



Waipapa
Taumata Rau
**University
of Auckland**

Using Digital Tools & AI for Transcription

➤ 3 July 2025

➤ Dr Kyle Hemming, Centre for eResearch

Who am I



```
Analysis v4.R* x | dat4 x | Untitled1* x | 04-linear_gls_models.R | Untitled2* x
Source on Save | Run | Source
⚠ Packages ape, fields, and MuMIn others required but are not installed. Install Don't Show Again
7 library(nlme)
8 library(sp)
9 library(tidyverse)
10
11 rm(list = ls())
12
13 # data -----
14 # family names
15 names22 <- as.matrix(read.csv("Results/csv/families and predictors/study families.csv")) %>% select
(family)[1:22]
16
17 # family names by status
18 current.list <- list.files(
19 path = "C:/Users/s436862/Dropbox/NNF/Results/rasters/scaled",
20 pattern = "_observed.grd")
21 names <- gsub(pattern = "\\_observed.grd$", "", current.list)
22
23 # family richness and environmental and anthropogenic variables
24 fpv <- read.csv("Results/csv/families and predictors/families predictors 1156.csv", header = T)
25
26 # identify spatial auto-correlation -----
27 # store all results for supplementary materials
28 moran_l <- list()
29 gls_l <- list()
30 model_list <- list()
31 ci_list <- list()
32 cor_m <- matrix(nrow = length(names))
33
1:1 (Top Level) | R Script
Console | Terminal | Background Jobs
R 4.4.2 - /Library/CloudStorage/OneDrive-TheUniversityofAuckland/Misc/General/
+ # Order the table for clarity
+ select(Direction, Description, `Before Lockdown`, `After Lockdown`) %>%
+ arrange(Direction, Description)
> print(change_counts_table)
# A tibble: 4 x 4
  Direction Description `Before Lockdown` `After Lockdown`
  <chr> <chr> <int> <int>
1 De-escalation LDA --> SVL 9 14
2 De-escalation Other --> LDA 7 13
3 Escalation LDA --> Other 0 16
4 Escalation SVL --> LDA 4 15
> View(dat4)
```



Experience



Senior eResearch Engagement Specialist

The University of Auckland · Full-time

Oct 2024 - Present · 9 mos

Auckland, New Zealand · Hybrid

💎 Community Engagement, Project Management and +1 skill



Intersect

Full-time · 2 yrs 11 mos

Canberra, Australian Capital Territory, Australia

Digital Research Analyst

Jan 2022 - Jul 2024 · 2 yrs 7 mos

On-site

Business analyst for the University of Canberra, translating requirements for digital research technologies between research groups, faculties, and business units. Developed, advertised, and delivered re: ...see more

💎 Project Management, Client Liaison and +3 skills

eResearch Analyst

Sep 2021 - Dec 2022 · 1 yr 4 mos

Liaison, relationship builder and technology admin and adopter, including driving AI/ML projects within Intersect Australia.

Learning outcomes

- Summarise transcription using AI
- Benefits of AI
- Apply responsible AI
- Understand AI tool considerations
- Identify next steps

Transcription using AI

1. Are you hoping to do transcription for your research?
2. What kind of research data are you collecting e.g., one-on-one interviews?
3. Have you used or are hoping to use AI to help transcribe/translate research data?

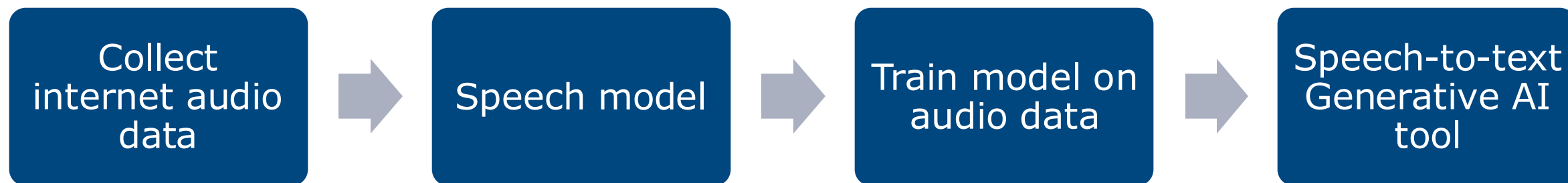
Transcription using AI

Transcription - Speech $\rightarrow$ Text

Translation - Language X $\rightarrow$ Y

Diarisation - Identify speaker A, B

Building an AI model for language recognition



AI transcription tools (searched Feb 2025)

