Embedding Fairness: A 7-Stage Blueprint for AI Governance

DESIGN & DEVELOPMENT
(Stages 1-3)

VALIDATION & DEPLOYMENT
(Stages 4-5)

CONTINUOUS SURVEILLANCE
(Stages 6-7)

Stage 1: Establish Fairness Criteria

Define project-specific fairness goals and metrics.

Data Steward

Data Card

Systematic review of data for bias.

AI Development Lead

Model Card

Detailed record of model design and decisions.

AI Validation Officer

Validation Report

Independent testing against fairness and performance criteria.

Stage 5: Human Oversight & Deployment

Manual review and controlled rollout with meaningful human control.

Stage 6: Track Performance Drift

Permanent monitoring officers track models for fairness regression.

Stage 7: Trigger Governance Review

Governance committee convened if threshoids are breached.

BUILDING ORGANIZATIONAL RESILIENCE

Eliminate Role Ambiguity

Assigning specific names to accountable roles prevents responsibility diffusion.

Break Down Functional Silos

Tech · Legal · Business

Mandatory cross-functional sign-off gates ensure shared visibility.

Ensure Regulatory Alignment

EU AI Act · NIST AI RMF

Every stage creates documentation artifacts that satisfy legal and framework obligations.

NotebookLM