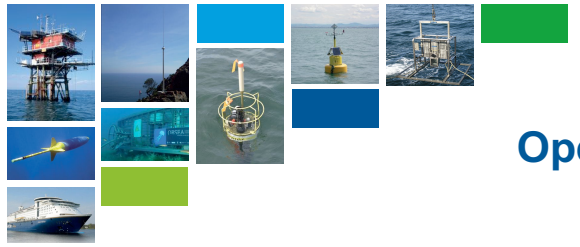


Malta Summer School 2018
Operational Oceanography for Blue Growth

Data, AI and Tokens: Ocean Protocol

John Enevoldsen

Blockchain

Blockchain: a confluence of technologies

Blockchain is a:

- Decentralized ledger
- A way of storing data
- Bringing together cryptography, proofs, consensus mechanisms, decentralization…

Blockchain 1.0: Bitcoin

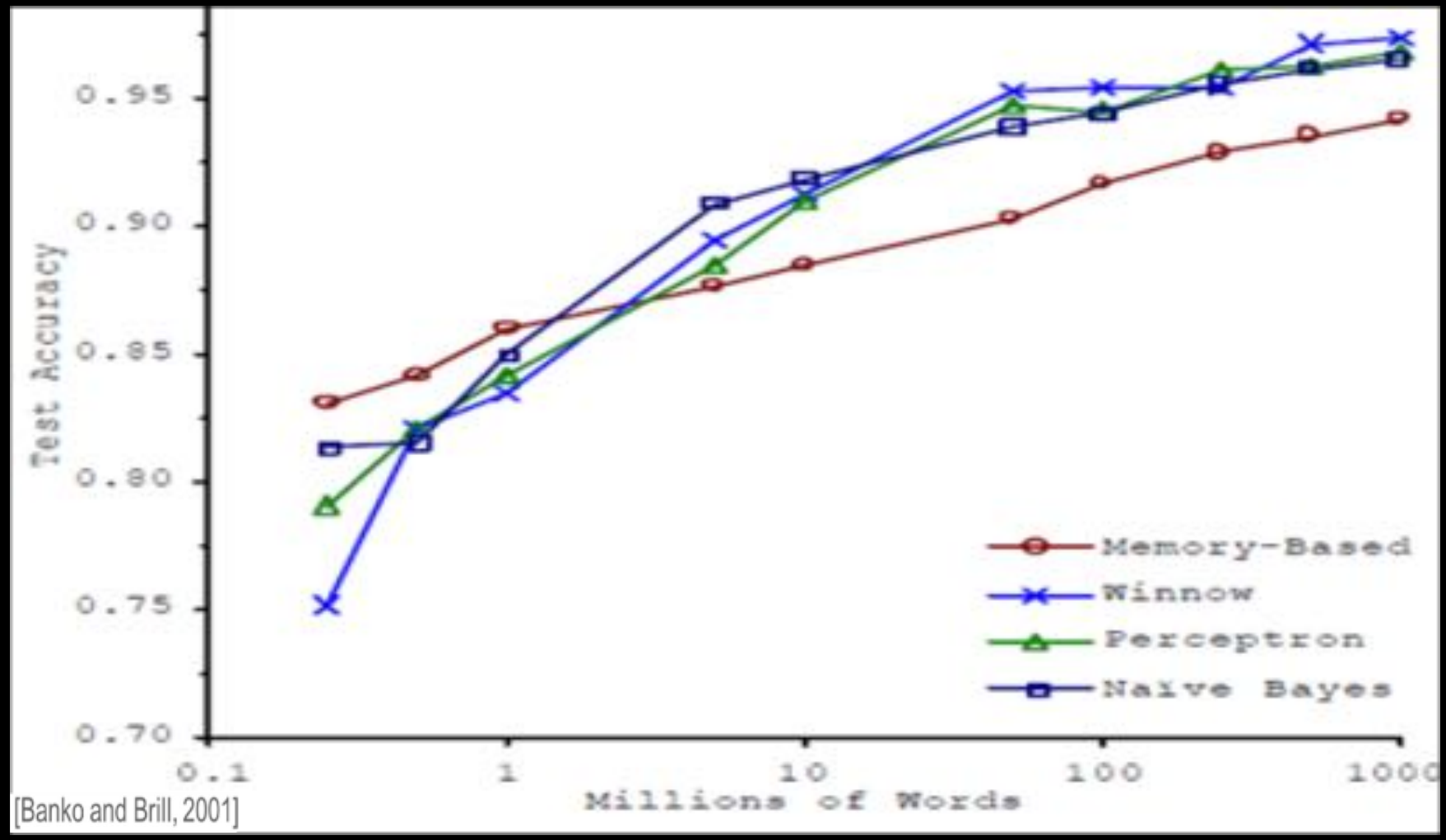
- The most famous blockchain white paper
- New way of exchanging and storing data (=value)
- Basic functionalities, but did it well (apart from energy consumption…)

