Divyavaahini Thyagarajan | EPITA School of Engineering and Computer science | Paris

ShopSmart: Privacy-Centric Customer Rewards Program

1.Dataset Overview

ShopSmart, a retail platforrm has introduced a customer rewards program to enhance customer loyalty through personalized offers, discounts, and reward points. The Dataset,**customer_rewards_data.csv**, contains information about the customer transactions, user details,purchase behavior and payment details. Each record provides valuable insights into user behavior, enabling analyses of shopping patterns, pricing strategies, and customer retention.However, from a **data privacy perspective**, it contains several **personally identifiable information (PII)** and **identifiers**, which require careful evaluation before any processing or sharing.

2.Dataset Composition

The dataset consists of **500 records and 16 columns**. Based on their data types and privacy categories the columns can be categorized as follows:

Column	Description	Data Type	Privacy Category
Name	Name of the customer	Categorical(Text)	Direct Identifier
Age	Age of the customer	Numerical	Quasi-Identifier
Location	Customer location or city	Categorical(Text)	Quasi-Identifier
Gender	Gender of the customer	Categorical(Text)	Quasi-Identifier
Transaction Date	Date of the transaction	Categorical(Date)	Indirect Identifier
Items Purchased	List of items bought	Categorical(Text)	Indirect Identifier
Quantity	Quantity of items purchased	Numerical	Indirect Identifier
Total Cost	Total cost of the transaction	Numerical	Indirect Identifier
Payment Method	Payment type used (e.g., card, cash)	Categorical(Text)	Indirect Identifier
Points Earned	Loyalty points earned in the transaction	Numerical	Indirect Identifier
Points Redeemed	Loyalty points redeemed in the transaction	Numerical	Indirect Identifier
Reward Chosen	Reward selected by the customer	Categorical(Text)	Indirect Identifier
Feedback on Reward	Customer feedback for the reward	Categorical(Text)	Indirect Identifier
Marketing Email Opened	Whether the marketing email was opened	Boolean	Indirect Identifier
Marketing Email Clicked	Whether the marketing email was clicked	Boolean	Indirect Identifier
Customer Service Inquiry	Whether the customer contacted support	Boolean	Indirect Identifier

Table1: Customer Transaction Data with Privacy Categories

3.Key Highlights

- **Variety of Data:** It includes numeric, categorical, text, and boolean fields, making it ideal for analytics, customer segmentation, and data privacy studies.
- **Data Privacy Considerations:** Some fields contain direct or quasi-identifiers, so anonymization is necessary before sharing to protect personal information.

4.Data Flow Diagram:

