

# ANZ INSTITUTIONAL DATA SCIENCE

## ANZ Banking Group

- Top 4 banks in Australia
- Largest banking group in New Zealand

## Our division

- No. 1 Institutional Bank for market penetration in AU and NZ.
- Top 4 corporate bank in Asia
- Clients with revenue over 400 million.

## Our team

- Started in 2016 with 3 people
- 20 people from 10 different countries

## Best Industry Skills Full stack developer

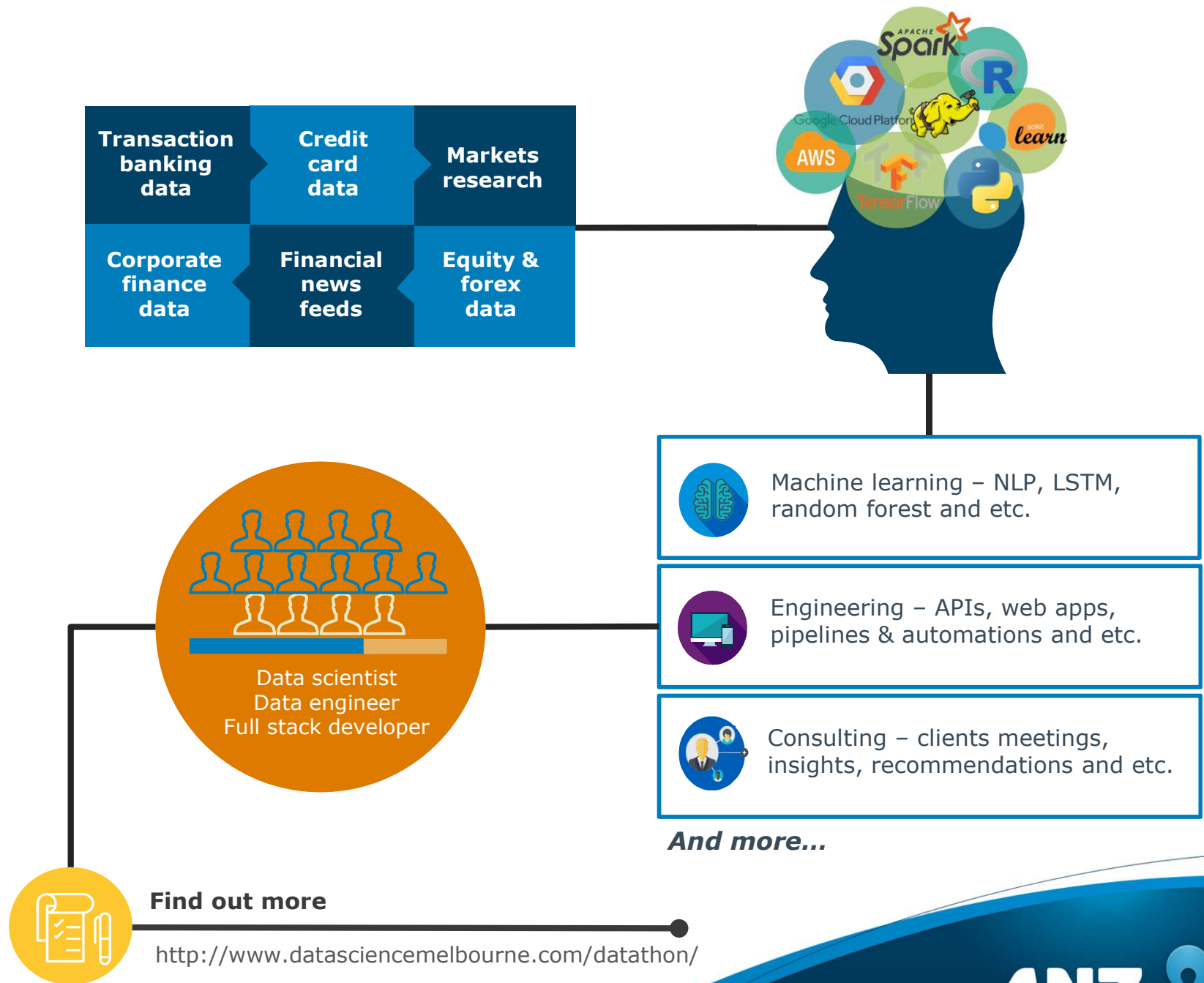
- Data scientist/ ML Engineer
- Data engineer/ Data Pipeline
- Front End/ Apps/ UI

