

JOAQUÍN DE FERRARI GONZÁLEZ

Computer Science Engineer, MSc · PhD Student in Engineering Sciences (Computer Science, UC Chile)

Santiago, Chile jdeferrari@estudiante.uc.cl LinkedIn ORCID 0009-0009-5885-888X github.com/Joacodef

RESEARCH PROFILE

Computer Science Engineer and MSc (UTFSM); PhD student in Engineering Sciences (Computer Science) at Pontificia Universidad Católica de Chile and doctoral fellow at the Millennium Institute iHealth. My research asks how to train **multimodal 3D vision-language models for chest CT under imperfect supervision**: incomplete voxel-level annotation, over-reliance on text over image evidence, and long-tailed findings. It builds on a track record in **LLM-based weak supervision** and clinical NLP in Spanish, from hospital imaging infrastructure (FONDEF ALPACS) to text-conditioned volumetric segmentation. First author with an oral presentation at **AIME 2025** (Springer LNCS); co-first author in a **WoS Q1** journal (BMC MIDM 2026) and at IEEE ICPRS 2024.

EDUCATION

PhD in Engineering Sciences, Computer Science 2026 – Present
Pontificia Universidad Católica de Chile (UC) Santiago, Chile

- **Thesis line:** text-conditioned 3D CT segmentation under imperfect supervision (partial annotation, image-vs-text reliance, long-tailed findings).
- **Co-advisors:** Prof. Jocelyn Dunstan (clinical NLP) & Prof. Denis Parra (multimodal systems). Doctoral fellow, Millennium Institute iHealth.

MSc in Computer Science Engineering (Magíster en Ciencias de la Ingeniería Informática) Mar 2025 – Apr 2026
Universidad Técnica Federico Santa María (UTFSM) Santiago, Chile

- **Thesis** (defended 15 Apr 2026; advisor Prof. Ricardo Nanculef): *Evaluating GPT as a Textual Oracle for Chest CT Classifiers*. Compared multi-label visual classifiers trained on LLM-generated labels against expert annotations over frozen CT-CLIP volumetric features (46,438 weakly labelled volumes). The LLM oracle reached 0.889 macro-F1 ($\kappa = 0.947$); synthetic supervision matched expert labels from $N \geq 100$ and outperformed them thereafter, plateauing at 0.568 AUPRC and isolating a representational bottleneck via linear probing and 5-fold CV over 5 seeds.
- UTFSM internal Master's scholarship (tuition and stipend, 2025).

B.Sc. & Computer Science Engineering 2018 – 2024
Universidad Técnica Federico Santa María (UTFSM) Santiago, Chile

- Graduated with Academic Distinction. Undergraduate thesis (within FONDEF Proximity): LLM weak supervision for annotating Spanish chest CT reports, published at AIME 2025 (first author, oral).

PEER-REVIEWED PUBLICATIONS

INDEXED JOURNAL ARTICLES

Godoy, E.*, **De Ferrari, J.***, Lazo, S., Pacheco, C., Mellado, D., Saez, A., Chabert, S., & Salas, R. (2026). A Framework for Extraction of Clinical Information from Radiological Mammography Reports Using Large Language Models and Retrieval Augmented Generation. *BMC Medical Informatics and Decision Making*, Vol. 26, Article 345. *Co-first author WoS / Scopus Q1
DOI: 10.1186/s12911-026-03633-8

Godoy, E., Mellado, D., **De Ferrari, J.**, Querales, M., Saez, A., Chabert, S., Parra, D., & Salas, R. (2025). Hybrid framework for automated generation of mammography radiology reports. *Computational and Structural Biotechnology Journal (CSBJ)*, Vol. 27, pp. 3229–3239. WoS / Scopus Q1
DOI: 10.1016/j.csbj.2025.07.018

INTERNATIONAL CONFERENCE PROCEEDINGS

Benoit-Cea, D., Nanculef, R., & **De Ferrari, J.** (2026). Multilingual Minimal Contrastive Editing. In *Proceedings of the 18th International Conference on Agents and Artificial Intelligence (ICAART 2026)*, Vol. 5, pp. 4590–4598. SCITEPRESS.
DOI: 10.5220/0014474800004052

De Ferrari, J., Nanculef, R., Benoit, D., Araya, M., & Solar, M. (2025). Assessing GPT as a Weak Oracle for Annotating Radiological Studies. *Artificial Intelligence in Medicine (AIME 2025)*, Pavia, Italy. In *Lecture Notes in Computer Science (LNCS)*, vol. 15734, pp. 98–109, Springer.
Oral Presentation DOI: 10.1007/978-3-031-95838-0_10

Godoy, E.*, **De Ferrari, J.***, Chabert, S., Mellado, D., & Salas, R. (2024). Automatic detection of contextual laterality in Mammography Reports using Large Language Models. *IEEE International Conference on Pattern Recognition Systems (ICPRS 2024)*, London, United Kingdom, pp. 1–6.
*Co-first author DOI: 10.1109/ICPRS62101.2024.10677842

RESEARCH DATA

De Ferrari, J., Araya, M., Solar, M., & Nanculef, R. (2025). Chilean Radiological Reports - AIME 2025 [Dataset]. 42,605 Spanish-language chest CT reports from HCUCH labelled with GPT, plus 217 reports labelled by four radiologists (condensation, nodules, cysts). *Zenodo*.
DOI: 10.5281/zenodo.14791071

* Indicates equal contribution / co-first authorship. Underlined author denotes candidate.

