PETER SKOMOROCH
Head of Data Products,
Workday

#DOMINOREV

Product Management for AI

Product Management for AI

Peter Skomoroch - @peteskomoroch
Rev 2 Data Science Leaders Summit, NYC - May 24, 2019

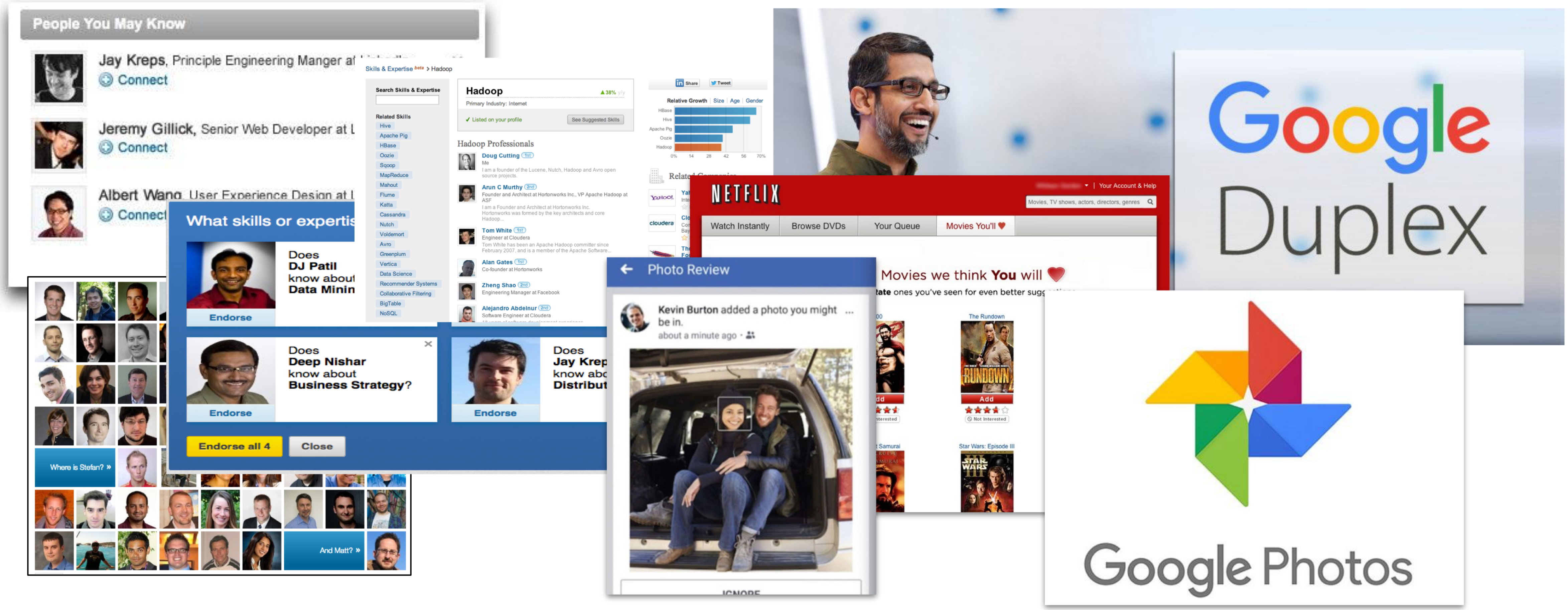
Background: Machine Learning & Data Products

Peter Skomoroch
@peteskomoroch

- Co-Founder and CEO of SkipFlag, Enterprise AI startup acquired in 2018 by Workday
- 18+ years building machine learning products
- Principal Data Scientist, ran Data Products team at LinkedIn. ML & Search at MIT, AOL, ProfitLogic
- Co-Host of O'Reilly AI Bots Podcast, Startup Advisor

