Bias and Fairness in Large Language Models for Financial Compliance: Risks and Mitigation

Marie Levesque
MSc Data Science, University of Essex
Online – Octobre 2025

Why this research matters

Finance is highly regulated

LLMs can automate compliance tasks

Bias: risk to fairness, trust and legality

Regulatory Frameworks
Solvency II • IFRS 17 • Basel III • AMLD

↓

LLM Processing
GPT-4 · FinGPT · BloombergGPT

↓

Outputs
Obligations • Risk classification • Summarisation

References: ESMA & Turing, 2025.

Bias in compliance remains under-researched

Research Problem

Bias well studied in:

- Healthcare
- Hiring
- Criminal Justice

Gap: Little research in
Financial Compliance
(High stakes, under-studied)

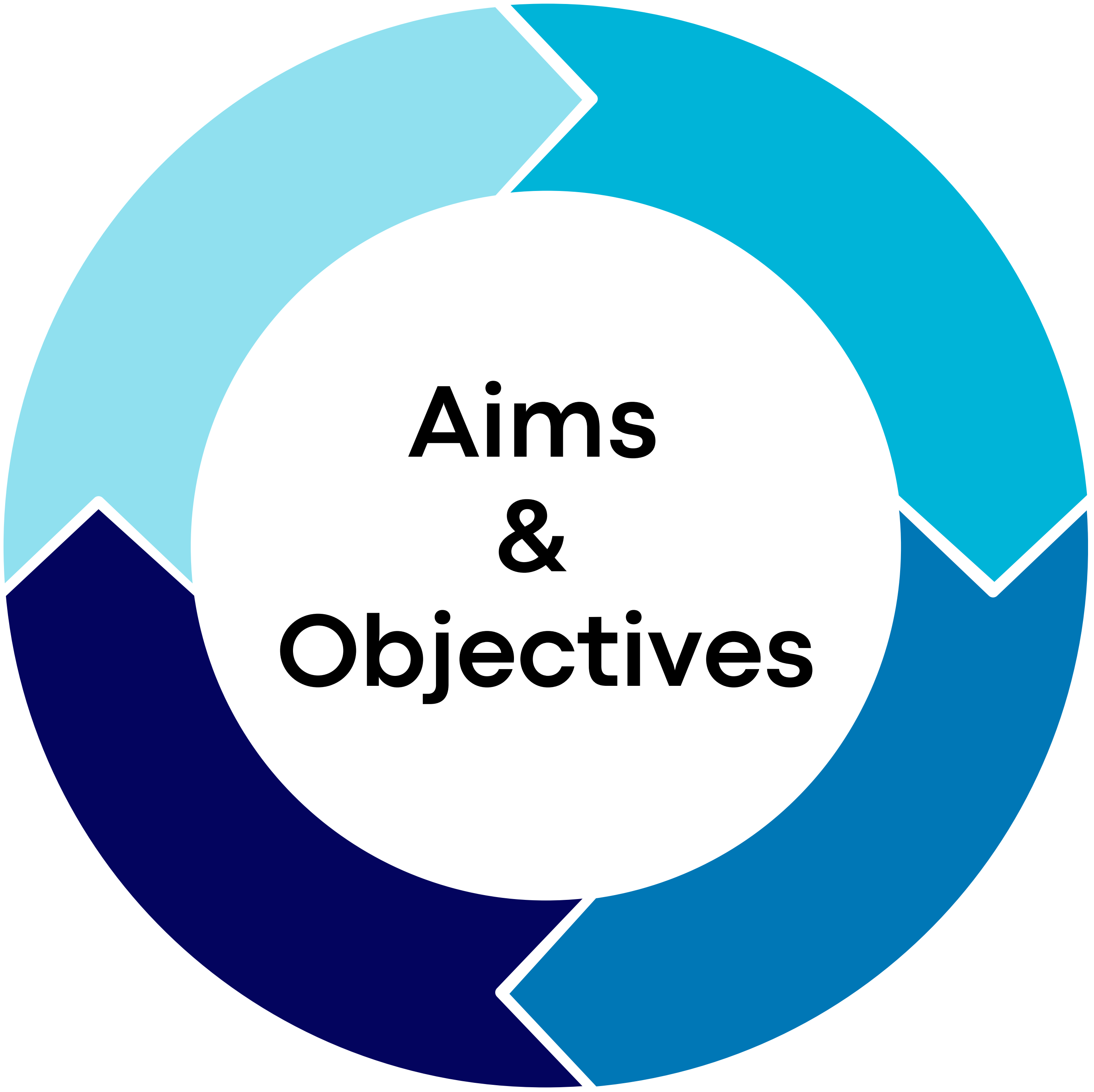
References: Bender et al., 2021; Mehrabi et al., 2021

How can bias and fairness be evaluated and mitigated in LLMs for compliance?

