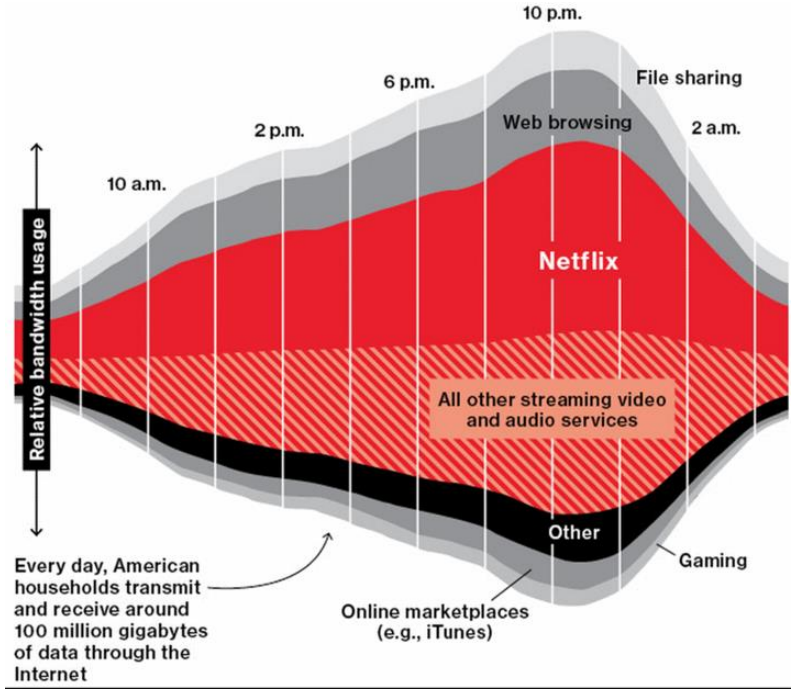
# Data Science + Content

Todd Holloway, Director of Content Science & Algorithms
*for* Smart Content Summit, 3/9/2017

NETFLIX

NETFLIX

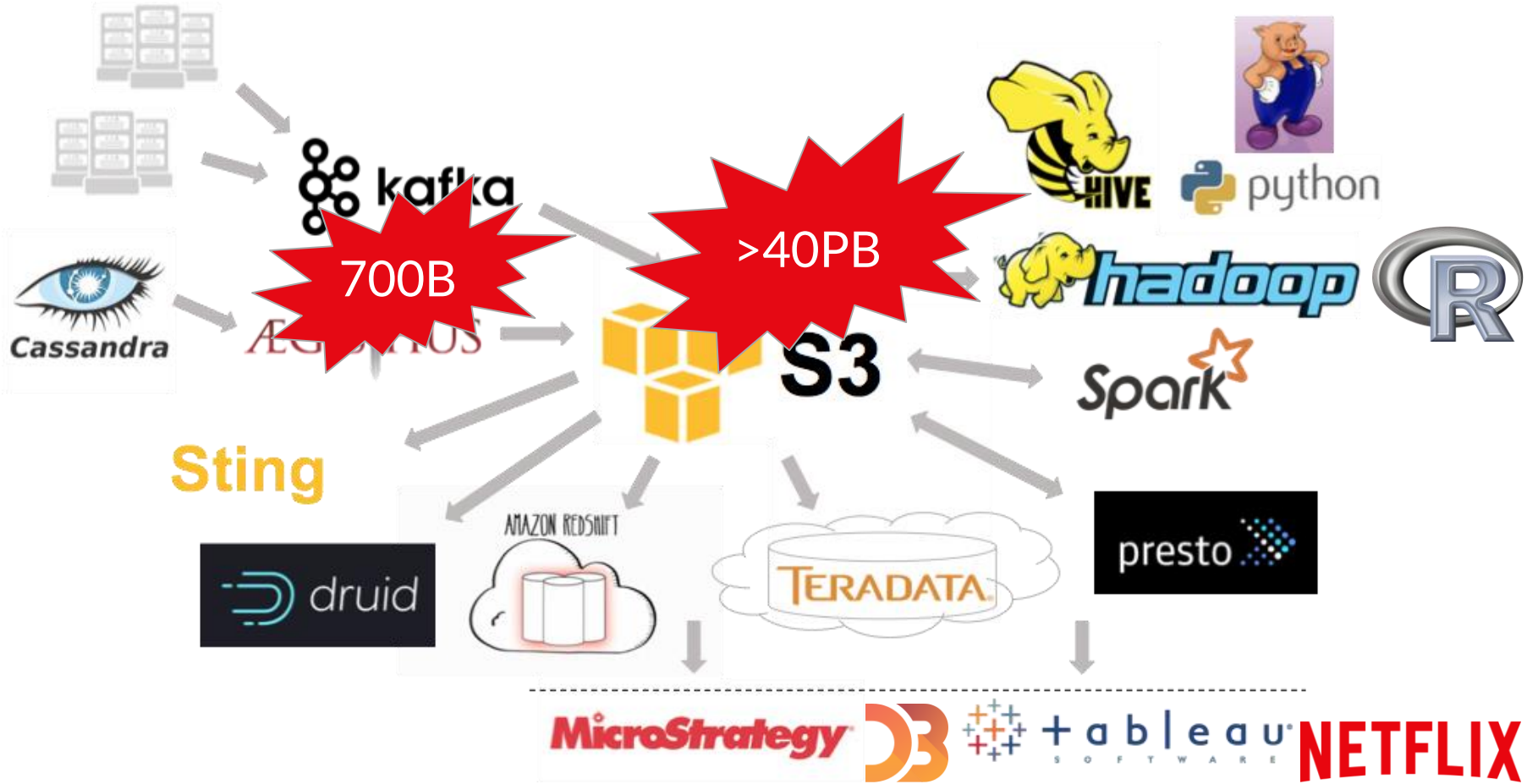
# Netflix *by the Numbers*...




