Muhammad Zahid

Digital Humanities Researcher | Urdu Computational Philology

An Independent Researcher.

Pakistan

RESEARCH PROFILE

Digital Humanities researcher developing reproducible computational frameworks for the analysis and digital preservation of classical Urdu poetry. Combines an M.A. in Urdu Literature with an MSc. in Computer Science to bridge low-resource NLP and literary scholarship. Author of a peer-reviewed publication and creator of an open-source Python toolkit for Urdu computational philology. Seeking a PhD to advance methodological rigor and accessibility in the computational analysis of non-Western literary traditions.

EDUCATION

Master of Computer Science | CGPA: 3.36/4.00
Virtual University of Pakistan | 2015 – 2017
Relevant Coursework: Data Structures, Algorithms, Foundations of AI & Data Science

Master of Arts in Urdu & Iqbaliat
The Islamia University of Bahawalpur | 2003 – 2005
Advanced study in: Classical & Modern Urdu Poetry, Literary Theory, Philology

Bachelor of Arts (Arabic, Persian & Islamic Studies)
The Islamia University of Bahawalpur | 1999 – 2001

PUBLICATIONS

Zahid, M. (2026). Computational Framework for the Descriptive Analysis and Digital Preservation of Classical Urdu Poetry.
Liberal Journal of Language & Literature Review, 4(1). HEC Recognized (Y-Category Journal)

- End-to-end, reproducible Python pipeline for Urdu poetry analysis
- Case Study: Comparative stylistic analysis of Diwan-e-Ghalib, Mir Taqi Mir, and Faiz Ahmed Faiz

View Publication

In Preparation: Computational Analysis of Classical Urdu Poetry: A Digital Humanities Framework for Ghalib, Mir, and Faiz. (Target: 2026)

TECHNICAL PROJECTS

Urdu Computational Philology Toolkit
Python, Git, Jupyter

An open-source framework for the computational analysis of classical Urdu poetry. Designed Python modules addressing the lack of tools for low-resource, RTL scripts in Digital Humanities.

- corpus_loader.py (cleaning & normalization)
- unicode_verifier.py (UTF-8 validation)
- urdu_tokenizer.py (rule-based, RTL-aware tokenization)
- corpus_analyzer.py (frequency, TTR, n-gram analysis)

RESEARCH INTERESTS

- Computational Literary Studies & Digital Philology
- Low-Resource NLP (Urdu, Persian, Arabic Scripts)
- Digital Archives & Cultural Heritage Preservation
- Stylometry & Corpus Linguistics for Non-Latin Scripts
- Reproducible Research Methodologies in the Humanities

TECHNICAL & METHODOLOGICAL SKILLS

Programming & Tools
Python, Regular Expressions, Git/GitHub, Jupyter, Command Line

Digital Humanities / NLP
Text Preprocessing, Corpus Linguistics, Rule-based Tokenization, Descriptive Text Analytics, Stylometry, RTL/Unicode Processing

Research Practice
Reproducible Pipeline Development, Scholarly Code Documentation, Open-Source Research, Data Visualization (Matplotlib/Seaborn)

Languages
Urdu (Native), English (Professional), Arabic (Intermediate Reading), Persian (Intermediate Reading)

PROFESSIONAL EXPERIENCE

Data Management & Systems Roles

Punjab Police Department & Government Institutions | 2005 – Present

- Systems Analysis & Workflow Design: Developed and maintained digital record-keeping systems
- Problem-Solving in Constrained Environments: Project management and independent initiative

CERTIFICATIONS

- Machine Learning, AI & Data Science – eHunar (2025)
- Python for Everyone – eHunar (2023)
- Web Design & Development – University of the Punjab Lahore
- Database Administration – University of the Punjab Lahore

SEEKING PhD IN DIGITAL HUMANITIES (2026)

Academic transcripts, writing samples, and letters of reference are available upon request.

CONTACT

zahidrana708@gmail.com LinkedIn GitHub