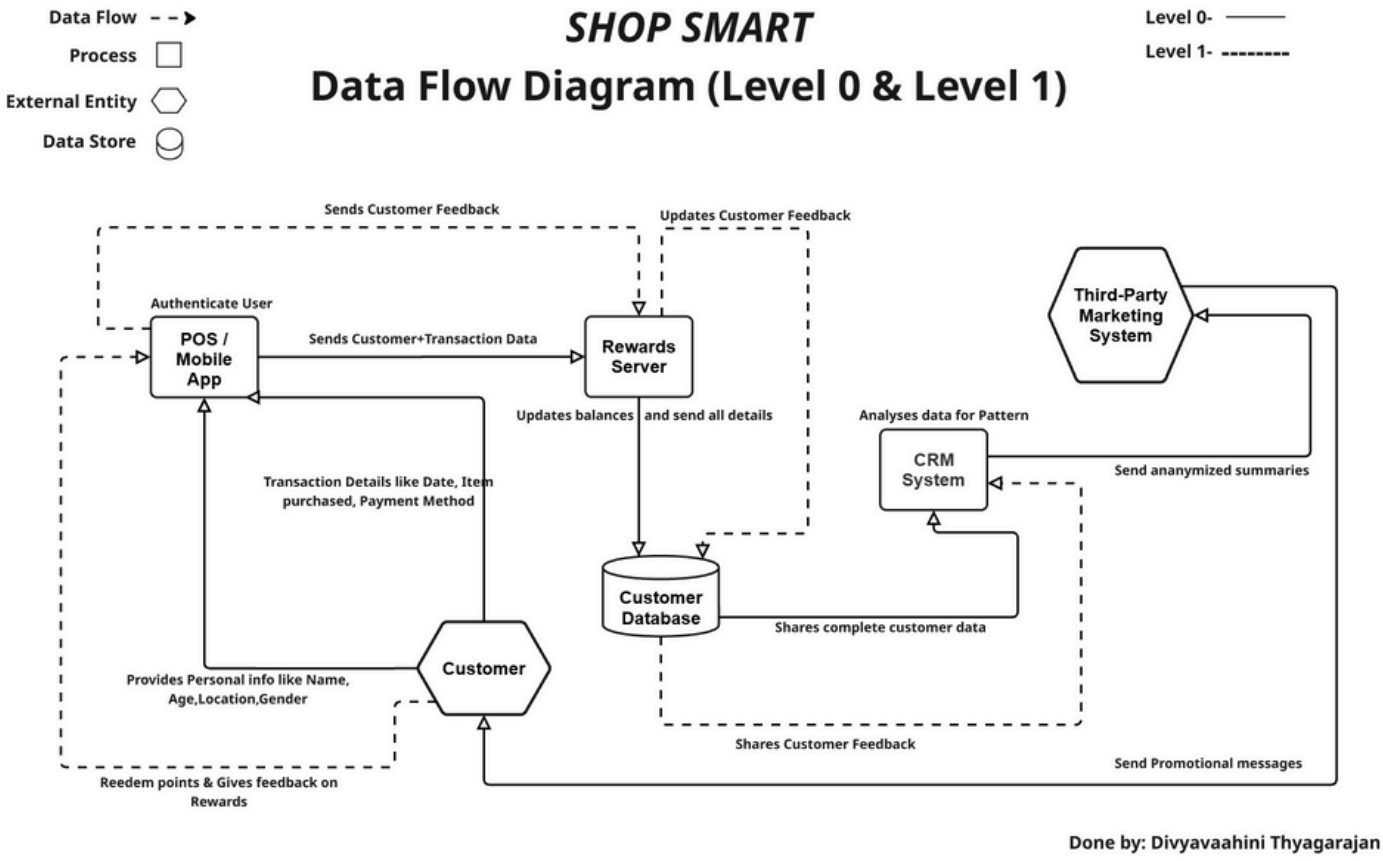

Figure 1: High- Level Data Flow Diagram for ShopSmart

The DFD illustrates how SmartShop manages customer purchases, rewards, and feedback through connected systems like the POS/Mobile app, Rewards Server, and CRM. It ensures secure data handling and privacy while enabling personalized marketing and customer engagement.

5.Threat Modelling Using LINDDUN Threat Categories

The **LINDDUN framework** is applied to the SmartShop Rewards System to identify and reduce privacy risks across data collection, processing, and other components.

Threat Category 1: **Linking**

- Customer activity (like purchases, app usage, or rewards redemption) could be linked across different visits. This leads to combine data and identify full profile of shopping habits in a period of time.
- **Affected Component:** Reward server Logs, Third party marketing system and CRM
- **Impact:** Additionally, data can be linked together using profiling and other sophisticated techniques to identify product prefernce and shopping pattern.

Threat Category 2: **Identifying**

- Other information, such as age, location, or purchasing patterns, can help identify customers even if their names are not kept in some systems.
- Consumers who have unusual purchasing habits or unusual product preferences might be identified in reports because they stand out.
- **Affected Component:** Customer Database, CRM, Third party marketing system
- **Impact:** People can know who you are and breaking law.