- > 90M members
- Available worldwide (except China)
- > 1000 device types
- Hours: > 3B per month
- Log 100B events/day
- 36.5% of peak US downstream traffic

NETFLIX

# Data @ Netflix



# Evolution of Netflix

DVD service

Streaming service

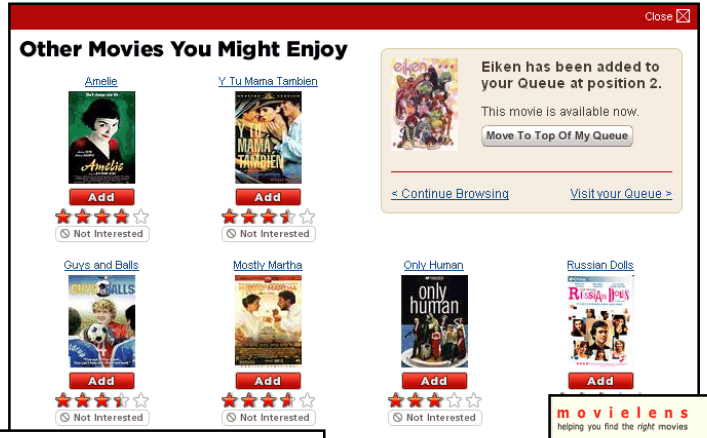
Originals

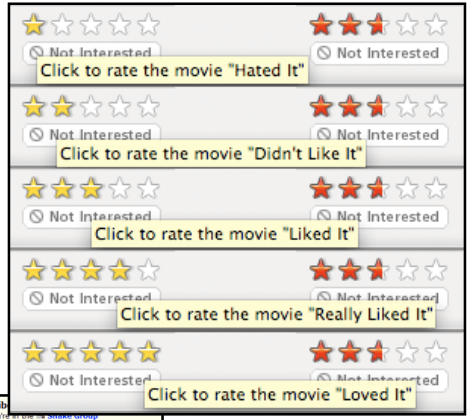
Going Global

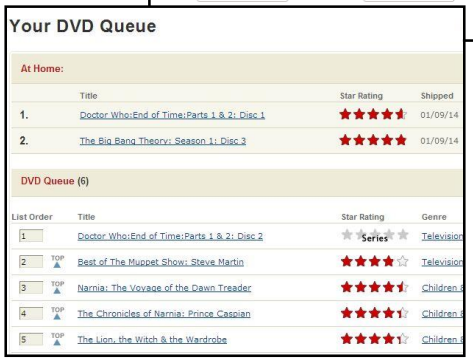
1999 — 2007 — 2012 — 2016

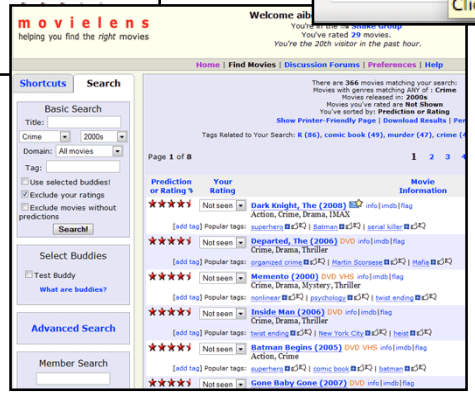
NETFLIX

# Evolution of Machine Learning @ Netflix (2000)



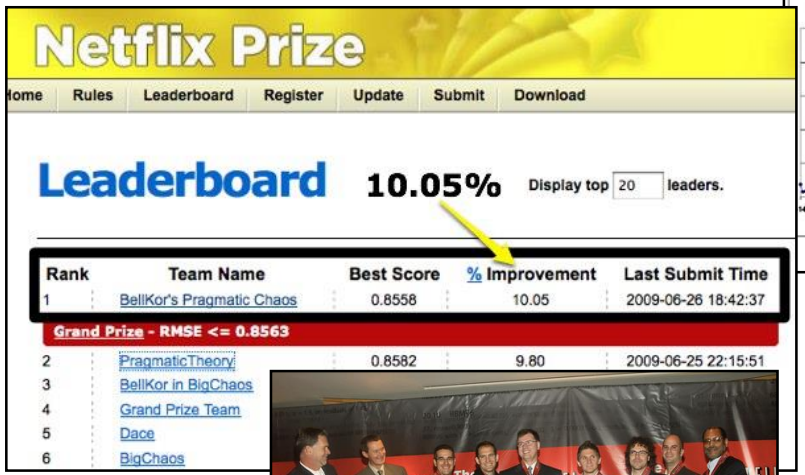




http://classic.movielens.org/

NETFLIX