## Data2App Datathon Category

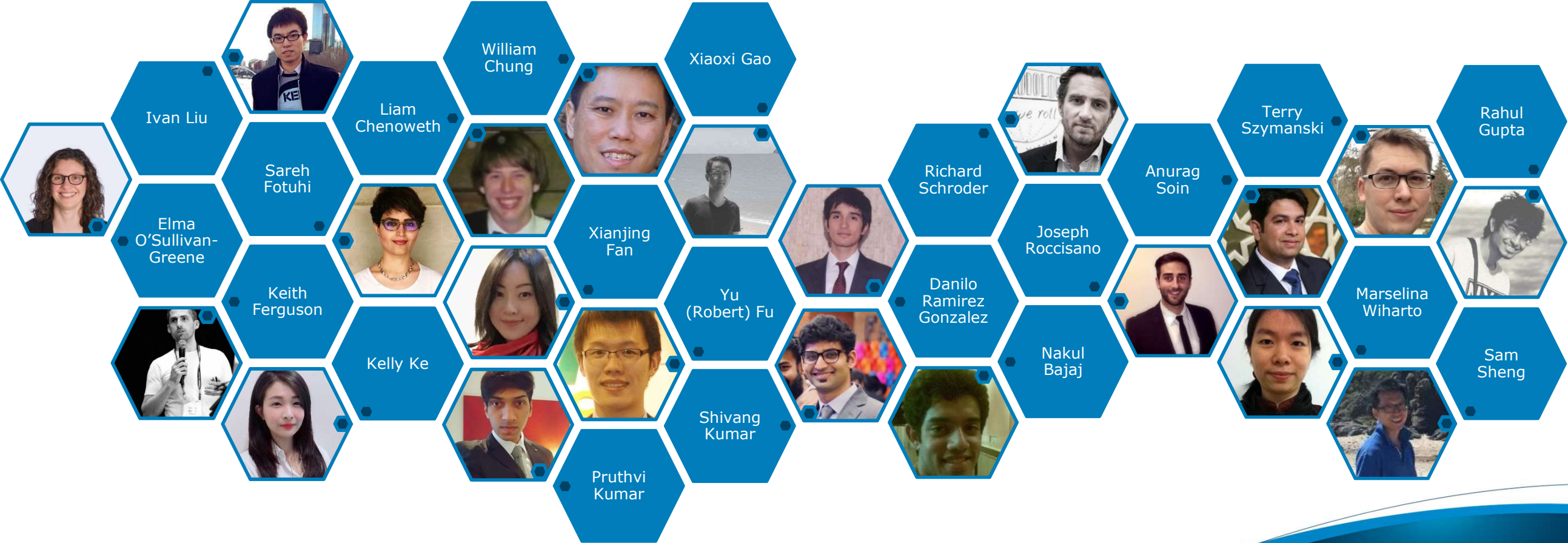
- Offering 3 internships for student registered team

Elma O'Sullivan-Greene

E: [osullive@anz.com](mailto:osullive@anz.com)



# ANZ INSTITUTIONAL DATA SCIENCE TEAM



# WHAT YOU NEED TO DO FOR DATA2APP?

- Use interesting real dataset: transport theme



Data  
Engineering

Traditional Data  
Science  
ML algorithms

Visualisation  
linked to action

## WHY DATA2APP?

What do we stand for?

- People want to interact with data the way they use apps in their personal lives: mobile-first, signals, sharing
- People want data insights to be linked to actionable workflows
- The Lean Startup (Eric Reis) / customer-led design process
- Data2App category in the datathon

# BEYOND KAGGLE: NETFLIX EXAMPLE

kaggle

A platform for data science competitions founded in 2010 now has over **536,000** data scientists across **194** countries.

Solve real-world problems from top companies and research institutions.

NETFLIX

Sep 2009, an algorithm competition with prize of **US\$1,000,000** bested Netflix's own algorithm for predicting ratings by **10.06%**.

(100,480,507 ratings that 480,189 users gave to 17,770 movies)

NETFLIX

Netflix never used its \$1 million algorithm

*Due to engineering costs and has no plan to use it in the future*



Netflix used a suboptimal solution

*For an 8.43% improvements (versus 10% improvements for the winner solution)*

kaggle

Most Kaggle winning solutions are impractical to be implemented

*Kaggle solutions are aiming at model accuracies (complexities) rather than scalabilities.*



suboptimal solutions are more likely to be used

*Simplified version of the winning solutions are more likely to be implement*



ANZ



# BEYOND KAGGLE: INTERPRETABLE VS POWERFULLY PREDICTIVE

Medium



Christophe Bourguignat  
Data enthusiast #BigData #DataScience  
#MachineLearning #FrenchData #Kaggle  
Sep 17, 2014 · 3 min read

