

Automatic Speech Recognition and the EHR: The HoMed project

Henk van den Heuvel
3 October 2022
Health RI & DTLs



And then there was speech !

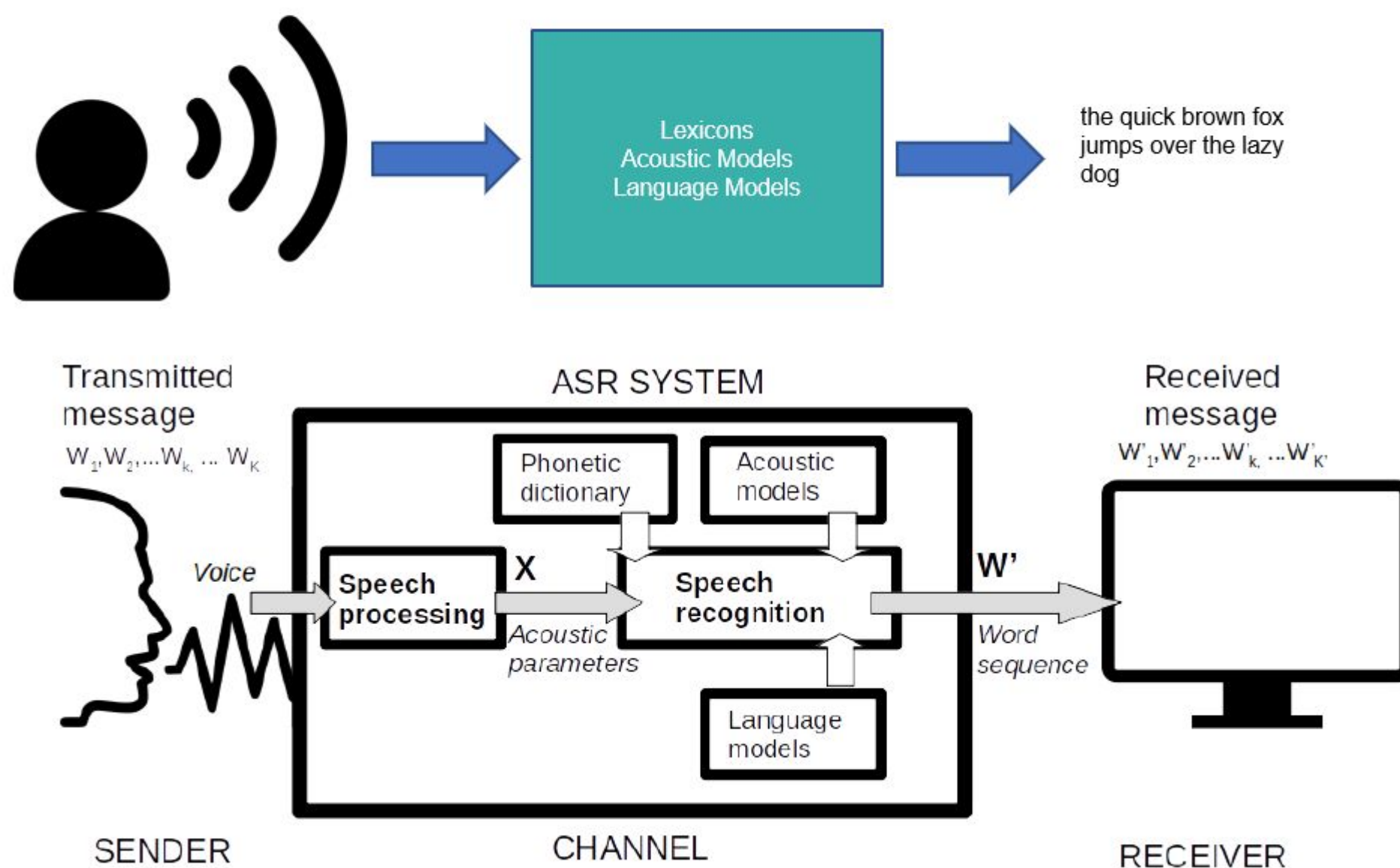
What makes speech special:

- Other modality, not text but audio
- Other conditions affecting quality of speech
 - Background (noises)
 - Number of speakers
 - Speaking context
 - Speaker accents
 - Vocabulary



And then there was automatic speech recognition!

