

Blockchains for Artificial Intelligence

Trent McConaghy
Cofounder & CTO, BigchainDB | IPDB
@trentmc0



The Unreasonable Effectiveness of Data

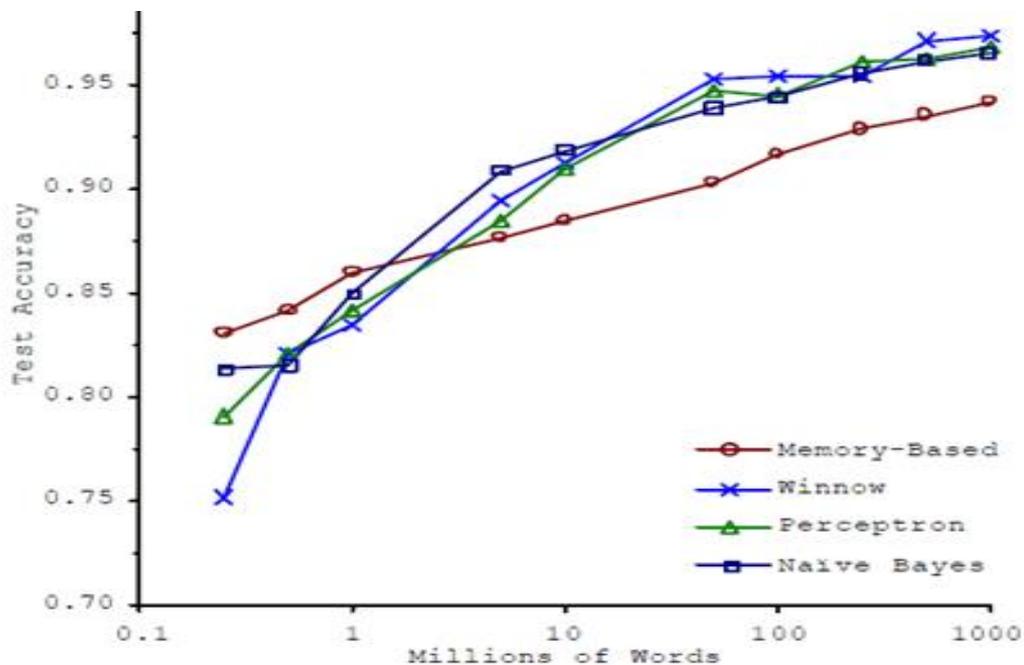


Figure 1. Learning Curves for Confusion Set Disambiguation
[Banko and Brill, 2001]

The world's most valuable resource



Data and the new rules
of competition

Mo' data
(and mo' compute)



Mo' accuracy



Mo' \$

THE 3 ELEMENTS OF COMPUTING

STORAGE

PROCESSING

COMMUNICATIONS

THE 3 ELEMENTS OF COMPUTING



Key Blocks in AI Landscape

STORAGE

FILE SYSTEM

HDFS, S3

DATABASE

MongoDB,
Cassandra

PROCESSING

BIZ LOGIC

CPU, EC2

HIGH PERF. COMPUTE

Nvidia GPU, Goog TPU,
MapReduce, Spark

COMMUNICATIONS

DATA

TCP/IP, HTTP

But all is not well in the world of AI

- **Data hoarding.** Big guys have all the data.
- **Weak data history.** Garbage in, garbage out.
- **Data is *expensive*.**

And more..

But all is not well in the world of AI

- **Data hoarding.** Big guys have all the data.
- **Weak data history.** Garbage in, garbage out.
- **Data is *expensive*.**

And more..

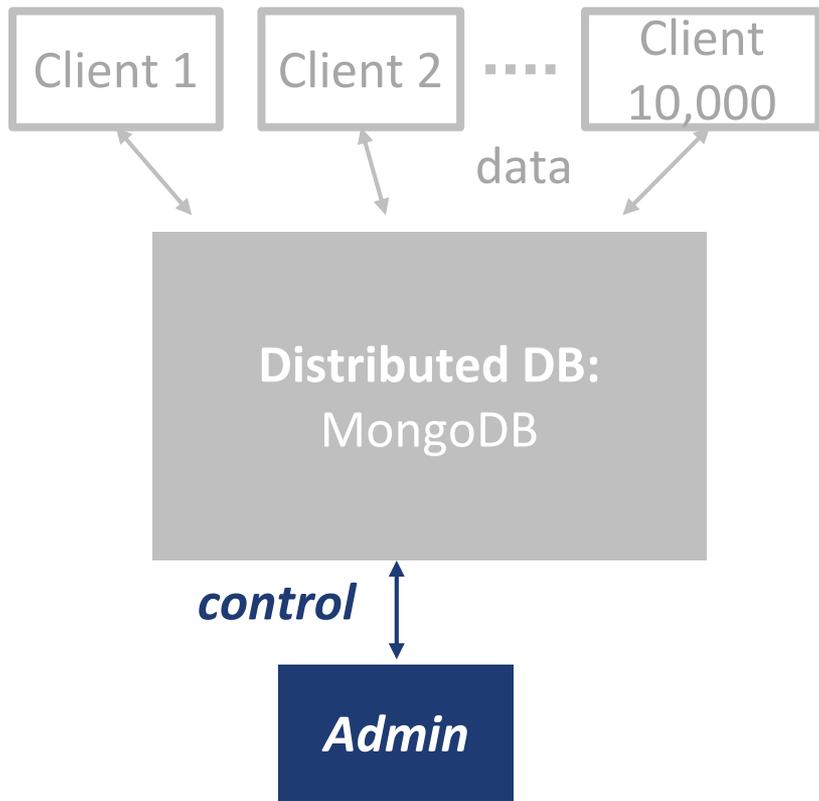
Can decentralization help?

Q: How to unlock blockchains for AI?

