Richard Evans

+44 773 638 1233 | r.j.3vans@gmail.com | RJ3vans.github.io | ORCID: [0000-0002-1220-8605](0000-0002-1220-8605) | *Google Scholar*

PROFESSIONAL SUMMARY

Goal-driven Applied Machine Learning Engineer with 25+ years of experience in developing, finetuning, and deploying ML models for practical applications in NLP, data analysis, and AI systems. Expertise in building scalable solutions using Python, PyTorch, and Transformers to solve real-world problems like information extraction, text simplification, and semantic parsing. Skilled in translating data insights into business strategies, collaborating cross-functionally, and contributing to team growth through knowledge sharing and mentorship. Eager to apply ML to optimize inventory, predict demand, and enhance customer experiences in the refurbished tech industry at The Big Phone Store.

KEY SKILLS

Machine Learning & AI: Finetuning LLMs (e.g., BERT, DeBERTa, ELECTRA), model development and deployment, CRF-based classifiers, A/B testing, RLHF frameworks, synthetic data generation.

Programming & Tools: Python (advanced), PyTorch, Transformers (Hugging Face), NLTK, spaCy, NumPy, pandas, RegEx, JSON, Git, GitLab, Conda, Google Colab, PyCharm, Bash.

Data Handling: Data cleaning, anomaly detection, feature extraction, annotation schemes, training data preparation, metrics/segments for optimization.

Collaboration & Communication: Cross-functional teamwork (e.g., with developers, stakeholders), presenting insights at conferences, peer review, supervising projects, explaining technical concepts to non-experts.

Other: Algorithms and data structures for efficient model design, ethical AI deployment, problem-solving in fast-paced environments.

EDUCATION

PhD in Computer Science: "Sentence Simplification for Text Processing" *University of Wolverhampton, UK*	Feb. 2012 – Jan. 2020
Msc in Cognitive Science and Natural Language *University of Edinburgh, UK*	Sep. 1995 – Jul. 1996
BA (Hons) Linguistics *University College of North Wales, UK*	Sep. 1992 – Jun. 1995

PhD focused on ML-driven syntactic analysis and simplification, including human evaluation via surveys and extrinsic testing in NLP applications.

PROFESSIONAL EXPERIENCE

AI Writing Evaluator (Contract)   May2024–Present

Smart Ecosystem, Inc., Remote

- Designed and evaluated training data for Generative AI models, incorporating complex reasoning and multimodal prompts to enhance model performance.

- Implemented RLHF frameworks by developing postgraduate-level MCQs on machine learning topics, reviewing/correcting items, and translating data insights into actionable improvements for AI developers.

- Collaborated with cross-functional teams via Slack to identify AI opportunities, ensuring data-driven strategies for model optimization.

- Contributed to team growth by sharing knowledge on data preparation and evaluation, supporting colleagues in building advanced ML solutions.

NLU English Language Consultant (Contract)   Feb. 2023 — Mar. 2024

Cerence Inc., Remote

- Built and deployed ML resources (e.g., JSGF grammars, gazetteers) for synthetic data generation to train cloud-based and embedded AI models parsing user commands.

- Developed Python scripts for data processing, annotation, entity canonicalization, and model evaluation, fixing bugs and using Gradle for build automation.

- Optimized model coverage for real-world queries (e.g., UEFA Euro 2024 events), turning semantic insights into business-aligned strategies for clients like Audi and BMW.

- Coordinated with teams via GitLab, reviewing code and discussing requirements to foster collaborative problem-solving and knowledge sharing.

Audio Attributes Annotator & Audio-Visual Segmentation (Contract)   Mar. 2025 — May 2025

RWS Holdings PLC, Remote

- Annotated multimedia data (videos) for ML training, identifying sound events, timings, and visual masks to support AI model development in acoustic environments.

- Used proprietary tools (SRT HALO) to prepare high-quality datasets, ensuring accuracy for downstream ML applications like object detection and segmentation.

Lecturer in Computational Linguistics   Feb. 2020 — Feb. 2023

University of Wolverhampton, UK

- Delivered lectures on ML topics including text simplification, sentiment analysis, and machine translation; supervised 2 PhD and 2 MA students on projects like biographical IE and ethical misconduct detection in legal texts using deep learning.

- Co-organized SemEval-2021 shared task on Lexical Complexity Prediction (198 teams), analyzing data to derive insights and contribute to ML benchmarks.