Converts received audio (oral messages) into its textual form in the source language



Typed text ≠ spoken text

Typed text input	ASR text output
Grammatical (?)	Ungrammatical
Spelling errors	No spelling errors
Interpunction	No interpunction

Homed Objectives

- In the first step an existing “generic” Dutch ASR, used in the CLARIAH infrastructure, is adapted to the MedPharm domain on the *semantic level*, using a domain adaptation component (language model).
- In the second step the ASR is adapted on both the *semantic* and the *acoustic level* using sensitive inhouse data of Nivel, whereby the AV-recordings themselves will not leave the Nivel building.
- The resulting ASR component will be made available at Nivel and in the CLARIAH Infrastructure.
- The resulting acoustic models (AMs) and the language model (LM) of the recogniser (which are as such anonymous), and the metadata of the AV consultations at Nivel will be made available to the research community at large (open source).
- <https://pdi-ssh.nl/nl/home/>



Partners

- Dr Henk van den Heuvel is director of the Center for Language and Speech Technology (CLST) at Radboud University Nijmegen.
- Prof. dr Toine Pieters is professor of the History of Pharmacy and Allied Sciences at Utrecht University and senior fellow of the Descartes Institute of the History and Philosophy of the Sciences and the Humanities.
- Prof. dr Sandra van Dulmen is program leader at Nivel (Netherlands institute for health services research) and professor of Communication in Healthcare at Radboud university medical center as well as at the University of South-Eastern Norway. She primarily conducts research in the field of communication in e.g. pediatrics, oncology, primary care, pharmacy, home care, palliative care and medical genetics.
- Dr Arjan van Hessen is researcher Human Language Technology (HLT) at the University of Twente, head of imagination at Telecats, and director user involvement of CLARIAH. His main interest lies in bridging the gap between academic research and applications in the real world.
- Dr Roeland Ordelman is product manager of the research and innovation infrastructure at NISV and technical coordinator of the CLARIAH Media Suite.
- [Stichting Open Spraaktechnologie](#)
- [Care2Report](#) (U Utrecht, Nivel)
- [CaireLab](#) (AI LUMC)

Background

- Every year there are more than 15,000 hospital admissions due to avoidable **misuse of medicines** in the Netherlands. Often, this has to do with the patient's **unintentional improper use** caused by either hard to understand information or cognitive problems. This may result in rather diverse and inappropriate forms of use, low levels of adherence and waste of scarce financial resources.
- In order to overcome these misunderstandings we need to **better understand the explicit and implicit attribution of meaning to medicines** as part of the information processing.

ASR and Homed

What makes doctor patient conversations special:

- Conditions affecting quality of speech
 - Background (noises)
 - Number of speakers
 - Speaking context
 - Speaker accents
 - Vocabulary
 - **Privacy !**



Steps in Homed

- To deal with the specific medical/pharmaceutical vocabulary, we propose a methodology in which the recogniser's LM is adapted in a **first step** by additional training on text material from:
 - Healthbase
 - ICH: MedDRA
 - NHG
 - Nictiz
 - Zorginstituut Nederland
 - Medicijnjournaals - Instituut Verantwoord Medicijngebruik

Step 1: Experiment

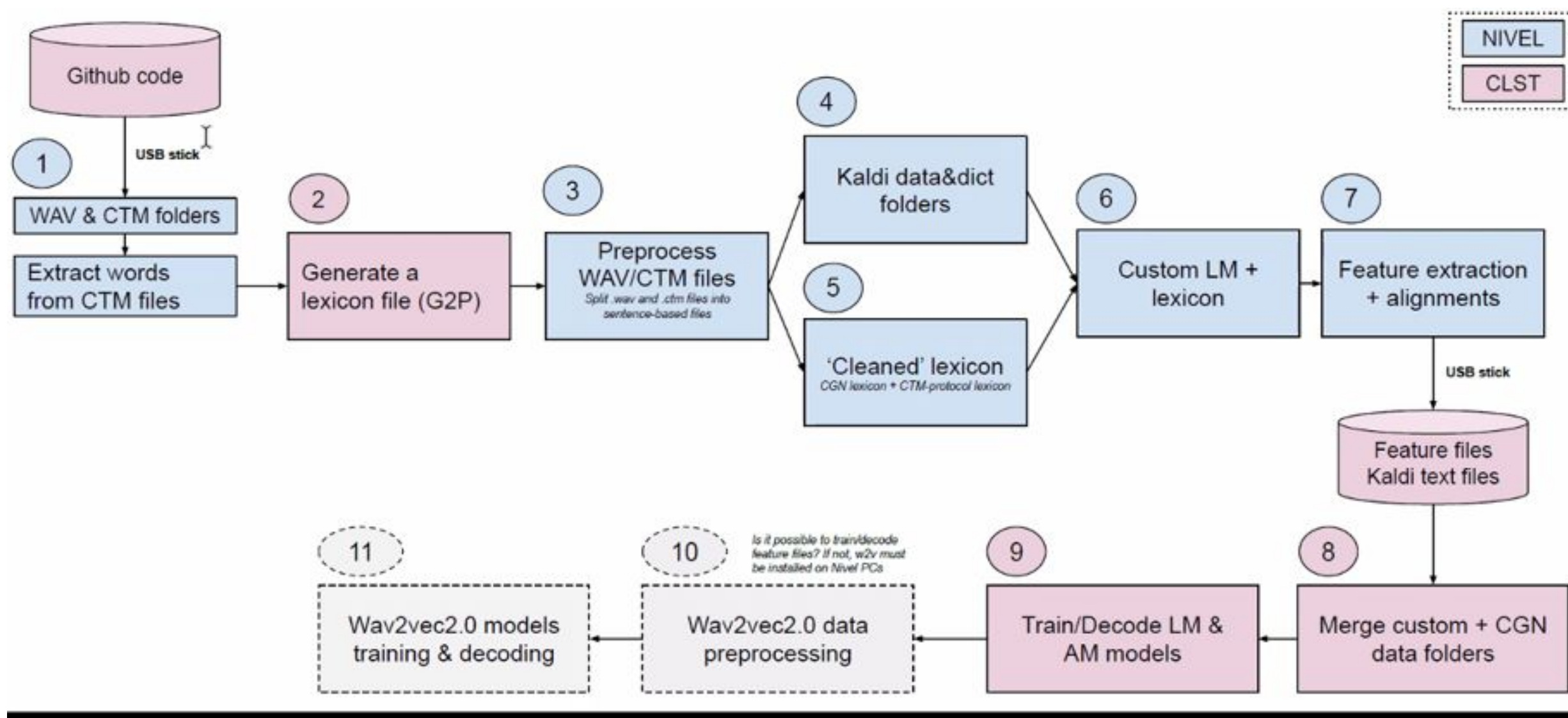
1. CGN: LM: General conversations, AM: Adult speech, Lexicon: 255000 tokens
 2. HoMedV1: LM: CGN+Lists of medical terms, AM: CGN, Lexicon: CGN + 13934 tokens
 3. HoMedV2: LM: CGN+Medicijnjournaal, AM: CGN, Lexicon: CGN + 5342 tokens
 4. Wav2Vec2: XLS-R-based CTC model with 5-gram language model from Open Subtitles tuned on CGN and the Mozilla Common Voice 8_0 NL dataset
- Test material: Medicijn journal (IVM): read out monologue news bulletin with medical terms

