

Martín Villalba



Köln, Germany



villalba@7c0h.com

EDUCATION

- Jan. 2013 **PhD in Computational Linguistics**
to Aug. 2019 *Saarland University, Germany*
- ▶ Researched, modeled and implemented models of user behavior using both traditional and neural Machine Learning techniques.
 - ▶ Taught classes on Language theory and modern NLP techniques.
 - ▶ Organized conferences, workshops, events and group activities.
- Mar. 2002 **Licentiate in Computer Science (M.Sc. equivalent)**
to Feb. 2011 *National University of Córdoba, Argentina*
- ▶ Selected courses: Artificial Intelligence, Probability and Statistics, Microcontrollers, Programming Paradigms, Operating System Implementation.
 - ▶ Thesis: Aspect-oriented web requirements engineering with model transformations.

WORK EXPERIENCE

- Nov. 2023 **Postdoctoral Researcher in Computational Linguistics**
to Present *University of Cologne, Germany*
- ▶ Application of NLP techniques and LLMs to the early detection of Psychosis.
 - ▶ Design, execution and analysis of crowdsourced multimedia experiments.
 - ▶ Use of Generative AI for the design of experiments.
- Jun. 2023 **Independent Consultant**
to Sep. 2023 *AllyMatch GmbH, Germany*
- ▶ Zero-Short classification of companies according to existing categorization.
 - ▶ Viability study on LLMs and Bert-based models over traditional ML methods.
- Feb. 2019 **NLP Expert Applied Mathematics**
to Dec. 2022 *Bayer AG, Germany*
- ▶ Development of system for classification of electronic laboratory notebooks.
 - ▶ Crowdsourcing of scientific judgments for highly-restricted data.
 - ▶ Cloud-based GUI development and deployment for document categorization.
 - ▶ Full-stack development and administration of web retrosynthesis platform.
- Oct. 2011 **R&D Internship on Natural Language Interfaces**
to Mar. 2012 *Silvergate Group – IBM, Argentina*
- Apr. 2007 **System Administration & Software Development**
to Feb. 2011 *Planning and Institutional Management Ministry, UNC, Argentina*
- Mar. 2011 **Engineering and Development of medical system for image reconstruction**
to Sep. 2011 *FiMe (Medical Physics), Argentina*
- Oct. 2009 **Independent Developer for VoIP system**
to Jan. 2010 *FastVoIP, The Netherlands*
- Nov. 2007 **Medical system development for classification of skin lesions**
to May. 2009 *FiMe (Medical Physics), Argentina*

PUBLICATIONS

ERP Signatures of Prediction Error in the Comprehension of Ambiguous Plural Pronouns in German

Derya Çokal, Markus Phillip, Gokben Konuk, Martín Villalba, Klaus von Heusinger, and Massimo Poesio
PsyArXiv Preprints (2025, under review)

Automated Detection of Referential Features in Schizophrenic Speech Using Large Language Models

Derya Çokal, Melike Filizer, Martín Villalba, Douglas Turkington, Nicol Ferrier, Klaus von Heusinger, Stuart Watson, Wolfram Hinzen, and Massimo Poesio
PsyArXiv Preprints (2025, under review)

What is the retest reliability of computationally extractable speech and language markers?

Derya Çokal, Martín Villalba, Rui He, Claudio Flores Palominos, Annkathrin Böke, Philipp Homan, Klaus von Heusinger, Joseph Kambeitz, and Wolfram Hinzen
PsyArXiv Preprints (2025, under review)

When a Dialog becomes a Monologue: A debate on custom-made literature with Generative AI

Maja T. Jerrentrup and Martín Villalba.
26th Annual Meeting of the SIGDial – August 2025, Avignon, France

Generating Diversity and Securing Completeness in Algorithmic Retrosynthesis

F. Mrugalla, C. Franz, Y. Alber, G. Mogk, M. Villalba, T. Mrziglod, and K. Schewior
Journal of Cheminformatics 17, Article number 72 (2025).

