Zhiqi Xu, PhD

Evaluation Specialist | AI Product Ownership | Policy and Behavioural Research

zhiqi.s.x@outlook.com · zhiqixu.com · November 2026

Professional profile

Evaluation specialist and interdisciplinary researcher combining development economics, behavioural research, mixed methods, and applied AI. At FAO's Office of Evaluation, develops analytical frameworks and evidence workflows and serves as product owner for an AI application that applies conceptual frameworks across document collections. Connects methodological design, human validation, software development, and evaluator needs. Experience spans rural development, seed systems, technical cooperation, social protection, and pandemic preparedness.

Professional experience

Evaluation Specialist and AI Product Owner

Food and Agriculture Organization of the United Nations • Office of Evaluation • Rome | November 2025-present

Joined as Evaluation Analyst; subsequently served as Evaluation Specialist. Responsibilities span evaluation design, portfolio analysis, evidence synthesis, reporting, and work within OED's Data Lab. Product ownership describes responsibilities for the Portfolio Mapping application.

AI product development and adoption

- Serve as product owner for Portfolio Mapping, a framework-guided qualitative analysis application. Identify evaluation needs, support analytical workflow design, test functionality, collect user feedback, and translate findings into feature changes and bug fixes in collaboration with an AI engineer.
- Collaborate on testing and improving working features that let users create analytical workflows with an agent, configure workflow-specific validation checks, and generate dashboards. Support evaluators in designing and running their own analyses.
- Onboarded TCP evaluation users and supported initial workflow design, enabling a user to conduct an independent analysis that classified project logframes into 15 output types and mapped them to effect and catalytic pathways.
- Collect feedback from the first cohort of test users and refine usability, with the aim of making the application intuitive for new evaluators. Draft maintenance and feedback-collection arrangements to support continued use.
- Explore applications beyond portfolio mapping, testing whether a workflow-building agent can inspect a connected database, assess whether it supports an analytical request, and propose an approach.

Analytical workflows and quality assurance

- Develop reusable AI-assisted extraction and classification workflows using Python and Azure OpenAI, translating research questions into explicit categories, evidence requirements, boundary cases, and structured outputs linked to source passages.
- Compare AI classifications with human-labelled examples, track precision, recall, and F1, and investigate disagreements to refine prompts, analytical rules, and inconsistent reference labels. Develop dashboards for reviewing classifications and supporting evidence.

- Designed a stratified random sampling framework for Seeds analysis, offering evaluation managers sample-size and expected-margin-of-error options at 95% confidence. Balanced category coverage and review effort within proposed samples of 15-30 projects; manual review and performance estimates remain pending.

Evaluation assignments at FAO

Technical Cooperation Programme evaluation

- Developed and refined rules for classifying evidence of catalytic pathways in technical-assistance projects, distinguishing policy drafting from adoption or implementation. Used calibration and independent human validation to assess performance and investigate errors.
- Supported application-based analysis linking 15 output types to effect and catalytic pathways. TCP analytical results have been used in evaluation reporting.

Evaluation of FAO seeds programming

- Lead portfolio-analysis methods for seed-system projects, including evidence from Afghanistan, Ethiopia, Sri Lanka, Venezuela, and Tajikistan; supported analysis of approximately 600 project documents.
- Translate evaluation questions into structured classification rules and traceable evidence outputs. Seeds analytical results have been used in evaluation reports; further quality assurance using the application is planned.

Evaluation of pandemic preparedness contributions

- Contribute to analytical frameworks, portfolio coding, and evidence synthesis for pandemic prevention, preparedness, and response. Develop approaches to assess workforce capacity development and workplace application of acquired competencies.
- Specify evidence requirements linking training, competencies, and use of new skills, supporting analysis against the evaluation’s theory of change.

Global thematic evaluation of Better Production 4

- Supported evaluation delivery from inception and evidence collection through analysis, stakeholder validation, and reporting, focusing on equitable access to resources and finance for rural producers.
- Structured and analysed approximately 750 project records and 632 results and achievement entries. Translated priority areas into an analytical framework, developed AI-assisted coding, and produced portfolio findings and visualizations.
- Contributed report sections and methodological annexes, prepared the first draft of the Programme Committee report for senior review, and supported consultations across technical, regional, and country offices.

Additional professional experience

Evaluation and Policy Consultant

Center of Consultation for Farmers’ Association • Beijing, China | September 2021-November 2025

- Delivered evaluation and policy assignments on rural development, poverty reduction, rural governance, microfinance, and institutional reform for government, foundations, NGOs, and international partners.
- Co-led a mixed-method assessment across 216 villages in 23 provinces. Developed instruments, trained enumerators, coordinated and co-supervised field teams, and applied data-quality protocols.

- Synthesized more than 180 submissions on draft rural legislation, identifying areas of agreement, competing stakeholder interests, and implications for local organizations; translated consultation evidence into policy recommendations.
- Initiated collaboration between CCFA and ISS, Erasmus University Rotterdam. Produced evaluation reports and policy recommendations for government and foundation partners and liaised with the International Cooperative Alliance in Brussels.

Staff Accountant and Tax Policy Advisory

Ernst & Young China Advisory Limited • Beijing, China | October 2015-March 2016

Conducted tax-policy research and analytical assessments, prepared regulatory briefs, and supported client-government engagement.

Research experience

Doctoral Researcher

International Institute of Social Studies • Erasmus University Rotterdam | December 2020-September 2026

Investigated behavioural and institutional agency in poverty alleviation, rural governance, and community-led service delivery through experimental and qualitative research in rural China.