## Interpretable vs Powerful Predictive Models

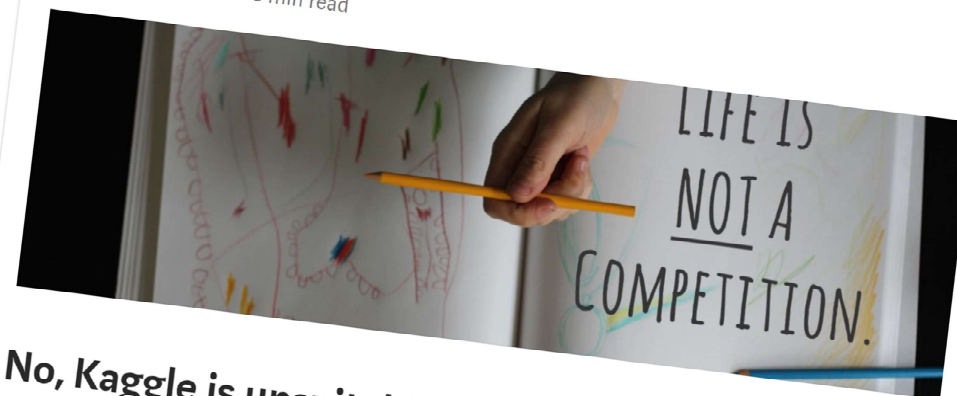
Unfortunately, the predictive models that are most powerful are usually the least interpretable.

**Julia Evans**

"Machine learning isn't  
Kaggle competitions"



Mostapha Benhenda in Hacker Noon  
Apr 27, 2017 · 6 min read



No, Kaggle is unsuitable to study AI & ML. A reply  
to **Ben Hamner**  
[Read more](#)

*"In a business, predictive modeling and accuracy is a means, not an end. What's better: A simple model that's used, updated, and kept running? Or a complex model that works when you babysit it but the moment you move on to another problem no one knows what the hell it's doing?"*

[www.john-foreman.com/blog](http://www.john-foreman.com/blog)

# WHY DATA2APP AND INTERPRETABILITY?

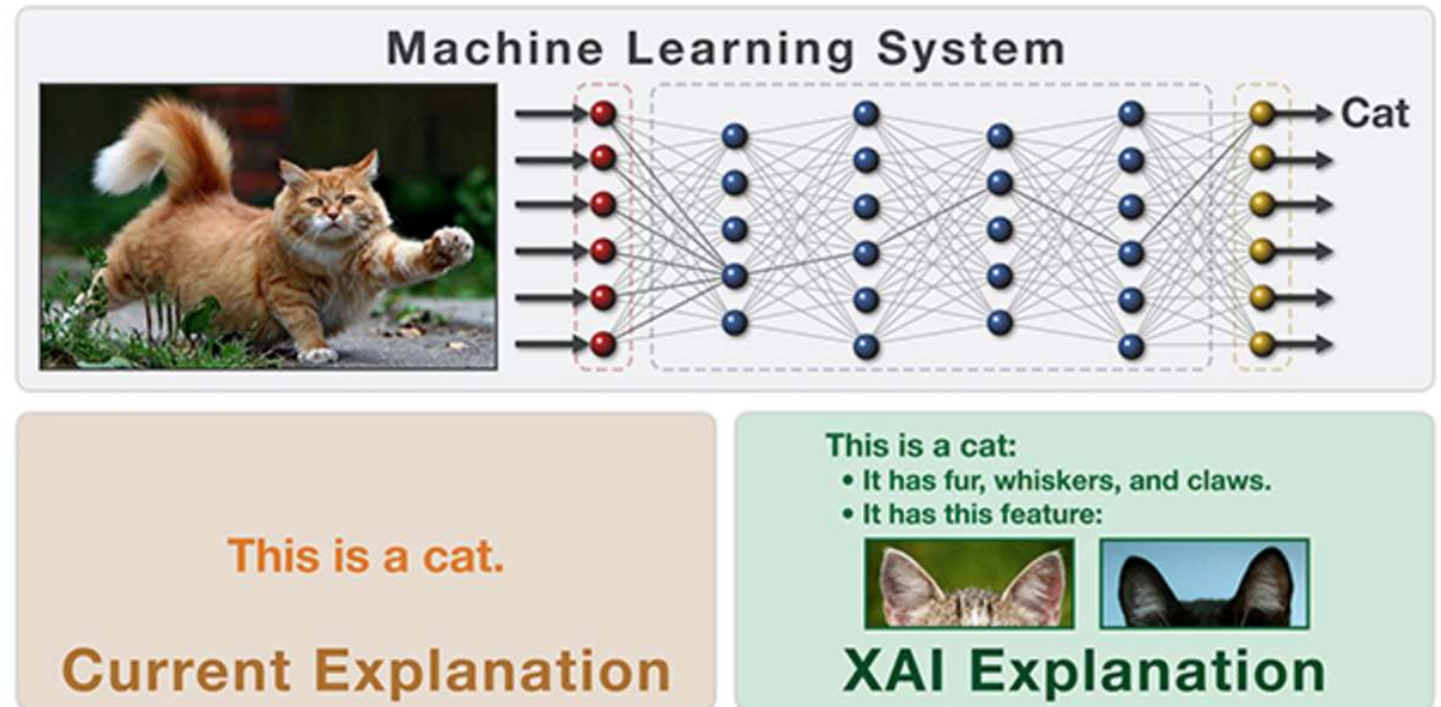
- James Bridle, TED 2018
- YouTube content suggestion algorithms hack the brains of young children in return for advertising revenue
- From "surprise egg" reveals to algorithmically created mashups of familiar cartoon characters in violent situations, these videos exploit and terrify young minds -- and they tell us something about where our increasingly data-driven world is headed.



