

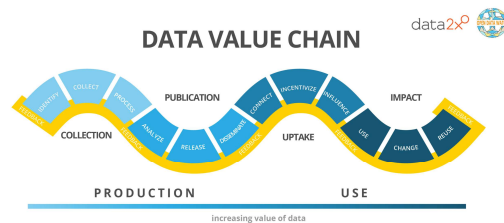
PRIVACY IMPLICATIONS OF SPEECH DATA PROCESSING

A.L. MAGNÚSDÓTTIR, DR. S.E. LENSINK

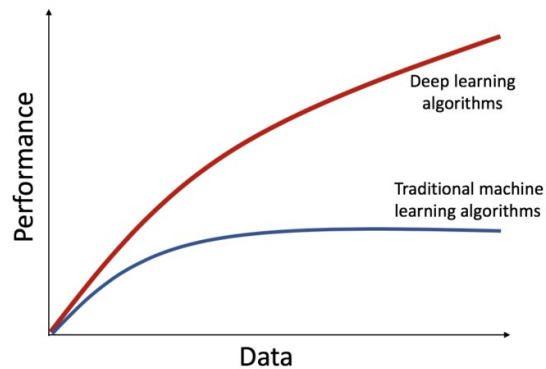
AI MODELS

POSSIBILITIES AND CHALLENGES

- › Huge potential in AI solutions



- › AI models, especially deep learning models, are very data-hungry

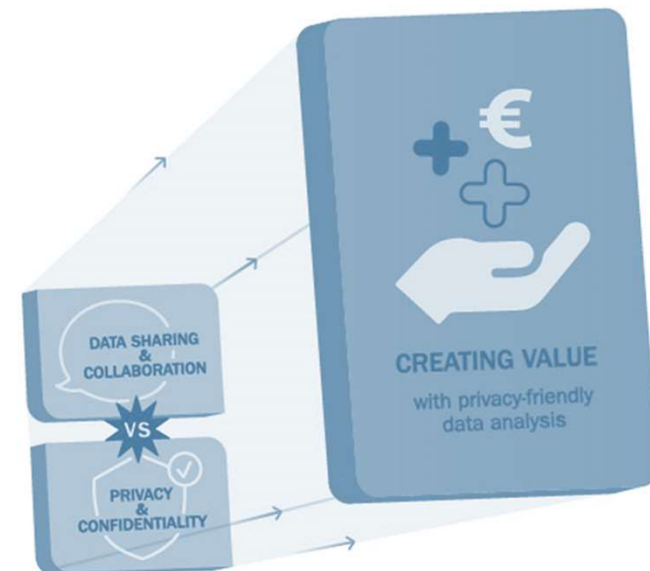


- › Data is often available, but cannot be shared and combined across organizations or even departments due to privacy and IP concerns



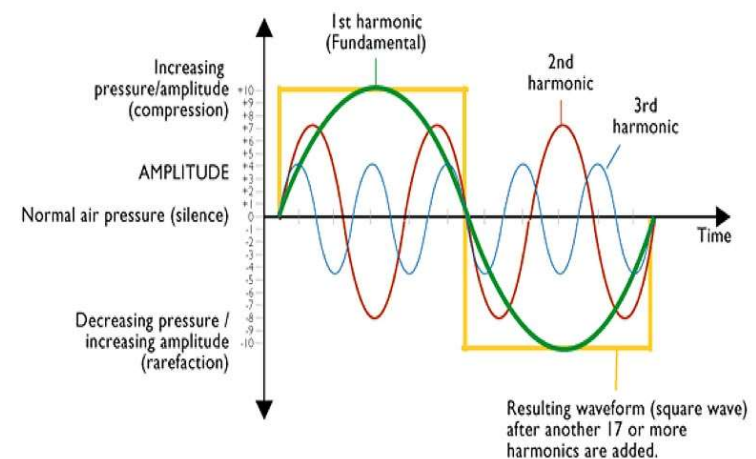
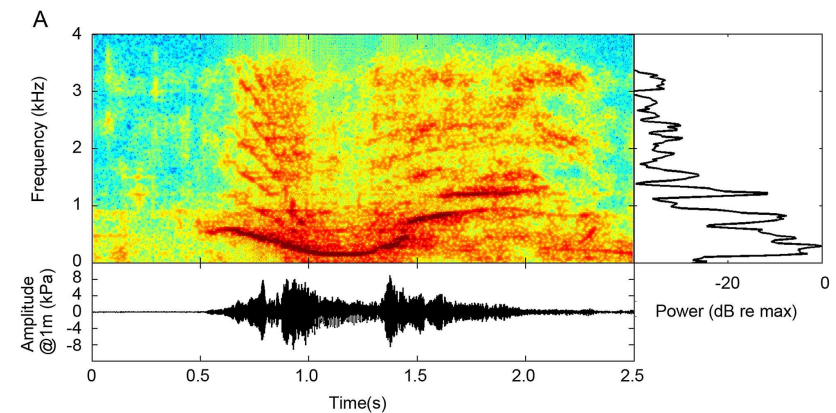
PRIVACY ENHANCING TECHNOLOGIES

