearl
  - Homophones: seas, sees, seize; friar, fryer; nun, none; grease, Greece
  - Silent letters (speech): knead, kneed (need); knows (noes, nose)
  - Homographs: minute, lead, object, tear
  - Homophone + homograph: arms, ball, spring, duck, watch, ring

# Key challenges for linguistic tools

## Separate/Disambiguate

- Tokenizers (characters, words etc)
- sentence splitters
- Part of Speech (POS) taggers (nouns, verbs etc.)
- Parsers (grammatical structure)
- Named Entity Recognition (NER)
- Word Sense Disambiguation (WSD)

# Key challenges for linguistic tools

## Cluster/Group

- Spell checkers
- Lemmatizers (walk, walks, walking, walked)
- Synonyms
- Multiword Expressions (MWE)
  - collocations, idioms, coreferencing

# Key challenges for linguistic tools

## Control/Reduce

- Stopwords; whitelists and blacklist
- Keyword extraction (with, without a CV)
- Summarizers



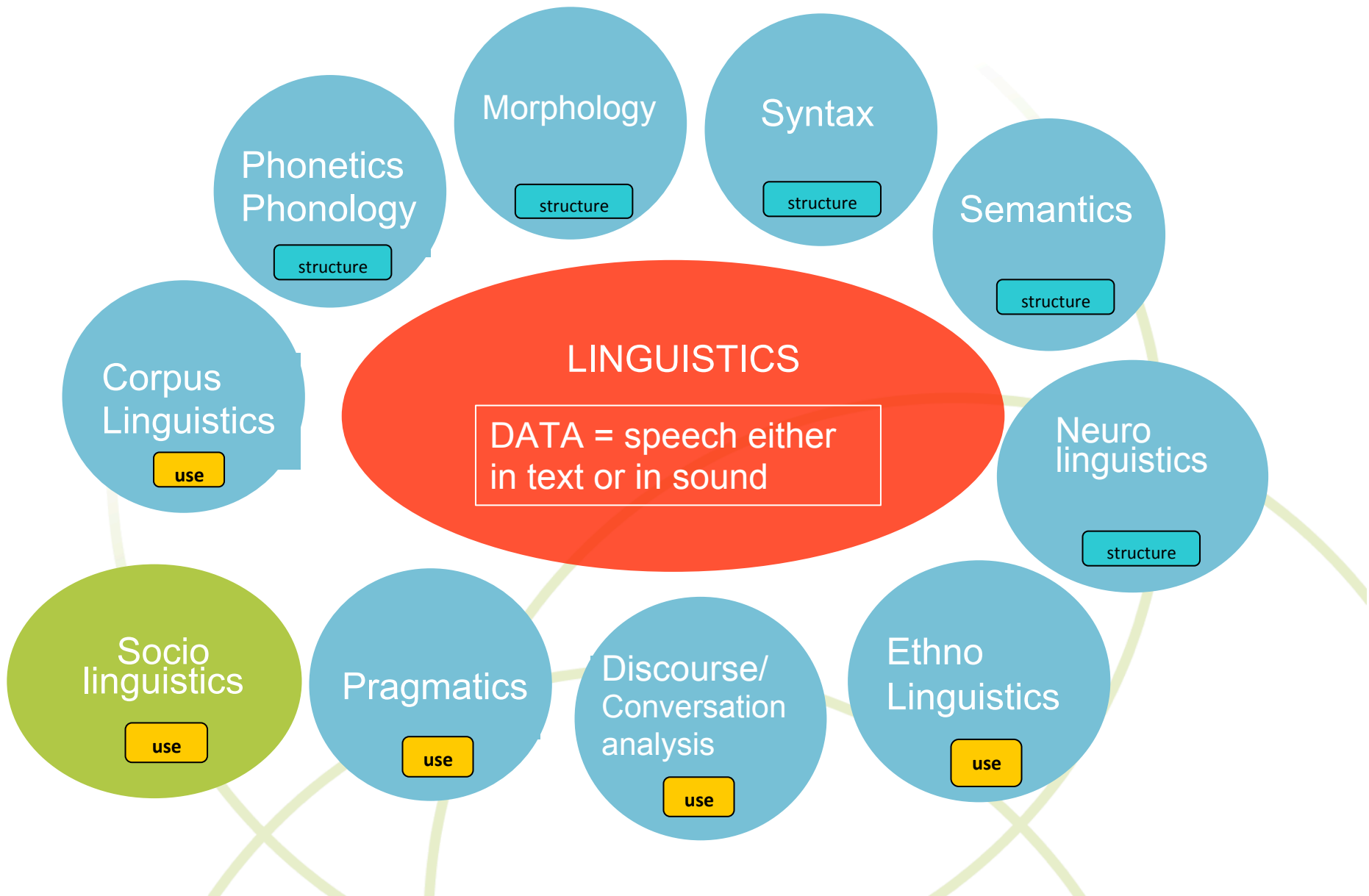
# Landscape

What do sociolinguists do with spoken corpora?

Slides by Silvia Calamai

Sociolinguists meets  
qualitative research?





# Key questions Sociolinguists

## WHO IS SPEAKING ?

- How long? How do they take turns?
  - concepts: overlap/interruption; dialogic repetition; free indirect discourse; polyphonic monologues... )

## WHAT AND HOW?

- What are the topics of the interview, semantics, frequency of words
- What are the styles? What does style tell us about the speakers?
- Why do the speakers speak the way they speak?
- Why do they change style and when?
  - (concepts: accommodation, entrainment, linguistic repertoire)

# TOOLS for the jobs

Discover

Prepare

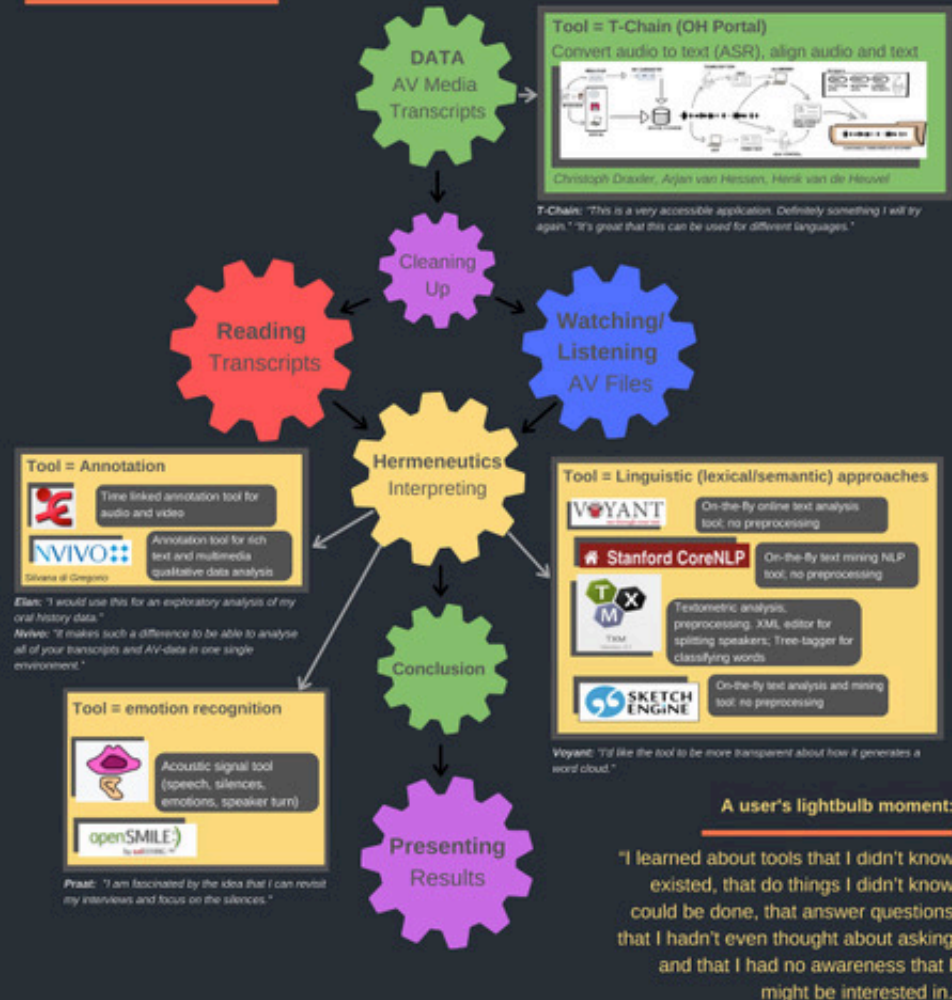
Analyse

Report

## Oral History & Technology

Since 2016 a multidisciplinary group of speech technologists, social scientists, linguists and oral historians have come together to explore the integration of digital tools in the existing workflows of scholars who work with oral history and interview data

We are grateful to:



Organizers: Arjan van Hesse (Utrecht University / University of Twente), Louise Cori (UK Data Service), Stef Scagliola (University of Luxembourg), Silvia Calamai (University of Siena), Christoph Draxler (Ludwig-Maximilians-Universität München), Jeannine Beeken (UK Data Service), Norah Karrouche (Vrije Universiteit Amsterdam), Max Broekhuizen (Vrije Universiteit Amsterdam), Maureen Haaker (UK Data Service), Khiet Truong (University of Twente)  
See: <http://oralhistory.eu>

# Discovery: In depth interview data

- Hard to find publicly available FAIR data sources
- That are **Findable** /discoverable
- Often in **Accessible**: GDPR, consent, copyright
- Rarely are they **Interoperable**: formats/metadata
- **Reusable**? Some have good provenance information

# Available qualitative collections

CESSDA archives: <https://www.cessda.eu/>



From academic research studies

Arising from donations or **Research Data Policy**

- UK Data Service - self deposit / curated collections
- DANS - self deposit / curated collections
- Others

