

Dr Emad A. Alghamdi

Assistant Professor of NLP, AI Consultant,
Founding Director of the ASAS AI Lab
Founding Director of the AI in Language Learning and Assessment Lab
www.emad.ai · eaalghamdi@kau.edu.sa · [Github](#)

Summary

Dr. Emad Alghamdi is a distinguished AI researcher specializing in Arabic Natural Language Processing (NLP) and computational linguistics. Holding a Ph.D. from the University of Melbourne, he received the Chancellor Award for Excellence in his doctoral thesis. As the founding director of ASAS AI research lab, he leads pioneering initiatives at the forefront of AI innovation, and develops AI technologies for Arabic and with special focus on the Saudi context, including Arabic Large Language Models. Dr Alghamdi has won numerous research awards and grants. He holds a prolific publication record in top journals and conferences. Additionally, he serves as an AI consultant for AI governance at the Center of Excellence in AI and Data Science at KAU, ensuring responsible and ethical AI practices.

Education

- 2017 – 2021 **University of Melbourne** – Melbourne, Australia
PhD in Computational Linguistics
Supervisors: Dr Paul Gruba (Linguistics) and Dr Eduardo Velloso (Computer Science).
PhD Title: "Automated Video Difficulty Assessment"
*Received **the Chancellor's Prize for Excellence in PhD Thesis** in 2022.*
- 2010 – 2013 **Missouri State University** – Missouri, USA
M.A. in English (Linguistics)
GPA: 3.97/4.

Work experience

- 2024 - Present **AI Ethics and Safety Consultant**
• Center of Research Excellence in Artificial Intelligence and Data Science at KAU
- 2023 - Present **Founding Director of ASAS AI Lab**
• Founded ASAS Lab and Research Community in May 2023 [Website](#).
• The lab has 40+ members who hold PhDs, Masters and Bachelors in various AI domains.
• The lab conducts fundamental research, offer consultancy, and build AI technologies and solutions for the Arabic language and its people.

- 2022 - Present **Founding director of the AI in Language Learning & Assessment Lab**
- Founded the AILLA Lab in January 2022. [Website](#)
 - The lab's projects are supported by grants from local and international organizations, including Educational Testing Service (ETS), Sage, Pearson, Ministry of Culture, and KAU.
- 2022 - Present **AI Consultant**
- Provide consultation services to local companies in areas such as data governance and AI strategy.
- 2022 - Present **Assistant Professor in Computational Linguistics**
King Abdulaziz University, Jeddah, Saudi Arabia
- Sep 2020 - Feb 2021 **Data Scientist Intern**
Melbourne Data Analytics Platform, University of Melbourne, Australia
- Worked on NLP projects using weak supervision machine learning techniques.
- Jul 2019 - Aug 2019 **Lecturer**
University of Melbourne, Australia
- Taught a Master-level course on Technology and Language Learning
- 2014 - 2016 **Director of e-training Unit**
Center for Teaching and Learning Development, King Abdulaziz University
- Founded KAUX, an Arabic MOOC platform
 - Led a team of faculty members, administrators, and on-site and remote technicians
 - Oversaw the development of two MOOCs on innovation and critical thinking

Research Grants

- Submitted **RDIA: SART (10M SAR)**
Arabic Large Language Model with Applications
Role: Co-Primary Investigator
Research Team: Prof Jinchao Xu (KAUST), Dr Emad Alghamdi, Prof Jürgen Schmidhuber (KAUST), Prof David Keyes (KAUST), Prof Bernard Ghanem (KAUST), Prof Abdulmohsen Althubaity (KACST), Dr Sakhar Alkhereyf (KACST), Dr Tariq Alhindi (KACST), Dr Benyou Wang (UHK-SZ), Dr Lian Zhang (SRIBD), Hani Bukhari (Nex-G)
- Submitted **RDIA: TDG (10M SAR)**
An AI-powered Arabic Tutoring Platform for STEM Education in the Kingdom.
Role: Co-Primary Investigator
Research Team: Prof. Sultan Albarakati (KAUST), Dr Emad Alghamdi, Prof Bernard Ghanem (KAUST), Dr Naeemullah Khan (KAUST), Dr George Turkiyyah (KAUST), Dr Younes Mourri (LiveTech.AI), Dr Nayyaf Aljabri (Director of Research, ETEC)

