Lingsoft and Speech recognition solutions

ELEC-E5500 - Speech Processing guest lecture
October 9, 2023

Tiina Lindh-Knuutila, solution architect
tiina.lindh-knuutila@lingsoft.fi

Lingsoft®

SUOMALAISTA PALVELUA

Today’s agenda

Introduction

Lingsoft’s solutions

Speech recognition and its uses

Human-in-the loop use case: Live subtitling

Time for questions

Lingsoft®

Lingsoft in a nutshell

The largest privately owned Finnish language management company

- Founded in 1986
- More than 100 employees
- More than 500 partners

The only Finnish company among the 100 largest language management companies in the world

- Over 1000 end customers
- Our language technology solutions have tens of millions of end users

Language services: translations, subtitling, transcription, medical transcription, term management, content production

Language solutions: accessibility and clarity, findability and text analysis, semantic compatibility

SUOMALAISTA PALVELUA
FINLÄNDSK SERVICE
®

Lingsoft Core Technologies

Text Analysis

- Hybrid rule-based/deep learning-based
- Lemmatization, disambiguation, named entity recognition, semantic enrichment
- Spelling and grammar checker

Automatic Speech Recognition (ASR)

- Based on deep learning
- Language and domain customization
- Pre- and postprocessing

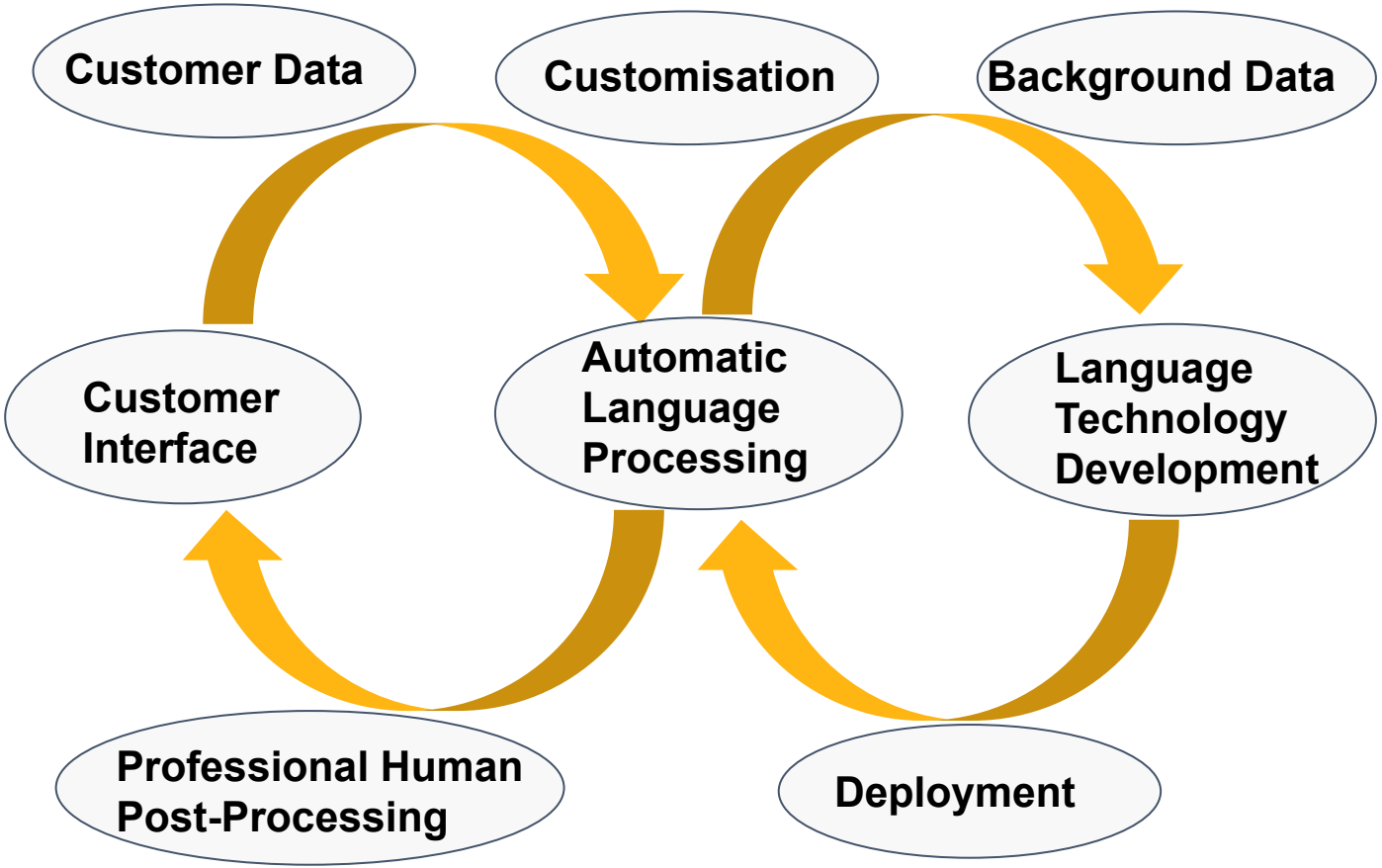
Neural Machine Translation (NMT)

