

Глоссарий — Financery

Ниже приведены ключевые понятия и определения, используемые в проекте *Financery*. Этот глоссарий можно использовать в SRS, README и другой проектной документации.

Транзакция (Transaction) - запись, отражающая движение денег: расход или доход. Содержит сумму, дату, категорию и прочую метаинформацию (вендор, теги, источник).

Чек (Receipt) - изображение покупательского чека, преобразованное с помощью OCR в структурированный набор данных (позиции, итог, дата, продавец).

Позиция (Item) - строка из чека: конкретный товар или услуга с количеством и ценой.

Категория (Category) - семантическая группа, присваиваемая транзакции (например 'Продукты', 'Транспорт'). Может быть глобальной (предустановленной) или пользовательской.

Индекс цен (PriceIndex) - агрегированное статистическое представление цен по наименованию товара и региону (средняя, медиана, количество наблюдений).

Цель (Goal) - целевое состояние, заданное пользователем (например, снижение трат по категории на X% за период); система отслеживает прогресс и генерирует рекомендации.

Автокатегоризация (Auto-categorization) - процесс назначения категории транзакции автоматически на основании текста позиции, контекста, истории пользователя и ML-модели.

OCR (Optical Character Recognition) - технология распознавания текста на изображениях, используемая для получения данных из чеков.

ML-модель (MLModel) - модель машинного обучения (классификатор, нормализатор наименований, recommender), используемая для предсказания категорий, нормализации наименований и выдачи рекомендаций/прогнозов.

Feedback-loop (Обратная связь модели) - механизм, через который пользователь корректирует автокатегоризацию, и эти исправления используются для дообучения модели и улучшения качества.

Агрегация данных (Data Aggregation) - процесс объединения анонимизированных данных от множества пользователей для получения статистики (например, средние цены по товарам). Производится только с согласия пользователя.

Source (источник транзакции) - поле, указывающее откуда пришла транзакция: ручной ввод (manual), распознавание чека (ocr), импорт (import).

Anonymized Analytics (анонимная аналитика) - аналитические данные, собранные без возможности идентификации отдельного пользователя.

Vendor (продавец) - название торговой точки или компании, указанное в чеке или введённое пользователем.

Receipt Parsing (парсинг чека) - процесс извлечения структурированных данных (items, total, date) из raw_text, полученного после OCR.

Normalization (нормализация наименований) - приведение наименований товаров к единому формату для корректной агрегации и сравнения цен (например, 'milk', 'молоко', 'Молоко' -> 'молоко').