

Evolving the Machine Learning Platform Organisation at Netflix:

A case study

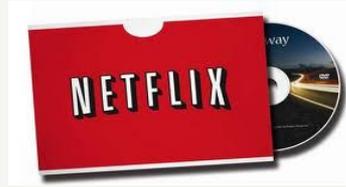
| Julie Amundson (@yakticus)
2021.05.11

NETFLIX

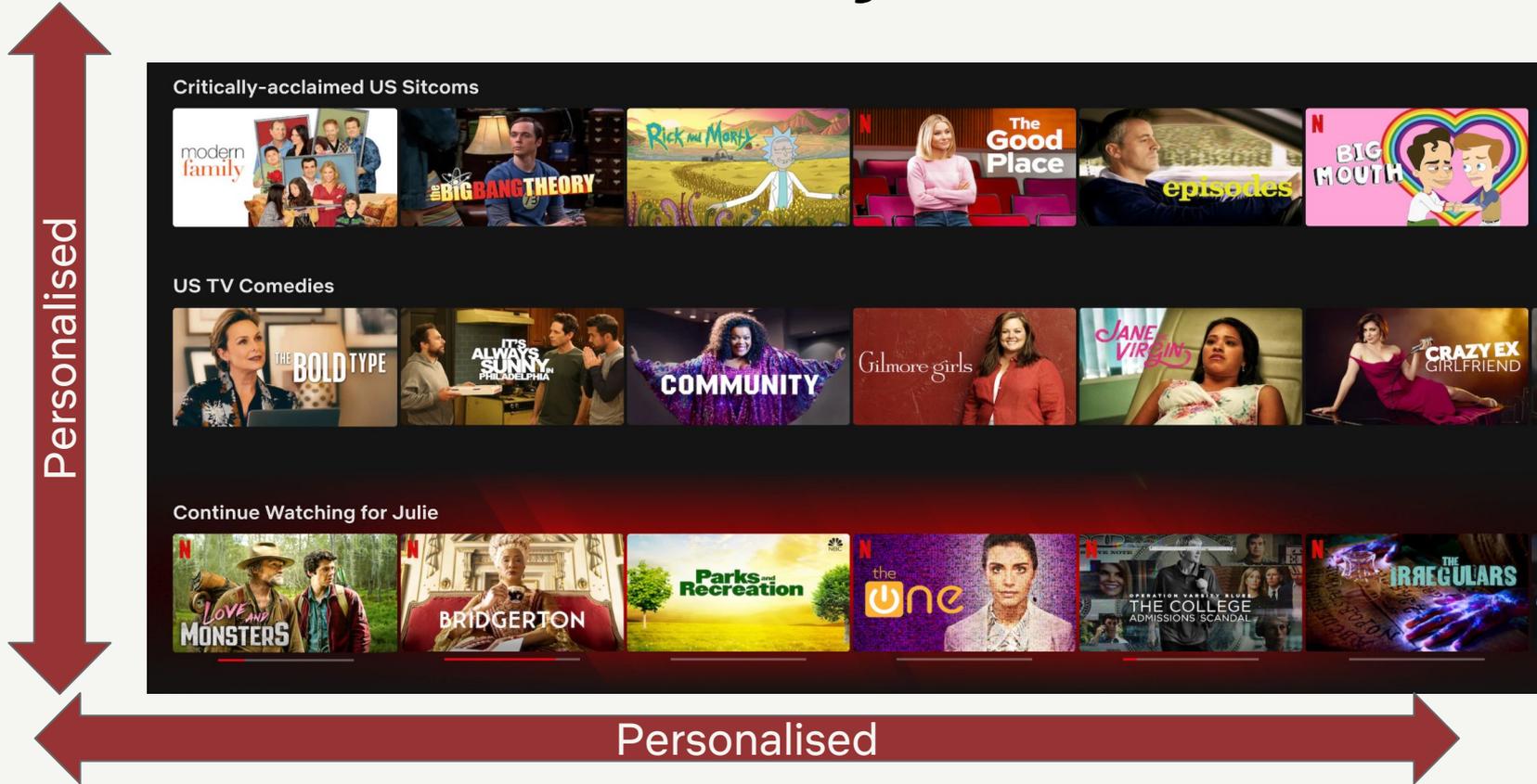
Machine Learning at Netflix



What comes to mind when you think of “Machine Learning at Netflix”?



The Netflix Prize inspired even more ML innovation internally





ROCKETMAN

TOP 10 No.5 in the UK Today

His fantastic music moves the world. But the glam rock glitz masks deep chasms of pain. At last, this is his song.

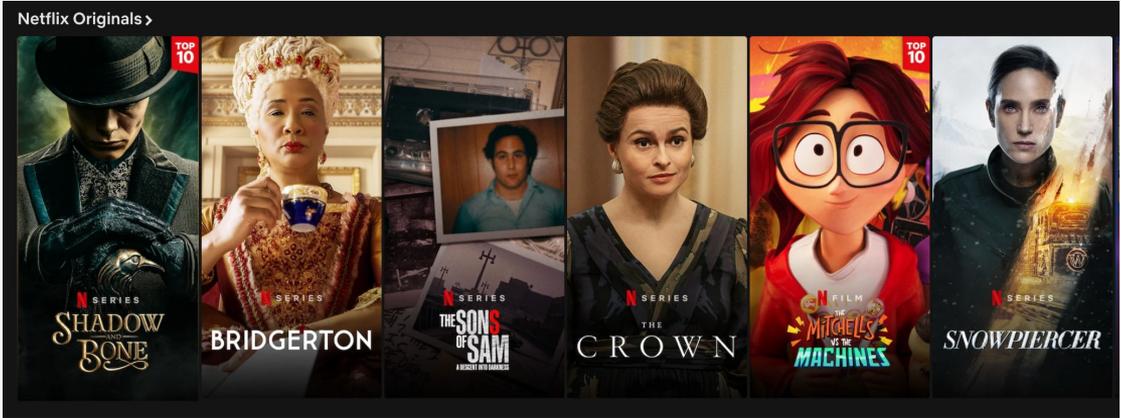
[▶ Play](#) [ⓘ More Info](#)

15

A promotional banner for the movie 'Rocketman'. It features a man in a shiny, sequined suit and a sequined hat, playing a piano on a stage in front of a large, cheering crowd. The background is a vibrant blue and purple. The title 'ROCKETMAN' is written in large, white, pixelated letters. Below the title, there is a 'TOP 10' badge and the text 'No.5 in the UK Today'. A short synopsis follows: 'His fantastic music moves the world. But the glam rock glitz masks deep chasms of pain. At last, this is his song.' At the bottom left, there are two buttons: '▶ Play' and 'ⓘ More Info'. A small '15' rating icon is visible in the bottom right corner.

Personalised

Netflix Originals >



TOP 10 **SHADOW AND BONE** N SERIES

BRIDGERTON N SERIES

THE SONS OF SAM N SERIES

THE CROWN N SERIES

MITCHELLS VS THE MACHINES N FILM

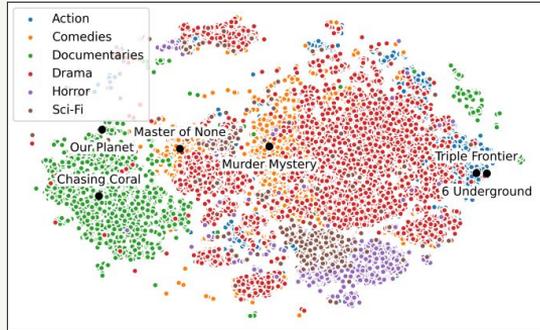
SNOWPIERCER N SERIES

A horizontal carousel of Netflix Originals. At the top left, it says 'Netflix Originals >'. Below this, there are six posters for different titles. From left to right: 1. 'Shadow and Bone' (N Series) featuring a man in a top hat and a woman, with a 'TOP 10' badge. 2. 'Bridgerton' (N Series) featuring a woman in a red dress holding a teacup. 3. 'The Sons of Sam' (N Series) featuring a man in a blue shirt, with a 'TOP 10' badge. 4. 'The Crown' (N Series) featuring a woman in a black dress. 5. 'Mitchells vs. the Machines' (N Film) featuring a woman with red hair and glasses. 6. 'Snowpiercer' (N Series) featuring a man in a black jacket. Each poster has the 'N' logo and the title.

