

# José Daniel GUTIÉRREZ NIETO

La Paz, México | +52-612-158-3817 | danezgj@gmail.com | linkedin.com/in/danezgj | github.com/ezejg

## PERFIL PROFESIONAL

Lingüista Computacional con formación en ciencia cognitiva, lingüística y procesamiento del lenguaje natural (NLP). Experiencia en investigación e implementación de tecnologías de NLP, curación automática, consultoría lingüística, y colaboración interdisciplinaria. Dominio de entornos Linux, machine learning, procesamiento de datos y redacción técnica. En todos mis proyectos mantengo un enfoque riguroso y con atención al detalle. Bilingüe en español (nativo) e inglés (fluido).

## EXPERIENCIA DE TRABAJO

### NeTS (IUSS Pavia)

Pavia, Italia

Investigador en NLP (Pasante)

2023-2024

- Diseñé un modelo experimental de parsing para fenómenos de adjunción complejos, integrando frameworks computacionales con modelos cognitivos modernos sobre el procesamiento del lenguaje (discontinuous parsing, teoría del acceso léxico, gramáticas generativas top-down).
- Colaboré con investigadores internacionales para presentar innovaciones en parsing en el *31st Colloquium on Generative Grammar* (Palma, España), contribuyendo a discusiones sobre metodologías interdisciplinarias en NLP y al avance en la comprensión de estructuras complejas de adjunción.

### Centro de Ciencias Genómicas (CCG-UNAM)

Cuernavaca, México

Especialista en NLP

2016-2019

- Implementé soluciones en NLP para RegulonDB, una base de datos de regulación genética en *E. coli* K-12, desarrollando métodos de minería de texto especializados para extraer datos de regulación genética de literatura biomédica (biocuración automática, Python, Linux/Bash, OpenIE).
- Diseñé componentes híbridos de minería de texto combinando análisis basado en reglas lingüísticas (simplificación de texto, identificación de roles temáticos) con machine learning (reconocimiento de entidades, extracción de relaciones), logrando un F-score de 0.84 en identificación de pares regulatorios, superando a los sistemas existentes comparables.
- Dichas contribuciones permitieron la adaptación transversal de pipelines de biocuración (de literatura sobre *E. coli* a *Salmonella*), demostrando la transferibilidad de los métodos desarrollados.
- Establecí vínculos entre los equipos de lingüística computacional y de genómica mediante mentoría técnica y presentaciones interdisciplinarias regulares.

Investigador en NLP (Pasante)

2016

- Brindé asesoría lingüística y contribuí al rediseño de un pipeline de biocuración automática en la recién formada división de NLP (análisis de estructura sintáctica, evaluación de métodos de simplificación de texto).
- Desarrollé prototipos de componentes para simplificación de texto que posteriormente se escalaron a sistemas de producción (Python, Bash).

### Semiosfera Marketing

Ciudad de México, México

Consultor Lingüístico

2016

- Asesoré sobre el diseño y la curaduría de corpus anotados de español mexicano con datos recopilados mediante entrevistas de campo orales, estableciendo lineamientos de anotación y análisis lingüístico.
- Colaboré en el desarrollo de corpus multimodales anotados (transcripción, análisis del discurso, anotación fonética y prosódica, etiquetado pragmático, Praat) para aplicaciones de NLP en investigación de mercado.

### Laboratorio de Estudios Fónicos (El Colegio de México)

Ciudad de México, México

Lingüista de campo | Especialista en datos de habla

2014

- Realicé grabaciones de campo sistemáticas de español mexicano dialectal con controles fonéticos y sociolingüísticos en múltiples regiones, generando corpus de habla para el análisis prosódico y variacionista.

## HABILIDADES TÉCNICAS

**Competencias principales:** NLP, machine learning, procesamiento de datos, redacción técnica, investigación, consultoría.

**Programación:** Python, Bash, C, SQL, LaTeX, Markdown.

**Herramientas y frameworks:** PyTorch, scikit-learn, spaCy, Pandas, NumPy, Praat, Audacity, Linux, macOS.

**Análisis lingüístico:** Procesamiento del habla, lingüística de corpus, gramática generativa y de dependencia, análisis pragmático-discursivo.

**Università degli Studi di Trento***M.Sc. in Cognitive Science**Trento, Italia**2022–2023***Universidad del País Vasco***M.Sc. in Computational Linguistics**San Sebastián, España**2020–2021***Escuela Nacional de Antropología e Historia***Licenciatura en Lingüística (mención honorífica)**Ciudad de México, México**2013–2017***PRESENTACIONES Y PONENCIAS**

*Construyendo la proyección extendida de N: Una perspectiva derivacional descendente.* Conferencia en el V Encuentro de Lingüística Formal en México (ELF). Benemérita Universidad Autónoma de Puebla, Puebla, 2024.

*Towards an Analysis of Ordering Restrictions within Noun Phrases: A Top-Down Minimalist Approach.* Presentación de póster en el 31st Colloquium on Generative Grammar. Universidad de las Islas Baleares, Palma, 2022.

*Sentence Simplification for Biomedical Literature.* Presentación para el Seminario de Genómica Computacional. CCG-UNAM, Cuernavaca, 2017.

*Extraction of Regulatory Interactions and Growth Conditions.* Presentación para el Programa de Genómica Computacional. CCG-UNAM, 2018.

*La sintaxis de los adjetivos en la teoría generativa.* Ponente invitado. ENAH, Ciudad de México, 2019.

*La frase nominal y sus modificadores.* Ponente invitado. ENAH, Ciudad de México, 2017.

*Problemas en el diseño de una interfaz morfosintáctica de traducción automática de FNs entre español e inglés.* Conferencia en el XI Coloquio de Lingüística en la ENAH. ENAH, Ciudad de México, 2016.