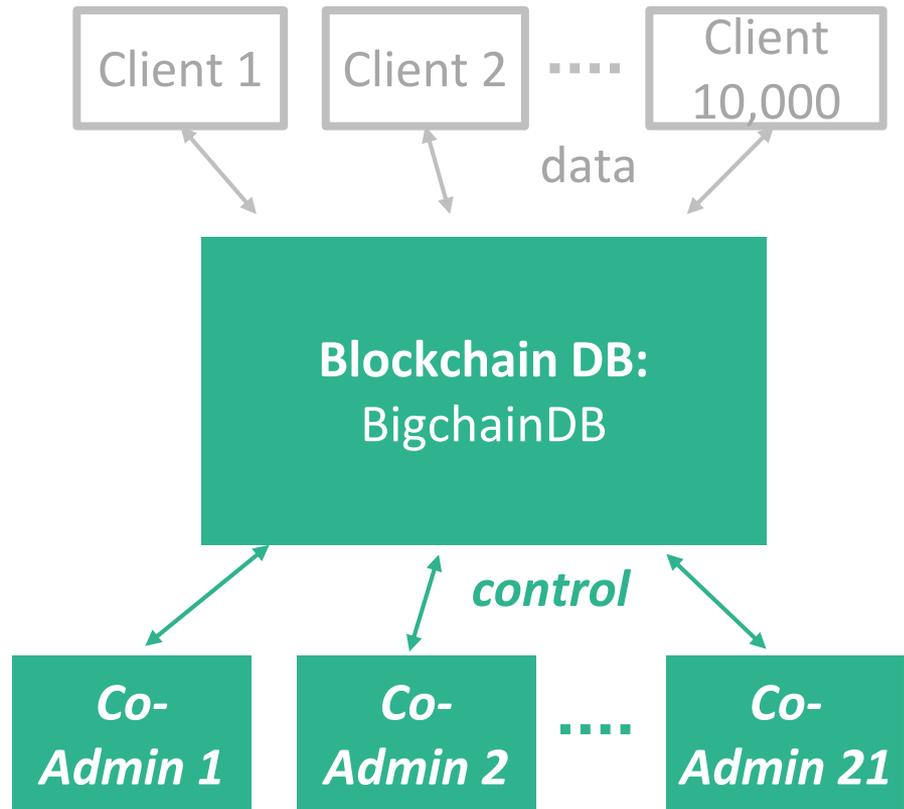
A: A shared database with planetary reach

+Query
+Open-source
+Scale
+Decentralized, Assets

1. Relational DB – Oracle
2. Website-ready DB – MySQL
3. “Big data” Distributed DB – MongoDB
4. “Blockchain” DB – BigchainDB + IPDB



A traditional database has **one administrator** (centralized)



A blockchain database has **many administrators** (decentralized)

THE 3 ELEMENTS OF COMPUTING, *DECENTRALIZED*

STORAGE

FILE SYSTEM

IPFS/FileCoin, Swarm

DATABASE

BigchainDB/IPDB

E-GOLD / E-CASH

Bitcoin, zcash, .*

PROCESSING

BIZ LOGIC

Ethereum, Hyperledger

HIGH PERF. COMPUTE

TrueBit, Golem, iExec,
VMs, client-side compute

COMMUNICATIONS

DATA

TCP/IP, HTTP

VALUE

ILP, Cosmos

State

PolkaDot, Aeternity

THE 3 ELEMENTS OF COMPUTING, *DECENTRALIZED*



Key Blocks in AI Landscape

STORAGE

FILE SYSTEM

IPFS/FileCoin, Swarm

DATABASE

BigchainDB/IPDB

E-GOLD / E-CASH

Bitcoin, zcash, .*

PROCESSING

BIZ LOGIC

Ethereum, Hyperledger

HIGH PERF. COMPUTE

TrueBit, Golem, iExec,
VMs, client-side compute

COMMUNICATIONS

DATA

TCP/IP, HTTP

VALUE

ILP, Cosmos

State

PolkaDot, Aeternity



Online platform for industrial 3d printing.

