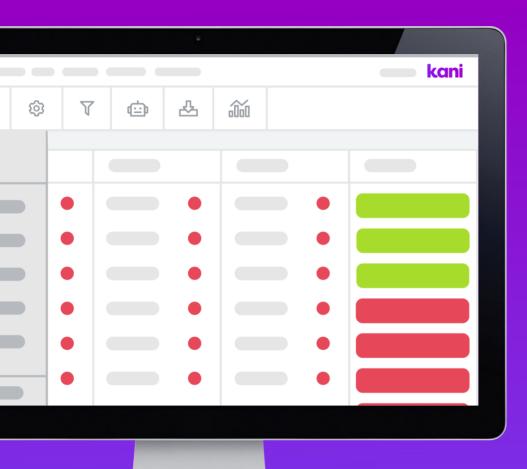
Data challenges in the payments industry

Sophie Harbisher

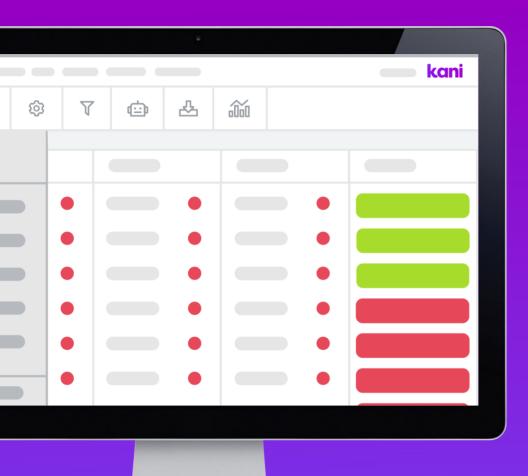




Talk Outline

- Who are Kani and what do we do?
- Data challenges in the fintech industry
- Reconciling data

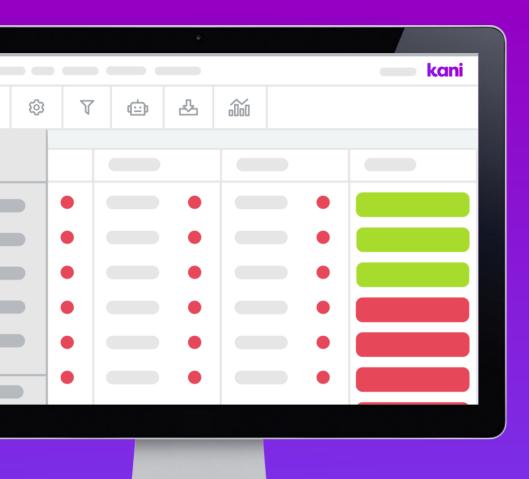




A short guide to Kani

- Kani is a Fintech based in Newcastle
- Offer SasS to fintech companies
- The platform ingests data from multiple sources, handles reconciliations and provides BI reporting

