



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG



LLMs im FID Romanistik: Webkategorisierung und Expertenfinder

Robert Geislinger

29.05.2026



Foto: UHH/Lutsch

Projekte

Crawled Pages

+ Crawl URL

34 records - Page 1 / 7

Search by URL or keyword...

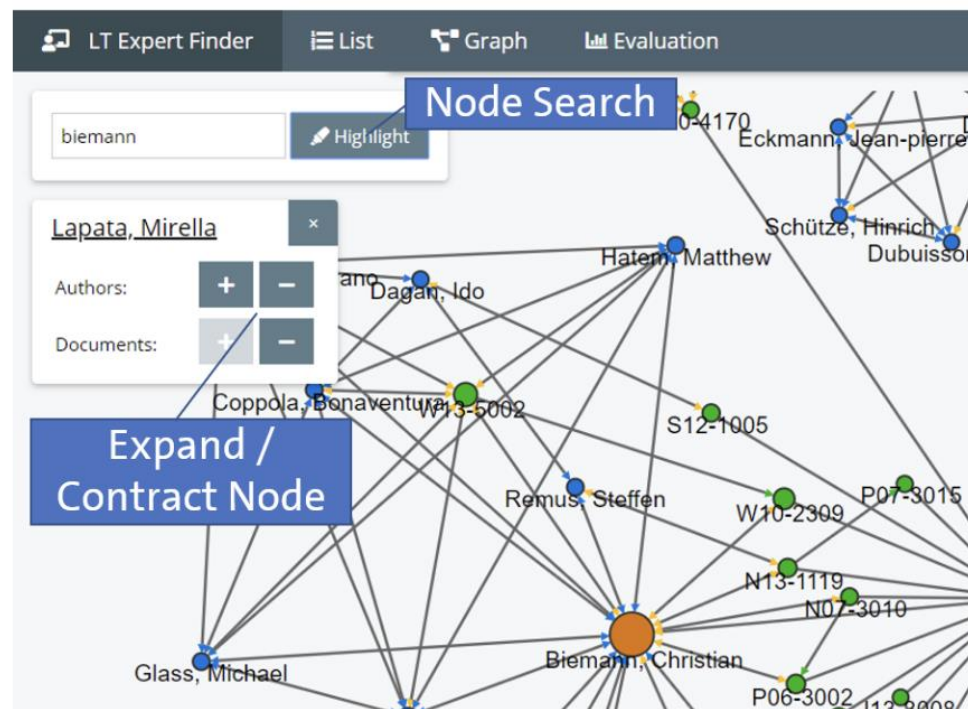
Newest first All liked/disliked pages All edited/not edited pages All reviewed/not reviewed pages

449.7 - Occitan (historical, geographic, social varieties)

Source: input file | Crawled: 15.02.2026

Overall page feedback

		👍 0	👎 0
url	https://www.cieldoc.com/lourtougrafi-prouvençalo-jules-ronjat/	👍 0	👎 0
meta_title	L'Ourtougràfi Prouvençalo, Jules RONJAT Ciel d'Oc - Trésor de la langue d'Oc (crawled from LLM)	👍 0	👎 0
meta_description	No value found (crawled from LLM)	👍 0	👎 0
abstract_en	This page discusses the Ourtougràfi Prouvençalo, an orthographic system for the Provençal language developed by Jules Ronjat. It explains the history and significance of this writing system within the context of the Occitan language. The page is part of Ciel d'Oc, a center dedicated to the Occitan language and culture. (generated/translated) (crawled from LLM)	👍 0	👎 0
abstract_de	Diese Seite behandelt die Ourtougràfi Prouvençalo, ein orthographisches System für die provenzalische Sprache, entwickelt von Jules Ronjat. Es wird die Geschichte und Bedeutung dieser Schreibweise im Kontext der okzitanischen Sprache erläutert. Die Seite ist Teil des Ciel d'Oc, einem Zentrum für die okzitanische Sprache und Kultur. (generated/translated) (crawled from LLM)	👍 0	👎 0



Webcrawler

- Pflegen einer kuratierten Linkliste
- Verarbeiten von Webquellen mit LLMs
- Automatische Erstellung von Metadaten
- Möglichkeit der händischen Pflege

Von der Website zu Metadaten



UHH → MIN → Department of Informatics → Institutes → Research groups
BA/MA Theses → **Completed Theses**

COMPLETED THESES

This page lists M.Sc./B.Sc. theses that have been supervised by the LT group. Theses are made available online with individual student's consent.