- Submitted **RDIA: YSF (1.5M SAR)**
A Robust and Culturally-aligned Benchmarking Suite for Arabic and Chinese Large Language Models
Role: Primary Investigator
Research Team: Dr Emad Alghamdi, Dr Benyou Wang (The Chinese University of Hong Kong, Shenzhen), Saied Alshahrani (University of Bisha), Khalid Almubarak (PSAU)
- Submitted **RDIA: YSF (1.5M SAR)**
Taakkad: An Arabic AI System for Fact Checking
Role: Co-Primary Investigator
Research Team: Dr Noha Alduaij (Majmaah University), Dr Emad Alghamdi, Dr Amal A. Al-Shargabi (Qassim University), Dr Nora Alturayef (IAU), Atheer Algherairy (IAU)
- Submitted **RDIA: YSF (1.5M SAR)**
Unlocking Insights: Constructing a Comprehensive Arabic Event Relation Corpus
Role: Co-Primary Investigator
Research Team: Dr Mohammad Aldawsari (PSAU), Dr Emad Alghamdi, Dr Sakhar Alkhereyf (KACST), Prof Sultan Almujaivel (KSU)
- Submitted **Duolingo (370K SAR)**
A Grammatical Error Correction System with Personalized Corrective Feedback using Retrieval Augmented Generation
Role: Primary Investigator
Research Team: Dr Zheng Yuan (King's College London, Cambridge University), Dr Afnan Alzahrani, Dr Staurt McLean (Kindai University, Japan)
- 2022 – 2023 **Pearson Inc (55k SAR)**
An investigation into the relationship between Knowledge-based Vocabulary Lists and PTE writing essay test
Role: Primary Investigator
For developing an optimization algorithm that best encapsulates the spread of vocabulary knowledge and predictive capacity in the PTE writing assessment.
- 2022 – 2023 **SAGE Concept Grant (10k SAR)**
Automated Video Analysis
Role: Primary Investigator
For developing automated video analysis software for analyzing videos using computer vision and AI.

- 2022 – 2023 **Ministry of Culture Grant (40k SAR)**
Enhancing Arabic MT Translation for Domain Adaptation: A Case Study of Financial Texts
Role: Primary Investigator
- 2021 – 2022 **KAU Excellence in Research Grant (40k SAR)**
 "The Relative Contribution of Verbal Complexity to Second Language Video Lectures Difficulty Assessment". Published in *Modern Language Journal*
 Ranked 1 out of 1101 in linguistics and language journals
- 2021 – 2022 **KAU Excellence in Research Grant (35k SAR)**
 "The Use of Lexical Complexity for Assessing Difficulty in Instructional Videos" Published in *Language Learning and Technology*
 Ranked 24 out of 1101 in linguistics and language journals

Honors and scholarships

- 2021 **The Chancellor's Prize for Excellence in PhD Thesis** (University of Melbourne)
My PhD thesis was the best thesis in the faculty of Art and one of 6 theses that won the prestigious Chancellor's prize for excellence in PhD thesis in the University of Melbourne.
- 2019 **Research Excellence Award** (The Saudi Cultural Attaché in Australia)
- 2017-2021 **PhD Scholarship** (King Abdulaziz University)
I received full scholarship from KAU to pursue my PhD research at the University of Melbourne.
- 2012, 2013 **Academic Excellence Award** (The Saudi Cultural Attaché in USA)
I was awarded Academic Excellence Award by the Saudi Cultural Attaché in USA for excelling academically throughout the school year
- 2012 **Certificate of Appreciation** (The Saudi Cultural Attaché in USA)
I was awarded a Certificate of Appreciation by the Saudi Cultural Attaché in Washington, D.C for leadership as a president of the Saudi Club at Missouri State University

Scholarly Work (Selected)

- In preparation **Alghamdi, Emad**, Alshahkih, R., Khalid, Alshahrani, S., Alduaiji, N, & Batool
 Evaluating pruning and post-training quantization methods for Arabic LLMs
- In preparation **Alghamdi, Emad**, Alghamdi, M., Albosais, R, & Loquman, H.
 ArabTalk: Arabic Talking Head Dataset and Baseline