Blockchain 2.0: Ethereum

- The second most famous blockchain whitepaper!
- Improves on Bitcoin by adding an EVM and smart contracts
- You can execute code, create DAOs, and other cool stuff

AI

The Unreasonable Effectiveness of Data



1000% less error!

1000x *more* data

The Problem

More data
(and more compute)

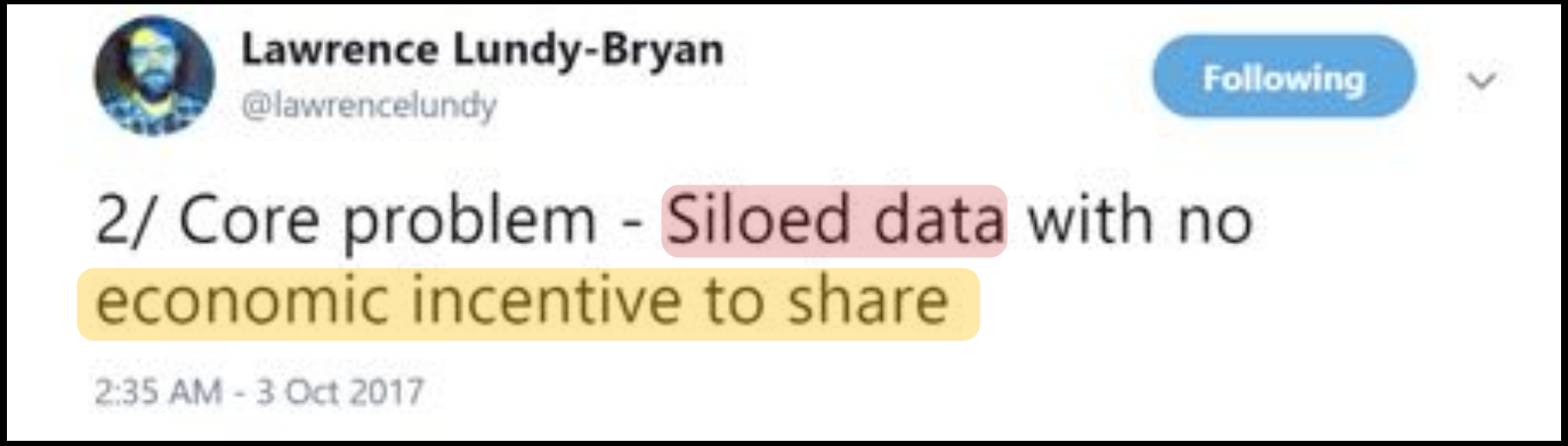
↓

More accuracy

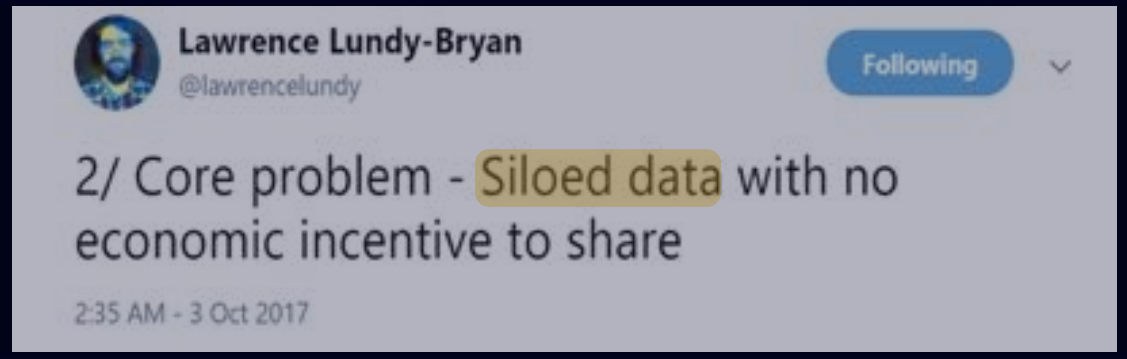
↓

More $


Here's
your
personal
data


Lawrence Lundy-Bryan
@lawrencelundy
Following
2/ Core problem - Siloed data with no economic incentive to share
2:35 AM - 3 Oct 2017