2025

- 04/2025 Anton Gustav Trappe. [Topic Modeling and Stance Detection for Accessible Legislative Sessions](#). B.Sc. Thesis, Universität Hamburg
- 01/2025 Robert Günzler. [Exploring Large Language Models in Argument Retrieval for Content- and Demographic-Relevance](#). B.Sc. Thesis, Universität Hamburg
- 01/2025 Ahmad Shallouf. Comparative Understanding and Generation Evaluation (CompUGE) A Benchmark for Comparative Question Answering. M.Sc. Thesis, Universität Hamburg
- 01/2025 Lennart Joshua Moritz. Towards Cross-modal Media Retrieval for Discourse Analysis. M.Sc. Thesis, Universität Hamburg
- 01/2025 Hina Samreen. Sentiment Analysis for German Language Texts: Data Collection, Model Development, Interpretability, and Evaluation. M.Sc. Thesis, Universität Hamburg

2024

- 12/2024 Levente Schoenherr. GerLUE: The German Language Understanding Evaluation Benchmark for Large Language Models. M.Sc. Thesis, Universität Hamburg
- 11/2024 Neil Hinrichs. Combined Generation of Images and Texts for Image Macro Memes. B.Sc. Thesis, Universität Hamburg
- 10/2024 Lennart Danielsen. Empowering KGQA Systems through Automated Corpus Generation: From Graph Traversal to SPARQL Queries. M.Sc. Thesis, Universität Hamburg
- 09/2024 Lars Lamberti. [Evaluating Dimension Based Configurable Similarity Search](#). B.Sc. Thesis, Universität Hamburg
- 08/2024 Daniel Djahangir. Retrieval Augmented Generation of Character Profiles. M.Sc. Thesis, Universität Hamburg
- 07/2024 Robert Geislinger. [Enhancing Sentiment Analysis: Model Comparison, Domain Adaptation, and Lexicon Evolution in German Data](#). M.Sc. Thesis, Universität Hamburg
- 07/2024 Angel Ontiveros. [Development and Evaluation of an IT Tool for ESG Compliance Consulting for Companies Using LLMs](#). M.Sc. Thesis, Universität Hamburg



```
[
  {
    "authors": [
      "Anton Gustav Trappe"
    ],
    "title": "Topic Modeling and Stance Detection for Accessible Legislative Sessions",
    "year": "2025",
    "error": false
  },
  {
    "authors": [
      "Robert Günzler"
    ],
    "title": "Exploring Large Language Models in Argument Retrieval for Content- and Demographic-Relevance",
    "year": "2025",
    "error": false
  },
  {
    "authors": [
      "Ahmad Shallouf"
    ],
    "title": "Comparative Understanding and Generation Evaluation (CompUGE) A Benchmark for Comparative Question Answering",
    "year": "2025",
    "error": false
  },
  {
    "authors": [
      "Lennart Joshua Moritz"
    ],
    "title": "Towards Cross-modal Media Retrieval for Discourse Analysis",
    "year": "2025",
    "error": false
  },
  {
    "authors": [
      "Hina Samreen"
    ],
    "title": "Sentiment Analysis for German Language Texts: Data Collection, Model Development, Interpretability, and Evaluation",
    "year": "2025",
    "error": false
  },
  {
    "authors": [
      "Daniel Djahangir"
    ],
    "title": "Retrieval Augmented Generation of Character Profiles",
    "year": "2024",
    "error": false
  },
  {
    "authors": [
      "Robert Geislinger"
    ],
    "title": "Enhancing Sentiment Analysis: Model Comparison, Domain Adaptation, and Lexicon Evolution in German Data",
    "year": "2024",
    "error": false
  },
  {
    "authors": [
      "Angel Ontiveros"
    ],
    "title": "Development and Evaluation of an IT Tool for ESG Compliance Consulting for Companies Using LLMs",
    "year": "2024",
    "error": false
  }
]
```

Oberfläche

Crawled Pages

+ Crawl URL

◀ Prev

Next ▶

34 records • Page 1 / 7

Search by URL or keyword...

Search

Newest first ▾

All liked/disliked pages ▾

All edited/not edited pages ▾

All reviewed/not reviewed pages ▾

449.7 - Occitan (historical, geographic, social varieties) ▾

Source: input file | Crawled: 15.02.2026

Edit

Overall page feedback

👍 0 👎 0

url

<https://www.cieldoc.com/lourtougrafi-prouvençalo-jules-ronjat/>
read-only

👍 0 👎 0

meta_title

L'Ourtougràfi Prouvençalo, Jules RONJAT | Ciel d'Oc - Trésor de la langue d'Oc
(crawled from LLM)

👍 0 👎 0

meta_description

No value found
(crawled from LLM)