We have expanded the use of ML to many other areas of our business

Content Catalog Planning & Optimization

Food & Travel



Adult Animation



We have expanded the use of ML to many other areas of our business

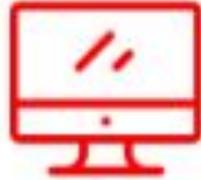
Studio Production



PRE-
PRODUCTION



PRODUCTION



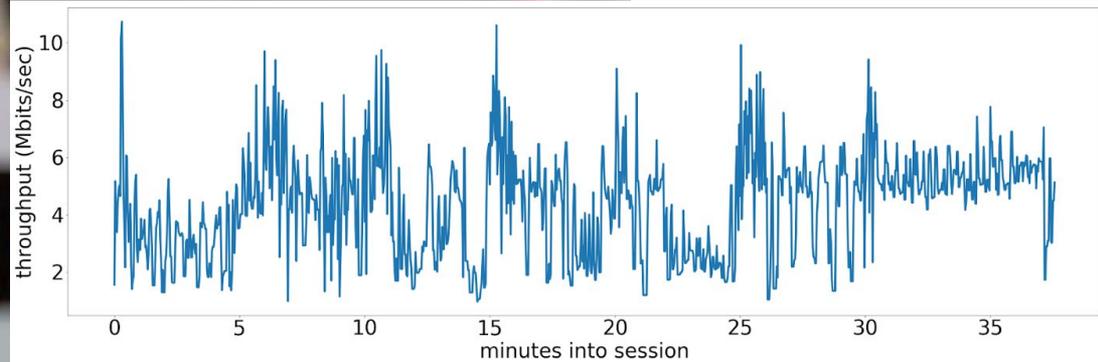
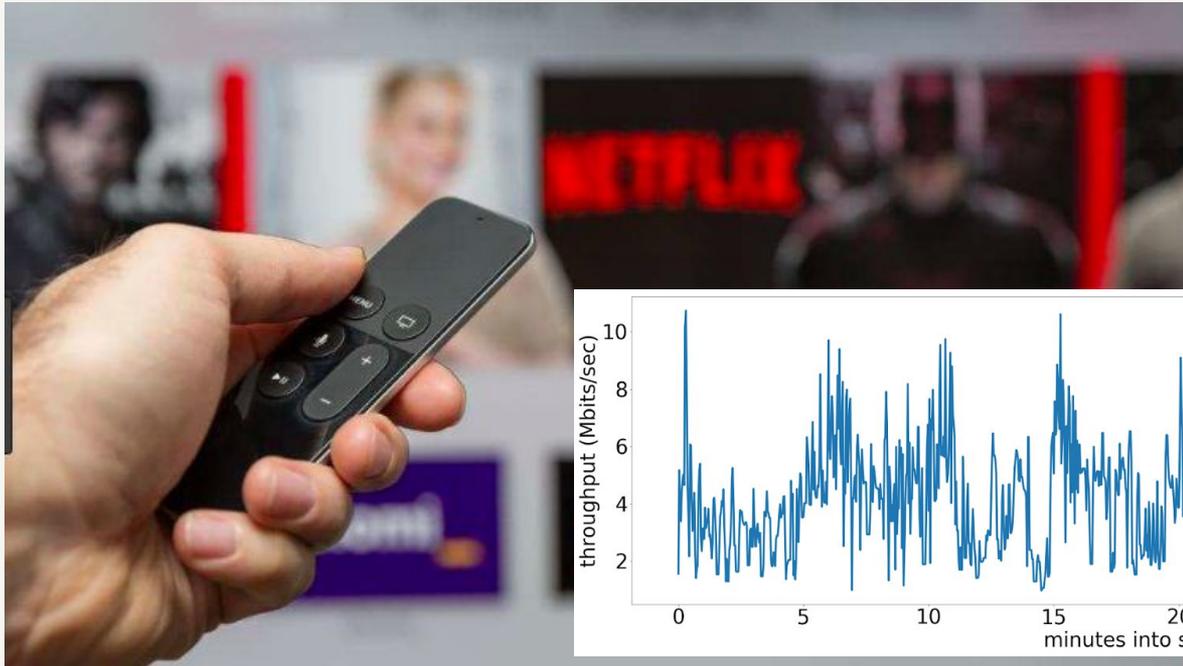
POST
PRODUCTION



LOCALIZATION
& QC

We have expanded the use of ML to many other areas of our business

Streaming Optimisation



A tale of two ML teams



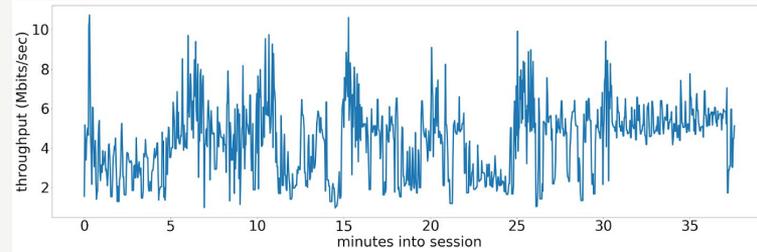
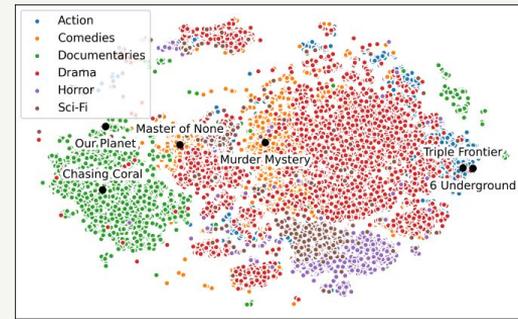
ROCKETMAN
No.5 in the UK Today
His fantastic music moves the world. But the glam rock glitz masks deep chasms of pain. At last, this is his song.

Critically-acclaimed US Sitcoms
modern family, #BIGMOUTH, Rick and Morty, The Good Place, episodes, BIG MOUTH

US TV Comedies
THE BOLD TYPE, ALL ABOUT SUNNY, COMMUNITY, Gilmore girls, GAME THEORY, CRAZY EX GIRLFRIEND

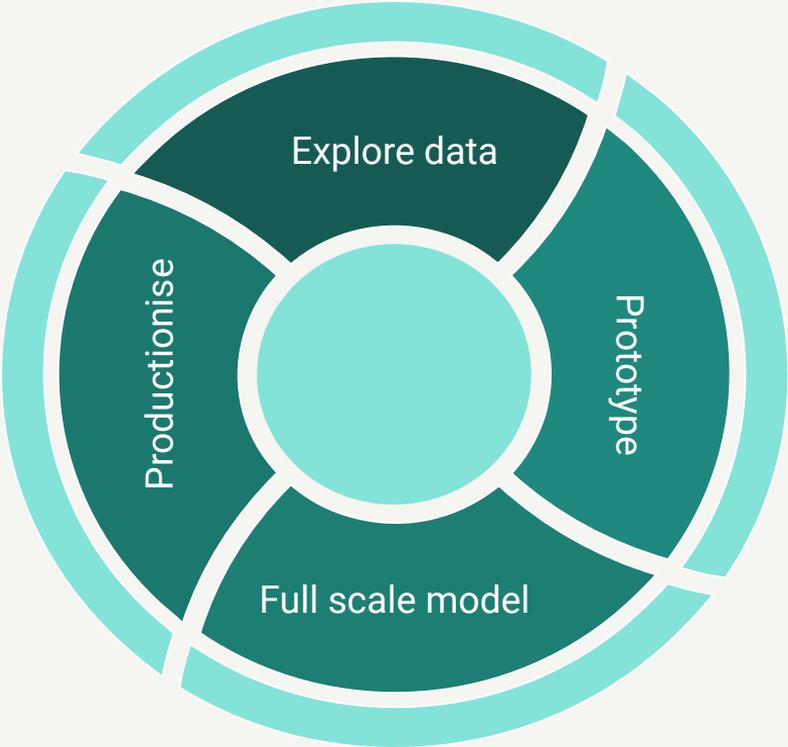
Continue Watching for Julie
MONSTERS, BRIDGERTON, Parks Recreation, one, THE COLLEGE ADMISSION SCANDAL, THE REGULARS

