

Richard Evans

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

PROFESSIONAL SUMMARY

Detail-oriented data curator with **25+ years of research and curation experience** in NLP, ensuring data is **FAIR-compliant, usable, and preserved** for open science reuse. Led annotation, metadata review, and quality control for corpora; guided authors on best practices. Proficient in **file formats, software tools, and customer service** for researchers. U.K.-based — **fully remote, 12-month contract ready**. Eager to support Dryad's mission in high-standard data publication.

KEY QUALIFICATIONS

- **Data Curation & Research Experience:** Curated quantitative/qualitative data (NLP corpora, AI training sets); ensured compliance with guidelines.
- **File & Software Proficiency:** Handled XML, JSON, text files; used Python, NLTK, spaCy, PyTorch, R-like tools (NumPy, pandas), ImageJ-equivalent (video annotation).
- **Time-Management & Organization:** Managed multiple projects; met daily/weekly curation targets.
- **Communication & Customer Service:** Corresponded with authors/reviewers; provided guidance for improvements.
- **Microsoft Office & Google Drive:** Intermediate Excel, Google Sheets for data tracking; quick navigation/Windows OS.
- **Ethical & Legal Compliance:** Chaired ethics committee; anonymised data per GDPR.

PROFESSIONAL EXPERIENCE

AI Writing Evaluator (Contract)

May 2024 — Present

Smart Ecosystem, Inc., Remote

- Curated AI training data: reviewed metadata, files for compliance; ensured FAIR principles (findable, accessible, interoperable, reusable).
- Guided authors on improvements via feedback; resolved issues for quality publication.
- Handled diverse file types (text, multimedia); used Google Sheets for tracking.

NLU English Language Consultant (Contract)

Feb. 2023 — Mar. 2024

Cerence Inc. – Remote

- Curated semantic data: evaluated utterances, gazetteers for accuracy/relevance; flagged ethical issues.
- Corresponded with stakeholders; provided best practice advice for data reuse.
- Managed JSON files; used Python for canonicalisation and workflow optimization.

Audio Attributes Annotator & Audio-Visual Segmentation (Contract)

Mar. 2025 — May 2025

RWS Holdings PLC – Remote

- Curated audio/video data: reviewed metadata, timing; ensured no legal/ethical concerns.
- Accessed proprietary software (SRT tools); prepared files for digital storage/reuse.
- Maintained high curation standards; assisted team with process improvements.

Lecturer in Computational Linguistics

Feb. 2020 — Feb. 2023

University of Wolverhampton

- Curated student/research data: supervised PhD/MA theses; reviewed submissions for quality.
- Chaired EMTTI Ethics Committee: ensured data integrity, compliance; trained on ethical curation.
- Developed annotation schemes; guided researchers on FAIR-like best practices.

Researcher / Research Associate

Dec. 2016 — Feb. 2020

University of Wolverhampton

- Curated NLP corpora: annotated texts, metadata for preservation/reuse in open science.
- Identified workflow optimizations; implemented Python tools for efficiency.
- Corresponded with collaborators; provided curation guidance.

Scientific Coordinator, EU-Funded FIRST Project (€2.5M)

Dec. 2013 — Dec. 2016

University of Wolverhampton

- Curated project data: reviewed files from 9 partners; ensured accessibility/reusability.
- Assisted in team training; maintained documentation for curation protocols.
- Optimized processes for data publication; focused on integrity and sustainability.

EDUCATION

PhD in Computer Science: "Sentence Simplification for Text Processing"

Feb. 2012 – Jan. 2020

University of Wolverhampton, UK

Msc in Cognitive Science and Natural Language

Sep. 1995 – Jul. 1996

University of Edinburgh, UK

BA (Hons) Linguistics

Sep. 1992 – Jun. 1995

University College of North Wales, UK

TECHNICAL & CURATION TOOLS

- **File/Software:** XML, JSON, text editors, Python (pandas, NLTK, spaCy), PyTorch, Google Colab, Win-SCP.
- **Office/Google:** Excel (formulas, filtering), Google Sheets/Drive (tracking, collaboration).
- **Navigation:** Windows OS, quick keys, rapid file access.
- **Standards:** FAIR principles, metadata (DOIs, identifiers), open science best practices.

SELECTED PAPERS

Evans, R. (2001). Applying machine learning toward an automatic classification of it. *Literary and linguistic computing*, 16(1):45–58.

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

Orăsan, C., Evans, R., and Mitkov, R. (2018). Intelligent text processing to help readers with autism. In Shaalan, K., Hassanien, A. E., and Tolba, M. F., editors, *Intelligent Natural Language Processing: Trends and Applications*, pages 713–740. Springer.

Pekar, V. and Evans, R. (2007). Discovery of language resources on the web: Information extraction from heterogeneous documents. *Literary and Linguistic Computing*, 22(3):329–343.

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](#) or my [Google Scholar](#) for a complete list of my peer-reviewed research papers.