ASR system	WER
Kaldi_NL	25.8
HoMedV1	24.7
HoMedV2	20.6
Wav2Vec	19.2

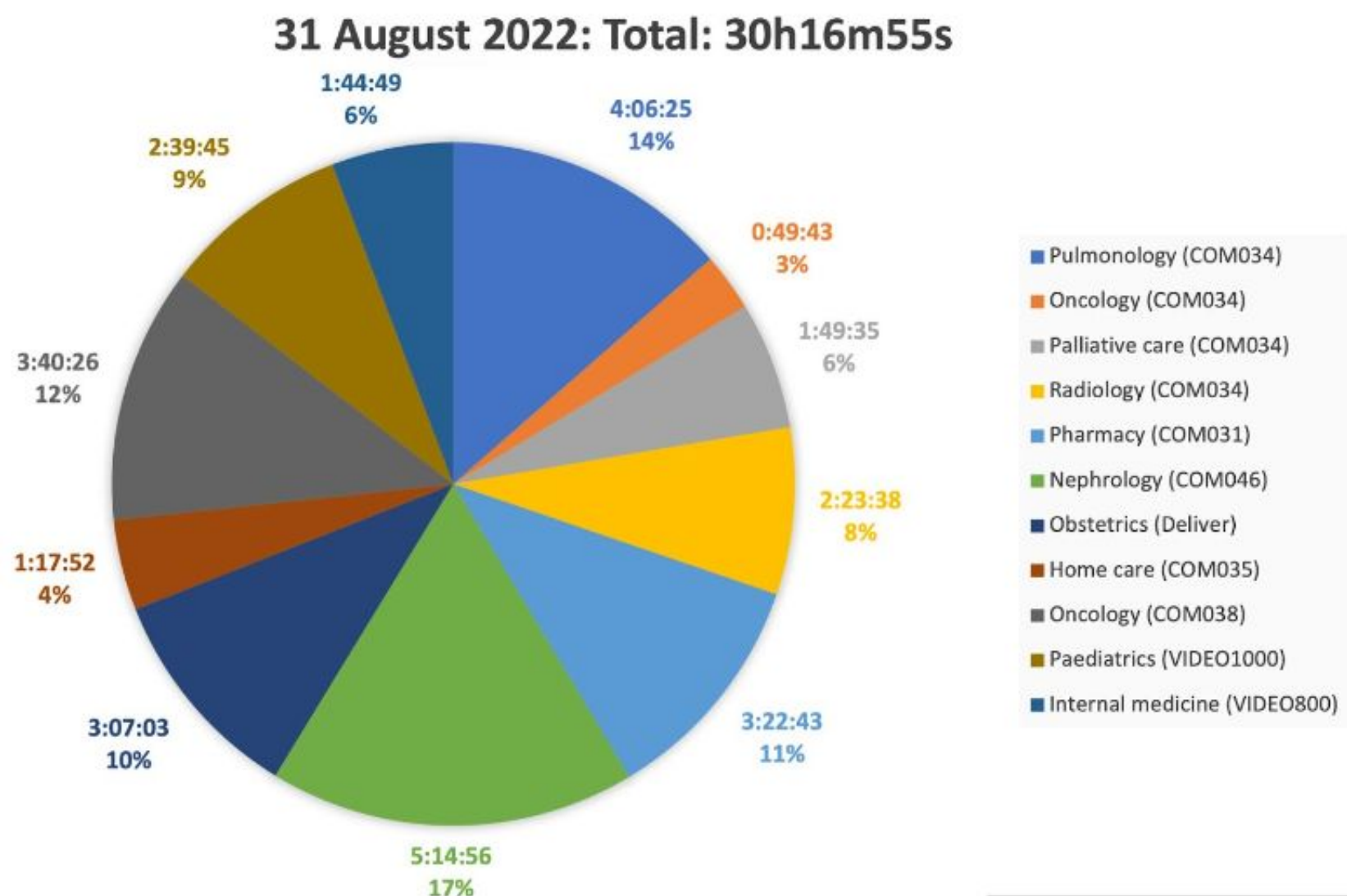
Steps in Homed

- In a **second step**, the resulting recogniser will be implemented in-house at Nivel and fine-tuned on +50 hours of transcribed real-life recordings of medical visits with patients with either COPD or cancer, and patient visits to the pharmacists.
 - Recordings may not leave Nivel
 - But what can?
 - Conundrum: make a specialized speech recognizer and training it without sensitive data leaving the Nivel building
 - After training/fine tuning the models are aggregates of speaking articulation behaviour which cannot be traced back to individual speakers

Step 2



Step 2



Expected Output

- A **methodology** for building special purpose speech recognisers for sensitive data domains built on evidence-based pilot projects
 - HoMed will already link other domains with similar questions and needs to the project by including the Dutch Police and Nederlandsche Bank into its Advisory Board so as to safeguard extensibility.
- **Special purpose ASR models** for the MedPharm domain which will be made available to the research community at large
 - the models will be distributed by the Stichting Open Spraaktechnologie
 - and used in other projects such as the Care2Report and CAIRE-lab projects
- **Anonymised metadata** of Nivel consultations
- **Access** to the domain-specific ASR via the CLARIAH Media Suite.
- An **inhouse** version of this ASR at Nivel.
- A methodology to deploy **ASR on premises** when sensitive data prohibit the use of ASR-as-a-Service via internet.
- Increased **awareness** for potential users inside and outside the academic arena of the possibilities of (co-)developing and using an ASR engine of their own even for sensitive and domain-specific AV data.

