Damian Yukio Romero Diaz
Charlotte, North Carolina, U.S.A., 28262 | damianiji@arizona.edu

Computational and analytical linguist experienced in data collection, annotation, management of digital assets, and AI. Able to work in English, Mexican/LATAM Spanish and US Spanish under high-demand situations.

TECHNICAL SKILLS

- **Linguistic data annotation:** Spoken data transcription and analysis using APA and IPA conventions | Morphology, syntax, phonology, and pragmatics | Experience with annotation and transcription tools.
- **Data collection:** Creation and management of annotation workflows, guidelines, and reliability metrics | Experience working with PII and ethical and government guidelines.
- **Stewardship of digital assets:** Project management for the creation, distribution, and reuse of datasets (data management plans) | Copyright and legal sanity checks for language datasets
- **Programming languages:** Python | Bash | SQL | PostgreSQL | Git & GitHub
- **Natural language proficiency:** Mexican Spanish (Native and Trained Linguist) | English (Bilingual)

WORK EXPERIENCE

- **Analytical Linguist. Freelance contractor. Remote.**

Several Technology Companies (Undisclosed by agreement) **2019, 2020, Today**
    - Language transcription, language data annotation, and Spanish data localization.

- **Explosion AI, GmbH, Berlin, Germany (Remote)**

Machine Learning Intern **05/2021 – 05/2023**
    - Creation of annotation guidelines and resources for natural language processing.
    - Cross-cultural communication with engineers and customers for machine learning projects.
    - Research into language models and human collaboration for the creation of robust datasets.

- **School of Information, The University of Arizona, U.S.A.**

NLP Research Assistant for Dr. Tracy E. Crane and Dr. Steven J. Bethard **08/2021 – 11/2022**
    - Created an [award-winning data management plan](#) focused on machine learned model sharing and reuse
    - Planned future data-centric ML research via [actionable data management foresight](#).
    - Developed resources for speech data annotation and managed the work of seven part-time annotators.

- **Department of Spanish and Portuguese, The University of Arizona, U.S.A.**

Teaching, outreach, research **2016 - 2020**
    - Designed workflows and conventions of two (private) oral language corpora for Dr. Julieta Fernandez.

EDUCATION

- **M.S. in Human Language Technology**
  The University of Arizona, Tucson, Arizona, U.S.A.
  Graduation date*:* May 2023, GPA: 3.889/4.0
- **M.A. in Hispanic Linguistics**
  The University of Arizona, Tucson, Arizona, U.S.A.
  Graduation date: Spring 2014, GPA: 3.909/4.0
- **B.A. in Linguistics**
  Autonomous University of Querétaro, Querétaro, México
  Graduation date: Summer 2012, Cumulative grade: 9.359/10