Personalization ML models:
Made by Algorithm Engineers



Business ML models:
Made by Data Scientists

Data Scientists & Algo Engineers are responsible from research to production



Different tech stacks

ML frameworks

Custom Java algorithms, XGBoost4j



PyTorch

dmlc
XGBoost

sklearn

Feature gen

Shared encoders



Custom ETL



Fact store

Custom fact store



Data Store



Dependency Management



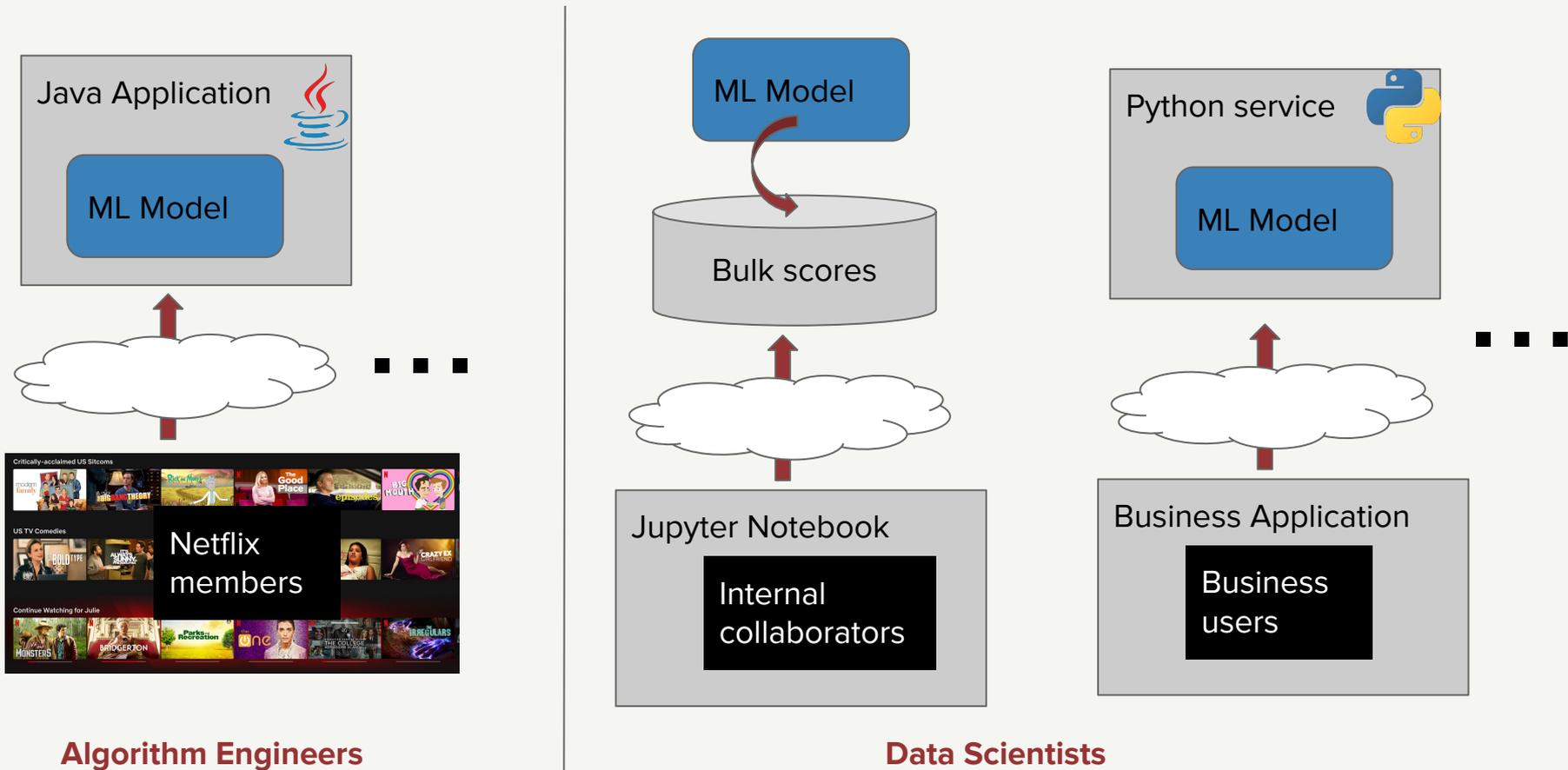
Languages



Algorithm Engineers

Data Scientists

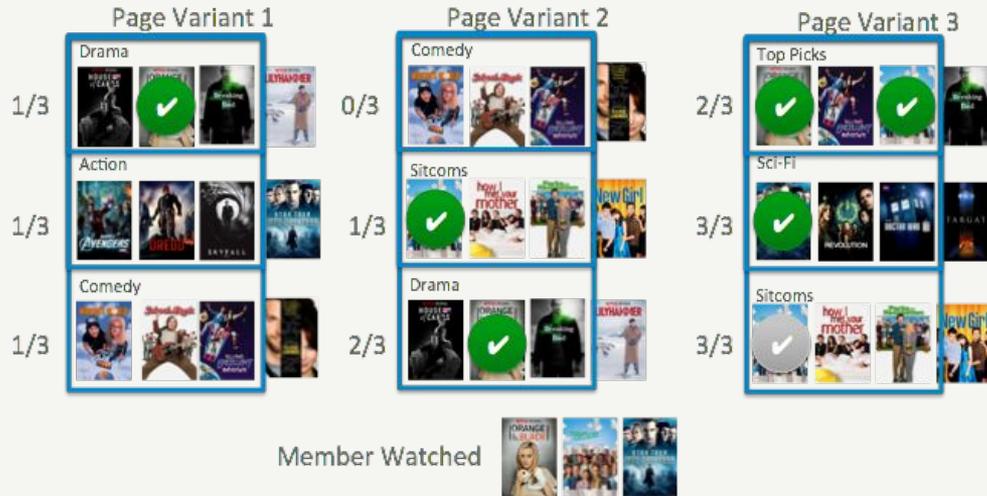
Many flavors of "production"



