

VIEUX VALCIN

NLP + AI + ENG

vieuxtiful@pm.me

EXPERIENCE

Positions

Lead Research Assistant

- MIIIS Mixed-Methods, Evaluation, Training, and Analysis (META) Laboratory

Research Assistant II

- BU Chobanian & Avedisian School of Medicine

Research Assistant

- BU Sargent Motor Development Laboratory

Technical Editor

- BU The Nerve Magazine

Member

- BU Mind & Brain Society

Applications

- Project Management
 - Waterfall, Agile
 - JIRA, Confluence
 - Asana
- Web Localisation
 - AEM, WordPress
- Graphics & UI Localisation
 - Adobe Ai, Id, Ps, Pr
- Software L10n/l18n
 - Mobile (iOS, Android)
 - Game
- Desktop Publishing (DTP)

© GPTZero



01.2026 - Present

Onki

Engineer, Computational Linguistics/Psychology

- Prototypes deep NLP architectures for VLM-based multilingual robotic agentic systems.
- Enables technical innovation and advanced product testing initiatives (>5,400 Python LoC {pydub, librosa, numpy, torch, scipy, pydantic} >2,100 JavaScript {React/TS/CSS} LoC).
- Built an automated CLI- and UI-based ground truth annotation framework, STT provider testing environment, and speech database across a broad non-/CCJK language spectrum to enable empirical evaluation, testing, and deployment strategies for enterprise agents.
- Met weekly with the executive team (CEO, CTO) and fellow engineers across to discuss, plan, and execute advanced voice-CV infrastructures.
- Performs statistical metric data analysis on structured audiovisual data.
- Creates an automated testing and speech data annotation framework for the GH repo that reduced evaluation time by 40%.
- Performs routine programming, translation/localisation, and dubbing tasks for STT/TTS model assessment.

11.2025 - Present

Handshake

AI Fellow, MOVE Programme @ Handshake (#1/290 Fellows)

- Provides expertise for complex VLM/LLM/MT system development as an research auxiliary for bleeding-edge AI research for a top AI lab in Silicon Valley.
- Audits 3,100+ datasets as an in-house QA specialist, composed of the top 5% {#1/290} MOVE Programme Fellows (2.11/3.00 QS).
- Scripts web UI {>14,000 React.js LoC} with HTML/CSS wrapping.
- Spends 1,450+ hours on RLHF evaluation responsibilities, particularly for STT/TTS provider testing, complex problem-solving, text/image/video, machine translation {en-US>fr-FR, en-US>zh-Hant}, and document template generation.
- Completes spot translations {en-US>fr-FR, en-US>zh-Hant} for select evals.
- Evaluates quality, fidelity, and reproducibility of translation AI outputs {fr-*, zh-Hant, ja-JP, ko-KR, and th-TH}.

EDUCATION

Boston University
Neuroscience

Middlebury Institute
Localisation Management

Affiliations

Member

American Translators Association (ATA)

Reviewer

American Journal of Managed Care (AJMC)

Skills

SWE & Dev

- Front-end: HTML, TypeScript, SASS/Less/Tailwind CSS, JavaScript, React, Vite, Vue.js, 3js, particles.js, WebGL, shadcn/ui, Radix UI
- Back-end: Node.js, Express.js, PyTorch 2.0+, TensorFlow, Django, Flask, Flask-RESTful
- Database: Snowflake, Google BigQuery, SuperSet, Qlik, LibreOffice Base, MongoDB, MySQL, PostgreSQL, Milvus, Chroma, Pinecone, pgvector, Google Looker
- NLP: Polyglot, spaCy, NLTK, Gensim, MeCab, Google Cloud, AWS, Azure OpenAI, Supabase, Firebase, LangChain
- Other: Command Line Interface (CLI), Git, Hugging Face, Docker, Kubernetes, AEM, WordPress, R, MathWorks MATLAB, YAML, JSON, XML, CI/CD, A/B Testing, Redis, Pydantic, SDK, ADK

Applications

- Project Management
 - Waterfall, Agile
 - JIRA, Confluence
 - Asana
- Web Localisation
 - AEM, WordPress
- Graphics & UI Localisation
 - Adobe Ai, Id, Ps, Pr
- Software L10n/l18n
 - Mobile (iOS, Android)
 - Game
- Desktop Publishing (DTP)

Publications

Manuscripts

- Valcin, V. (2025, October). A systematic statistical-neural architecture for multilingual terminology recognition and cross-lingual processing.
<https://doi.org/10.13140/RG.2.2.2737.8.67529>

Reviews

- Higher Risk, Lower Reimbursement: Medicare Payment Paradox in Cranial Neurosurgery. (2025). The American Journal of Managed Care.