- Adobe Creative Cloud
- Amazon Transcribe
- Audacity
- Azure AI Speech
- Fireflies.ai
- Google Cloud
- IBM Watson Speech
- MS Teams
- MS Word
- Otter.ai
- PANOPTO (via Canvas)
- Rev
- Vibe with Whisper AI
- Zoom (with/without AI Companion)

What do you use, and why?

Benefits of AI Transcription tools

- Speed
- Accuracy
- Scope
- Reduce cost
- Higher accessibility
- Data preparation

Responsible use of AI

What topics do you ask research participants about?

Terms of use example Otter.ai

2. HOW WE USE YOUR PERSONAL INFORMATION

We use your Personal Information to:

- **Set up your account.** We use your registration information, device information and information received from third parties (such as your username, email address) in order to set up an account for you to use our Services. We do so in accordance with our contractual and precontractual obligations to you in order to provide you with an account to use the Services.
- **Provide you with the Services.** We use your audio recordings, usage information and platform information in order to provide you with the Services. In addition, we use your communication information to facilitate support (e.g. retrieval of a forgotten password). We do so in accordance with our contractual obligations to you in order to provide you with the Services.
- **Improve and monitor the Services** We use information we automatically collect or generate about you when you use the Services, as well as information about your device such as device manufacturer, model and operating system, and the amount of free space on your device, to analyze the use of and improve our Services. We train our proprietary artificial intelligence technology on de-identified audio recordings. We also train our technology on transcriptions to provide more accurate services, which may contain Personal Information. We obtain explicit permission (e.g. when you rate the transcript quality and check the box to give Otter.ai and its third-party service provider(s) permission to access the conversation for training and product improvement purposes) for manual review of specific audio recordings to further refine our model training data.

Responsible AI

- Many institutions don't have new, specific AI policies
- Many AI guidelines fit into existing policies
- This area is new, & moving
- But there are useful considerations that already exist