👍 0 👎 0

abstract_en

This page discusses the Ourtougràfi Prouvençalo, an orthographic system for the Provençal language developed by Jules Ronjat. It explains the history and significance of this writing system within the context of the Occitan language. The page is part of Ciel d'Oc, a center dedicated to the Occitan language and culture.
(generated/translated)
(crawled from LLM)

👍 0 👎 0

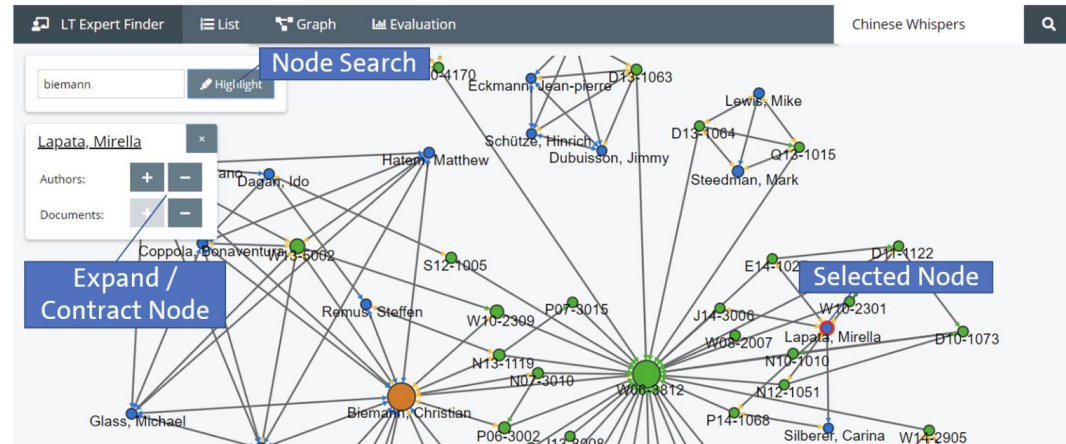


Expertenfinder

Forschende an der Universität Hamburg
mit Schwerpunkt Romanistik

Publikationen über französische
Grammatik des 19. Jahrhunderts

Chris Biemann



Quellen

- Fachspezifischen Quellen
 - romanistik.de
 - Forschungsgruppen
- Öffentlichen Quellen
 - OpenAlex
 - Gemeinsame Normdatei (GND)
 - ORCID
- Ausgehende Links

Arten der Quellen

Forschung

PROJEKTE

PUBLIKAT

Sie finden hier Neuersche
romanistischer Fachzeitsc

SORTIERUNG: ^ JA

✓ Belén Almeida, Del

✓ Romanische Forsch

✓ Stephan Leopold, N

✓ Jörg Dünne: Quiltrc

✓ Prof. Dr. Susanne G
siècle. . Paris 2026

✓ Angela Fabris, Stefi

HOME

PEOPLE

RESEARCH

PUBLICATIONS

RESOURCES

TEACHING

Q

LANGUAGE TECHNOLOGY

Photo: alt left

UHH → MIN → Department of Informatics → Institutes → Research groups → Language Technology Group → Home

HOME

Language Technology

Welcome to the Language Technology Group

We are researching on all aspects of technologies for the processing of natural (human) language, and apply language technology in practice. In particular, we are interested in unsupervised, large-scale methods for natural language semantics.

Statistical Semantics

The Language Technology research group explores statistical methods that reflect natural language

Home

Job Opportunities

News

Directions

NEWS

18 MAY 2026

ACL 2026 accepts 3 papers from LT members

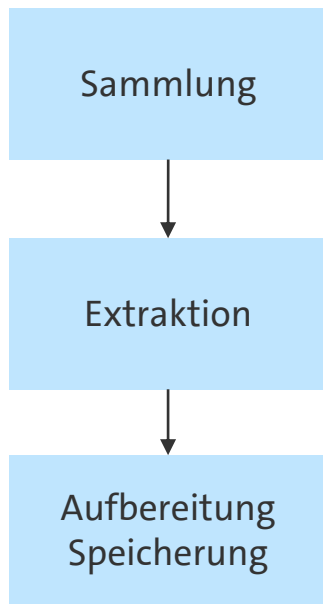
UHH

Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

8

Verarbeitung



- Analyse der Seed-Liste und sammeln zusätzlicher Quellen
- Filtern der Ergebnisse (Relevanz, Deduplikation)
- Extraktion mit LLMs
- RAG & Graphdatenbank

Beispiele - ARDIAS

ARDIAS

Works

Hello,

Publications

Co-author Network

Research Focus

Sort by

Showing 10 results of 186. Fetched in 55ms.

Carbon quantum dots and their applications

Shi Ying Lim, Wei Shen, Zhiqiang Gao, 2015 - **Chemical Society Reviews**

This review covers the progress in research and development of carbon quantum dots their applications chemical sensing, biosensing, bioimaging, nanomedicine, photocatalysis electrocatalysis.