AI Products

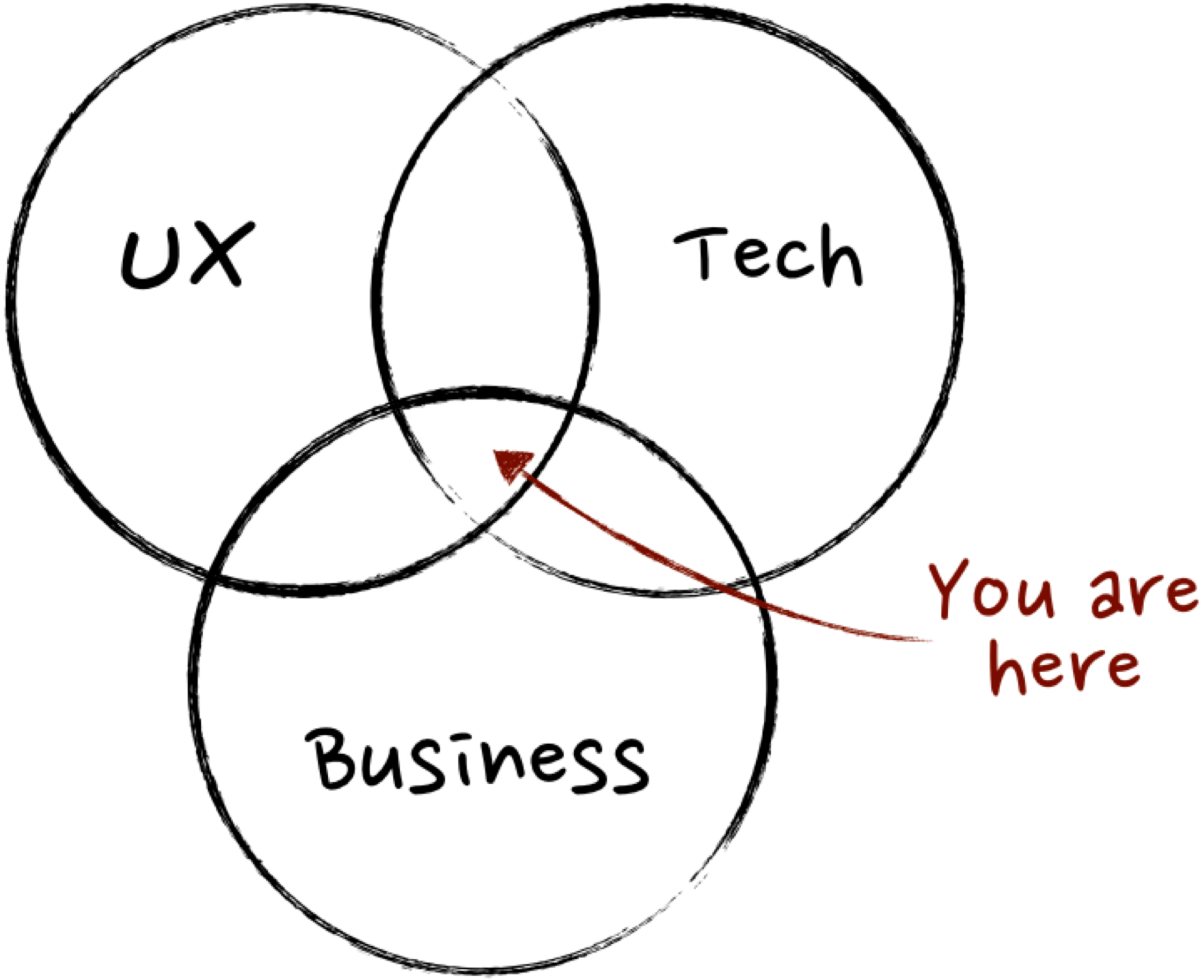
Automated systems that collect and learn from data to make user facing decisions with machine learning



Machine Learning Projects are Hard

- The transition to machine learning will be about 100x harder than the transition to mobile apps
- Some of the biggest challenges are organizational, not technical
- Data driven companies like Google and Facebook have a strategic advantage building ML products based on their data & compute assets, large user population, tracking & instrumentation, and AI talent

The Role of an AI Product Manager



- An AI Product Manager (PM) has core product skills (strategy, roadmaps, prioritization, etc.) along with an intuitive grasp of ML
- They help identify and prioritize the highest value applications for machine learning and do what it takes to make them successful

Image source: Martin Eriksson https://www.mindtheproduct.com/2011/10/what-exactly-is-a-product-manager/

