Names
Participants

- Mark Davis
- Souheil Ben Yacoub
- Richard Ishida
- Doug Lawrence
- Gary Lefman
- Christian Lieske
- Juan Pane
- Kerstin Steffen
- Feiyu Xu
Types of names

• Personal names:
  – Full name
  – Alternative names
Use cases

• Recognition
  – NER, Machine translation, business intelligence, search, identity resolution, segmentation

• Display
  – Sorting, contextual usage: Short, Informal, Full name (formal, postal name), inflexions, auto-completion, segmentation, automatic abbreviation, text-to-speech

• Capturing
  – Transliteration, speech-to-text, input-form-input
Use cases

• Business intelligence: Recognize the name in the text
  – NER, NED (identity resolution)

• Display: generating names
  – Short, Informal, Full name (formal, postal name), inflexion

• Segmentation of names (line-wrapping)

• Matching

• Auto-completion

• Translation, transliteration (between alphabets, using same alphabet)

• Sorting (given names, last names)

• Text-to-speech
Problems

• Input, capturing, forms (keep the context to be able to reproduce the sound again), representation (endonym, exonym, alternative names, preferred, ...)

• Display
  – Familiarity, formality, context, inflexion (language dependent)

• Recognition
  – NED, NER, matching, normalization, canonization

• Search (input + recognition)

• Aliases, alternative names, abbreviations, preferred names
Information out of the names

• Structure of the name;
  – What are the semantics of each component
    • Gender,
    • Origin (chinese, roman, islandic)
Problem

• Sorting:
  – How to recognize the tokens, and use them to sort

• Segmentation:
  – how to recognized suitable break points

• Alternative names: recognition, capturing

• Component names:

• Usage: formal, informal
Possible approaches

• Name structure is language (cultural?) dependent
  – Changes in time
    • German, when getting a university degree
    • Spanish, when marrying