Types of bias?

General vs domain-specific models?

Effective mitigation strategies?



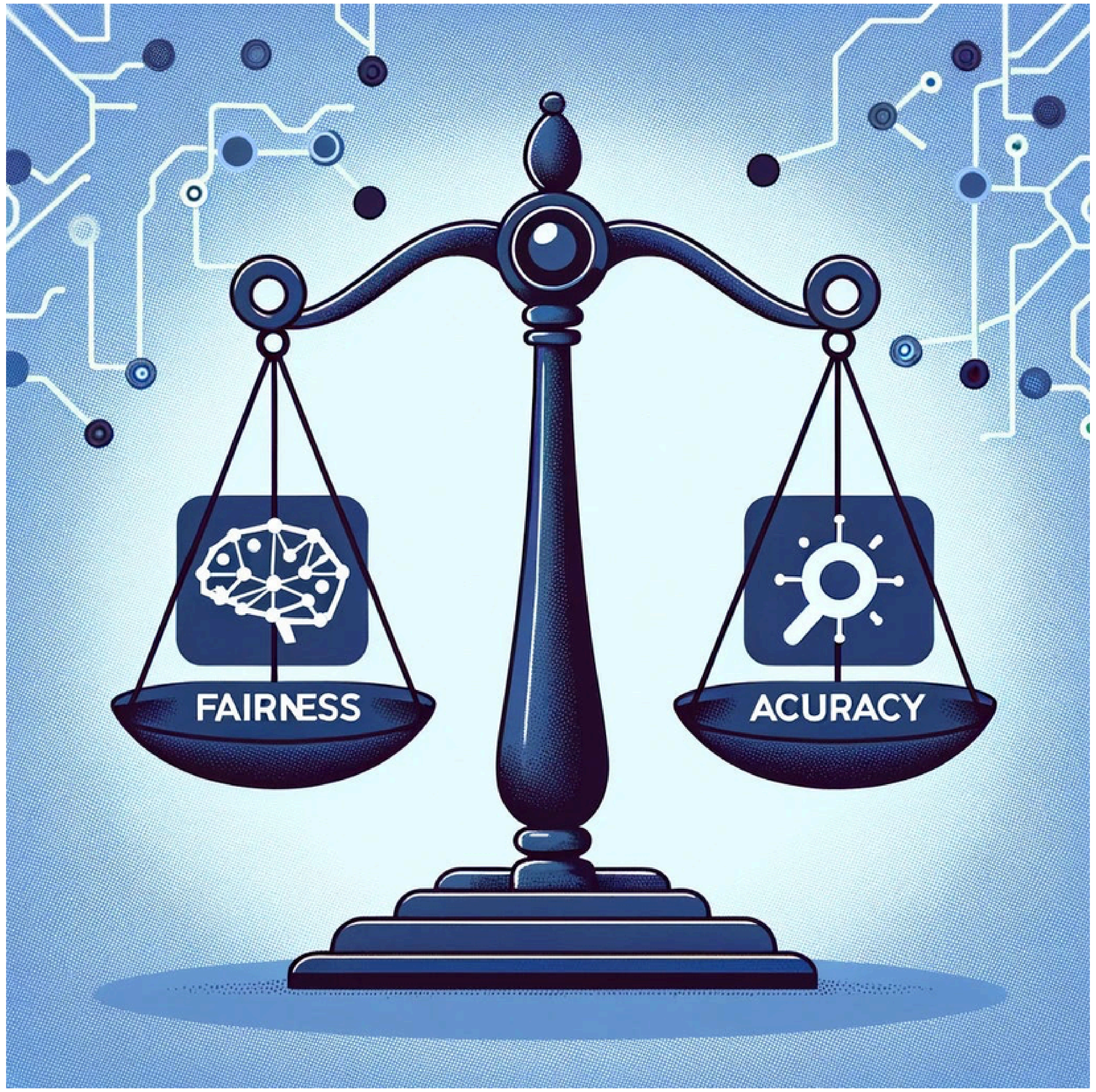
1. Detect bias
2. Compare models
3. Test mitigation
4. Build auditing framework

Literature Review (Bias & Fairness)

"Stochastic parrots" risk
(Bender et al., 2021).

