


✉ translatjohn@gmail.com

☎ +34 670 087 316

JOHN HANDLEY

NLP Developer | Data Scientist

linkedin.com/in/john-handley 

github/johnhandleyd 

PROFILE

NLP Developer leveraging a strong linguistics background and a passion for data and research. Detail-oriented, versatile and a team player, driven by new opportunities and challenging environments. I speak Spanish (native) and English fluently.

SKILLS AND LANGUAGES

- **Languages (technical):** Python (OOP); PyTorch and R (intermediate)
- **Data:** ETL, SQL (intermediate) (PostgreSQL, MySQL), MongoDB
- **Misc.:** Git, GitHub; MS Office; Agile (Scrum, Kanban), Jira
- **Libraries:** spaCy, NLTK, pandas, NumPy, Matplotlib, scikit-learn; HuggingFace, OpenNMT, OpenAI (ChatGPT), LLMs
- **Computational Linguistics:** Stanford-NLP, WordNet, JSpeech Grammar Format, FOMA, XML, XPath
- **Languages (spoken):** Spanish (native); English (C2); French (A2)

PROJECTS

Therapy Chatbot | Python (HuggingFace, pandas, JSON)

github.com/johnhandleyd/thesa

- Datasets normalised with custom functions. Fine-tuned with a GPTQ (weight quantisation) version Zephyr-7B
- Inference achieves generalisation (i.e., depression, family issues); suggests next steps and resources with empathy

Impact of Noise in Neural MT | Python (OpenNMT, NLTK, spaCy, pandas)

- Wikipedia Corpus (+1,8M segments) trained using the Transformer architecture, paired with BPE and tokenisation
- Findings showed that abnormal length and foreign language content decreased translation quality

EXPERIENCE

QA Manual Tester

Saal Digital

Aug. 2021 – Feb. 2023

- Performed daily regression and usability tests, and hundreds of integration and volume test cases on iOS and Android (app and website), Chrome and Safari (website), Windows and macOS (desktop software)
- Search, documentation and prevention of bugs, reported via the internal ticketing system (Kanban)
- Results enhanced user experience and allowed faster release of new features (1x/month to 1x/week, or 2x/month)
- Teamed with developers to conduct paired testing on newly developed features, and with senior management to reach performance goals and adopt optimal QA practices

NLP Developer: Chatbot | Python (NLTK, spaCy, pandas)

Sept. – Dec. 2022

- Custom data (+1M segm.) cleaning and fine-tuning to HuggingFace's Transformer, and GPT-3's DaVinci via OpenAI
- Conversations solved standard customers' questions with context-appropriate information

EDUCATION AND COURSES

Google Data Analytics Professional Certificate

Coursera

In Progress

NLP with Classification and Vector Spaces | Python (NumPy, NLTK, Naïve Bayes, Logistic Regression)

DeepLearning.AI

Dec. 2023

Databases with SQL | Querying, designing, writing, optimising and scaling databases with SQL

CS50, Harvard University

Nov. 2023

Master in Language Analysis and Processing, spec. in Machine Translation

Universidad del País Vasco

Sept. 2020 – 2022

Donostia

- Machine Learning, Deep Learning, neural networks (RNN, LSTM/GRUs, CNN, Transformers), word embeddings, topic modelling, recommendation systems, text classification and stance detection
- Computational Morphology, Syntax, Semantics, Pragmatics; Corpus Linguistics

Degree in Translation and Interpreting, spec. Technologies and Science

Universidad Complutense de Madrid

Sept. 2016 – June 2020

Madrid