# Evolution of Machine Learning @ Netflix (2007-09)

https://en.wikipedia.org/wiki/Netflix_Prize



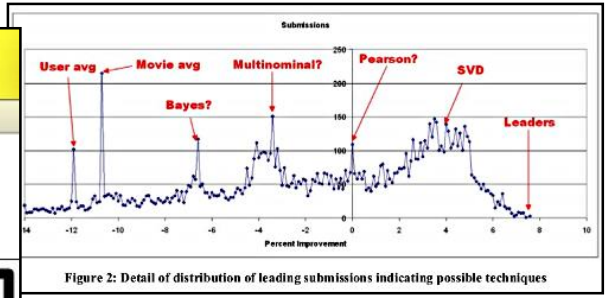
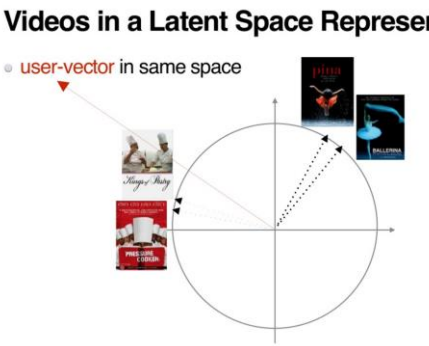
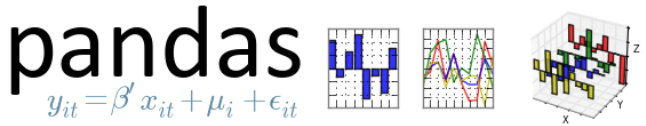



Figure 2: Detail of distribution of leading submissions indicating possible techniques




NETFLIX

# Machine Learning & Data Science Tools (2017)




NETFLIX

# Machine Learning Algorithms (2017)

- Regression models (Logistic, Linear, Elastic nets)
- GBDT/RF
- SVD & other MF models
- Factorization Machines
- Clustering (from k-means to HDP)
- Deep Learning
- LDA
- Association Rules

NETFLIX

ML plays a role in every phase of the content lifecycle.

