ML at Zopa

Peadar Coyle – Senior Data Scientist

Zopa's vision

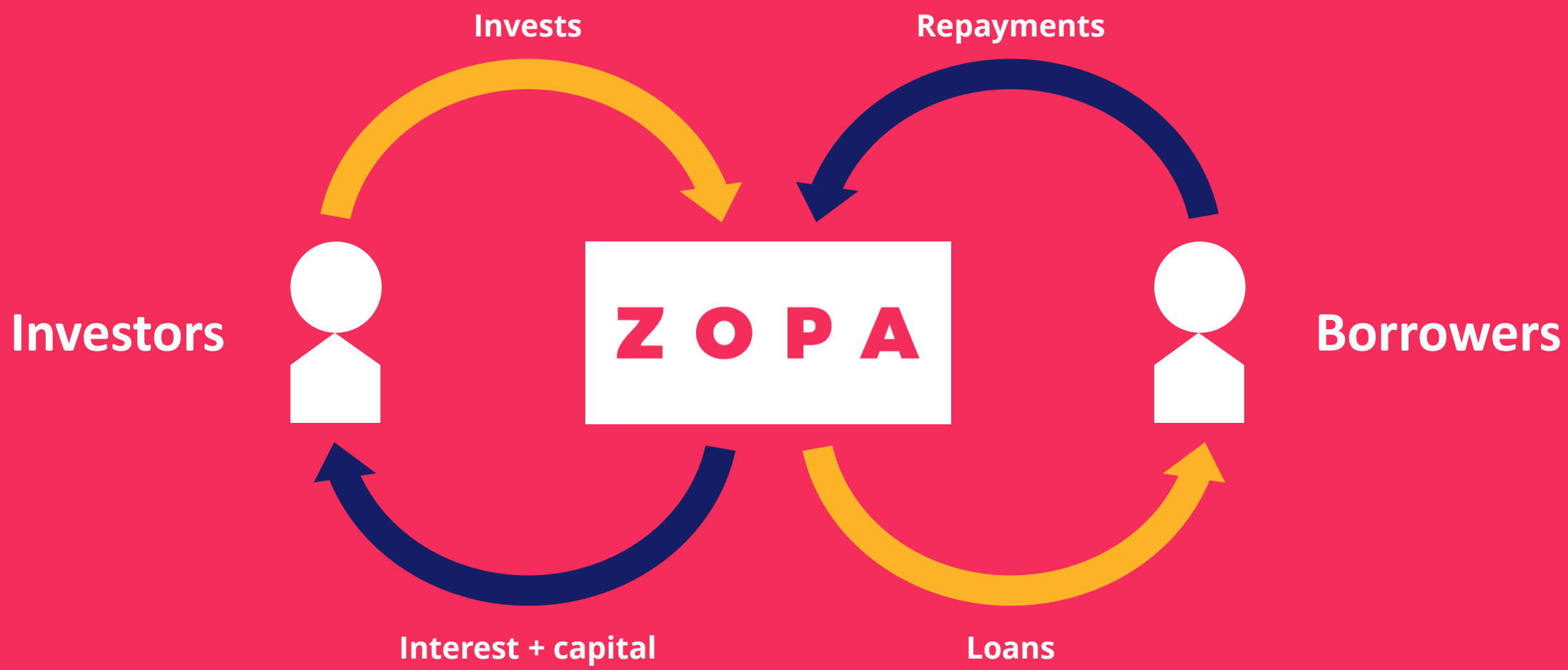
We will be the best place for money

ZOPA

What we do

Simple loans.
Smart investments.