Threat Category 3: **Non-repudiation**

- Detailed activity of the customer transaction is maintained indefinetly.
- Permanent Linking of transaction to the customer identity.
- **Affected Component:** Reward Server Logs
- **Impact:** Unwanted persistent records, Privacy risks

Threat Category 4: **Detecting**

- Even if data is encrypted, someone who is observing network traffic could identify when a customer is active on the app, checking products, or redeeming the rewards.
- Patterns of this activity could reveal the customers frequent purchasing time or program participation
- **Affected Component:** POS/App, Third party marketing System
- **Impact:** Behavioral inference, Targeted exploitation

Threat Category 5: **Data Disclosure**

- If the database or app is not adequately secured, sensitive information such as past purchases, reward points, or personal information may be made public.
- Private information may also be leaked if data is shared with unapproved third parties, such as marketing partners or loyalty programs.
- **Affected Component:** Customer Database, CRM, Third party marketing system
- **Impact:** Financial loss/ Reputational loss can occur

Threat Category 6: **Unawarness**

- Many customers do not know what data ShopSmart collects, how long it is stores, or who with whom it shares.
- Privacy settings may be confusing, hidden, or hard to adjust, which makes it difficult for customers to control how their data is used.
- **Affected Component:** POS/App
- I**mpact:** Failure to follow the law or legal requirements, Loss of Customer Trust

Threat Category 7: **Non-compliance**

- Keeping data too long and using it without consent
- Privacy laws may broke when failed to handle data subject request
- **Affected Component:** Customer Database, CRM, Reward Logs
- **Impact:** Fines and Reputational loss

Section 4: ARX Certificate Summary

The anonymization was performed in ARX for the project ShopSmart – Customer Reward Program.The dataset contained 500 records and 16 attributes with customer demographics, transactions, and feedback data.The attributes in ARX were categorized as follows:

- **Identifying:** *Name* (fully removed)
- **Quasi-identifying:** *Age, Gender, Location, Transaction Date, Total Cost*
- **Sensitive:** *Points Earned, Points Redeemed, Feedback on Reward*
- **Insensitive:** Marketing and transactional fields (unchanged)

Transformations used:

- *Age, Location, Gender,* and *Transaction Date* -**Generalization**
- *Total Cost* - **Microaggregation (arithmetic mean)**
- Equal weights were used (0.5) to balance influence, and suppression was restricted to 4.5%.
- All attacker models (prosecutor, journalist, marketer) had a risk level of 0.2.

Output summary:

- Records: 500 | Attributes: 16 | Anonymity: Achieved
- Highest individual risk: **11%** | Information loss: **0.3953 (~60% utility)**
- Privacy models: **k-Anonymity (k=5)**, **l-Diversity (l=2)**
- Generalization levels: Age 2/3, Location 2/2, Gender 0/2, Date 4/4

Section 5: Explanation and Decision Justification

Input Specifications

The input data contained customer demographic, behavioral, and transactional attributes collected through the ShopSmart rewards program.Some fields were classified based on their ability for identifying individuals. Identifiers like Name were fully masked, while attributes like Age, Gender, and Location were treated as quasi-identifiers. This ensures that attributes that could indirectly reveal an identity were properly protected without the loss of analytical value of the dataset.

Attributes and Transformations

Each type of data was anonymized based on its sensitivity:

- Name was removed entirely to eliminate direct identification.
- Age, Gender, Location, and Transaction Date were generalized according to hierarchical levels.For example, dates were generalized from exact day to broader time frames, and age values were converted into ranges.
- **Total Cost**: Microaggregated to mask individual spending while keeping overall patterns intact.
- **Sensitive attributes** (Points Earned, Points Redeemed, Feedback): Protected using **l-diversity (l=2)**, ensuring that each anonymized group contains varied values to prevent inference of sensitive details.

Configurations

- **Equal weights (0.5)**for all quasi-identifiers to ensure balanced generalization.
- **Suppression limit of 4.5%**, minimizing data loss.
- **No monotonicity assumption**, allowing flexible generalization strategies.
- **Risk threshold of 0.2**, meaning no individual record has more than a 20% chance of being re-identified—an intentionally cautious privacy stance.