- Based on deep learning
- Language and domain customization
- Pre- and postprocessing

Process Automation and Integration

- Tools available primarily as cloud services
- APIs based on industry standards

Lingsoft Process Design Model



Other Lingsoft solutions

Lingsoft Text Analysis

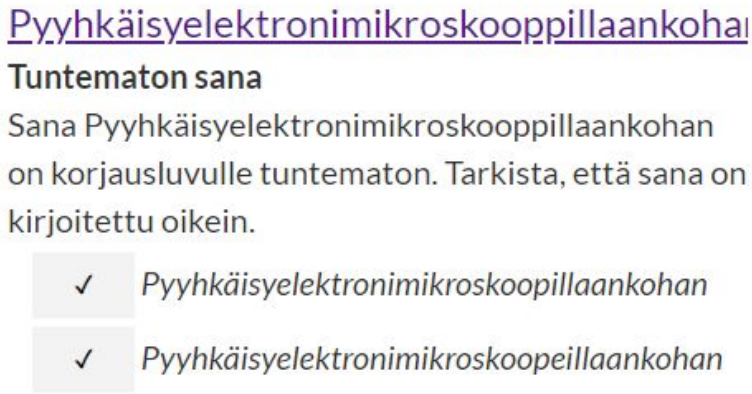
- **Morphological analysis**: fundamentals for the rest, e.g. lemma (dictionary form) identification and disambiguation
- **Proofing tools**: spelling and grammar checking
- **Terminology mark-up**: detect and mark-up terminology in running text
- **Named Entity Recognition**: detect names and related entities in text, important for e.g. **anonymization, keywords** and **metadata**
- An important part of speech recognition post processing

Proofing Tools

Pyyhkäisyelektronimikroskoopillaankohan

Pyyhkäisyelektronimikroskooppillaankohan

maybe with their scanning electron microscope



Supported languages:

- Finnish (incl. general and medical domain)
- Swedish (incl. standard and Finland Swedish)
- Norwegian (bokmål and nynorsk)
- Some support also for a number of other languages, e.g. English, Russian, German, Northern Sami, …

Terminology and Keywords

Valtiontalouden tarkastusvirasto valvoi vuoden 2022 aluevaaleissa ehdokkaan vaalirahoituksen ja vaalikampanjakulujen ilmoitusvelvollisuuden noudattamista laissa säädetyllä tavalla. Kaikki ilmoitusvelvolliset jättivät vaalirahoitusilmoituksen.

ilmoitusvelvollinen (vtv)

sv: redovisningsskyldig
en: discloser

Given a termbase or ontology, then terms/concepts can be identified and marked-up in running text

Lingsoft Machine Translation

Automatically generates translations for text

Based on **open source neural machine translation** (NMT) tools (same as behind Microsoft Translator)

Trained with **tens of millions of segments** from internal and public data sets (e.g. DGT TM)

Used by **hundreds of translators** via MTLab - our plugin for Trados Studio

Translates about **two million words every month** for Lingsoft post-editing jobs

Terminology Control in NMT

Henkilötunnus = Social security number? Personal identification number? Personal identity code? …?

source + term list + NMT engine = machine translation with the right terminology

Source	**termNMT**
<SRCTERM> henkilötunnuksen <TRANS> personal identity code </TRANS> saaminen edellyttää henkilökohtaista käyntiä <SRCTERM> palvelupaikassamme <TRANS> service location </TRANS> .	Obtaining a personal identity code requires a personal visit to our service location.

