Evolving the Machine Learning Platform Organisation at Netflix: A case study

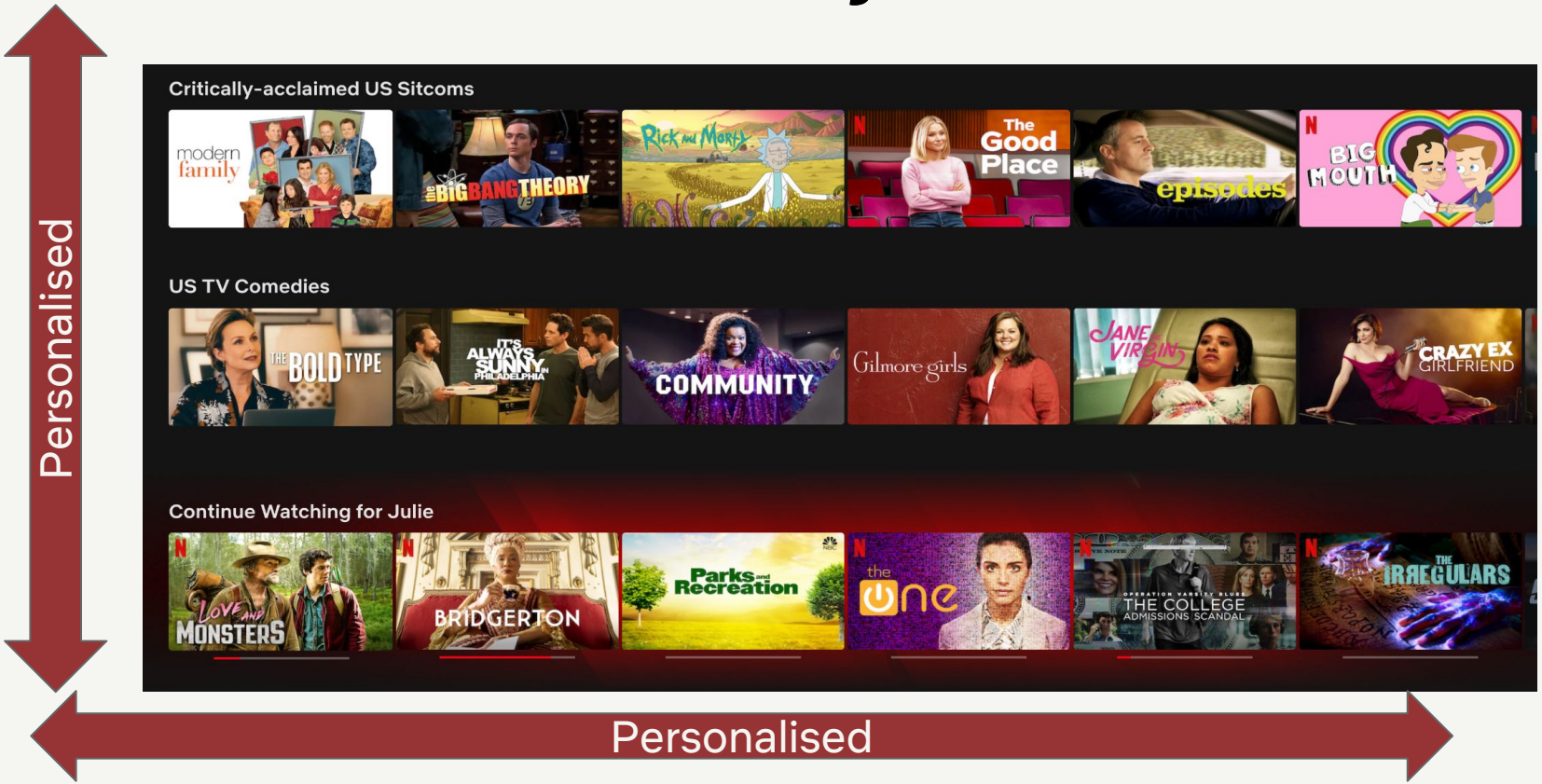
Julie Amundson (@yakticus)
2021.05.11

NETFLIX

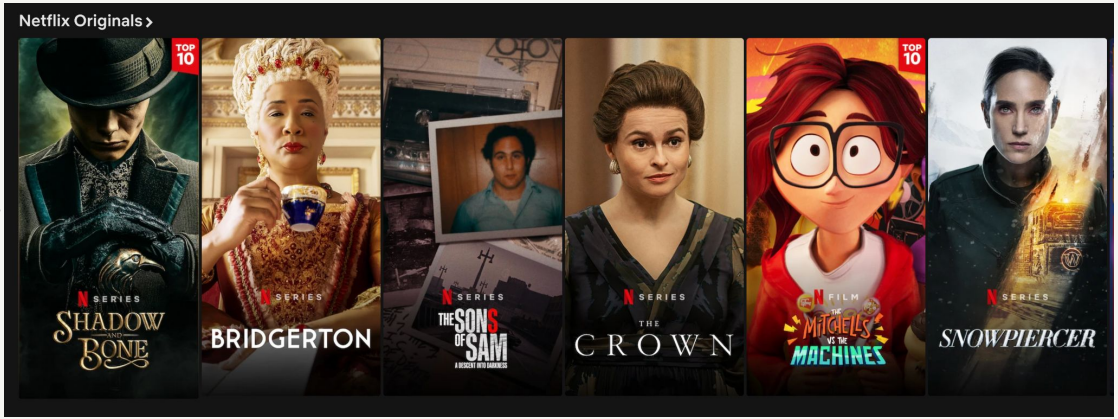
Machine Learning at Netflix

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

The Netflix Prize inspired even more ML innovation internally

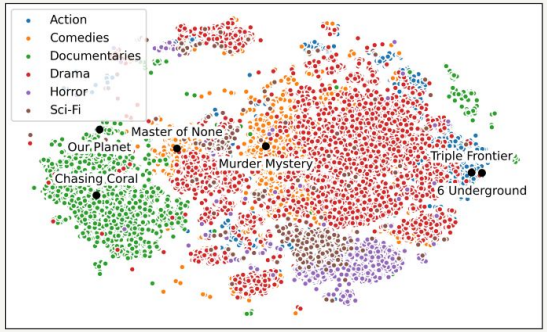


Personalised



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

Content Catalog Planning & Optimization



Food & Travel

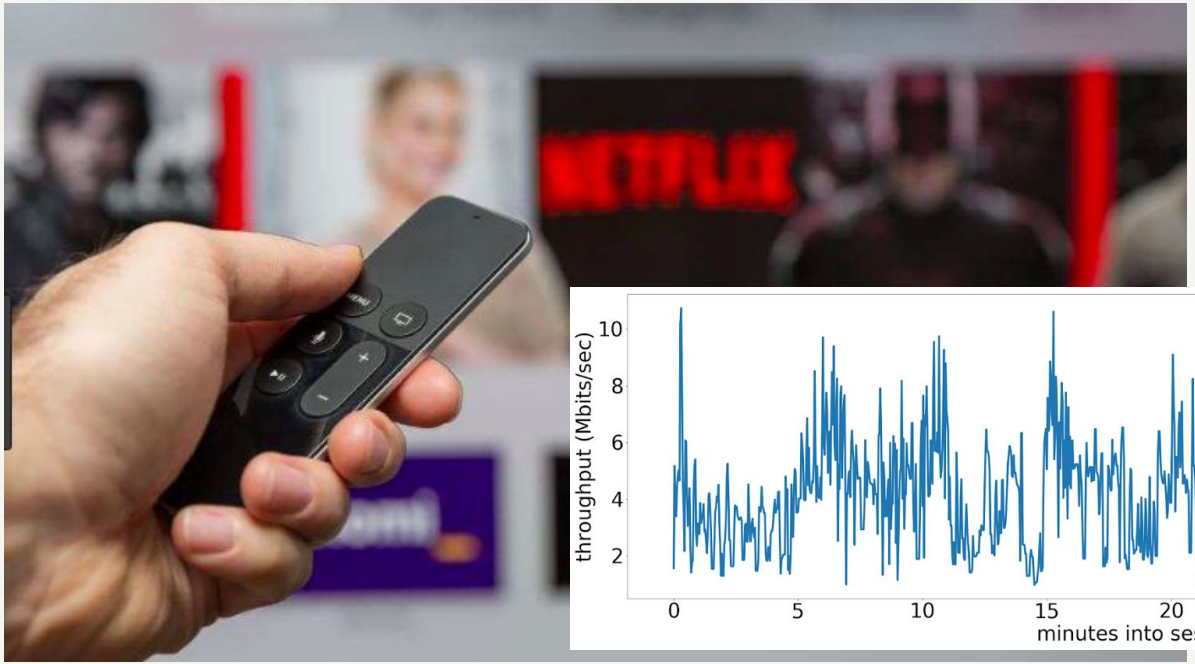
Adult Animation

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

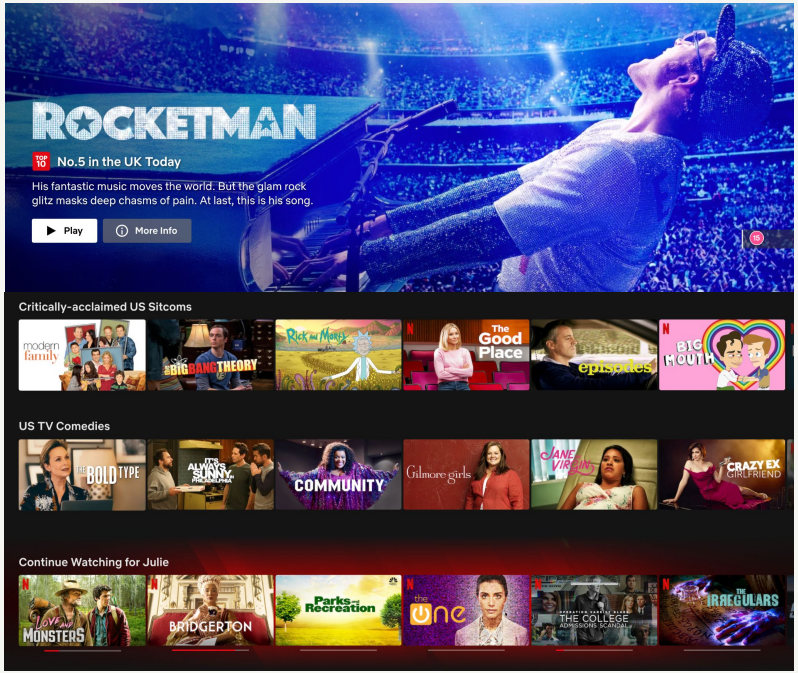
Studio Production

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

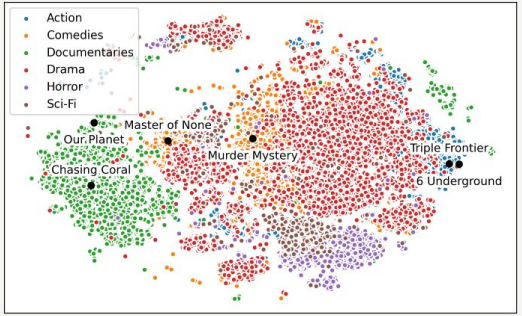
Streaming Optimisation

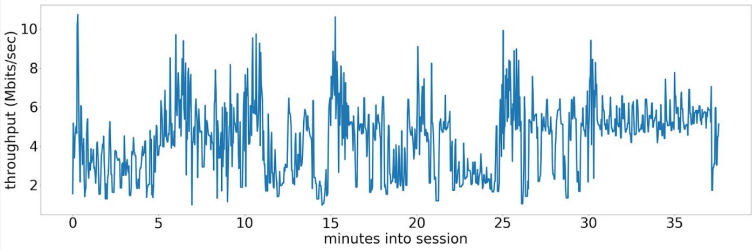


A tale of two ML teams



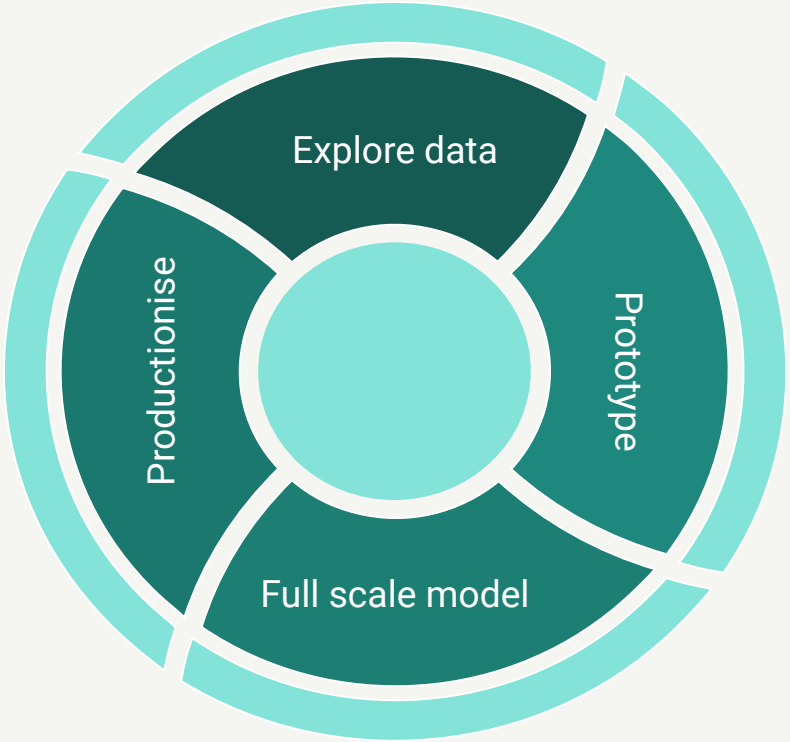
Personalization ML models:
Made by Algorithm Engineers