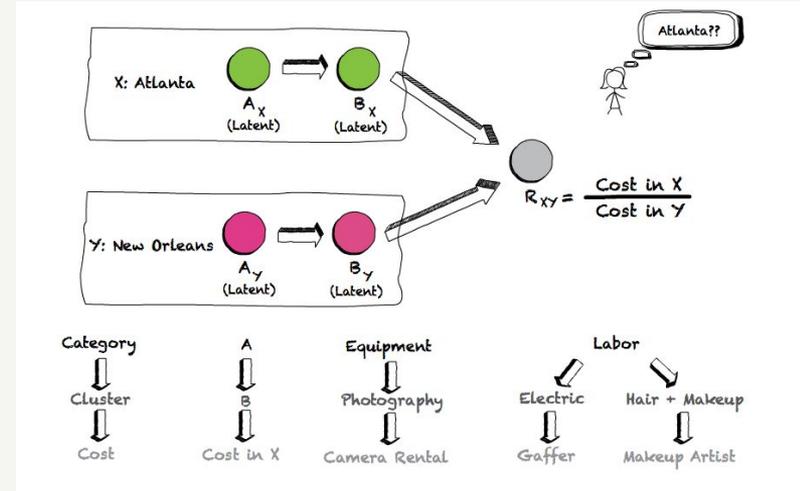
Algorithm Engineers

Data Scientists

Different objectives

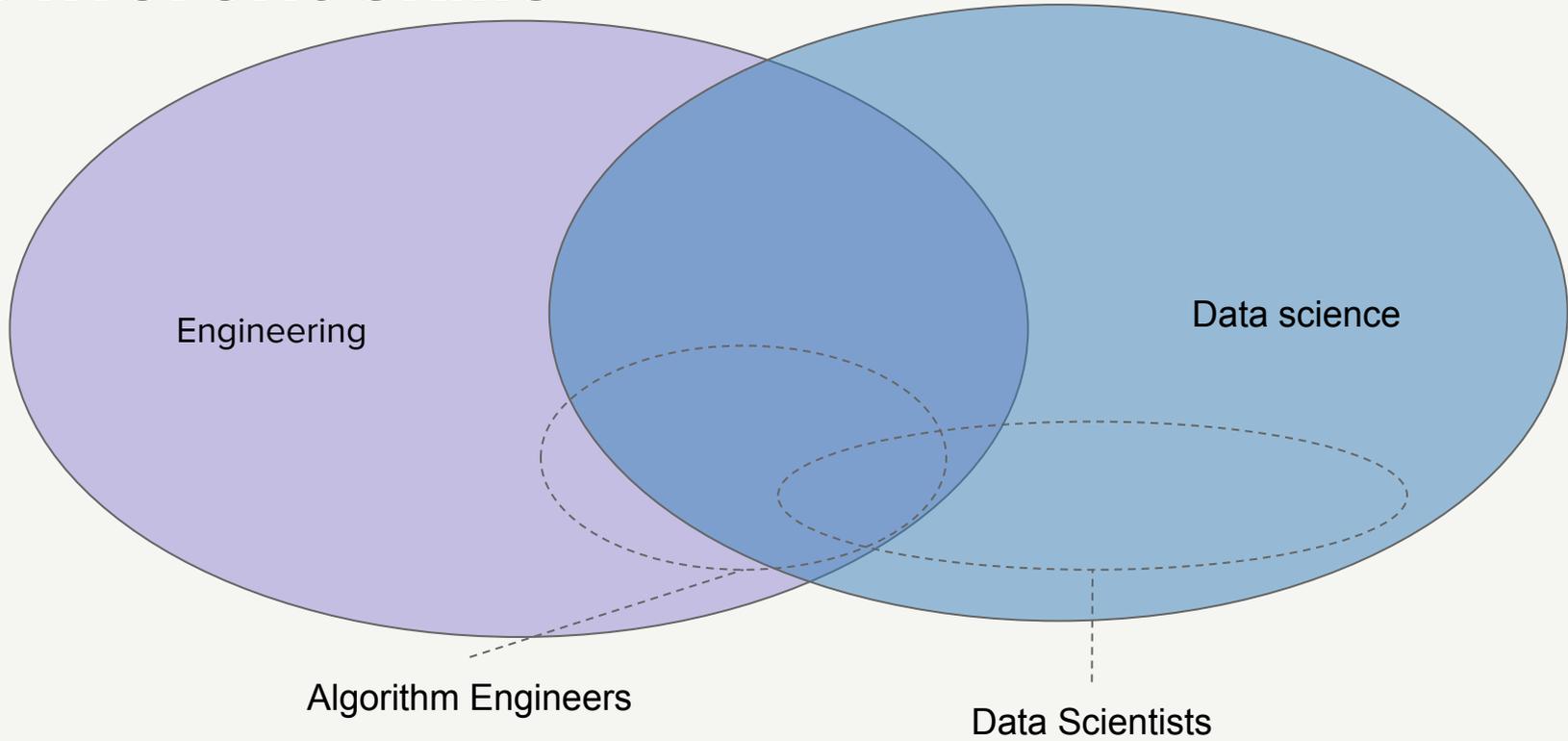


Iteratively increase member joy through A/B tested algorithm improvements



Reach a business objective or improve on current methods using data science

Different skills



Different organisations

Product Engineering

AE team 1

AE team 2

AE team 3



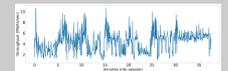
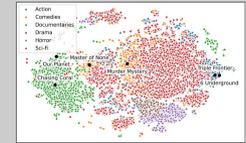
Algorithm Engineers

Data Science & Engineering

Business vertical 1

Business vertical 2

Business vertical 3



...

