



THE UNIVERSITY OF
CHICAGO
BIOLOGICAL
SCIENCES

Postdoctoral (Scholar or Fellow) Position Open in the Thirty Million Words Center for Early Learning + Public Health.

Areas of expertise: Computer Science, Economics, Statistics, Computational Linguistics, Bioinformatics, Cognitive Science

Start date: Flexible, ideally by Fall 2021

Application Deadline: April 1, 2021

The TMW Center for Early Learning + Public Health is a joint venture between The University of Chicago Biological Sciences Division and the Division of the Social Sciences. It implements and evaluates evidence-based interventions designed to promote young children's cognitive and social-emotional development, with a priority on families living in poverty. To that end, the TMW Center has developed an expertise in the analysis of children's early experiences and environments. Using innovative technological tools and natural language processing methods, the Center collects and analyzes thousands of hours of audio and video recordings of parent-child interactions. Additionally, the Tech team is currently developing a new wearable technology to better measure the properties of children's natural linguistic environments and assess the impact of parent talk on brain development. Along with those unique data, the Center also collects extensive sociodemographic data, parental belief data and assessments of children's skills, allowing for a broad range of analyses of early disparities and their drivers.

The role of the Postdoctoral scholar in the TMW Center will be twofold. It will first consist in building a research agenda that leverages new advances in machine learning areas such as Natural language processing, Sentiment analysis, Gesture, Emotion and Activity recognition to explore the audio and video recordings of parent-child interactions. The definition of this new agenda will entail close collaboration with the Research Analysis team - an interdisciplinary group of researchers from Developmental Psychology, Education, Medicine, Economics, and Computer Science – and will be supervised by Dr. Dana Suskind and Dr. John List. The Postdoctoral scholar will also be in charge of developing state-of-the-art processing tools to

characterize the variability of parenting styles and children's experiences and relate it to the variability of early trajectories. The position requires strong skills in machine learning and deep learning methods, application, and architecture. Familiarity with classification and prediction models is expected. It is anticipated that this postdoctoral opportunity will necessitate at least two years to garner successful proficiency.

Responsibilities

- 1) Develop possible solutions using standard and custom procedures to prepare and analyze the audio and video data with a moderate level of direction.
- 2) Keep abreast of new advances in machine learning, with a focus on deep learning, including but not limited to Natural language processing, Sentiment analysis, Semantic analysis, Gesture, Emotion and Activity recognition.
- 3) Provide guidance in the creation of research experiments relying on recordings of children's linguistic environments.
- 4) Assist in the management of audio and video recordings to ensure the rigor and quality of the data.
- 5) Work collaboratively with TMW Center's Project Managers, Research and Tech teams to conduct studies.
- 6) Write research articles and contribute to their dissemination via publications and participation in conferences.
- 7) Collaborate across disciplines and communicate methods and findings to non-specialized team members.

Education, Experience, and Qualifications

Applicants must have a Ph.D. in Computer Science, Economics, Statistics, Computational Linguistic, Bioinformatics, Cognitive Science or other related fields, and a strong commitment to research. The ideal candidate is a creative thinker, problem solver and excellent programmer with experience in the processing and analysis of audio, video and natural language data specifically. (S)he needs to be able to learn the most important components of a subject matter, understand the open questions in the field, and identify the data and experiments required to find answers to those questions.

Applicants must be able to handle multiple research projects simultaneously, take initiative and exercise independent judgment. Interdisciplinary experience is a plus, especially in education, developmental psychology, linguistics and applied public health fields.

Technical Knowledge/Skills

- 1) Programming Languages: Python
- 2) Large scale tools: Spark or Dask (rapids.ai is a bonus)
- 3) Core ML: Jupyter, Scikit-learn, Numpy, Pandas
- 4) NLP Tools: Spacy, NLTK or CoreNLP
- 5) Deep learning frameworks: Tensorflow or Pytorch (preferred)
- 6) Experience with Deep Learning Video and Image Analysis
- 7) Experience with Linux command line and High-Performance Computing is a plus
- 8) Experience with cloud resources such as Amazon Web Services is a plus
- 9) Experience with GIT and experiment version control approaches is a plus

Motivated candidates should submit a curriculum vitae and a statement of research goals to Andy Lewis at andrewlewis@bsd.uchicago.edu. Compensation in the Biological Sciences Division follows the NIH NRSA Stipend scale. Additional information on benefits and being a postdoc in the University of Chicago Biological Sciences Division can be found at bsdpostdoc.uchicago.edu.