ZOPA



A pioneering financial services company

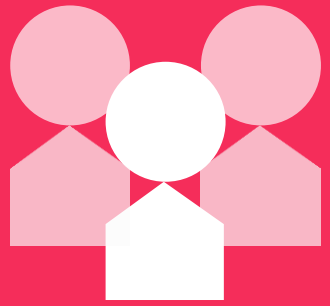
ZOPA

World's 1st

peer-to-peer lending platform in 2004

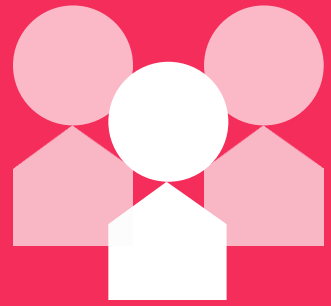
£3 billion

lent to date, and our growth is accelerating



300,000+

people have taken a Zopa loan



60,000 +

actively invest through Zopa

A disruptive, award-winning company

ZOPA

Personal loans

Money Super Market .com
SUPERS 2016
BEST PERSONAL LOANS PROVIDER

Investments

LendIt
Industry Awards
2017

Customer service

Moneywise
CUSTOMER SERVICE AWARDS
MOST TRUSTED
2010 to 2016

Technology

THE SUNDAY TIMES
TECH TRACK 100
2016
HISCOX

Company

altfi
WINNER
2016
ZOPA
EDITOR'S CHOICE AWARD

TRUSTPILOT
Excellent
Based on 2,333 reviews

AWARDED FOR QUALITY - RELIABILITY - DISTINCTION
UK Consumer Superbrands
2017

peer2peer
FINANCE ASSOCIATION

British Bank Awards
WINNER
2017

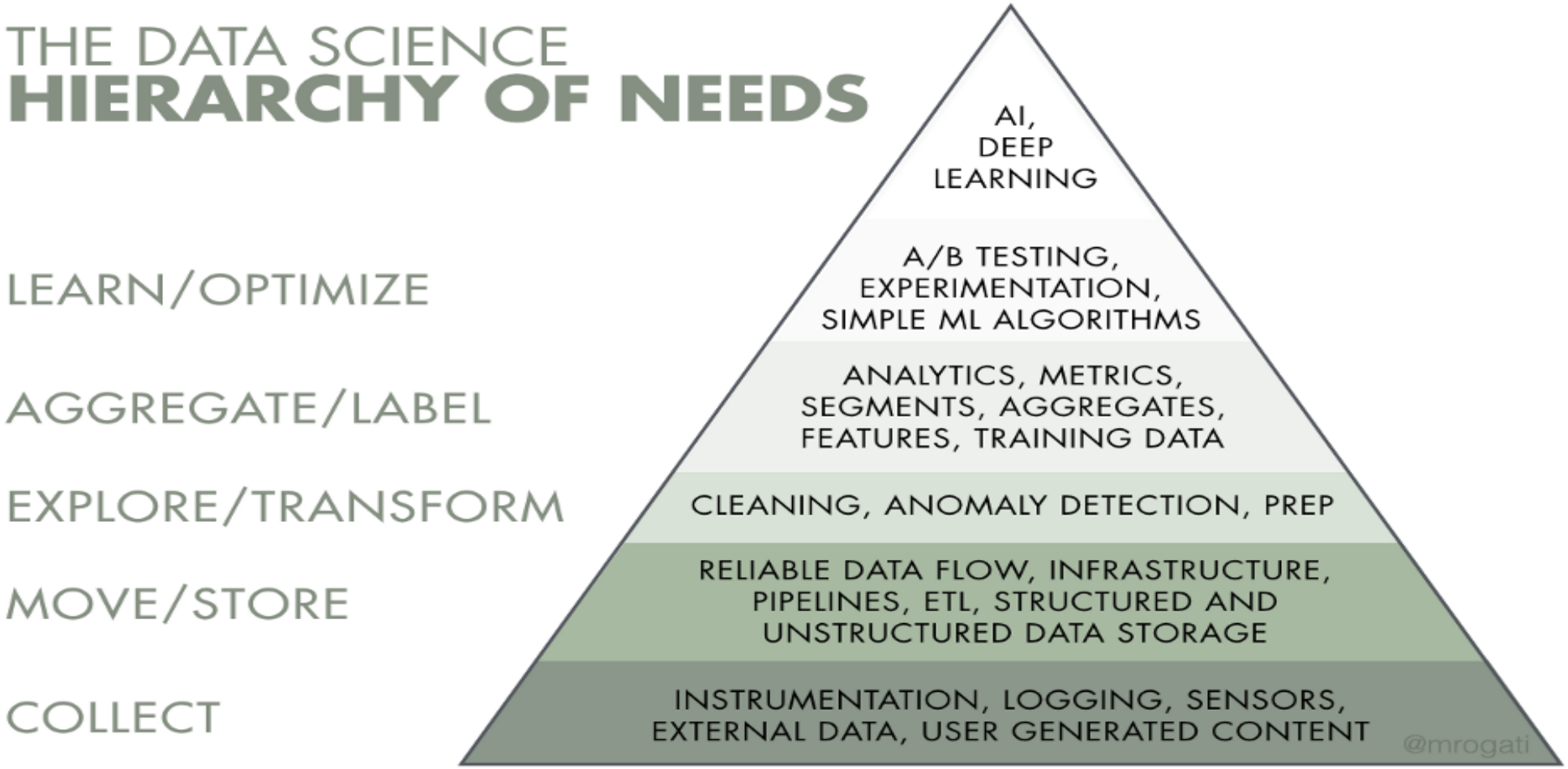
INNOVATE|FINANCE
MEMBER

ZOPA

At Zopa we've invested heavily in Data

- Credit Risk assessment
- Onboarding Funnel Simulation
- Pricing Optimization
- Fraud Identification
- Document forensics
- Optimized Prospect and Existing Customer Marketing
- Income and Rent estimation for Affordability Evaluation

AI/ML Hierarchy of needs



Started our journey in 2014

- Hired a Chief Data Scientist

- Created a Data Science function and invested heavily in Data Analytics

How do you make ML systematic?

ZOPA

Wanted to be able to produce models that were:

- Rapidly generated
- Easily vettable/ auditable
- Highly predictive
- Easily deployable

Several considerations

- **Common codebase** or personal choice of tools?
- Buy or **build**?
- Which language? Which package?

How do you make ML systematic? (2014)

Wanted to be able to produce models that were:

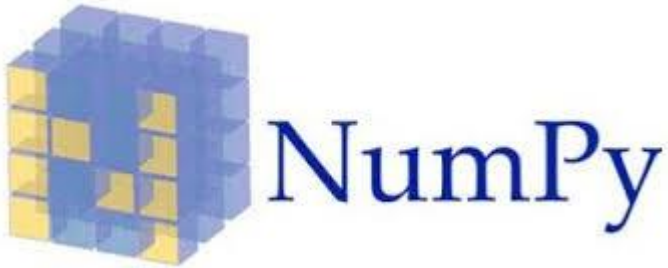
- Rapidly generated
- Easily vettable/ auditable
