Fan Xuejiao (Ariel)

Beijing, China • fanlaplace369@gmail.com

Accountable ML & AI governance: protocol compilation (YAML), rationale/change logs, dispute/appeal workflows, calibration & abstention for trustworthy deployment.

Profile

Research applicant focusing on making LLM system behaviour legible: combining learning with reasoning, calibrated uncertainty, selective abstention, and rationale logging; prototypes study UIs (Figma/HTML/CSS/JS) and runs fully reproducible analyses (Docker, seeds, artefact hashes).

Research Experience

School of Computer Science, Beijing Institute of Technology — Research Assistant (part-time) | Beijing, China • 2024–Present

- Prototyped conversational-agent UIs (identity cues; uncertainty/refusal messaging) with full interaction logging (Figma, HTML/CSS/JS).
- Implemented calibrated uncertainty and selective abstention; analysed accuracy, time-on-task, and trust metrics (Python/R).
- Compiled stakeholder preferences into executable YAML protocols; tuned operating points (τ for precision at minimum coverage; source-weighting w).

Institute of Intelligent Rule of Law, Tsinghua University — Research Assistant | Beijing, China • 2022–2023

- Translated legal/policy requirements into audit checklists and machine-readable protocols; stress-tested thresholds to balance FP/FN risk.
- Mapped fairness, verifiability, and reliability to concrete evaluation metrics and sampling rules; maintained rationale/change logs.
- Used diagrammatic/argument mapping to externalise reasoning and inform protocol design.

ByteDance — Policy/UX Intern, Content Governance | Beijing, China • 2021

- Converted platform rules and appeal data into monitoring metrics (misclassification, dispute tickets, latency) and decision logs.
- Facilitated cross-functional reviews (Policy, Legal, Product) with structured rationales to justify rule updates.

Research Projects & Grants (Selected)

AI Governance & Legal-Tech Integration — Major Projects, National Social Science Foundation of China

- **Focus:** Developing legal frameworks and model evaluation standards for ethical, fair, accurate, and reliable AI deployment, particularly in law enforcement and consumer protection contexts.
- **Contributions:** Spearheaded the translation of high-level legal principles into **auditable protocol rules and operational evaluation criteria**; drafted standards for ongoing oversight (e.g., calibration, selective abstention, error handling); conducted comparative analysis of global AI legislation trends.

- **Key Outputs:** Contributions to research on **"Addressing the Digital Divide in AI Access"**, **"AI in Online Dispute Resolution (ODR)"**, **"Legislative Responses to AI in Law Enforcement"**, and **"Global AI Legislation Trends"**, focusing on equitable access, accountability, and verifiability.

Participatory Audit Mini-Bench for LLM-Assisted Writing

- **Goal:** Investigate whether value-derived operational protocols and calibrated parameters (τ, w) can reduce user over-trust and errors without compromising coverage.
- **Methods:** Preregistered study design; analysis via risk–coverage and verifiability–diversity curves; fully containerized pipeline with artifact hashing for reproducibility; included a small-scale bilingual pilot.
- **Outputs:** Developed a reproducible testing framework comprising a protocol schema, a rule compiler, a comprehensive report with figures, and a make-reproduce script.

Skills

- HCI & Methods: user-centred design; consent/debrief; task & survey design; within-subject and A/B studies; basic thematic analysis.
- Quant & Tools: Python (pandas; basic scikit-learn); R (t-tests, bootstrap CIs, effect sizes); Git; Docker; OSF/Zenodo; preregistration; artefact hashes & manifests.
- Front-end & Prototyping: Figma; HTML/CSS/JS for study UIs and lightweight dashboards; instrumentation & logging.
- Governance & Knowledge Representation: GDPR/PIPL; algorithmic accountability; policy → metrics/protocols; argument maps → YAML rule templates.
- Reproducibility: locked dependencies; deterministic seeds; containerised runs; make-reproduce scripts.

Education

Beijing Institute of Technology — JM (Juris Master) | Beijing, China

Specialization: *Interdisciplinary Law and Computer Science* — GPA 3.73/4.00

Relevant Coursework: Big Data and Visualization； jurisprudence; Data science.

- **Thesis:** "Implementation of the Right to Data Portability" (comparative GDPR/PIPL).

Xi'an International University — B.Econ. International Economics & Trade (Science Track)— GPA 3.70/4.00 | Xi'an, China

- **Core courses:** Calculus; Linear Algebra; Probability & Statistics; Econometrics; C Programming; Image Processing; Web Development.

Awards & Activities

Innovation & Entrepreneurship First Prize (2023); PKU Fabao Outstanding Intern (2023); Gold Award—"Jingkai Cup" Innovation & Entrepreneurship Challenge (2020); Internet+ Beijing Prize (2021); China Outstanding Volunteer Award(2013); Math Modeling Second Prize (2010)

Keywords (ATS-Friendly)

HCI, XAI, calibration, selective abstention, uncertainty, risk-coverage, diagrammatic reasoning, argument mapping, YAML protocols, governance, auditability, reproducibility, trust/reliance studies, rationale logging, operating-point selection (τ, w)