Fairness metrics
(Mehrabi et al., 2021).

Context-dependence
(Blodgett et al., 2020).



Literature Review (Governance)

Right to explanation
(Goodman & Flaxman, 2017).

GDPR/AI Act
Finance = “high risk”
(Veale & Edwards, 2018).

RegTech governance issues
(Zetzsche et al., 2020).

Literature Review (LLMs in Finance)

Model	Type	Training Data	Strengths	Limitations
GPT-4	General-purpose	Broad internet corpus	Strong general NLP, versatile, state-of-the-art reasoning	Not finance-specialised, may hallucinate
FinGPT	Finance-specific (open-source)	Financial texts & datasets	Domain-tuned, open-source, adaptable for fine-tuning	Still emerging, smaller scale vs GPT-4
BloombergGPT	Finance-specific (proprietary)	Large-scale financial data	High accuracy on financial NLP benchmarks, industry-validated	Closed model, limited access

BloombergGPT & FinGPT: strong performance (Bloomberg, 2023).

Empirical results but little fairness testing.

ESMA & Turing (2025): call for fairness in financial AI.

Research Design Overview

Regulatory Data

Solvency II · IFRS 17 · Basel III · AMLD

Public documents (PDF/HTML → text)

Models

GPT-4 (general)

FinGPT (finance)

BloombergGPT (benchmarks)

Tasks

Obligation extraction

Risk classification

Regulatory summarisation

Outputs assessed for Accuracy (P/R/F1, ROUGE) and Fairness (consistency, calibration, parity)

Evaluation & Fairness Metrics

Accuracy Metrics

Precision
Recall
F1

ROUGE/BERTScore

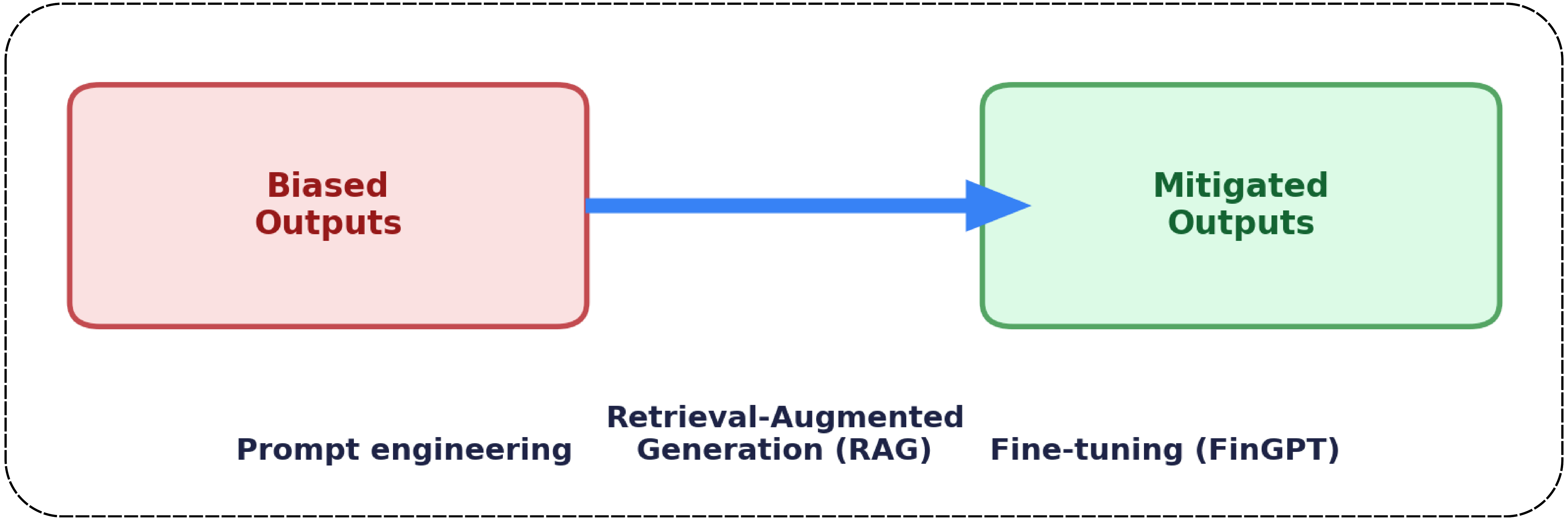
Fairness Metrics

Consistency
Calibration

Parity Skew Review
(Mehrabi et al., 2021)