- Highly predictive
- Easily deployable

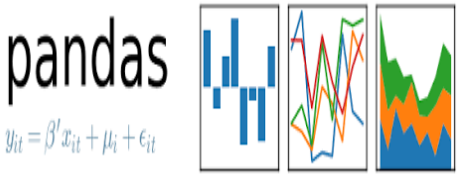
Several considerations

- **Common codebase** or personal choice of tools?
- Buy or **build**?
- **Python**

Predictor – Zopa's ML Toolkit (2014)

- Streamlined, Automated ML Application
- Implements all stages of producing an ML model, requiring minimal input
- Leverages PyData Ecosystem
- 9K lines





How is Predictor Used?

ZOPA

- **Same exact code used by DS and in Production**
  - No model translation overhead – no ‘code it all again in Java’
  - No restrictions on which ML techniques can be used

- **Training**
  - Needs unprocessed data and a simple config file
  - Driven via CLI, interacting with Python codebase, or Jupyter GUI

- **Querying Mode**
  - Driven by all of the above +
  - Rest API (Flask-based microservice) in Production, <1 -2 s/call

Strictly Private & Confidential

Learnings?

- We have full control over how we do ML
  - Methodology
  - Deployment
  - Features

- Competitive advantage, selling point to DSs

- One day we'll open source Predictor

- Need to maintain at least 2 people familiar with the code base

- Need to keep improving it – otherwise we'll miss developments

Strictly Private & Confidential

Not just ML

- We've been experimenting with Bayesian models, Survival Analysis, Optimization models, Natural Language Processing and Deep Learning.

- We're strong believers that we should develop an understanding of more-than-just supervised learning

- We do regular workshops internally both to the Data team and to the wider organisation

Thank you

Further Reading

- https://blog.zopa.com/2016/10/21/the-birth-of-predictor/
- https://blog.zopa.com/2016/12/02/data-democratization/
- https://blog.zopa.com/2017/12/13/open-source-zopa/

We're hiring Data Scientists, Risk Analysts, Product Analysts and Data Analysts

- https://jobs.lever.co/zopa/