Responsible AI - considerations

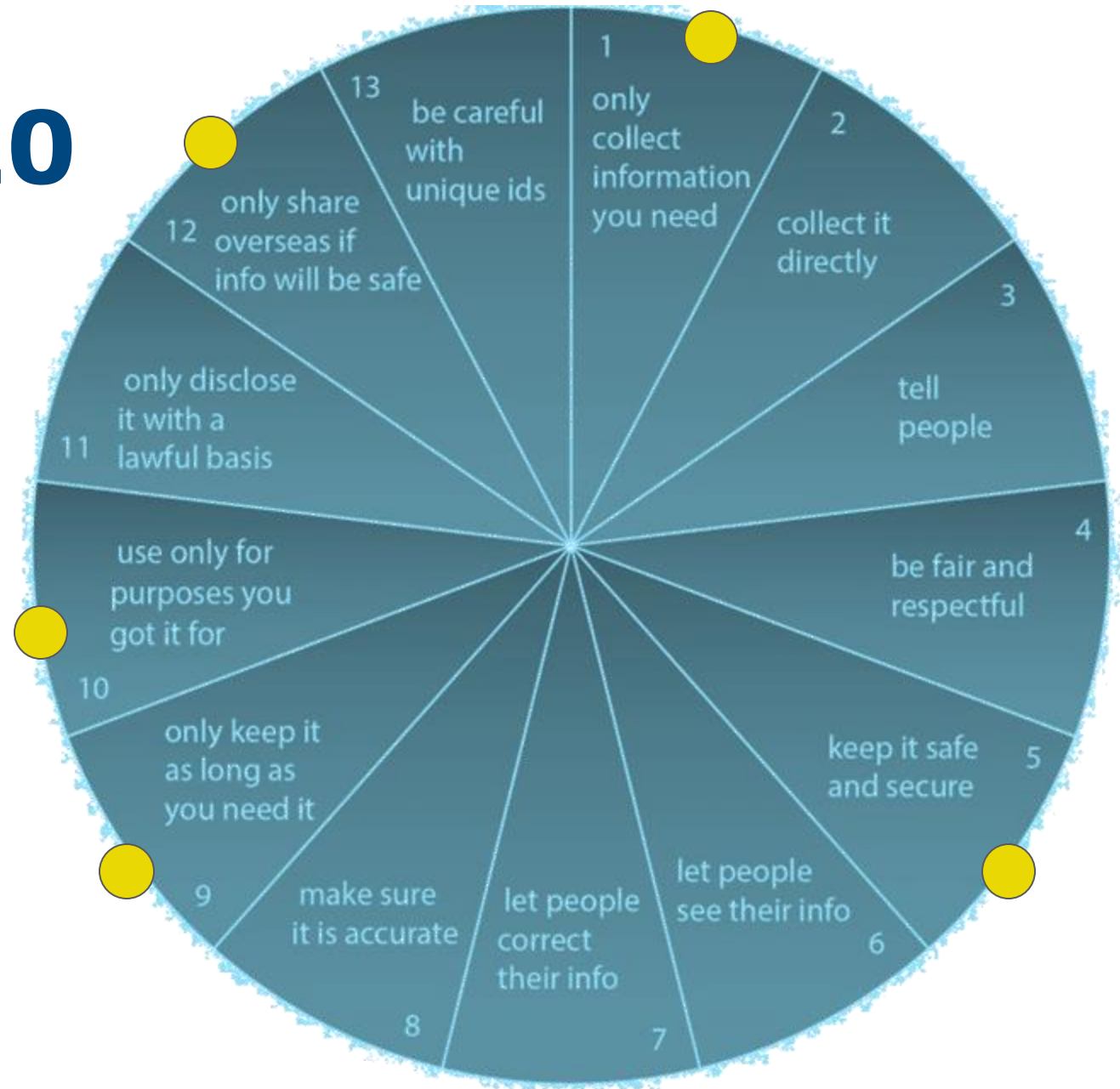
Examples for safe AI adoption

- Privacy Act
- Data classification
- Ethics
- Institutional guidance
- “Approved” tools



Privacy Act 2020 principles

- Assume interview or group discussion recordings contain personally identifiable information



Data classification

- Minimum safety standards for data
- Informs: sharing, storing, archiving, ...

UoA levels:

- Public
- Internal
- Sensitive
- Restricted

Sensitive data considerations

Four levels

- Public
- Internal
- **Sensitive**
- Restricted

Considerations

- Sensitive is common, *normal*
- Personally identifiable information
PII, deidentified health data
- Only suitable for certain tools
- May need to disclose tools in
Ethics application

AI usage guidelines might exist

1. Select Data classification
2. Use approved tool
3. Undertake Privacy Assessment
6. Understand AI limitations and biases
7. Disclose usage

Standards

1. The [data classification](#) of any Inputs submitted to GenAI tools must be established.
2. The choice of GenAI tools must be restricted to those suitable for the data classification level of inputs submitted:
 - Public Data: Any appropriate GenAI tool may be used.
 - Internal and Sensitive Data: Must only be used where a negotiated contract and service agreement exists between the University and the GenAI provider that establishes adequate protection for Inputs. For the avoidance of doubt, adequate protection will ensure that Inputs are not used for any other purpose by the provider, including further training of their public GenAI.
 - Restricted Data: Only services solely controlled by the University may be used.
3. A Privacy Impact Assessment must be completed before a GenAI tool is used with Personal Information.
4. The designated owner of a business function within the University is accountable and responsible for validating GenAI output prior to use of that output to inform business processes within their remit.
5. Users of GenAI tools should consult with the Office of the Pro-Vice Chancellor Māori where Māori data may be used in a GenAI tool, or use may impact Māori.
6. Users of GenAI tools should familiarise themselves with the limitations and/or the possibility of inherent bias within the tool prior to use.
7. Any content (including text, image, or video) intended for publishing or distribution where a substantial portion of the content has been created by a GenAI tool should be labelled as such.

What is an “approved” tool?