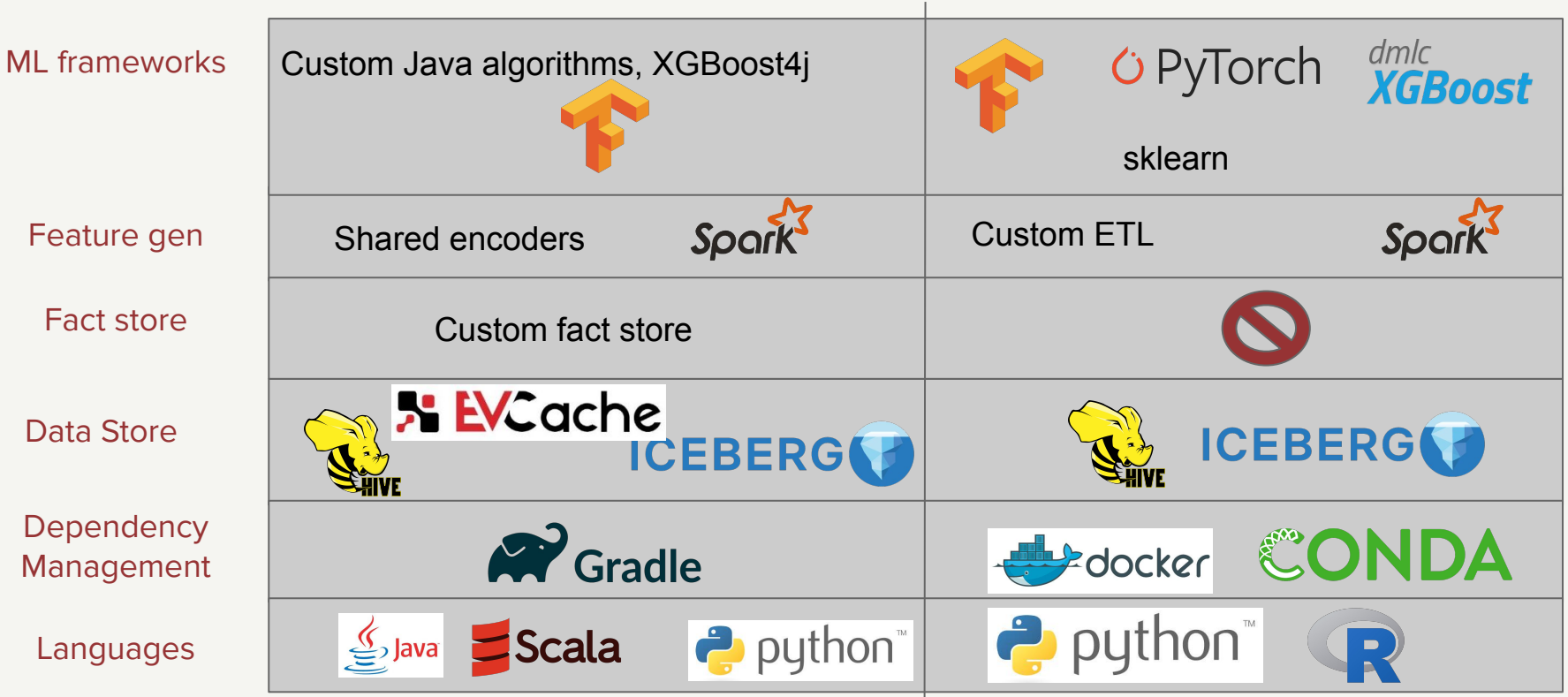


Business ML models:
Made by Data Scientists

Data Scientists & Algo Engineers are responsible from research to production



Different tech stacks



Many flavors of "production"