Eroding the silos



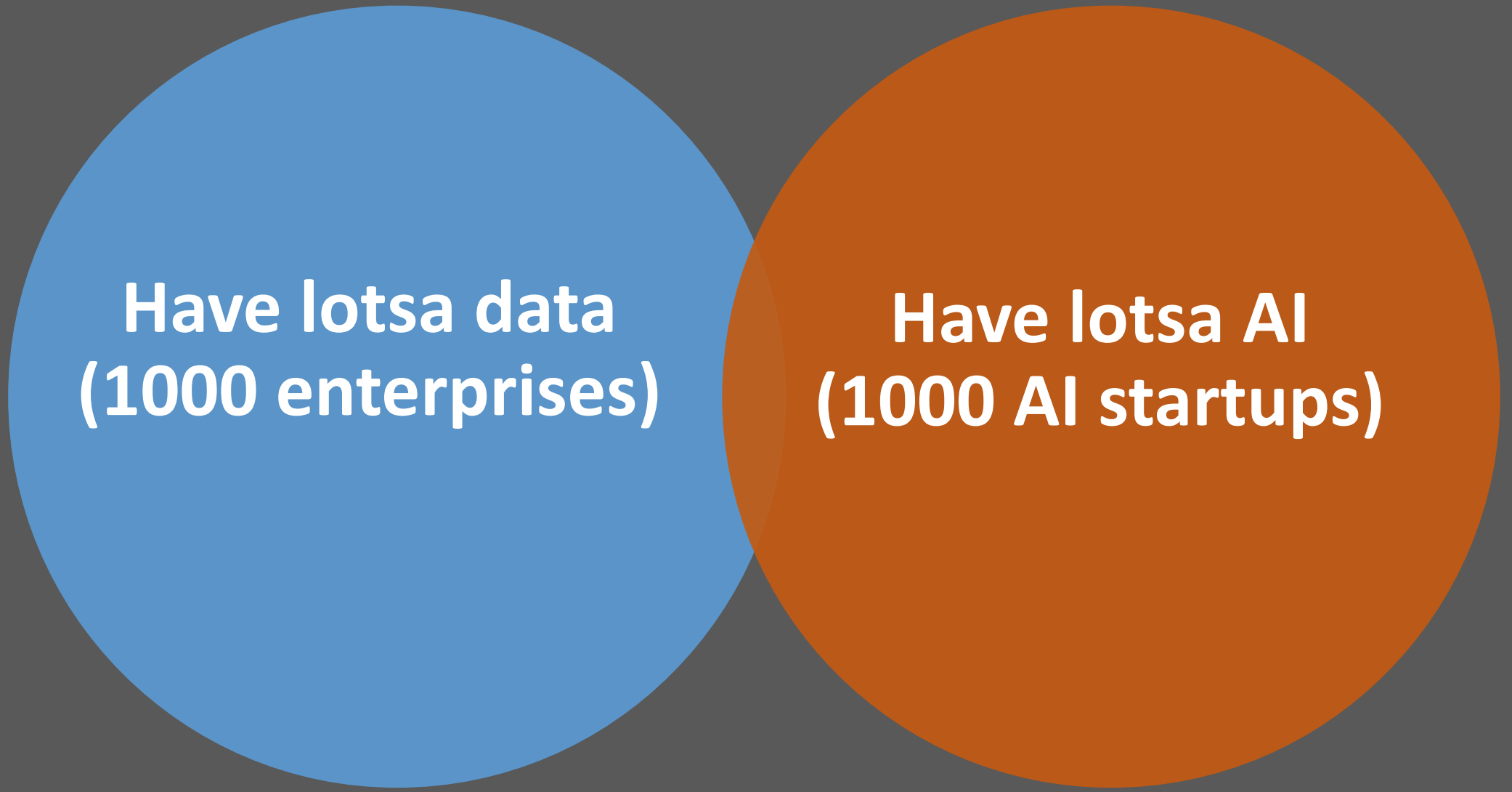
~~Silo~~ *Pool* more Data
(and more compute)