Increasingly, institutions are security-assessing tools

Cyber-security, Ethics based

Approved tools list – based on data classifications

- Worth investigating if this exists, or
- What processes are for tool selection

University of Auckland-approved AI transcription systems

MS Word

- Office 365 with University credentials

MS Teams

- with University credentials

Zoom

- without AI Companion, with University credentials

Vibe

- with AI model e.g., Whisper AI installed and run locally on University device

Approved Transcription Software

Members of the University community have access to the following enterprise software that has been approved by the Chief Information Security Office for use with **internal and sensitive data**, in compliance with the [Research data classification standard](#) and [Research Data Management Policy](#) and [guidance](#).

Transcribe audio files with:

- **Microsoft Word** (web and desktop application). Instructions are provided [here](#).

Record and transcribe with:

- **Zoom** (auckland.zoom.us). Go to [Zoom Support](#) for more information about automated transcription of audio from Zoom meetings or webinars.
- **MS Teams**. Go to [MS Teams Recording options](#) for more information about automated transcription and editing of audio from Teams meetings.

Transcribe and translate audio files using AI with:

- **Vibe**. Vibe is an open-source desktop client that uses machine-learning (AI) models (e.g. Whisper developed by OpenAI) to transcribe and translate audio files. When **installed and run on your University-managed device** (laptop, virtual machine) AND **used offline** (Vibe defaults to processing the audio on the computer without sharing data with the internet or sending data to a cloud server), it is able to be used to transcribe and translate internal and sensitive research data audio files (see, [Research data classification standard](#)). Go to [Transcription using AI](#) for more information.



Responsible AI summary

Considerations

- Data classification
- Privacy Act 2020
- University policies
- Approved tools
- Data Sovereignty

Benefits of acting responsibly

- Robust & defensible research practices
- Participant trust & safety
- Institutional reputation
- Streamlined processes and approvals e.g. Ethics

AI tool feature considerations for your research

Showcase: Nectar Virtual Desktop

- Access a virtual computer for 2 weeks - renewable
- Free and available to staff and postgraduate students
- Sign-in using Tuakiri, use Institution log in
- Desktop Library - select the **TranscriptionDesktop**

[Nectar virtual desktop](#)



Desktop Library

 GeoDesktop OPERATING SYSTEM GeoDesktop View Details >	 ubuntu OPERATING SYSTEM Ubuntu 22.04 (Jammy) View Details >	 Neurodesktop OPERATING SYSTEM Neurodesktop View Details >
 Rocky Linux™ OPERATING SYSTEM Rocky 9 View Details >	 TranscribeDesktop OPERATING SYSTEM TranscribeDesktop (beta test) - Selected View Details >	 ubuntu OPERATING SYSTEM Ubuntu 20.04 (Focal) View Details >

Nectar Virtual Desktop: Transcription

 Virtual Desktop Service

ARDC Newsletter Subscribe Contact Us    

Home About Help

WORKSPACE 

abc test

abc test

Good Morning Laura!

Your Virtual Desktop is

 **TranscribeDesktop**

[View Details >](#)

OPERATING SYSTEM	CURRENT SIZE	ZONE
TranscribeDesktop (beta test)	8 VCPUs 16 GB RAM	auckland

Open Desktop >

Reboot

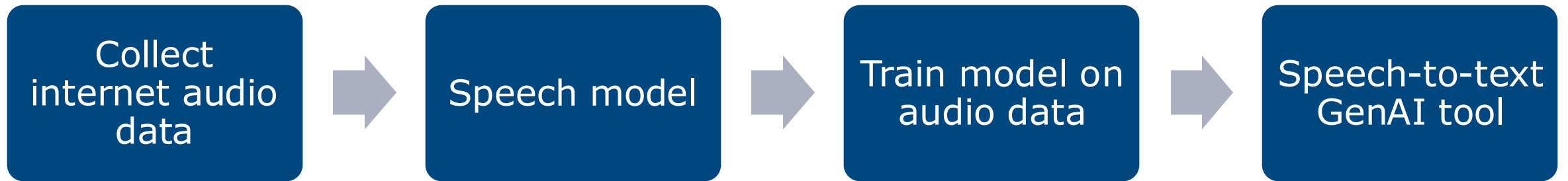
Shelve

Extend

Boost

Delete

Model considerations



Vibe

- Runs Whisper models developed by OpenAI
- Transcription, translation, diarisation
- Trained on 1300 hours of audio
- Translates 100 languages
- Software
 - Run via virtual service
 - Can be installed and run locally
 - Data remains on the device, doesn't 'leak' to cloud