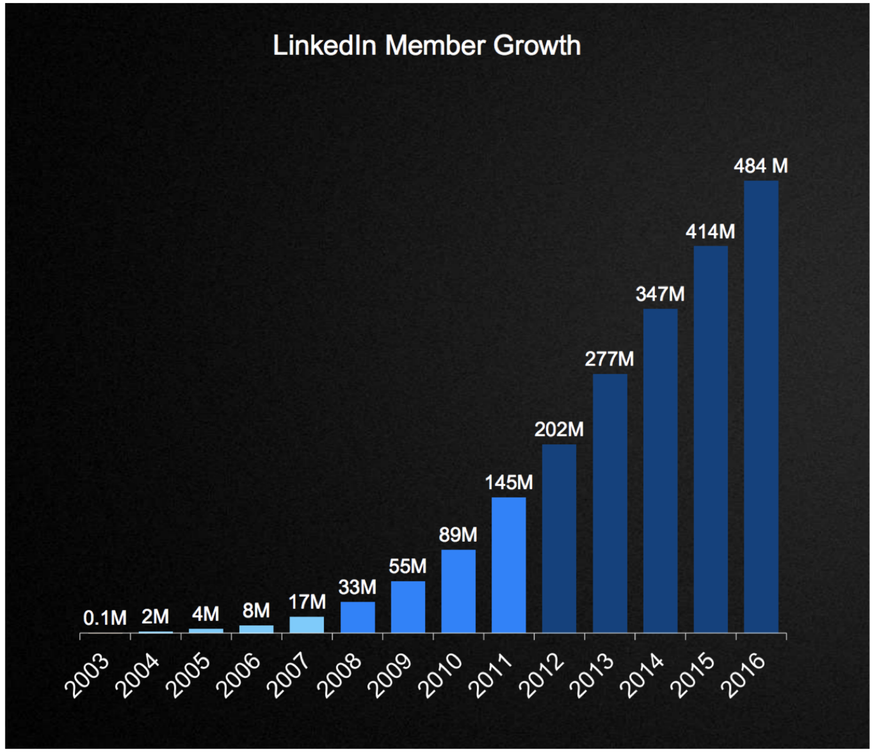
Good AI Product Managers Have Data Expertise

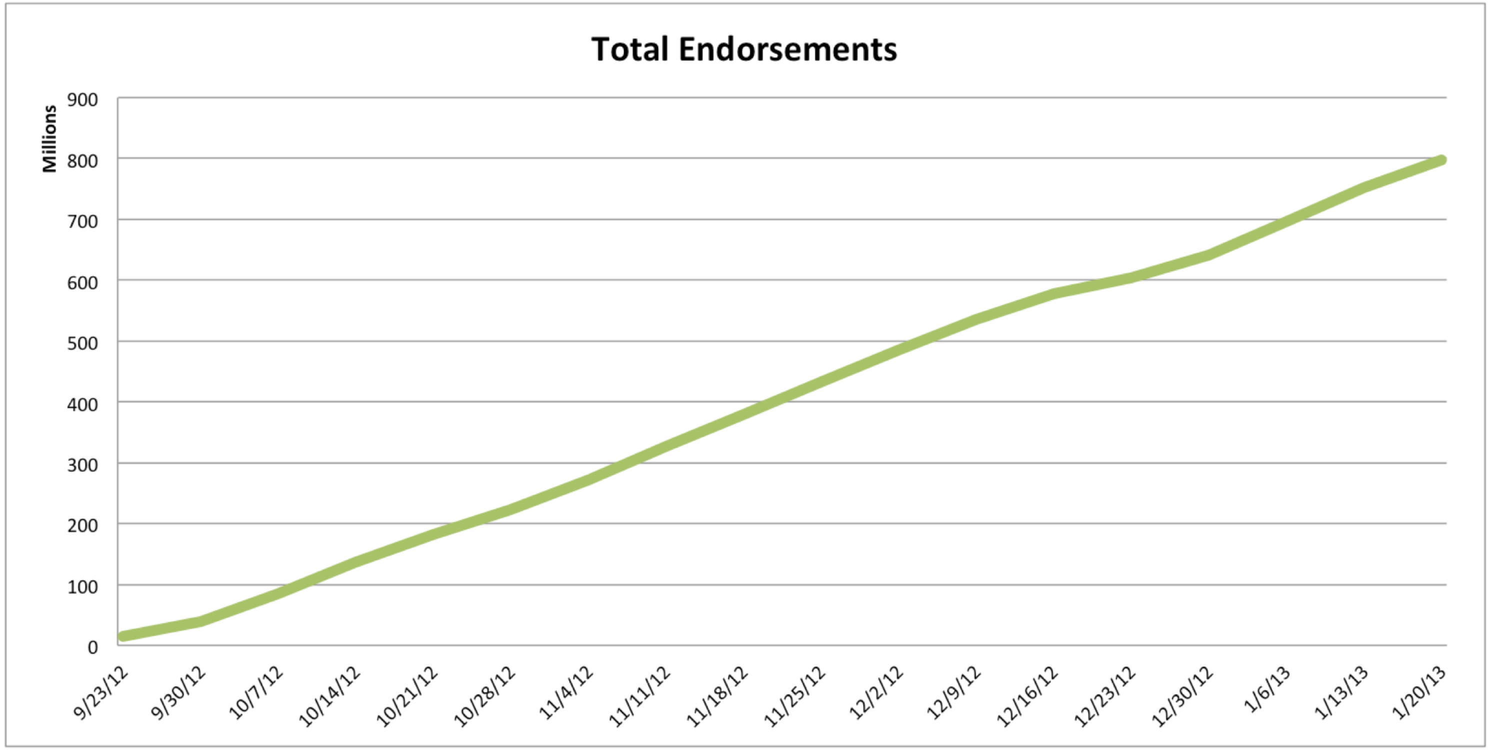
- Know the difference between easy, hard, and impossible machine learning problems
- Even if something is feasible from a machine learning perspective, the level of effort may not justify building the feature
- Know your company’s data inside and out including quality issues, limitations, biases, and gaps that need to be addressed
- Develop an intuitive understanding of your company’s data and how it can be used to solve customer problems