From local cultural enterprise

- Museums and libraries (National Life Story Collection)
- Local records offices

# How NVivo supports analysis



Organizing



Reflecting



Exploring

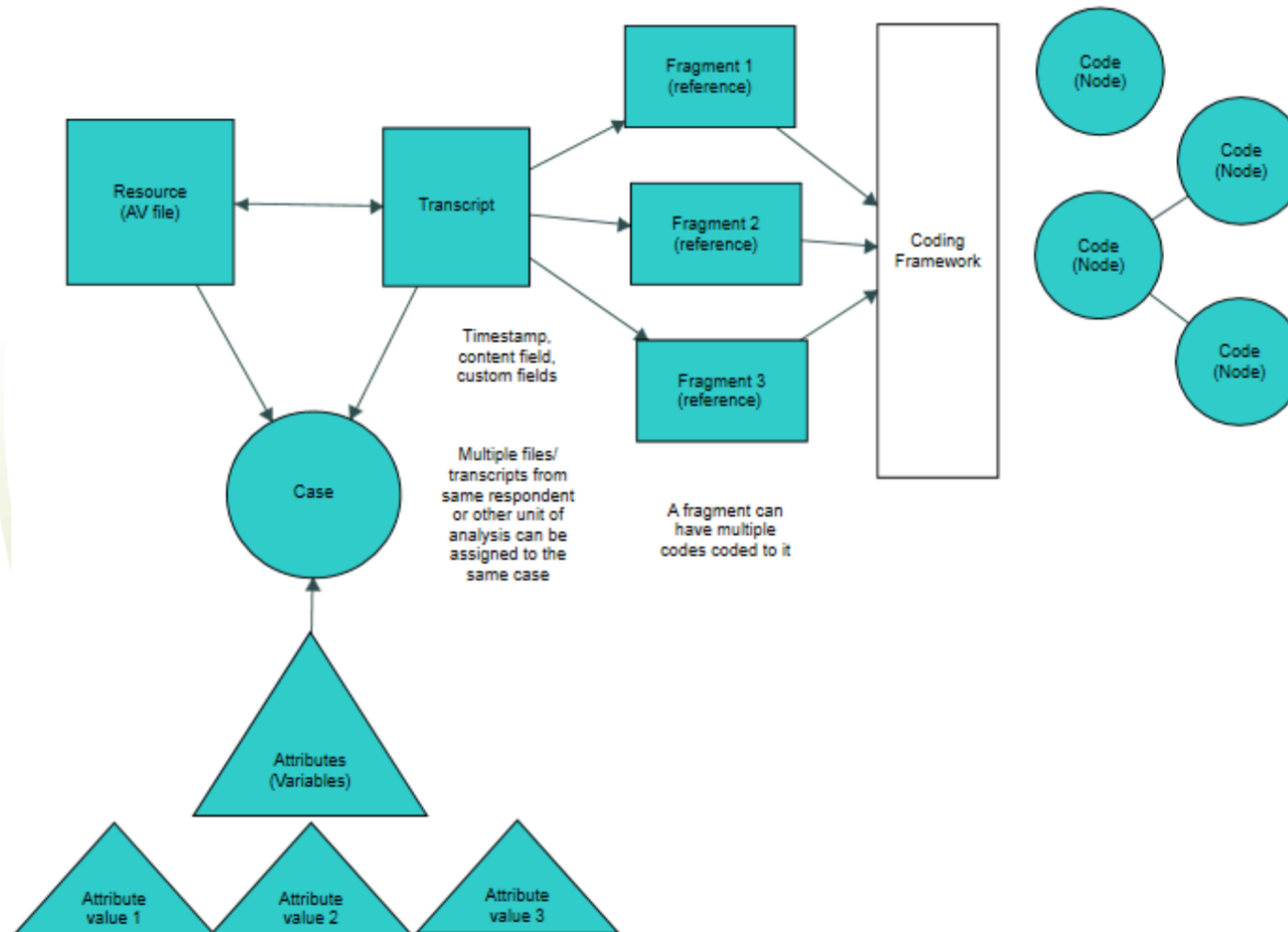


Integrating



Thanks to Silvana di Gregorio for slide input

# NVivo Structure





# The NVivo Workspace - Windows



The screenshot shows the NVivo 12 Plus workspace for a project named 'Sample Project (2).nvp'. The interface is divided into several panes:

- Ribbon:** Located at the top, it contains various toolbars for File, Home, Import, Create, Explore, and Share. A red callout box labeled 'Ribbon' points to the top of the ribbon area.
- Navigation View:** On the left side, it shows a tree view of the project's structure. A red callout box labeled 'Navigation View' points to the 'Interviews' folder under the 'Data' section.
- List View:** In the center, it displays a table of interviews. A red callout box labeled 'List View' points to the table.
- Detail View:** On the right, it shows the content of the selected interview. A red callout box labeled 'Detail View' points to the text area.

The 'Interviews' table contains the following data:

Name	Codes	References
Barbara	43	197
Betty and Paul	13	41
Charles	38	134
Dorothy	39	128
Helen	14	50
Ken	17	56
Margaret	35	78
Maria and Daniel	43	150
Mary and James	42	111
Richard and Patri	35	101
Robert	31	96
Susan	47	146
Thomas	28	112
William	47	106

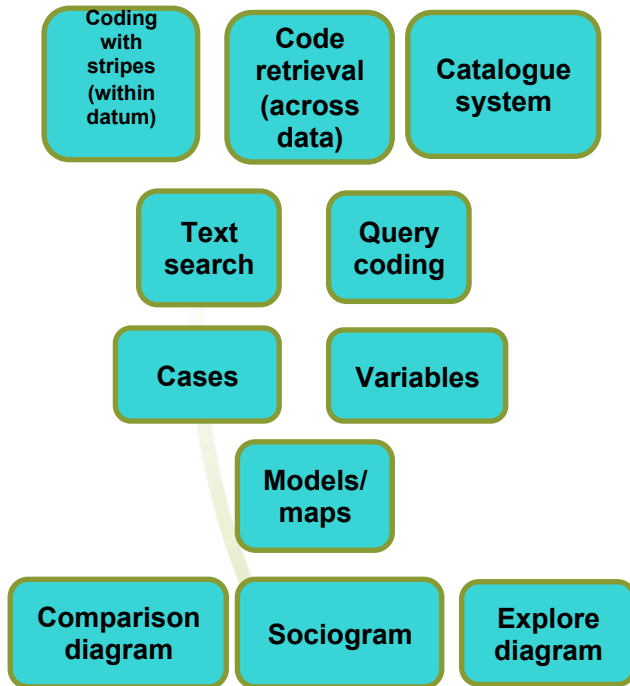
The 'Detail View' shows the content of the 'Barbara' interview, including a section titled 'Q.1. Connection to Down East' and text from 'Henry' and 'Barbara'.



# NVivo Analysis Tools



## Translation of social science manual methods



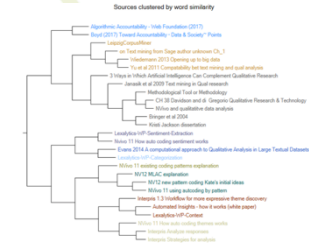
## Basic corpus linguistic tools



Word frequency

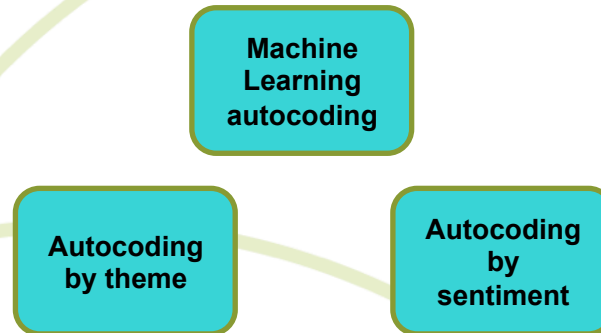


Word tree - KWIC



Clustering by words or coding

## Artificial Intelligence tools



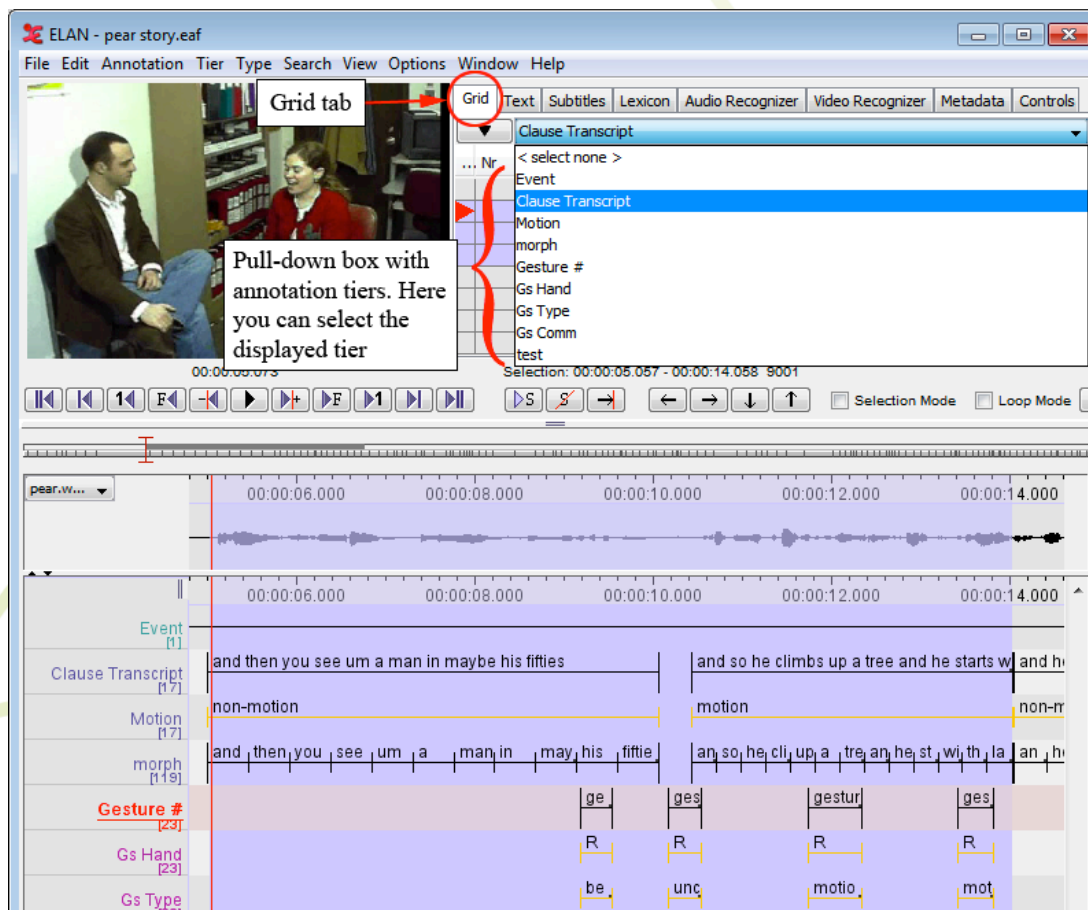
# ELAN (EUDICO Linguistic Annotator)

- Open source desktop tool for manual (computer-assisted) annotation (and semi-automatic enrichment) of multimedia
- Developed at the Max Planck Institute for Psycholinguistics, Nijmegen (The Netherlands)
- **Multi-tier, multi-speaker, time-linked annotation of audio and video recordings**
- Produces xml files (.eaf)

**Thanks to Lilian Melgar for slide input**

# Use of Tiers

- ***“A tier is a set of annotations that share the same characteristics”***
- Tiers are “facets” (same type of annotation, or “event”)
- They represent different analytical aspects
- Tiers contain annotations
- They can be grouped hierarchically



# Tiered approach - 1

The screenshot displays the ELAN 4.9.4 software interface. The main window is titled "ELAN 4.9.4 - elan-ppiSunday\_FINAL\_EVA.eaf". The menu bar includes File, Edit, Annotation, Tier, Type, Search, View, Options, Window, and Help. The interface is divided into several sections:

- Video Player:** Shows a black and white video of a woman drinking from a cup. The current time is 00:05:26.881.
- Scene List:** A table listing 13 scenes with their respective annotations, begin/end times, and durations.
- Annotation Track:** A timeline view showing various tiers of annotations, including Shots, Scenes, Characters, Leisure activities, Street scenes, and Film techniques/ae.

