

Rajaswa Patil


Research Fellow | Microsoft PROSE

[Website](#) [@ E-mail](#) [Github](#) [Google Scholar](#) [Twitter](#) [LinkedIn](#)

Experience

Present Oct 2022	Microsoft PROSE [🌐] <i>Research Fellow Advisors: Gustavo Soares, Sumit Gulwani</i>	Bengaluru, India
	<ul style="list-style-type: none">Working on developing specialized LLM-based code suggestion systems in collaboration with the Visual Studio and .NET teams.	
Sep 2022 Jul 2021	Tata Consultancy Services (TCS) - Research Deep Learning & AI Group [🌐] <i>Pre-Doctoral Research Fellow Advisors: Gautam Shroff, Lovekesh Vig</i>	Pune, India
	<ul style="list-style-type: none">Worked on broad research topics: NLP for Programming, Text-to-SQL Semantic Parsing, and Chart Question-Answering.	
Jun 2021 Apr 2021	Cognizer AI [🌐] <i>Data Scientist (NLP) - Intern Manager: Vineeth T M</i>	Bengaluru, India
	<ul style="list-style-type: none">Developed APIs & data resources for Question-generation, Grammar-correction, and Text-simplification for the Cognizer Genius platform.	
Jun 2021 Apr 2021	Max Planck Institute for Human Cognitive & Brain Sciences [🌐] <i>Summer Research Intern Advisor: Lars Meyer</i>	(Remote) Leipzig, Germany
	<ul style="list-style-type: none">Experimented with Combinatory Categorical Grammar - CCG parsers (left-branching-based incremental parsing) for modelling human-like language processing.	
Mar 2021 Aug 2020	IIIT Delhi Multimodal Digital Media Analysis (MIDAS) Lab [🌐] <i>Research Assistant (Bachelor Thesis) Advisors: Rajiv Ratn Shah, Yaman Kumar Singla</i>	Delhi, India
	<ul style="list-style-type: none">Worked on modelling information structure and coherence in spontaneous spoken discourse.Developed annotation pipelines & data resources for disfluency and content relevance & sufficiency for ESL speech data.	
Dec 2019	<i>Research Intern Advisors: Rajiv Ratn Shah, Yaman Kumar Singla</i>	
	<ul style="list-style-type: none">Contributed to the Berlitz SLTI automatic language proficiency - speech scoring system.	
May 2021 Aug 2019	BITS Pilani (Goa Campus) Cognitive Neuroscience Lab [🌐] <i>Research Assistant Advisor: Veeky Baths</i>	Goa, India
	<ul style="list-style-type: none">Developed a benchmark for syntactic evaluation of multilingual language models for Indic languages.Lead projects on Mental Lexicon Development and EEG-based brain imagery.Organized and volunteered for various events under the ERASMUS iBRAIN project.Designed and implemented an experiment for EEG data collection (6 subjects with 200 trials each) to inspect the role of low-level visual representations (BA-18 & BA-19) in spoken word processing.	
May 2021 Dec 2019	Indian Institute of Technology, Bombay Department of CS&E [🌐] <i>Summer Research Intern Advisor: Ganesh Ramakrishnan</i>	Mumbai, India
	<ul style="list-style-type: none">Implemented a NLP pipeline for open-domain opinion-mining & aspect-based sentiment-analysis in natural language text.Built various text-based utilities for entity resolution, topic extraction, theme extraction, and lexicon-based sentiment polarity extraction.	

Education

Jun 2021 Aug 2017	Birla Institute of Technology and Science (BITS) Pilani <i>Bachelor in Engineering, Electrical & Electronics</i> CGPA: 8.26/10.00 Thesis: Computational Modelling for Coherence in Spoken Discourse  Study in Advanced Topics: [1] Meta Learning; [2] Computational Neuroscience	Goa, India
----------------------	--	------------

Publications

C=Conference, P=Poster, S=In Submission, T=Thesis, W=Workshop (* = equal contribution)

