

Hennessy Ng

Technical Product Manager, Localization

AI-native localization platforms — LLM translation, workflow automation, full-stack delivery

📍 Hong Kong · Remote / Hybrid ✉ hi@hennessyng.com 🌐 hennessyng.com 📁 github.com/hennessyng

📄 linkedin.com/in/hennessyng

Technical Product Manager with 8+ years building localization automation, now owning an AI-native localization platform end to end. I work the full cycle — discovery, PRDs and roadmap through to shipping Django/Vue services, RAG systems and LLM translation workflows myself. Background spans terminology and translation QA, NMT and OCR pipelines, and cloud-native platform engineering on AWS, GCP and Azure.

📁 EXPERIENCE

Product Manager, Localization

Aug 2024 – Present

Crypto.com · Hong Kong · Remote

Product owner and builder for an AI-native localization management platform — owning discovery, roadmap and delivery while writing production code across the stack.

- Own the roadmap and product definition for an AI-native localization management platform, translating localization operations pain points into shipped capability.
- Design and build end-to-end Django / Vue.js workflow pipelines, delivered through GitHub Actions CI/CD.
- Design RAG systems and AI agents that automate the translate–edit–proofread (TEP) workflow, keeping terminology and style consistent across locales.
- Develop and optimise localization prompts, and build LQA tooling to score machine-translated output for accuracy.
- Evaluate cloud and open-source LLM options against cost, latency and quality to keep AI localization workflows efficient.
- Architect cloud infrastructure on AWS, running microservices orchestrated with Kubernetes.

Django Vue.js Python RAG AI agents Prompt engineering LLM evaluation AWS Kubernetes GitHub Actions

Language Technology Consultant

Nov 2017 – Aug 2024

DataSource International · Hong Kong

Built the NLP, OCR and machine translation automation that moved a traditional language services business onto cloud-native infrastructure.

- Designed and shipped an OCR + machine translation pipeline processing 500+ scanned PDFs (~5,000 pages) daily from Japanese to English.
- Built bilingual term extraction using BERT embeddings and semantic similarity, generating termbases that fed both MT training and linguist consistency.
- Developed multilingual web crawling and language detection pipelines to acquire and clean parallel corpora for MT training.
- Migrated desktop-bound language tooling to cloud-based RESTful APIs, integrating with CMS platforms and cloud MT vendors.
- Trained custom MT engines on validated post-edited output, recursively improving quality with roughly 40% of production data validated and reusable.

Python EasyOCR Tesseract spaCy KeyBERT Sentence-BERT Scrapy GCP AutoML Azure Translator RESTful APIs

Senior Manager of Quality Assurance (Translation)

Oct 2016 – Oct 2017

Mind Your Language · Hong Kong

Owned translation quality assurance, and introduced neural machine translation and OCR into the QA workflow.

- Led the translation QA function, defining the standards and review process applied across delivery.
- Introduced neural machine translation and OCR into production workflows, shifting the team toward post-editing rather than translation from scratch.
- Ran linguist recruitment and training, building post-editing capability across the team.

Neural MT OCR MTPE QA process design

DataSource International · Hong Kong

Computer-aided translation tooling, terminology management and hands-on linguistic delivery.

- Administered and supported computer-aided translation tooling across the delivery team.
- Owned terminology management — termbase creation, curation and maintenance for corporate accounts.
- Delivered translation and proofreading work directly, grounding later automation in real linguist workflow.

CAT tools Translation memory Terminology management

SELECTED PROJECTS

Otoflow

Personal project

Full-stack piano practice app turning PDF sheet music into an interactive falling-notes waterfall. PDF upload, AI/rule-based recognition, synced sheet + keyboard UI, MIDI input, hand-separate practice, and wait-for-correct-note mode. React/TypeScript + WebGL frontend, Python/FastAPI backend, async workers, PostgreSQL/object storage, multi-service Docker.

React TypeScript WebGL Python FastAPI PostgreSQL Docker MIDI

Senji (せんじ)

Personal project

Self-hosted knowledge ingestion platform converting URLs, HTML, PDFs and images into structured Obsidian markdown, with AI wiki generation and semantic embeddings.

Python Embeddings LLM Markdown Obsidian OCR

OCR + MT Workflow

DataSource International

Japanese-to-English OCR and machine translation pipeline processing 500+ scanned PDFs (~5,000 pages) daily into deliverable DOCX, with validated post-edited output recycled into custom MT engine training.

EasyOCR Tesseract PDF2Image CUDA / GPU GCP AutoML Translation Azure Translator REST API

Bilingual Term Extraction

DataSource International

Chinese-English term extraction from parallel corpora using KeyBERT embeddings and cross-lingual semantic similarity, producing termbases for MT training and linguist consistency.

KeyBERT Jieba EasyNMT Sentence-BERT Manhattan LSTM Python

Web Crawling + Language Detection

DataSource International

Scrapy-based crawler acquiring multilingual PDFs at scale, with OCR fallback, spell-check-based OCR quality gating and document-level language detection to build clean parallel corpora for MT training.

Scrapy requests PyMuPDF pyenchant langdetect spaCy Regex

Cloud-Based System Integration

DataSource International

Migration of desktop-bound language technology to cloud-native RESTful services, with vendor MT integration, CMS connectors and centralised linguistic database maintenance.

RESTful API GCP Azure AWS Drupal / CMS integration SQL

SKILLS

Product	Product ownership · Roadmap & PRDs · Discovery & user research · Agile / Scrum · Stakeholder management · Success metrics & SLAs
AI / LLM	RAG systems · AI agents · Prompt engineering & optimisation · LLM evaluation · MT quality estimation (LQA) · Cloud & open-source model selection
Backend	Python · Django · FastAPI · Pandas · NumPy · PyTorch · TensorFlow · SQL
Frontend	Vue.js · Nuxt.js · TypeScript · Tailwind CSS
NLP	spaCy · NLTK · BERT / Sentence-BERT · Hugging Face · KeyBERT · Jieba · EasyNMT
Platform & DevOps	AWS · GCP · Azure · Docker · Kubernetes · GitHub Actions CI/CD · RESTful APIs · Linux · Microservices
Localization Engineering	Machine translation · TEP workflow design · Post-editing (MTPE) · Terminology & termbase management · Translation QA · Subtitling · CAT tooling

EDUCATION & CERTIFICATION

Professional Certificate in Artificial Intelligence — Stanford University	2021
MA, Computer-Aided Translation — The Chinese University of Hong Kong	2012
BSocSc, Sociology — Hong Kong Shue Yan University	2010

🇮🇹 LANGUAGES

🇮🇹 Cantonese (Native) 🇬🇧 English (Full professional) 🇨🇳 Mandarin (Full professional) 🇯🇵 Japanese (Beginner)