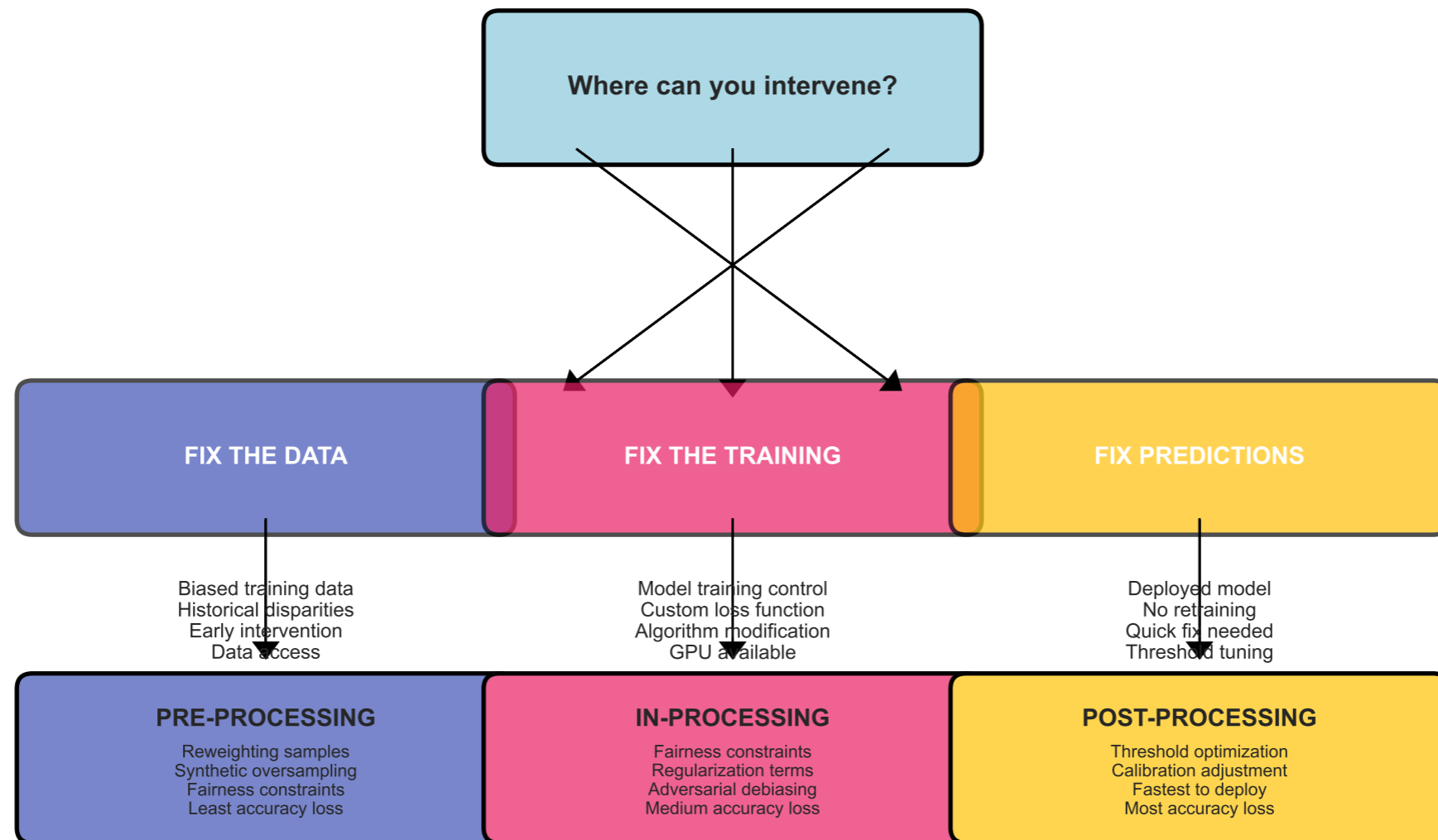


# When to Use Which Fairness Intervention: Decision Framework



## Additional Considerations

Intervention Stage: Have data access → Pre-processing; Model training → In-processing; Deployed model → Post-processing  
Accuracy Trade-off: Minimize loss → Pre-processing (best); Balance → In-processing; Accept loss → Post-processing  
Fairness Definition: Demographic parity → Pre-processing; Equal opportunity → In-processing; Equalized odds → Post-processing  
Computational Cost: Limited compute → Pre-processing (once); GPU available → In-processing; Inference only → Post-processing  
Transparency: Need audit trail → Pre-processing (data changes visible); Black box OK → In/post-processing  
Stakeholders: Data scientists → Pre/in-processing; ML ops/deployment → Post-processing

*Principle: Fix bias at the earliest stage possible - pre-processing preferred, post-processing as last resort*