Output Properties

Output Data:
The anonymized dataset contained all 500 records and 16 attributes, preserving the dataset size and analytical consistency. The final checksum uniquely identifies this version of the anonymized output, ensuring reproducibility.

Solutions and Transformation:
The ARX tool explored 180 possible transformation solutions and materialized 19, finally selecting the transformation that provided the best balance between privacy and utility.

- *Age:* 2 of 3
- *Location:* 2 of 2
- *Gender:* 0 of 2
- *Transaction Date:* 4 of 4

The resulting transformation achieved a **maximum individual re-identification risk of 11%**, meeting the set privacy threshold and confirming **k-anonymity**.

Data Quality Models

The data quality was measured using a method that looks at how much information was lost when the data was changed to protect privacy. It considered two types of changes like generalization and suppression in equal range. The final score was 0.3953, which means the data still keeps about 60% of its usefulness. In other words, the dataset is still good enough for analyzing trends and building models, while keeping people's private information safe.

Privacy Models

Three privacy models were implemented in the data:

1. **k-Anonymity (k = 5):** Ensures that each anonymized record appears in at least five different forms, protecting against identity disclosure.
2. **l-Diversity (l = 2):** Ensures diversity within sensitive attributes to prevent attribute disclosure.
3. **Risk-based thresholding:** Maintains all attacker model risks below 0.2, guaranteeing a strong privacy guarantee.

Divyavaahini Thyagarajan | EPITA School of Engineering and Computer science | Paris

ShopSmart: Privacy-Centric Customer Rewards Program

Section 6: Proposed solutions for threats

1. Linking

Problem: The customer's actions can be linked together to form a complete profile.

Solution:

• Make use of tokenisation and pseudonymization in cryptography.

• Use PET to securely share data.

• Follow OWASP A02 – Cryptographic Failures to protect stored data.

---

2. Identifying

Problem: Patterns or unique data can be used to identify customers.

Solution:

• To hide specific actions, add Differential Privacy to reports.

• To safeguard identity, use anonymisation and data minimisation.

• To reduce the risk of re-identification, adhere to OWASP A04 – Insecure Design.

---

3. Non-repudiation

Problem: Customer identities are permanently preserved in the transaction log.

Solution:

• Employ cryptography by using hashed IDs with rotating keys.

• Set limits on data retention and have outdated records automatically deleted.

• Use OWASP A09 – Security Logging and Monitoring Failures as a guide when performing secure logging.

---

4. Detection

Problem: By network traffic attackers can easily detect user activity.

Solution:

• Use Forward Secrecy (Cryptography) with TLS 1.3.

• Add traffic padding or batching (PET) to mask real activity.

• To ensure secure communication, adhere to OWASP A02 – Cryptographic Failures.

---

5. Information Disclosure

Problem: Sensitive information about the customer might be leaked or shared.

Solution:

• Use AES-256 encryption at rest and TLS 1.3 in transit.

• Apply role-based access control and secure APIs.

• Follow OWASP Top 10 controls, such as A01 – Broken Access Control and A03 – Injection.

---

6. Unawareness

Problem: Customers are not aware of what information is collected and shared.

Solution:

• Implement visible notices and opt-in/opt-out mechanisms.

• Use privacy dashboards and consent management tools (PETs).

• Refer to OWASP A04 – Insecure Design, which supports users in controlling their privacy.

---

7. Non-compliance

Problem: Data retained for too long or used without permission is a violation of laws.

Solution:

• Apply data retention and deletion policies.

• Automate Data Subject Requests handling in line with GDPR.

• Follow OWASP A09 – Security Logging & Compliance Controls for audits.