- In preparation Khalid, A., Alshahrani, S., & **Alghamdi, Emad**,
ArbBench: Arabic Open LLM Leaderboard
- In review **Alghamdi, Emad** Reem Masoud, Deema Alnuhait, Afnan Alomairi, Ahmed Ashraf, Mohamed Zaytoon. "AraTrust: An Evaluation of Trustworthiness for LLMs in Arabic"
<https://arxiv.org/abs/2403.09017>
- In review Radwan, Ahmed, Alasmari, Khaled, Abdulbagi, Omar, **Alghamdi, Emad**. "SARD: A Human-AI Collaborative Story Generation."
The ACM CHI conference on Human Factors in Computing Systems.
- In review Abdelaty, Yara, Alnajashi, Alaa, **Alghamdi, Emad**. "How Accurate is GPT-4 in Detecting and Providing Written Corrective Feedback on L2 Grammatical Errors"
- In review **Alghamdi, Emad**, Zakraoui, Z. & Albanmy, F
"Enhancing Arabic MT Translation for Domain Adaptation: A Case Study of Financial Texts" <https://doi.org/10.48550/arXiv.2309.12863>
- In review **Alghamdi, Emad**
"Leveraging ensemble machine learning and multimodal video complexity for better prediction of video difficulty in second language"
- 2024 Shivalika Singh, Freddie Vargus, Daniel D'souza, ... **Alghamdi, Emad**, ... Sara Hooker. "Aya Dataset: An Open-Access Collection for Multilingual Instruction Tuning"
<https://arxiv.org/pdf/2402.06619.pdf>
- 2024 Mina Lee, Katy Gero, John Chung, Simon Buckingham Shum, Vipul Raheja, Hua Shen, Subhashini Venugopalan, Thiemo Wambsganss, David Zhou, **Alghamdi, Emad**, Roy Pea, ... Pao Siangliulue. "A Design Space for Intelligent and Interactive Writing Assistants."
The ACM CHI conference on Human Factors in Computing Systems.
- 2024 **Alghamdi, Emad**
"What makes video-based academic lectures difficult for language learners to comprehend? The role of multimodal complexity" <https://doi.org/10.1111/flan.12747>. *Foreign Language Annals*.
- 2023 Yousef Altaher, Ali Fadel, Mazen Alotaibi, Mazen Alyazidi, Mishari Al-Mutairi, Mutlaq Aldhbuiub, Abdulrahman Mosaibah, Abdelrahman Rezk, Abdulrazzaq Alhendi, Mazen Abo Shal, **Alghamdi, Emad**, ... Zaid Alyafeai.
"Masader Plus: A New Interface for Exploring+ 500 Arabic NLP Datasets."
<https://doi.org/10.48550/arXiv.2208.00932>

- 2023 Matthews, Joshua., Masrai, Ahmad, Lange, K., McLean, S., **Alghamdi, Emad**, Kim, K., Shinhara, Y., Tada, S.
"Using meaning recall to explore links between aural lexical knowledge and L2 listening: A replication and extension study." *TESOL Quarterly*
- 2023 **Alghamdi, Emad**, Paul Gruba, Ahemd Mersi, and Eduardo Velloso
"The Use of Lexical Complexity for Assessing Difficulty in Instructional Videos." *Language Learning and Technology*.
- 2022 Nawaz, Sadia, **Alghamdi, Emad**, Namrata Srivastava, Jason Lodge, & Linda Corrin.
"Understanding the role of AI and learning analytics techniques in addressing task difficulties in STEM education." A book chapter in *Artificial Intelligence in STEM Education: The Paradigmatic Shifts in Research, Education, and Technology*, Springer
- 2022 **Alghamdi, Emad**, Paul Gruba, and Eduardo Velloso
"The Relative Contribution of Verbal Complexity to Second Language Video Lectures Difficulty Assessment."
The Modern Language Journal.
- 2022 **Alghamdi, Emad**
"Book review: Computational psychometrics: new methodologies for a new generation of digital learning and assessment"
- 2021 **Alghamdi, Emad**, Eduardo Velloso, and Paul Gruba
"AUVANA: An Automated Video Analysis Tool for Visual Complexity."
<https://doi.org/10.31219/osf.io/kj9hx>
- 2018 **Alghamdi, Emad**
Effects of Using Multimodal Glosses in Video Games to Enhance Incidental Vocabulary Learning and Retention
A book chapter in *Handbook of Research on Integrating Technology Into Contemporary Language Learning and Teaching*, Chap.2
- 2018 **Alghamdi, Emad**, Fahad Otaif, and Paula Gruba
Designing a video playing interface for second language learners.
The 35th International Conference on Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education
- 2018 **Alghamdi, Emad** and Sayyed Rashid Shah
Exploring the Effects of Mobile-Based Audience Response System on EFL Students' Learning and Engagement in a Fully Synchronous Online Course.
International Journal of English Linguistics
- 2016 **Alghamdi, Emad** and Hussam Rajab
Unmonitored students self-created WhatsApp groups in distance learning environments: A collaborative learning tool or cheating technique
International Journal of Research