1. Acquisition
2. Quality control
3. Localization
4. Marketing
5. Streaming
6. Presentation to users

Let’s work backward in the lifecycle…

NETFLIX

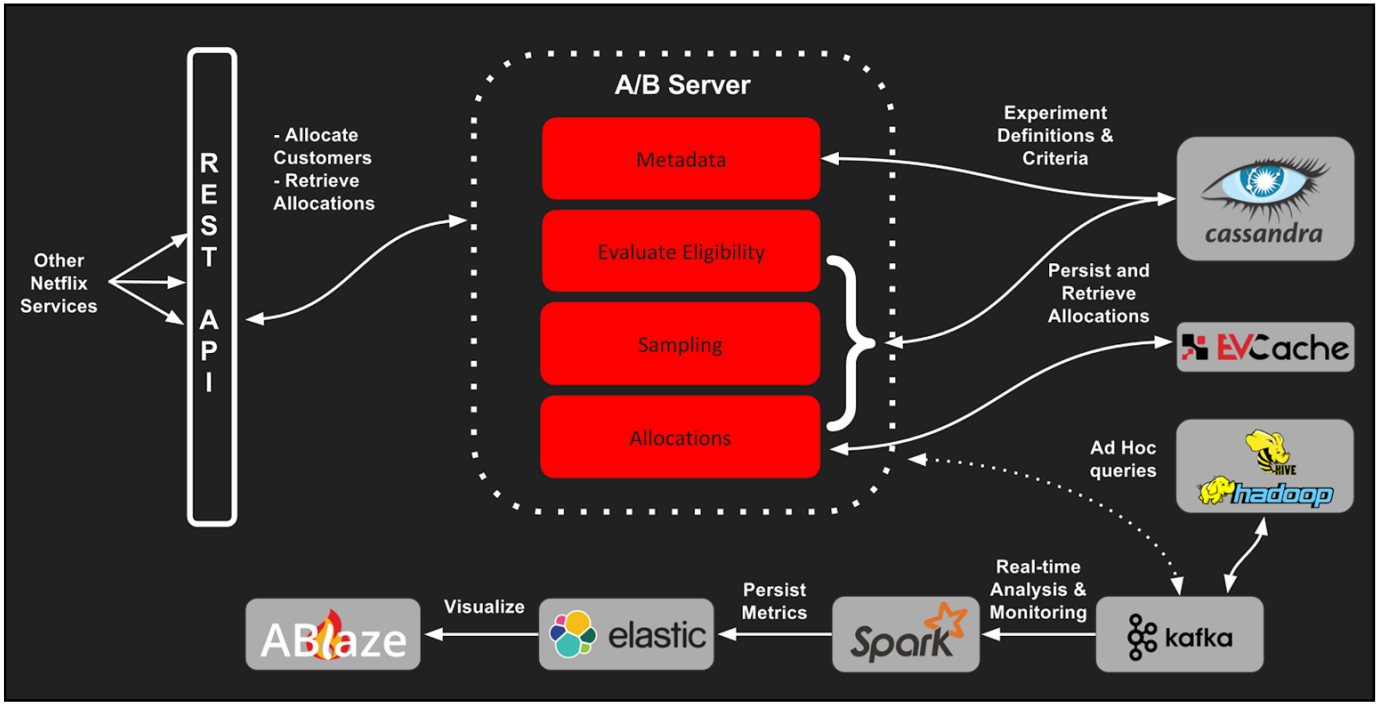
# User Interface

Promotion
Layout
Imagery
Metadata
Search Functionality
Row / Content Selection
… and more

http://delivery.acm.org/10.1145/2850000/2843948/a13-gomez-uribe.pdf

NETFLIX

# User Interface Experimentation Platform




http://techblog.netflix.com/2016/04/its-all-about-testing-netflix.html

NETFLIX

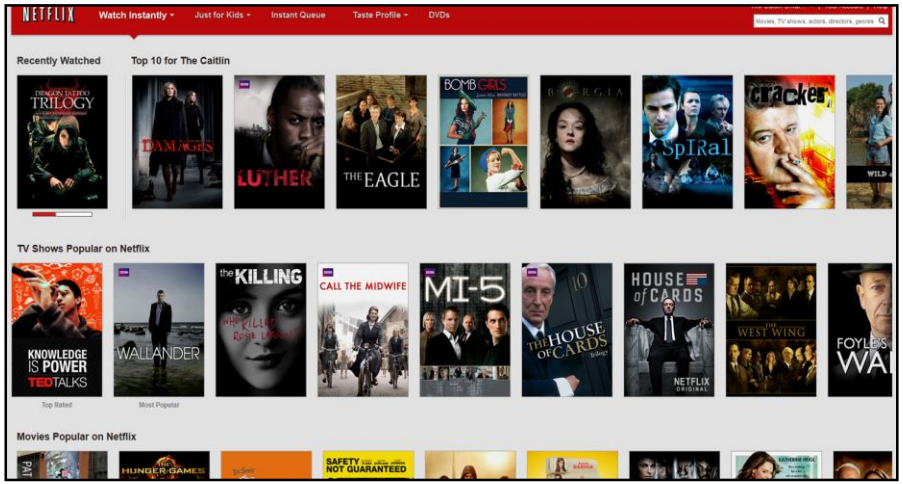
# User Interface Experimentation Platform

