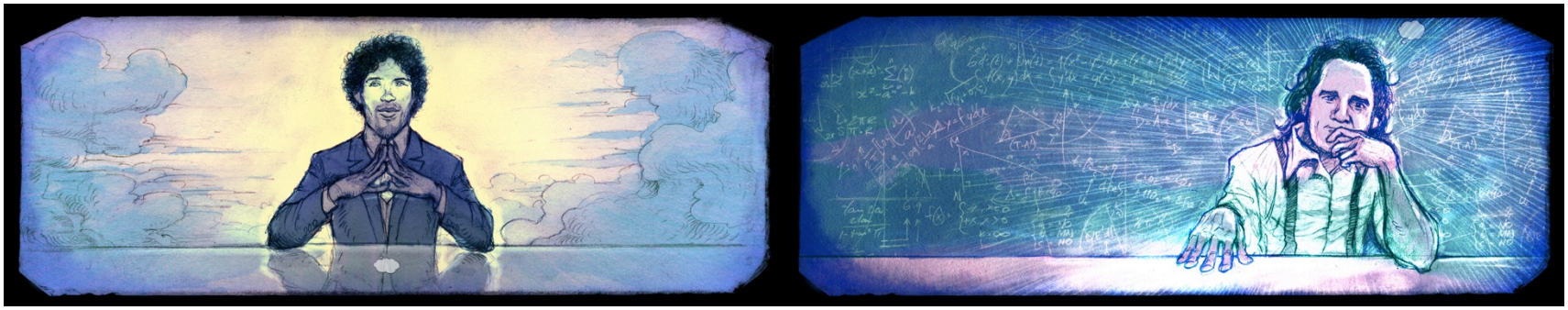
Presentatie Fontys

Introductie

- Over orikami
- Examples
- Challenge(s)

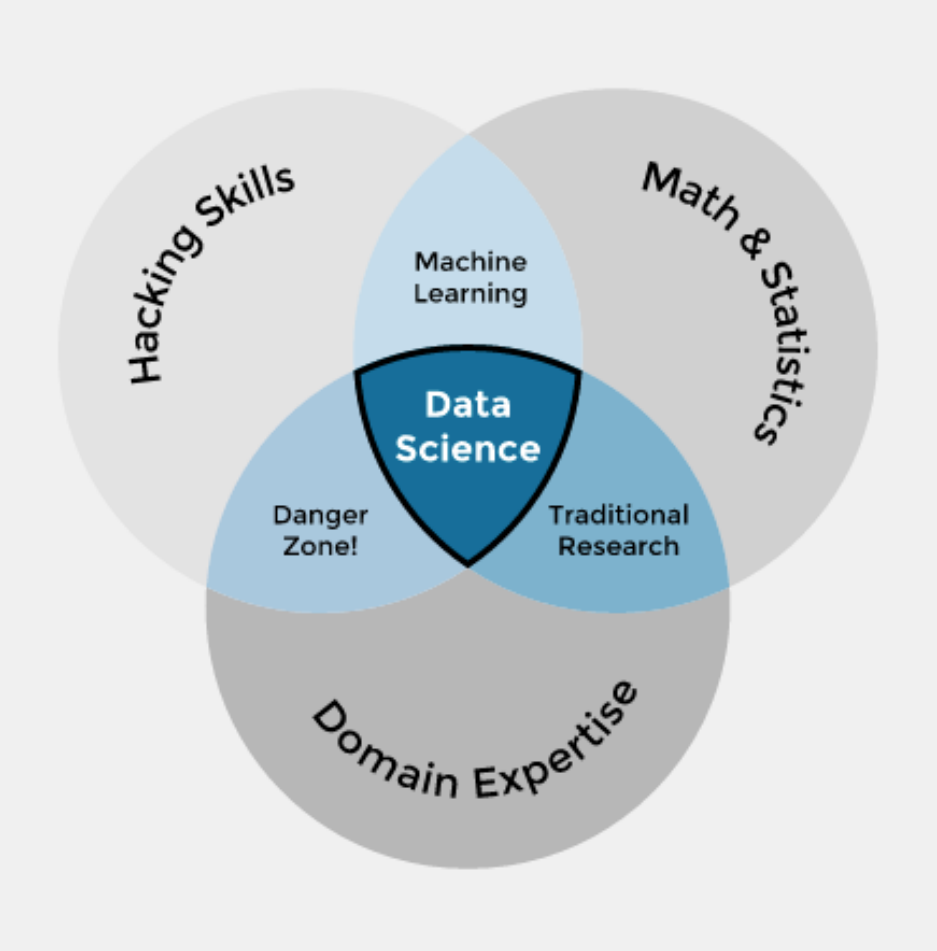
orikami

data science boutique



Wij geloven dat er meer uit data te halen is!