RESEARCH EXPERIENCE

Doctoral Fellow — Project “Multimodal AI for Computed Tomography” Mar 2026 – Present
Millennium Institute iHealth · Pontificia Universidad Católica de Chile Santiago, Chile

- Train end-to-end 3D vision-language segmentation models on full CT volumes: given a volume and a free-text finding from its report, the model predicts the corresponding 3D mask (ReXGroundingCT benchmark, MICCAI 2026).
- Built a multi-host distributed training and evaluation framework (SLURM clusters at iHealth and CENIA, mixed precision) with reproducible cross-host experiment tracking; public release pending.
- Research line: learning under imperfect supervision — partial voxel-level annotation (positive-unlabeled learning), image-vs-text reliance measured through counterfactual input interventions, and long-tailed finding categories.

NLP Engineer — FONDEF Project VIU23P0082 (AIMA) Apr 2024 – Jan 2025
Universidad de Valparaíso Valparaíso / Santiago, Chile

- **Clinical information extraction with RAG (BMC MIDM 2026, WoS Q1, co-first author):** developed the methodology and ran the experiments for a retrieval-augmented framework that uses annotated reports as structured in-context demonstrations for joint NER and relation extraction on 453 Spanish mammography reports. GPT-4.1-mini (3-shot) reached 0.93 strict-NER F1 vs. 0.97 for fine-tuned BETO, with a robustness study under annotation noise.
- **Contextual laterality extraction (IEEE ICPRS 2024, co-first author):** laterality span categorization with boundary-correction heuristics, reaching 0.77 F1 (P 0.83 / R 0.72) with GPT-4o few-shot on 377 clinical reports.
- Packaged NLP inference services as deployment-ready Docker containers designed for PACS/RIS interoperability.

Undergraduate Thesis Researcher — FONDEF Project IT21I0019 (Proximity) Mar 2024 – Jul 2024
Universidad Técnica Federico Santa María (UTFSM) Santiago, Chile

- **LLM weak supervision for report annotation (AIME 2025, first author, oral):** curated 42,605 Spanish chest CT reports from HCUCH, released on Zenodo, plus a 217-report expert-annotated set (DR217).
- Distilled GPT annotations into Spanish BETO student models (0.884 Micro-F1). Large-scale weak supervision significantly outperformed scarce expert supervision (0.87 vs. 0.60 F1, Wilcoxon $p = 6e-8$), and student models showed higher inter-model reliability (Fleiss $\kappa \geq 0.78$) than the teacher LLM ($\kappa = 0.40$).

Research Intern Dec 2023 – Mar 2024
Millennium Institute for Intelligent Healthcare Engineering (iHealth) Santiago, Chile

- Benchmarked zero-shot and few-shot LLMs against fine-tuned transformers (Hugging Face, spaCy) for clinical NER on mammography notes.
- Implemented automated structured clinical information extraction and prompt-chaining pipelines using LangChain.

Software Developer & Assistant Researcher — FONDEF ID19I10023 (ALPACS) Jun 2021 – Dec 2021
Universidad Técnica Federico Santa María (UTFSM) Santiago, Chile

- Developed TypeScript/Node.js/Docker microservices for patient de-identification and DICOM metadata indexing of chest CT scans within the ingestion pipeline of an interoperable (DICOM, HL7/FHIR) medical imaging repository; contributed to the project's first integrated implementation and its closing functional prototype (milestones 4–5).

PARTICIPATION IN FUNDED RESEARCH PROJECTS

- **Millennium Science Initiative (ANID ICN2021_004):** *Millennium Institute iHealth* — Doctoral Fellow (2026 – present); Research Intern (2023 – 2024).
- **FONDEF VIU23P0082 (ANID):** *"AIMA: AI Assistant for Mammography"* — NLP Engineer (2024 – 2025). Clinical information extraction and RAG.
- **FONDEF IT21I0019 (ANID):** *"Proximity: Content-Based Medical Image Retrieval"* — Undergraduate Thesis Researcher (2024). LLM weak supervision for report annotation.
- **FONDEF ID19I10023 (ANID):** *"ALPACS: Interoperable Medical Imaging Repository"* — Developer & Assistant Researcher (2021). De-identification and DICOM metadata microservices.

FELLOWSHIPS, SCHOLARSHIPS & AWARDS

- **Doctoral Fellowship**, Millennium Institute iHealth (ANID Millennium Science Initiative), 2026 – 2027.
- **UC Doctoral Scholarship (VRIP)**, tuition and stipend, Pontificia Universidad Católica de Chile, 2026 – present.
- **UTFSM Master's Scholarship**, tuition and stipend, 2025.
- **Conference travel support** for AIME 2025 (Pavia, Italy): UTFSM conference scholarship and AC3E (ANID-Basal AFB240002).
- **Academic Distinction**, Computer Science Engineering, UTFSM, 2024.

TECHNICAL SKILLS & COMPETENCIES

Medical AI & Deep Learning: PyTorch, Hugging Face, 3D vision-language models, text-conditioned 3D medical segmentation, weak supervision & distillation, PEFT / LoRA, clinical NER & information extraction, RAG (LangChain), spaCy, scikit-learn.

HPC & Training Infrastructure: SLURM clusters, multi-GPU distributed training, mixed-precision training, reproducible experiment tracking, NIFTI / DICOM processing.

Languages & Data: Python (NumPy, Pandas, SciPy, Matplotlib), SQL, TypeScript / JavaScript, C / C++, R.

DevOps & Tools: Docker, Linux / Bash, Git, GitHub Actions (CI/CD), REST APIs (FastAPI, Flask, Express.js).

Languages: **Spanish:** native · **English:** C1 (Versant English Test, 2026).