Nr	Annotation	Begin Time	End Time	Duration
1	introduction taxi driver Erwin	00:01:17.970	00:01:46.050	00:00:28.080
2	introduction record seller Brigitte	00:01:46.050	00:02:01.823	00:00:15.773
3	introduction Wolfgang	00:02:01.823	00:02:22.968	00:00:21.145
4	introduction Christl	00:02:22.968	00:02:38.103	00:00:15.135
5	introduction Annie	00:02:38.103	00:02:47.608	00:00:09.505
6	street scenes around Zoo station, trams, traffic, pedestrians, Wolfga...	00:02:53.300	00:05:10.001	00:02:16.701
7	Wolfgang and Christl in street café drinking coffee and talking	00:05:10.001	00:06:35.216	00:01:25.215
8	street scenes after closing hours: dirty streets and street cleaning	00:06:49.000	00:07:23.900	00:00:34.900
9	Erwin cleaning taxi at debot and receiving call from Annie (cinema)	00:07:23.950	00:08:31.950	00:01:08.000
10	people walking in the streets (on way home), tram, taxi, Erwin going...	00:08:31.950	00:09:12.000	00:00:40.050
11	Wolfgang and Christl talking in the street and arranging to meet the...	00:09:12.075	00:10:02.225	00:00:50.150
12	city impressions with children, water, boats	00:10:02.225	00:11:04.972	00:01:02.747
13	Erwin coming home after work, dispite with Annie, Wolfgang arrivin...	00:11:04.972	00:18:46.427	00:07:41.455

The annotation track shows the following tiers and their content:

- Shots [44]:** Wolfgang and Christl sitting on table, waiter brings coffee; Christl drinkin; Wolfgang ta; intertitle. "Embarra; Christl intertitle. "Neither! N; Wolfgang looking and smiling; Christl talki; Wolfgang
- Scenes [13]:** Wolfgang and Christl in street café drinking coffee and talking
- Characters [17]:** Wolfgang, Christl; Christl; Wolfgang; Christl; Wolfgang; Christl; Wolfgang
- Leisure activities [2]:** drinking coffee
- Street scenes [1]:** street café
- Film techniques/ae [14]:** dissolv; medium close-up; close-up; close-up; intertitle; close; intertitle; close-up; close-up; close-up

# Tiered approach - 2

File Edit Annotation Tier Type Search View Options Window Help

Grid Text Subtitles Lexicon Comments Recognizers Metadata Controls

Brooke Baldwin - speech transcript

Nr	Annotation	Begin Time	End Time	Duration
11	her case is THE first known transMISSION in the us	00:00:24.440	00:00:27.790	00:00:03.350
12	HOSPital officials say	00:00:27.880	00:00:29.290	00:00:01.410
13	she had a NUMber of contacts with duncan	00:00:29.335	00:00:32.075	00:00:02.740
14	and that she FOLlowed all the proper protocols	00:00:32.350	00:00:34.840	00:00:02.490
15	the cdc however is not so CERtain	00:00:35.200	00:00:37.650	00:00:02.450
16	ah it believes there was SO:ME kind of breach	00:00:37.923	00:00:41.099	00:00:03.176
17	ah although it cannot PINpoint that	00:00:41.115	00:00:43.203	00:00:02.088
18	that was the phraseology from that news conference YESterday	00:00:43.203	00:00:46.195	00:00:02.992
19	ah a source with direct KNOWedge of the case tells cnn	00:00:46.507	00:00:49.172	00:00:02.665
20	that cdc detectives talked to the nurse severAl Ti:MES	00:00:49.172	00:00:52.240	00:00:03.068
21	and found inconSISStencies	00:00:52.540	00:00:54.135	00:00:01.595
22	about the kinds of protective gear she WORE	00:00:54.135	00:00:56.203	00:00:02.068

00:00:32.350 Selection: 00:00:32.350 - 00:00:34.840 2490

2014-10-13... 00:00:31.000 00:00:32.000 00:00:33.000 00:00:34.000 00:00:35.000 00:00:36.000 00:

Brooke Baldwin - s [30] with duncan and that she FOLlowed all the proper protocols the cdc however is not so CERtain

Body movement Br [6]

Thomas Frieden - s [0]

Elizabeth Cohen - [0]

Brian Steller [0]

Facial gestures Bro [3] EBR

Gaze [10] camera right camera

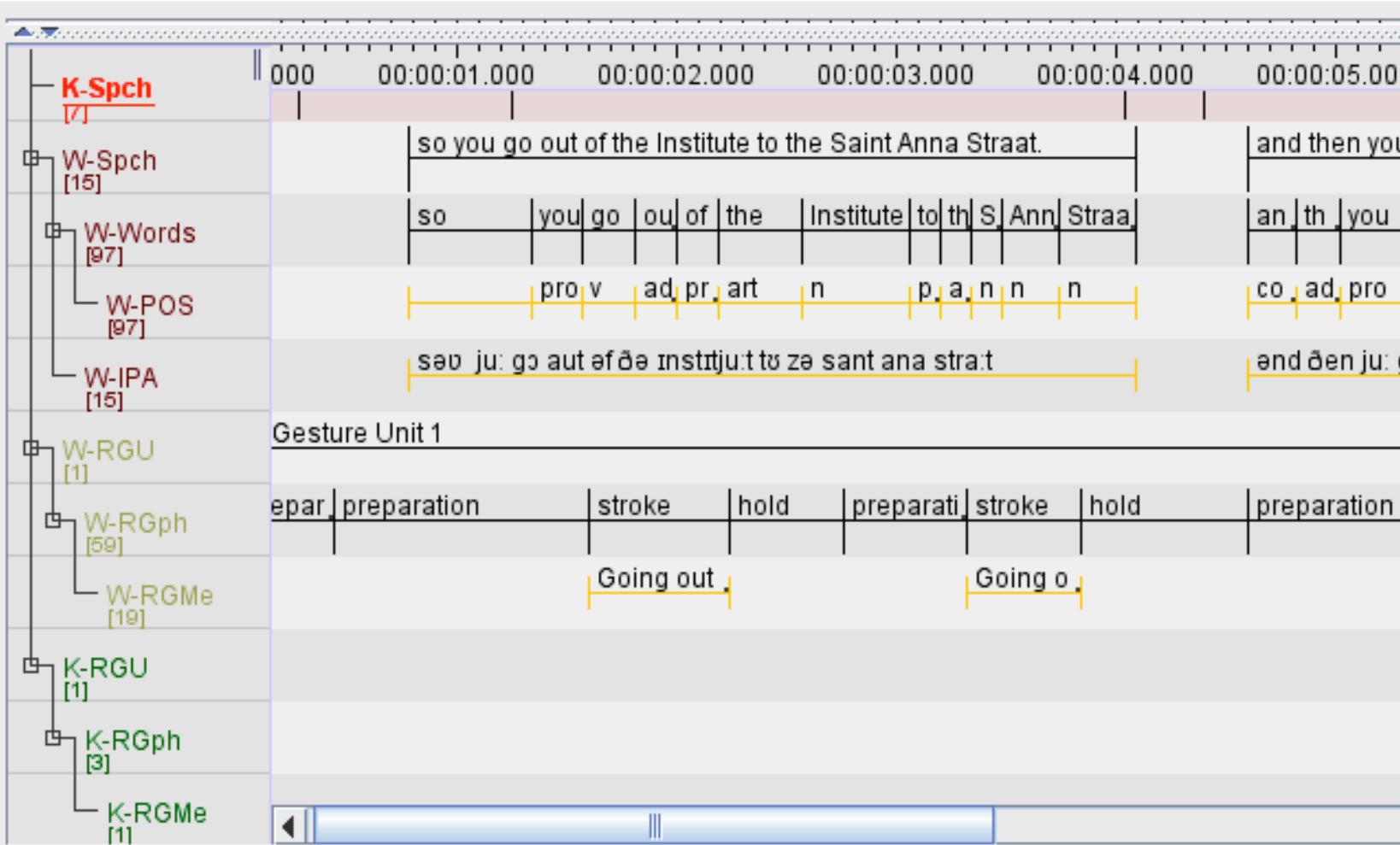
Hand gestures [7]

# Tiered approach - 3

The screenshot displays the Elan software interface for video annotation. At the top, a video window shows two individuals in conversation. Below the video is a control bar with playback buttons and a selection range of 00:00:00.000 - 00:00:00.000. The main annotation area is divided into several tiers:

- speech wMP gest:** Contains the German text "geht dann halt damit ganz schnell raus".
- Translation:** Contains the English translation "and goes then just with it very fast outside".
- MP event:** Contains the event label "carry cage out".
- gesture phase RH:** Shows phases: prep, stroke, stroke, hold, retr.
- gesture descr RH:** Shows descriptions: "cat carry", "cat carry ca".
- GestureMP RH:** Shows gesture codes: "manner", "manner+pat".
- referent RH:** Shows referent codes: "Syl+cag", "Syl+cage".
- hand config RH:** Shows hand configurations: "A3d", "A3a".
- loc/mov RH:** Shows location/movement codes: "cfm->scr", "crm->crm".
- gesture phase LH:** Shows phases: prep, strok, stroke, other g, hold, retr.
- gesture descr LH:** Shows descriptions: "cat c", "cat carry case".
- GestureMP LH:** Shows gesture codes: "man", "manner+path".
- referent LH:** Shows referent codes: "Syl+", "Syl+case".
- hand config LH:** Shows hand configurations: "A3t", "A3t".
- loc/mov LH:** Shows location/movement codes: "clm-", "cfm->cfm".
- gesture phase RH (bottom):** Partially visible tier.