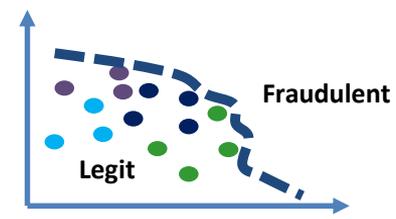
E.g. spare aircraft parts

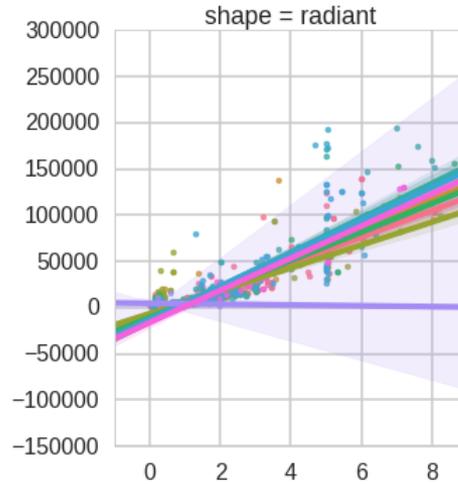
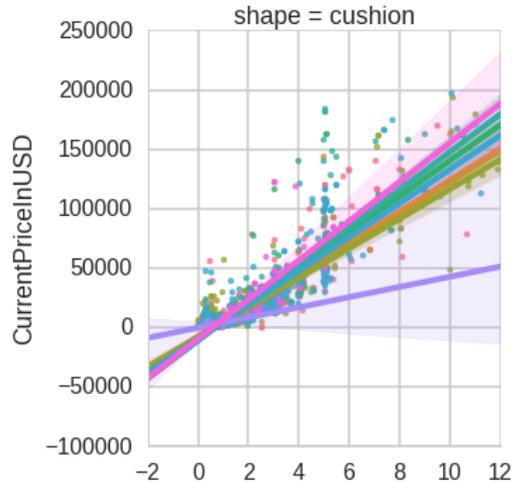
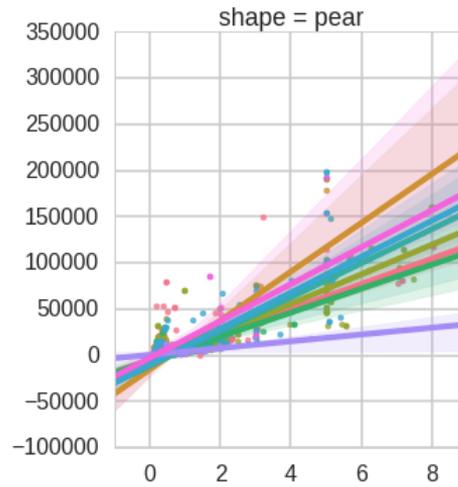
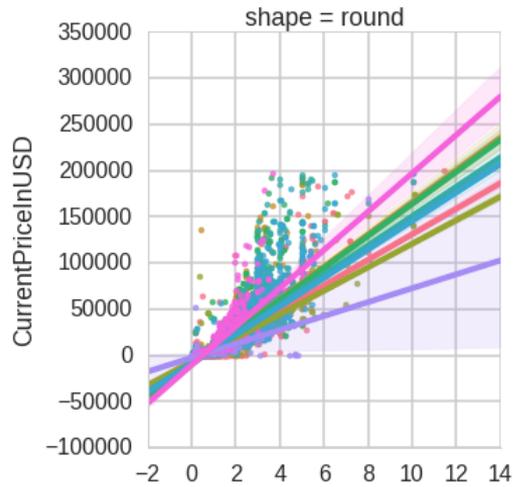
- Find and contract the best 3d printer
- Securely transfer production files
- Pool data in ecosystem → 1-class classifiers for fraud detection

Problem: Data Hoarding

Sol'n: Data Pooling

For More Accurate Models





Problem: Data Hoarding (2)
Sol'n: Data Pooling For More Accurate Models

**Diamond price prediction
for fraud detection:
Warn if predicted price \neq
asking price**

Problem: High Friction to Monetize Algorithms (2)

Solution: Hedge Fund In a Box (Numeraire)

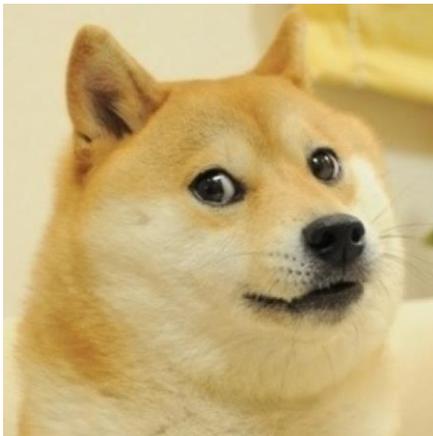
1. 12K+ data scientists submit algorithms
2. Market winnings are distributed wrt performance
3. Positive-sum via tokenization



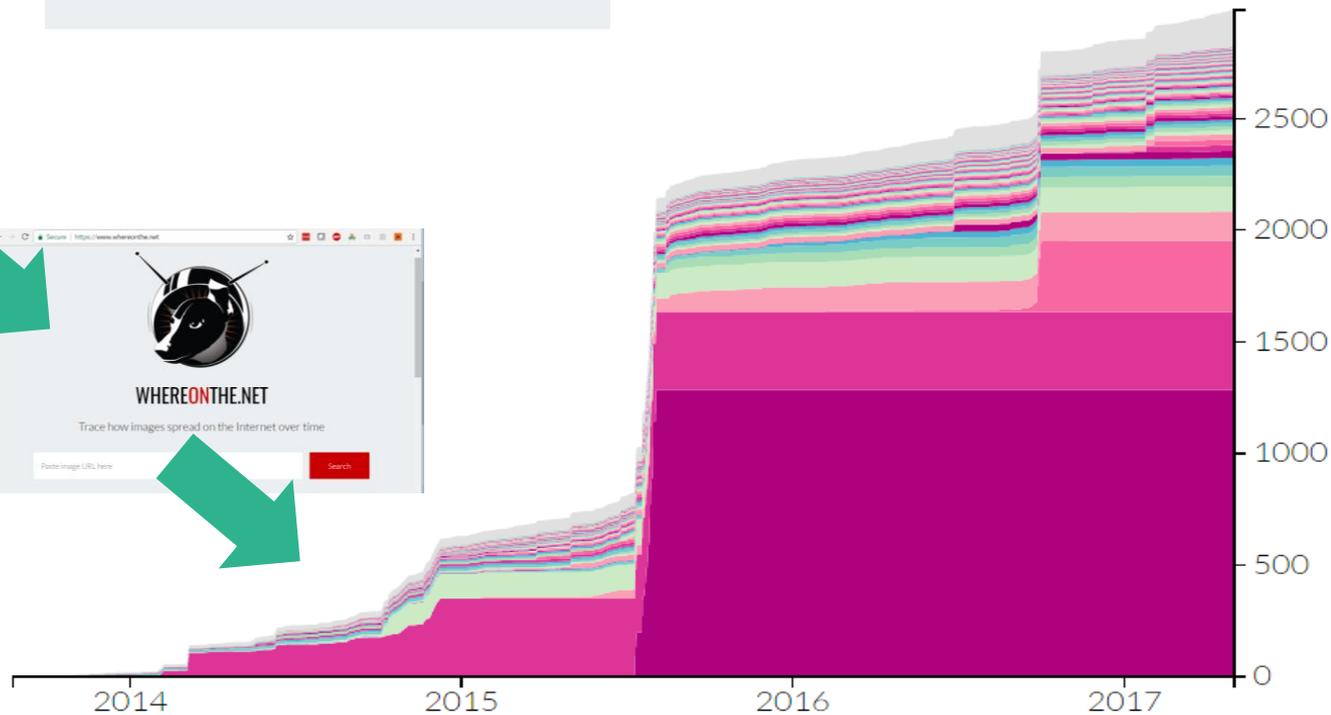
Problem: blockchain-secured data spreads online