Vakmanschap



Team

Vera Rutten
interaction designer

Niklas Weber
Bachelor of Science
junior data hacker

Bram den Teuling
Master of Science
data scientist

Syed Saiden Abbas
PhD Graduand Computer Science
data hacker

Hanneke van den Akker
Master of Science
testing and support

Andres Lamont
Master of Science
data scientist

Marijn Martens
PhD Graduand Neuroinformatics
data scientist

Stijn Albers van der Linden
junior developer

Joris Cuppen
Master of Science
educational data wizard

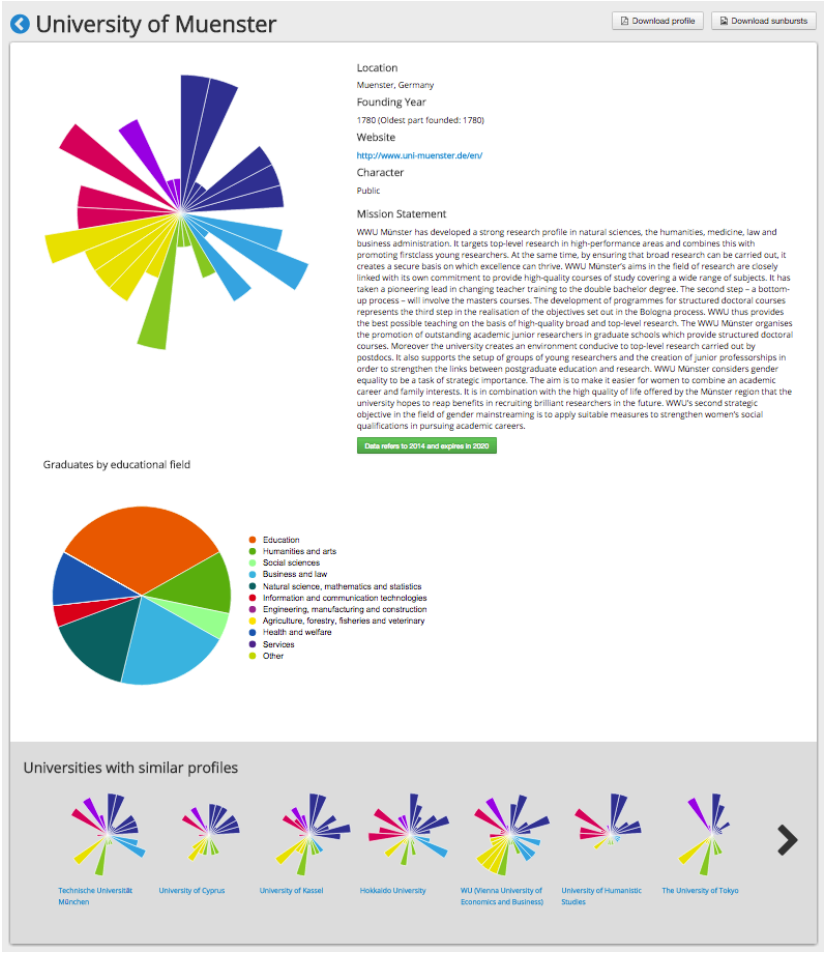
Big Data

- Cluster on commodity hardware (low cost)
- Real-time pipelines
- Availability of Super Computer
- All open-source! No vendor lock-in…

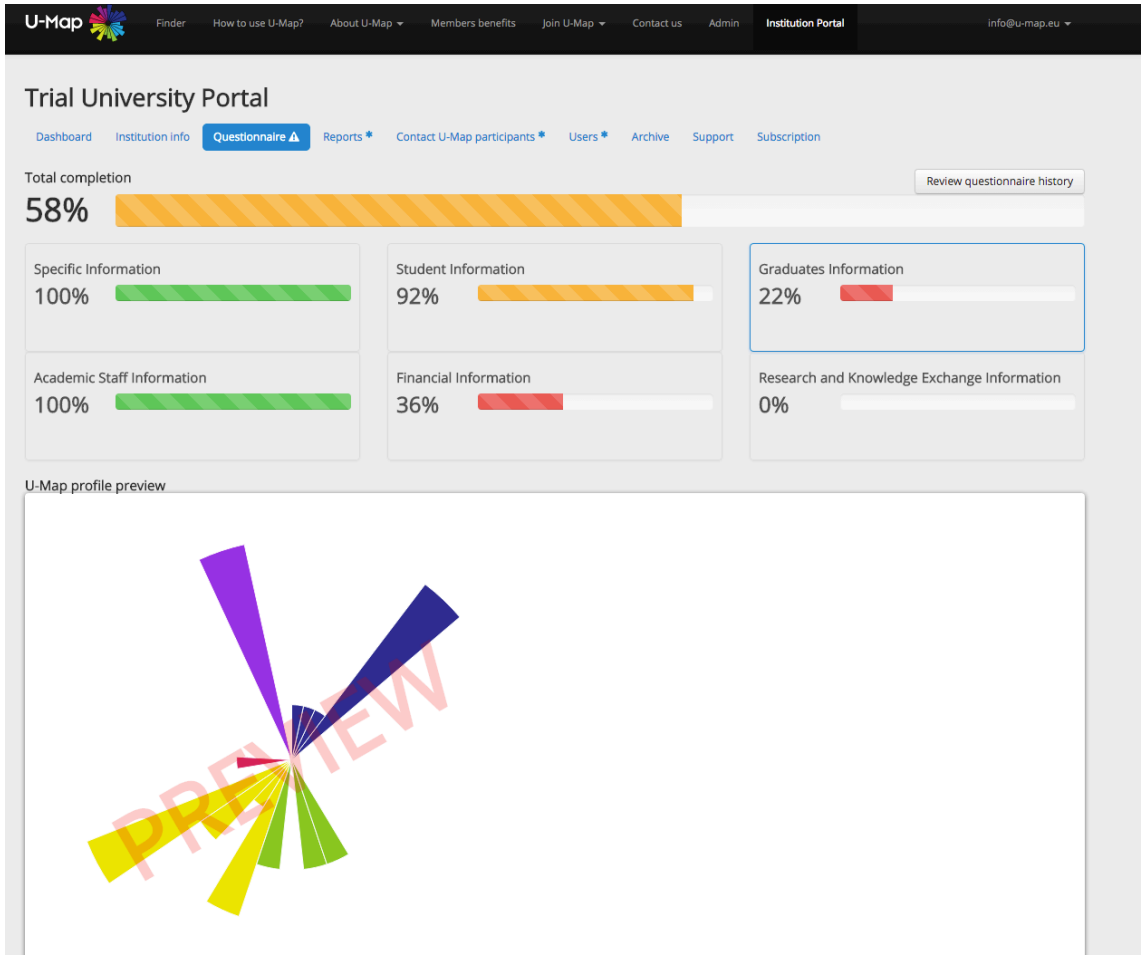
Overview

- Over orikami
- Examples
- The challenge

U-map institutional Profile



U-map questionnaire completion overview



U-Map

U-Map is an ongoing project in which the European classification of higher education institutions is further developed. Using a multidimensional ranking, you can compare all participating higher education institutions across Europe.

