LORELEI: Low Resource Language Indexing
Sponsor: IARPA
Consortium Lead: IBM
Partners: Columbia University, City University of New York, Cambridge University
RWTH Aachen University, New York University, and University of Southern California

We propose to achieve the Babel program goals with a three-pronged approach: we will develop language-independent architectures and systems for Reliable Keyword Search; we will push the boundaries on Learning with Limited Labeled Training Data; and we will greatly enhance the Robustness to Noise, Channel, and Speaker Variability of the underlying speech recognition systems for keyword search.

The work on these three thrusts correspond to the three major tasks of our proposal. The work will cut across four classes of technologies. We will enhance our core speech recognition systems and search capabilities; we will develop a novel framework for language-independent speech recognition; we will develop new ways of exploiting linguistic information; and we will develop novel machine learning techniques specially for the program goals.