How to evaluate and prioritize your AI projects

- Start with your mission and strategic objectives, and select projects that align well with those goals
- LinkedIn mission: *“Connect the world's professionals to make them more productive and successful”*
- Example strategy: *“To be the professional profile of record”*
- Get everyone in a room, group project ideas by theme and make “T-shirt sized” estimates (L/M/S) of impact and difficulty for each idea.
- Rank and prioritize projects by ROI where possible

Apply ML to a Metric the Business Cares About





ML Adds Uncertainty to Product Roadmaps

- PMs are often uncomfortable with expensive ideas that have an uncertain probability of success
- Many organizations will struggle to justify the expense of projects that require significant research investment upfront
- Some ML products may need to be split into time boxed projects that get to market in a shorter time frame
- What can you productize now vs. much later on?
- Keep track of dependencies on other teams and have a “Plan B”

Experimental Culture

- Machine Learning shifts engineering from a deterministic process to a probabilistic one
- Take intelligent risks
- Most successful ML products are experiments at massive scale
- Companies driven by analytics and experimental insights are more likely to succeed

“

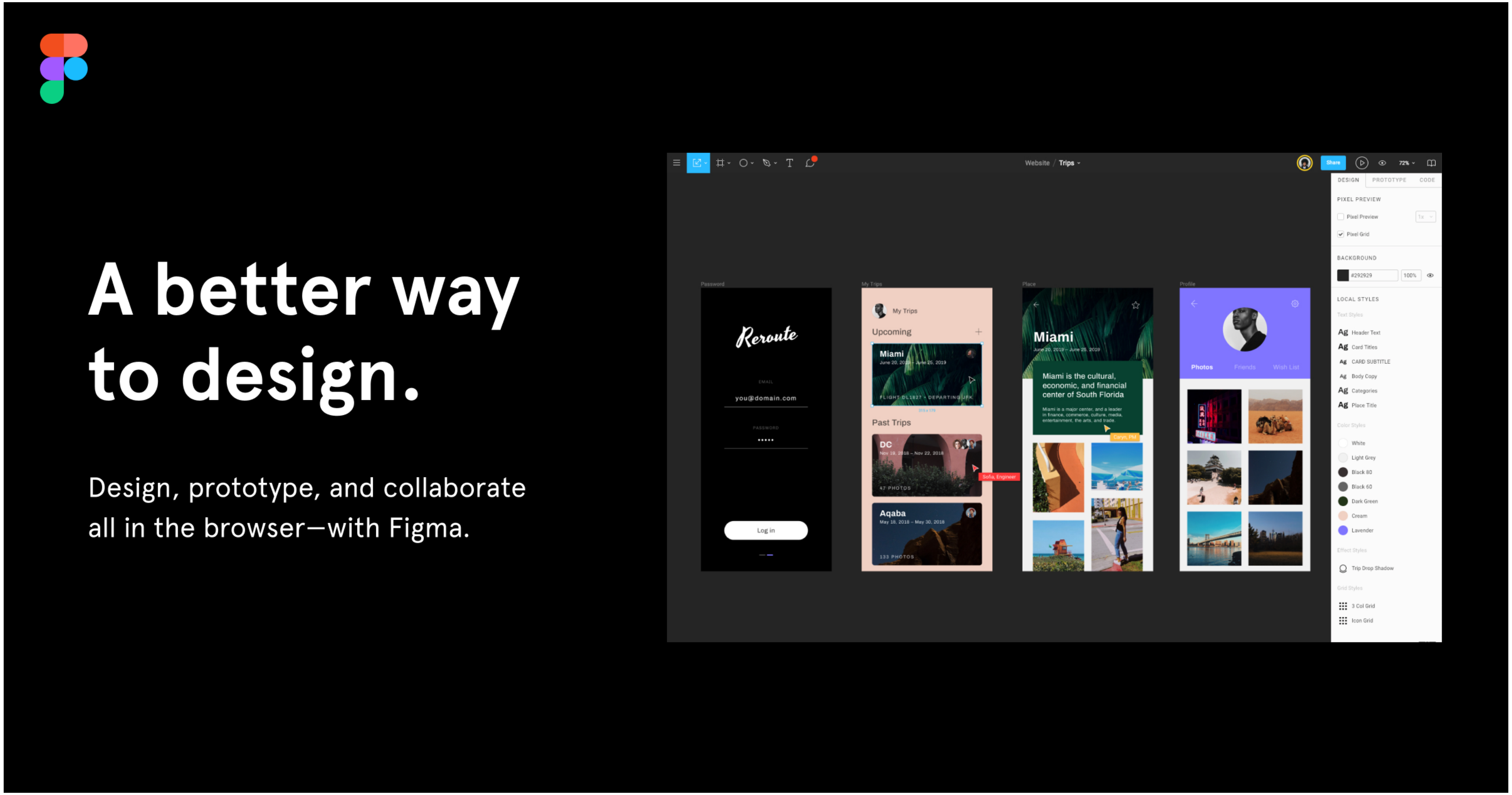
If you only do things where you know the answer in advance, your company goes away.