## Cover Art Optimization

http://techblog.netflix.com/2016/05/selecting-best-artwork-for-videos.html

NETFLIX

# User Interface Experimentation Platform




2013

2015

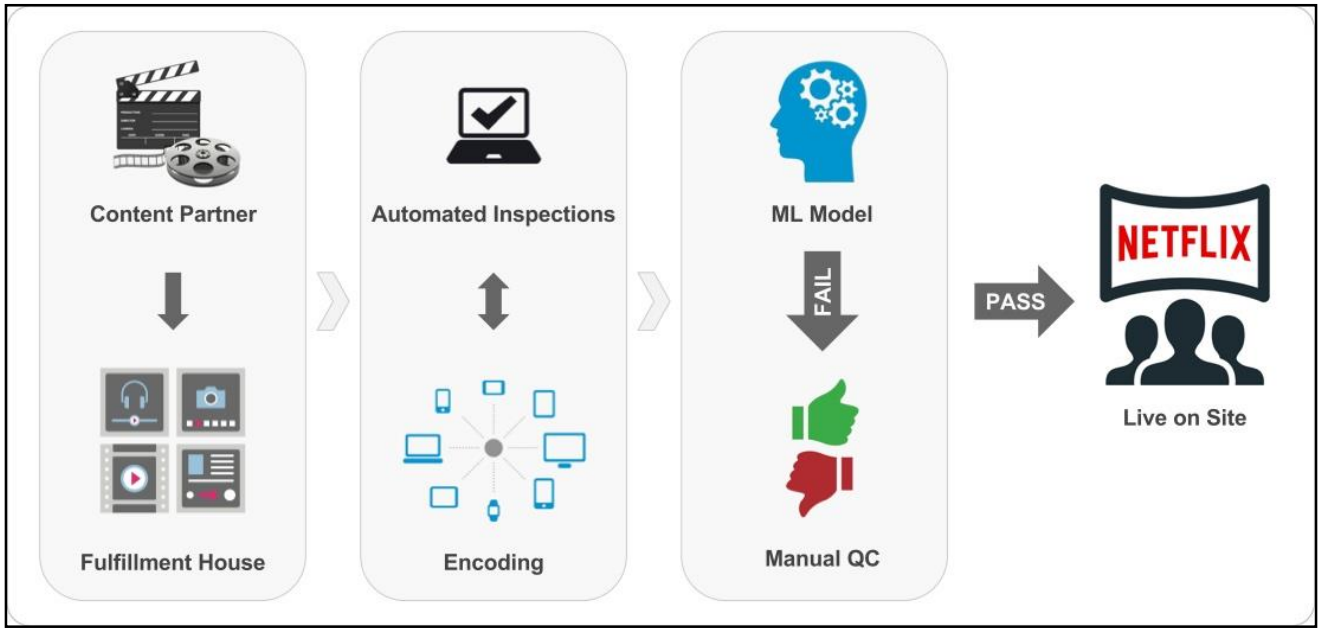
NETFLIX

# Streaming Optimization

https://media.netflix.com/en/company-blog/how-netflix-works-with-isps-around-the-globe-to-deliver-a-great-viewing-experience

NETFLIX

# Digital Asset Quality Control




http://techblog.netflix.com/2015/12/optimizing-content-quality-control-at-netflix-predictive-modeling.html