Languages supported

Languages translated with >50% accuracy

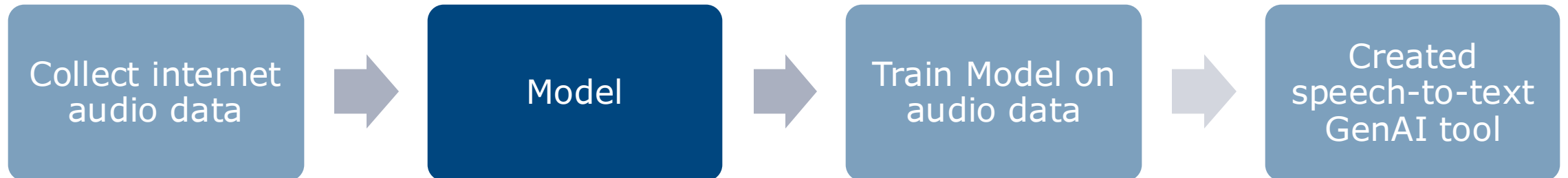
Afrikaans, Arabic, Armenian, Azerbaijani, Belarusian, Bosnian, Bulgarian, Catalan, Chinese, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Greek, Hebrew, Hindi, Hungarian, Icelandic, Indonesian, Italian, Japanese, Kannada, Kazakh, Korean, Latvian, Lithuanian, Macedonian, Malay, Marathi, **Māori**, Nepalese, Norwegian, Persian, Polish, Portuguese, Romanian, Russian, Serbian, Slovak, Slovenian, Spanish, Swahili, Swedish, Tagalog, Tamil, Thai, Turkish, Ukrainian, Urdu, Vietnamese, Welsh

- **Still need to check transcription**
- **Still need someone who knows the language**

Speed vs Accuracy



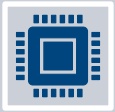
Whisper model sizes: consider them



Speed vs Accuracy



Whisper model sizes: consider them



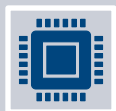
Tiny, Base (39M, 74 M parameters)
requires ~1-2GB of RAM

Less accurate but fastest

Speed vs Accuracy



Whisper model sizes: consider them



Tiny, Base (39M, 74 M parameters)
requires ~1-2GB of RAM

Less accurate but fastest



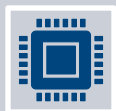
Small, Medium (244M, 769 M parameters)
requires ~2-8GB of RAM

More accurate but slower

Speed vs Accuracy



Whisper model sizes: consider them



Tiny, Base (39M, 74 M parameters)
requires ~1-2GB of RAM

Less accurate but fastest



Small, Medium (244M, 769 M parameters)
requires ~2-8GB of RAM

More accurate but slower



Large (1550 M parameters)
requires ~10-16GB of RAM

Most accurate but slowest

Speed vs Accuracy



Whisper model sizes: consider them



Tiny, Base (39M, 74 M parameters)
requires ~1-2GB of RAM

Less accurate but fastest



Small, Medium (244M, 769 M parameters)
requires ~2-8GB of RAM

More accurate but slower
– **Medium default**



Large (1550 M parameters)
requires ~10-16GB of RAM

Most accurate but slowest

Transcribe + translate accuracies

Scenario

Accuracy

English (US / UK)

English (NZ / accented)