Data product

Together with the Center for Higher Educational Policy Studies, an institute linked to the University of Twente, we´ve developed a web application that maps universities on 29 dimensions.

The data application lets you search, filter and compare universities on all of these dimensions. Furthermore, the back-end is fully equiped for gathering standardized information about higher educational institutions in countries with different traditions and ways of presenting their data.

The U-Map application validates data input thoroughly.

Education profile			
Doctorate degrees awarded	0%	5.31%	0.51%
Master degrees awarded	7.04%	65.49%	17.19%
Bachelor degrees awarded	92.96%	29.2%	81.73%
Career oriented programmes	100%	63%	97%
Income from knowledge transfer	2.92%	7.07%	3.17%
Total enrolment	7035	565	11815
Exchange students incoming	3.24%	1.42%	1.7%
Research expenditure	21%	41.05%	3%
Graduates in the region	96%	13%	88%

Insight

The multidimensional classification is presented in an interactive and simple way.

Intelligence

There is a lot of information in the 29 dimensional data of all universities. This data encloses connections between universities never seen before. We have developed a simple but efficient algorithm to discover and compare universities based on similarity.

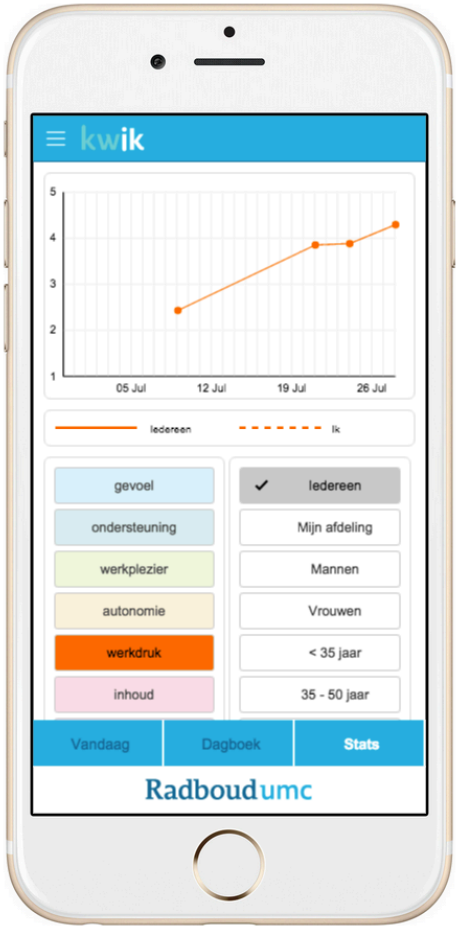
Techniques used

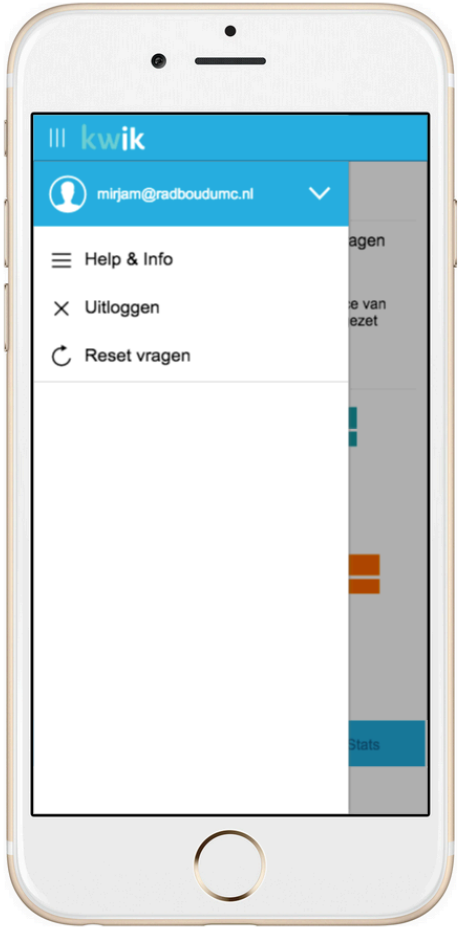
Ruby on Rails, Javascript, SVG, raphael.js, Nearest Neighbour algorithm, data visualisation, mongoDB, postgreSQL

