

# DANZENG CAIRANG

[cdanzeng@ucsd.edu](mailto:cdanzeng@ucsd.edu) | (858) 428 - 4683 | [LinkedIn](#) | [GitHub](#) | [Personal Website](#)

## EDUCATION

### University of California, San Diego (UCSD)

Jun 2025 – Jun 2026

M.S. in Computational Social Sciences, GPA: 4.0/4.0

**Coursework:** Machine Learning, Deep Learning, Large Language Models (LLMs), Data Analysis, Human-Computer Interaction, Multimodal Data Analysis and Visualization, Social Science Machine Learning.

### Southwest Minzu University

Sep 2018 – Jun 2022

B.A. in Sociology, GPA: 3.88/4.0

**Coursework:** Statistical Analysis, Sociological Programming, Social Research Methods, Social Network Analysis, Social Psychology, Linguistics, Anthropology, Project Management.

## EXPERIENCE

### Independent Developer, La Jolla, CA – AI Application Developer

Dec 2025 – Present

- Independently built and shipped 10+ cross-platform applications, several AI-powered, owning each end-to-end from architecture to release.
- Flagship: Sandbox Oasis, a human-AI collaboration platform that orchestrates multiple models such as Claude and GPT via API and CLI through capability and quality guardrails, isolated worktrees, and a human-in-the-loop patch-review pipeline; also shipped an AI photo editor built on a third-party inpainting model.

### University of California, San Diego, La Jolla, CA – Machine Learning Researcher

Sep 2025 – Jun 2026

- Built end-to-end ML pipelines for naturalistic fMRI analysis—data preprocessing, unsupervised modeling, and cross-subject validation—under a reproducible, version-controlled workflow spanning multiple modeling approaches.

### Southwest Minzu University, Chengdu, CN – Research Assistant

Feb 2020 – May 2025

- Coordinated interdisciplinary field research—collecting, structuring, and analyzing linguistic, survey, and spatial datasets while tracking deliverables across a multi-year project—resulting in three peer-reviewed publications.

## PROJECTS

### Neural Data Decomposition for Hidden Population Discovery in Language and Social Perception – Capstone Project

Sep 2025 – Jun 2026

- Built a reproducible preprocessing pipeline that transformed 540K+ naturalistic fMRI timepoints from 4 subjects into full-resolution timepoint-by-1,000-parcel Schaefer matrices, with automated quality checks and validation procedures.
- Implemented consensus Bayesian Non-Negative Matrix Factorization across 200 repeated model runs, generating 80 subject-level latent components and developing temporal-similarity and stability workflows for cross-subject component matching.
- Validated component interpretations through Yeo 7-network mapping, brain-surface visualization, permutation testing, and multimodal alignment of neural activation peaks with transcripts, video frames, 6,000+ structured visual-coding labels, and robustness-check samples.
- Identified reproducible cross-subject components associated with dialogue and social-action contexts, visible human activity, and a candidate visual food/dining response pattern consistent with established functional-network organization.

### Multimodal Brain-Encoding Transformer – Algonauts 2025 Project

Sep 2025 – Feb 2026

- Independently reproduced the challenge end-to-end — designed and trained a custom multimodal Transformer encoder (PyTorch, mixed-precision, custom correlation loss) fusing visual/audio/language features from pretrained backbones (slow\_r50, BERT) to predict fMRI responses; benchmarked against MLP and linear baselines.

### Dictionary of Common Words in Chinese, Tibetan, Qiang, and Yi – National Social Science Project

Oct 2023 – May 2025

- Constructed and standardized a multilingual language corpus of ~30,000 vocabulary entries across four languages from primary field-recorded audio, producing a cleaned, consistently structured dataset for analysis.
- Ran cross-language comparative and statistical analysis on lexical and phonetic features, quantifying inter-language correlations to surface dialect relationships and inherited-vs-loanword patterns.

### Study of the Kham Dialect in Tibetan Language – National Social Science Project

Dec 2021 – Dec 2023

- Collected, cleaned, and analyzed survey and geospatial data on income, per-capita distribution, and regional settlement patterns using statistical and GIS-based methods, identifying regional socioeconomic trends to support reporting.
- Built and standardized a structured acoustic dataset from primary field recordings, processing raw audio into clean, analysis-ready data; conducted field interviews to contextualize quantitative findings.

## PUBLICATIONS

- |                                                                                                                           |                                                    |          |
|---------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------|----------|
| • <i>Phonological and Lexical Profiles of the Kham Dialect Muya</i>                                                       | Journal of Studies on Languages in China, Volume 6 | Jan 2025 |
| • <i>The Swan Song of Matrilineal Clans: The Walking Marriage Customs and their Evolution Among the Zhaba Zang People</i> | Journal of Sichuan Minzu College                   | Apr 2025 |
| • <i>The Ghee Flowers Art of Lingque Temple in North Kham Area</i>                                                        | China Tibetology, No. 2(2020)                      | Mar 2020 |

## SKILLS & LANGUAGES

- Programming & Systems:** Python (PyTorch, TensorFlow, NumPy, pandas, scikit-learn), TypeScript, JavaScript, SQL, Git, Linux.
- Machine Learning & NLP:** Supervised & Unsupervised Learning, Neural Networks, Transformers, Feature Engineering, Model Evaluation, NLP Pipelines, LLMs, Prompt Engineering, AI Agents, Workflow Automation.
- Data Analysis:** Statistical Modeling, Predictive Modeling, Trend Analysis, Visualization, GIS Analysis, Data Governance, Tableau, Power BI.
- AI App & Tooling:** Model Integration (API & CLI), React, Tauri, Electron, OpenCV, FFmpeg.
- Languages:** English (Fluent), Chinese (Native), Tibetan (Native), Japanese (Working Proficiency).