Java Application

ML Model

Algorithm Engineers

ML Model

Bulk scores

Jupyter Notebook

Internal collaborators

Python service

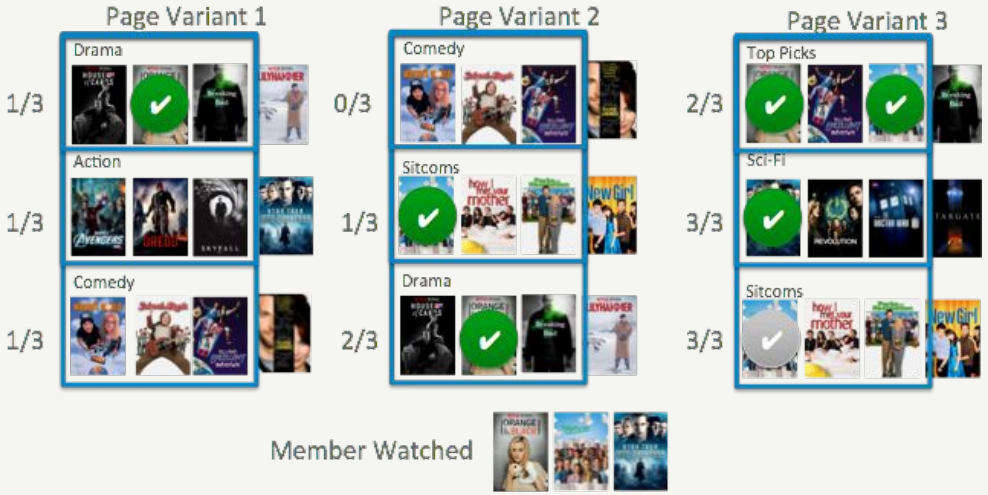
ML Model

Business Application

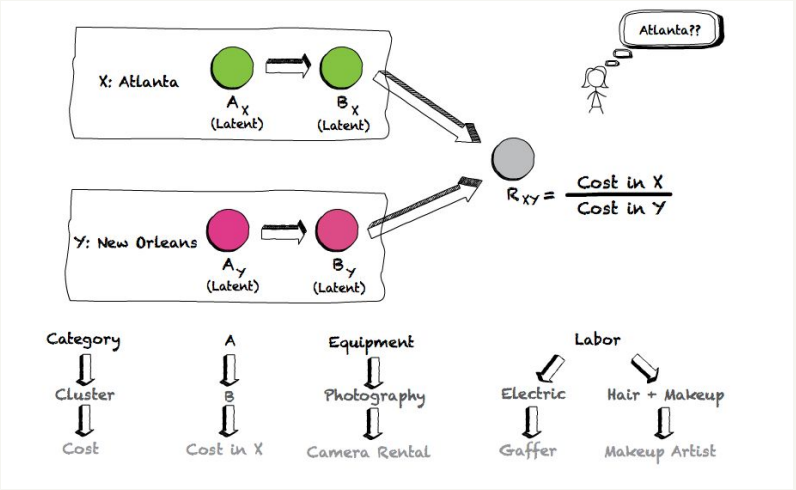
Business users

Data Scientists

Different objectives

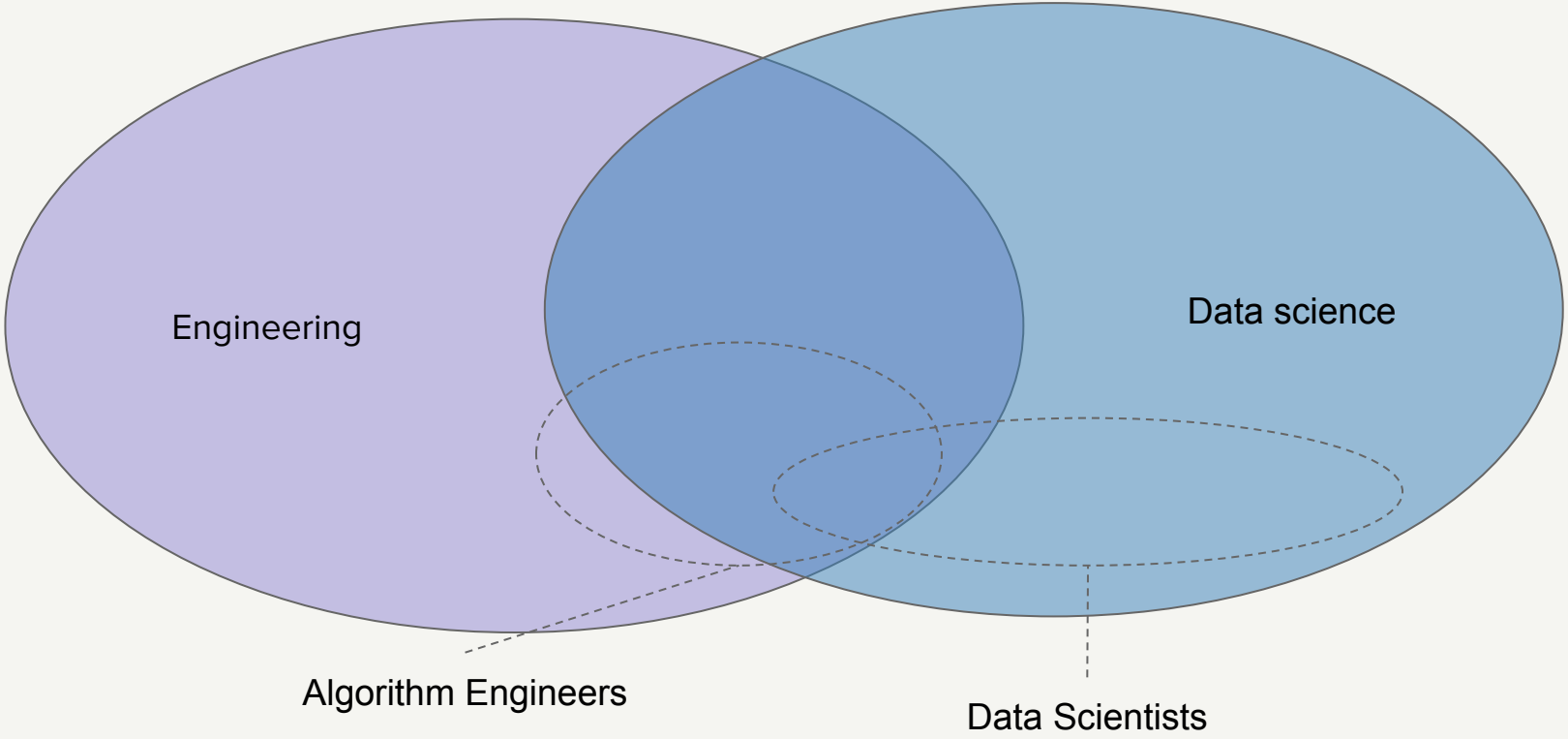


Iteratively increase member joy through A/B tested algorithm improvements



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

Different skills