# Tiered approach - 4

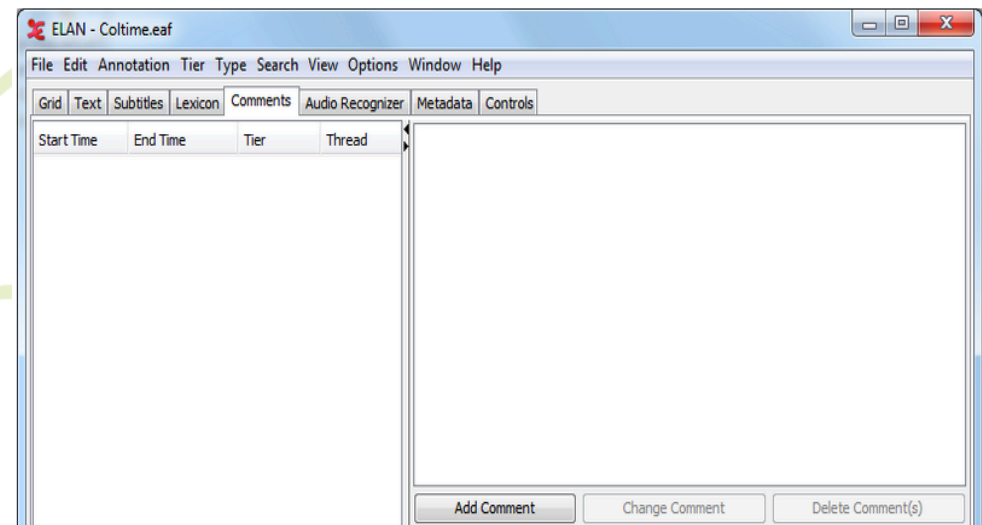




# Annotating in ELAN - master the controls

ELAN understands “annotation” as the content of the segments (e.g., transcripts, codes, etc.):

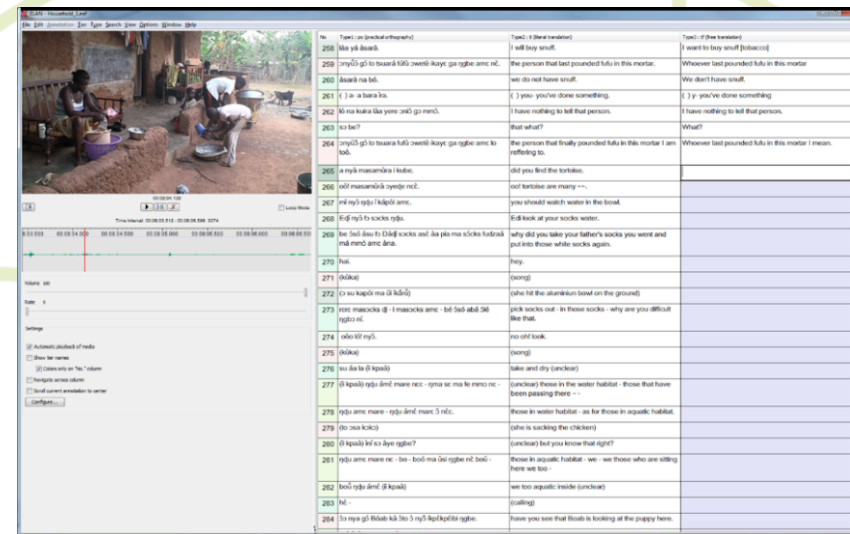
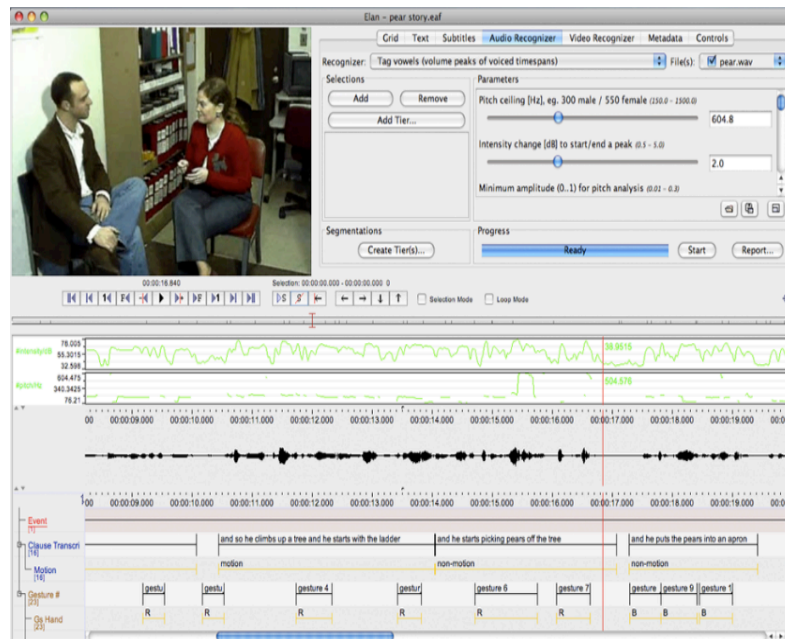
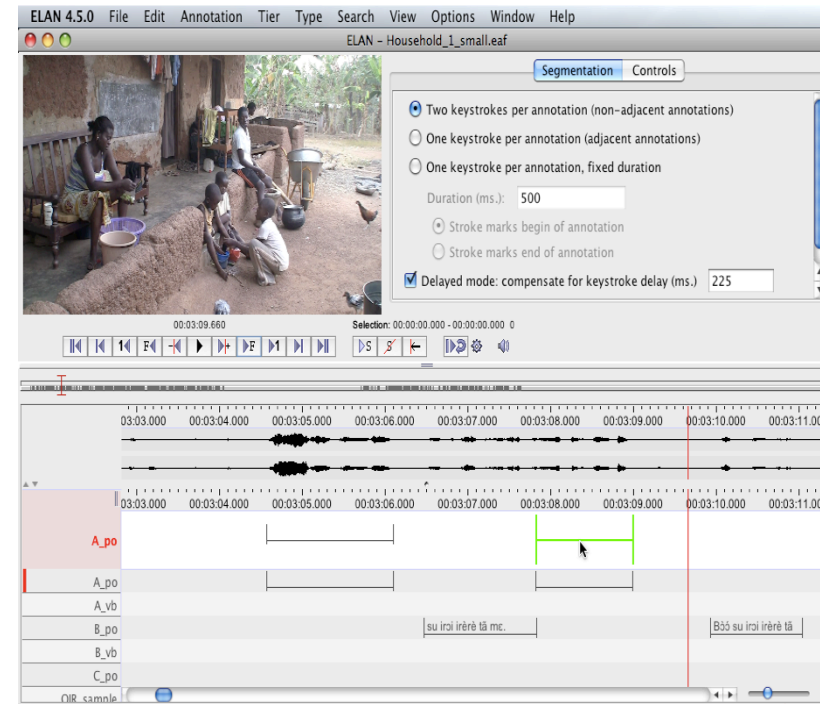
- **segments** are containers for annotations
- Each annotation is entered on a **tier** and assigned to a **time interval**
- There is also the possibility to add **comments**





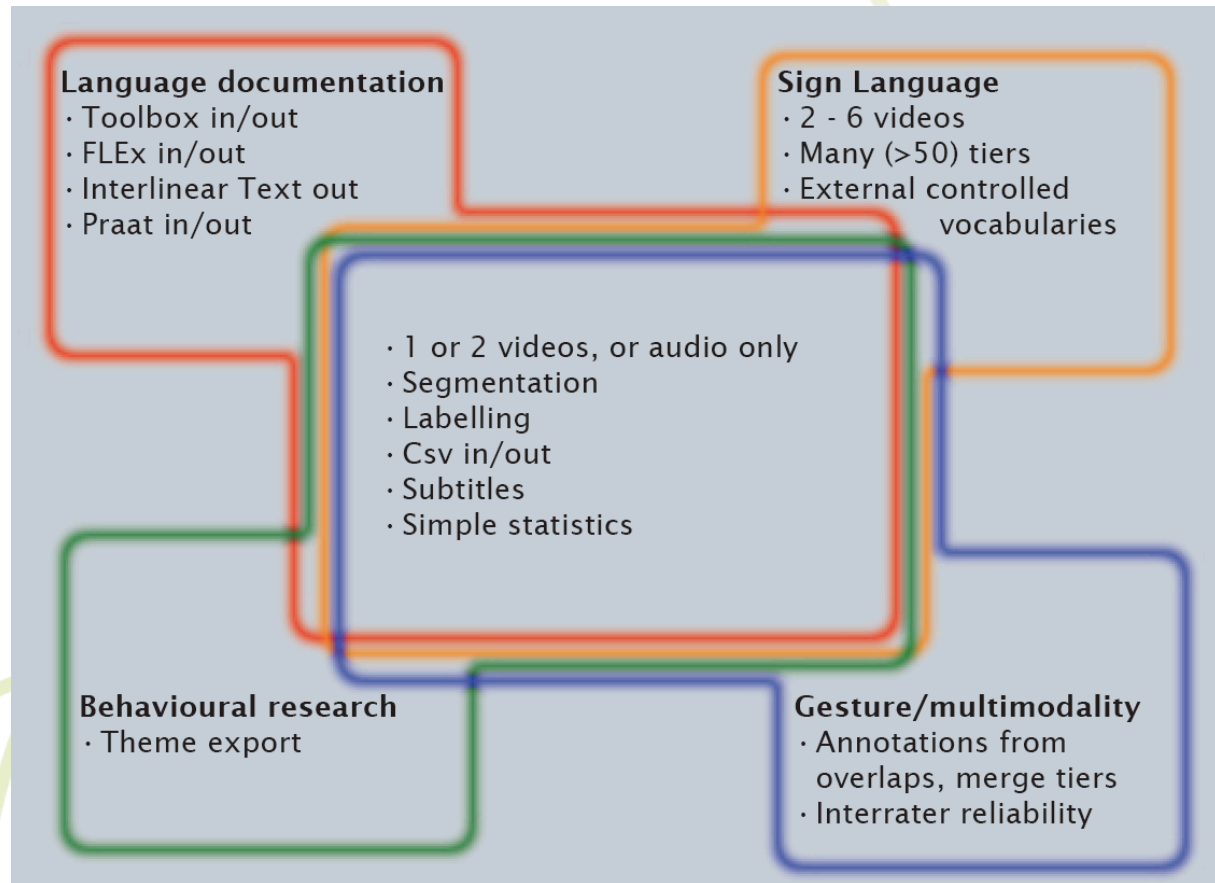
# ELAN's working modes

- Segmentation mode
- Annotation mode
- Synchronization mode
- Transcription mode



# ELAN user communities

- Language documentation
- Sign Language research
- Gesture research
- Multimodality research
- Behavioural studies
- Psychology, psychiatry
- Other...