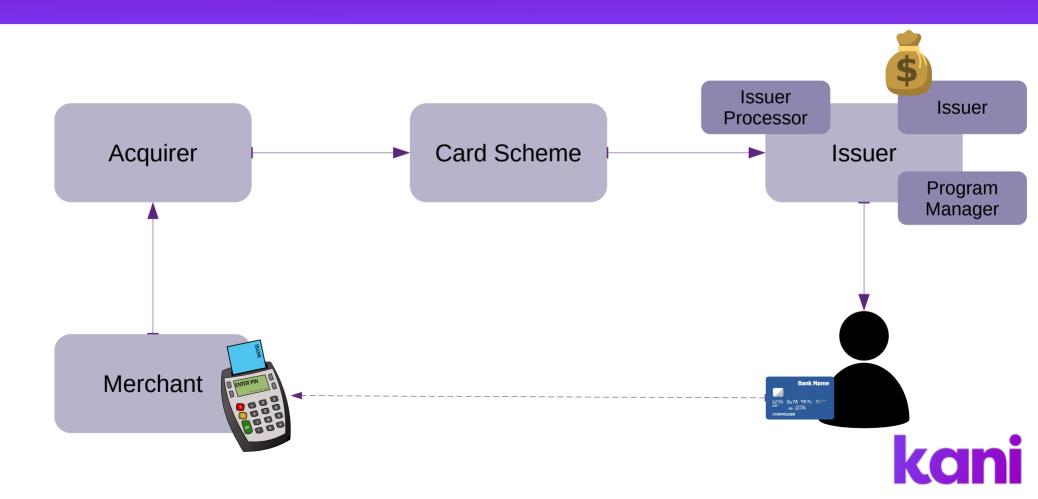




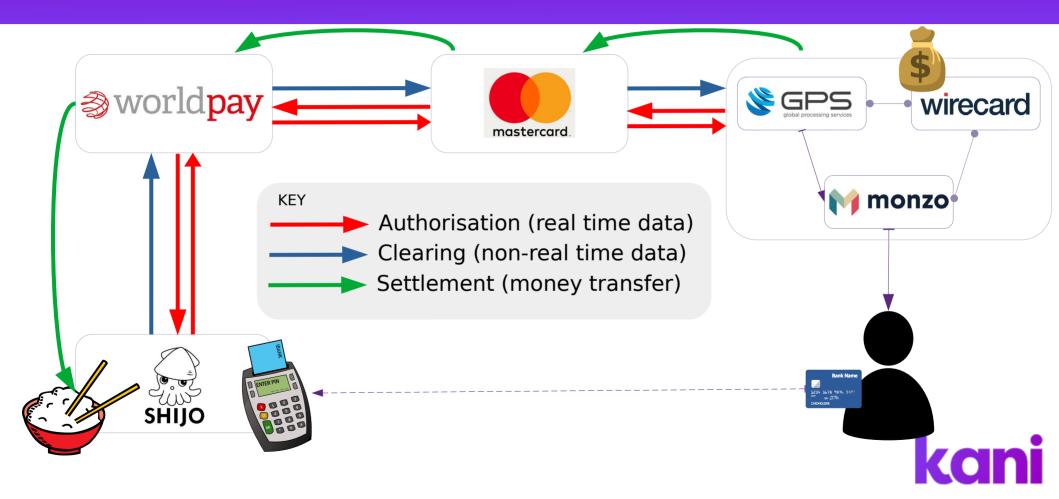
Payments Data



Transaction Lifecycle



Transaction Lifecycle



Transaction Lifecycle

- Messages are passed between acquirers, card schemes and issuers(/issuer processors)
- Settlements are aggregated (daily/over a clearing cycle) so that money transfers are made periodically rather than for every transaction
- Data passed around the transaction lifecycle isn't necessarily consistent in format or interpretation
- Lack of understanding about various data elements



Transaction Data





IT

build project needed; unfamiliar with what finance & compliance need to achieve



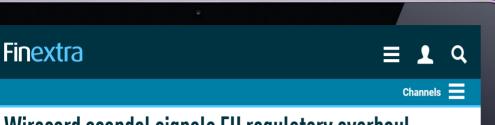
Finance

regular requests to IT due to lack of control over data, often fallback to MS Excel



Compliance

regular requests to IT due to lack of control over data, often fallback to MS Excel



Wirecard scandal signals EU regulatory overhaul

27 July 2020



European Union officials have suggested that there will be changes in the supervision of fintech firms in the wake of the £1.8bn accounting scandal at German payments firm Wirecard.

Earlier this month, the European Commission's (EC's) executive vice president Valdis Dombrovskis announced that the European Securities and Markets Authority (ESMA), the EU's financial services regulator would conduct a review into the regulation of the sector.

Of particular concern is a potential grey area in regulation for those firms, such as Wirecard, that start life as non-bank startups but then acquire banking licences and services.