Kwik: moderne teamklimaat meting

Kwik – de ‘smell’ van de dag

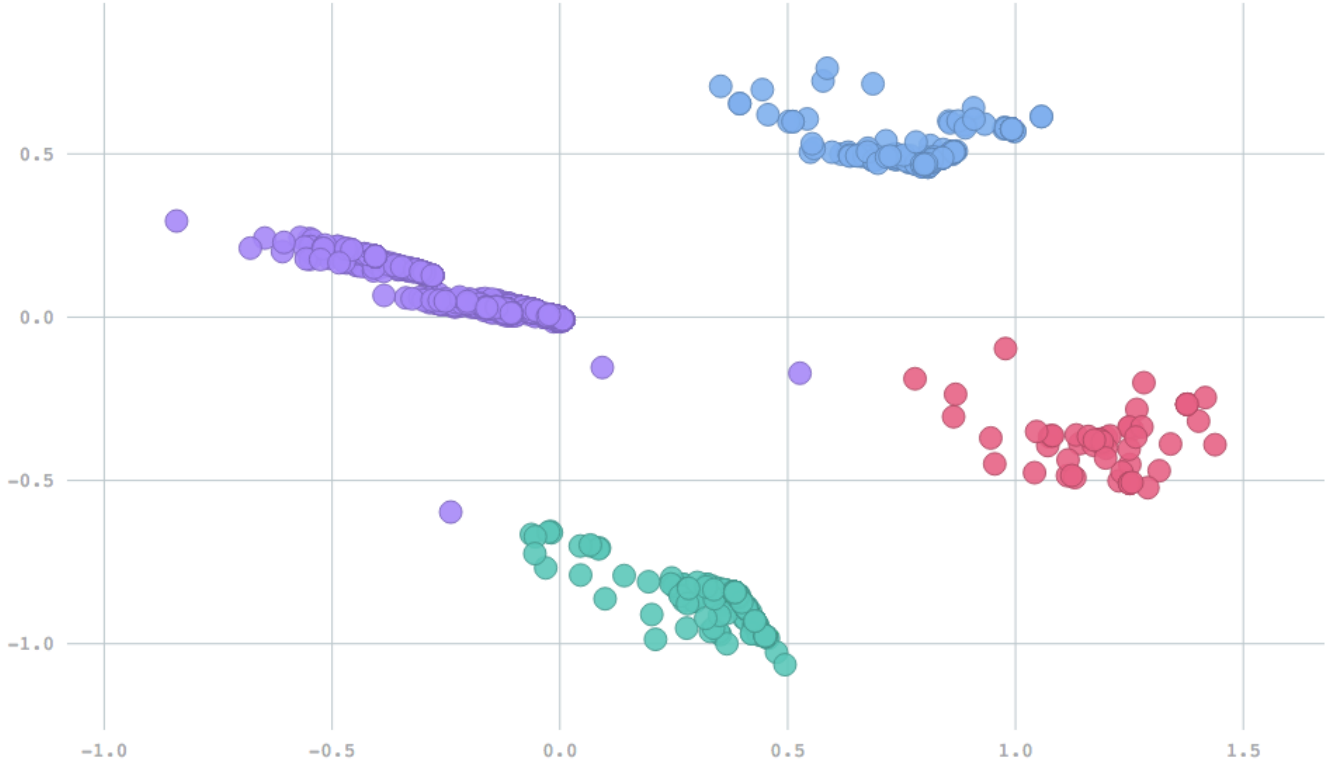
Kwik – dagboek, statistieken





Kwik: visual feedback with lights

Cluster analysis beslist.nl



Categories: level 1+2
PCA: 2 components
K-Means: 4 clusters

Sampled 5000

Mozart

Popularity of music should be measured in terms of what people really think about music. Mozart taps into society to pick-up signals about what's hot and trending.

Data product

We are developing Mozart in collaboration with market research company DVJ insights.

Mozart keeps an eye on open sources like Twitter, Facebook, YouTube and Last.FM to learn what people are saying about music.

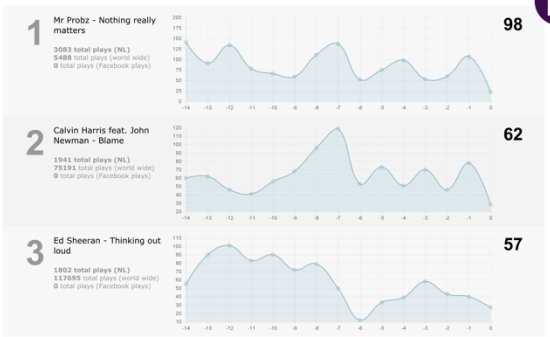
We combine this with specialized information from radio airplay, Spotify etcetera.

The final piece of information comes from a large user panel. We sample feedback from users about what they like.

Date	Artist	Title
ma. 28 jul. 2014 13:55	Big Mountain	Baby I Love Your Way
ma. 28 jul. 2014 13:55	múm	We Have a Map of the Piano
ma. 28 jul. 2014 13:55	The J. Geils Band	Land of a Thousand Dances
ma. 28 jul. 2014 13:55	Lasgo	Out of My Mind
ma. 28 jul. 2014 13:55	Within Reach	Enemy
ma. 28 jul. 2014 13:55	Brooklyn Express	Sixty-Nine
ma. 28 jul. 2014 13:55	Cheap Trick	California Man
ma. 28 jul. 2014 13:55	Quantic	Life in the Rain

Insight

Popularity of music should be measured in terms of what people think about music. Mozart keeps an ear on society to pick-up signals about what's hot and trending.



Intelligence

Mozart combines information about music from different sources for a particular market segment.

Our state of the art algorithms intelligently track music and make predictions about what is going to be popular.

Furthermore, our algorithms calculate scores from the historic data of music and efficiently keep on updating these scores.

Last but not least, Mozart presents trends and predictions about music to users with simple graphs.