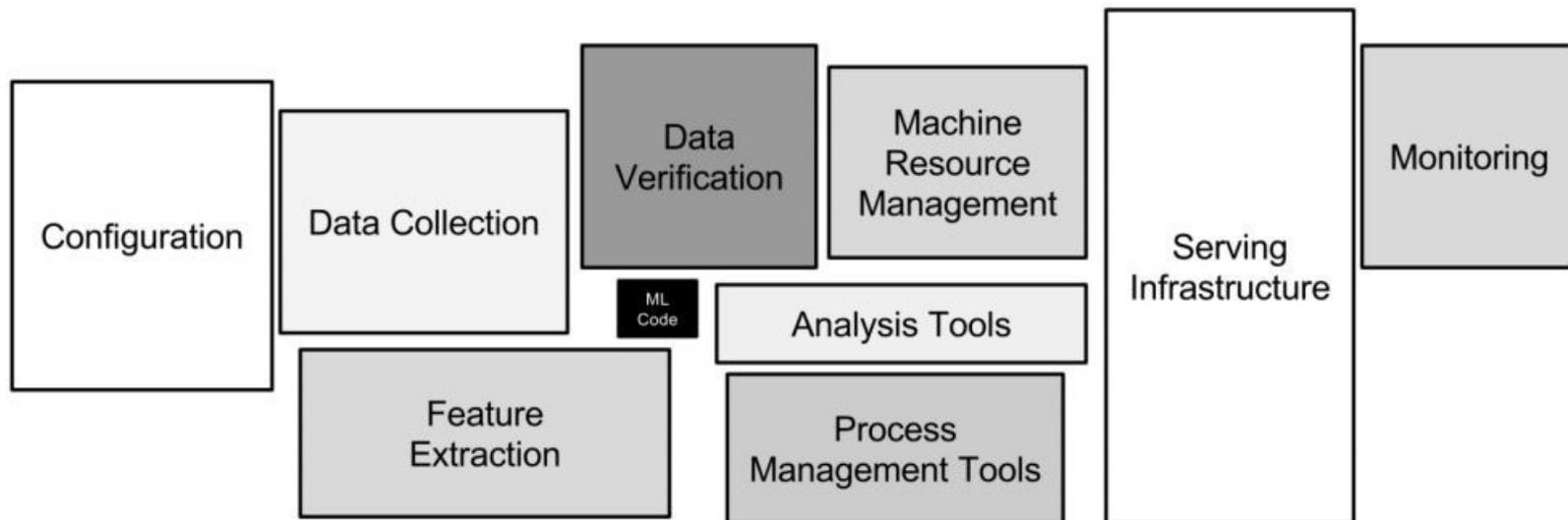
Data Scientists

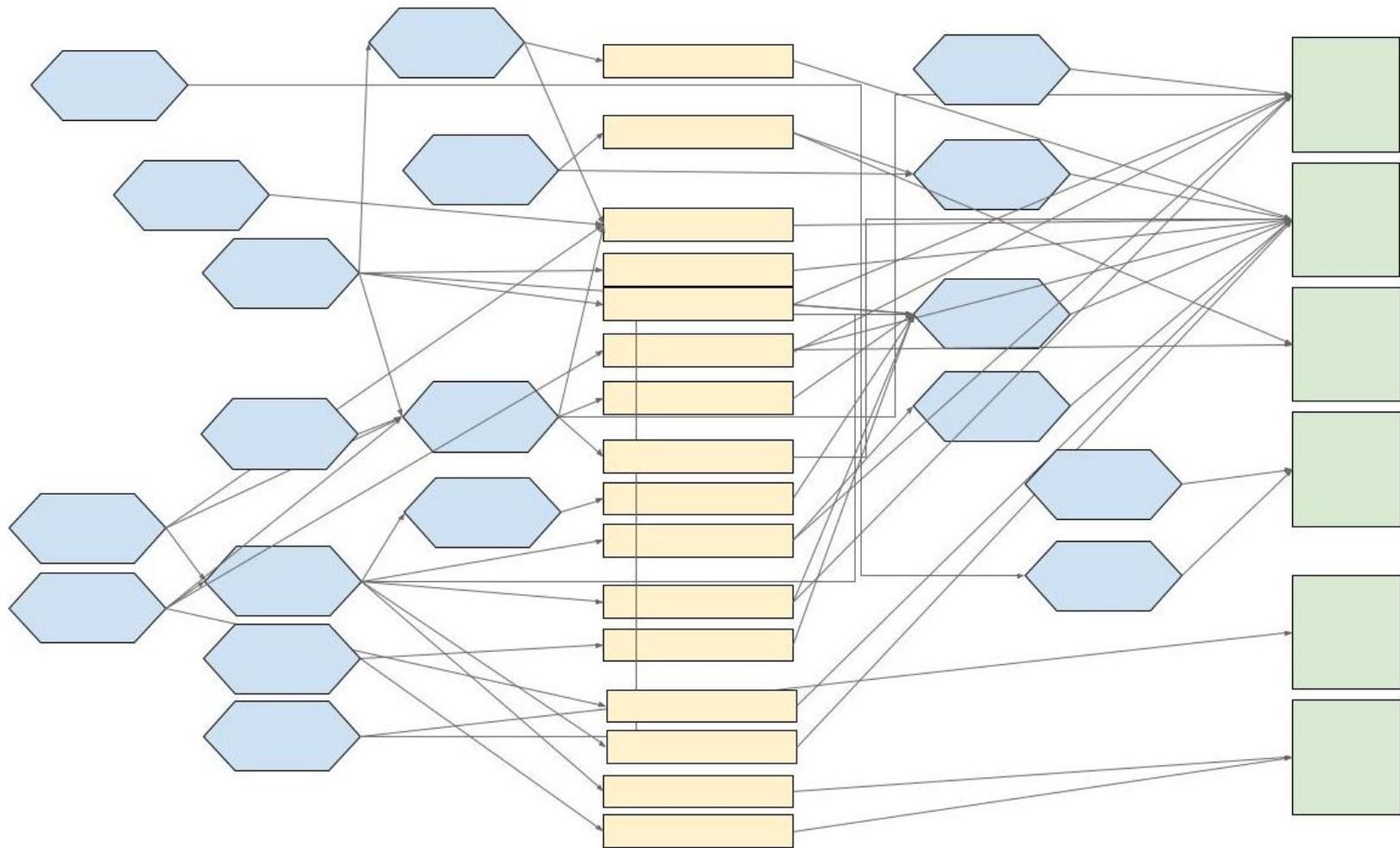
The case for ML Infrastructure



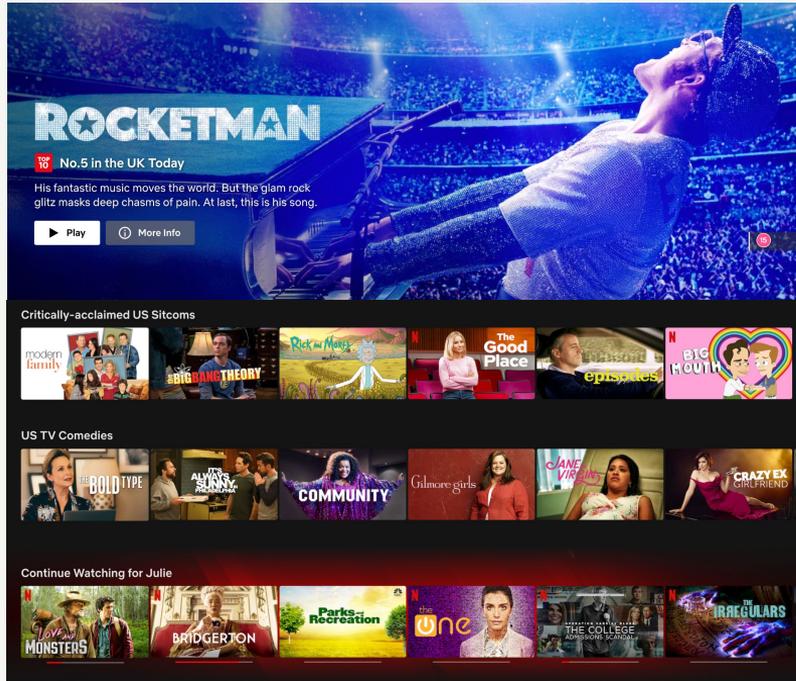
Hidden Technical Debt in Machine Learning Systems

D. Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips
{dsculley, gholt, dgg, edavydov, toddphillips}@google.com
Google, Inc.





Algorithm Engineering pain points

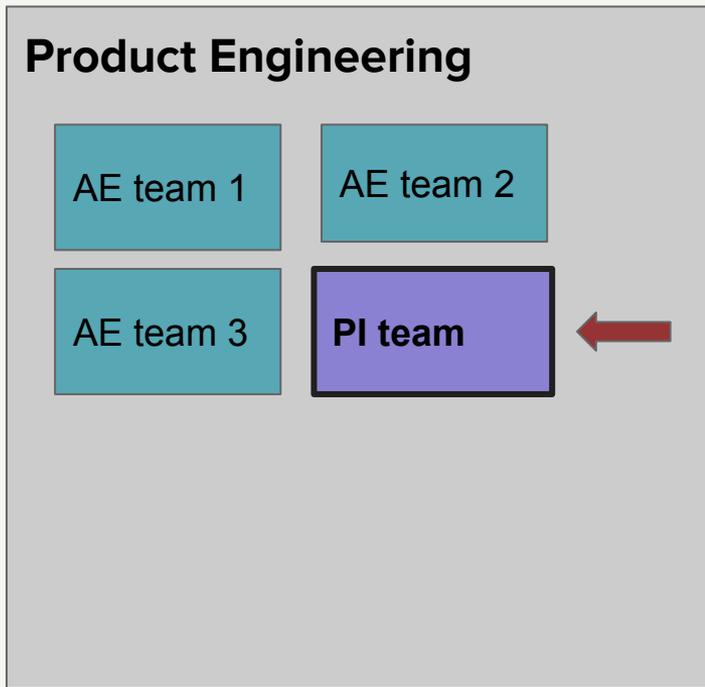


- Too much time spent on preventable bugs (e.g., online/offline feature disparity)
- Individual researchers building similar bits of infra to solve local problems
- Need to operationalize hundreds of A/B tests per year (e.g., standard way to version, publish and score models)

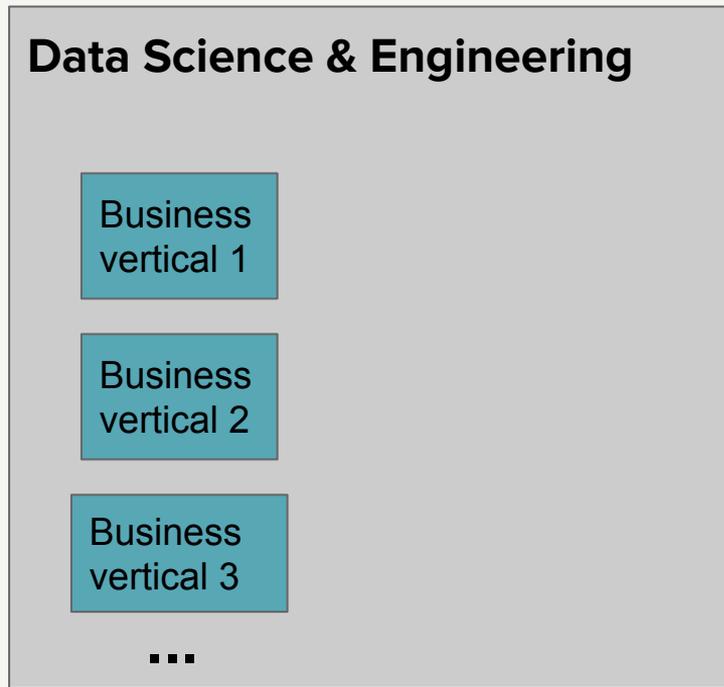
Personalisation Infrastructure (PI) team mission

*Innovate on **large-scale Machine Learning Infrastructure** to help the Algorithm
Engineering teams deliver more moments of joy for Netflix members.*

Where should the PI team go?