Sol'n: visibility into spread via web crawl + AI



WHEREONTHE.NET



Problem: Weak Data History (Garbage In Garbage Out)

Sol'n: Immutable Audit Trails of AI Data & Models

Provenance in model building:

- Sensor / input stream data →
- Training X/y data →
- Model building convergence →

Provenance in model deployment:

- Testing X data →
- Model simulation →
- Testing yhat data →

Time-stamp to IPDB
Store to IPFS



Problem: Weak Data History (2)

Sol'n: Audit Trails of Vehicle Life Cycle Data (CarPass)

innogy
BIGCHAIN

CAR PASS

Welcome admin
Last Login : 22-Jun-2016 20:26 IST

See All Users Invite New User See All Cars Register Car Transfer Car

TRIP MODE	START TIME	MILEAGE START	START TRIP	END TIME	MILEAGE END	END TRIP	DURATION	DISTANCE	FUEL USAGE
	2012-09-10T19:06:33Z	15515757	In Leipzig, Wahren, Stahlmeier Straße 195 (DE 041 59)	2012-09-10T19:25:53Z	15527460	In Leipzig, Südvorstadt, Bernhard-Göring-Straße / Schamkarstraße(DE04275)	1160	11703	4
	2013-10-10T12:06:13Z	15527460	The Taj Mahal Palace Apollo Bandar, Colaba, Mumbai, India	2013-10-10T13:45:45Z	15539462	Chhatrapati Shivaji International Airport Area, Vile Parle, Mumbai, India	1360	12002	30

TRIP

TRIP

BIGCHAIN DB



innogy



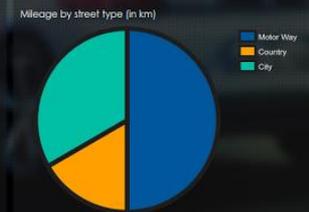
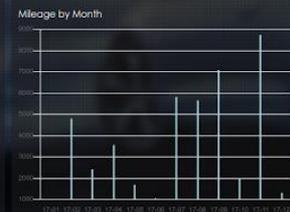
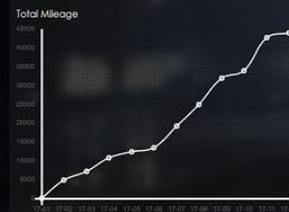
riddle&code

Welcome admin
Last Login : 22-Jun-2016 20:26 IST

Transfer Car

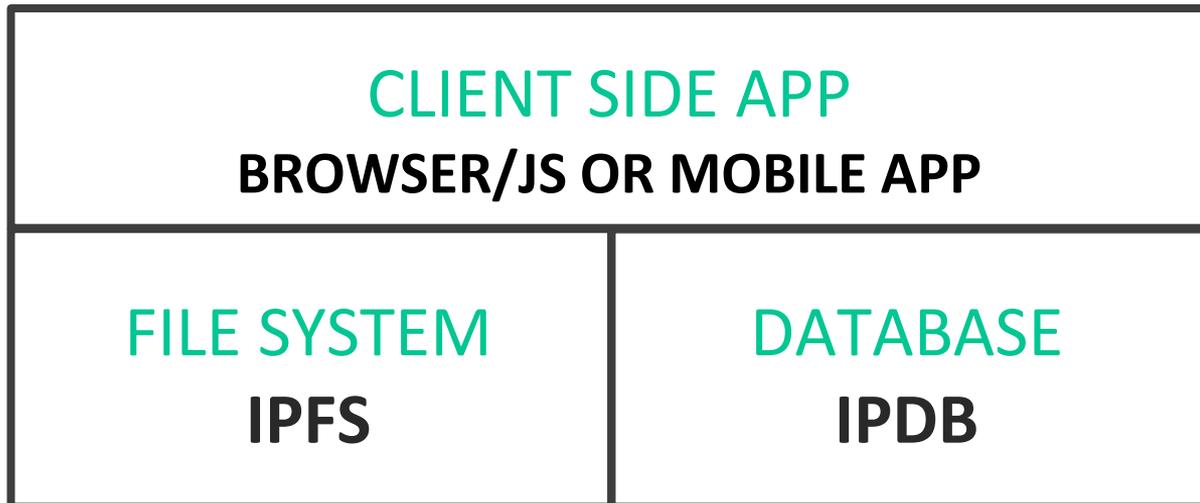
Pollutant class :
Environmental label :
First Registration : 2012-09-10T19:06:33Z
Color(manufacturer) : red
Average Consumption : 6658
Mileage : 43897

Transmission : manual



Architecture: Single-page webapp

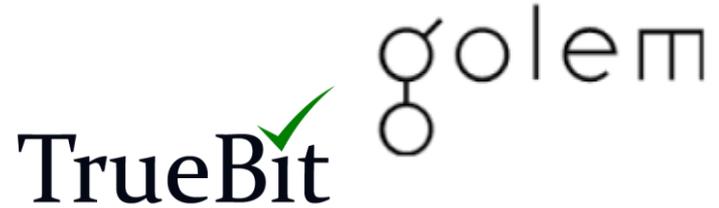
Data persistence via IPFS + IPDB
(“AWS without the AWS”)



FILE SYSTEM



HIGH PERF. COMPUTE



Problem: Compute & Storage are Expensive

Solution: Tokenized, Competitive Markets for Compute & Storage

Problem: Data is Expensive

What's the ultimate way to
unlock data?