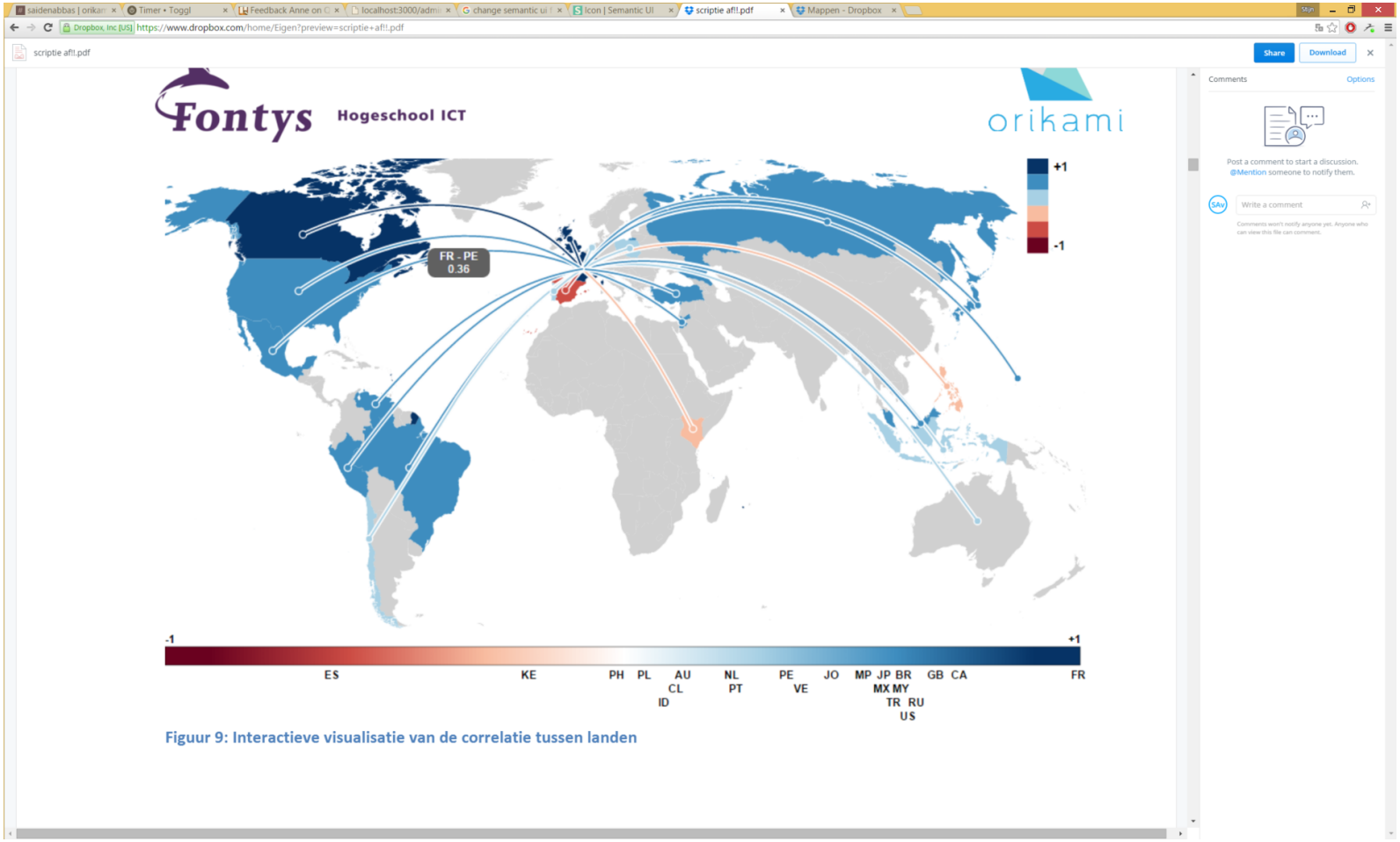
Techniques used

Twitter streaming API, Hadoop, Pig, Julia, Python, mongodb, Facebook API, Last.FM API, meteor.js

dvj insights BRANDING ADVERTISING MEDIA orikami

http://mozart-admin.herokuapp.com/top40

Mozart: country correlations



Figuur 9: Interactieve visualisatie van de correlatie tussen landen

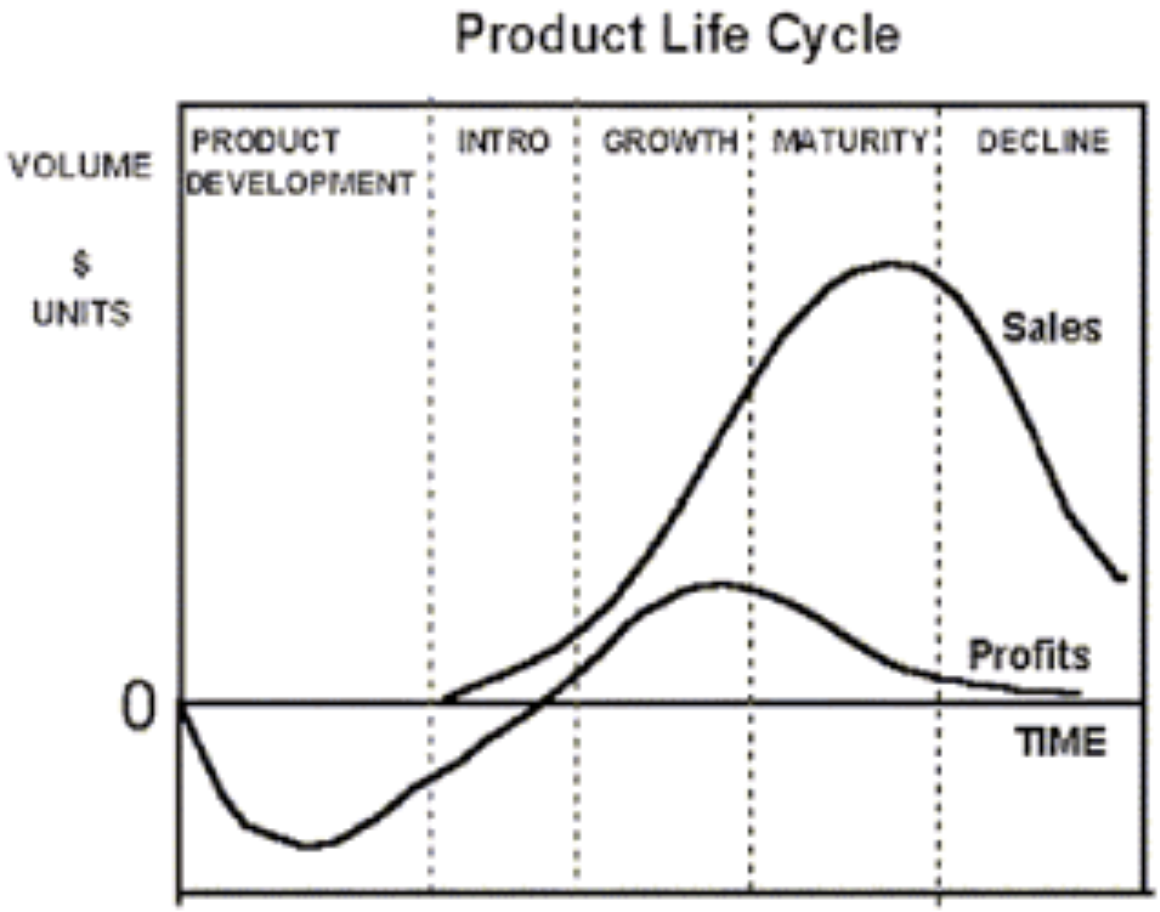
Overview

- Over orikami
- Examples
- The Challenge(s)