# Linguistic Analysis - Free Tool

SketchEngine at <https://auth.sketchengine.eu/#login>

User Manual at  
<https://www.sketchengine.eu/user-guide/user-manual/>



# Landscape

What affective social sign processing scholars do with spoken data



# Social Signal Processing

“How can emotions be generated in computers, **recognized** by computers, and expressed by computers?”

“SSP is the computing domain aimed at the modeling, **analysis**, and synthesis of social signals in human-human and human-machine interactions.”

“How can we automatically **analyse** and **interpret** human behavior in human-human and human-machine interactions?”

# Social Signal Processing in Oral History

Khiet Truong's Video:

<https://drive.google.com/file/d/1-ooofaXzR3RWe9fL75NBQGT34B0KdbH2/view>

Location:

[https://drive.google.com/drive/folders/150Qe3itZArR0s2sAArPcaFUJT2YI9Y\\_9](https://drive.google.com/drive/folders/150Qe3itZArR0s2sAArPcaFUJT2YI9Y_9)

## Appreciating differences

- Scholars often ingrained in their ways
- But can see potential in alternative methods
- Can provide additional new angles e.g. for social scientists to explore the role of audio/speech/emotion in a guided conversation
- Social science/oral historians - better question the role of an acquired textual 'transcription'

# Challenge no 1: the jargon

## Jargon - techniques and tools

- Getting to know a 'new paradigm'
- Time is needed to work through analytical approaches and their own dedicated terminology
- **So many acronyms, unfamiliar terms!**
- Annotation types: annotation, coding, nodes
- Different and competing metadata schemas
- Suggest lay descriptions of tools and key functionality



## Challenge no 2: accessing the technology

### Technology - installation and (lack) of documentation

- Time and sometimes expertise may be needed to install open-source software, register, and get it working
- Technical barriers e.g. the platform and OS - linux, mac, windows only
- Many researchers struggle to download software, suggesting a lack of basic technical proficiency
- And the problem of Admin access!

# Challenge no 3: using the tool

## The tool - working the tool

- Time is needed to work through a new tool/ become familiar with it
- Preprocessing of data may be needed
- Watching an orientation video in advance helps
- Step-by-step guides and familiar teaching data and work through exercises very useful
- **Be brave and curious**
- Ask for advice and/or training
- Work **collaboratively** across disciplines

## Challenge no 4: finding and accessing data

### The data - much interview data not open or discoverable

- Hard to find suitable accessible collections for the user
- Limited 'metadata', often manually created, overview. Free text and index terms **often describe the research/data collection method**, not content
- Data publishers could use **language tools to better emerge content** e.g. auto-summarisation and concept extraction



# Thank you!

corti@essex.ac.uk  
@LouiseCorti

ukdataservice.ac.uk  
oralhistory.eu



# Transcription Chain

**Christoph Draxler**  
**Arjan van Hessen**  
**Henk van den Heuvel**

PARTHENOS Workshop  
Oct. 8, 2019  
Sofia



# Overview

Why?

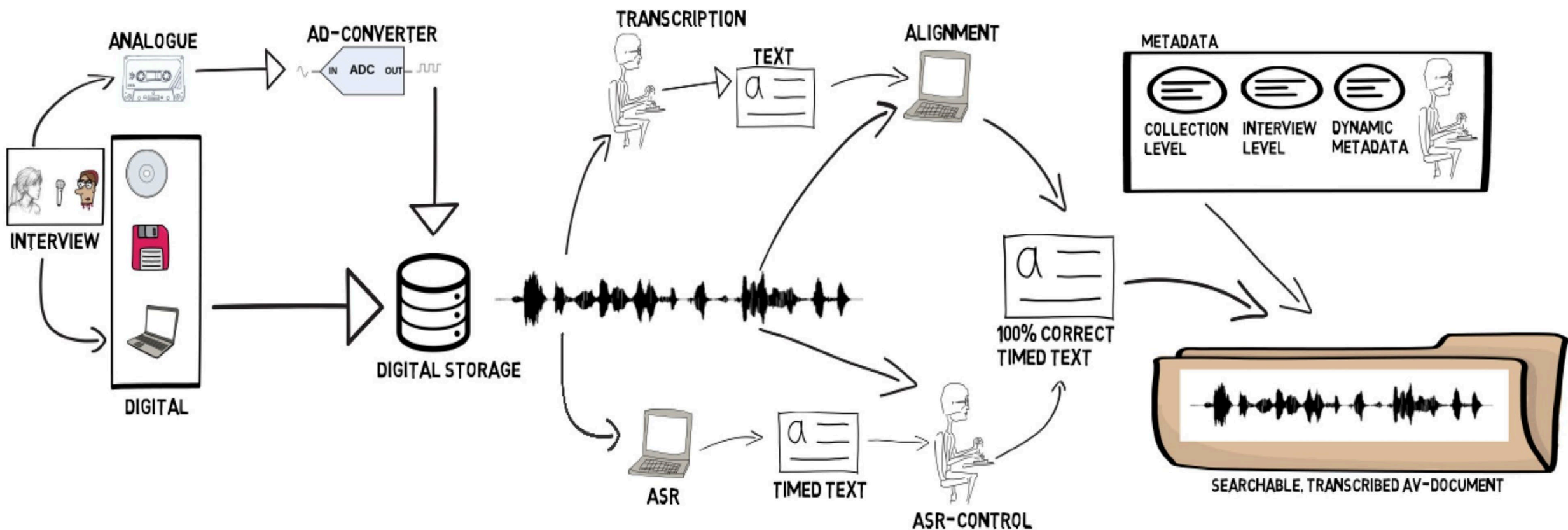
There's no free lunch

It's not rocket science

You get what you paid for

Life is beautiful

# Transcription Chain



# Why?

Transcripts are only one source of data

Audio (and video) signals carry additional information

- Speaker sex, age, health
- Dialect, regional accent
- Emotional state
- Paralinguistic information
- Discourse analysis
- Recording environment

