

# Abdelrahman Abdallah

PH.D. CANDIDATE · University of Innsbruck

Tyrol, Austria

✉ [abdelrahman.abdallah@uibk.ac.at](mailto:abdelrahman.abdallah@uibk.ac.at) | 🏠 [abdoelsayed2016.github.io](https://github.com/abdoelsayed2016) | 🔗 [abdoelsayed2016](https://github.com/abdoelsayed2016) | 🐦 [perdactor](https://twitter.com/perdactor)

## About

Abdelrahman Abdallah is a final-year Ph.D. candidate advised by Professor Adam Jatowt in the Department of Computer Science at the University of Innsbruck, Austria. His research interests primarily lie in natural language processing (NLP) and information retrieval (IR), focusing on advancing retrieval-augmented generation systems, document reranking with reasoning capabilities, and temporal question answering. In pursuit of more effective and efficient information access, he has worked on (1) developing LLM-based reranking and reasoning systems, (2) creating large-scale evaluation benchmarks and open-source tools for the research community, and (3) building multilingual QA resources for underrepresented languages.

## Education

### University of Innsbruck

Ph.D. in Computer Science

*Tyrol, Austria*

*Oct 2022–Present*

- Advisor: Adam Jatowt
- Focus: Information Retrieval, Document Reranking, Retrieval-Augmented Generation, Temporal QA
- Expected completion: June/July 2026

### Satbayev University

M.S. in Data Science and Machine Learning

*Almaty, Kazakhstan*

*Sep 2019– July 2021*

- Advisor: Daniyar Nurseitov
- Focus: Deep Learning, Computer Vision, NLP
- Thesis: Handwritten Recognition and Table Detection using Deep Learning

### Assiut University

B.S. in Computer Science

*Assiut, Egypt*

*Sep 2011–July 2015*

- Summa Cum Laude, Rank: 2st in the department

## Employment

### University of Innsbruck

Research Assistant

*Tyrol, Austria*

*Oct 2022–Present*

- Advisor: Adam Jatowt

### Digital Science Center (DiSC), University of Innsbruck

Research Assistant

*Tyrol, Austria*

*Oct 2022–Present*

- Knowledge extraction and information retrieval from unstructured text documents
- Development of Rankify framework (500+ GitHub stars) and RankArena platform
- Research on information retrieval, document reranking, and RAG systems
- Host: Adam Jatowt

### DISCO App

Machine Learning Engineer

*Remote*

*Aug 2021–May 2025*

- Built and improved OCR system for receipt extraction and information classification
- Application live on [Google](https://play.google.com/store/apps/details?id=com.disco) and [Apple](https://apps.apple.com/us/app/disco/id1544444444) Play Store

### Università Ca' Foscari Venezia

Machine Learning Researcher

*Venezia, Italy*

*Mar 2024–Aug 2024*

- Development and evaluation of graph neural networks and random forest for climate change tasks
- Host: Sebastiano Vascon

### KMG Engineering

*Remote*

- Construction and management of Image inpainting using GANs
- Hosts: Daniyar Nurseitov

## National Open Research Laboratory, Satbayev University

Undergraduate Research Assistant

Almaty, Kazakhstan

Nov 2019–July 2021

- Study of deep Learning model handwritten Kazakh and Russian database
- Advisor: Daniyar Nurseitov

## Assiut University

Research and Teaching Assistant

Assiut, Egypt

July 2016–June 2019

- Teaching assistant for computer science courses
- Research on machine learning and computer vision applications