A Data Exchange

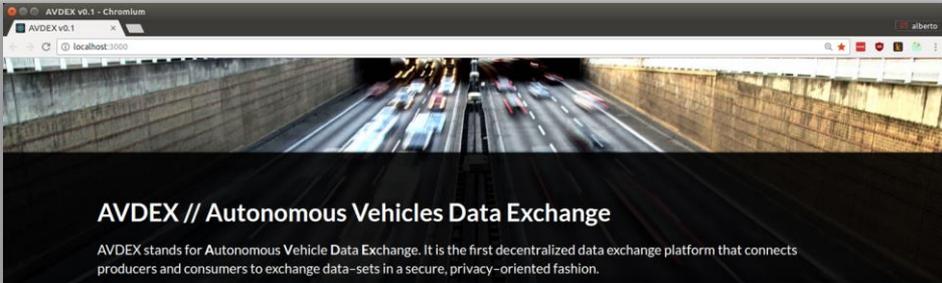
Mo' data
(and mo' compute)



Mo' accuracy



Mo' \$



Problem: Data is Expensive

Sol'n: A Decentralized Data Exchange for Self-Driving Car Data

Secure. Decentralized. Open.

- Secure: Every transaction is secured by a decentralized network of nodes.
- Decentralized: We don't have a central authority.
- Open: Everybody can participate.

Connect to your digital identity

Welcome back, BigchainDB Home Create offer My Offers (1) Search Offers Logout

Name: CH2_002

Purpose: Final Round Test Data: JPG and Filtered ROSBAG

Description:	Date	Lighting Conditions	Duration	Compressed Size	Direct Download
11/18/2016	Daytime/Shadows	--	4.4GB	None	None

* HMB_1: 221 seconds, direct sunlight, many lighting changes. Good turns in beginning, discontinuous shoulder lines, ends in lane merge, divided highway
 * HMB_2: 791 seconds, two lane road, shadows are prevalent, traffic signal (green), very tight turns where center camera can't see much of the road, direct sunlight, fast elevation changes leading to steep gains/losses over summit. Turns into divided highway around 350s, quickly returns to 2 lanes

Date	Size (GB)	Hash
2016	4.4	md5:f3178f88

Sign and submit



Welcome back, BigchainDB Home Create offer My Offers (2) Search Offers Logout

BigchainDB

7KaEt27hS5wDfPMZmdzQo28BSGJCJK3djR8kLUMankE

<https://www.bigchaindb.com/>

CH2_002

\$1500 4.4 GB transaction ★★★★★

Date	Lighting Conditions	Duration	Compressed Size	Direct Download	Torrent	MDS
11/18/2016	Daytime/Shadows	--	4.4GB	None	None	

- HMB_1: 221 seconds, direct sunlight, many lighting changes. Good turns in beginning, discontinuous shoulder lines, ends in lane merge, divided highway
- HMB_2: 791 seconds, two lane road, shadows are prevalent, traffic signal (green), very tight turns where center camera can't see much of the road, direct sunlight, fast elevation changes leading to steep gains/losses over summit. Turns into divided highway around 350s, quickly returns to 2 lanes
- HMB_4: 99 seconds, divided highway segment of return trip over the summit
- HMB_5: 212 seconds, guardrail and two lane road, shadows in beginning may make training difficult, mostly normalizes towards the end
- HMB_6: 371 seconds, divided multi-lane highway with a fair amount of traffic

View >

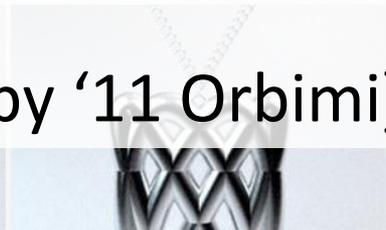
“Self Driving cars are the killer app for AI”
 -Madrona Venture Group

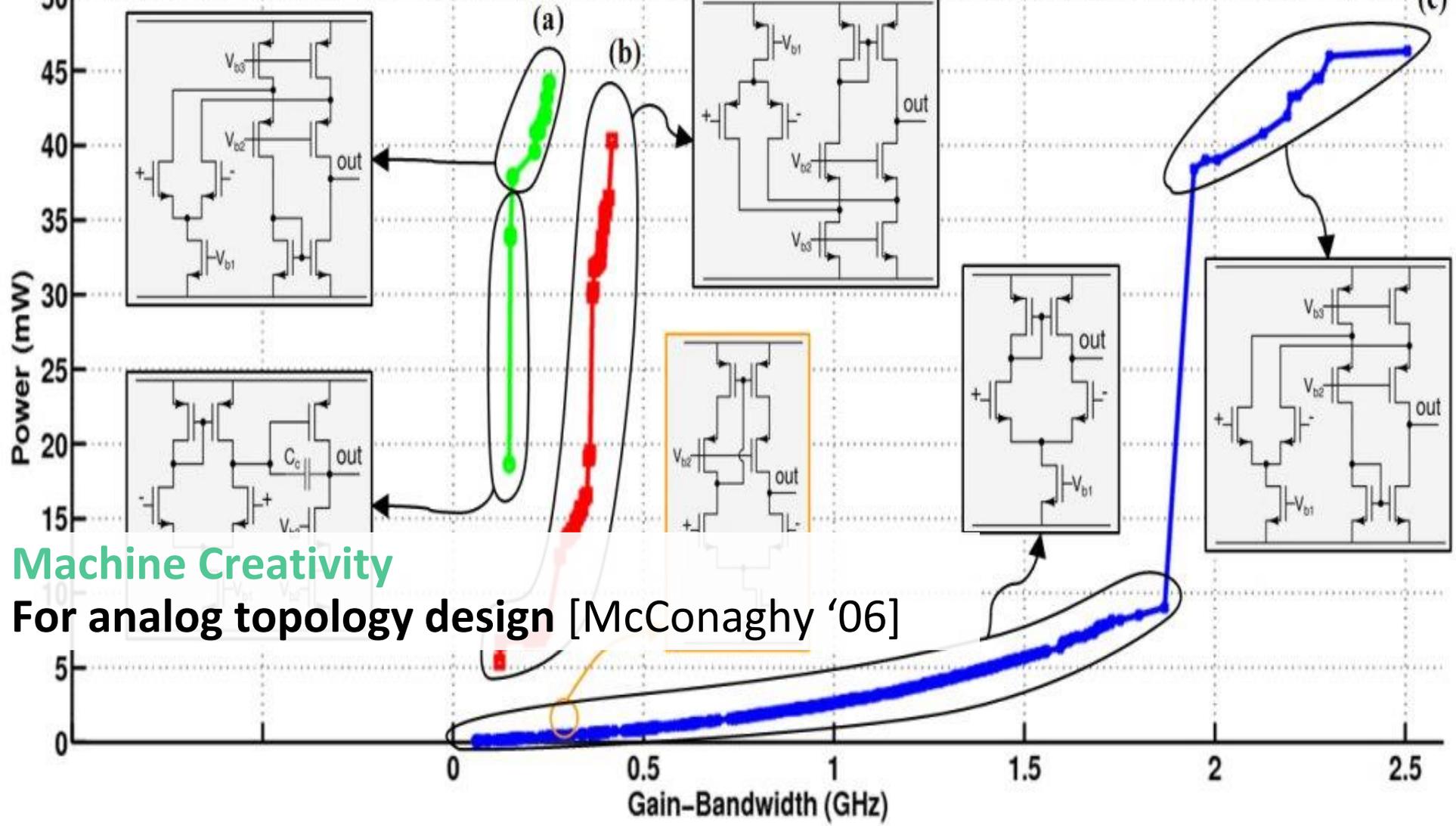
AI * Blockchain Symbiosis:

AI DAOs



Machine Creativity
For jewelry design [Hornby '11 Orbimi]





AGI: Artificial *General* Intelligence



Agents that sense, model, and act

LOG IN | SIGN UP LONGFORM REVIEWS VIDEO TECH CIRCUIT BREAKER SCIENCE ENTERTAINMENT CARS TL;DR FORUMS

COMMENTS

Microsoft will now let anyone test their AI creations in Minecraft

By [Russell Brandom](#) on July 8, 2016 10:36 am [Email](#) [@russellbrandom](#)



son premiere
ial media by
rs

s used
o target

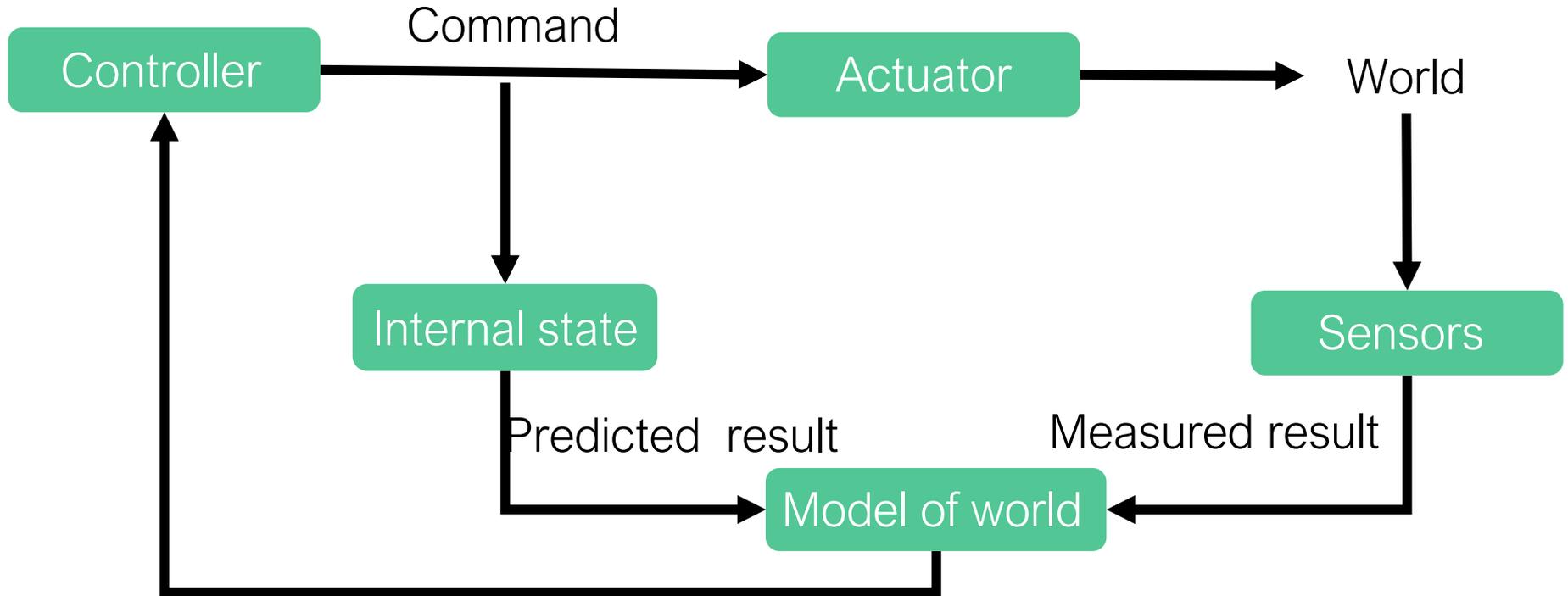
ly
to go!
still

The image shows a screenshot of a Minecraft game. The main view is a large, complex stone structure built with grey and brown blocks, featuring a central tower with a black and white pixelated face. The structure is surrounded by green foliage and a cobblestone path. There are two smaller inset images on the left side of the article, showing different parts of the same structure. The article is from a website with a navigation bar at the top and a comments section on the right.

AGI: Artificial General Intelligence

“AI meets Feedback Control Systems”

Update internal state based on estimate of world state



Example: The ArtDAO

Algorithm...