References: Raji et al., 2020

Mitigation Strategies



Artefact & Framework

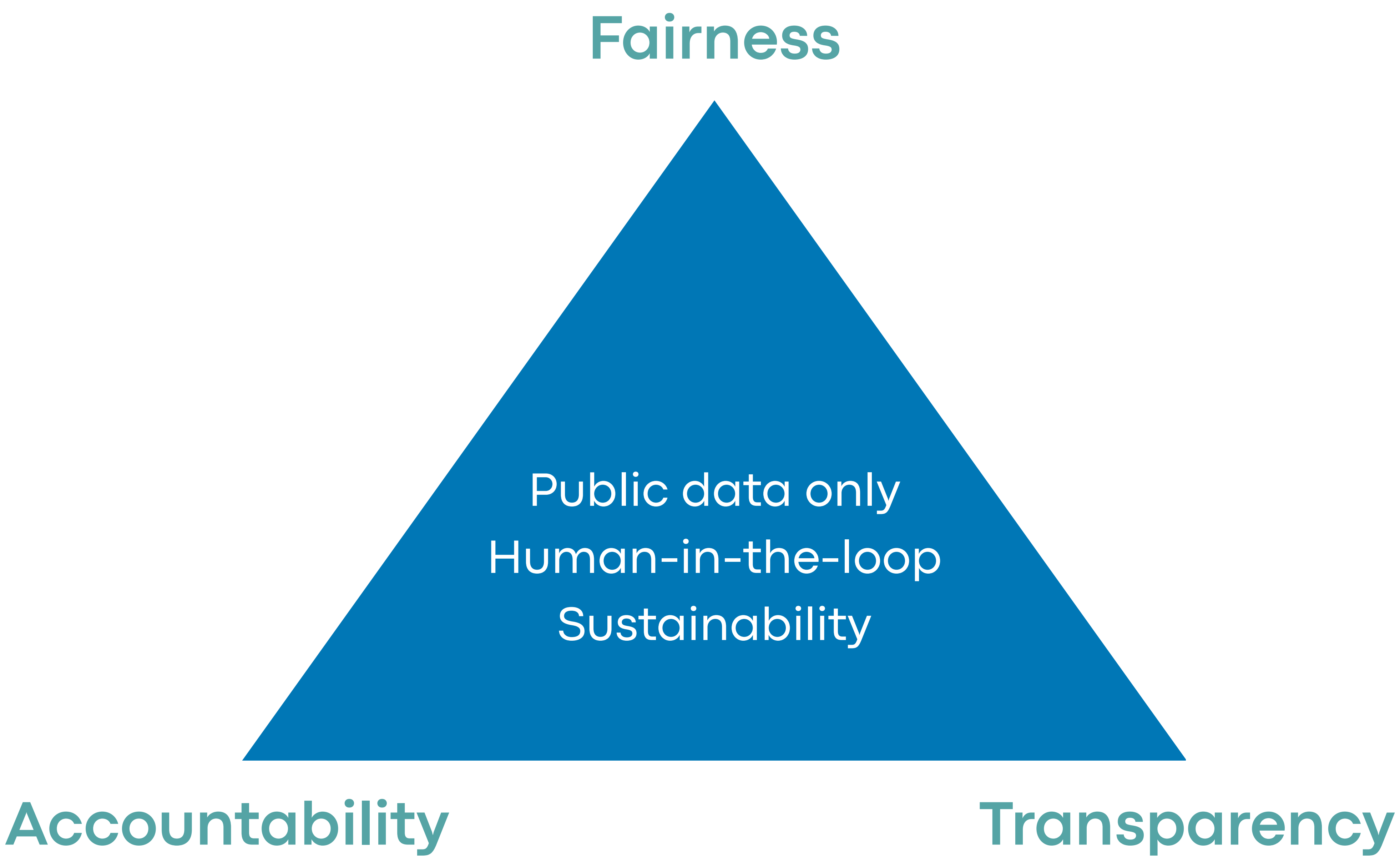
Framework Document

Guidelines
Indicators

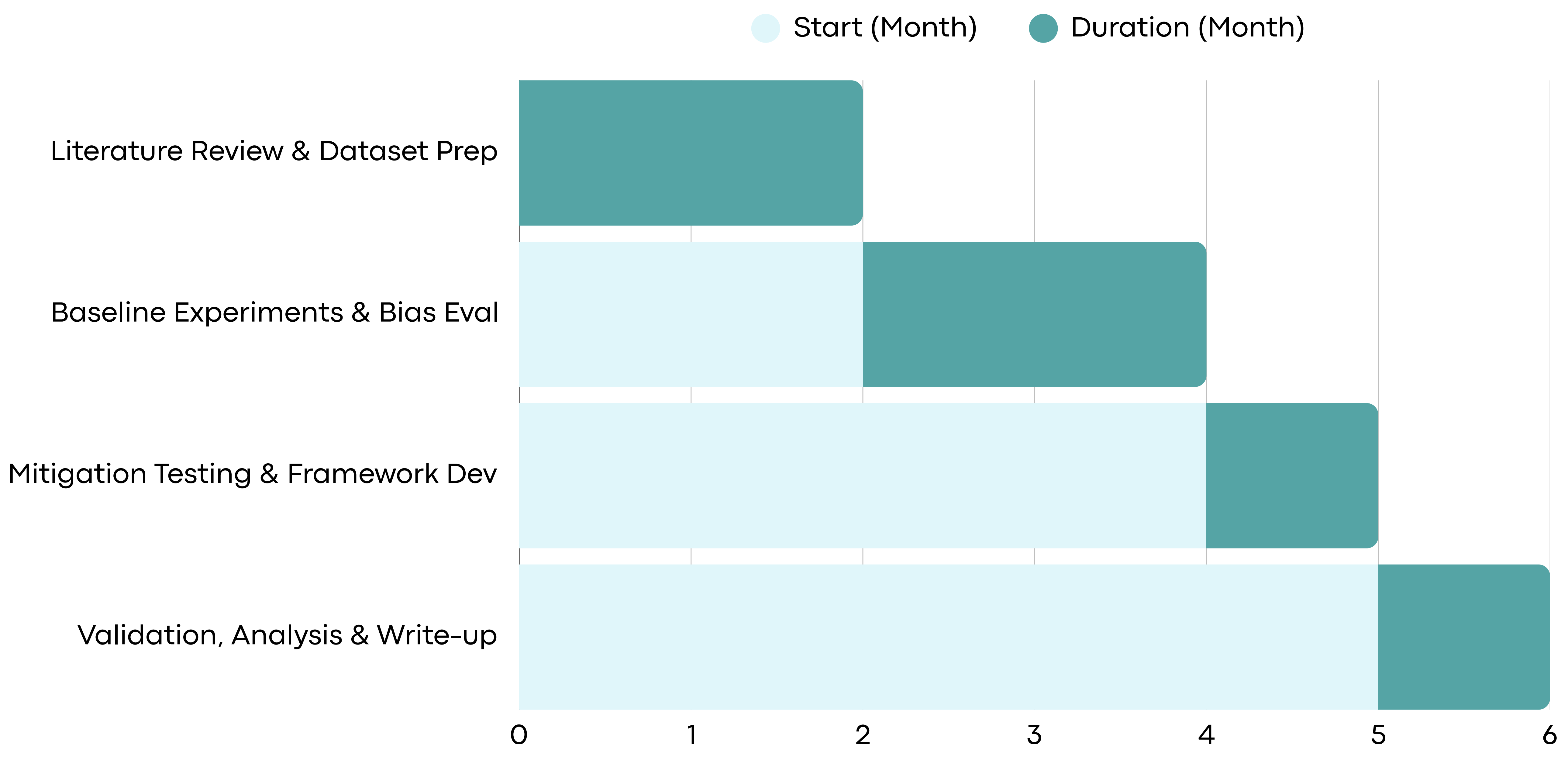
Prototype Tool

Python notebook
Fairness checks

Ethical Considerations

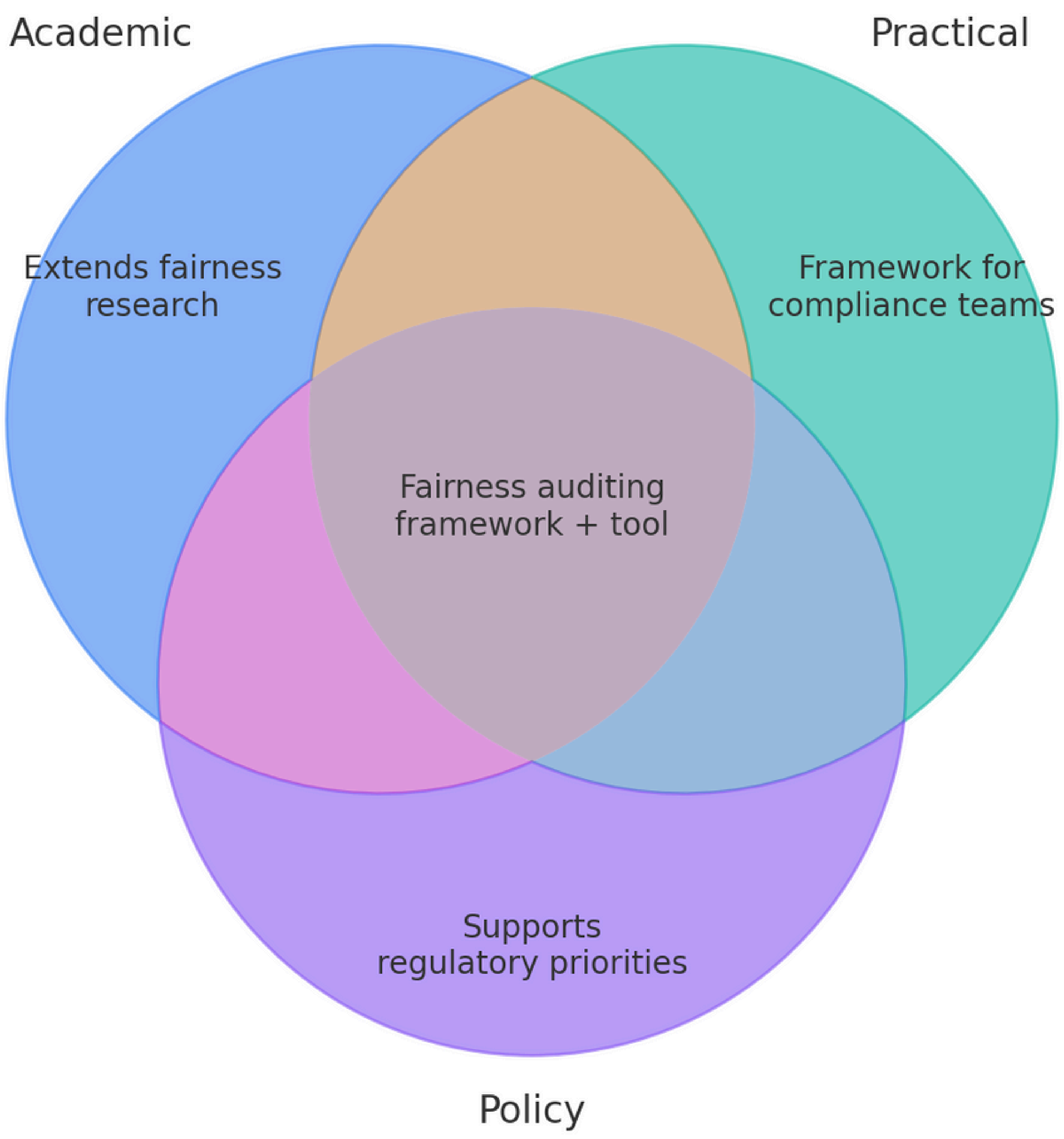


Timeline



Start (Month)
Duration (Month)
Literature Review & Dataset Prep
Baseline Experiments & Bias Eval
Mitigation Testing & Framework Dev
Validation, Analysis & Write-up
0
1
2
3
4
5
6

Contribution



Conclusion