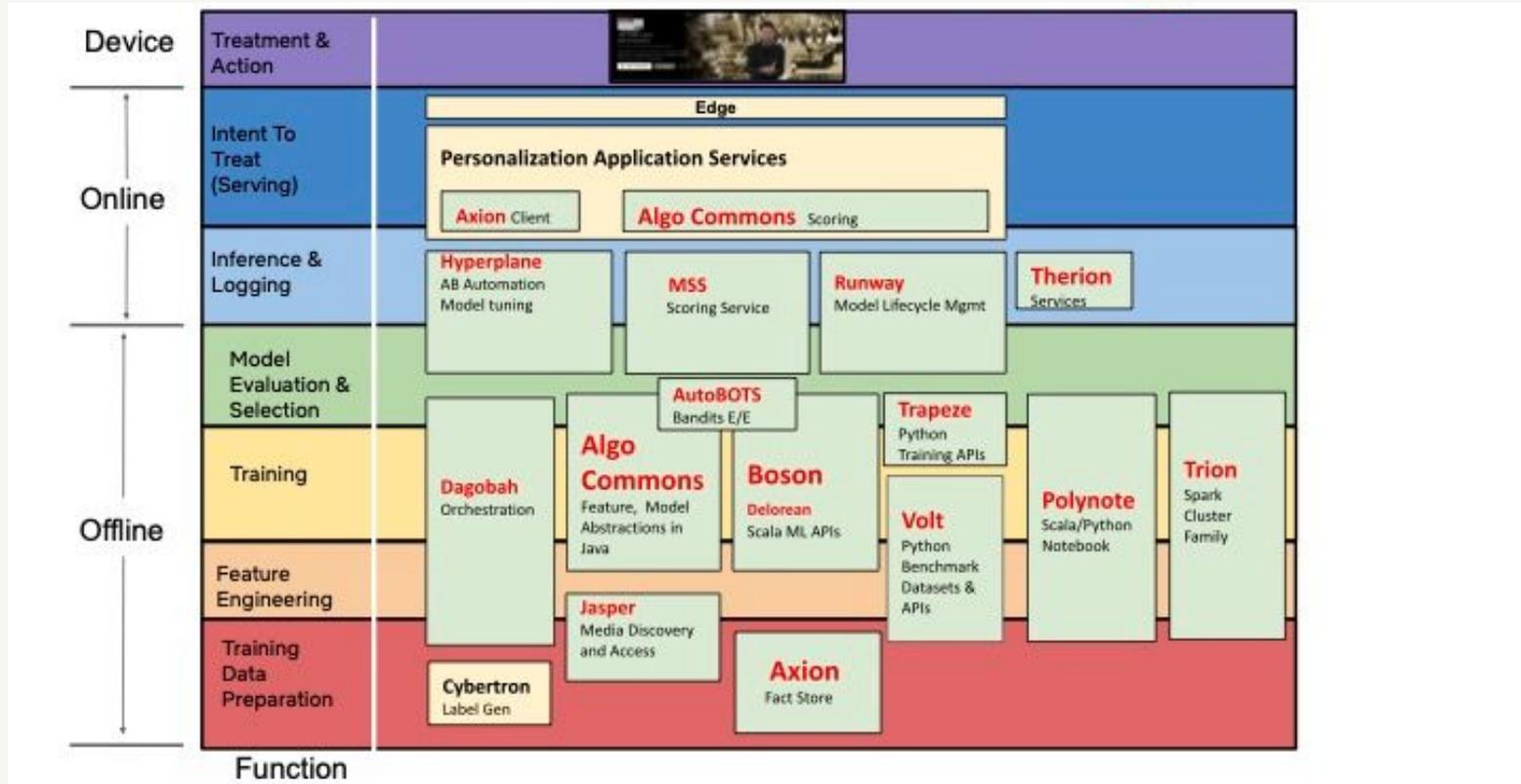


Algorithm Engineers



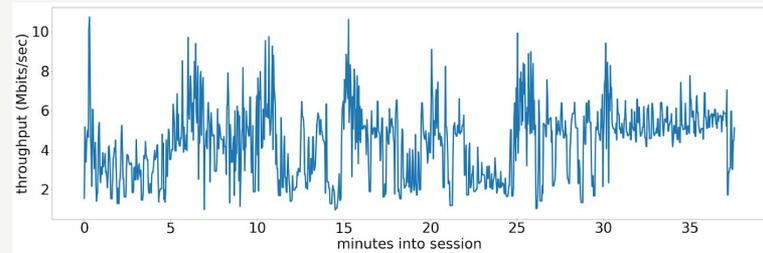
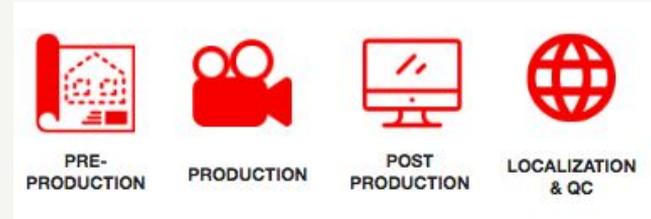
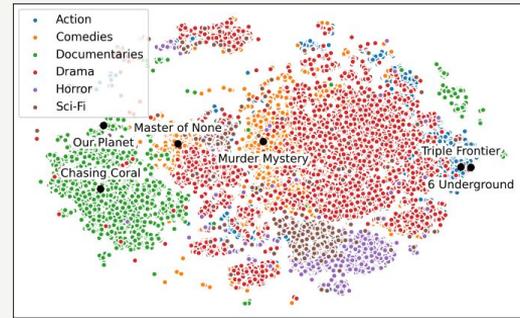
Data Scientists

Personalisation Infrastructure stack



Data Science pain points

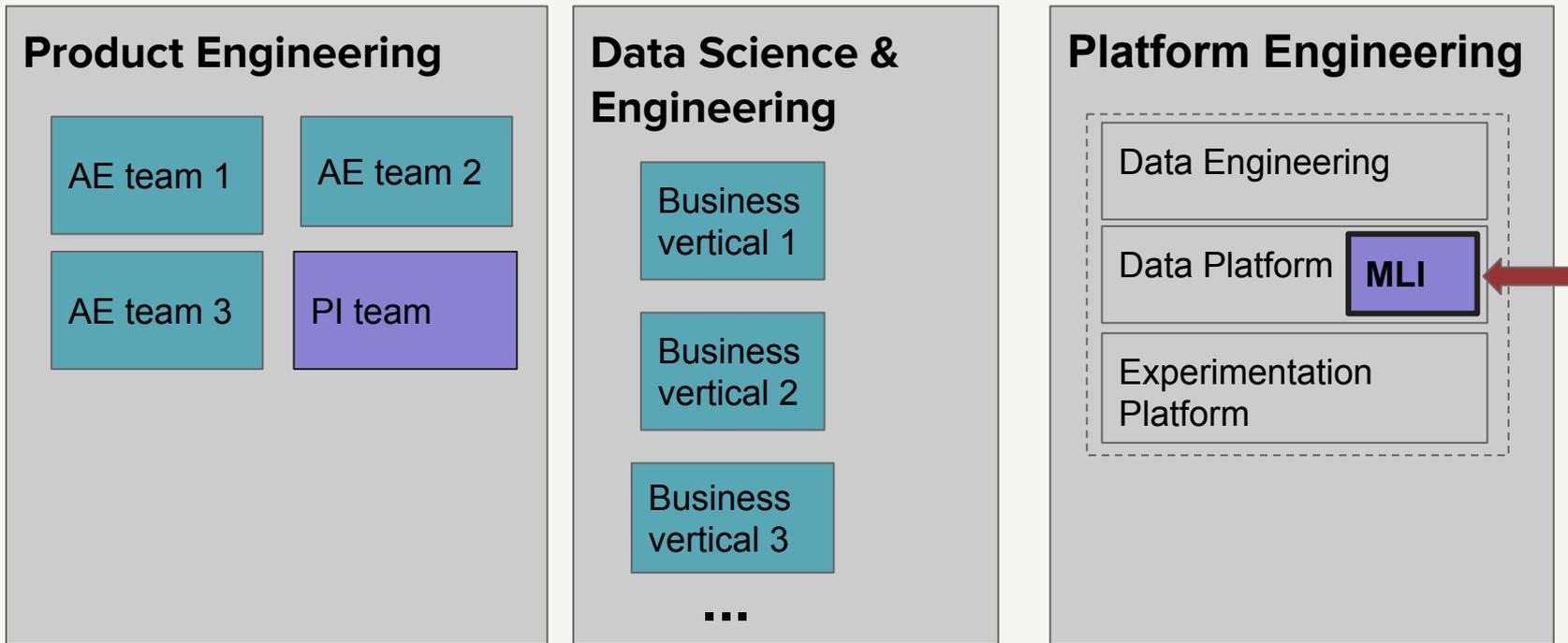
- Getting to production is slow and painstaking
- Reproducing a colleague's work is difficult or impossible
- Handing off a project requires lots of hand-holding



Machine Learning Infrastructure (MLI) team mission

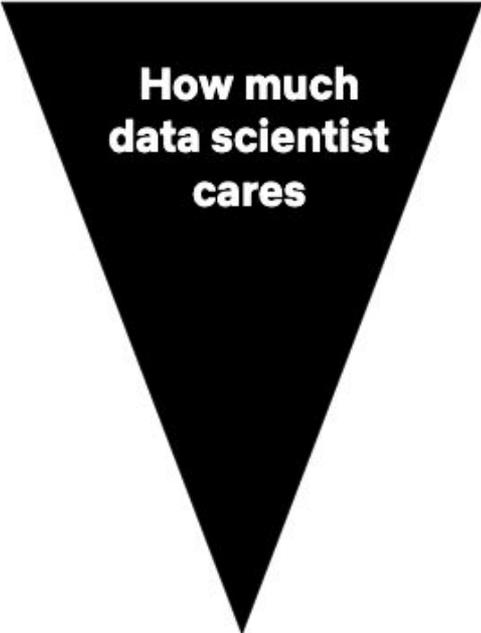
*Help **Data Scientists become** as **productive** as possible from prototype to production.*

Where should the MLI team go?