Generating Contrastive Referring Expressions

Martín Villalba, Christoph Teichmann, and Alexander Koller
55th Annual Meeting of the ACL – August 2017, Vancouver, Canada

The Impact of Listener Gaze on Predicting Reference Resolution

Nikolina Koleva, Martín Villalba, Maria Staudte, and Alexander Koller
53rd Annual Meeting of the ACL – July 2015, Beijing, China

Predicting the Resolution of Referring Expressions from User Behavior

Nikos Engonopoulos, Martín Villalba, Ivan Titov, and Alexander Koller
EMNLP 2013 – October 2013, Vancouver, Canada

Interpreting Natural Language Instructions Using Language, Vision, and Behavior

Luciana Benotti, Tessa Lau, and Martín Villalba
ACM T.I.I.S. – Multiple Modalities in Interactive Systems and Robots. October 2014, Volume 4 Issue 3

Corpus-based Interpretation of Instructions in Virtual Environments

Luciana Benotti, Tessa Lau, Julián Cerruti, and Martín Villalba
50th Annual Meeting of the ACL – April 2012, Jeju, Korea

Inference of Strategic Points in Virtual Worlds (Spanish)

Luciana Benotti and Martín Villalba
Argentinean workshop on videogames (WAVI) – October 2011, Buenos Aires, Argentinien

Aspect-Oriented web requirements engineering with model transformations (Spanish)

Juan Durán and Martín Villalba
Latin-American Conference on Informatics (CLEI) – October 2010, Asunción, Paraguay

SELECTED PUBLIC PRESENTATIONS

"Ich hab's auf Parship versucht und auf Tinder ...": eine Veranstaltung der R-KOM

A presentation on the impact of Generative AI and Chatbots for therapy and dating.

#digit-R Veranstaltung – October 2025, Regensburg, Germany.

Predicting new Chemistry: Impact of High-Quality training data on Prediction of Reaction Outcomes

A presentation on the results obtained when applying a specific dataset to a neural retrosynthesis model.

Pistoia Alliance Virtual Conference, 2021.

Bringing your NLP research to the World: Good practices on sharing your research with end-users

A talk to new researchers on good practices for writing code that can be easily shared with other researchers.

LatinX Social – NAACL 2021.

VOLUNTEERING

Jan. 2017 **System Administrator (2019) and Conference Organizer**

to Dec. 2023 *Association for Computational Linguistics (ACL)*

- ▶ Installation, daily maintenance, and upgrade of the ACL online Anthology.
- ▶ Member of several Virtual Infrastructure Committees.

Sep. 2012 **Workshop Organizer**

to Jun. 2014 *Young Researchers' Roundtable on Spoken Dialog Systems – Endorsed by SIGDial*

- ▶ Contact with sponsors, reviewer, and graphic designer.

TECHNICAL SKILLS

Software development

Advanced: Python, Java, PHP.

Professional: HTML/CSS, SQL, R, JavaScript.

Knowledgeable: C/C++, Rust, Perl, Haskell.

Other tools: Git, Jenkins, REST.

Natural Language Processing (NLP)

Natural Language Generation (NLG), Automatic Summarization, Large Language Models (LLMs), and Neural Networks/Deep Learning (PyTorch).

Scientific Research

Knowledge of AI, Crowdsourcing experiments, Statistics, Data Processing and Visualization.

System Administration

Professional experience with Linux, Bash, Docker, Azure Cloud, Web servers (Apache, NGINX), Email (Exim, Dovecot), Databases (PostgreSQL, MySQL).

OTHER SKILLS

Public Communication

Experienced in technical writing (LaTeX), scientific presentations, and public speaking.

Graphic Design

Experience in Poster, Booklet, Logo, and Web Design, Digital Illustration, and Generative AI.

LANGUAGES

English

Full Proficiency – Certified by ETS TOEFL

German

B1 level – Certified by Goethe-Institut.
B2 conversational level.

Spanish

Native proficiency