Different organisations

Product Engineering

AE team 1

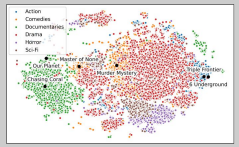
AE team 2

AE team 3

Algorithm Engineers

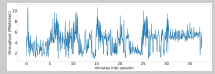
Data Science & Engineering

Business vertical 1



Business vertical 2

Business vertical 3



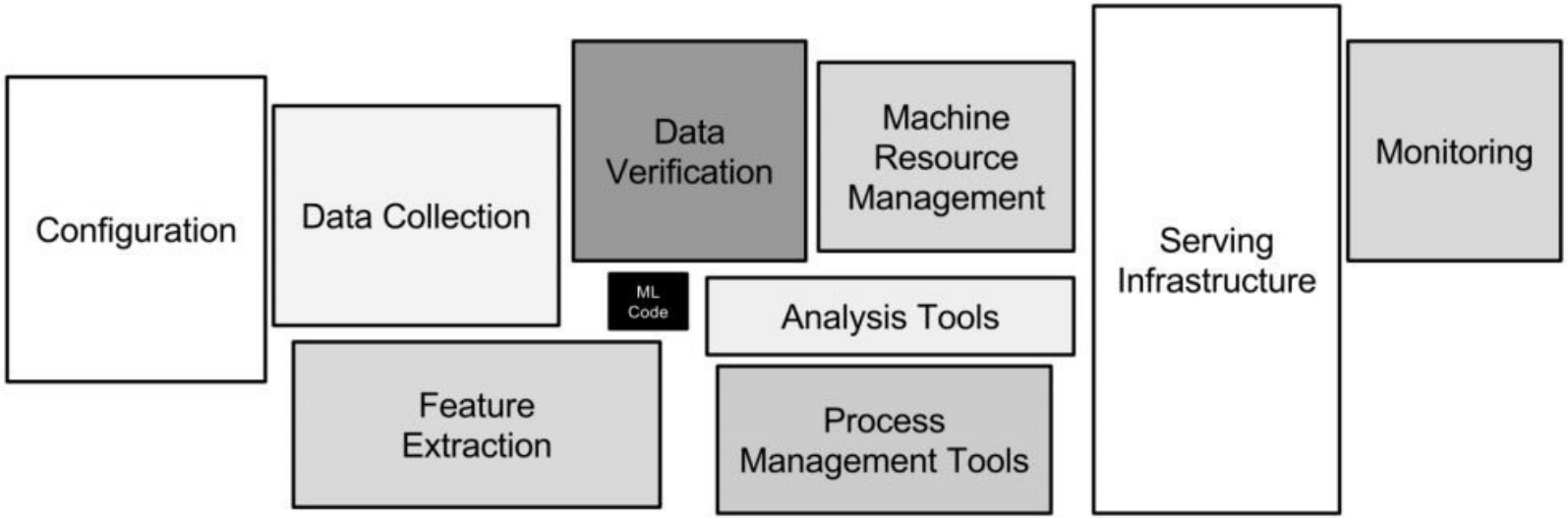
...

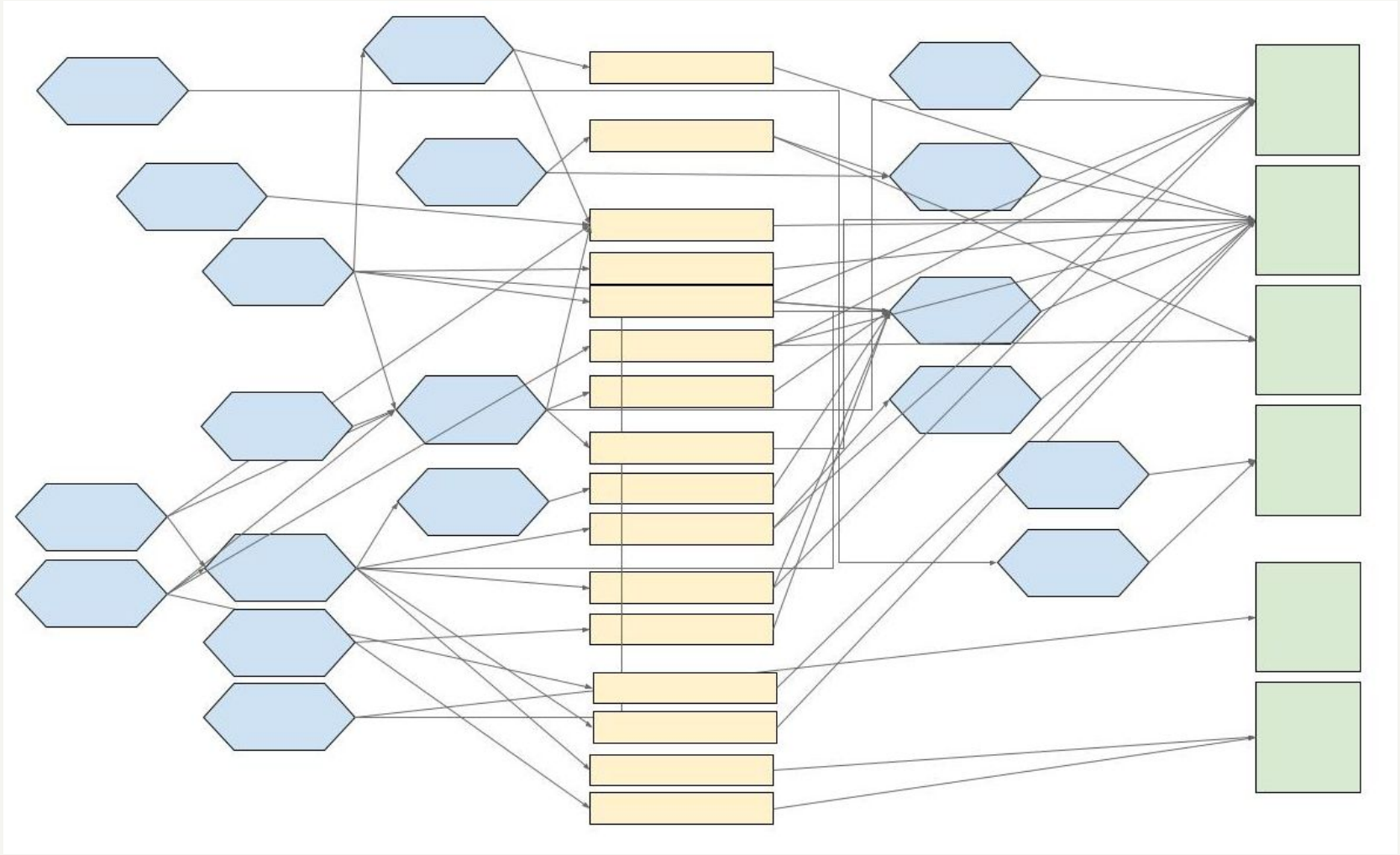
Data Scientists

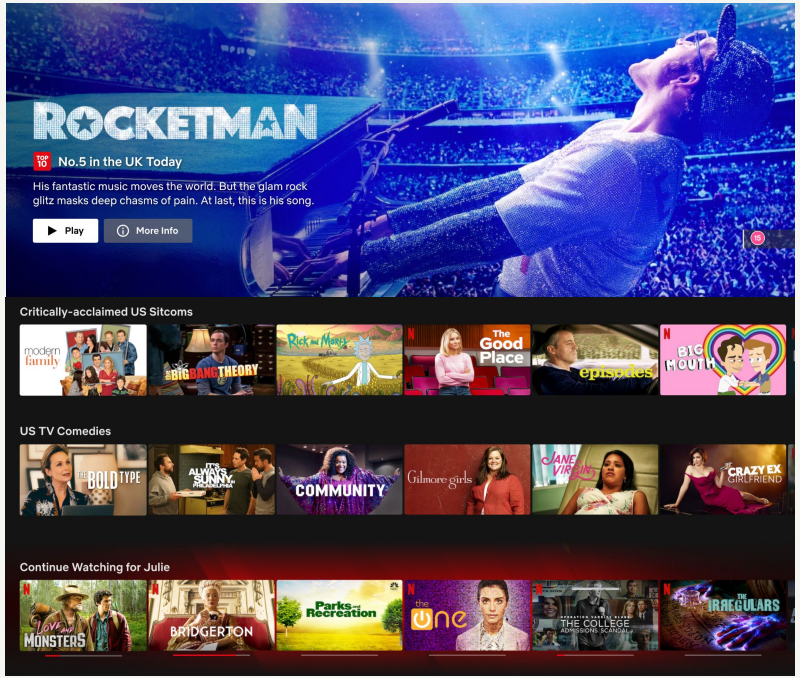
The case for ML Infrastructure

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.