↓

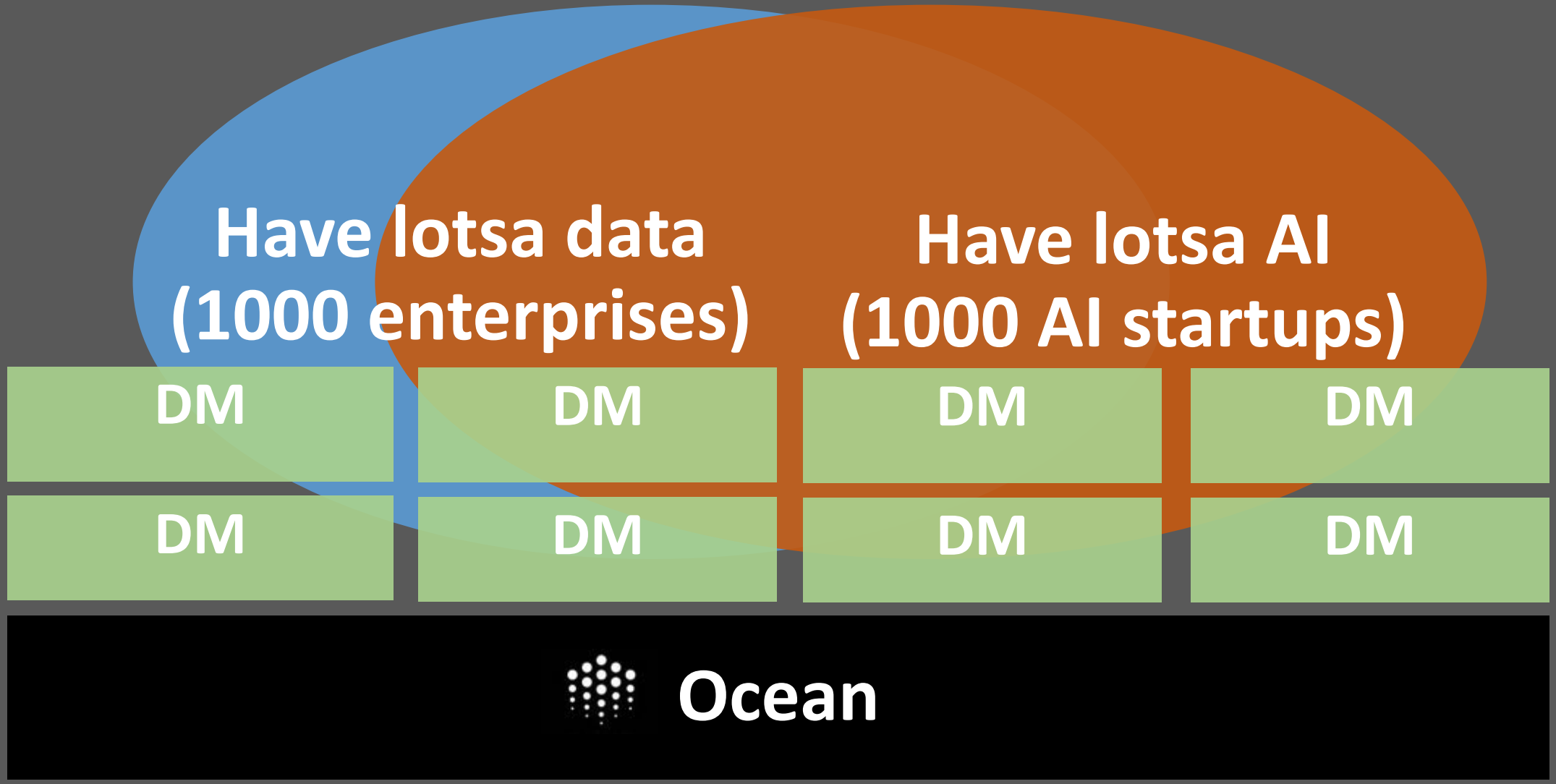
More accuracy

↓

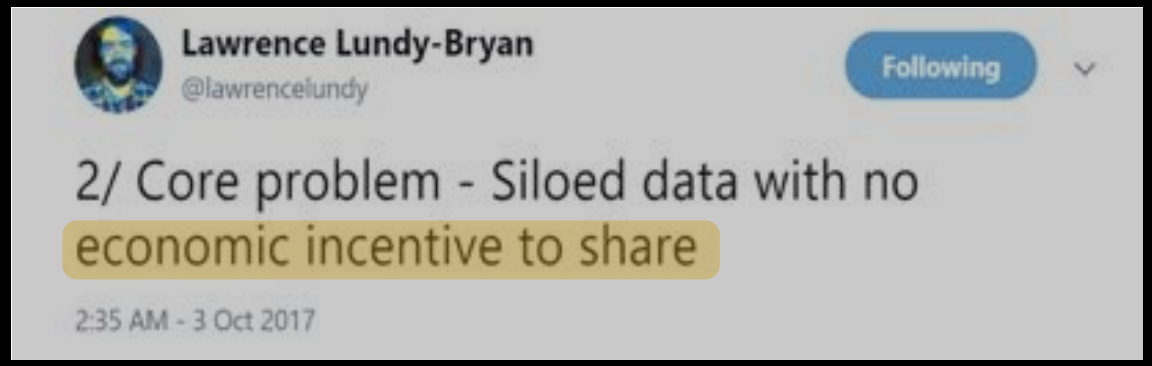
More $


Have lotsa data
(1000 enterprises)
Have lotsa AI
(1000 AI startups)

A new data economy



Economic incentives



“Show me the incentive
and I will show you the outcome.”

-Charlie Munger

Blockchain Superpower:

Get people to do stuff

By rewarding with tokens

Case Study: Analysis of Bitcoin

Bitcoin goal: maximize security of network

Token rewards if: run compute to secure network

Bitcoin objective function

Objective: Maximize security of network

- **Where “security” = compute power**
- **Therefore, super expensive to roll back changes to the transaction log**

Bitcoin objective function

Objective: Maximize security of network

- **Where "security" = compute power**
- **Therefore, super expensive to roll back changes to the transaction log**

$$E(R_i) \ \alpha \ H_i * T$$

E() = expected value

block rewards

hash power of actor = contribution to "security"

 # tokens (BTC) dispensed each block

Case Study: Design of Ocean

Steps in *Token* Design

1. **Formulate the problem.** Objectives, constraints, design space.
2. **Try an existing pattern.** If needed, try different formulations or solvers.
3. **Design new pattern?**

1. Formulate the Problem:
(a) Who are stakeholders? What do they want?

Key stakeholders in Ocean ecosystem

Stakeholder	What value they can provide	What they might get in return
Data/service provider, data custodian, data owner	Data/service (market's supply)	Tokens for making available / providing service