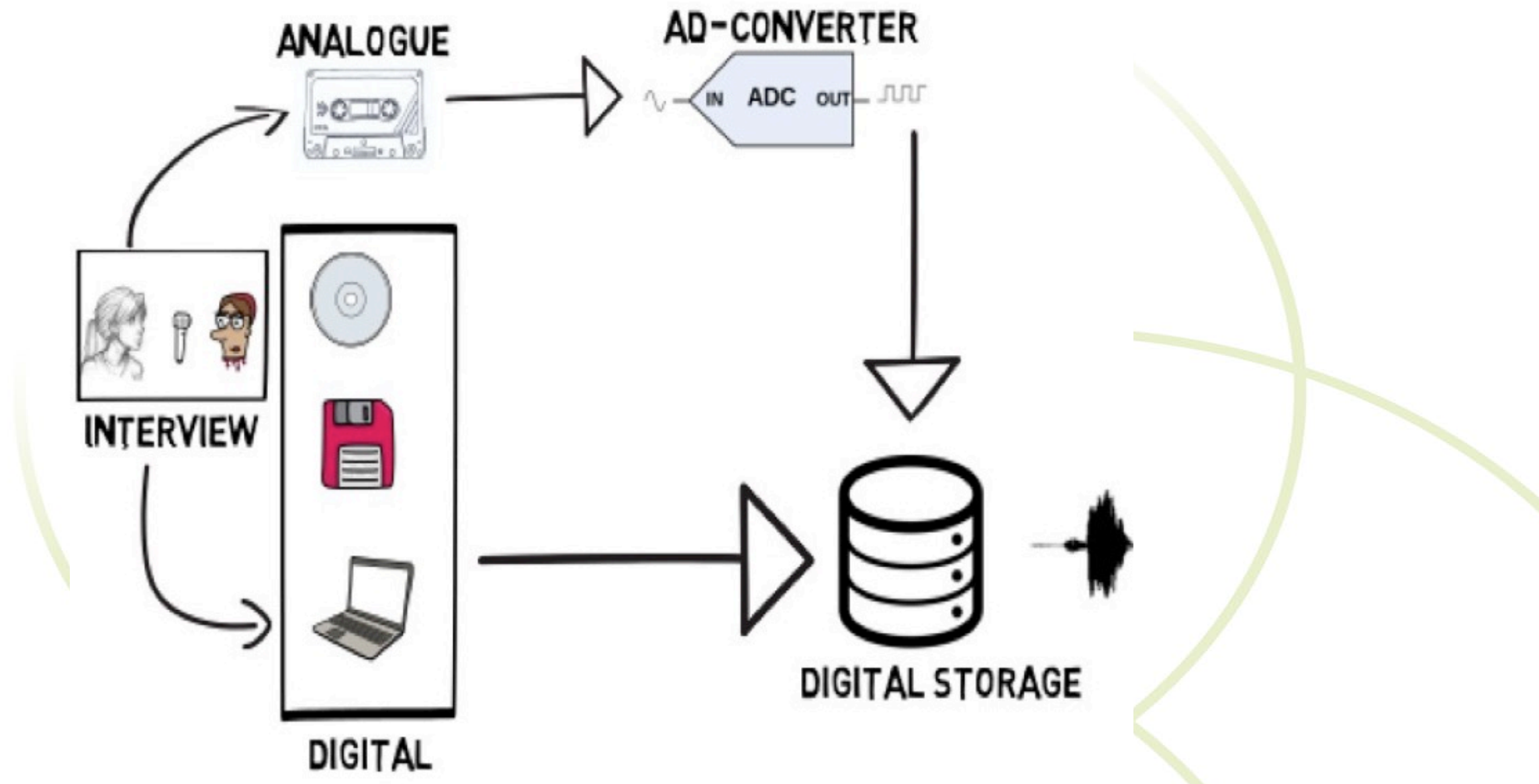






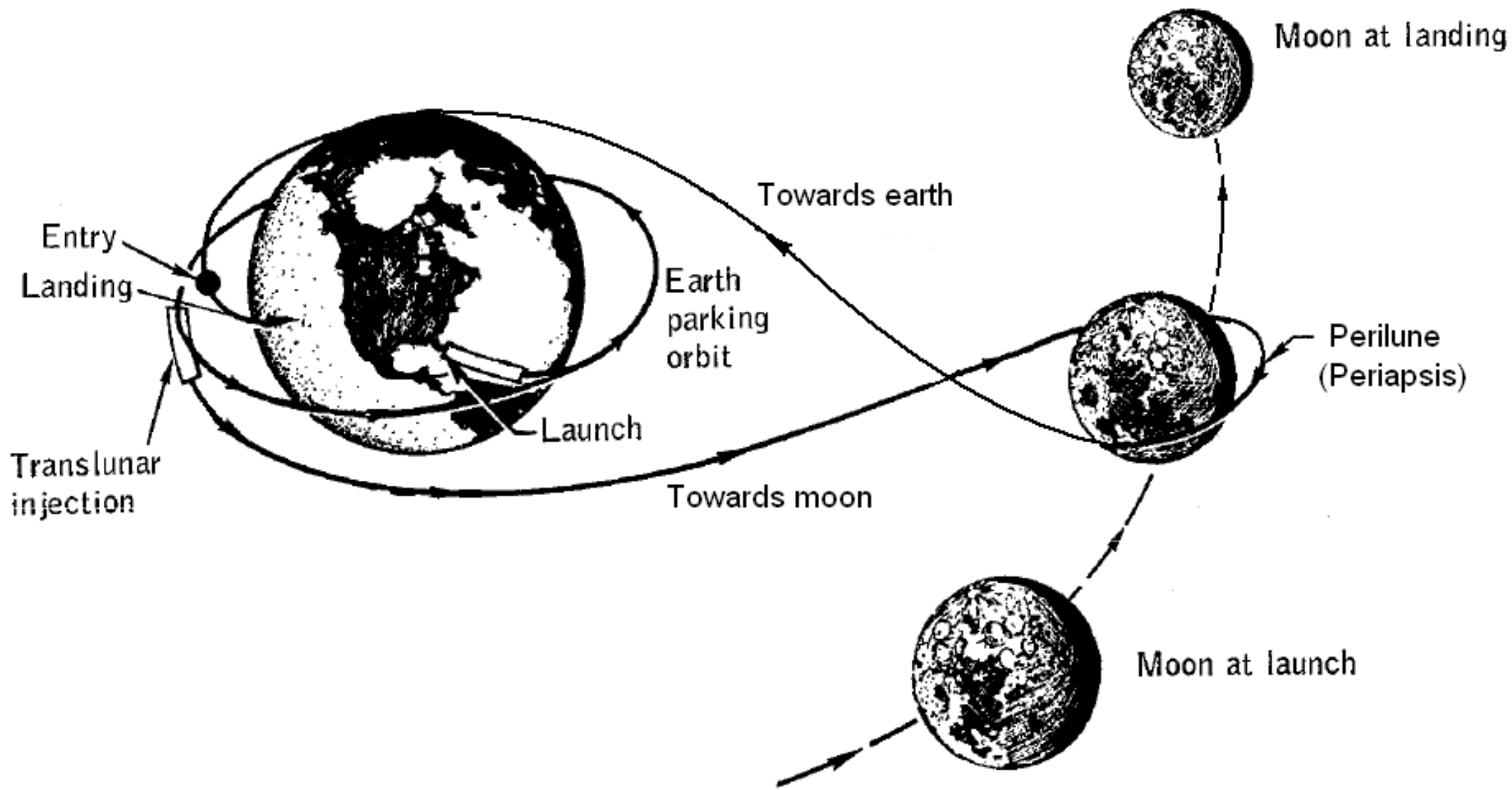
**There's no free lunch**

# Data Collection and Preprocessing

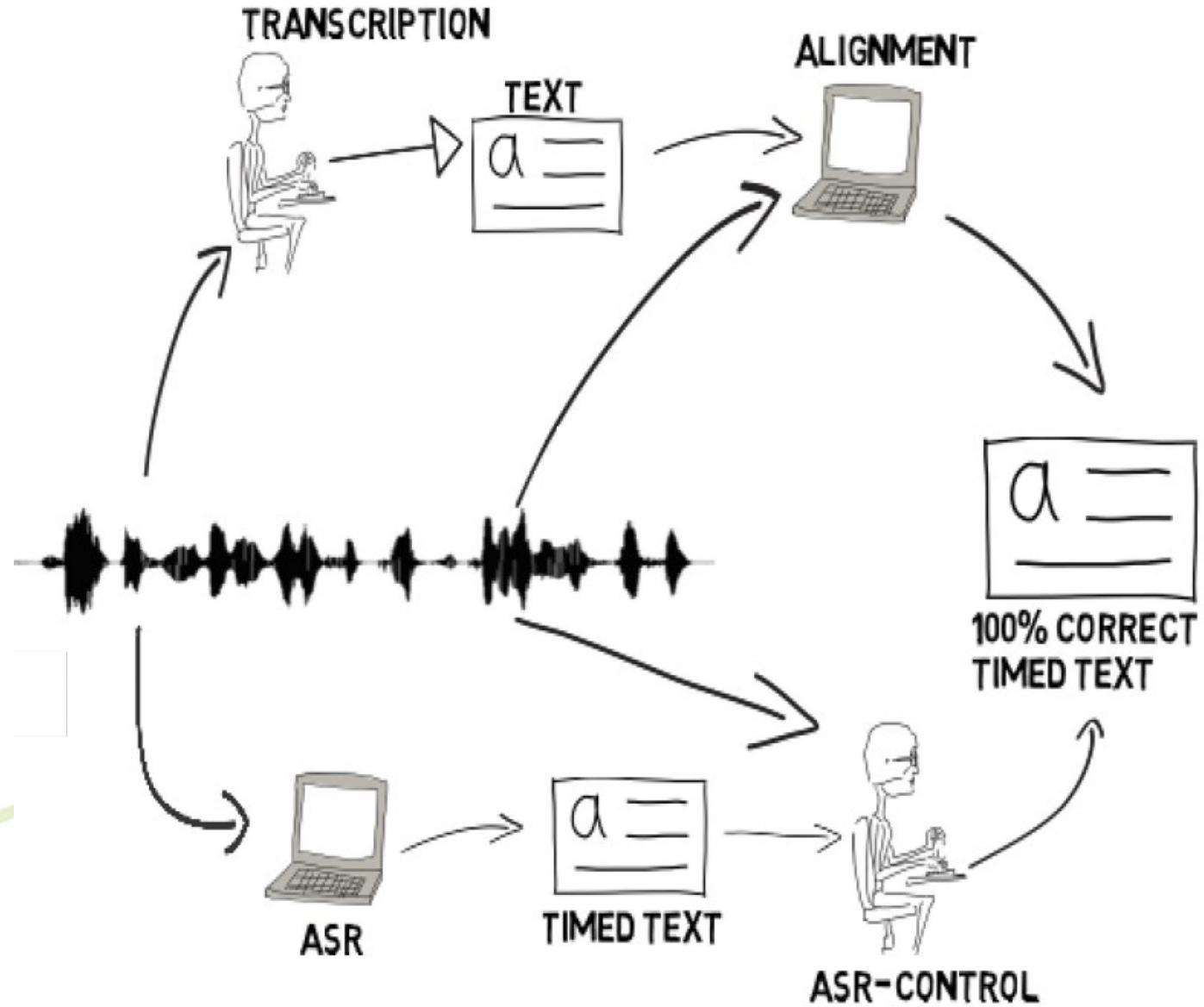




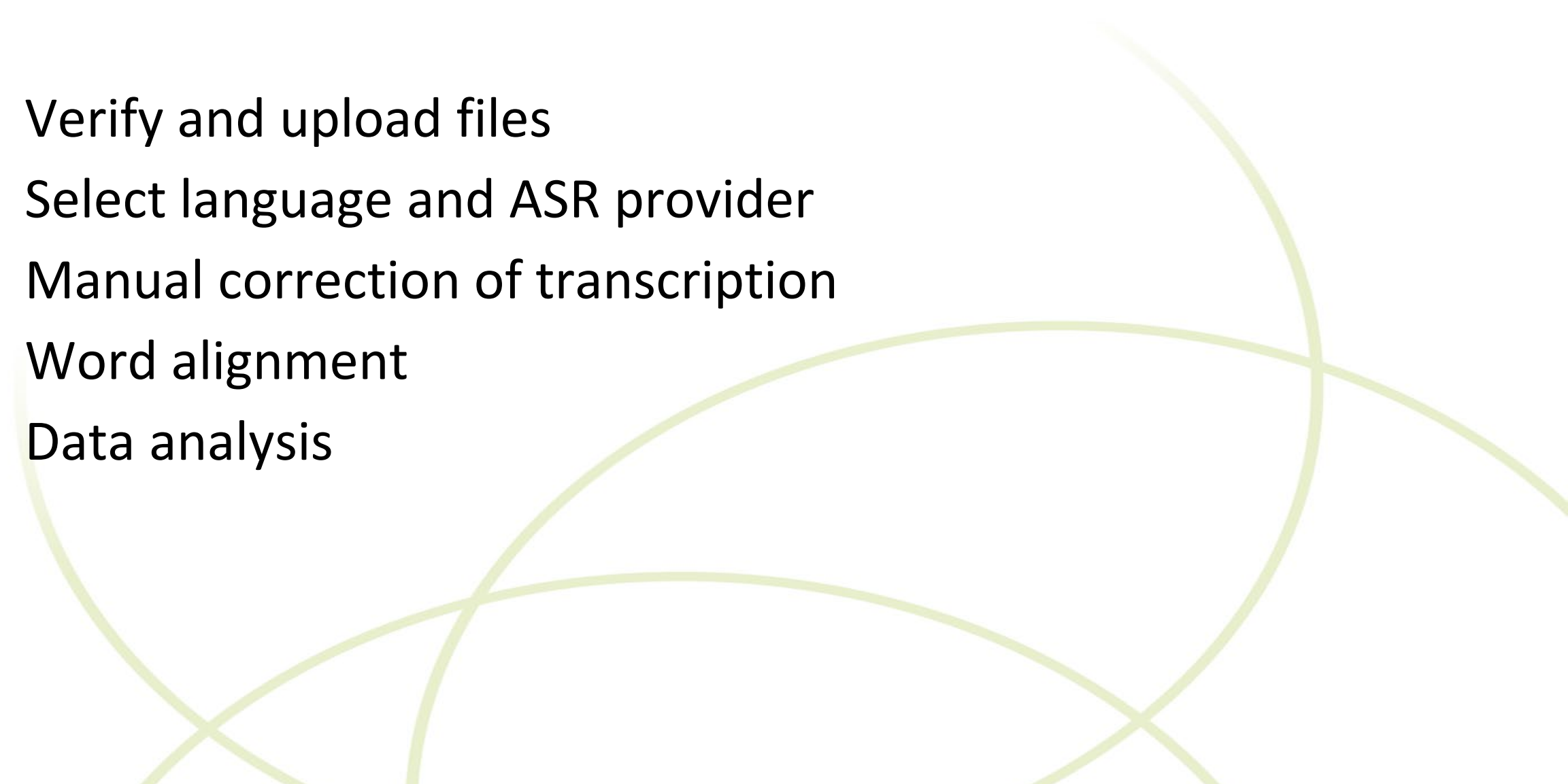
**It's not rocket science**



# ...it's much more difficult



# Processing Steps

- Verify and upload files
  - Select language and ASR provider
  - Manual correction of transcription
  - Word alignment
  - Data analysis
- 

OH-Portal v1.0.0

Ready

+ 1. ADD FILES   2. VERIFY   3. START PROCESSING

File   Upload   Speech Recognition   Manual Transcription   Word alignment   Phonetic detail

## Quickstart

1.

