



Machine-Learning-Forscherin
mit Schwerpunkt NLP und
angewandten
Embedding-Methoden, in der
finalen Phase der Promotion
und strebt den Übergang in
eine Position als Applied
Scientist / ML Engineer in der
Industrie an.

INFO & KONTAKT

📞 0177 3948481
✉️ ritagonzalez97@gmail.com
🌐 ritagonmar.github.io
📍 Nauklerstraße 11
72074 Tübingen

🐙 [ritagonmar](https://github.com/ritagonmar)

🌐 [in/ritagonmar/](https://in.ritagonmar/)

🏠 [Rita González-Márquez](#)

FACHBEREICH

Maschinelles Lernen
NLP
LLMs
KI-Agenten
Text embeddings
RAG-systeme
Retrieval
Self-supervised learning
Contrastive learning
Neighbor embedding
Visualisierungsmethoden

Rita Gonzalez Marquez

BERUFSERFAHRUNG

Wissenschaftliche Mitarbeiterin

04.2022 - heute

Hertie Institute for AI in Brain Health

- Interaktive Visualisierung von 21 Mio. wissenschaftlichen Publikationen zur Trendidentifikation in der Biomedizin.
- Benchmarking und Training von LLMs für Text-Embeddings (Retrieval, Nearest-Neighbor-Anwendungen).
- Datengetriebene Schätzung der LLM-Nutzung in biomedizinischen Fachtexten.
- Lehre im Kurs „Neural Data Analysis“; Mitentwicklung der Übungsaufgaben.
- Betreuung von Masterarbeiten und Studierenden.

Forschungspraktikantin

04.2020 - 09.2020

Werner Reichardt Centrum für Integrative Neurowissenschaften

Skalierung bestehender Dimensionsreduktionsverfahren auf Millionen von Textdaten.

Forschungspraktikantin

02.2019 - 06.2019

Institute of Microelectronics of Sevilla (IMSE CSIC)

Untersuchung der Halbleiterphysik im Kontext von CMOS-Technologien.

AUSBILDUNG

Promotion in ML, Informatik

04.2022 - heute

*Universität Tübingen,
International Max Planck Research School for Intelligent Systems
„Learning representations of scientific text data“*

M. Sc. in Neural Information Processing

10.2019 - 12.2021

*Universität Tübingen,
International Max Planck Research School*

Note: 1.59

„Two-dimensional visualization of document libraries“

B. Sc. in Physik

09.2015 - 07.2019

Universität Sevilla, Universität Münster

Note: 1.43

„A study on fundamental physical limitations of CMOS technologies“

B. Sc. in Materialwissenschaft

09.2015 - 07.2018

Universität Sevilla

Parallelstudium, ohne Abschluss



Rita Gonzalez Marquez

Machine-Learning-Forscherin
mit Schwerpunkt NLP und
angewandten
Embedding-Methoden, in der
finalen Phase der Promotion
und strebt den Übergang in
eine Position als Applied
Scientist / ML Engineer in der
Industrie an.

KENNTNISSE

Python
Pytorch
Scikit-learn
Matplotlib
Uv
Git + GitHub
Docker
LaTeX

SPRACHEN

German	fließend (C1)
English	verhandlungssicher
Spanish	Muttersprache

SKILLS

Kommunikationsstärke
Selbstgesteuertes Lernen
Anpassungsfähigkeit
Selbstständigkeit
Ressourcenmanagement

STIPENDIEN UND AUSZEICHNUNGEN

Forschungsstipendium 09.2019 - 12.2019

Universität Sevilla

Projekt zur Untersuchung des Einsatzes neuartiger Nano-devices
für die Bildverarbeitung.

Erasmus Program 09.2017– 08.2018

Universität Münster

Internationales Mobilitätsstipendium.

Matricula de Honor 07.2015

Sevilla, Spain

Bestnote im spanischen Schulabschluss Bachillerato, verbunden
mit Erlass der Studiengebühren im ersten Hochschuljahr.

Comenius-Programm 07.2015

Aydin, Turkey

Mobilitätsstipendium zur Förderung der Zusammenarbeit
europäischer Bildungseinrichtungen

PUBLIKATIONEN

Holzwarth, **González-Márquez**, et al.
Most biomedical publications show signs of LLM-assisted
writing.
In Begutachtung, 2026.

González-Márquez, et al.
Cropping outperforms dropout as an augmentation strategy for
training self-supervised text embeddings.
Transactions on Machine Learning Research, 2026.

Kobak, **González-Márquez**, et al.
Delving into LLM-assisted writing in biomedical publications
through excess vocabulary.
Science Advances, 2025

González-Márquez, et al.
The landscape of biomedical research.
Patterns, 2024.

González-Márquez, et al.
Learning representations of learning representations.
ICLR Workshop DMLR, 2024.