[Read more](#)

11 3097

SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation

Wei Shen, Shuai Le, Yan Li, Fuquan Hu, 2016 - **PLOS ONE**

FASTA and FASTQ are basic ubiquitous formats for storing nucleotide protein sequences. Common manipulations of FASTA/Q file include converting, searching, filtering, deduplication, splitting, shuffling, sampling. Existing tools only implement some these manipulations, not particularly efficiently, available certain operating systems. Furthermore, the complicated installation process...

[Read more](#)

[Open Access](#) 11 722

Wei Shen
Jiangsu University of Science and Technology , CN

186
 9704
Openalex: A2602767326
Orcid: 0000-0001-8159-5728
Mag: 2602767326
Scopus: authorDetails.url?authorID=57001188100&partnerID=MN8TOARS

Beispiele - ARDIAS

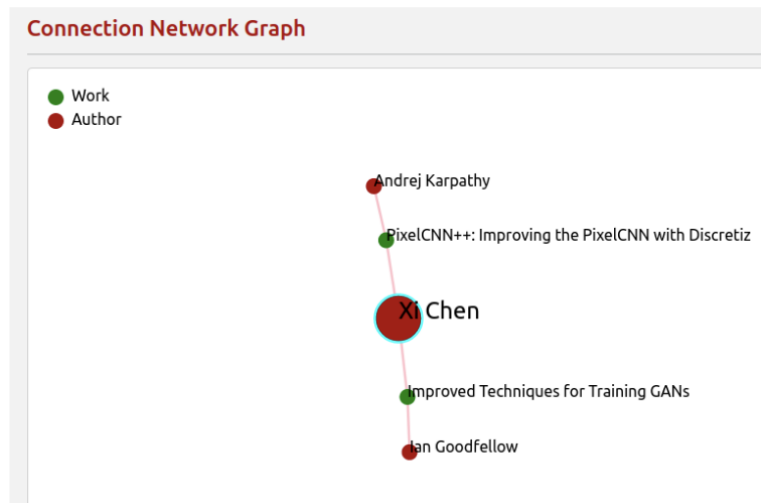


Figure 3: Shortest authorship path between two authors

Login

Full Name

Affiliation

ORCID

OpenAlex ID

Find

Lei Zhang

Xi'an Jiaotong University, CN

0000-0001-5301-4658

Feng Zhang

National Center for Biotechnology Information, US

0000-0001-5254-2264

Wei Zhang

Chongqing Institute of Green and Intelligent Technology, CN

0000-0001-5465-3267

Jian Zhang

Hangzhou Dianzi University, CN

0000-0001-5000-6142

Figure 4: Author disambiguation for first-time login.

Beispiele - LT Expertfinder

LT Expert Finder | List | Graph | Evaluation

named entity recognition

Advanced Search

Ranking

#	Global Citations	Model 2	Relevance Propagation	Weighted Relevance Propagation
1	Och, Franz Josef	Manning, Christopher D.	Grishman, Ralph	Grishman, Ralph
2	Manning, Christopher D.	Sasano, Manabu	Manning, Christopher D.	Manning, Christopher D.
3	Klein, Dan	Utsuro, Taketoshi	Tjong Kim Sang, Erik F.	Tjong Kim Sang, Erik F.
4	Ney, Hermann	Finkel, Jenny Rose	Suczeran, Silvio	Suczeran, Silvio
5	Collins, Michael John	Tsujii, Jun'ichi	Tsujii, Jun'ichi	Tsujii, Jun'ichi
6	Knights, Kevin	Dien, Dinh	Meulder, Fien	Meulder, Fien
7	Callison-Burch, Chris	Ngo, Quoc Hung	Roth, Dan	Roth, Dan
8	Pereira, Fernando	Winawarter, Werner	Yarowsky, David	Yarowsky, David

Various Expert Finding Methods

LT Expert Finder | List | Graph | Evaluation

Manning, Christopher D. is assumed to be an expert for Named Entity Recognition because of the following publications:

Statistics

Statistics	Local	Global
H-Index		
Publications	22	140
Citations	153	2975
Collaborations	41	171

Documents as Evidence

File	Relevance	Title
N09: 1037	1.209652	Named Entity Recognition
W12: 4403	4.8378321573012514E-9	Accurate Unsupervised Joint Named-Entity Extraction from Unaligned Parallel Text
D09-1015	3.460151019347847E-9	Nested Named Entity Recognition
N13: 1006	2.9665168215543155E-9	Named Entity Recognition with Bilingual Constraints

Relevance

Any



Questions?

