

# ACL 2025 Industry Track: Overview

**Georg Rehm**

Deutsches Forschungszentrum für  
Künstliche Intelligenz GmbH (DFKI)  
georg.rehm@dfki.de

**Yun Yao Li**

Adobe  
yunyaol@adobe.com

## Abstract

For the 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025), it was decided once again to organise a dedicated Industry Track. Similar to the main research track of the conference, the industry track attracted an unprecedented number of 421 paper submissions. In total, 453 reviewers and 19 area chairs participated in the evaluation of these papers. After a thorough, double-blind peer-review evaluation with three reviews for each submission followed by reviewer discussions and additional deliberations, 108 papers were selected for presentation at the ACL 2025 Industry Track. Large language models were front and center of almost all submissions with trustworthiness, domain-adaptation, retrieval-augmented generation, and agentic architectures – across domains such as medical, legal, and finance – being popular topics.

## 1 Introduction

Language technologies and their applications are an integral and critical part of our daily lives. Many of these technologies have their roots in academic and industrial research laboratories where researchers invented a plethora of algorithms, benchmarked them against shared datasets and perfected their performance to provide plausible solutions to real-world applications. While a controlled laboratory setting is vital for a deeper scientific understanding of the problems underlying language technologies and the impact of algorithmic design choices on their performance, transitioning the technology to real-world industrial strength applications raises a different, yet challenging, set of technical issues.

We acknowledge the challenges when adapting language technologies for building novel and robust real-world applications as the journey from theoretical research to practical deployment can be difficult. Challenges can include technical aspects

of system deployment and optimizing for efficiency, making informed design choices or methodological considerations of incorporating human feedback and oversight. The Industry Track provides a forum to address these multifaceted issues. We were seeking submissions that not only delve into research but also demonstrate the application of systems in real-world scenarios, irrespective of whether they involve proprietary data.

## 2 Call for Papers

We invited submissions describing innovations and implementations in all areas of speech and natural language processing (NLP) technologies and systems that are relevant to real-world applications. The primary focus of the ACL 2025 Industry Track was on papers that advance the understanding and demonstrate the effective handling of practical issues related to the deployment of language processing or language generation technologies, including those of large language models (LLMs), in non-trivial real-world systems. By “non-trivial real-world system” we mean an application deployed for real-world use, i. e., outside controlled environments such as laboratories, classrooms or experimental crowd-sourced setups, and that uses NLP and/or speech technology, even if not state of the art in terms of research. There was no requirement that the system be made by a for-profit company, but the users of the system are most likely outside the NLP research community.

This track provided an opportunity to highlight the key insights and new research challenges that arise from real-world implementations.

Relevant areas included system design, efficiency, maintainability and scalability of real-world applications, with topics including, but not limited to (in alphabetical order):

- Benchmarks and methods for improving the latency and efficiency of systems

- Continuous maintenance and improvement of deployed systems
- Efficient methods for training and inference
- Enabling infrastructure for large-scale deployment
- Handling unexpected user behaviour
- Human-in-the-Loop approaches to application development
- Implementation at speed, scale and low-cost
- Negative results related to real-world applications
- System combination

Novel applications and use cases, with topics including, but not limited to (in alphabetical order):

- Best practices and lessons learned
- Case studies, from design to deployment
- Description of an application or system
- Design of application-relevant datasets
- Development of methods under system constraints (model or data size)
- Novel, previously unsolved NLP problems and novel NLP applications

Methods for deployed systems, with topics including, but not limited to (in alphabetical order):

- Ethics, bias, fairness, harmlessness and trustworthiness in deployed systems
- Interpretability
- Interactive systems
- Offline and online system evaluation methodologies
- Online learning
- Robustness
- In addition, opinion/vision papers related to real-world applications were also welcome.

Submissions had to clearly identify one of the following three areas they fall into:

**Deployed** Must describe a system that solves a non-trivial real-world problem. The focus may include describing the problem related to actual use cases, its significance (against opportunity size, value proposition, and ideal end state), design/formulation of methods, tradeoff design decision for solutions, deployment challenges, and lessons learned.

**Emerging** Must describe the development of a system that solves a non-trivial real-world problem (it need not be deployed or even close, but

there needs to be evidence that this development is intended for real-world deployment). Papers that describe enabling infrastructure for large-scale deployment of NLP techniques also fall in this category.

**Discovery** Must include results obtained from NLP applications in real-world scenarios that result in actionable insights. These discoveries should reveal promising directions in their application areas, leading to further system or societal enhancements. For example, an actionable discovery from an analysis of call center transcripts may reveal that certain language choices negatively impact customer experience, leading to better training of service representatives and improved customer experience.

### 3 Submissions and Results

The call for Industry Track papers attracted an unprecedented number of 421 paper submissions. A total of 453 reviewers and 19 area chairs participated in the evaluation of these papers. After a thorough, double-blind peer-review evaluation with three reviews for each submission, we eventually selected a total of 108 articles for presentation within the Industry Track at ACL 2025, with 35 oral and 73 poster presentations.

### 4 Research Trends

Nearly all submissions (approx. 90%) revolve around LLMs, indicating the prevalence of their adoption in real-world applications. More specifically, we observe the following five research trends based on this year's submissions.

**Evaluation and Prompt Engineering** Many submissions focus on the evaluation of LLM responses and improving their quality through prompt engineering, reflecting a broader push toward trustworthiness and safety in outputs. Hallucination detection and mitigation are particularly popular among such submissions.

**Retrieval-Augmented Generation (RAG)** RAG remains dominant, indicating continued interest in bridging static LLM knowledge with dynamic external data, especially in enterprise use cases such as enterprise document QA and domain-specific knowledge mining.

**Domain Adaptation** Domain adaptation (e. g., finance, medical, legal) is prominent, with an emphasis on techniques such as fine-tuning and reinforcement learning, underscoring the commercial push to tailor general models for domain-specific performance.

### **Agentic Workflows and Multi-Agent Systems**

LLM-powered agents and multi-agent systems are being developed to automate workflows and enhance user experience. The growing focus on agent-based architectures indicates a sharp industry shift toward LLM-as-a-service ecosystems.

**Medical Applications** The medical domain is particularly popular among the submissions, covering a wide range of use cases from ICU monitoring, diagnostics, to medical coding, a sector with high impact and regulatory sensitivity.

With the growing adoption of LLMs and agent-based architectures, we expect that the above trends will continue and rapidly evolve in the near future.

## **5 Programme Co-Chairs**

- Georg Rehm, Deutsches Forschungszentrum für Künstliche Intelligenz GmbH and Humboldt-Universität zu Berlin, Germany
- Yunyao Li, Adobe, USA

### **Acknowledgments**

The ACL 2025 Industry Track Programme Co-Chairs would like to thank the authors of all Industry Track submissions as well as the reviewers and area chairs for their hard and dedicated work under very tight deadlines. We would also like to thank the General Chair and various ACL 2025 committees with which we interacted between the summer of 2024, when this endeavour started, and the summer of 2025, when we finally have been able to have the Industry Track at the ACL 2025 conference in Vienna, Austria. Finally, we would also like to thank our keynote speakers and the whole ACL team, especially Jennifer Rachford.

Georg Rehm was supported through the project NFDI for Data Science and Artificial Intelligence (NFDI4DS) as part of the non-profit association National Research Data Infrastructure (NFDI e. V.). The NFDI is funded by the Federal Republic of Germany and its states.