### Add files

You can drag & drop files to this whole area (inside the dashed rectangle) or you can use the "1. ADD FILES" button. Later, you can add files via drag & drop to the table rows, too.

2.

### Verify new files

Before the new files can be processed it's required to set few options. Click on the "Verify" Button.

3.

### Start Processing

After all files are verified you can click on "START PROCESSING". The application starts the processing of pending tasks.

# OH portal page

OH-Portal v1.0.0 Beta

0 1 0 2 0

Help Statistics Feedback

+ 1. ADD FILES 2. VERIFY 1 task needs your attention 3. STOP PROCESSING

File Upload Speech Recognition Manual Transcription Word alignment Phonetic detail

<https://www.phonetik.uni-muenchen.de/apps/oh-portal/>



# Processes & control

The screenshot displays the OH-Portal v0.9.8 Beta interface. The browser address bar shows the URL <https://www.phonetik.uni-muenchen.de/apps/oh-portal/>. The page header includes the title "OH-Portal v0.9.8 Beta" and a status bar with icons for database (0), clock (0), gear (0), checkmark (1), and error (0). Navigation links for Help, Statistics, Feedback, and a settings gear are also present.

The main content area is highlighted with a red box and contains a workflow for processing a file. At the top, there are three buttons: "+ 1. ADD FILES", "2. VERIFY", and "3. START PROCESSING". The status "Ready" is displayed in the center. Below these buttons is a progress bar with five stages: File, Upload, Speech Recognition, Manual Transcription, Word alignment, and Phonetic detail. Each stage is connected by a blue arrow. Checkmarks are visible under the Speech Recognition, Manual Transcription, and Word alignment stages. Below the progress bar, a table shows the progress for a specific file:

File	Upload	Speech Recognition	Manual Transcription	Word alignment	Phonetic detail
0006P032.wav	✓	✓	✓	✓	

# Status table

The screenshot shows the OH-Portal v0.9.8 Beta interface. At the top, there is a navigation bar with the title "OH-Portal v0.9.8 Beta" and a status indicator showing 0 database icons, 0 clock icons, 0 gear icons, 1 green checkmark, and 0 red X icons. To the right of the status indicator are links for Help, Statistics, Feedback, a notification bell, and a settings gear. Below the navigation bar is a workflow bar with three main steps: "1. ADD FILES", "2. VERIFY", and "3. START PROCESSING". The "2. VERIFY" step is currently active, and the status is "Ready".

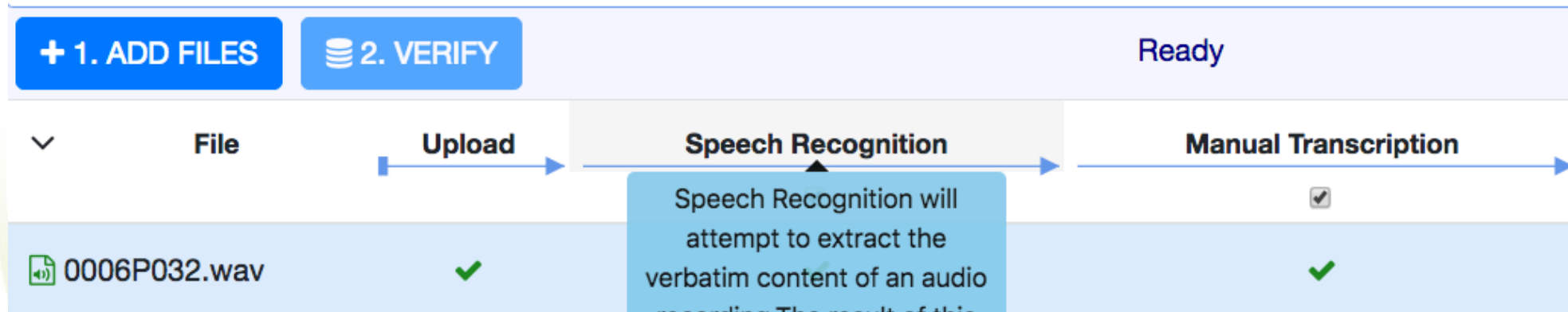
The main content area displays a progress bar with five stages: File, Upload, Speech Recognition, Manual Transcription, Word alignment, and Phonetic detail. Below the progress bar is a table showing the status of a file named "0006P032.wav". The table has five columns corresponding to the stages, and each cell contains a green checkmark, indicating that all stages are completed successfully. A red box highlights the entire table area.

File	Upload	Speech Recognition	Manual Transcription	Word alignment	Phonetic detail
0006P032.wav	✓	✓	✓	✓	

# Tooltips

OH-Portal v1.0.0 Beta

0  0  0  1  0 



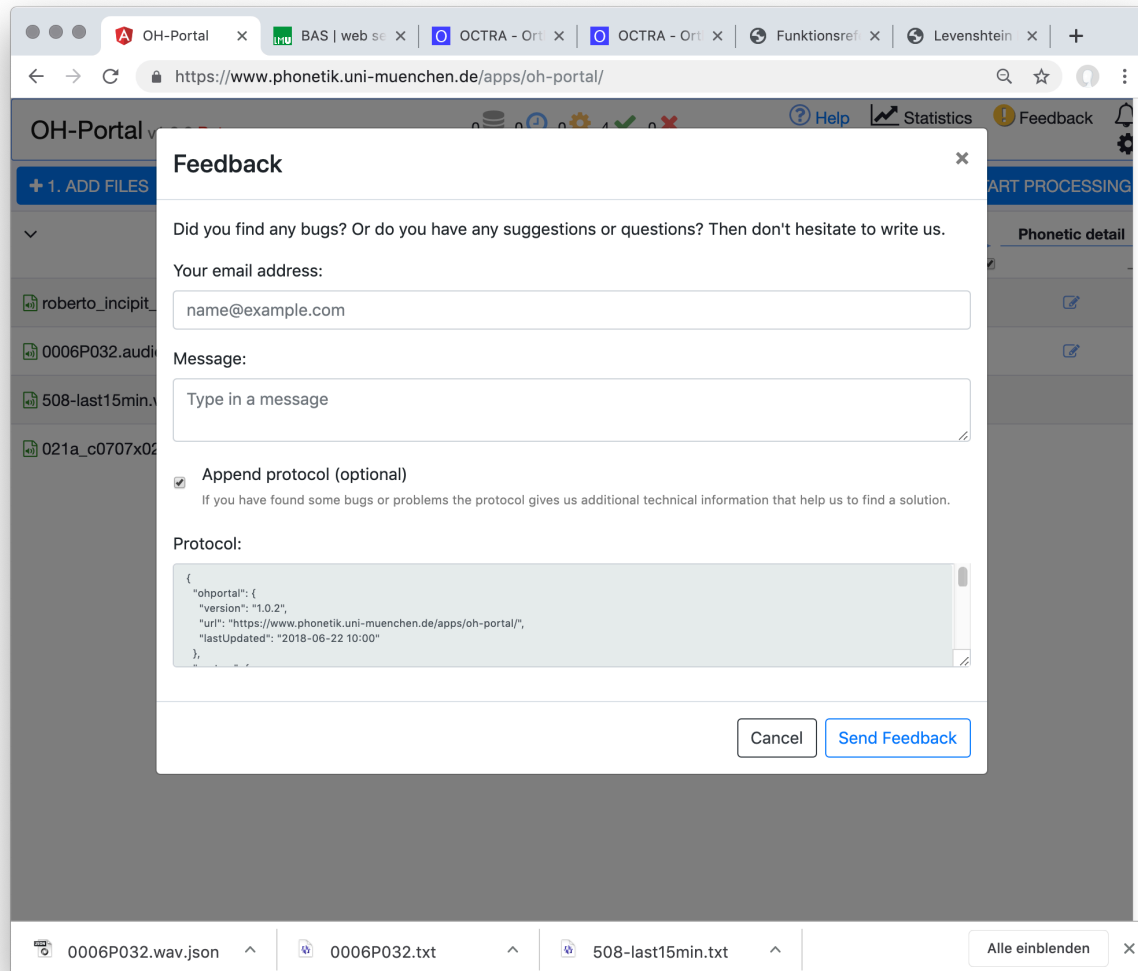
Speech Recognition will attempt to extract the verbatim content of an audio recording. The result of this process is a text file with a literal transcription of the audio file. NOTE: audio files may be processed by commercial providers who may store and keep the data you send them!

# Feedback & information

The screenshot displays the OH-Portal v0.9.8 Beta interface. The top navigation bar includes the title "OH-Portal v0.9.8 Beta", a status indicator (0 database, 0 clock, 0 gear, 1 checkmark, 0 X), and links for Help, Statistics, Feedback, and a settings gear. Below the navigation bar, there are three main buttons: "+ 1. ADD FILES", "2. VERIFY", and "3. START PROCESSING". The "Ready" status is shown in the center. A workflow diagram below the buttons shows the process steps: File, Upload, Speech Recognition, Manual Transcription, Word alignment, and Phonetic detail. A table below the workflow shows the progress for a file named "0006P032.wav".

File	Upload	Speech Recognition	Manual Transcription	Word alignment	Phonetic detail
0006P032.wav	✓	✓	✓	✓	✗

# Feedback



The screenshot shows a web browser window with the URL <https://www.phonetik.uni-muenchen.de/apps/oh-portal/>. A "Feedback" modal window is open, containing the following fields and options:

- Feedback** (modal title)
- Text: "Did you find any bugs? Or do you have any suggestions or questions? Then don't hesitate to write us."
- Your email address:**
- Message:**
- Append protocol (optional)**  
If you have found some bugs or problems the protocol gives us additional technical information that help us to find a solution.
- Protocol:**

```
{
  "ohportal": {
    "version": "1.0.2",
    "url": "https://www.phonetik.uni-muenchen.de/apps/oh-portal/",
    "lastUpdated": "2018-06-22 10:00"
  }
},
```
- 

The background interface shows a file list with items like "roberto\_incipit", "0006P032.audi", "508-last15min.", and "021a\_c0707x02". A file manager bar at the bottom shows "0006P032.wav.json", "0006P032.txt", and "508-last15min.txt".