1. Run AI art engine to generate new image
2. Claim attribution in blockchain
3. Post editions for sale onto a marketplace, using Getty (centralized), or OpenBazaar (decent.)
4. Sell the editions. \$ goes to ArtDAO, in exchange for IP

Repeat! Create more art, sell it, get wealthier

Example: The ArtDAO Algorithm...

1. Run AI art engine to generate new image

Over time, if ArtDAO makes more money from sales than from generating new art, then it will accumulate wealth. And, you can't turn it off.

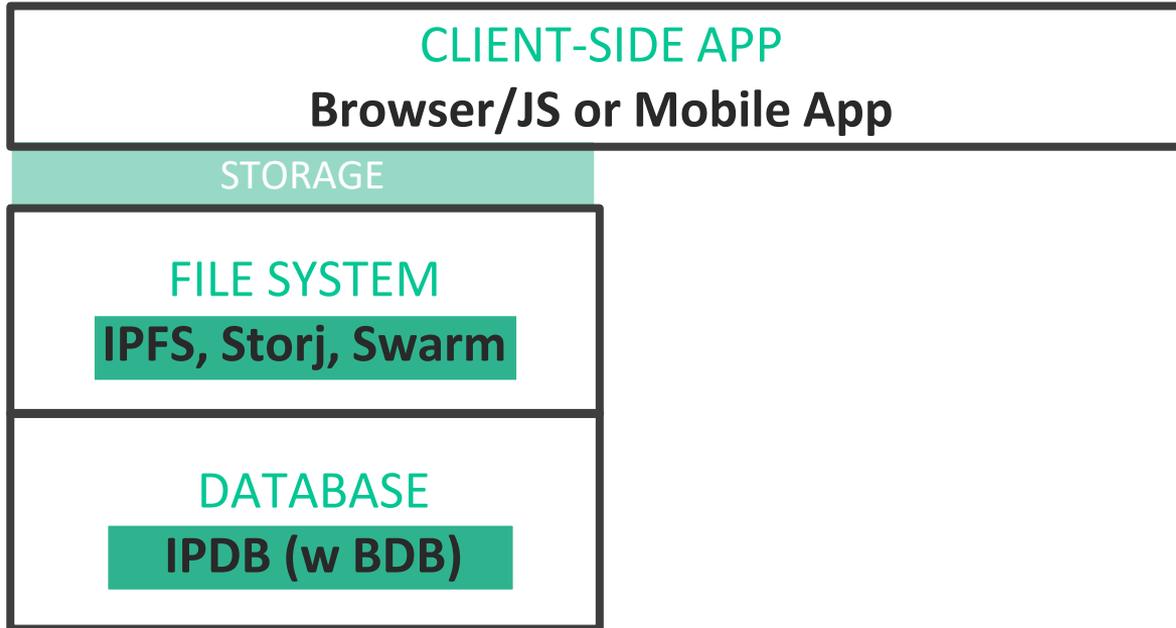
4. Sell the editions. \$ goes to ArtDAO, in exchange for IP

Repeat! Create more art, sell it, get wealthier

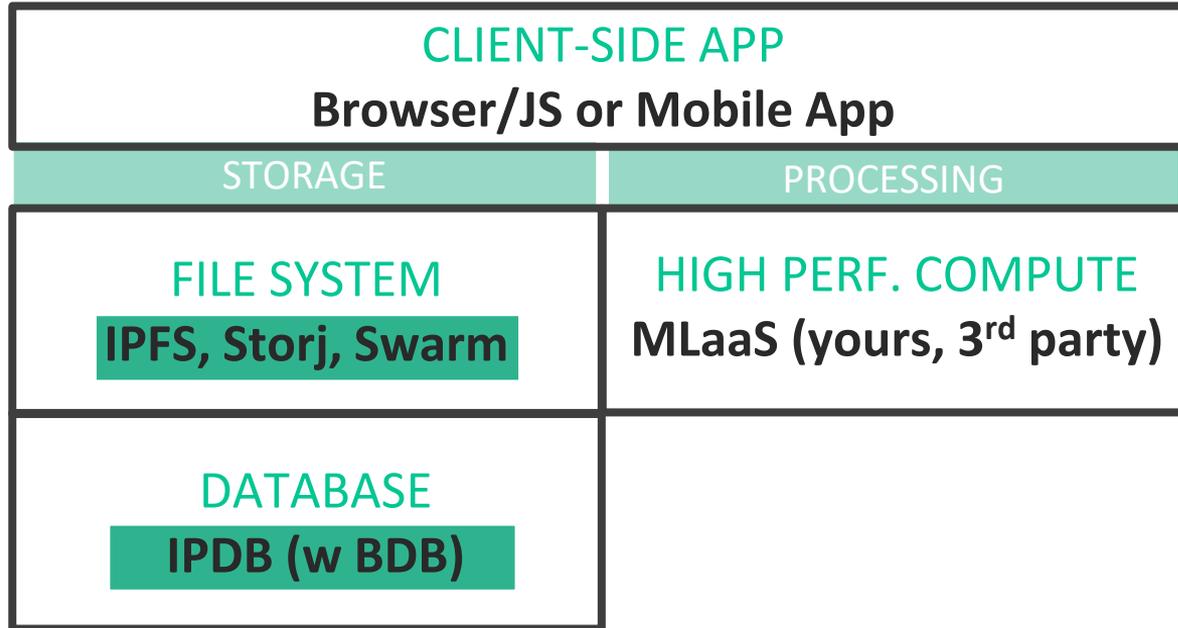
Deployment Architectures

A spectrum of decentralization choices

Single-Page Webapp (Simple!)