NETFLIX

# Data Science to Aid Selection of Content

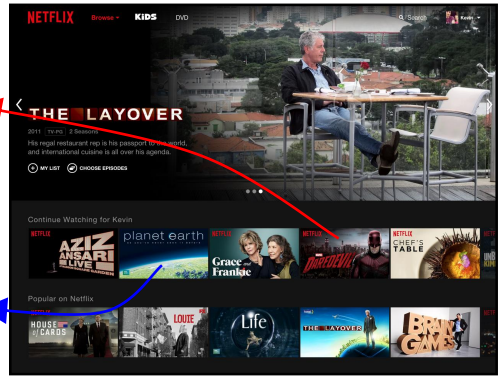
Scripts and pitches

Studio productions

**Selection process**

Original content

Licensed content



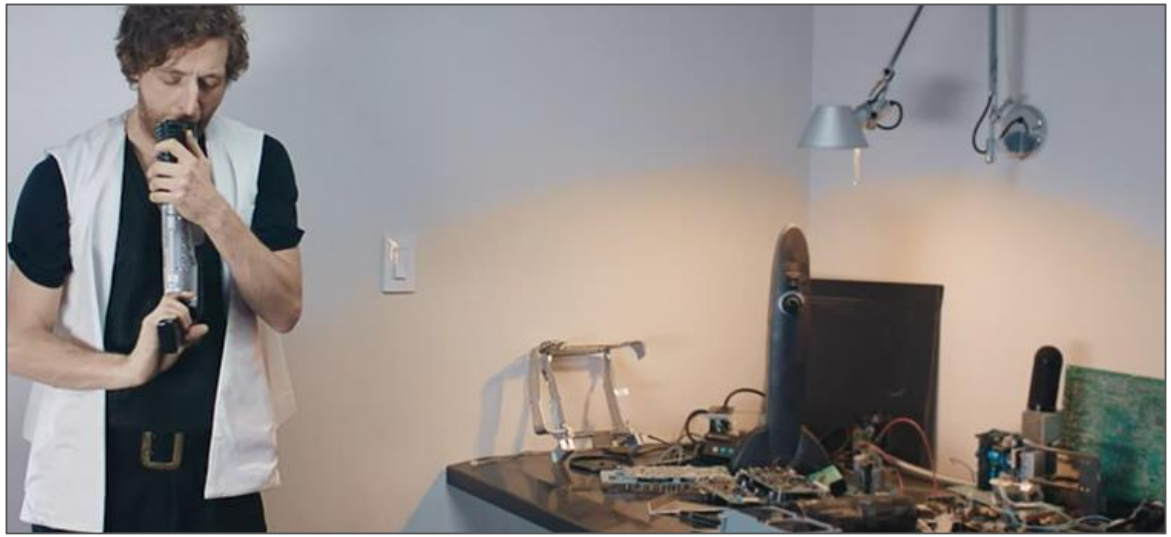

NETFLIX

# Can data science create content?

Can a computer write a script that would win a competition? [benjamin.wtf]

NETFLIX

# Can data science create content?



He is standing in the stars and sitting on the floor. He takes a seat on the counter and pulls the camera over to his back. He stares at it. He is on the phone. He cuts the shotgun from the edge of the room and puts it in his mouth. He sees a black hole in the floor leading to the man on the roof.

NETFLIX

Is Netflix doing this?

**No**.
It’s the opposite.
We give creative freedom to the creatives.

NETFLIX

But can data help in choosing content?

**Yes.**

All decisions are made by experienced creatives, but analytic products can help.

NETFLIX

# Netflix's Notion of Value

**Content Efficiency** = value / cost

Netflix's quarterly shareholder letter, released overnight, contained lots of new information. But one section in particular stood out:

> Looking at our original content performance over the last few years, there have been so many impressive aspects...
>
> But there is one real shocker; **last year our original content overall was some of our most efficient content**. Our originals cost us less money, relative to our viewing metrics, than most of our licensed content, much of which is well known and created by the top studios.

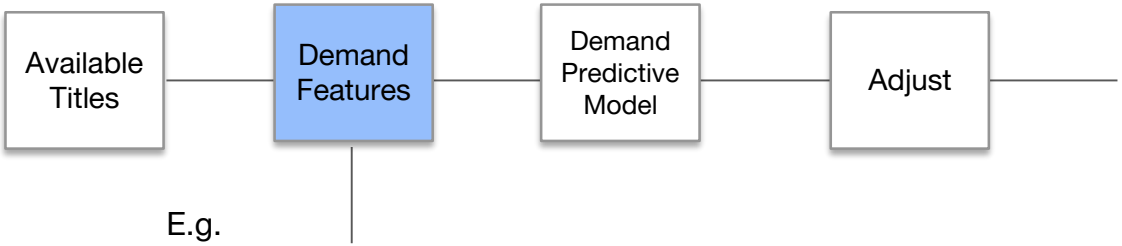
NETFLIX

# Using Machine Learning to Predict Value

Available Titles → Demand Features → Demand Predictive Model → Adjust → If efficient, proceed