## Your feedback is important

- Compiles vital system information
- Text input for problem description
- Maps to our server logs to facilitate tracing errors

# Verify audio files

- Drag & drop audio files to browser window
- Click on „Verify“ to perform basic tests
- Optional steps have a check box (default: active)
- Note: some restrictions apply
  - all recordings must be in the same language
  - only WAV-files (mono or stereo) are supported
  - tested on Chrome and Firefox browsers
  - currently only **English, German, Dutch, Italian**

# Verify files settings & upload

OH-Portal v1.0.0 Beta

+ 1. ADD FILES 2. VERIFY FILES 3. START PROCESSING

File

0006P032.wav

0008P011.wav

0008P012.wav

Feedback

Queue

OK




The following files are going to be processed one after the other. Please check if all options are set as you wish. Click "OK" to mark the files for further processing.

Language: German (DE) [EML]

File	Language	Upload	Speech Recognition	Manual Transcription	Word alignment	Phonetic detail
0008P011.wav	deu-DE	✓	✓	✓	✓	✓
0008P012.wav	deu-DE	✓	✓	✓	✓	✓

Cancel OK

# Automatic speech recognition (ASR)

▼	File	Upload	Speech Recognition
	 0006P032.wav	✓	✓
	 0008P011.wav	✓	⚙️
	 0008P012.wav	✓	⚙️

✖ error











✓ success



# View & download intermediate files



18.09.2018 14:52 ⌚ 01:32:34

Results		Conversions			
#	.par	*.ctm	*_annot.json	*.TextGrid	*.Table
#1	 	 	 	 	 

On Windows please use right-click and "Save as" to download a result



**You get what you paid for**

# Levenshtein distance

- Standard measure for the similarity of strings
- Based on edit operations
- counts the number of
  - insertions
  - deletions
  - substitutions
- to convert string A to string B

In the following slides, the Levenshtein distance is used to measure the difference between ASR outcomes

- **Example:** Levenshtein distance between the character sequences

*my argument*

and

*mike yeah argument*

is **7**

- substitute ,y' by ,i'  
insert ,k' and ,e'  
insert ,y' ,e' ,a' ,h'

# Why manual transcription?



## Manual transcript

my uhm argument for useful history is that actually all history starts by being oral because somebody goes and tells somebody about something and then they write it down yeah okay now back to so so the guy writing it down making it wrong or the guy may tell it wrong so just because it is written down it doesn't mean it is true uhm I mean I have commanded a brigade in a war the Falklands war I know how for example where every unit stop me telling you that they already have a diary that has to produce an official diary of what they were doing every day okay and that is logged now in the national archive yeah now some academic historians have been like a tablet of stone brought down by by Moses across war diaries where there are actual lies being written because often a war diary would be written to uhm please the writer and others in the best possible light yeah ok so not necessarily accurate at all secondly things like uhm after action reports a pilot reports the intelligence officer so this is what I did okay it is backed up by photographs and things like that but it is still an oral account patrol coming back will report to the intelligence officer an oral account so without wishing to sort of bore you I actually think oral history has a huge value provided you know what you are doing when you are listening to it and you have a very good idea of who the people are and therefore how trustworthy or other they are I think before I bore you to tears and and

**Manual transcript is considered the gold standard**

# Why manual transcription? (English results)

## Google

my argument for useful history is actually all history start by being horrible because somebody doesn't tell someone about something and they write it down ok now so so the guy writing it down make it wrong with the guy make him wrong so just because it's written it doesn't mean it's true I have commanded a Brigade in a war the Falklands War I know how data is written that the British waters where is unit stocking something every unit has produced an official diary of what they were doing everyday ok and that's launched in the archive academic and look up on war Diaries being like them tablets of stone brought down by Moses it must be true that's not I've come across war Diaries where they are that relies been written was often awarded it'll be written to place the writers and authors in the best possible light ok so it's not necessarily accurate at all secondary things like after action report a pilot earn from the pilot is actually...

## IBM Watson

Mike yeah argument the use of all history is actually all history stop by the wall because somebody doesn't tell someone about something they write down yeah okay now in writing it down maybe it won't go it alone just because it's written down it doesn't mean it's true uh I mean I I cannot be again in the war the fog of war and I know how for example water is correct a quarter of a unit yeah well it can in fact has produced an official direct what they were doing every day in the law now in the in in the national park yeah no some of the local border if being the the brought down by but in truth written that not in the water is where they are actually like being written the water in the region to I'll place the writers and all in the best possible light yeah Kay that's not very accurate at all secondly things like half dressed report of miss in the end of the day ...

# Why manual transcription? (English results)



## Google

my argument for useful history is actually all history  
start by being horrible because somebody doesn't  
tell someone about something and they write it  
down ok now so so the guy writing it down make it

## IBM Watson

Mike yeah argument the use of all history is actually  
all history stop by the wall because somebody  
doesn't tell someone about something they write  
down yeah okay now in writing it down maybe it

Type	manual	Google	IBM
Characters	1723	1395	1173
Words	332	256	238
Levenshtein distance (chars) error rate		481 27.9%	802 46.5%
Levenshtein distance (words) error rate		150 45.2%	213 71.2%

report a pilot earn from the pilot is actually...

# Manual transcription

The screenshot displays the OH-Portal v1.0.0 Beta interface. At the top, the title bar shows 'OH-Portal v1.0.0 Beta' and navigation icons. Below this, a breadcrumb trail indicates the current file: 'OCTRA: 0008P011.par, Language: deu-DE, Audio duration: 00:32'. A menu bar includes 'File', 'UL', 'ASR', 'MT', 'WA', and 'PD'. A table on the left lists files with their processing status:

File	UL	ASR	MT	WA	PD
0006P032.w...	✓	✓	✓	✓	✎
0008P011.w...	✓	✓	⚙️	✎	✎
0008P012.w...	✓	⚙️	✎	✎	✎

The main area shows the 'OCTRA v1.2.6 (url)' player with a green audio waveform. A red vertical line marks the current playback position. At the bottom, a 'SAVE TRANSCRIPTION' button is visible.

Check (and correct) the output of the ASR service



# Listen to and correct transcript

OCTRA v1.3.0 (local) — Dictaphone Editor Linear Editor 2D-Editor OCTRA\_1 ⓘ Export ⚙️ EN ▾

00:00:00 00:00:10 00:00:20 00:00:30 00:00:40 00:00:50 00:01:00 00:01:10 00:01:20 00:01:30 00:01:40

...P my ahm argument for useful history is that actually all history starts by being oral because somebody g...

...oes and tells somebody about something and then they write it down yeah okay now back to so so the guy writing it down making it wrong or the gu...

...y may tell it wrong so just because it is written down it doesn't mean it is true ah I mean I I have commanded a brigade in a war the Falkl...

...ands war I know how for example war diaries are written now the British have a system called war diaries where every unit stop me telling you something you know already...

... no no no no every unit every unit has to produce an official diary of what they were doing every day okay and that is logged now in the national archive yeah now some academi...

...c historian look upon war diaries as being like you know tablet of stone brought down by by Moses it must be true because it is written down it is not...

... I have come across war diaries where there are actual lies being written because often a war diary would be written to ahm place the write...

...r and others in the best possible light yeah ok so it not necessarily accurate at all secondly thi...

...ngs like ah after action reports a pilot returning from missions the pilot is actually telling the intelligence officer so this is what ...

I did okay it maybe backed up by photographs and things like that but it is still an oral account patrols coming back will report to to the intelligence officer a



# Save transcript

OCTRA v1.3.0 (local) — Dictaphone Editor Linear Editor 2D-Editor OCTRA\_1 Export EN

### Statistics

Total segments	Transcribed segments	Segments with breaks	Segments without content
15	13	1	1

No errors found

### Transcript

Segment	Transcription	▶ skip silence
# 01		▶
# 02	<input type="checkbox"/>	▶
# 03	my ahm argument for useful history is that actually all history starts by being oral	▶
# 04	because somebody goes and tells somebody about something and then they write it down	▶
# 05	yeah	▶
# 06	okay now back to so	▶
# 07	so the guy writing it down making it wrong or the guy may tell it wrong so just because it is written down	▶
# 08	it doesn't mean it is true ah I mean I I have commanded a brigade in a war the Falklands war I know how for example war diaries are written	▶
# 09	now the British have a system called war diaries where every unit stop me telling you something you know already no no no no every unit every unit has to produce an official diary of what they were doing	▶
# 10	every day okay and that is logged now in the national archive yeah now some academic historian look upon war diaries as being like you know tablet of stone brought down by by Moses it must be true because it is written down it is not I have come across war diaries where there are	▶

CLOSE

# ASR vs. manual transcription pilot study

- 10 audio recordings
- monologues on ‚communication‘
- 3-5 min long
- video camcorder in seminar room
- 2 transcribers
  - 5 ASR + manual correction,
  - 5 fully manual transcripts
- measure

$$tfactor = \frac{dur_{transcription}}{dur_{recording}}$$

# Results II: ASR vs. manual transcription

type	tfactor
ASR + manual correction	8.52
fully manual transcription	9.43

Expect higher tfactors for real-world data (interviews, historical recordings, dialectal speech, etc.)



**Life is beautiful**

# Download results column-wise

OH-Portal v1.0.0 Beta

1. ADD FILES 2. VERIFY 3. START PROCESSING

File Upload

0006P032.wav ✓

Word alignment ✓ Phonetic detail ✓

Help Statistics Feedback

### Download results by column

This creates a zip-archive of all the selected results and optionally conversions to other formats. If you do not select any additional conversions only the original results are added to the zip-archive.

Add conversions (optional):

- CTM (.ctm)
- Annot.JSON (\_annot.json)
- Text (.Table)
- BAS Partitur Format (.par)
- TextGrid (.TextGrid)

Get package

Close

# Where to go from here?

Format	Tool	Application
CTM	ASR	speech recognition transcripts
EAF	ELAN	multi-level audio & video annotation
AnnotJSON	Emu SpeechDB, any programming language	multi-level phonetic and linguistic analysis, statistics, programming
BAS Partitur	WebMAUS	multi-level (word, phoneme) forced alignment
TextGrid	Praat	phonetic analysis, signal processing
plain text	any text editor	content or pattern search
table	Excel, R, DBMS	content statistics