Wirecard began as a company processing electronic payments before acquiring a bank and adopting

Issues

- Finance companies need accurate and verifiable information
- Reporting and reconciliation is essential for payments companies
- This is difficult to achieve, time consuming and often inaccurate









Wirecard scandal signals EU regulatory overhaul

27 July 2020



European Union officials have suggested that there will be changes in the supervision of fintech firms in the wake of the £1.8bn accounting scandal at German payments firm Wirecard.

Earlier this month, the European Commission's (EC's) executive vice president Valdis Dombrovskis announced that the European Securities and Markets Authority (ESMA), the EU's financial services regulator would conduct a review into the regulation of the sector.

Of particular concern is a potential grey area in regulation for those firms, such as Wirecard, that start life as non-bank startups but then acquire banking licences and services.

Wirecard began as a company processing electronic payments before acquiring a bank and adopting

Record Matching





Record Matching

- If a unique reference exists then reconciling across different data sources is easy
- However this is not always the case and we have to try to match based on similarity
- E.g. bank statement vs store receipt



Record Matching

The record matching problem is often broken down into three main stages:

- 1. Similarity scores for individual record elements
- 2. Overall similarity score between records
- 3. Suggested matches



Data Element Similarity

- There are 3 different data types we may want to assess similarity over:
 - Numeric quantities
 - Dates
 - Text
- The aim is to get a value which indicates which records are similar to each other and which are not



Text scoring

- Classical comparisons between strings often belong to the family of edit distances (e.g. Levenshtein, Jaro-Winkler,...)
- E.g. the Levenshtein distance between canny and kani is 3

$$canny \xrightarrow{del} cann \xrightarrow{sub} cani \xrightarrow{sub} kani$$

 More recent developments in string comparisons leverage natural language processing



Record Similarity & Matching

- Consideration of how to calculate an overall score given similarity scores of individual components
 - Logistic Regression
 - Naive Bayes Classification
 - Other...
- Suggesting which records are matches is difficult



Record matching challenges

- Scalability & curse of dimensionality
- Ideally would want a probability of a match between records – lack of training data for supervised models, still issue of how to recommend matches
- Matching over different aggregations of numeric quantities
- Record matching/entity resolution is an active research area with some interesting approaches



Application



HOW TO PLAY

×

Guess the WORDLE in six tries.

Each guess must be a valid five-letter word. Hit the enter button to submit

After each guess, the color of the tiles will change to show how close your guess was to the word.

Examples



The letter W is in the word and in the correct spot.



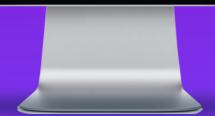
The letter I is in the word but in the wrong spot.



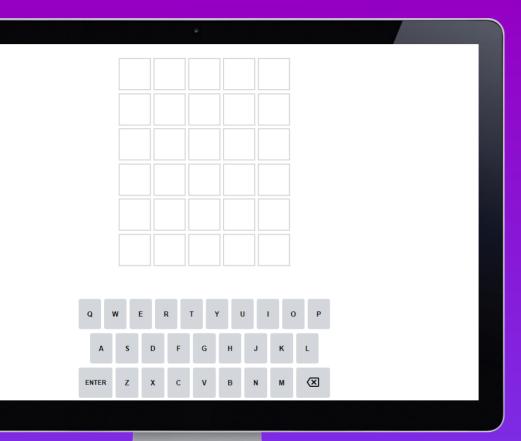
The letter U is not in the word in any spot.

A new WORDLE will be available each day!