Section 7.1: Data Inventory Spreadsheet

PERSONALLY IDENTIFIABLE INFORMATION		Purpose (Legal basis for processing e.g., consent, contract necessity, legal obligation)
DIRECT IDENTIFIERS	QUASSI IDENTIFIERS	
Name		**Contract Necessity**-To identify the customers account for Customer reward program
	Age	**Legitimate Interests-** To analyze customer demographics and Preferences to improve products
	Location	**Legitimate Interests**- To understand regional trends and customize offers
	Gender	**Consent**- To customize the customer experience and audience segmentation.
	Transaction Date	**Contract Necessity-** To store the purchase history, link transaction to reward points and identify buying patterns
	Total Cost	**Contract Necessity & Legal Obligation –** To calculate reward points, analyze spending behavior, and maintain financial records for accounting compliance.
	Points Earned	**Contract Necessity** – To record loyalty points gathered from customer purchases and manage reward balances.
	Points Redeemed	**Contract Necessity** – To record the redemption of points and maintain accurate reward tracking.
	Feedback on Reward	**Legitimate Interests** – To evaluate customer satisfaction and enhance future reward offerings.

Figure 7.1 Data Inventory Table

7.2 Data Processing Register

Processing details				Purpose of the data processing	Special categories of personal data?
Name of the processing operation	N° / REF	Date of creation of the record form	Last update of the record form		Yes/No
Customer Account & Loyalty Program	PR-001	11-Nov-25	11-Nov-25	To create and manage customer accounts and loyalty rewards.	No
Transaction Processing	PR-002	11-Nov-25	11-Nov-25	To process purchases and calculate reward points.	No
Marketing & Communication	PR-003	11-Nov-25	11-Nov-25	To share offers and measure customer engagement.	No
Customer Feedback & Support	PR-004	11-Nov-25	11-Nov-25	To improve services and respond to customer issues.	No
Analytics & Personalization	PR-005	11-Nov-25	11-Nov-25	To understand behavior and personalize rewards.	No
Fraud Detection & Security	PR-006	11-Nov-25	11-Nov-25	To prevent and detect fraud.	No
Backup & Disaster Recovery	PR-007	11-Nov-25	11-Nov-25	To ensure data recovery if systems fail.	No
Data Subject Rights Management	PR-008	11-Nov-25	11-Nov-25	To comply with GDPR and protect user rights.	No

Figure 7.2 Data Processing Register Table

7.3 Privacy Policy: Customer Rewards Program- ShopSmart (Last updated: 12 November 2025)

1. Who we are?: We operate the Customer Rewards Program, which allows customers to earn and redeem points for purchases and rewards. For the purposes of data protection law, we act as the data controller, meaning we determine how and why your personal data is processed.This policy explains how we collect, use, and protect your data. By using the Program, you agree to this policy.

2.Information we collect: We collect the following information from the customer. We do not collect any *sensitive information* (like Health or religious information)

- *Personal details:* Name, Age, Gender, and Location
- *Transaction data:* Items Purchased, Quantity, Total Cost, Transaction Date, and Payment Method
- *Loyalty information:* Points Earned, Points Redeemed, and Reward Chosen
- *Feedback and marketing activity:* Feedback on Reward, Marketing Email Opened, and Marketing Email Clicked

3.Why do we collect:

- We utilize your information to create and handle your loyalty account.
- To calculate and apply your reward points for the valid purchases.
- Send you promotions and special offers with your consent for marketing purpose.
- Enhance our products and services based on the customer feedback.
- Prevent fraud and preserve the security of our system.

4.Legal Bases for Processing: We use your personal data based on

- **Contract necessity(Article 6(1)(b))**– to manage your reward account and process purchases.
- **Consent(Article 6(1)(a))** – for sending promotional offers and updates about the rewards program.
- **Legitimate interests(Article 6(1)(f))** – to analyze spending patterns and improve service.
- **Legal obligation(Article 6(1)(c))** – to maintain financial and transactional records.

5. Data Sharing: We may disclose limited information with our trusted service providers - payment gateways, email systems, and IT hosting partners. Such service providers are under a contractual agreement to keep your information secure and process it only as per our instructions. We do not sell or rent personal data.

6. Data Retention and Security: Data are kept securely and only retained for a period necessary to meet program needs or legal requirements. We use encryption, secure servers, and limited access controls to protect your information.