NETFLIX

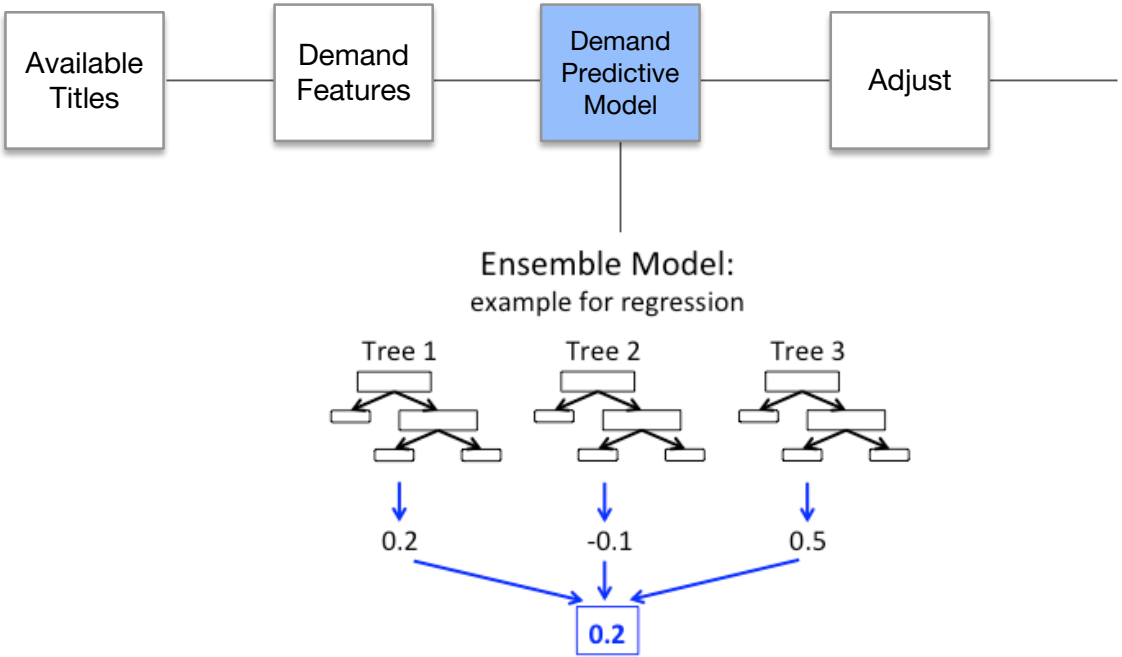
# Using Machine Learning to Predict Value




- Past performance on Netflix (if previously licensed)
- Past performance of similar titles on Netflix
- Broadcast ratings
- Theatre ticket sales
- Talent involved
- Reviews
- Awards