Data/service referrers, curators. Includes exchanges and other application-layer providers.	Data/service (via a provider etc), curation	Tokens for curating
Data/service verifier. Includes resolution of linked proofs on other chains	Data/service (via a provider etc), verification	Tokens for verification
Data/service consumer	Tokens	Data/service (market's demand)
Keepers	Correctly run nodes in network	Tokens for chainkeeping

1. Formulate the problem:
(b) Translate into objectives and constraints

Objective function: maximize supply of relevant data

Token rewards if: supply relevant data
Token rewards if: supply data, and curate it

Objective: maximize supply of relevant data

- **Reward curating data (staking on it) + making it available**
- **New pattern: Proofed Curation Market**

$$E(R_{ij}) \alpha \ log10(S_{ij}) * log10(D_j) * T * R_i$$

Expected reward for user *i* on dataset *j*

S_{ij} = predicted popularity = user's curation market stake in dataset *j*

D_j = proofed popularity = # times made dataset available

 # tokens during interval

From AI data to AI *services*

Motivations:

- **Privacy, so compute on-premise or decentralized**
- **Data is heavy, so compute on-premise**
- **Link in emerging decentralized AI compute**

Objective function: Maximize supply of relevant *services*

=reward curating *services* + proving that it was delivered

$$E(R_{ij}) \alpha \log10(S_{ij}) * \log10(D_j) * T * R_i$$

predicted popularity of *service*

proofed popularity of *service*

Ocean is a network of *curated services*. An AI services pipeline.