- 2016 **Alghamdi, Emad**
 Untangling Multimedia Effects on EFL Incidental Vocabulary Learning via Playing an Online Hidden-Object Game
International Journal of Computer-Assisted Language Learning and Teaching
- 2015 **Alghamdi, Emad** Aljohani, N, Alsaleh, A, Bedwi, W, Basher, M
 CrowdQ: A virtual crowdsourcing platform for question items development in higher education.
The 17th International Conference on Information Integration and Web-Based Applications and Services
- 2014 **Alghamdi, Emad**
 Gulf Pidgin Arabic: A Descriptive and Statistical Analysis of Stability
International Journal of Linguistics

Talks

- 2024 **Alghamdi, Emad A,**
 [Workshop] Introduction to AI Ethics and Governance.
King Abdulaziz University, Saudi Arabia
- 2024 **Alghamdi, Emad A,**
 [Invited Panelist] Revolutionizing Education for New Generation.
ESC24 conference
- 2023 **Alghamdi, Emad A,**
 [Invited Seminar] The Current State of Art in Arabic Large Language Models.”
The Computer, Electrical, Mathematical Sciences and Engineering Division at King Abdullah University of Science and Technology (KAUST).
- 2023 **Alghamdi, Emad A,**
 Towards More Predictive Video Difficulty Assessment Estimation Using Ensemble Machine Learning” *Australian Learning Analytics Summer Institute, University of South Australia, Adelaide SA, Australia*
- Sep 2023 **Alghamdi, Emad A,** and Lallmamode, S. P
 ”Where Humans Meet the Machines: An Innovative Approach to Language Test Development at Scale.”
Artificial Intelligence Research in Applied Linguistics (AIRiAL), Teachers College, Columbia University
- 2023 Saleh, Alaa, and **Alghamdi, Emad A,**
 ”But No One Else in My Room: Preventing Cheating in At-home Tests Using 360 Camera”
Symposium on Applied Linguistic Research, Prince Sultan University

- Nov 2022 [Invited Colloquium] The Current and Potential Applications of NLP in First and Second Language Learning and Research
The 3rd International Symposium on Applied Linguistics Research
- Nov 2022 Current and potential applications of NLP in L2 language learning and assessment
The 3rd International Symposium on Applied Linguistics Research
- Jun 2019 Video difficulty assessment
Human-Computer Interaction Seminars, School of Computing and Information Systems, The University of Melbourne, Melbourne, Australia
- Dec 2018 Video complexity and language learners: Making the match
The Third Saudi Scientific Symbolism, University of Sydney, Sydney, Australia

Projects (Selected)

- 2023 – 2024 **Aya Dataset: An Open-Access Collection for Multilingual Instruction Tuning**
Role: Arabic Language Ambassador
The project aims develop and release a high-quality multilingual dataset for more than 100 languages.
- 2023 – 2024 **Arabic Font Optimization**
Role: Primary Investigator
The project aims to develop an innovative approach to optimize Arabic font for better reading experience. In this project, I collaborate with expert researchers from **Google Font Team**.
- 2023 – 2024 **Automatic Text Difficulty Adjustment**
Role: Primary Investigator
The project aims to develop an innovative web reading system that automatically adjusts reading difficulty for readers. In this project, I collaborate with expert researchers from **Adobe Inc, Google, and Pennsylvania University, Melbourne University**
- Dec 5 2022–
Dec 19 2022 **Hugging Face X Lambda: Whisper Fine-Tuning Event**
Role: Contributor
We fine-tuned OpenAi’s Whisper models for several Arabic dialects and ranked number one on the competition [See the results](#)
- 2022 -2023 **Automated Video Analysis Software**
Role: Primary Investigator
A cross-platform desktop for automated video analysis using deep learning and advanced computer vision algorithms. The project is funded by **SAGE Publishing**.