Jeff Bezos
Founder, Chairman & CEO of Amazon.com

ML Product Development Process

1. Verify you are solving the right **problem**
2. Theory + **model design** (in parallel with UI design)
3. **Data collection**, labelling, and cleaning
4. **Feature engineering**, model training, offline validation
5. **Model deployment**, monitoring & large scale training
    - Iterate: repeat process, refine live model & improve
    - 80% of effort and gains come from iterations after shipping v 1.0
    - Use derived data from the system to build new products

Bridging the worlds of design and machine learning



Data Quality & Standardization

- Guide user input when you can
- Use auto suggest fields
- Validate user inputs, emails
- Collect user tags, votes, ratings
- Track impressions, queries, clicks
- Sessionize logs
- Disambiguate and annotate entities (company names, locations, etc.)

“

Every single company I've worked at and talked to has the same problem without a single exception so far — poor data quality, especially tracking data

Ruslan Belkin
VP of Engineering, Salesforce.com

Testing Machine Learning Products

- Algorithm work that drags on without integration in the product where it can be seen and tested by real users is risky
- Ship a complete MVP in production ASAP, benchmark, and iterate
- Beware unintended consequences from seemingly small product changes
- Remember the prototype is not the product - see what happens when you use a more realistic data set or scale up your inputs
- Real world data changes over time, ensure your model tests and benchmarks keep up with changes in underlying data
- Machine learning systems tend to fail in unexpected ways

Look at Your Input Data & Prediction Errors

Suggested Skills

Enter a member id or name to get skills suggestions

Search

Random member

Peter Skomoroch

Principal Data Scientist at LinkedIn

Production

Feedback

Suggestion	Score	PeopleRank	
Machine Learning	1.000	0.889	cs_20120106
Hadoop	0.631	0.847	cs_20120106
Data Mining	0.500	0.882	cs_20120106
R	0.431	0.892	cs_20120106
Natural Language Processing	0.387	0.761	cs_20120106
MapReduce	0.356	0.861	cs_20120106
Information Retrieval	0.333	0.803	cs_20120106

Explicit Skills
Text Classification
Web Scraping
Sentiment Analysis
Biodefense
-

Flywheel Effects & AI Products

The technology developed for Amazon's family of voice-activated devices, including the Echo Spot, spurred a larger AI renaissance at the company. IAN C. BATES

BACKCHANNEL 02.01.18 02:53 PM

INSIDE AMAZON'S ARTIFICIAL INTELLIGENCE FLYWHEEL

How deep learning came to power Alexa, Amazon Web Services, and nearly every other division of the company.

BY STEVEN LEVY

- Users generate data as a side effect of using most software products
- That data in turn, can improve the product's algorithms and enable new types of recommendations, leading to more data
- These "Flywheels" get better the more customers use them leading to unique competitive moats
- This works well in platforms, networks or marketplaces where value compounds

* https://medium.freecodecamp.org/the-business-implications-of-machine-learning-11480b99184d

Final Thoughts

- Machine learning products are hard to build, but within reach of teams who invest in data infrastructure
- Some of the biggest challenges are organizational, not technical
- Good product leaders are a key factor in shipping successful ML products
- Find a machine learning application with a direct connection to a metric your organization values and ship it

Send me questions! **@peteskomoroch**