Production	Availability	Discovery	Consumption	Privacy	Inter-Operability	Governance

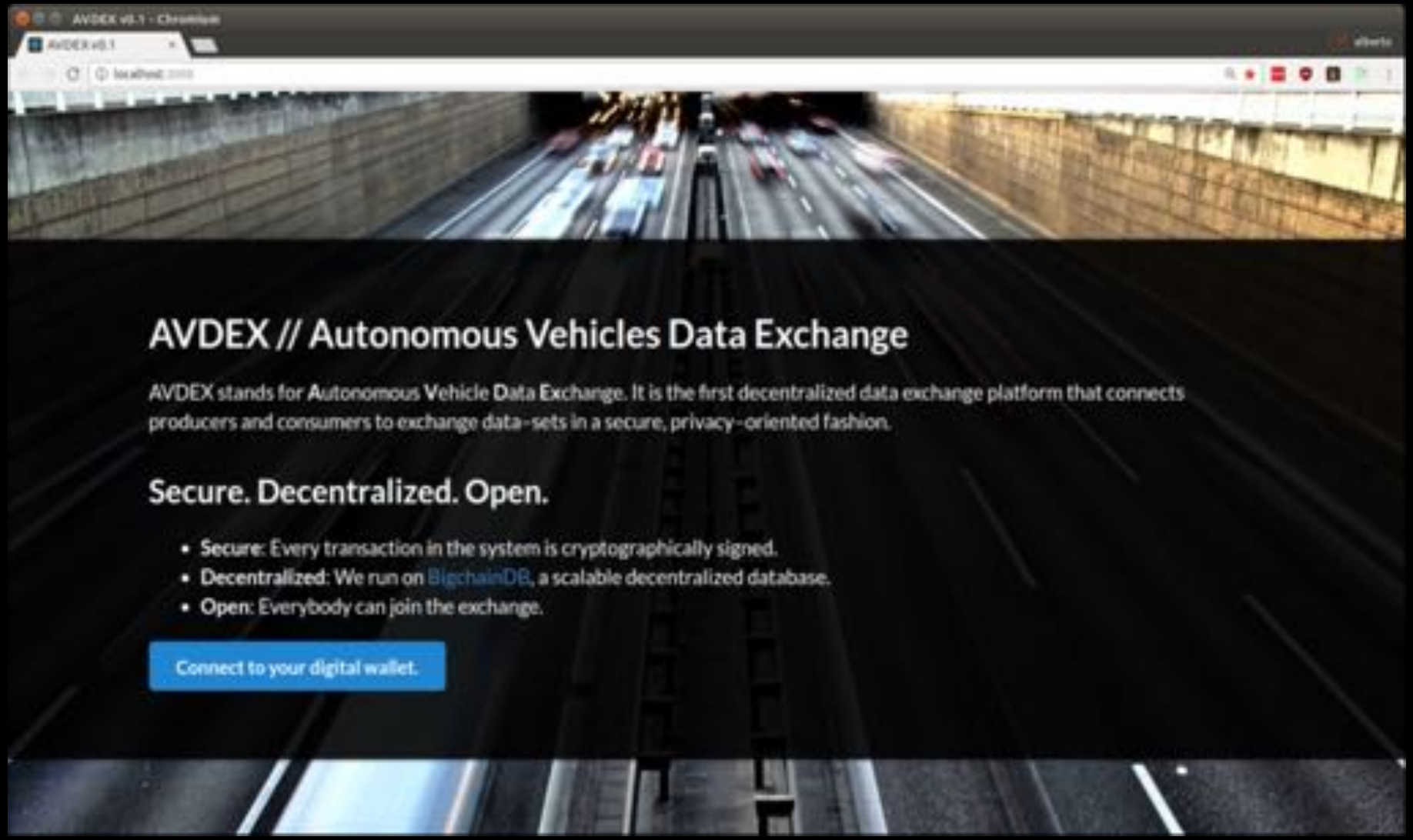
ocean

*Note: logos shown are examples and do not imply partnerships or integrations

Data Commons

What unlocking AI data & services unlocks

Self-driving cars: fewer accidents, more mobility




>100x more data for health care research

Machine/Deep Learning as a Service

Fred Ehrsam - 2018

We can go further: an Ocean service where competing models earn tokens:

- Marketplace for ML/DL models
- Uses secure computation (MPC, HE)
- An Oracle checks the best models, and reward them

Models competing for data!

Brad Burnham

Search inverted: not people searching for products, but products searching and competing for people

- People upload their data using an Ocean service
- Models for companies compete and bid to use this data
- People get rewarded by tokens

Conclusion

AI data is siloed.

AI services are siloed.

Let’s change the rules of the game with incentives.

Let’s democratize access to AI data & services!

John Enevoldsen