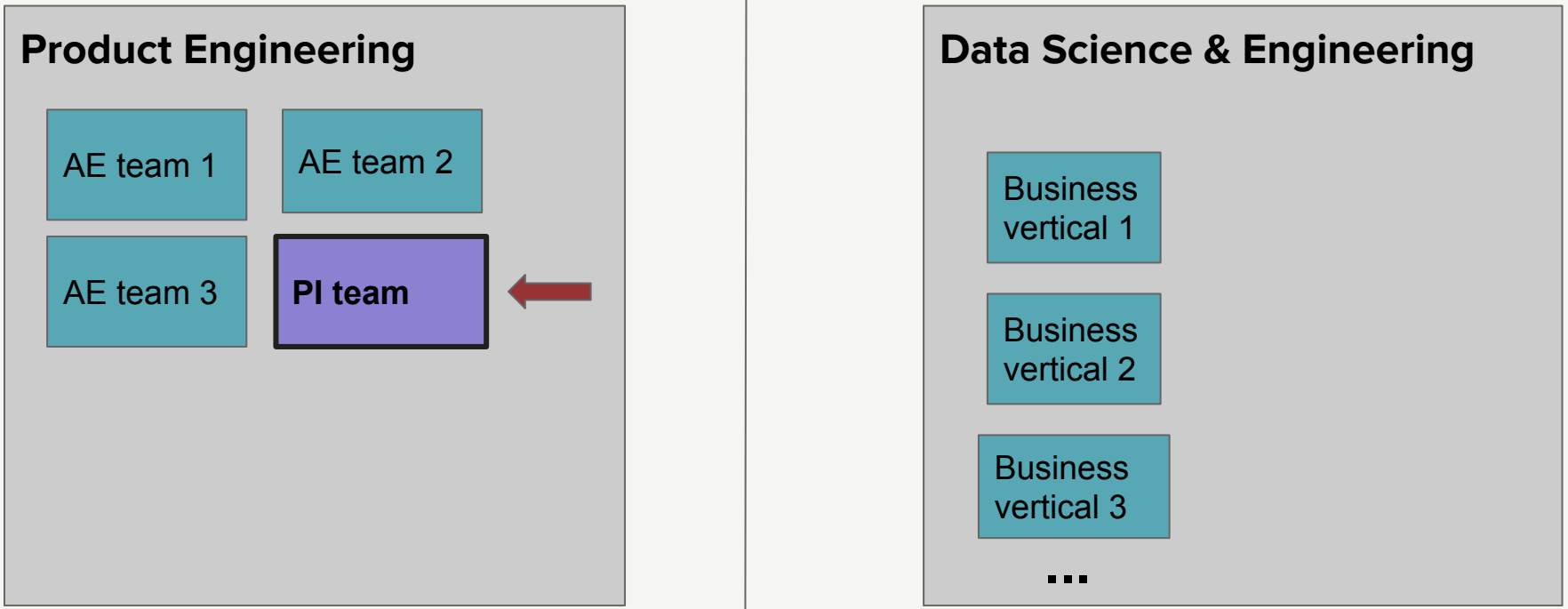
Algorithm Engineering pain points

- 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.*

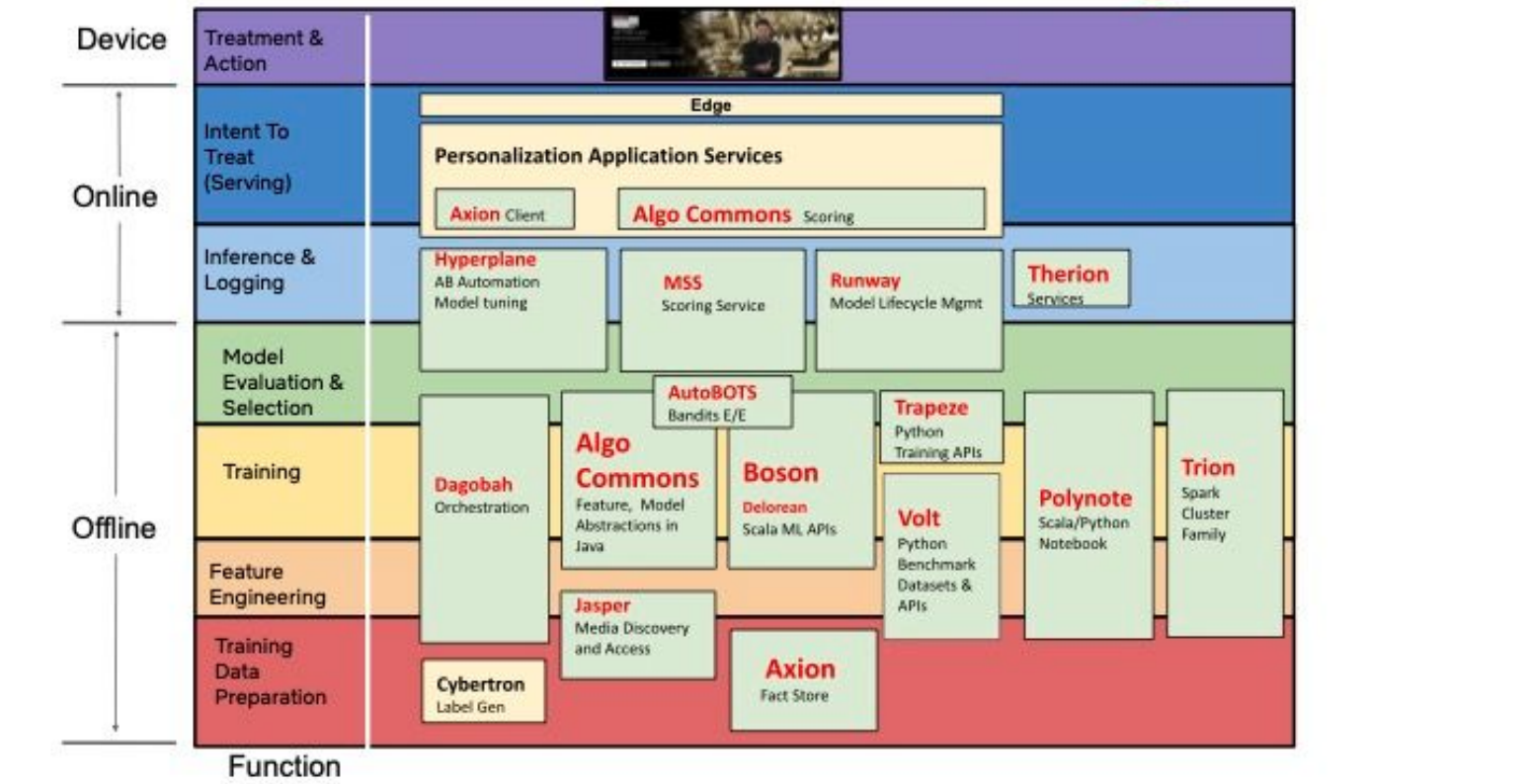
Where should the PI team go?



