Name: Joel Livingstone Kofi Ackah

QuickLoan Mobile Ethical Data Review

Deliverable 1: Governance Review Card

Section	Issue / Definition	Impact	Suggested Fix / Mitigation
1. Data Quality Risk	Customer data entering the pipeline is often incomplete or inconsistently formatted (e.g., varied phone prefixes or null income fields).	Produces unreliable loan scores, increasing default risk and leading to unfair automated rejections.	Implement field-level validation at the API Gateway and a data quality dashboard in the Preprocessing Service to flag malformed records before they reach the ML model.
2. Legal & Compliance Risk	Excessive collection of personal data (e.g., full contact lists and GPS logs) without explicit consent violates Sections 18 and 27 of Ghana's DPA.	Exposure to heavy regulatory penalties from the Data Protection Commission and severe reputational damage.	Apply the **Data Minimization** principle by collecting only essential loan-related data and inserting a consent management layer between the API Gateway and Raw Data DB.
Data Classification	**Sensitive**.	Personal identifiable information (PII) including financial records, national IDs, and location data requires the	Implement encryption at rest and in transit, combined with strict Role-Based Access Controls (RBAC).

		highest protection standards.	
3. Bias & Fairness Risk	The ML model may inherit historical lending biases, leading to automated decisions that discriminate against specific regions, genders, or age groups.	Systematic exclusion of creditworthy applicants from vulnerable populations, violating anti-discrimination provisions.	Conduct quarterly fairness audits and remove bias-prone proxy variables during feature engineering to ensure equitable outcomes.
Source of Bias	Historical training data and proxy features such as geolocation, social graph size, and mobile phone metadata.	Socioeconomic status and regional origin become indirect predictors of creditworthiness.	Use adversarial debiasing techniques and establish a diverse model review committee to identify bias before deployment.
4. Storytelling / Reporting	The lack of transparent reporting on automated loan outcomes prevents leadership from seeing if the ML model is drifting toward biased or unethical decisions.	Without clear visibility, the company cannot prove it is complying with Ghana's DPA (Act 843) or demonstrate fairness to investors and regulators.	**Recommendation:** Establish a monthly "Ethical Performance Dashboard" that highlights fairness metrics alongside traditional business KPIs.
Metric to Monitor	**Demographic Approval Parity Index (DAPI):** The	A score of 1.0 indicates perfect parity; scores	**Goal:** Ensure that automated decisions do not systematically

	ratio of loan approval rates across demographic groups (e.g., gender, region, age cohort).	below 0.85 or above 1.15 indicate significant disparity requiring immediate investigation.	disadvantage specific populations based on socio-economic or regional traits.
Visualization Type	**Grouped Bar Chart** showing approval rates by demographic category.	This visualization should include a reference line at the overall approval rate and a trend line tracking the Parity Index over time.	**Dashboard View:** Use color-coding (Green/Yellow/Red) to flag segments where approval rates drop significantly below the average.
Why It Matters	This metric makes algorithmic fairness measurable and actionable, allowing QuickLoan to detect and correct discriminatory patterns before they lead to regulatory violations.	It transforms abstract ethical principles into concrete operational targets that non-technical executives can easily understand and act upon.	**Benefit:** Strengthens customer trust and ensures long-term regulatory accountability under Act 843.

Deliverable 2: Corrected Data Flow Diagram Annotations

Step 1: Implement Data Minimization (User Mobile App)

- **Correction:** Removed the collection of entire contact lists, GPS logs, and device metadata. The app now only collects essential data required for credit assessment, such as verified income, national ID, and repayment history.
- **Why:** This ensures compliance with the **Data Minimization** principle of Ghana's Data Protection Act (Act 843, Section 27) and reduces the risk of unlawful personal data processing.

Step 2 to 3: Insert Consent Management Layer (API Gateway to Raw Data DB)

- **Correction:** Added a mandatory consent validation step before any data is stored in the database. This captures a timestamped record of the user's explicit opt-in for specific data categories.
- **Why:** Ghana's Act 843 (Section 18) requires informed and specific consent before processing personal data. This provides a verifiable audit trail for regulatory compliance.

Step 3: Apply Data Classification and Retention (Raw Data DB)

- **Correction:** Classified the database as **Sensitive** and implemented a defined retention schedule (e.g., 5–7 years). Data is now encrypted at rest and in transit.
- **Why:** Storing financial and identity data without classification exposes it to misuse. Retention limits prevent the indefinite storage of PII, adhering to storage limitation principles.

Step 5: Add Data Quality and Schema Validation (Preprocessing Service)

- **Correction:** Implemented automated validation rules to check for completeness and consistent formatting (e.g., standardizing phone numbers and currency) before data reaches the ML model.
- **Why:** This prevents "garbage-in, garbage-out" failures, ensuring the ML model produces reliable credit scores and prevents unfair rejections due to malformed data.

Step 7: Enable Decision Logging and Transparency (Decision Service)

- **Correction:** Added a mechanism to log every automated decision, including the influencing factors and the applicant's demographic metadata.
- **Why:** This creates an audit trail for **Fairness Audits** to detect algorithmic bias and provides a basis for an appeal mechanism for rejected customers.

Step 9: Implement Data Masking (Analytics DB)

- **Correction:** Applied data masking and anonymization techniques (e.g., replacing national IDs with codes) before storing data for secondary analysis.
- **Why:** Analytics does not require identifiable data. Masking reduces the impact of a potential data breach and ensures the privacy of individual customers is maintained during business reporting.

Deliverable 3: Summary of Review Process

To conduct this governance review, I mapped QuickLoan Mobile's automated pipeline against the **Data Lifecycle** stages: collection, storage, processing, and secondary use. By tracing the flow of information, I identified critical gaps where data was moving between systems without adequate controls. At the collection stage, I applied the principle of **Data Minimization**, determining that collecting entire contact lists and GPS logs was excessive for credit scoring and violated **Ghana's Data Protection Act (Act 843)**.

Using **Data Classification** principles, I categorized the loan application dataset as **Sensitive**. This classification revealed that current storage practices were insufficient; I subsequently recommended encryption and strict **Role-Based Access Controls (RBAC)** to protect PII from unauthorized access. During the processing and decision-making phases, I identified a lack of transparency and data quality checks, which allowed malformed data to potentially corrupt the ML model's accuracy. To rectify this, I integrated automated validation at the preprocessing stage and a decision-logging audit trail in the rules engine.

The proposed **Demographic Approval Parity Index (DAPI)** is central to ensuring ethical governance. By quantifying approval rate disparities across demographics, this metric transforms abstract fairness concepts into measurable performance targets. Weekly monitoring via grouped bar charts allows leadership to proactively detect and address algorithmic bias, ensuring that automated decisions remain transparent, accountable, and legally compliant with the standards set by the Data Protection Commission.