2022 **Adwat**
Role: Group leader
Adwat the first online catalogue and interface for Arabic NLP tools. The catalogue contains more than 60 tools for various NLP tasks. [Visit the project site](#)

2022 **Masader Plus**
Role: Contributor
Masader Plus is the first online catalogue for Arabic NLP datasets. The catalogue contains 500+ datasets with more than 25 metadata annotations for each dataset. [Visit the project site](#)

Teaching Course

Natural Language Processing: Professional Master in Artificial Intelligence

Student Supervision

Ahmed Radwan, Khaled Alasmari, Omar Abdulbagi: Sard: Image-inspired Narratives

Shrifah Al-shehri: Evaluating GPT-4 ability to Design English Reading Test Items for Language Learners

Haifa Fayez Alhusaini: Exploring AI-language learners interactions in second language writing

Technical skills

Programming languages

Python, JavaScript, R

Machine Learning

PyTorch, Keras, Scikit-learn, spaCy, OpenCV, Huggingface, NLTK

Web and Desktop Technologies:

HTML, CSS, AWS, Firebase, D3, Tableau, Vue, Nuxt, Electron, Django, Flask, Streamlit

Statistics:

RStudio, SPSS

Tools and software

2023 **Floow**

Floow is a GPT-powered writing assistant for language learners. I designed and developed the interface, coded the backend. Floow is in the beta version and it has more than 1k registered users. The system is available at www.floow.ai

2021 **The Automated Video Analysis (AUVANA)**

I developed the first version of The Automated Video Analysis Software during my PhD. AUVANA is an open-source, cross-platform computer vision software. It helps AI researchers and scientist to expedite video annotation. A new updated version of the software will be available at www.auvana.ai

Personal Development

2018 **Machine Learning with TensorFlow on Google Cloud Platform**

Coursera– A 5-course specialization by Google Cloud

• Introduction to TensorFlow • How Google does Machine Learning • Feature Engineering • Launching into Machine Learning • Art and Science of Machine Learning

2016 **International License for Entrepreneurship**

75-hour Training Course, Jeddah, Saudi Arabia.

2014 **Design and Development of Educational Technology, MIT**

EdX

2014 **Entrepreneurship 101: Who is your customer?, MIT**

EdX

2014 **Statistics: Making Sense of Data, University of Toronto**

Coursera

2013 **Justice, Harvard University**

EdX

2014 **Inspiring Leadership through Emotional Intelligence, Case Western Reserve University**

Coursera

2013 **Internet History, Technology, & Security, University of Michigan**

Coursera

Volunteer

2024 Judge in AI in Transpiration (TransAI) Competition
Engineering Day, KAU

2015 – 2016 **Founding Director**

Afaqi Learning Hubs

2012 – 2013 **President of the Saudi Student Club**

Missouri State University, MO, USA

2012 – 2014 **English-Arabic Translator**
TED

2012 – Present **Journal and Conference Reviewer**

- Arabic Conference (EMNLP)
- CHI
- Computer Assisted Language Learning
- Behavior Research Methods
- Language Learning & Technology
- Interactive Learning Environments
- Computer Assisted Language Learning
- Sage Open
- System

Languages

Arabic (fluent), English (advanced)

Communities and Professional Memberships

2024 – Present Consultation Committee for Applied Linguistics Program, Prince Norah bin Abdulrahman University

2023 – Present forhumanity <https://forhumanity.center>

2022 – Present ACL Special Interest Group on Arabic Natural Language Processing

2022 – Present The ACM Special Interest Group on Computer-Human Interaction (SIGCHI)

2021 – Present The Association of Computational Linguistics

2021 – Present ARBML (Arab Machine Learning) - Core member

2020 – Present The Readability Consortium

2020 – Present American Association of Applied Linguistics

2020 – 2021 Melbourne Data Analytics Platform

2019 – 2021 Human-Computer Interaction Group, the University of Melbourne