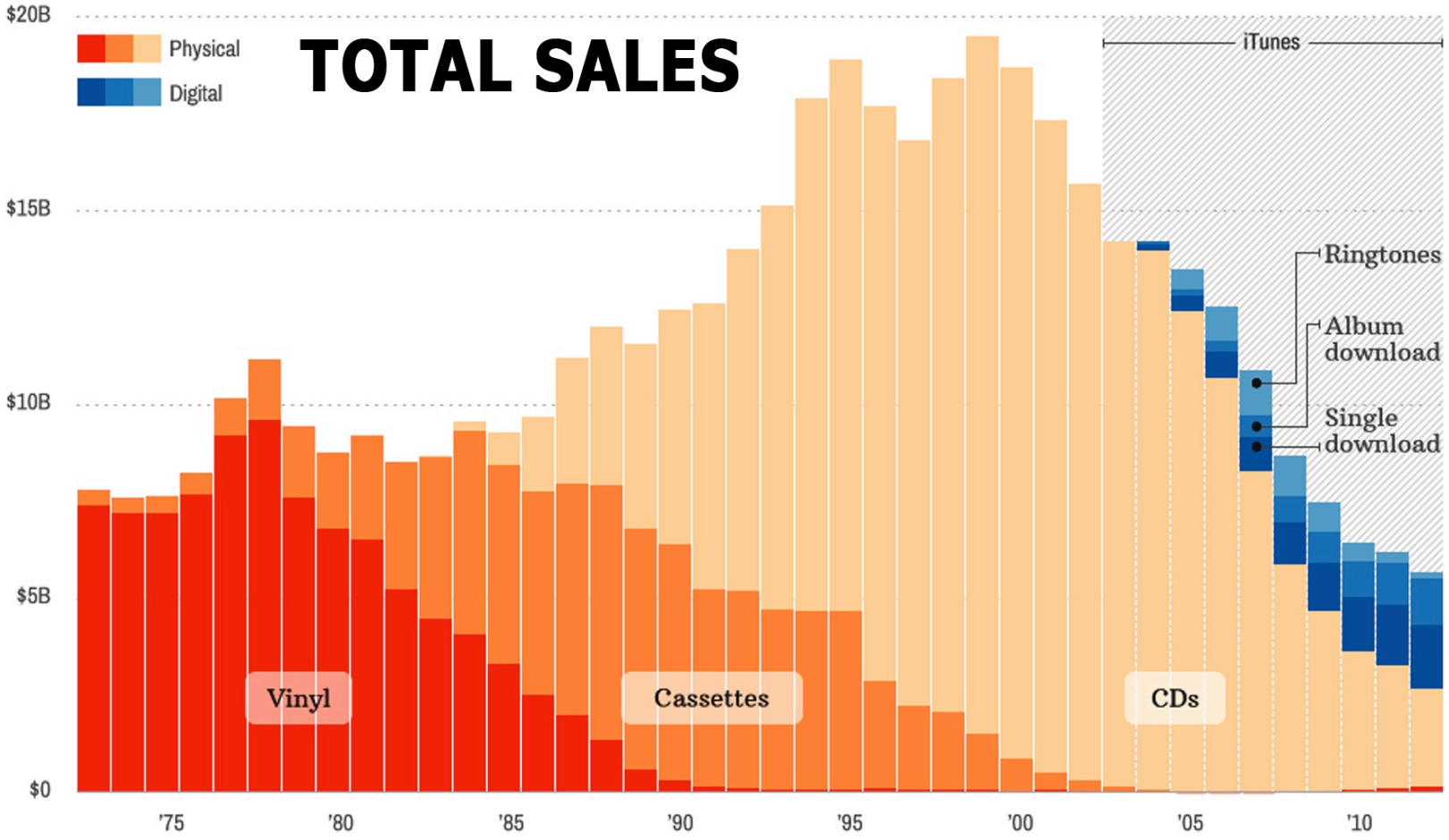
Challenge

Music life cycle

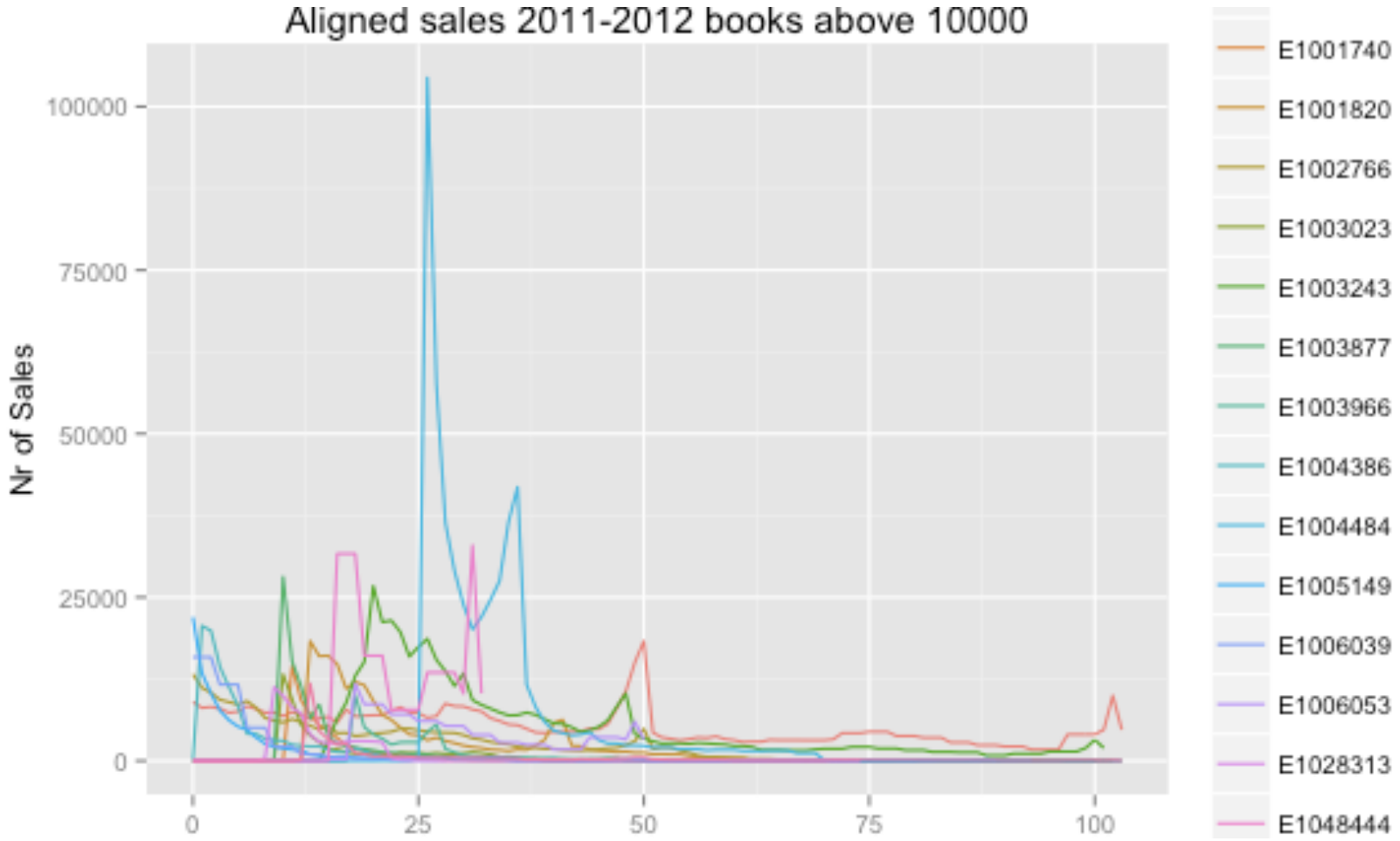
Life cycle


Product Life Cycle
VOLUME
$
UNITS
PRODUCT DEVELOPMENT
INTRO
GROWTH
MATURITY
DECLINE
Sales
Profits
0
TIME

