Speech Technology Research
at ICSI

Arlo Faria
ICSI

• Independent non-profit research institute

• Groups:
  • Algorithms
  • Architecture
  • Artificial Intelligence
  • Networking & Security
  • Speech
  • Vision

• Open House: 1947 Center Street, 6th floor
The Speech Group

• **Group leader:** Steven Wegmann
• **Faculty advisor:** Prof. Nelson Morgan
• **Research staff:** Jordan Cohen, Dan Ellis, Jim Hieronymus, Adam Janin, Nikki Mirghafori, Roberto Pieraccini, Liz Shriberg, Andreas Stolcke, Chuck Wooters
• **Post-docs:** Howard Lei, Hari Parthasarathi
• **Ph.D students:** Shuo-Yiin Chang, Arlo Faria, Mary Knox, Suman Ravuri, TJ Tsai, Oriol Vinyals
• **Visitors:** Hai Do, Korbinian Riedhammer, Mirco Ravanelli
Research areas

- Speech recognition
  - also: keyword search
- Speaker recognition
  - ID, verification
- Multimedia processing
  - e.g. Web videos
Speech Recognition

- Diagnosing flawed HMM assumptions
- Acoustic features for noisy data
- Revisiting “deep” neural networks
- Portability to low-resource languages
Speaker Recognition

- Robustness
  - to noise (AFRL)
  - to signal degradation (DARPA)
- Speaker Diarization (ParLab)
Other research areas

• Speech activity detection
• Language identification
• Conversational, dialog systems
• Models based on measured brain activity
  - auditory cortex of ferrets (UCLA)
  - neurosurgery patients (UCSF)
Scientific Goals

- Robust acoustic processing
- Better understanding of principles
- Significant (non-incremental) progress
- http://www.icsi.berkeley.edu