- › Privacy Enhancing Technologies (PET) offer ways to efficiently and safely learn models on distributed datasets
- › TNO has broad expertise in secure multi-party computation (MPC), Federated Learning (FL), and synthetic data generation (SDG)
- › Speech data comes with its own characteristics – which PET to use?



› SPEECH DATA

- › Speech + its transcriptions can be used to learn powerful Automatic Speech Recognition (ASR) systems
- › Unstructured data
- › Time-series data
- › Many layers of data in speech signal
 - › Wavelength
 - › Amplitude
 - › Time period
 - › Frequency (Hz)
 - › Velocity



› FEATURES THAT CAN BE EXTRACTED FROM SPEECH DATA

- › Body measure (height and weight)
- › Mood and emotional state
- › Age
- › Gender
- › Personality trait, particularly the “Big Five” – openness, conscientiousness, extroversion, agreeableness and neuroticism.
- › Deception, are you trying to lie
- › Sleepiness
- › Intoxication
- › Accent and dialect indicate geographical origin
- › Health, relating to the vocal cord and beyond such as Alzheimer’s and Parkinson’s
- › Mental illnesses such as schizophrenia and severe depression
- › Interpersonal perception, that is how are they perceived by other people (relating to personality trait, for example fast talkers are considered more competent)
- › Socioeconomic status such as education
- › Acoustic scenes and events, such as the location of the speaker
- › Biometric identity, speakers can be uniquely identified by a voice sample
- › ...

Kröger, J. L., Lutz, O. H. M., & Raschke, P. (2020). Privacy implications of voice and speech analysis—information disclosure by inference. In *IFIP International Summer School on Privacy and Identity Management* (pp. 242-258). Springer, Cham.



› THREE TYPES OF SENSITIVE FEATURES

Who, where, when - metadata

- Date/time
- Location
- Names of speakers
- ...

How - Acoustic characteristics speakers

- Gender
- Age
- Native language
- Emotional state
- Health status (dementia, Parkinsons, covid-19...)
- Deception
- ...

What (Content) - Intellectual property

- Content of television show
- Speech/talk/class
- Discussion of medical information
- Police hearing
- Court case
- ...

› WHICH ASPECTS MATTER MOST TO YOU?

Questionnaire:

- › What sensitive features are present in the speech data of your organization?
- › What needs to be protected most? Metadata-level, content-level, or acoustic-level features?
- › Is there a prioritization between the different sensitive features?
- › Do you currently have an AI solution that was learned on your speech data, and if yes, what is its performance?
- › When gain in performance for a speech model is your organization aiming for?

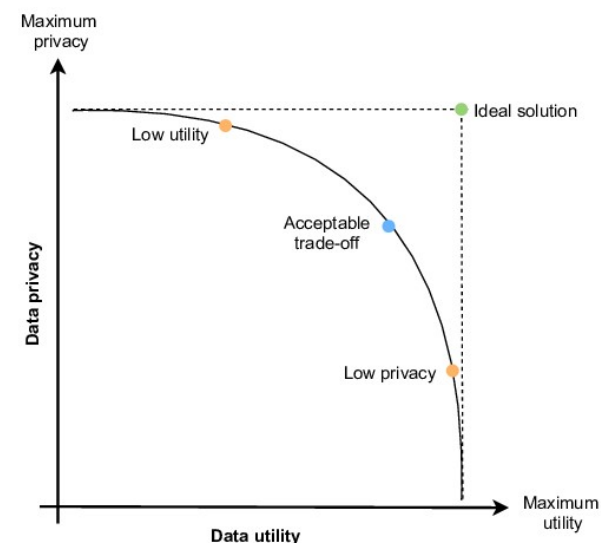


Fig 1 | Privacy-utility trade-off – the more you protect, the more difficult it will be to learn a high-performing AI model



› **BEDANKT VOOR**
UW AANDACHT

TNO innovation
for life