Algorithm Engineers

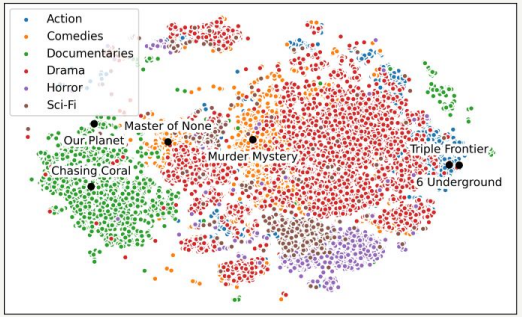
Data Scientists

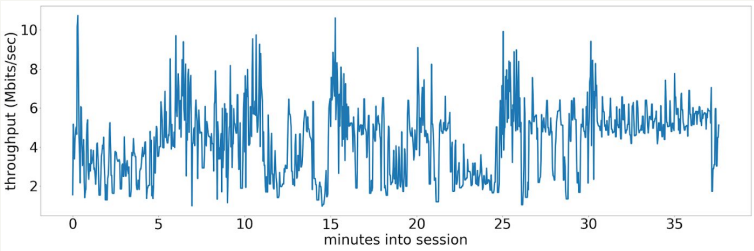
Personalisation Infrastructure stack



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

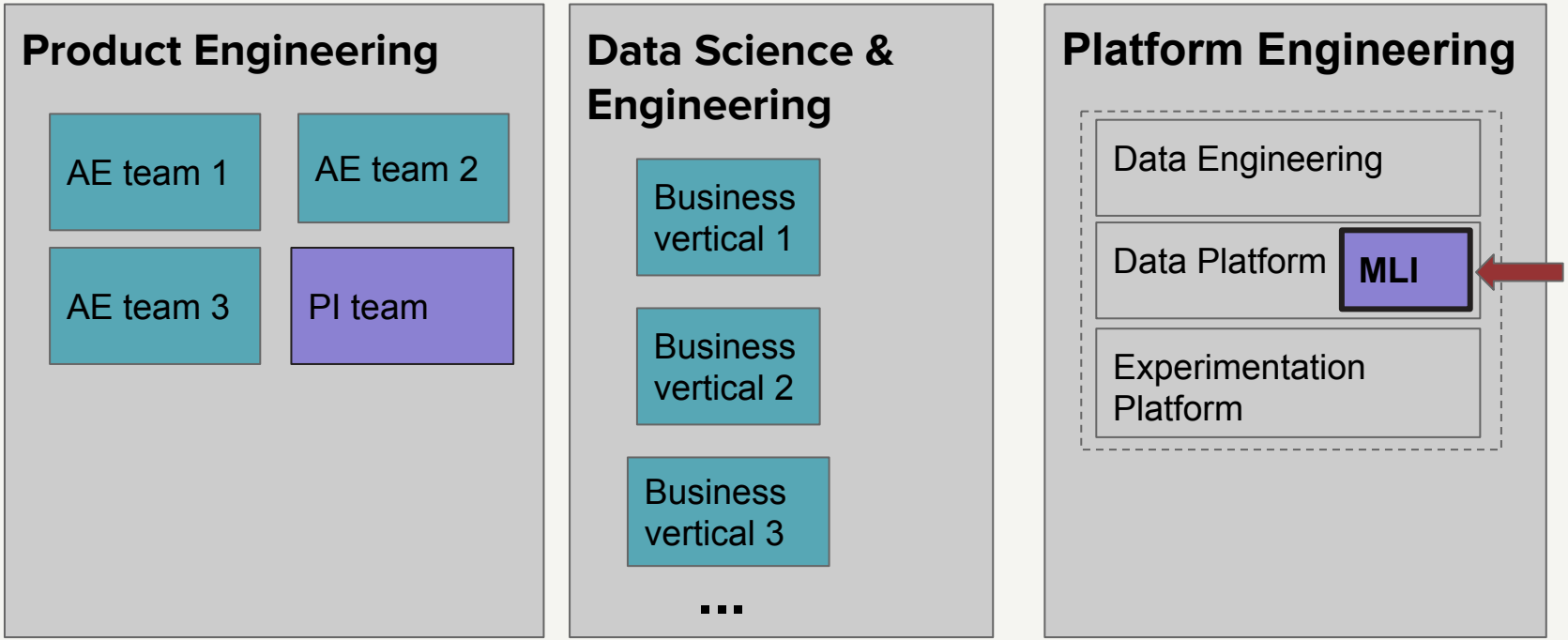




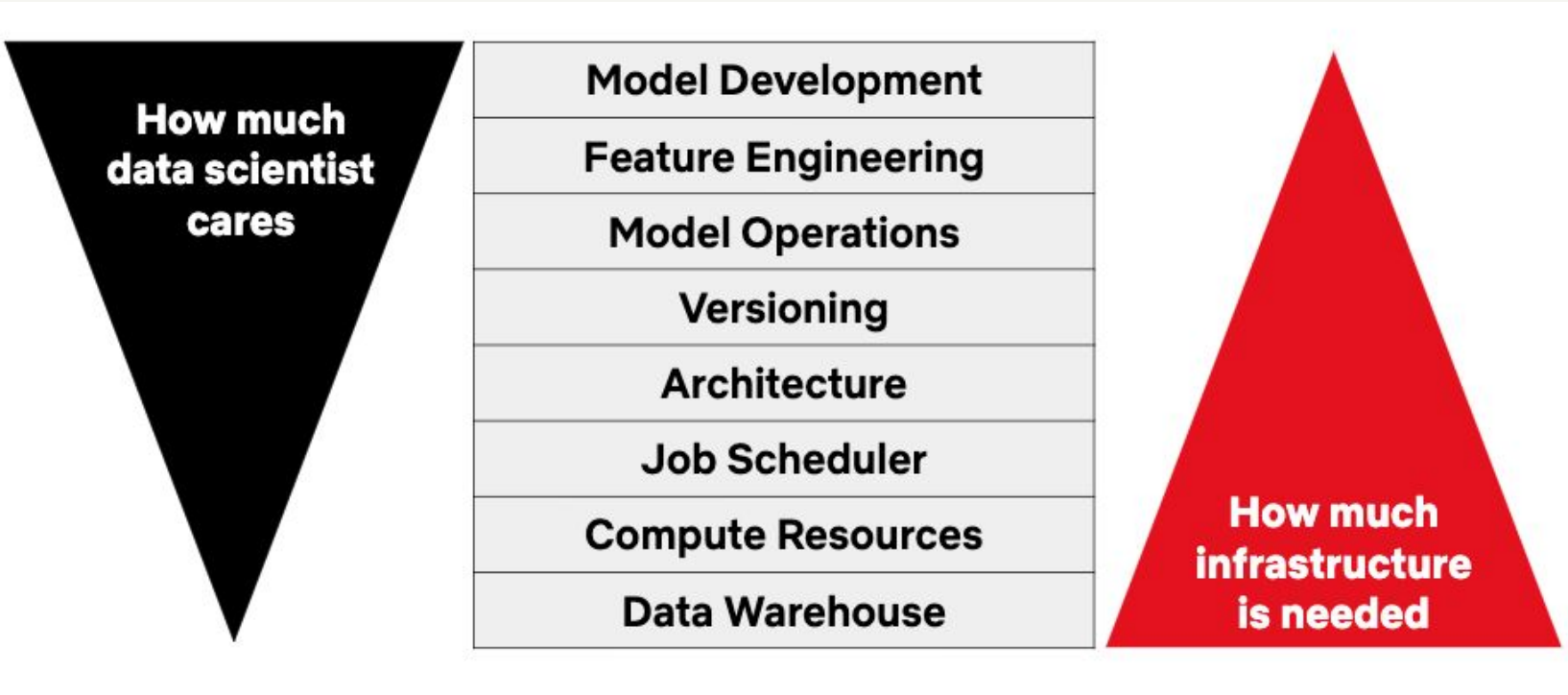
Machine Learning Infrastructure (MLI) team mission

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

Where should the MLI team go?