Lingsoft speech solutions

Lingsoft Speech Service

Our platform for serving speech recognition clients

Speech-to-text or automatic speech recognition (ASR)

- Automatically transcribes speech
- Based on **open source** tools
- Trained with a few **thousands of hours of** transcribed **speech plus hundreds of millions of sentences**
- Near **real-time** transcription and live subtitling
- Used by **Yle, The Finnish Institute for Health and Welfare** and **the Finnish Parliament** (Eduskunta)

Tailored post-processing capabilities available

Features

Supports file-based / real-time streams / voice commands

Industry standard APIs and two UIs

- Transcription, subtitling, speaker segmentation
- Helper APIs for data I/O, statistics

Speaker-independent, requires no individual speaker training process

General language models, as well as customer/domain-specific language models development

- Supported languages
- Finnish: general, medical, politics
- Swedish: general
- English: general (pilot use)

Secure cloud connections (based in Finland), on-premise possible

Domain and Customer Adaptation

Problem

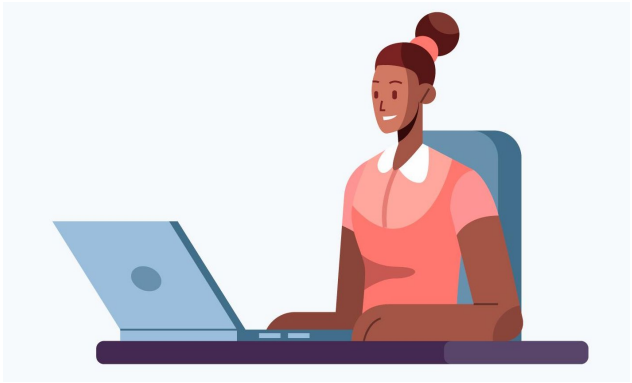
Words that were not seen in training are poorly recognised

Solution 1

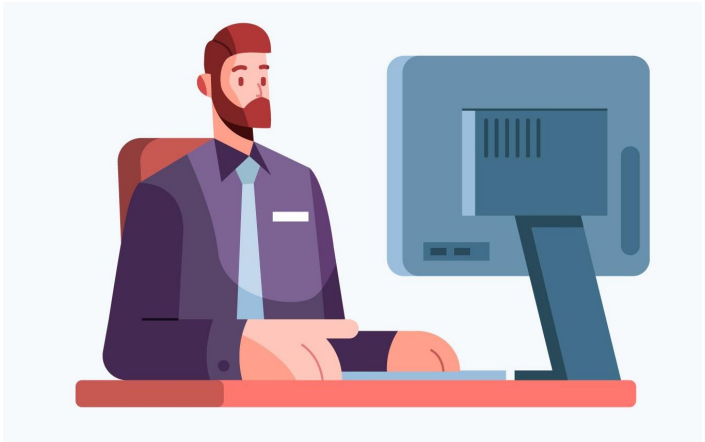
Adapt the model a bit more with text from the customer. The more text with the right content, the better the result. Works well for applications where the customer has a lot of pre-existing text with correct terminology, e.g. pathology.

Solution 2

Collect sentences with words that are poorly recognised that you believe might be useful in the future, e.g. names of (in)famous people. Then: re-train.



Who are our users?



Designed by Freepik

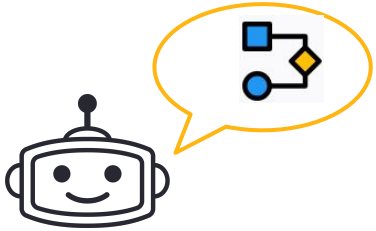
Example use case

Use case: From speech to subtitles - live

Good afternoon
and welcome!

?