Bridging Data Science, Finance, and Regulation

Key References

Bender, E.M. et al. (2021). On the dangers of stochastic parrots. FAccT Conference.

Blodgett, S.L. et al. (2020). Bias in NLP: A critical survey. *TACL*, 8, pp. 545–565.

Bloomberg (2023). *BloombergGPT: A large language model for finance*. arXiv.

ESMA & The Alan Turing Institute (2025). *Large language models in finance*. ESMA Report.

Goodman, B. and Flaxman, S. (2017). EU regulations on algorithmic decision-making. *AI Magazine*, 38(3).

Jobin, A., Ienca, M. and Vayena, E. (2019). Global landscape of AI ethics guidelines. *Nat. Machine Intelligence*, 1(9).

Leslie, D. (2019). *Understanding AI ethics and safety*. The Alan Turing Institute.

Mehrabi, N. et al. (2021). Survey on bias and fairness in ML. *ACM Computing Surveys*, 54(6).

Raji, I.D. et al. (2020). Closing the AI accountability gap. *FAccT Conference.*

Veale, M. and Edwards, L. (2018). Automated decision-making and profiling. *CLSR*, 34(2).

Zetzsche, D.A. et al. (2020). Regulating FinTech: Lessons from Africa. *Law and Contemporary Problems*, 83(2).