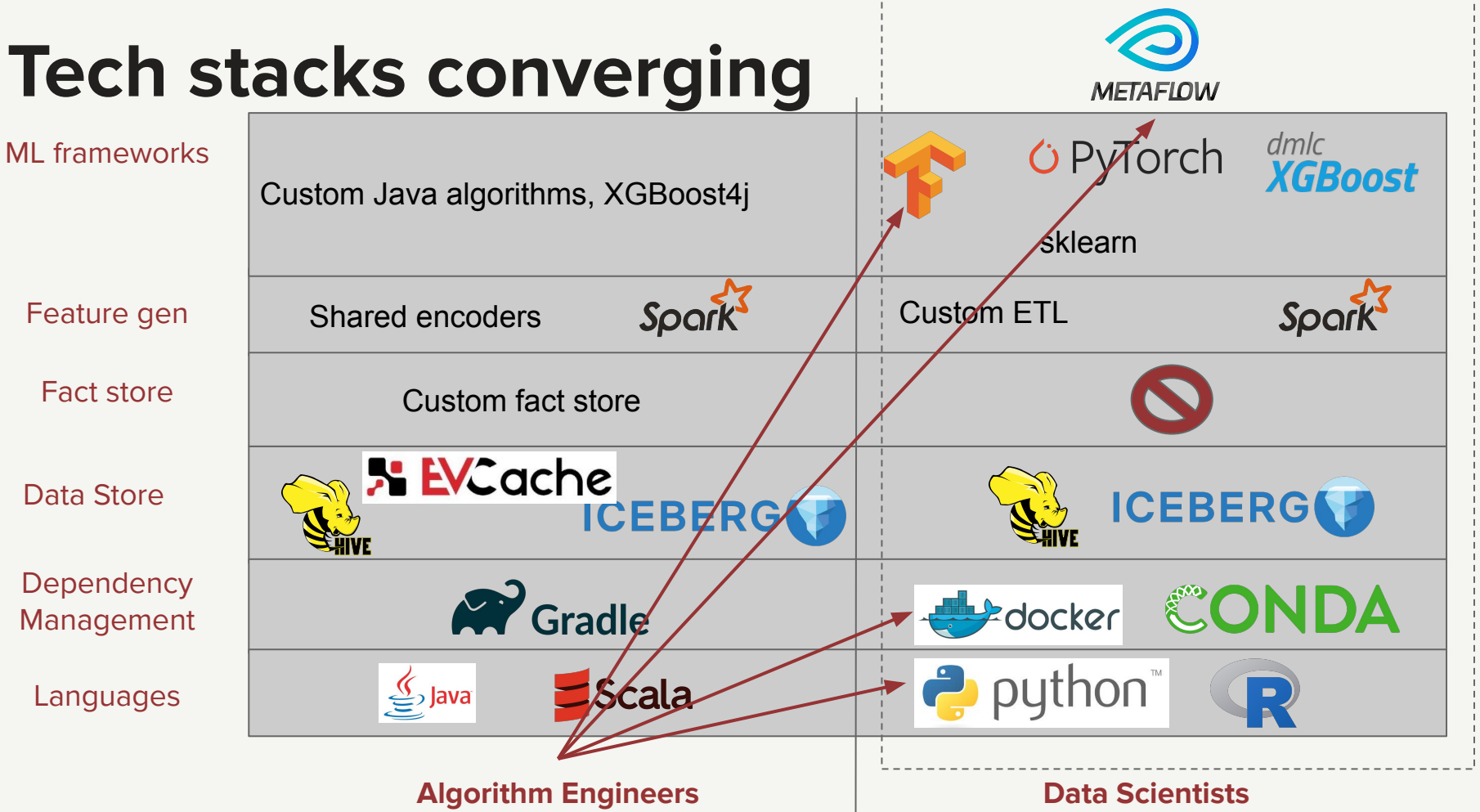
MLI built Metaflow for Data Scientists



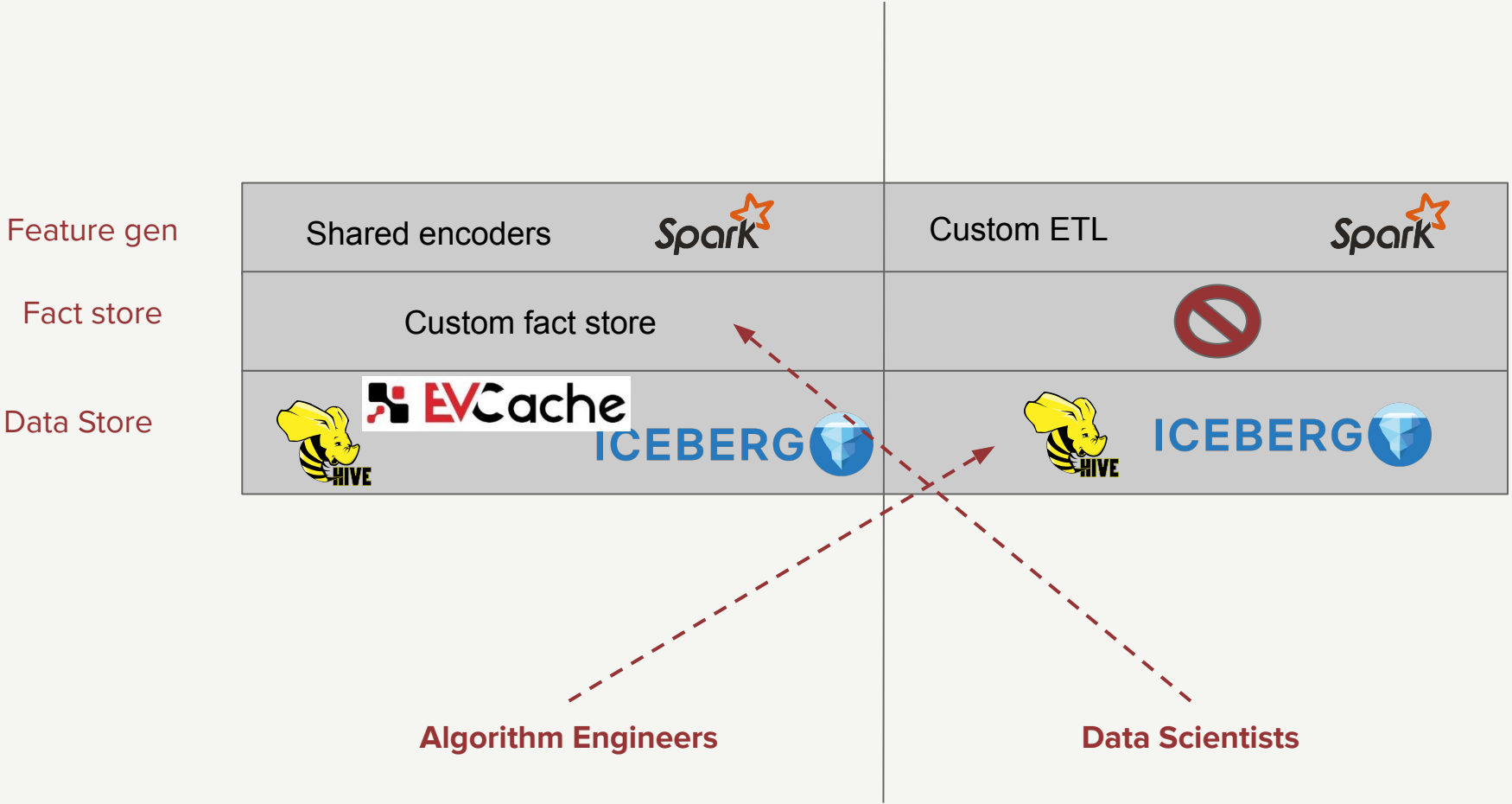
MLI built Metaflow for Data Scientists

Merging into ML Platform

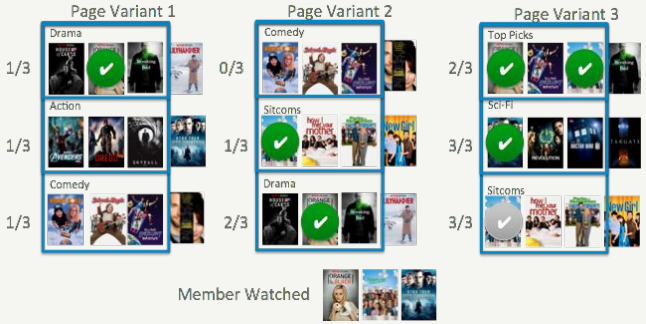
Tech stacks converging

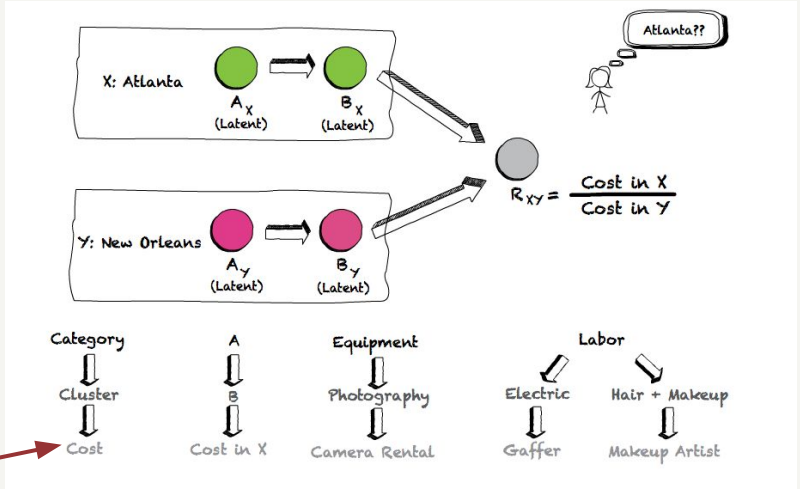