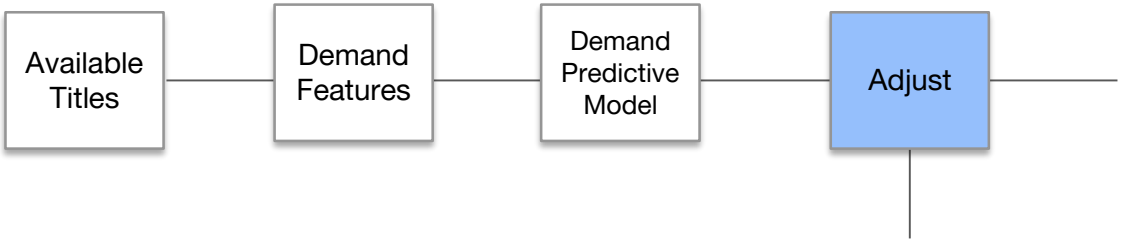
NETFLIX

# Using Machine Learning to Predict Value




NETFLIX

# Using Machine Learning to Predict Value

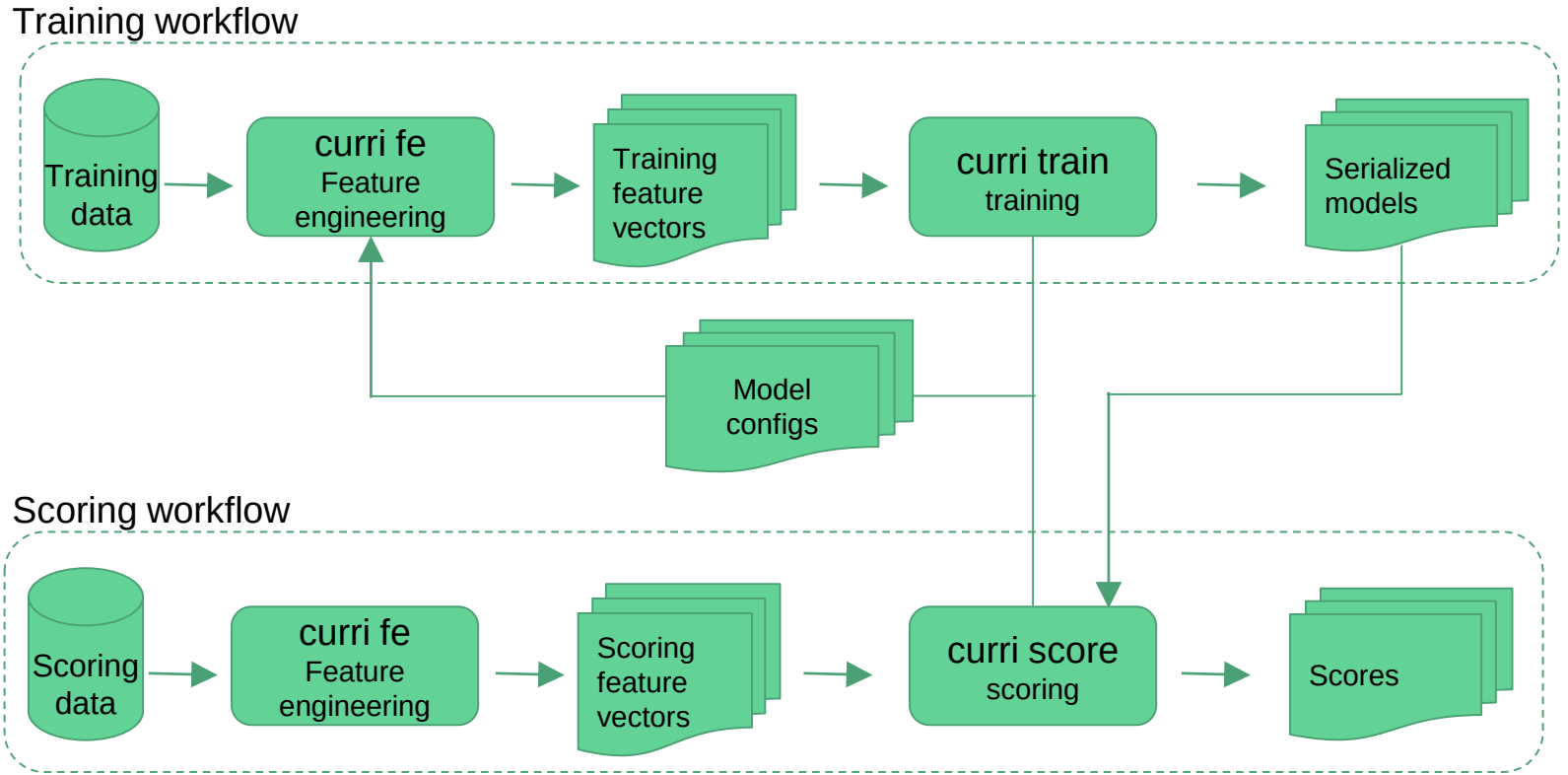



E.g.

- Buyer judgements
- Deal term adjustments

NETFLIX

# Custom Machine Learning Framework




NETFLIX

# Using Machine Learning for Originals

Originals is a more difficult problem than licensing

- Less data - no box office or reviews
- Moving target - ideas and scripts can evolve
- Fungible - execution varies with talent and budget

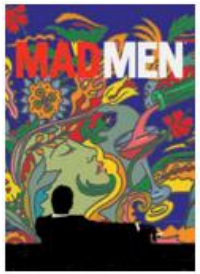
NETFLIX

# Finding Comparable Titles

‘House of Cards’ script =

X


Those members also watch:



NETFLIX

What's like **Twister** meets **Shark Week?**

NETFLIX

# Programming to Tastes

(Something for Everyone)

NETFLIX

# Programming to Tastes

NETFLIX

# Data Science and Tech are in the DNA of Netflix

**(and we'll keep looking for ways to leverage that DNA for content)**

Thank You.  Questions?

NETFLIX