Navigating Human Bias: A Value Statement and Strategies from Financial Services Engineering

Manas Maheshwari | July 2026

Personal Value Statement

I am a backend engineer in financial services, where the systems my teams build decide things about real people: whether a transaction looks fraudulent, whether an account gets flagged, what permissions a customer holds. In this field, human bias enters twice. It enters the systems, through the data we select, the labels we trust, and the thresholds we set. And it enters the engineering process itself, through anchoring on the first hypothesis in an incident, deferring to the most senior voice in a design review, and over-trusting the output of automated tools. My core value is that consequential decisions, whether made by a person or a model, must be explainable, contestable, and correctable. Fairness is not something I assume about a system; it is something I have to be able to demonstrate, and something a customer or colleague must be able to challenge.

Field-Specific Strategies

1. Design AI tools that resist automation bias.

- **Show evidence, not verdicts:** I build AI tooling for engineers, and my rule is that a tool must show its reasoning (the query plan, the metric, the log line) alongside its conclusion, so the human stays a reviewer rather than becoming a rubber stamp.
- **Treat overrides as data:** When users ignore or reverse a tool's recommendation, capture why. Override patterns are the earliest and cheapest signal of a biased or miscalibrated system.

2. Treat training data and labels as human artifacts, not ground truth.

- **Audit the labels' origin story:** In fraud detection, labels arrive late, skewed, and shaped by past enforcement choices; a model trained on them inherits every historical decision. Document where labels come from and who decided them before trusting what they teach.
- **Evaluate by segment, not just in aggregate:** A model with strong overall accuracy can still fail specific customer groups. Fairness checks across segments belong in the deployment gate, not in the retrospective.

3. Engineer the human process against its own biases.

- **Blameless, anchor-aware postmortems:** In incident reviews, we name the first hypothesis the team anchored on and what evidence finally broke it. Making anchoring visible is how a team learns to distrust it.
- **Structure reviews so juniors speak first:** In design reviews, collecting written input before discussion and inviting the least senior reviewers to speak first counters authority bias. Some of the best catches I have seen came from engineers who would have stayed quiet had a senior voice gone first.

4. Make decisions contestable by default.

- **Log for the challenge, not just the audit:** Every automated decision about a customer or an account should leave a record sufficient for a human to reconstruct and dispute it. In a regulated industry this is compliance; everywhere it is simply fairness with an interface.
- **Treat compliance and operations as bias detectors:** The teams fielding customer disputes see a model's blind spots before its builders do. A standing feedback channel from them to engineering closes the loop.

Leading well amid human bias, in my field, is less about eliminating bias, which no amount of training fully does, than about building systems and team habits that catch it early, surface it visibly, and make it correctable.