- Eldercare and collective action: designed and led a field experiment and household survey in 14 villages using a two-round public-goods game. Examined hypothetical monetary and time contributions, government and community messages, responses after message removal, and differences in beliefs about responsibility for care.
- Led instrument design, field-team training, implementation, data preparation, and statistical analysis. Used latent profile analysis to examine heterogeneity in participants' beliefs and responses.
- Rural governance: designed a randomized survey experiment with 210 village leaders across 98 counties in 23 provinces to examine how state-centred and individual-agency narratives influence judgments about local implementation.
- Policy entrepreneurship: conducted a qualitative case study of a community-driven tourism programme using interviews, document analysis, and the Multiple Streams Framework. Examined how focusing events, organizational affiliations, and policy alignment shaped programme expansion and local ownership.

Mixed methods evaluation scoping review

Co-author • Ongoing

Mixing the (un)mixable in mixed methods evaluations: a scoping review protocol. The planned review examines quality criteria and their use across evaluation design, implementation, and reporting. It proposes mapping how and when methods are integrated and developing a mixed-methods evaluation quality index to inform practice.

Education

PhD in Development Studies

ISS, Erasmus University Rotterdam • The Netherlands | December 2020-September 2026

Main subject: Development Economics. Successfully defended on 29 September 2026. Dissertation: Co-producing Development: Behavioural and Institutional Agency in China's Poverty Alleviation.

Master of Law with a focus on Social Psychology

Central University of Finance and Economics • Beijing, China | September 2016-June 2019

Thesis: Economic Preferences of Farmers and Pastoralists in China: An Experimental Economics Approach.

Bachelor of Economics in Taxation

Central University of Finance and Economics • Beijing, China | September 2011-June 2015

Selected publications and work in progress

- First author (2026). Policy windows in an authoritarian system: How alignment, leadership signals, and disasters enabled policy entrepreneurship in China's poverty alleviation. Review of Policy Research.
- Co-first author (2025). Open Science in the Developing World: A Collection of Practical Guides for Researchers in Developing Countries. Advances in Methods and Practices in Psychological Science. Collaboration involving 45 authors across 20 countries.
- Xu, Z., & Xin, Z. (2019). Differences in risk preferences under farming and nomadic cultures: Evidence and analysis frameworks. Advances in Psychological Science, 27(9), 1631-1642.
- Xu, Z., & Xin, Z. (2018). Statistical analyses in single-subject experiment research: Effect sizes based on non-overlap methods. Psychology: Techniques and Applications, 6(2), 89-99.

Work in progress: rural eldercare manuscript examining incentives, beliefs, and willingness to contribute to community services; mixed-methods evaluation scoping review protocol.

Inter agency collaboration and professional service

UNEG ethical AI checklist

Co-lead • Ongoing outline and partial draft

Co-lead development of a practical checklist for ethical AI use across evaluation phases. Collaborate with colleagues from UNITAR, ILO, UNICEF, IFAD, WFP, UNEP, OHCHR, ITC, and the Adaptation Fund's Technical Evaluation Reference Group (AF-TERG).

Policy brief on AI adoption and evaluator capabilities

Lead drafter • Ongoing outline and partial draft

Lead drafting of a policy brief reflecting on AI use cases in collaboration with IFAD, WFP, GCF, and GEF. Proposed themes include practical steps for evaluation offices at different levels of AI maturity, institutional readiness, capacity building, adoption arrangements, and maintaining critical inquiry as evaluators delegate more tasks to AI.

Professional service

- Journal reviewer for the Journal of Development Studies and the European Journal of Development Research since 2024.
- Ex officio member of the Executive Board, Society for the Improvement of Psychological Science, since 2025; Outreach and Impact Committee member, Chinese Open Science Network, since 2023.

Teaching and research dissemination

- EvalforEarth roundtable panelist, November 2026: contributed methodological guidance on sampling and integration in mixed-methods evaluation, including recruitment beyond survey respondents, shared exclusions and biases, and limits to generalization.
- Presented How People Shape Development: Behaviour, Institutions and Evidence from Rural China at an FAO OED division meeting, November 2026, connecting three doctoral studies through their mixed-methods approach.
- Lecturer and examiner, ISS, Erasmus University Rotterdam, 2023-2024: contributed to Evaluation of Development Policies, Programmes and Projects and Human Behaviour and Experiments in Development; assessed MA-level terms of reference, logical frameworks, and evaluation methodologies.
- Delivered meta-research methods workshops and training at the Busara Center in Nairobi in collaboration with SIPS, 2024.
- Supervised a master's student from the Government of Suriname through the FHR Lim A Po Institute for Higher Education on migration policy in Suriname and Central America, 2025-2026.

Methods and technical capabilities

Evaluation and research: mixed-methods design; randomized field and survey experiments; qualitative policy research; portfolio analysis; theory-of-change frameworks; evidence synthesis; survey design; field-team coordination; stratified random sampling; human-AI validation; precision, recall, and F1; econometric analysis; latent profile analysis.

Applied AI and product work: conceptual frameworks and classification rubrics; structured LLM extraction; source-linked evidence; Azure OpenAI; ReAct and validator-based workflows; text embeddings and BERTopic; product requirements; usability testing; evaluator onboarding; feedback-led iteration with engineers.

Tools: Python, R, Stata, NVivo, Excel, Power Query, Git, GitHub, and VS Code. AI-assisted programming with Codex and Claude; initial hands-on experience with Claude Code.

Selected professional training

Theory-based Impact Evaluation, FAO, 2026; Impact Evaluation of Development Policies, Italian Center for International Development, 2025; Advanced Statistics with STATA, LinkedIn and NASBA, 2026; NVivo, 2025; Google Cloud Generative AI Leader, 2026; AI Ethics, Governance and Prompting, PoliSync, 2026.

Languages

Chinese Mandarin: native | English: professional working proficiency | Dutch: A2