Fair Contextual Bandits for Equitable Diagnostic Decision-Making Under Missing Context

Piter Z. Garcia Bautista
MS Data Science / Bioinformatics
Rochester Institute of Technology
pizg8794@g.rit.edu

Dr. Daniel Krutz, Travis Desell
Department of Software Engineering, Data Science
DSCI 601 Project Advisors
dxkvse@g.rit.edu, tjdvse@g.rit.edu

This project will develop a practical, reproducible framework for **contextual multi-armed bandits** (iC-/C-MABs) to make sequential decisions under uncertainty and limited resources while treating **algorithmic fairness** as a first-class objective. The work on this generic framework focuses on measuring and improving fairness across the clinical (diagnostic-like) and quantum domains, through the following deliverables: i) a simulation-first diagnostic-like sequential decision environment with controllable context missingness/uneven measurement and distribution shift, ii) a fairness aware and contextual MABs evaluation stack (baselines + contextual policies) with time-evolving group disparity reporting and at least one mitigation mechanism, and iii) a transfer study demonstrating the same policy stack in a quantum-network routing + qubit allocation simulator.

Background: Many diagnostic workflows involve sequential choices about tests, models, and retesting under uncertainty, budgets, and distribution shift. These settings can amplify inequities when some groups systematically have lower-quality context or different error profiles. This proposal frames this as a contextual bandit problem: at each step, choose an action (an "arm") given observed context to maximize utility while controlling fairness gaps. Multi-armed bandits formalize online decision-making with exploration–exploitation tradeoffs, and contextual bandits condition decisions on side information (context) that can improve sample-efficiency and stability. In diagnostics, context may include patient features, test and sample-quality indicators, and operational constraints, but access to context can be incomplete or systematically noisier for some populations. This creates both performance risk and fairness risk: aggregate optimization can hide subgroup error spikes (e.g., false-negative gaps) unless the system is explicitly monitored and constrained.

This proposal draws on contextual bandit methods for sequential decision-making with side information [3, 4] and on algorithmic fairness work that defines and measures group-based error disparities [5]. Practically, it builds on an existing quantum-network routing + qubit allocation simulator and on prior work, *Equitable Bioinformatics* [2], operationalizing fairness audits and mitigation for diagnostic-relevant pipelines. The novelty is the **integration**: evaluating context aware MABs policy choices in a diagnostic-like sequential environment while reporting and mitigating fairness disparities **over time** rather than only post-hoc. Unlike prior work that often reports utility-only bandit performance or post-hoc fairness for static predictors, this project makes the time-evolving utility–fairness tradeoff explicit and tests transfer across both testbeds.

Scientific Merit: The core scientific question is: *when does informative context (and how it is modeled) materially reduce disparity and error in sequential decision-making under shift?* This is challenging in modern contextual-bandit settings because (1) feedback is partial and delayed (bandit feedback), (2) distributions shift, and (3) fairness constraints can conflict with pure utility optimization. The key innovative component is that a **quantum-network routing and qubit-allocation environment** is treated as a first-class testbed: the same fairness-aware contextual bandit stack is evaluated in (i) a diagnostic-like simulation and (ii) a quantum routing + qubit allocation simulator to test robustness and transfer across radically different domains.

Broader Impacts: If successful, this work provides a concrete, reproducible framework for fairness-aware sequential decision systems in diagnostics and other public-health settings (including COVID-style test allocation) where both resources and context quality are limited. It also advances trustworthy learning-based control in quantum networking by adding explicit disparity monitoring and constraints to online routing/allocation policies, producing reportable utility–fairness tradeoffs under shift. For this project, the result is a reusable evaluation harness (two testbeds + shared policy API) that supports advisor-driven experimentation and future publication-quality benchmarking of fairness-aware bandit policies.

Approach: The goal is to implement and evaluate fairness aware and contextual MABs policies across two testbeds (diagnostic-like sequential decision simulation + quantum-network routing/qubit allocation simulator), and to report utility–fairness tradeoffs under shift. The work is structured to be feasible without sensitive clinical data access (simulation-first), while remaining extensible to approved open datasets. The tasks are:

- **Build diagnostic testbed:** implement a sequential decision simulation (test/model choice, retesting, pipeline choice) with distribution shift and controllable context missingness/uneven measurement.
- **Build quantum testbed:** integrate policies into the quantum routing + qubit allocation simulator; define flow/node groups and track service equity across groups.
- **Implement bandit policies:** non-contextual baselines (epsilon-greedy, UCB, Thompson), a contextual baseline (LinUCB-style), and one informed contextual (iCMAB-style) policy under a shared interface.
- **Define reward + metrics:** utility with explicit cost/latency and safety weighting; report regret/utility plus time-evolving fairness metrics (group-wise FNR/FPR gaps and at least one additional criterion).
- **Add fairness mitigation:** implement one fairness-aware mechanism (constraint/penalty or calibration) and quantify utility–fairness tradeoffs across both testbeds.
- **Package for reproducibility:** Python 3.11+ modular scripts (not one notebook), fixed seeds + configs, and a minimal test suite (sanity checks for simulations and metric computation).

Phase 1 includes environment + baselines + initial fairness audit/mitigation; Phase 2 includes robustness under shift + ablations + packaging for code review and final report/demo.

References:

[1] P. Z. Garcia Bautista. *Qubit Allocation in a Quantum Network using Stochastic Bandits*. Internal manuscript (course/research draft), 2026.

[2] Garcia Bautista, P. Z. (2025). *Equitable Bioinformatics: Enhancing Diagnostic Decision-Making through RNA and Biomarker Data*. Unpublished paper, Rochester Institute of Technology.

[3] L. Li, W. Chu, J. Langford, and R. Schapire. A contextual-bandit approach to personalized news article recommendation. In *WWW*, 2010.

[4] Y. Abbasi-Yadkori, D. Pál, and C. Szepesvári. Improved algorithms for linear stochastic bandits. In *NeurIPS*, 2011.

[5] M. Hardt, E. Price, and N. Srebro. Equality of opportunity in supervised learning. In *NeurIPS*, 2016.