

Opportunity Title: Non-traditional annotation of realistic speech

Opportunity Reference Code: IC-16-33

Organization Office of the Director of National Intelligence (ODNI)

Reference Code IC-16-33

How to Apply **Create and release your Profile on Zintellect** – Postdoctoral applicants must create an account and complete a profile in the on-line application system. **Please note: your resume/CV may not exceed 2 pages.**

Complete your application – Enter the rest of the information required for the IC Postdoc Program Research Opportunity. The application itself contains detailed instructions for each one of these components: availability, citizenship, transcripts, dissertation abstract, publication and presentation plan, and information about your Research Advisor co-applicant.

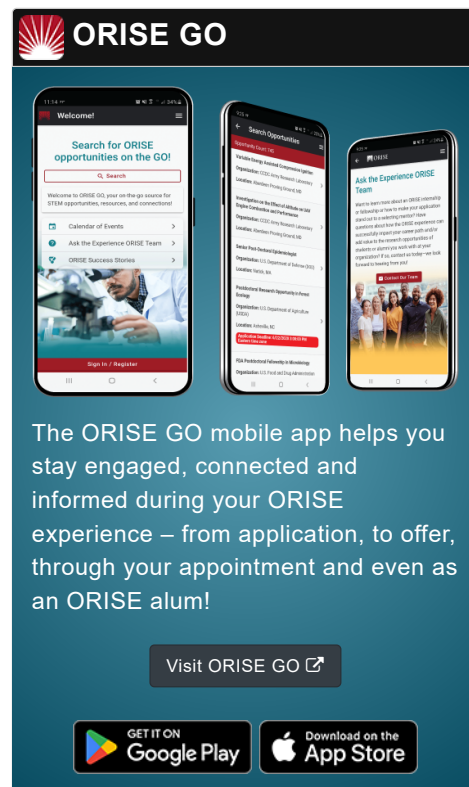
Application Deadline 4/15/2016 6:00:00 PM Eastern Time Zone

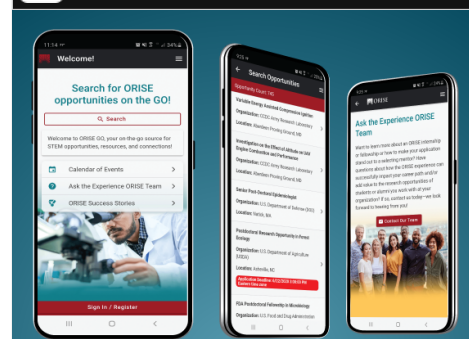
Description Annotated corpora for training, development, and evaluation are the backbone of human language technology (HLT) research. Traditional annotation tasks used to create speech recognition corpora require annotators with sufficient language knowledge to apply well-defined labels to speech data. These tasks can include labeling the regions in an audio file that contain speech, labeling the language(s) of the speaker(s), and transcribing the words that are spoken. Labels can be applied to an entire file, the individual speech utterances of the file, or each token in a speech utterance.

Consequently, many areas of speech recognition research are hindered by the lack of realistic data with necessary labels because the required labels are non-traditional and difficult to define. Examples of non-traditional labels include prosody (e.g. loudness, rhythm) and emotion (e.g. anger, happiness). The imprecise variations in these non-traditional labels are the source of the labeling difficulty, e.g. when speech loudness is considered, a set of labels could include [very quiet, quiet, normal, loud, very loud, extremely loud], but defining these labels and the boundaries between them is challenging.

Organizations that manage traditional annotation rely on documented guidelines that define the labels and provide direction for proper annotation with them. These guidelines can be used to determine qualifications for annotators, train recruited annotators, and ensure the resulting corpora are accurate and useful for the intended research. However, when the labels are not well-defined, it is difficult to write sufficient annotation guidelines to describe how to label the data, and subsequently determine annotator qualifications; hence the resulting corpus will not adequately support the intended research.









The ORISE GO mobile app helps you stay engaged, connected and informed during your ORISE experience – from application, to offer, through your appointment and even as an ORISE alum!

[Visit ORISE GO](#)





Opportunity Title: Non-traditional annotation of realistic speech

Opportunity Reference Code: IC-16-33

One solution to the non-traditional labeling problem is to create speech corpora in which speakers are directed to produce speech with the required labels, often by reading a prepared script or reciting meaningless phrases. However, the unrealistic nature of the resulting data is not satisfactory for researchers trying to solve real problems. Solving real problems requires speech from real environments, with speakers whose intent is to convey information to real listeners.

The goal of this project is to design and implement an annotation process to label speech data with non-traditional, less-defined labels, both as a standalone task and to incorporate it within an on-going traditional annotation task, e.g. along with transcription.

Example Approaches:






One possible approach to this project would be to divide it into three phases:

- Phase one could catalog all existing speech corpora containing non-traditional annotations of any kind and to evaluate their realism.
- Phase two could design a repeatable, language-independent annotation process that overcomes the deficiencies and includes the merits identified in the corpora evaluation of phase one.
- Phase three could use the process developed in phase two to annotate an existing speech corpus (to be defined by the IC adviser) with a non-traditional label.

A straightforward example process is to annotate the presence of a speech characteristic in each speech utterance, i.e. for each utterance, the annotator must answer the question “Does this utterance sound sarcastic?” This process could be combined with a transcription task, or as a follow-on to the transcription task. Transcriptions could also be first annotated without the voice data, then paired with the voice data in a second round.









An example combining both human and technical solutions is to ask an annotator to compare segments against one another and label which is more like a predefined characteristic, e.g. “Which segment is louder?” Results from the comparison can then be ordered and labeled based on mathematical analysis.

Eligibility Requirements

- **Citizenship:** U.S. Citizen Only
- **Degree:** Doctoral Degree.
- **Discipline(s):**
 - **Business** (11 )
 - **Chemistry and Materials Sciences** (12 )
 - **Communications and Graphics Design** (6 )
 - **Computer, Information, and Data Sciences** (16 )
 - **Earth and Geosciences** (21 )

Opportunity Title: Non-traditional annotation of realistic speech

Opportunity Reference Code: IC-16-33

- **Engineering** (27 )
- **Environmental and Marine Sciences** (14 )
- **Life Health and Medical Sciences** (45 )
- **Mathematics and Statistics** (10 )
- **Other Non-Science & Engineering** (13 )
- **Physics** (16 )
- **Science & Engineering-related** (1 )
- **Social and Behavioral Sciences** (28 )