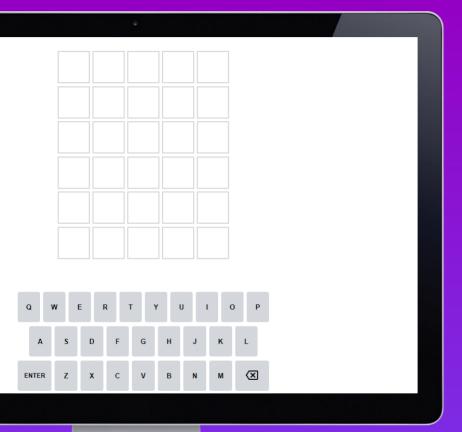
- Aim is to guess what the five letter word is within 6 tries
- Every guess gives you some information
- Question: What is the best initial guess?





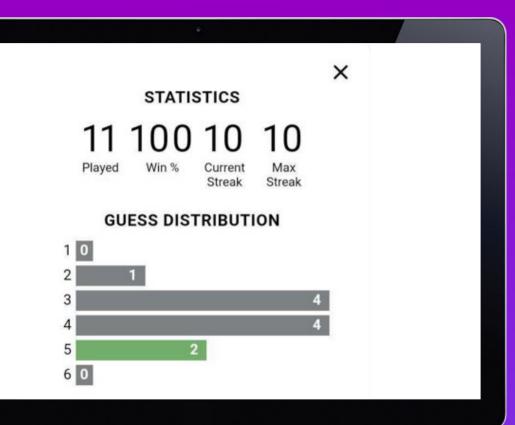


- Text similarity settings:
 - Correct letter, correct position gets similarity of 1
 - Correct letter, wrong position gets score of 0.5
 - Incorrect letter gets score of 0
- Aim to find the word that is on average closest to any other word

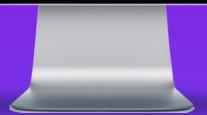


- Best initial guess is TARES
- Other good starting words:
 - tares 1.427
 - lares 1.421
 - rales 1.398
 - nares 1.397
 - cares 1.396

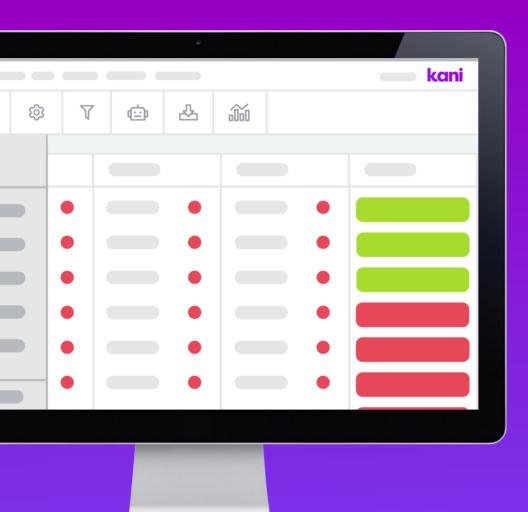




- Best initial guess is TARES
- Other good starting words:
 - tares 1.427
 - lares 1.421
 - rales 1.398
 - nares 1.397
 - cares 1.396



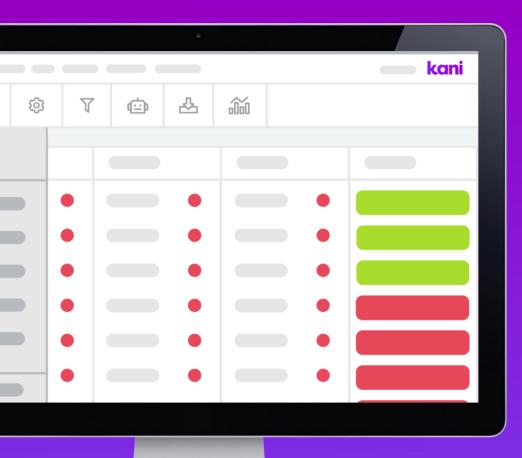




Summary

- Understand the transaction lifecycle
- Reviewed the difficulties fintech companies face
- Brief overview of record matching and how we can attempt to match data
- How to make a good start to Wordle





Thanks for listening

sophie@kanipayments.com hello@kanipayments.com