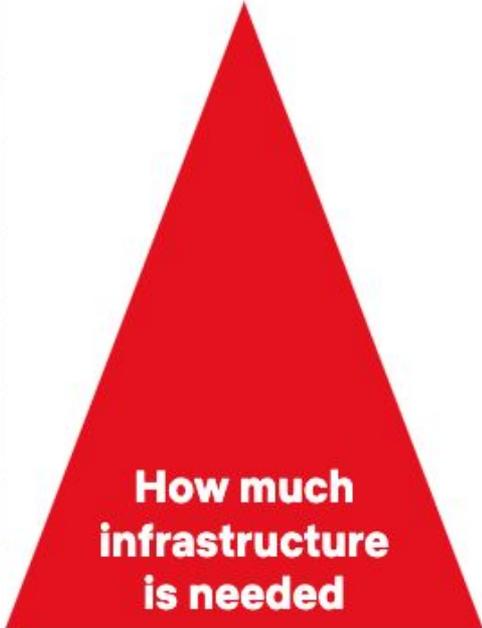
Data Scientists

MLI built Metaflow for Data Scientists



**How much
data scientist
cares**

Model Development
Feature Engineering
Model Operations
Versioning
Architecture
Job Scheduler
Compute Resources
Data Warehouse



**How much
infrastructure
is needed**

MLI built Metaflow for Data Scientists

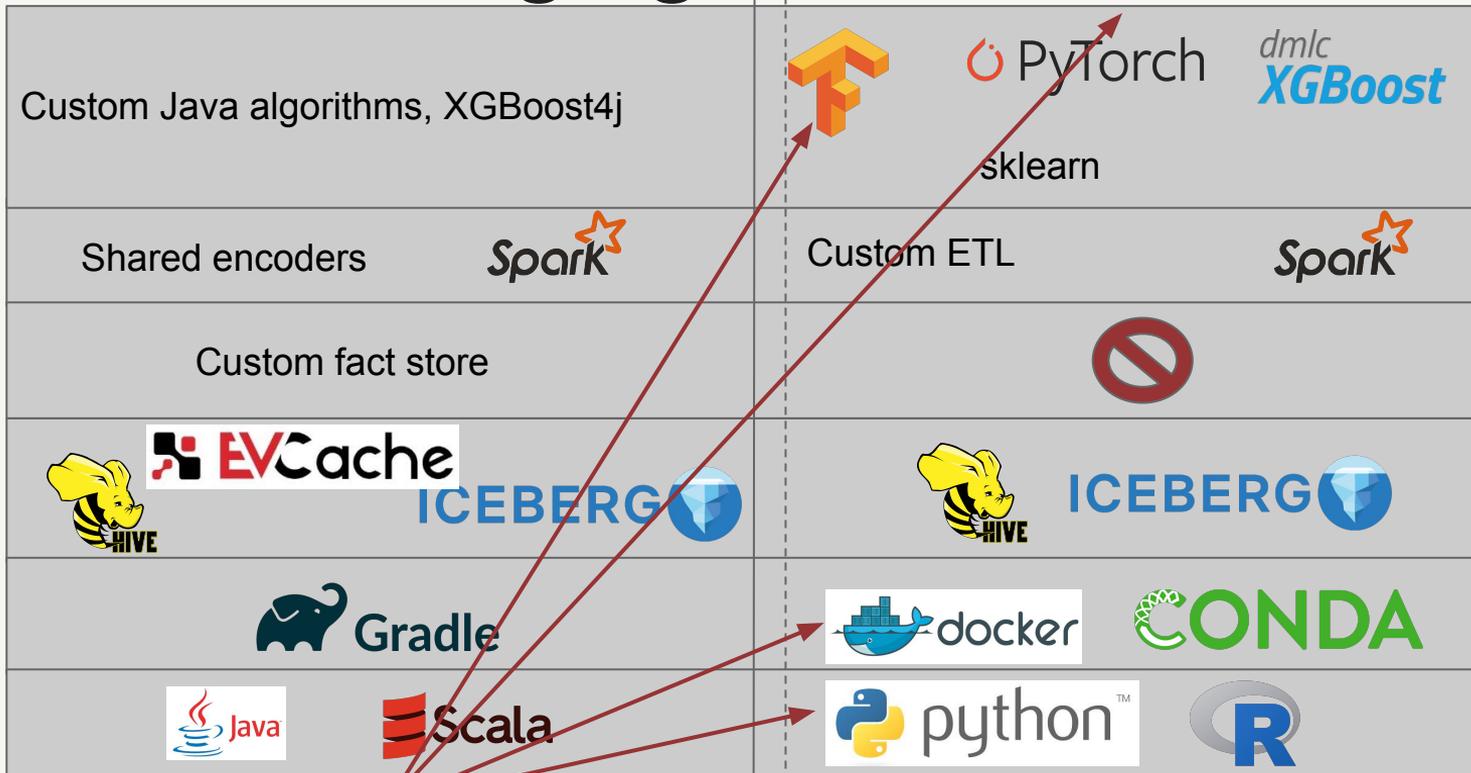


Merging into ML Platform



Tech stacks converging

ML frameworks



Algorithm Engineers

Data Scientists

Opportunity to share facts & features

Feature gen

Shared encoders



Custom ETL



Fact store

Custom fact store

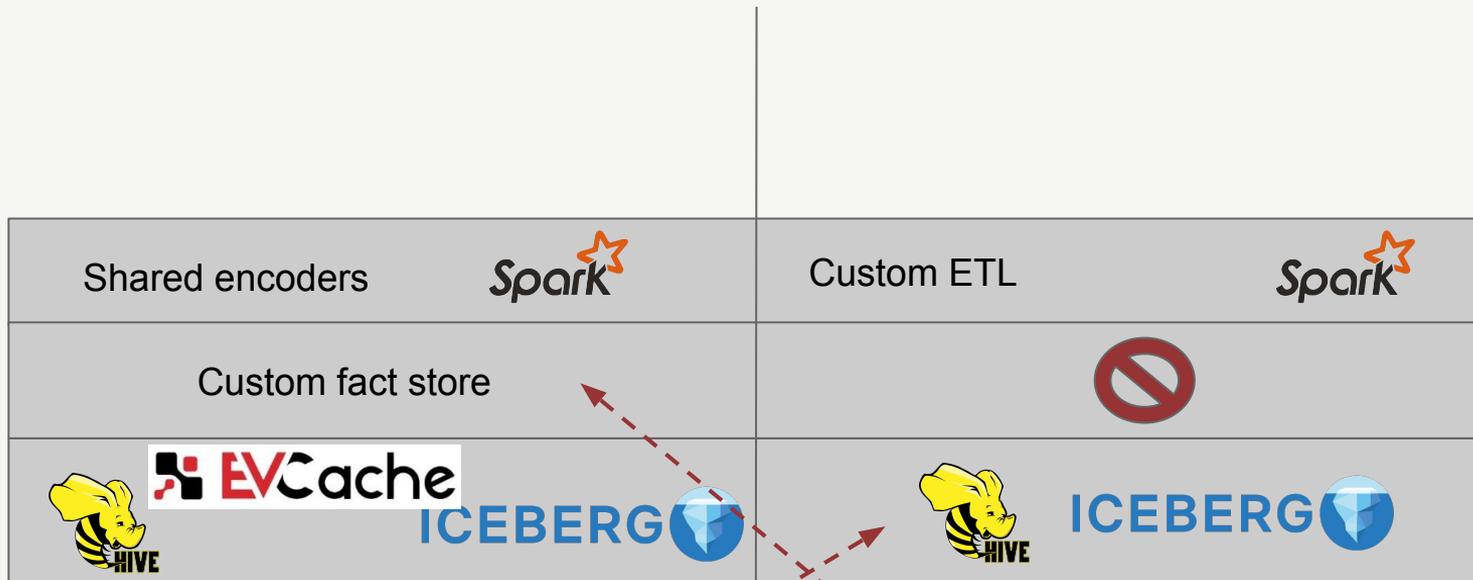


Data Store

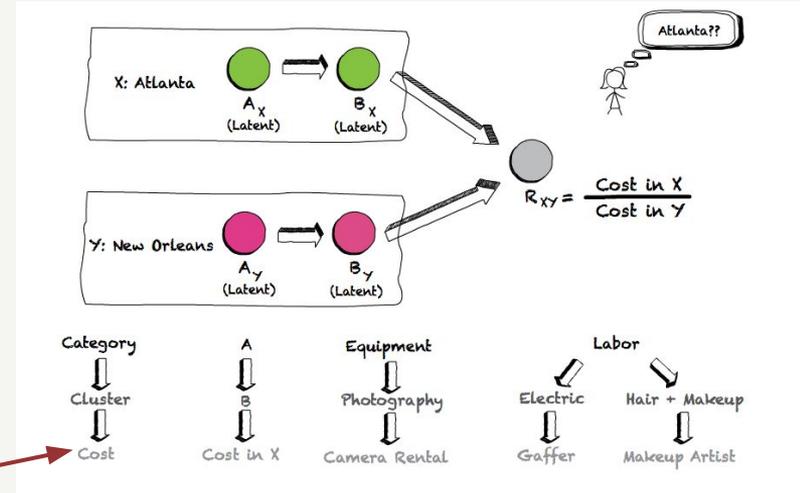
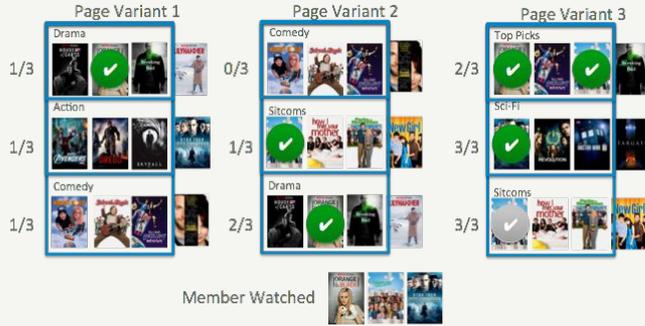