Publications

- Homed:

Van der Molen, Berrie, Tejedor-García, Cristian, van den Heuvel, Henk, Ordelman, Roeland, Pieters, Toine, van Dulmen, Sandra, & van Hessen, Arjan. (2022, May 4). *Challenges on the Promising Road to Automatic Speech Recognition of Privacy-Sensitive Dutch Doctor-Patient Consultation Recordings*. DH Benelux 2022 - ReMIX: Creation and alteration in DH (hybrid), Belval Campus, Esch-sur-Alzette, Luxembourg and online. <https://doi.org/10.5281/zenodo.6517157>

Tejedor-García, C., Van der Molen, B., Van den Heuvel, H., Van Hessen, A., Pieters, T. (2022). *Towards an Open-Source Dutch Speech Recognition System for the Healthcare Domain*. Proceedings of the 13th International Conference on Language Resources and Evaluation (LREC2022), pp. 1032-1039.

<http://www.lrec-conf.org/proceedings/lrec2022/pdf/2022.lrec-1.110.pdf>

The website of the project is: <http://homed.ruhosting.nl/>

- Other initiatives:

Van Buchem, M.M., Neve, O.M., Kant, I.M.J. *et al.* Analyzing patient experiences using natural language processing: development and validation of the artificial intelligence patient reported experience measure (AI-PREM). *BMC Med Inform Decis Mak* **22**, 183 (2022). <https://doi.org/10.1186/s12911-022-01923-5>

Brinkkemper, S. (2022). Reducing the Administrative Burden in Healthcare: Speech and Action Recognition for Automated Medical Reporting. In R. A. Buchmann, G. C. Silaghi, D. Bufnea, V. Niculescu, G. Czibula, C. Barry, M. Lang, H. Linger, & C. Schneider (Eds.), *Information Systems Development: Artificial Intelligence for Information Systems Development and Operations (ISD2022 Proceedings)*. Cluj-Napoca, Romania: Babeş-Bolyai University.