French

Mandarin

Māori

Collect internet
audio data



Model



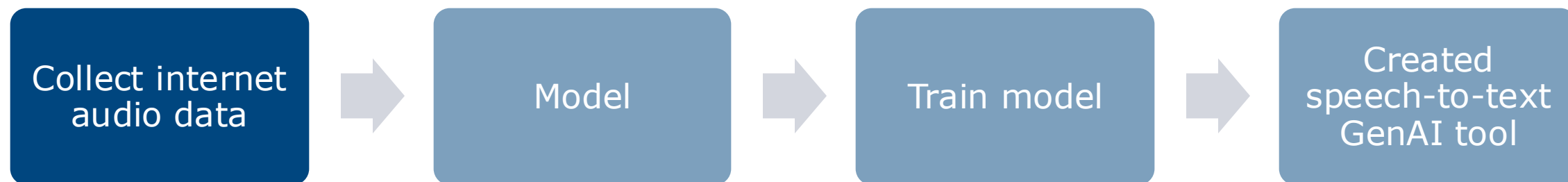
Train Model on
audio data



Created
speech-to-text
GenAI tool

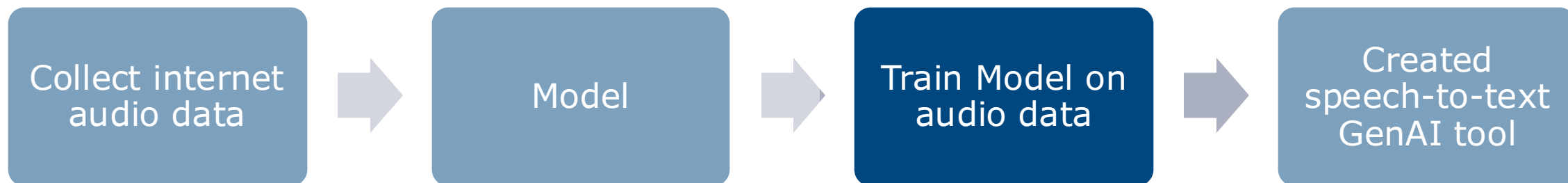
Transcribe + translate accuracies

Scenario	Accuracy
English (US / UK)	Best
English (NZ / accented)	Pretty good
French	Pretty good
Mandarin	Pretty good
Māori	OK



Transcription/translation accuracy

Scenario	Accuracy	Subjective scores
English (US / UK)	Best	>90%
English (NZ / accented)	Pretty good	+80
French	Pretty good	+70
Mandarin	Pretty good	+70
Māori	OK	~50



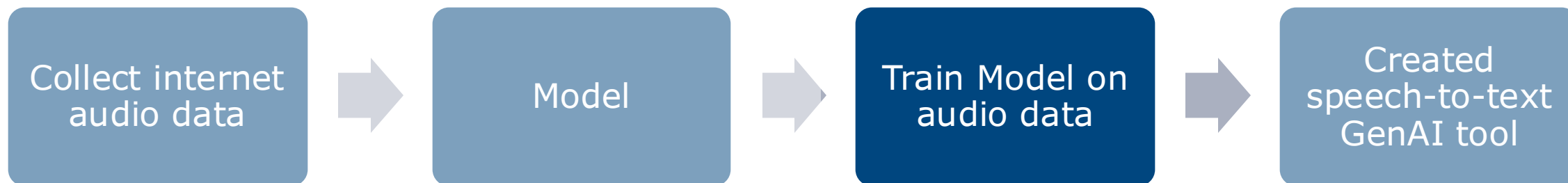
What else affects accuracy

Scenario	Accuracy	Subjective scores
English (US / UK)	Best	>90%
English (NZ / accented)	Pretty good	+80
French	Pretty good	+70
Mandarin	Pretty good	+70
Māori	OK	~50

Collecting recordings:

- Bad microphone
- Background noise
- Speaker clarity
- Overlapping voices
- Language rarity

Researchers NEED to check transcription and translation outputs



Takeaways & next steps

- AI can help research
- Be aware of “free” AI tools
- Understand data, project context
- Find “approved” tools for your research
- Seek approval for your research (e.g. Ethics)
- Use tool safely (e.g. Virtual Desktop)
- Check, verify output
- Enjoy benefits



Waipapa
Taumata Rau
**University
of Auckland**

***Hacky hour
Tuesdays, 3-4pm***

***Strata Café, L4 Kate Edgar Information
Commons***

&

<https://auckland.zoom.us/my/hackyhour>



Waipapa
Taumata Rau
**University
of Auckland**

**Thank you joining today,
we hope this was helpful
for your research**