## Publications

\* denotes equal contributions among authors

Google Scholar: <https://scholar.google.com/citations?user=uoXhF1QAAAAJ&hl=en>

### CONFERENCE PROCEEDINGS

- [C15] **Abdelrahman Abdallah**, Bhawna Piryani, Jonas Wallat, Avishek Anand, Adam Jatowt. "TempRetriever: Fusion-based Temporal Dense Passage Retrieval for Time-Sensitive Questions." **WSDM**. 2026.
- [C14] Bhawna Piryani, Jamshid Mozafari, **Abdelrahman Abdallah**, Antoine Doucet, Adam Jatowt. "Evaluating Robustness of LLMs in Question Answering on Multilingual Noisy OCR Data." **CIKM**. 2025.
- [C13] **Abdelrahman Abdallah**, Mahmoud Abdalla, Bhawna Piryani, Jamshid Mozafari, Mohammed Ali, Adam Jatowt. "RerankArena: A Unified Platform for Evaluating Retrieval, Reranking and RAG with Human and LLM Feedback." **CIKM**. 2025.
- [C12] Raphael Gruber, **Abdelrahman Abdallah**, Michael Färber, Adam Jatowt. "ComplexTempQA: A 100m Dataset for Complex Temporal Question Answering." **EMNLP**. 2025.
- [C11] **Abdelrahman Abdallah**, Bhawna Piryani, Jamshid Mozafari, Mohammed Ali, Adam Jatowt. "How Good are LLM-based Rerankers? An Empirical Analysis of State-of-the-Art Reranking Models." **Findings of EMNLP**. 2025.
- [C10] **Abdelrahman Abdallah**, Jamshid Mozafari, Bhawna Piryani, Adam Jatowt. "DeAR: Dual-Stage Document Reranking with Reasoning Agents via LLM Distillation." **Findings of EMNLP**. 2025.
- [C9] Jamshid Mozafari, **Abdelrahman Abdallah**, Bhawna Piryani, Adam Jatowt. "Wrong Answers Can Also Be Useful: PlausibleQA - A Large-Scale QA Dataset with Answer Plausibility Scores." **SIGIR**. 2025.
- [C8] Jonas Wallat, **Abdelrahman Abdallah**, Adam Jatowt, Avishek Anand. "A Study into Investigating Temporal Robustness of LLMs." **Findings of ACL**. 2025.
- [C7] **Abdelrahman Abdallah**, Jamshid Mozafari, Bhawna Piryani, Adam Jatowt. "ASRank: Zero-Shot Re-Ranking with Answer Scent for Document Retrieval." **Findings of NAACL**. 2025.
- [C6] **Abdelrahman Abdallah**, Jamshid Mozafari, Bhawna Piryani, Mohammed M. Abdelgwad, Adam Jatowt. "DynRank: Improve Passage Retrieval with Dynamic Zero-Shot Prompting Based on Question Classification." **COLING**. 2025.
- [C5] **Abdelrahman Abdallah**, Mahmoud SalahEldin Kasem, Ibrahim Abdelhalim, Norah Saleh Alghamdi, Sohail Contractor, Ayman El-Baz. "CascadePLS-ViT: Cascade with Patch-Level Self-Supervised Vision Transformers for Breast Cancer Classification in Mammography." **ISBI**. 2025.
- [C4] **Abdelrahman Abdallah**, Mahmoud Kasem, Mahmoud Abdalla, Mohamed Mahmoud, Mohamed Elkasaby, Yasser Elbendary, Adam Jatowt. "ArabicaQA: A Comprehensive Dataset for Arabic Question Answering." **SIGIR**. 2024.
- [C3] Mahmoud SalahEldin Kasem\*, **Abdelrahman Abdallah**\*, Ibrahim Abdelhalim\*, Norah Saleh Alghamdi, Sohail Contractor, Ayman El-Baz. "IHRRB-DINO: Identifying High-Risk Regions of Breast Masses in Mammogram Images Using Data-Driven Instance Noise." **MICCAI**. 2024.

- [C2] Bhawna Piryani, **Abdelrahman Abdallah**, Jamshid Mozafari, Adam Jatowt. "Detecting Temporal Ambiguity in Questions." **Findings of EMNLP**. 2024.
- [C1] Jamshid Mozafari, **Abdelrahman Abdallah**, Bhawna Piryani, Adam Jatowt. "Exploring Hint Generation Approaches for Open-Domain Question Answering." **Findings of EMNLP**. 2024.

