

# DATA COLLECTION AND AI IMPROVEMENT AGREEMENT

This Data Collection and AI Improvement Agreement ("Agreement") is entered into as of [DATE] by and between:

**OKO Investments** ("DataMigrate AI", "Provider", "We", "Us")

Address: [Provider Address]

CVR: [Company Registration Number]

AND

**[CUSTOMER NAME]** ("Customer", "Client", "You")

Address: [Customer Address]

CVR: [Customer Registration Number]

## 1. DEFINITIONS

**"Schema Data"**: Database structure information including table names, column names, data types, stored procedure logic, and SQL transformation patterns. Schema Data does NOT include actual row-level data or business content.

**"Anonymized Data"**: Schema Data that has been processed to replace all customer-specific identifiers (table names, column names, procedure names) with generic tokens (Table\_A, col\_001, etc.).

**"Named Data"**: Schema Data where original table and column names are preserved, but no actual business data values are included.

**"Training Data"**: Anonymized or Named Data used to improve DataMigrate AI's machine learning models.

**"Row-Level Data"**: Actual data values stored in database tables. This is NEVER collected.

## 2. SCOPE OF DATA COLLECTION

### 2.1 What We Collect

- SQL stored procedure structures and transformation logic
- CREATE TABLE statements and column definitions
- JOIN patterns and table relationships
- Data type mappings between MSSQL and dbt
- Successfully generated dbt model SQL code
- Migration success/failure metadata

### 2.2 What We NEVER Collect

- Actual row-level data (customer records, transactions, etc.)
- Personal identifiable information (names, addresses, SSNs, emails)
- Financial data (account numbers, balances, credit card numbers)
- Health information (medical records, diagnoses)
- Authentication credentials (passwords, API keys)
- Any data classified as sensitive under GDPR Article 9

### **3. DATA COLLECTION OPTIONS**

Customer may select one of the following data collection options:

#### **OPTION A: ANONYMIZED COLLECTION (Default)**

All Schema Data is automatically anonymized before storage. Table names, column names, and procedure names are replaced with generic identifiers. This option provides maximum privacy while still contributing to model improvement.

Example: 'CustomerMaster.credit\_score' becomes 'Table\_A.col\_001'

#### **OPTION B: NAMED COLLECTION (Opt-in)**

Table and column names are preserved in collected data. This improves model quality for industry-specific patterns. Customer receives a 15% discount on services in exchange for this contribution. No actual data values are ever collected.

Benefit: Priority access to industry-specific model improvements

#### **OPTION C: NO COLLECTION (Opt-out)**

Customer may opt out of all data collection. No Schema Data will be retained after migration completion. Service pricing remains unchanged.

## **4. CUSTOMER SELECTION**

Customer hereby selects the following data collection option (check one):

[ ]     OPTION A: Anonymized Collection (Default)

[ ]     OPTION B: Named Collection (15% discount)

[ ]     OPTION C: No Collection (Opt-out)

## **5. DATA SECURITY AND COMPLIANCE**

- 5.1 All collected data is encrypted at rest (AES-256) and in transit (TLS 1.3).
- 5.2 Data is stored in EU-based data centers compliant with GDPR requirements.
- 5.3 Access to training data is restricted to authorized ML engineers only.
- 5.4 Annual third-party security audits are conducted.
- 5.5 Data retention: Training data is retained indefinitely for model improvement. Customer may request deletion at any time per Section 7.
- 5.6 Provider maintains SOC 2 Type II compliance for data handling.

## 6. USAGE RIGHTS

- 6.1 **Training Purpose:** Collected data may be used solely for training and improving DataMigrate AI's machine learning models for SQL migration tasks.
- 6.2 **No Sale:** Provider will not sell, license, or transfer raw training data to third parties.
- 6.3 **Model Ownership:** All trained models remain the intellectual property of OKO Investments.
- 6.4 **Aggregated Insights:** Provider may publish aggregated, anonymized statistics about migration patterns (e.g., '60% of migrations involve stored procedures').

## 7. CUSTOMER RIGHTS

**7.1 Right to Withdraw:** Customer may withdraw consent and request deletion of their contributed training data at any time by written notice.

**7.2 Right to Access:** Customer may request a summary of what data has been collected from their migrations.

**7.3 Right to Change Option:** Customer may change their collection option at any time for future migrations.

**7.4 Right to Audit:** Customer may request an audit of data handling practices with 30 days notice.

## 8. GDPR COMPLIANCE

8.1 **Data Controller:** Customer is the Data Controller for their database content. Provider is the Data Processor for Schema Data collected during migrations.

8.2 **Legal Basis:** Data collection is based on legitimate interest (improving AI services) with customer consent as documented in Section 4.

8.3 **Data Protection Officer:** Provider's DPO can be contacted at [dpo@okoinvestments.com](mailto:dpo@okoinvestments.com).

8.4 **Cross-Border Transfer:** Training data is processed within the European Economic Area. No transfer to third countries without adequate safeguards.

## 9. SIGNATURES

**FOR PROVIDER:**

---

Name: Alexander Garcia Angus

Title: CEO, OKO Investments

Date: \_\_\_\_\_

**FOR CUSTOMER:**

---

Name: \_\_\_\_\_

Title: \_\_\_\_\_

Date: \_\_\_\_\_