Multi-Language & Multi-Domain

- Goal:
  adapting an existing and running system to a new Language and/or to a new Domain

Language spotting

- how do we know the language?
- how to do it in real-time?
- running a large number of recognizers at the same time might not be an option...

Language challenges & issues

- getting data
- avoid getting lost in translation
  - different words in one language translate to the same in other language
  - code-switching
  - not just ASR & TTS
  - different styles at multi-levels
    culture, politeness, friendliness, …
  - spoken dialogue systems might not be appropriate for the particular culture or might require different dialect

How to deal with Language issues?

- collect data
- awareness: understanding the problem
  - people talking the language locally – accent
  - knowing the target culture and the real people
  - describing the community
    from an external point of view
    (better if it is someone exposed to different cultures)
- test with the actual community

Domain spotting

- how do we know the domain in use?
- do keywords help?
  - are they sufficient?
- domain definition: how to define it?
- are task, topic and domain the same thing?

Domain challenges & issues

- getting data... again!
- there are some new words and dialects: ASR & TTS
- systems to the elder are different from those for young students
- changes in the core system
  - new things that were never tested
  - things that no one has thought about
- do we want multi-domain at the same time?
  - buying a ticket, change to home control, and then get back to the ticket we were buying...

Dynamic domains

- can we learn online from what the user is saying about a new domain?
- can we learn Language dynamically as well?
- can we have systems that automatically learn a completely different new task?

How to deal with Domain issues?

- databases and ontologies
  and the related translation
- collecting speech and language model data:
  most have to be updated
- task-related issues:
  the domain model, and things that don't make sense in the new domain...