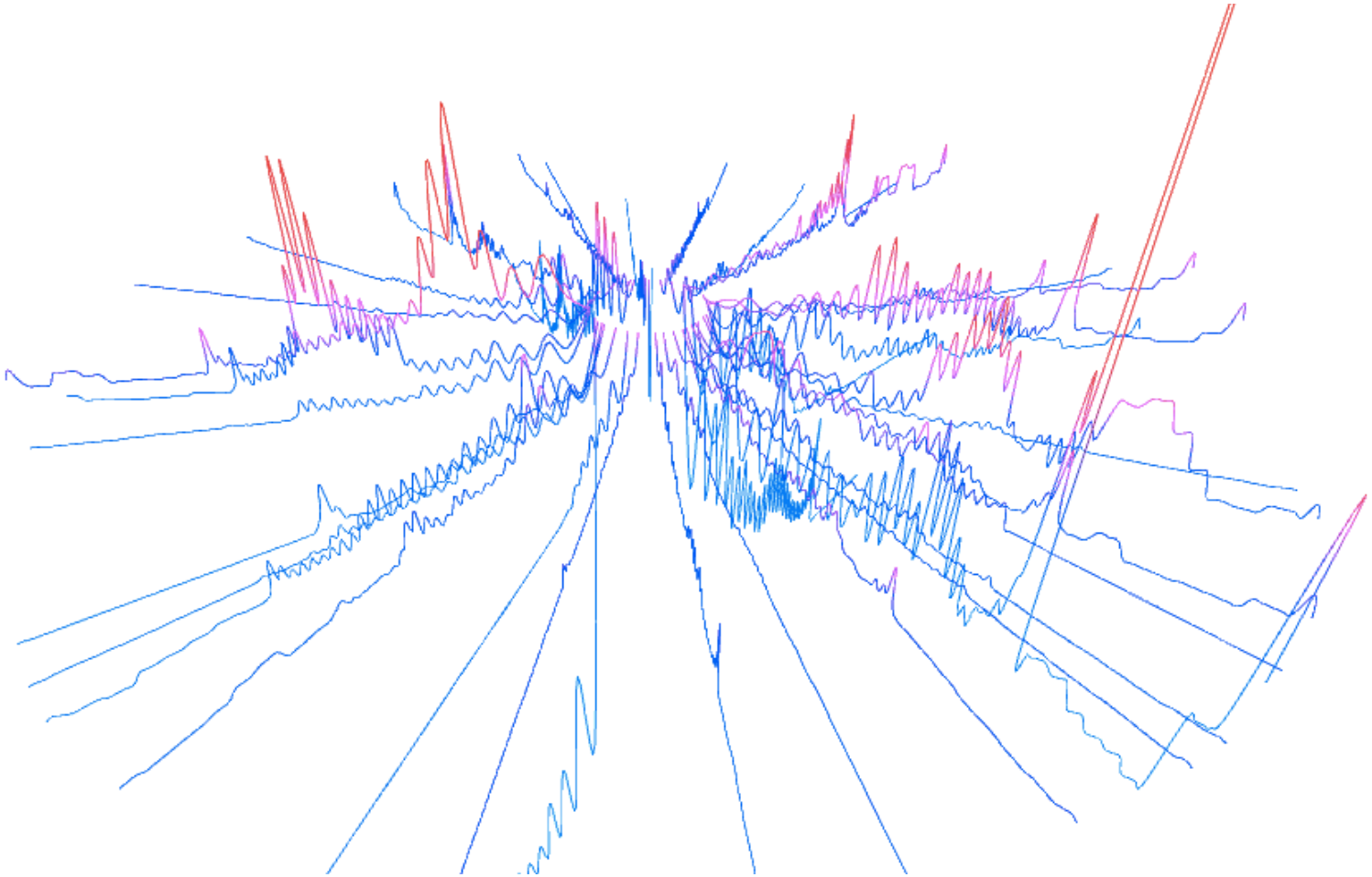
Product Life cycle



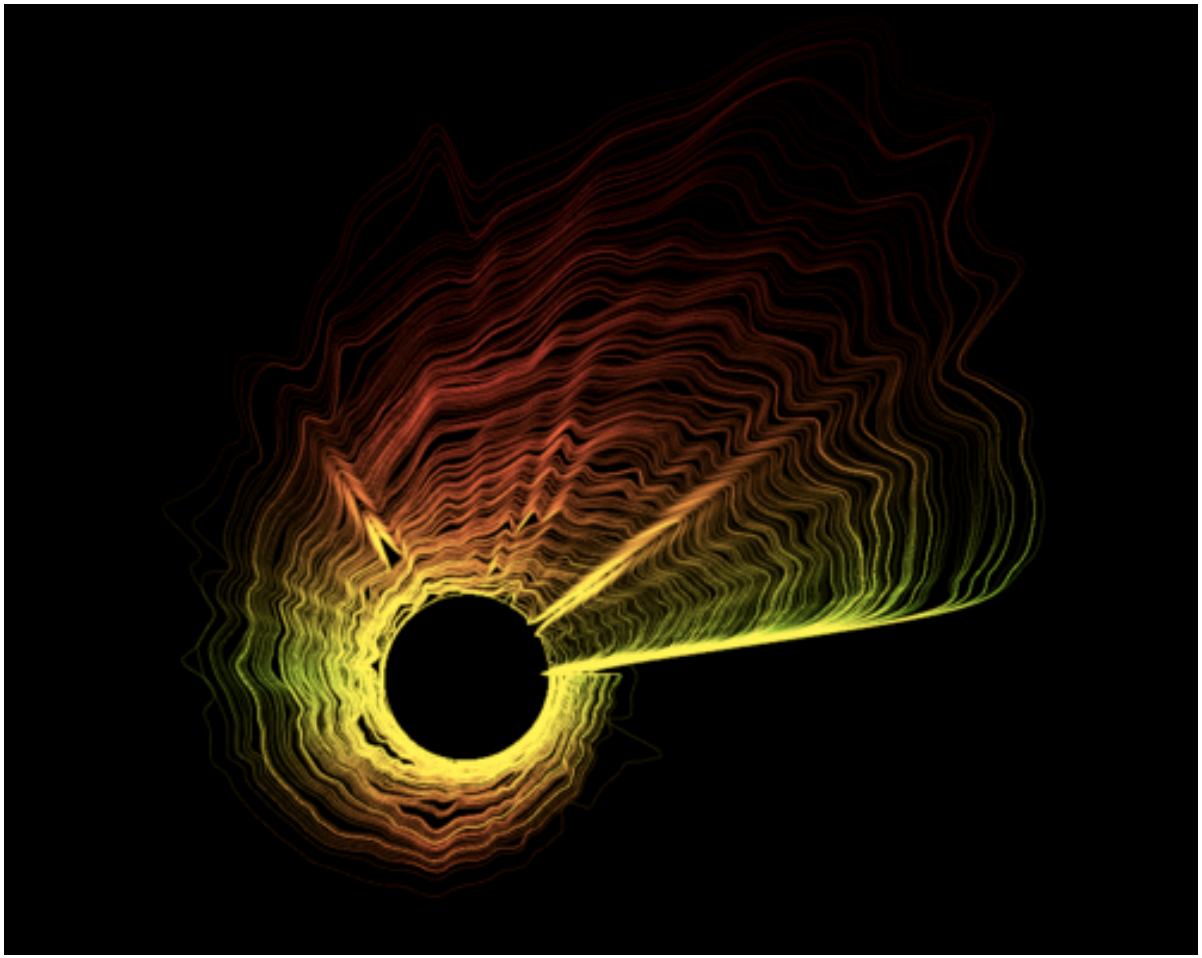
Life Cycle bestseller books



Life Cycle bestseller books



Life cycle bestseller books



Ingredienten

- 25 weeks panel data, 250 responses, 40 songs
- Twitter (#now playing) 20 weeks **65Miljoen**
  - <1% dutch!!
  - Aggregaties per dag, week en uur [WIP]
- Weekly charts
- MongoDB

Ideas

- Sentiment of tweets in life cycle?
- Getting tired of songs?
- Recognition saturation
- Template matching
- Clustering on different templates
  - E.g. popular artist vs non-popular artists

Ten slotte

- Meedere groepen?!
- Wees creatief
- Wees nieuwsgierig!

- Wij leveren support en toegangcodes tot databases!

- Contact: bram@orikami.nl / 0243010100

Vragen?

orikami

data science boutique