7. Retention and Your Rights(Under GDPR): You have the following rights in relation to your personal data:

- **Right of access:** Obtain a copy of the data we keep about you
- **Right to rectification:** Correct inaccurate or incomplete information
- **Right to erasure ("right to be forgotten")**
- **Right to restriction of processing**
- **Right to data portability:** Receive your data in a structured, usually used format
- **Right to object:** To process your information based on legitimate interests or for direct marketing
- **Right to withdraw consent:** Where processing is based on consent.

To exercise your rights, email **privacy@customerrewards.example -** our team will respond as soon as possible.You also have the right to lodge a complaint with local Data Protection Authority if you believe that your rights have been violated.

8. Changes to This Policy: We may sometimes update this Privacy Policy to consider improvements or regulatory changes. The "Last Updated" date above shows when it was most recently revised.

9. Contact : If you have any questions about this privacy Policy or how we manage your data, please contact us at **privacy@customerrewards.example**.

7.3 Terms of Use: *Customer Reward program - Shopsmart*

Last Updated: 12 November 2025

Welcome to ***"Customer reward program"****("the Program", "we", "our", "us")* by Shopsmart. By joining or using the Program, you agree to these Terms of Use and our Privacy Policy. *Please read them carefully before joining the program.*

1. Eligibility & Registration: Membership is free and open to any person aged 18 years or older. You must provide correct information and make sure that your username and password are kept safe. You are solely responsible for any activity occurring under your account.

2. Programming Usage: You earn points for every purchase that is valid and eligible, which can be redeemed for rewards. Points do not have any value in cash, are not transferable, and may lapse, as described on the Platform. Any misuse, fraud, or violation of these Terms may suspend or cancel your account at the discretion of Meridio and will result in the loss of your reward points.

3. Promotions & Rewards: Please note that special offers and rewards are subject to additional terms and may be changed or withdrawn at any time when operational or legal reasons necessitate such action.

4. Data & Privacy: We process your personal data to handle your account, calculate rewards, and send offers. We do this only with your consent. For a detailed explanation of how we collect, use, and protect your personal data, please read our Privacy Policy.

5. Intellectual Property: All the materials, branding and software published on the Platform are either owned by, or licensed to, us and are protected by intellectual property laws. You may use them only for lawful and personal purposes.

6. Limitation of Liability: We are not responsible for any indirect, incidental, or consequential losses which arise from your participation in the program. Nothing in these Terms limits our liability for death, injury, or fraud caused by our negligence.

7. Termination: You can stop using our Services at any time. We may temporarily suspend or terminate the accounts of users who violate these Terms, who commit fraudulent activities, or whose accounts are inactive for an extended period.

8. Governing Law: These Terms are governed by the laws of your EU member state of residence. Disputes may be brought before its competent courts.

9. Updates & Contact: We update these Terms at a regualar interval; the latest version is always available on the Platform. For questions contact us at **support@customerrewards.example**.

7.4 Privacy-Focused Strategy for Obtaining and Managing Consent

This section describes how the ShopSmart Customer Rewards Program complies with the General Data Protection Regulation (GDPR) by obtaining and managing user consent through a privacy-focused approach. The integrated wireframe diagram which follows shows how consent is collected, handled, and tracked throughout the whole ShopSmart ecosystem by combining the system data flow and the user interface (App flow). In line with the GDPR principle of Privacy by Design, the design guarantees responsibility, transparency, and user control.