*"If there's one thing that we can do to start to improve these systems, it's to make them more legible to the people who use them, so that all of us have a common understanding of what's actually going on here."*

# GOOD PRACTICE FOR INTERPRETABLE ML

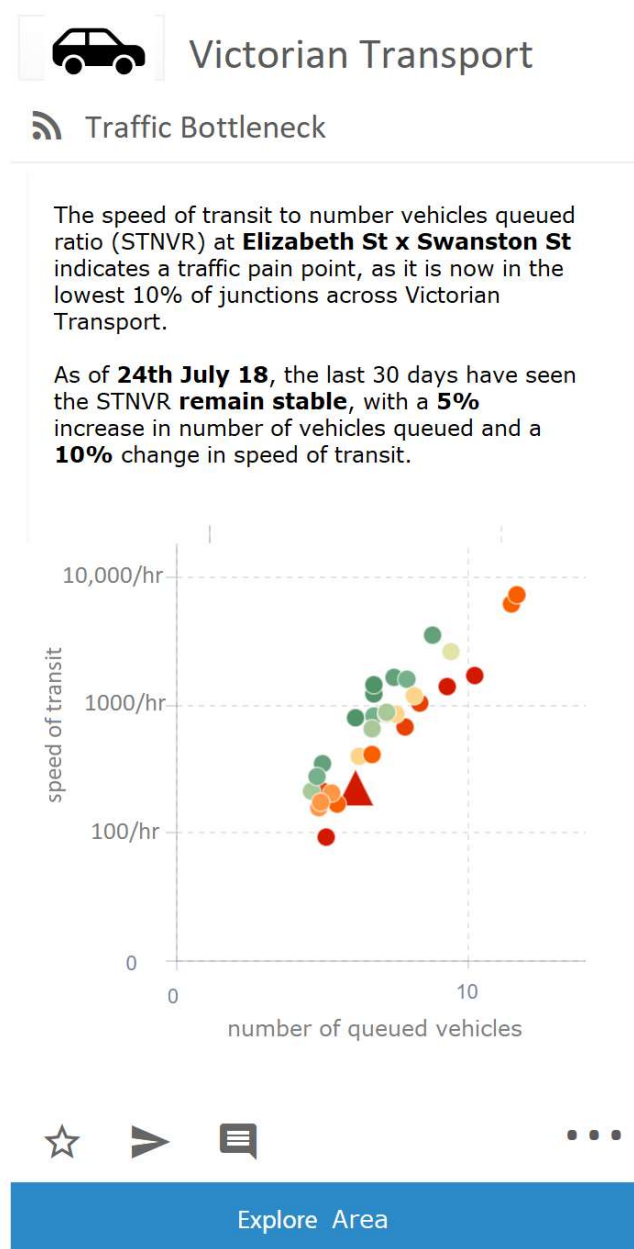
## DARPA's XAI program



- “Produce more explainable models, while maintaining a high level of learning performance (prediction accuracy)”
- “Enable human users to understand, appropriately trust, and effectively manage the emerging generation of artificially intelligent partners.”



# DATA2APP?



- People want to interact with data they way they use apps in their personal lives: mobile-first, signals, sharing
- People want data insights to be linked to actionable workflows
- The Lean Startup / customer-led design process
- A 'good-enough' ML Algorithm linked to interpretable insights visualised on the app

**GOOD LUCK!!**

**SEE YOU AT THE HACKDAY TUTORIALS  
TO FIND OUT MORE ON HOW TO BUILD AN APP**