## JOURNAL ARTICLES

- [J9] Ibrahim Abdelhalim, Yassir Almalki, **Abdelrahman Abdallah**, et al. "A Deep Learning Framework for Accurate Mammographic Mass Classification Using Local Context Attention Module." **Medical Physics**. 2025.
- [J8] **Abdelrahman Abdallah**, Mahmoud SalahEldin Kasem, Ibrahim Abdelhalim, Norah Saleh Alghamdi, Ayman El-Baz. "Improving BI-RADS Mammographic Classification With Self-Supervised Vision Transformers and Cascade Learning." **IEEE Access**. 2025.
- [J7] Mahmoud Abdalla, Mahmoud SalahEldin Kasem, Mohamed Mahmoud, Bilel Yagoub, Mostafa Farouk Senussi, **Abdelrahman Abdallah**, Seung Hun Kang, Hyun Soo Kang. "ReceiptQA: A Question-Answering Dataset for Receipt Understanding." **Mathematics**. 2025.
- [J6] **Abdelrahman Abdallah**, Daniel Eberharter, Zoe Pfister, Adam Jatowt. "A Survey of Recent Approaches to Form Understanding in Scanned Documents." **Artificial Intelligence Review**. 2024.
- [J5] Mahmoud Salaheldin Kasem, **Abdelrahman Abdallah**, et al. "Deep Learning for Table Detection and Structure Recognition: A Survey." **ACM Computing Surveys**. 2024.
- [J4] Daniyar B. Nurseitov, Galymzhan Abdimanap, **Abdelrahman Abdallah**, et al. "ROSID: Remote Sensing Satellite Data for Oil Spill Detection on Land." **Engineered Science**. 2024.
- [J3] **Abdelrahman Abdallah**, Bhawna Piryani, Adam Jatowt. "Exploring the State of the Art in Legal QA Systems." *Journal of Big Data*. 2023.
- [J2] Nazgul Toiganbayeva, Mahmoud Kasem, Galymzhan Abdimanap, Kairat Bostanbekov, Abdelrahman Abdallah, Anel Alimova, Daniyar Nurseitov. "KOHTD: Kazakh Offline Handwritten Text Dataset." **Signal Processing: Image Communication**. 2022.
- [J1] **Abdelrahman Abdallah**, Alexander Berendeyev, Islam Nuradin, Daniyar Nurseitov. "TNCR: Table Net Detection and Classification Dataset." **Neurocomputing**. 2022.

## Talks

- |            |  |              |
|------------|--|--------------|
| Nov 2025   | <b>Dual-Stage Document Reranking with Reasoning Agents via LLM Distillation</b><br>Invited Talk (Host: Prof. Han Xiao)   | <b>UIBK</b>  |
| April 2025 | <b>Advancing information retrieval: zero-shot re-ranking, and dynamic prompting for passage retrieval</b><br>Invited Talk at ifi Lunchtime Seminar (Host: Prof. Adam Jatowt) | <b>UIBK</b>  |
| Mar 2024   | <b>Generate-Retrieve-Generate: A Novel Approach to Open-Domain Question Answering</b><br>Invited Talk at ifi Lunchtime Seminar (Host: Prof. Adam Jatowt)                     | <b>EMNLP</b> |

## Teaching Experience

**Lecturer** University of Innsbruck (2022–2026)

- Computer Networks and Internet Technology (Winter 2021/22)
- Computer Networks and Internet Technology (Winter 2025/26)
- Machine Learning (PS Machine Learning) (Summer 2024)
- Data and Probabilities (PS Data and Probabilities) (*Summer 2024*)
- Data and Probabilities (PS Data and Probabilities) (Winter 2025/26)
- Advanced Data Engineering and Analytics – Understanding Natural Language (VU Advanced Data

## Engineering and Analytics) (Summer 2025)

### TA Assiut University (2015–2019)

- Web Programming
- Introduction to Programming (C++)
- Data Structures
- Machine Learning
- Project Management
- Network Management

### Mentor Mentoring 2 MSc and BSc students and (2022–Present)

- Baraa Ahmed (Winter 2024–present), MSc, Co-authored.
- Andreas Herzinger (Fall 2024–Summer 2025), MSc, Co-authored.

## Open-Source Contributions

---

### Rankify Library | Creator and Maintainer

- Developed and maintain comprehensive Python package for information retrieval and reranking evaluation
- Supports evaluation of retrieval, reranking, and RAG systems with automated metrics and human feedback

### RankArena Platform | Lead Developer

- Created unified evaluation platform for retrieval and reranking systems
- Enables researchers to benchmark models against state-of-the-art baselines

## Awards and Honors

---

Satbayev University Scholarship to study for a master's degree

Oct 2019

Satbayev University Best Master Thesis

July 2021

## Academic Services

---

### Conference Reviewer

- EMNLP (2024)
- ACL Rolling Review (2023, 2024, 2025)
- COLING (2025)
- ECIR (2025)
- CIKM (2024, 2025)
- SIGIR (2024,2025)

## References

---

Prof. Adam Jatowt **UIBK** (Adam.Jatowt@uibk.ac.at)

Prof. Islam A.T.F. Taj-Eddin **Assiut University** (itajeddin@aun.edu)

Prof. Daniyar Borisovich Nurseitov **KMG Engineering** (nurseitovdb@gmail.com)