~2 seconds



The broadcast and speaking begins

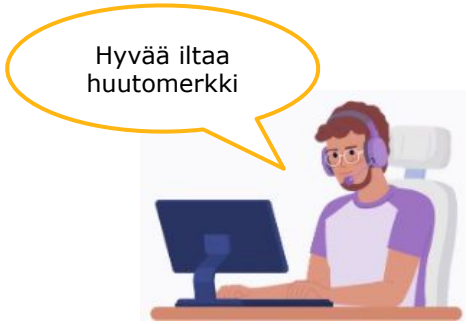
Subtitler respeaks

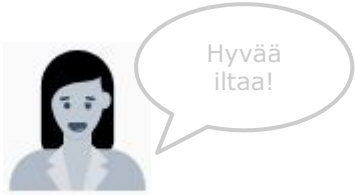
Speech recognition

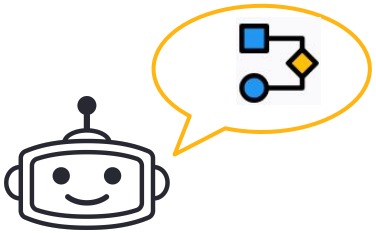
Subtitling software

Subtitles embedded to broadcast

Live subtitled TV or internet broadcast







The broadcast and speaking begins

Subtitler respeaks

Speech recognition

Subtitling software

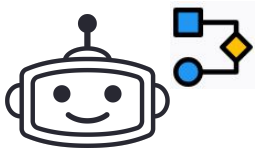
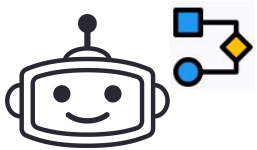
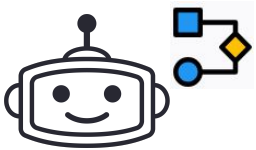
Subtitles embedded to broadcast

Live subtitled TV or internet broadcast

Hyvää huomenta ja tervetuloa Aalto-yliopiston puheen käsittelyn kurssille 9.10.2023!

Good morning and welcome to the Speech Processing course of Aalto University on October 9, 2023!

From speech to subtitles - live

Speech to text | Post-processing | Voice commands

Hyvää huomenta ja tervetuloa **aaltoyliopiston** puheen käsittelyn kurssille **yhdeksäs kymmenettä kaksituhattakaksikymmentä kolme** huutomerkki

Hyvää huomenta ja tervetuloa **Aalto-yliopiston** puheen käsittelyn kurssille **9.10.2023** huutomerkki

Hyvää huomenta ja tervetuloa Aalto-yliopiston puheen käsittelyn kurssille 9.10.2023**!**

Speech models

Text analysis and tailored rules

Voice commands to edit text

Speech to text

Post-processing

Voice commands

Large amounts of speech and even larger amounts of text needed for training the models

[aaltoyliopisto] > Aalto-yliopisto
[yhdeksäs kymmenettä kaksituhattakaksikymmentä kolme] > 9.10.2023

“Repla miksi kysymysmerkki”
-Miksi?

The goal: quickly readable text

*”Äsdeepee **kaksikymmentäneljä pilkku viisi** lisäystä **yksi pilkku yksi**.”*
➔ **SDP 24,5** lisäystä **1,1**.

*”**Vuonna seitsemänkymmentäkuusi** koneisto alkoi yskiä.”*
➔ Vuonna **1976** koneisto alkoi yskiä.

*”**Hoohetki** oli **kello seitsemän kolmekymmentä**.”*
➔ **H-hetki** oli **kello 7.30**.

Examples of acronym, number and unit shrinking

But preferred format depends on the use case

Subtitling

1500 kilometrin laajuisessa Rheassa näkyy muinainen pinta.
OR?
1 500 km:n laajuisessa Rheassa näkyy muinainen pinta

Medical transcription

RR 145/84 **mmHg**
OR?
Verenpaine 145/84 **elohopeamillimetriä**

Even more helpful tools for video files

Legibility can be automatically improved when not working in a live broadcast:

Automated recognition for **sentence borders** and **speaker changes**

Automatic recognition of **subclauses** as cut points for **splitting longer sentences**

Automatic recognition for **multi-word terms and collocations** that should not be split: the White House, president Niinistö

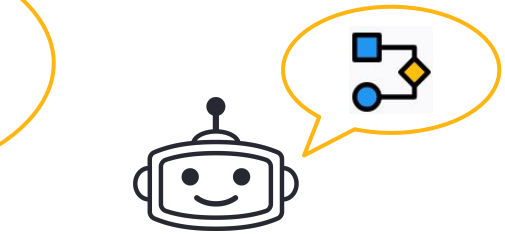
Automatic punctuation

Adaptable and developing system

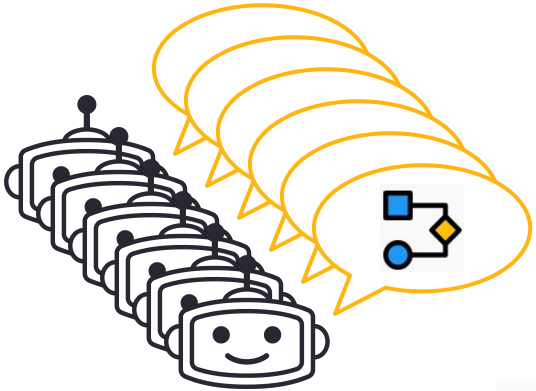
- Regularly trained for customer use with relevant and quality material
  - Vocabulary changes over time
- Models can be adapted for each organization if required
  - Customers can add material to the system and adapt their own model
- Customer can add individual words
  - pronunciation guidelines
  - examples of sentence context
  - Proper noun or not?
- Human in the loop - also in training

Future of live accessibility solutions?