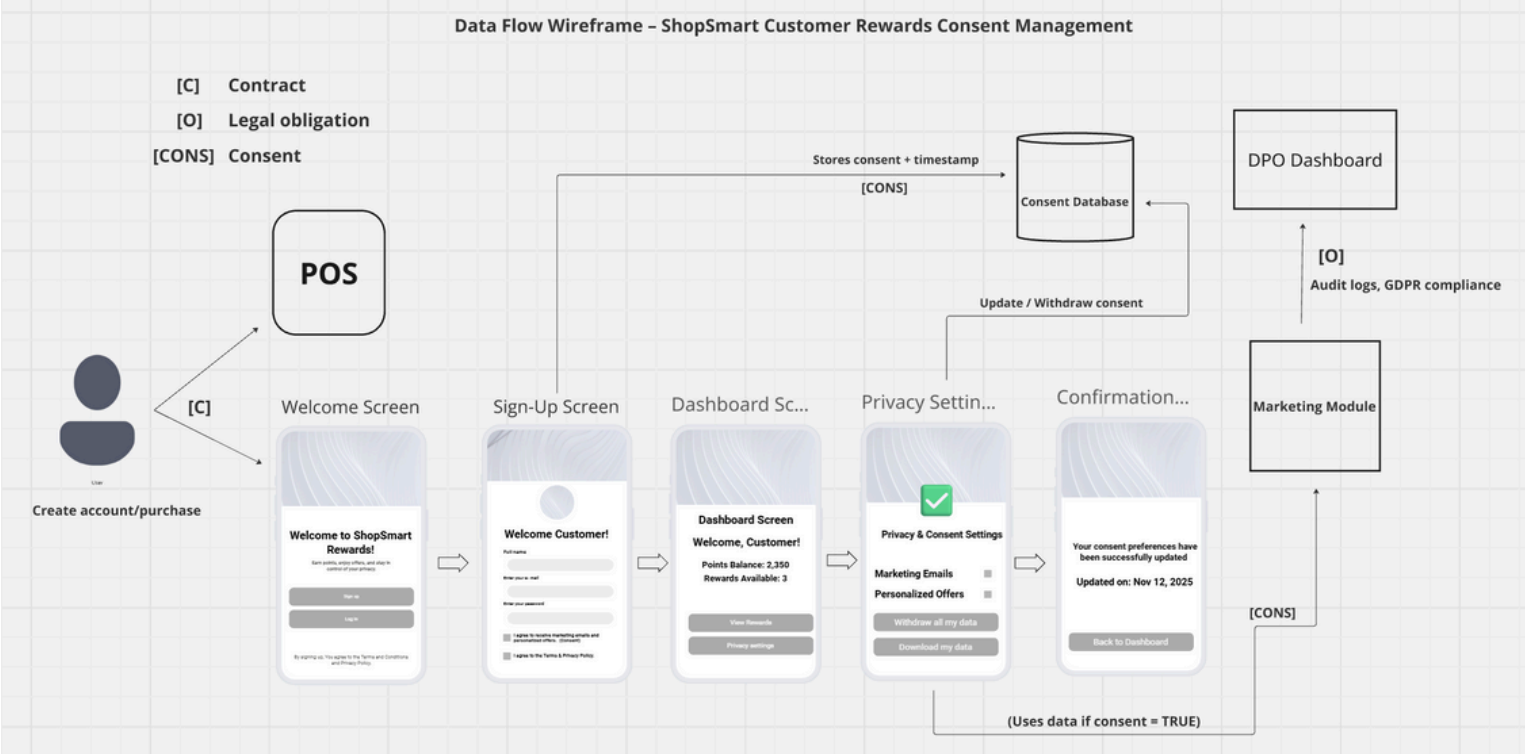

Figure 7.3 Integrated Wireframe Diagram – ShopSmart Customer Rewards Consent Management

The diagram shows how customers interact through the ShopSmart app and POS system to provide and manage consent, while backend components such as the Consent Database, Marketing Module, and DPO Dashboard ensure lawful and transparent data processing.

Important GDPR Principles Used

- Lawfulness and Transparency: Clear information on data use is provided, and consent is specifically required during registration.
- User Control: Through the app, users may simply change or revoke their consent.
- Accountability: The DPO Dashboard keeps track of and timestamps every consent action.
- Data Minimization: Rewards and offers only process the data that is required.
- Privacy by Design: The system design and user experience are directly integrated with consent and privacy management.

The integrated wireframe illustrates how ShopSmart integrates data processing and user consent acquisition into a single, compliant solution. ShopSmart maintains continuous GDPR compliance with building user confidence through control and transparency by connecting the app UI to backend consent management and auditing features.