- [C.3] **Program Synthesis for Complex QA on Charts via Probabilistic Grammar Based Filtered Iterative Back-Translation**
Shabbirhussain Bhaisaheb, Shubham Paliwal, [Rajaswa Patil](#), Manasi Patwardhan, Lovekesh Vig, Gautam Shroff
The 17th Conference of the European Chapter of the Association for Computational Linguistics EACL '23
- [W.6] **Exploring Dimensions of Generalizability and Few-shot Transfer for Text-to-SQL Semantic Parsing**
[Rajaswa Patil](#), Manasi Patwardhan, Shirish Karande, Lovekesh Vig, Gautam Shroff
Transfer Learning for NLP Workshop 2022 NeurIPS '22
- [C.2] **Probing Semantic Grounding in Language Models of Code with Representational Similarity Analysis**
Shounak Naik*, [Rajaswa Patil](#)*, Swati Agarwal, Veeky Baths
The 18th International Conference on Advanced Data Mining and Applications ADMA '22
- [T.1] **Towards Modelling Coherence in Spoken Discourse** 
[Rajaswa Patil](#)*, Yaman Kumar Singla*, Rajiv Ratn Shah, Mika Hama, Roger Zimmermann
Bachelor Thesis BITS Pilani
- [W.5] **Towards Using Diachronic Distributed Word Representations as Models of Lexical Development** 
Arijit Gupta*, [Rajaswa Patil](#)*, Veeky Baths
The 60th Annual Meeting of the Association for Computational Linguistics - Student Research Workshop ACL '22
- [C.1] **DRIFT: A Toolkit for Diachronic Analysis of Scientific Literature**  
Abheesht Sharma*, Gunjan Chhablani*, Harshit Pandey*, [Rajaswa Patil](#)
The 2021 Conference on Empirical Methods in Natural Language Processing (Virtual) EMNLP '21
- [W.4] **Vyākaraṇa: A Colorless Green Benchmark for Syntactic Evaluation in Indic Languages**  
[Rajaswa Patil](#), Jasleen Dhillon, Siddhant Mahurkar, Saumitra Kulkarni, Manav Malhotra, Veeky Baths
The First Workshop on Multilingual Representation Learning (Virtual) EMNLP '21
- [W.3] **BPGC at SemEval-2020 Task 11: Propaganda Detection in News Articles with Multi-Granularity Knowledge Sharing and Linguistic Features Based Ensemble Learning**  
[Rajaswa Patil](#), Somesh Singh, Swati Agarwal
The Fourteenth Workshop on Semantic Evaluation (Virtual) COLING '20
- [W.2] **CNRL at SemEval-2020 Task 5: Modelling Causal Reasoning in Language with Multi-Head Self-Attention Weights Based Counterfactual Detection**  
[Rajaswa Patil](#), Veeky Baths
The Fourteenth Workshop on Semantic Evaluation (Virtual) COLING '20
- [W.1] **LRG at SemEval-2020 Task 7: Assessing the Ability of BERT and Derivative Models to Perform Short-Edits Based Humor Grading**  
Siddhant Mahurkar*, [Rajaswa Patil](#)*
The Fourteenth Workshop on Semantic Evaluation (Virtual) COLING '20
- [P.2] **Word2Brain2Image: Visual Reconstruction from Spoken Word Representations**  
Ajay Subramanian, [Rajaswa Patil](#), Veeky Baths
The Sixth Annual Conference of the Association for Cognitive Science, 2019 ACCS '19
- [P.1] **Citta: A Lite Semantic Recommendation Framework for Digital Libraries**  
[Rajaswa Patil](#), Siddhant Mahurkar
The NDLI-UNESCO International Symposium on Knowledge Engineering for Digital Library Design 2019 KEDL '19

Program Committee	DL4C (ICLR'23), DL4C (ICLR'22), MRL (EMNLP'21)
Reviewer	ACL-SRW'23, NAACL-SRW'22, ACL-SRW'22, ICON'21, EAAI'21*
Volunteer	ACL'21 (Social Media), Neuromatch-DL'21 (Tutorial Creator)

Talks

“Vyākaraṇa: A Colorless Green Benchmark for Syntactic Evaluation in Indic Languages”

- > MRL EMNLP 2021 Presentation [📺] November 2021 (Remote)
- > Invited Talk - BITS Goa September 2021 (Remote)
- > ACL-IJCNLP SRW 2021 Presentation [📺] August 2021 (Remote)

“DRIFT: A Toolkit for Diachronic Analysis of Scientific Literature”

- > EMNLP 2021 Presentation [📺] November 2021 (Remote)
- > Invited Talk - Aon Research October 2021 (Remote)

“Syntactic Processing and Language Comprehension”

- > Round Table Discussion - CCCP Conference [📺] December 2020 (Remote)

“Getting Started with ML/NLP/Undergraduate Research”

- > Getting Started with ML - ACM BITS Goa Chapter [📺] July 2021 (Remote)
- > Thesis Talks - CTE BITS Goa [📺] May 2021 (Remote)
- > Introduction to NLP Research - TIP BITS Goa September 2020 (Goa, India)

Teaching and Leadership Roles

Language Research Group (LRG) - BITS Goa <i>Founder & Lead</i> [🌐]	Aug'19 - Present
<ul style="list-style-type: none"> > Established a student research group for NLP & Psycholinguistics > The group has currently published 11 papers & three projects 	
Birla Institute of Technology and Science (BITS) Pilani, Goa Campus <i>Teaching Assistant</i>	Aug'20 - May'21
<ul style="list-style-type: none"> > BITS F315: Introduction to Cognitive Neuroscience > BITS C382: Neuroscience and AI 	
QSTP & CTE (student-run clubs at BITS Goa) <i>Instructor</i>	May'19 - Dec'19
<ul style="list-style-type: none"> > Introduction to Machine Learning & Deep Learning > Deep Learning with PyTorch 	

Software and Open Source Contributions

- > **ML4Code**: Machine Learning for Big Code and Naturalness - [commits](#), [website](#)
- > **Semantic Parsing**: Modeling and Data resources for Semantic Parsing - [commits](#), [website](#)
- > **DecepticoNLP**: Python Library for Robustness Monitoring and Adversarial Debugging of NLP models - [commits](#)

References

- > Dr. Sumit Gulwani *Partner Research Manager - Microsoft* [🌐]
- > Dr. Gautam Shroff *Senior Vice President, Chief Scientist, and Head - TCS Research* [🌐]
- > Dr. Gustavo Soares *Principal Research Manager - Microsoft* [🌐]
- > Dr. Manasi Patwardhan *Senior Scientist - TCS Research* [🌐]
- > Vineeth T M *Data Science Architect - Cognizer AI* [🌐]
- > Dr. Lars Meyer *Research Group Leader - Max Planck Research Group Language Cycles* [🌐]
- > Prof. Veeky Baths *Associate Professor - Dept. of Biological Sciences, BITS Goa* [🌐]
- > Prof. Rajiv Ratn Shah *Head of Department - Dept. of Human-Centered Design, IIIT Delhi* [🌐]
- > Prof. Swati Agarwal *Assistant Professor - Dept. of Computer Science & Information Systems, BITS Goa* [🌐]