- Finetuned BERT models for anaphora resolution, translating research into practical NLP tools; chaired ethics committees to ensure data-driven, responsible AI deployment.

- Collaborated with cross-functional groups (e.g., ethics subcommittees) to identify opportunities for ML in accessibility and legal analysis.

Researcher / Research Associate   Dec. 2016 — Feb. 2020

University of Wolverhampton, UK

- Developed CRF-based ML models for partial parsing and syntactic simplification, including annotation tools and training data preparation; achieved F_1 scores of 0.72–0.97.

- Performed extrinsic evaluations of ML models via applications like information extraction and summarization, turning data into strategic insights for accessibility tools.

- Contributed to team growth by writing peer-reviewed papers and supporting grant proposals, fostering a collaborative environment for ML innovation.

Earlier Roles   Dec. 2003 — Feb. 2016

University of Wolverhampton, UK

- Scientifically coordinated EU-funded FIRST project (€2.5M) on ML-driven text simplification for autism accessibility, developing prototypes and evaluating impacts with end-users.

- Built ML frameworks for named entity recognition, animacy detection, and IE from clinical/legal texts, using tools like Perl, CRF++, and WordNet.

- Collaborated with multidisciplinary teams (IT, healthcare) to deploy models and share knowledge, optimizing processes for real-world deployment.

Selected Projects

Partial Parsing Using BERT for Sentence Simplification   Feb. 2020 — Present

- Finetuned pretrained models (BERT, DeBERTa, ELECTRA) using PyTorch for sequence tagging and partial parsing, outperforming baselines ($F_1 = 0.97/0.86$).

- Deployed models on Hugging Face for text simplification applications, demonstrating practical ML for efficiency gains.

A Flexible Interactive Reading Support Tool (FIRST)   Oct. 2011 — Sep. 2014

- Developed CRF-based classifiers and annotated corpora for syntactic analysis, enabling ML-driven text conversion for better accessibility.

- Collaborated with 9 organizations to evaluate and optimize models, translating insights into user-focused strategies.

Sentence Simplification for Information Extraction   Feb. 2009 — 2011

- Built ML classifiers (TiMBL) and rule-based systems to simplify complex sentences, improving IE recall in clinical data.

- Integrated models into pipelines, showcasing data-driven problem-solving for business-like applications.

Additional Information

- **Publications**: 30+ peer-reviewed papers on ML/NLP topics (e.g., syntactic simplification, lexical complexity prediction). Samples on GitHub and Google Scholar.

- **Peer Review & Presentations**: Reviewer for ACL, EMNLP; presented at international conferences, communicating complex ML ideas effectively.

- **Interests**: Ethical AI in business contexts, big data processing (eager to explore Hadoop/Hive for scalable ML).

Selected Papers

Evans, R. (2011). Comparing methods for the syntactic simplification of sentences in information extraction. *Literary and Linguistic Computing*, 26 (4):371–388.

Evans, R. and Orasan, C. (2013). Annotating signs of syntactic complexity to support sentence simplification. In Habernal, I. and Matousek, V., editors, *Text, Speech and Dialogue. Proceedings of the 16th International Conference TSD 2013*, pages 92–104. Springer, Plzen, Czech Republic.

Evans, R. and Orasan, C. (2019a). Identifying signs of syntactic complexity for rule-based sentence simplification. *Natural Language Engineering*, 25 (1):69–119.

Evans, R. and Orasan, C. (2019b). Sentence simplification for semantic role labelling and information extraction. In *Proceedings of the International Conference "Recent Advances in Natural Language Processing '2019" (RANLP-2019)*, pages 285–294, Varna, Bulgaria.

Parodi, G., Evans, R., Ha, L. A., Mitkov, R., c. J. J. Vergara, and Olivares-López, R. (2021). A sequence labelling approach for automatic analysis of *ello*: tagging pronouns, antecedents, and connective phrases. *Language Resources and Evaluation*, 56 (1):139–164.

Shardlow, M., Evans, R., and Zampieri, M. (2022). Predicting lexical complexity in English texts: the Complex 2.0 dataset. *Language Resources and Evaluation*, 56:1153–1194.

Yaneva, V. and Evans, R. (2015). Six good predictors of autistic text comprehension. In *Proceedings of Recent Advances in Natural Language Processing (RANLP 2015)*, pages 697–706, Hissar, Bulgaria.

Please visit this link for a complete list of my peer-reviewed research papers.