Opportunity to share facts & features



Increasing need for collaboration

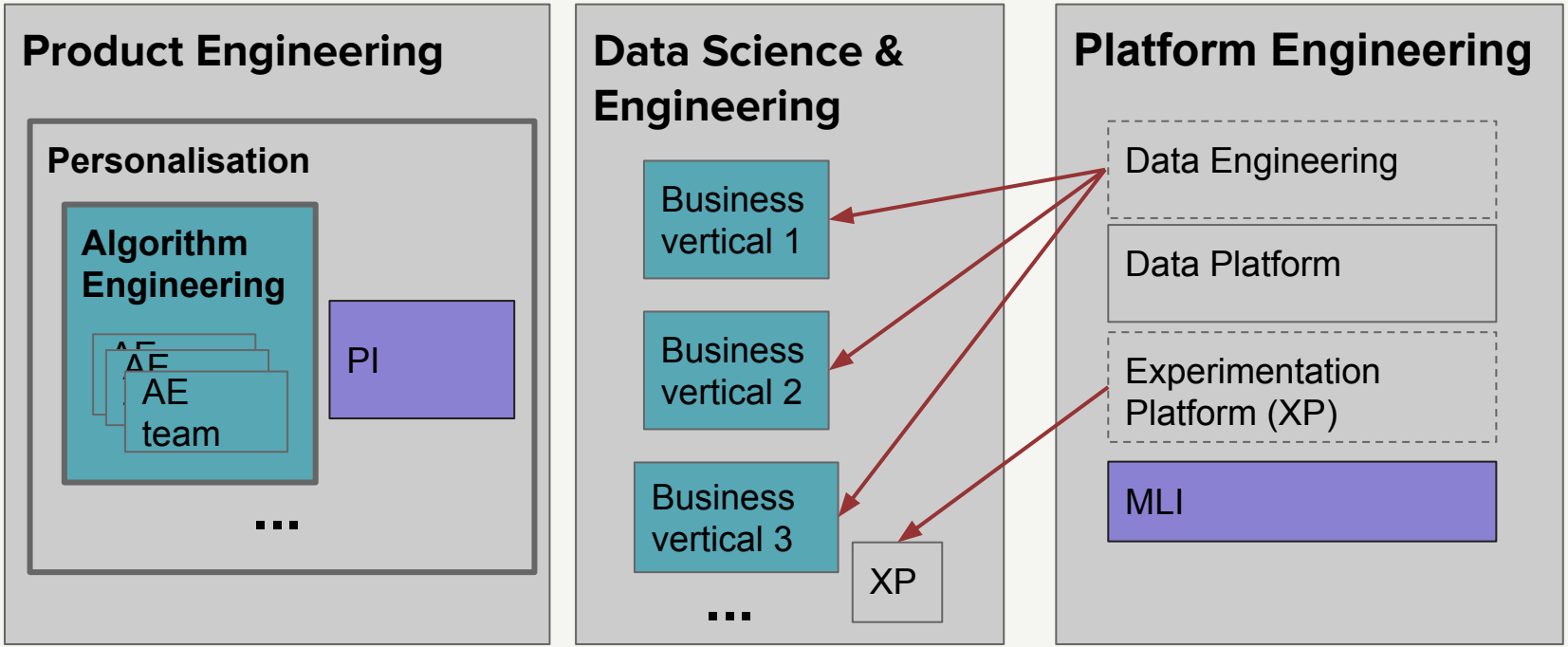




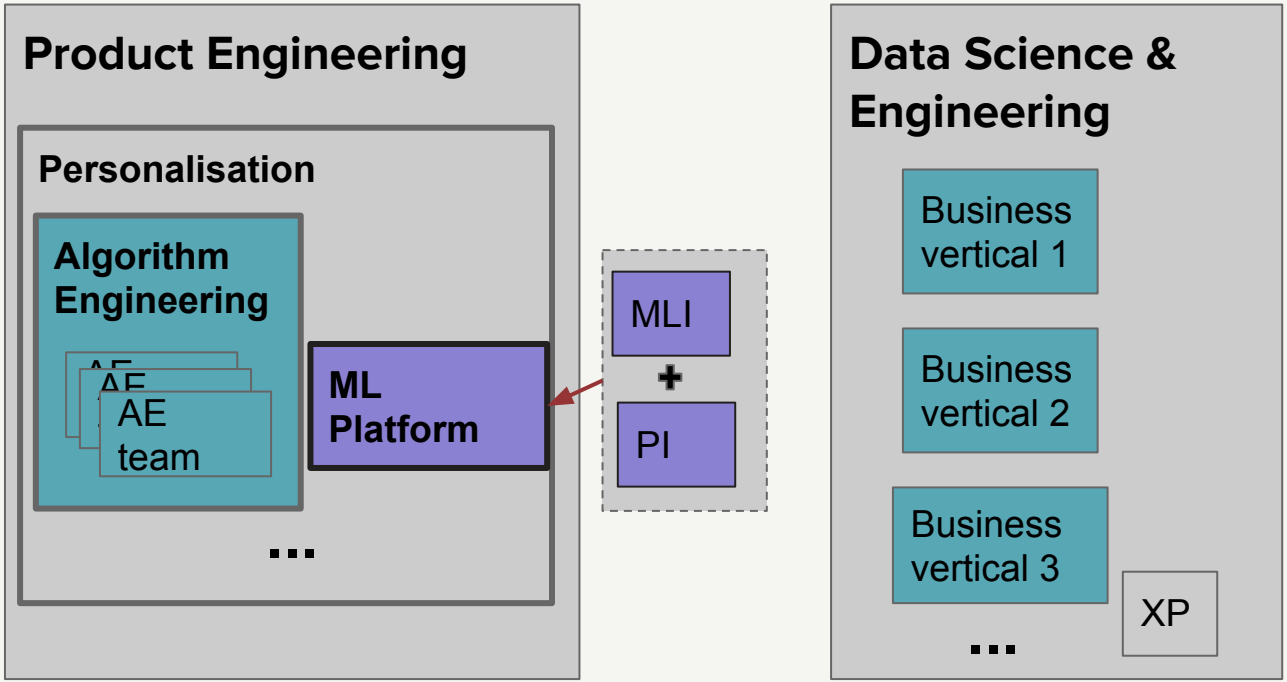
Algorithm Engineers working in more areas of the business

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