Algorithm Engineers

Data Scientists



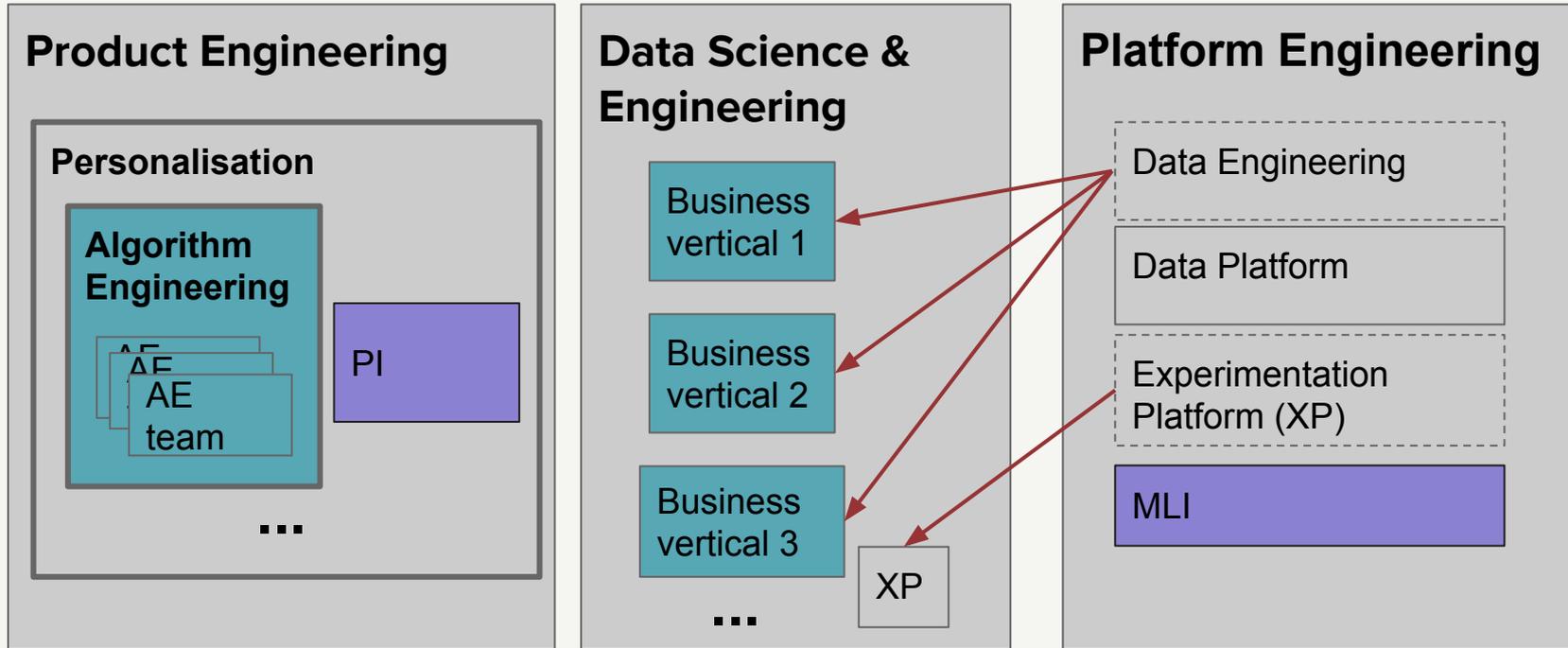
Increasing need for collaboration



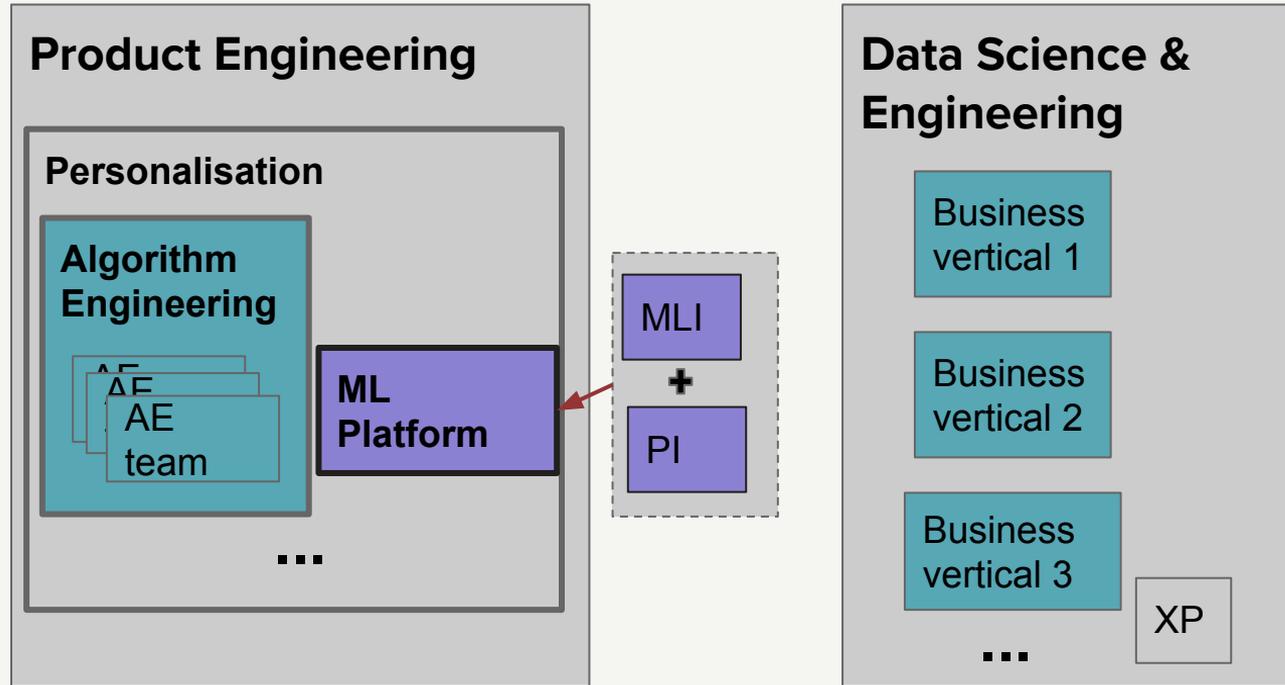
Algorithm Engineers working in more areas of the business

Data Scientists experimenting with different areas of the product

Organisational shifts



Forming Machine Learning Platform



Takeaways for ML Platform teams



1. Know your users

- What skills do they have?
- What technologies do they use?
- What environments are they working in and what are their daily struggles?
- Where do their responsibilities begin and end?
- Who do they partner with?

2. Know the business

- What business problems are being solved with ML?
- How is success defined?
- What is the next big milestone and what's in the way of achieving that?

3. Anticipate and adapt

- Look out for signs that the current organisational paradigm is no longer working. Look for shifts in:
 - a. Business objectives
 - b. Technology usage
 - c. Surrounding organisations
 - d. Collaboration models
- Look as far ahead as possible when making organisational changes

Thank you.

We're hiring!

<https://bit.ly/3uSpeTB>

<https://jobs.netflix.com/jobs/72787678>

<https://jobs.netflix.com/jobs/77169084>

NETFLIX