10.2025 - Present

MIIS META Laboratory

Technical Lead

- Built LexiQ™ AI via a React.js/TSX (>31,000 LoC) frontend and Python (>6,500 LoC) backend stack.
- Computed bilingual file parsers for XLIFF, TMX, SDLXLIFF, MQXLIFF, and PO, enabling direct ingestion from Trados, memoQ, and Phrase into an intelligent web- and app-based LQA platform.
- Constructed a SQL (>1,300 LoC) serverless file processing pipeline with advanced JSON (>700 LoC) detection and a session validation gating mechanism with fuzzy matching, configurable thresholds, and cross-segment deviation flagging.
- Developed lexiq.ai {JS >380 LoC, CSS >400 LoC, TSX >7,300 LoC} via Vercel host.
- Initialised a 4-language, multi-domain labeling platform {en-US, fr-*, ja-JP, zh-CN} for use in partnership with a California-based academic institution.

12.2024 - 11.2025

Labelbox

Expert Linguist

- Created multimodal (image, video, CV) data labels for fintech, health, and general agents.
- Wrote 500+ Python scripts (torch/tensorflow) for spatiotemporal feature encoding, function calling, and model inference during UAT and A/B testing for ambient computer vision models.
- Established data engineering pipelines and audited complex datasets for production-level AI systems, ensuring model quality and fidelity in specialised domains.
- Conducted QA assessments, localisation tests, and fine-tuning evaluations for multilingual {fr-*} LM and MT output development and refinement.
- Participated in in-house skill development via proprietary Zara™ tool, enabling written/verbal linguistic skills enhancement for future projects.

11.2017 - 04.2018

BU Chobanian & Avedisian School of Medicine

Research Assistant II

- Transcribed French-language scripts from clinical evaluation focus group session audio files for a gynecology & obstetrics study, led by P.I. Dr. Somphit R. Chinkam, C.N.M.
- Contributed to a manuscript addressing the patient/physician relationship in the local French-speaking maternal community at Boston Medical Center.
- Attended physician and P.I.-led focus groups for open discussion on the state of the physician-patient relationship and birth outcomes.

10.2016 - 04.2017

BU Sargent College Motor Development Lab

Research Assistant

- Evaluated Roux-en-y bariatric surgery candidates by acquiring and analysing multimodal data from 3D motion capture systems, wearable inertial sensors, and digital pressure mats.
- Executed structured video-coding for fraternal twin phenotypic correlation data, contributing to a behavioral modality database for a multi-institutional study.
- Coordinated with Dr. Simone Gill (P.I.) and Dr. Jan Stasinski (post-doctoral fellow) to amalgamate research data for large-scale motor development studies.

FREELANCE**AI Linguist @ BeSimple.ai**

- Creates multi-specialty (medical, legal, technical) French {fr-FR, fr-CA, fr-CH}, and English {en-US} text-to-speech (TTS) datasets to train voice AI systems.

AI Linguist @ Uber AI Solutions

- RLHF accuracy and cultural alignment evaluation for multimodal (image) LLM outputs {ja-JP} in focused leading-edge AI development encompassing data labeling, product testing, and localisation initiatives.

Translator @ Translated

- Performs MTPE across linguistic codes {en-US>fr-*, ja-JP>fr-FR}.
- Participated in a top client portfolio translation initiative, guided by Senior Localisation Project Manager Eleonora Ruggiu.

PROJECTS**LexiQ™ AI xLQA | lexiqai.vercel.app***Technical Lead, Full-Stack***Accolades**

- Booth/Stage Pitch, LocWorld54 Process Innovation Challenge (PIC) #19 Sneaks #1, 1st cohort (\$X,XXX Full Funding)
 - Oct 15, 2025 | Monterey, CA
- 3cubed (3^3) Preliminary Rounds Qualifier, LocWorld55 Process Innovation Challenge (PIC) #20
 - Apr 17, 2026 | Dublin, IE
- Produced the LexiQ™ AI Data Annotation platform across four linguistic codes {en-US, fr-FR, zh-CN, ja-JP}, for graduate cohort use at the Middlebury Institute of International Studies at Monterey.

Word2World™ UI | github.com/vieuxtiful/Word2WorldUI*Technical Lead***Accolades**

- Model Presentation, Manus AI (Meta)
 - Interview Panel
 - Parker Lyman (Chief of Staff), Michelle Tan (GTM), Lantao (Victor) Li (GTM)
 - Sep 22, 2025 | San Mateo, CA
- Programmed a production-ready pipeline (torch) capable of scaling to 1M+ users with 97% memory reduction versus SOTA approximation models (e.g., faiss).
 - Front-end: React 19, TypeScript, Vite 7, Tailwind CSS 4, Canvas API, Framer Motion, 3js
 - Back-end: Python 3.11+ , PyTorch 2.0+ , Transformers, SentenceTransformers, CLIP, Node2Vec
- License: MIT

Project Apollo*Back-end Engineer*

- Programmed a multifarious, computer vision-based functional ambient prototype intuitively designed to ameliorate T2DM patient QoL and daily life experience.
 - Front-end: Radix UI
 - Back-end: pyaudio, deepspeech, textblob
- ML Lib: tensorflow, scikit-learn