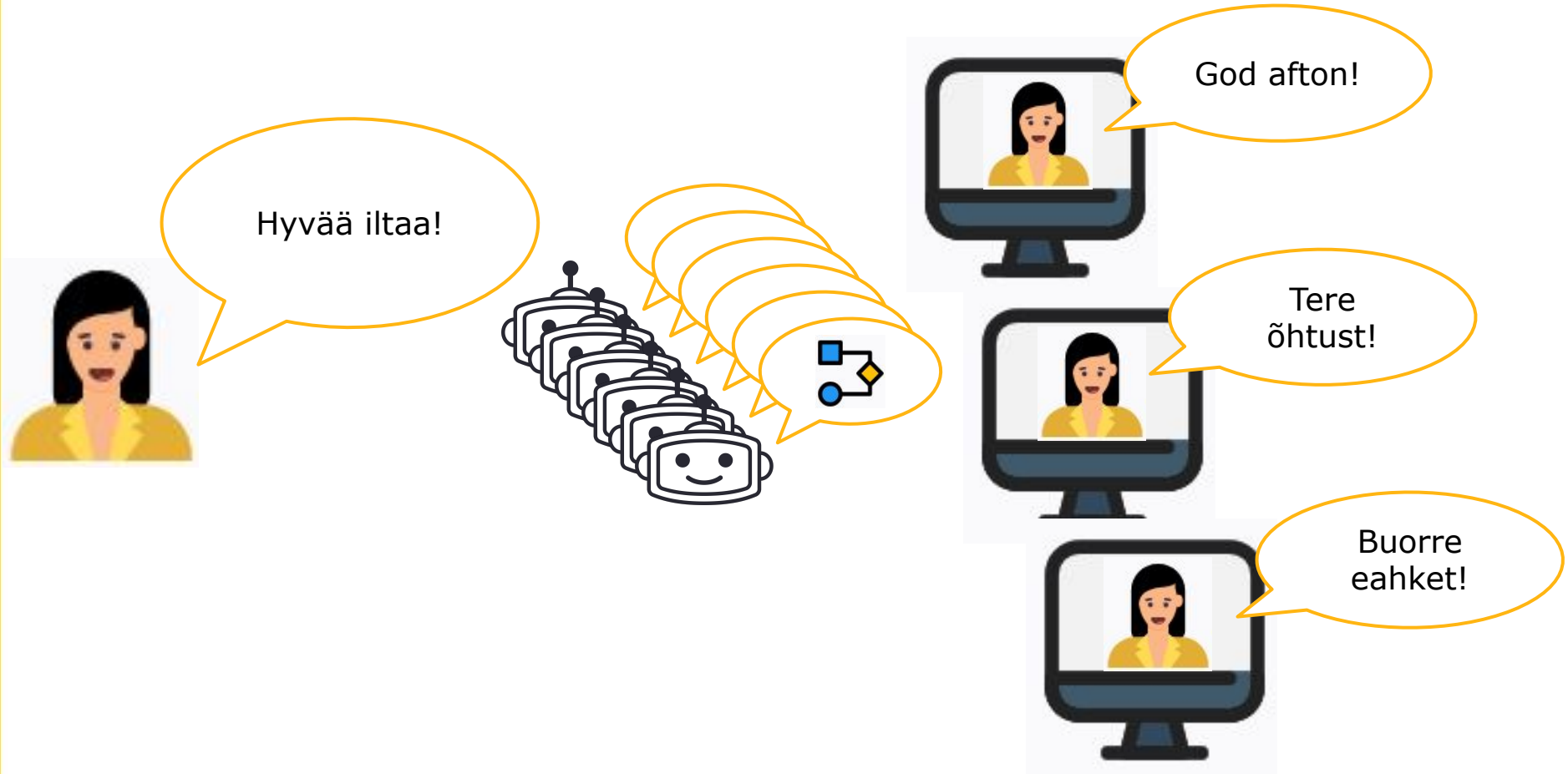
Live subtitling on all platforms



Multilingual subtitling



Multilingual speech to speech



Thank you!

www.lingsoft.fi

Lingsoft®