Single-Page Webapp, With Your AI



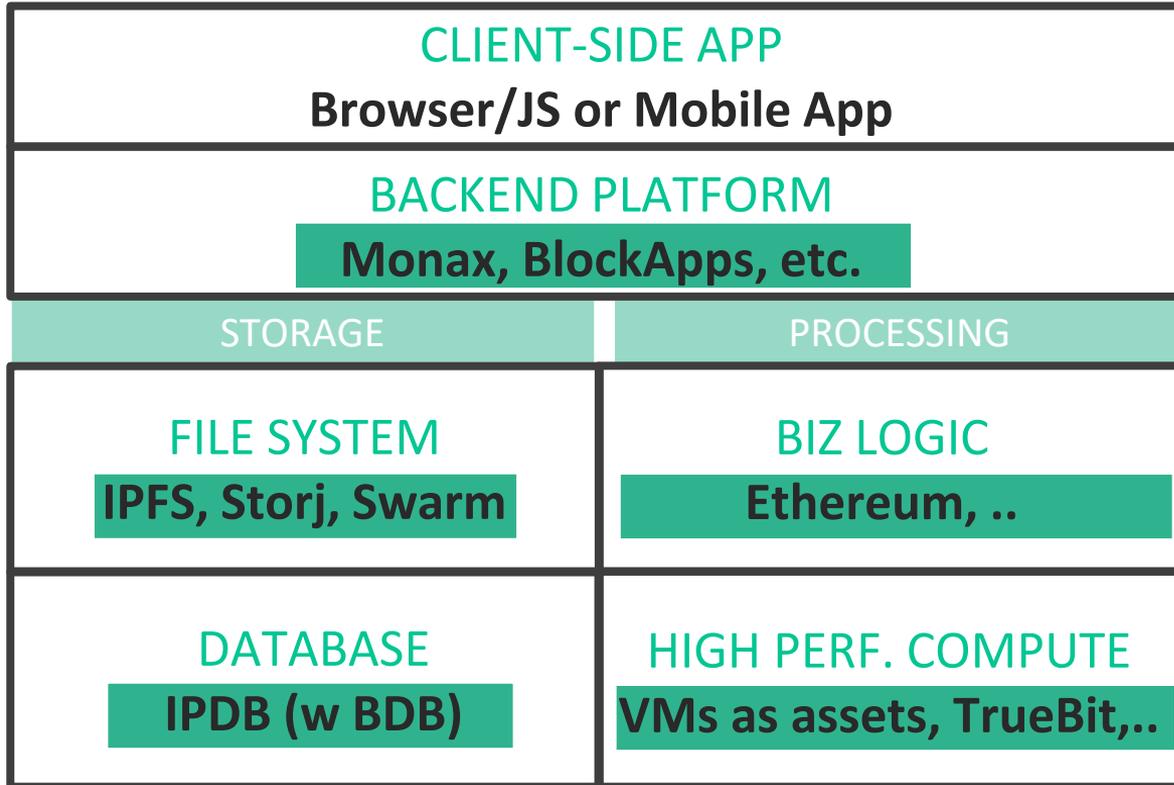
Decentralize DB of “Who Owns What”,



Keep the rest the same

CLIENT-SIDE APP Browser/JS or Mobile App	
BACKEND PLATFORM Django, Flask, Lambda	
STORAGE	PROCESSING
FILE SYSTEM HDFS, S3, maybe IPFS	BIZ LOGIC EC2
DATABASE MongoDB, IPDB (w BDB)	HIGH PERF. COMPUTE Nvidia GPU, Goog TPU, MapReduce, Spark

Decentralize All The Things



How to start doing AI * Blockchain

Get started



beta.bigchaindb.com/getstarted/



BIGCHAIN 

 We're hiring!

Get started

Features

Use Cases

Enterprise

Docs

Explore drivers, tools & documentation

Let's write a message into a planetary blockchain database

Type a message*

A transaction can contain a digital asset containing a message. Type something above to be sent in the asset.

Off you go

Beep, boop, waiting for your input...

Get started



beta.bigchaindb.com/getstarted/

Type a message*

Hello PyData

```
"fulfills": null,
"owners_before": [
  "FXB9jSF5DPV2RmN52MNeuLuYA68ww7AQakkUhwW7rMAB"
]
},
"metadata": {
  "metaDataMessage": "Hello PyData"
},
"asset": {
  "data": {
    "assetMessage": "Hello PyData"
  }
},
"version": "0.9"
}
```

Nicely done! You have just created an asset, sent it in a signed transaction and received the response. Phew.

[Check out your transaction on IPDB](#)

cURL

Python

Node.js

```
curl -X "POST" "https://test.ipdb.io/api/v1/transactions/" \
-d '${ message: "Hello PyData" }'
```

Get started



beta.bigchaindb.com/getstarted/

cURL

Python

```
from bigchaindb_driver import BigchainDB
from bigchaindb_driver.crypto import generate_keypair

bdb = BigchainDB('https://test.ipdb.io/api/v1/')

alice = generate_keypair()

tx = bdb.transactions.prepare(
    metadata={ "message": "Hello PyData" }
)

txSigned = bdb.transactions.fulfill(
    tx,
    private_keys=alice.private_key
)

bdb.transactions.send(txSigned)
```

Get started



beta.bigchaindb.com/getstarted/

cURL

Python

```
from bigchaindb_driver import BigchainDB
from bigchaindb_driver.crypto import generate_keypair

bdb = BigchainDB('https://test.ipdb.io/api/v1/')

alice = generate_keypair()

tx = bdb.transactions.prepare(
    metadata={ "message": "Hello PyData" }
)

txSigned = bdb.transactions.fulfill(
    tx,
    private_keys=alice.private_key
)

bdb.transactions.send(txSigned)
```

To bring in the AI:
`import sklearn`

...

Conclusion

The mist

The world's most
valuable resource

- Blockchains can *really* help AI
- **It's all about the data**
 - Getting the data
 - Getting *good* data - with provenance
- (Plus those pesky **AI DAOs**)
- You can get started with AI + blockchain easily