Blockchains for Artificial Intelligence

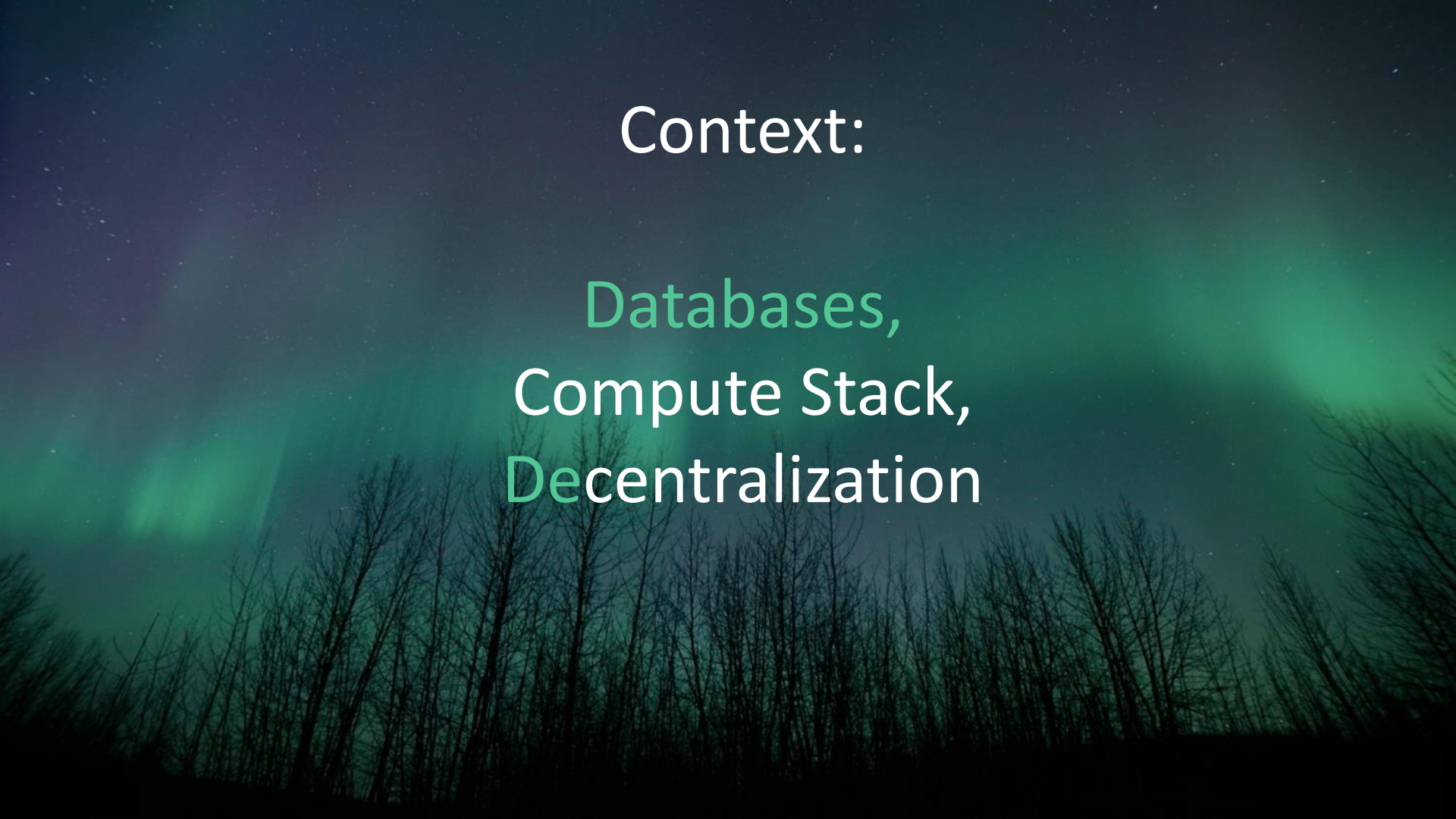
Trent McConaghy

@trentmc0

BIGCHAINDB

IPDB

INTERPLANETARY DATABASE

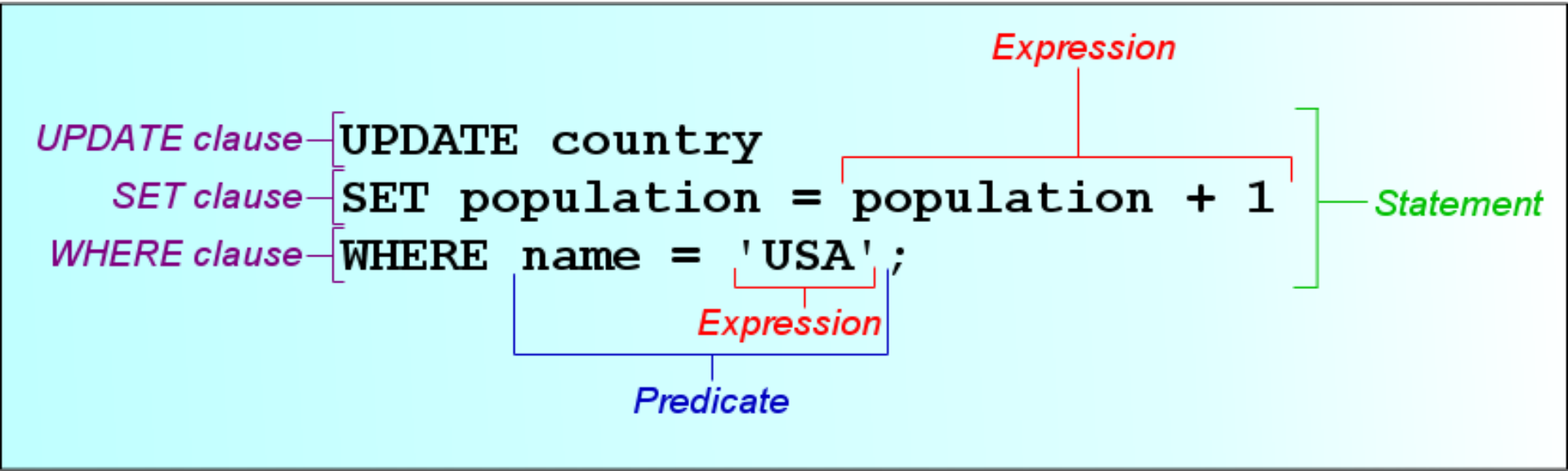

Context:
Databases,
Compute Stack,
Decentralization

What's the difference between a database and a spreadsheet? Querying. From M's of records, find the relevant ones.

1 Line of standard code, optimized

Vs

50-500 lines of slow custom code, unoptimized



The first “Blue Ocean” DBs: Relational DBs

Benefits: powerful structured querying

Winner: Oracle, 80s and 90s

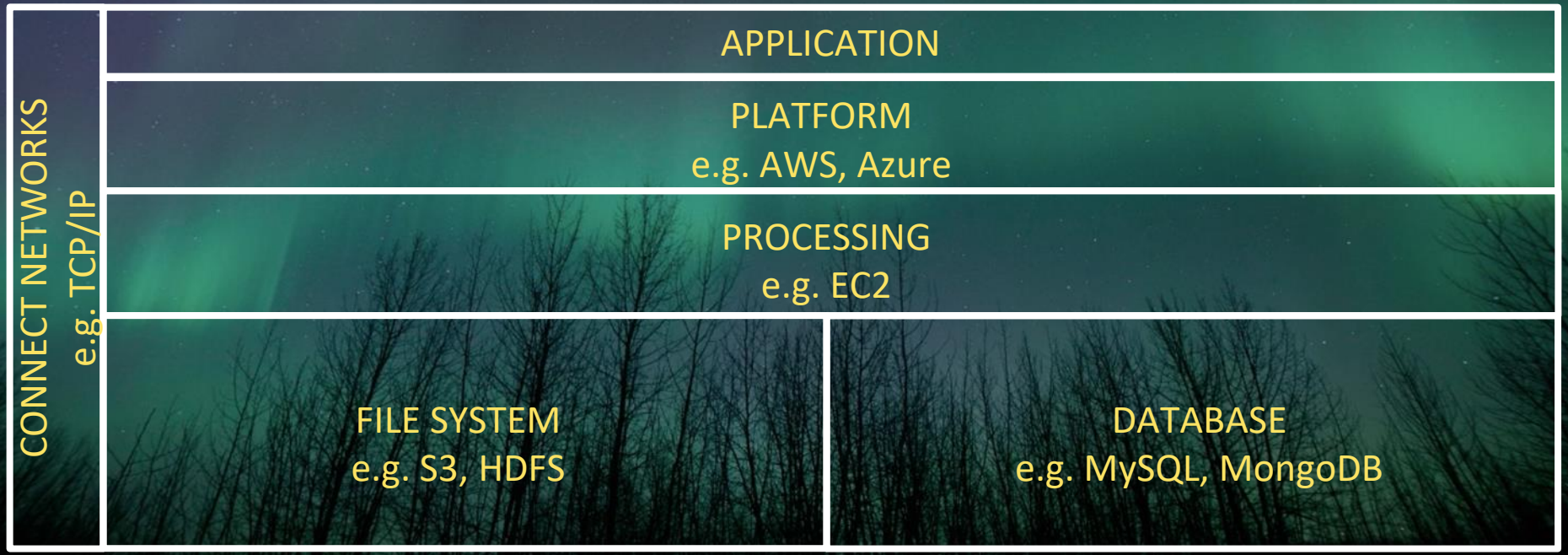

The next “Blue Ocean” DB:
Website-ready DBs
New benefits: lightweight for startups
Winner: MySQL, early 2000s

The next “Blue Ocean” DB: Distributed / NoSQL DBs

New benefits: “Big data” scale, flexible schemas

Winner: MongoDB, late 2000s-now

Databases in status quo compute stack



Towards a **decentralized** compute stack

CONNECT NETWORKS
e.g. TCP/IP

APPLICATION

PLATFORM
e.g. AWS, Azure

PROCESSING
e.g. EC2

FILE SYSTEM
e.g. S3, HDFS

DATABASE
e.g. MySQL, MongoDB
Bitcoin blockchain?

Towards a **decentralized** compute stack

Towards a **decentralized** compute stack

CONNECT NETWORKS
e.g. TCP/IP, **Interledger**

APPLICATION

PLATFORM
e.g. AWS, Azure, **Eris/Monax, BlockApps**

PROCESSING
e.g. EC2, **Ethereum, Hyperledger, Tendermint, Lisk**

FILE SYSTEM
e.g. S3, HDFS, **IPFS**

DATABASE
e.g. MySQL, MongoDB
???

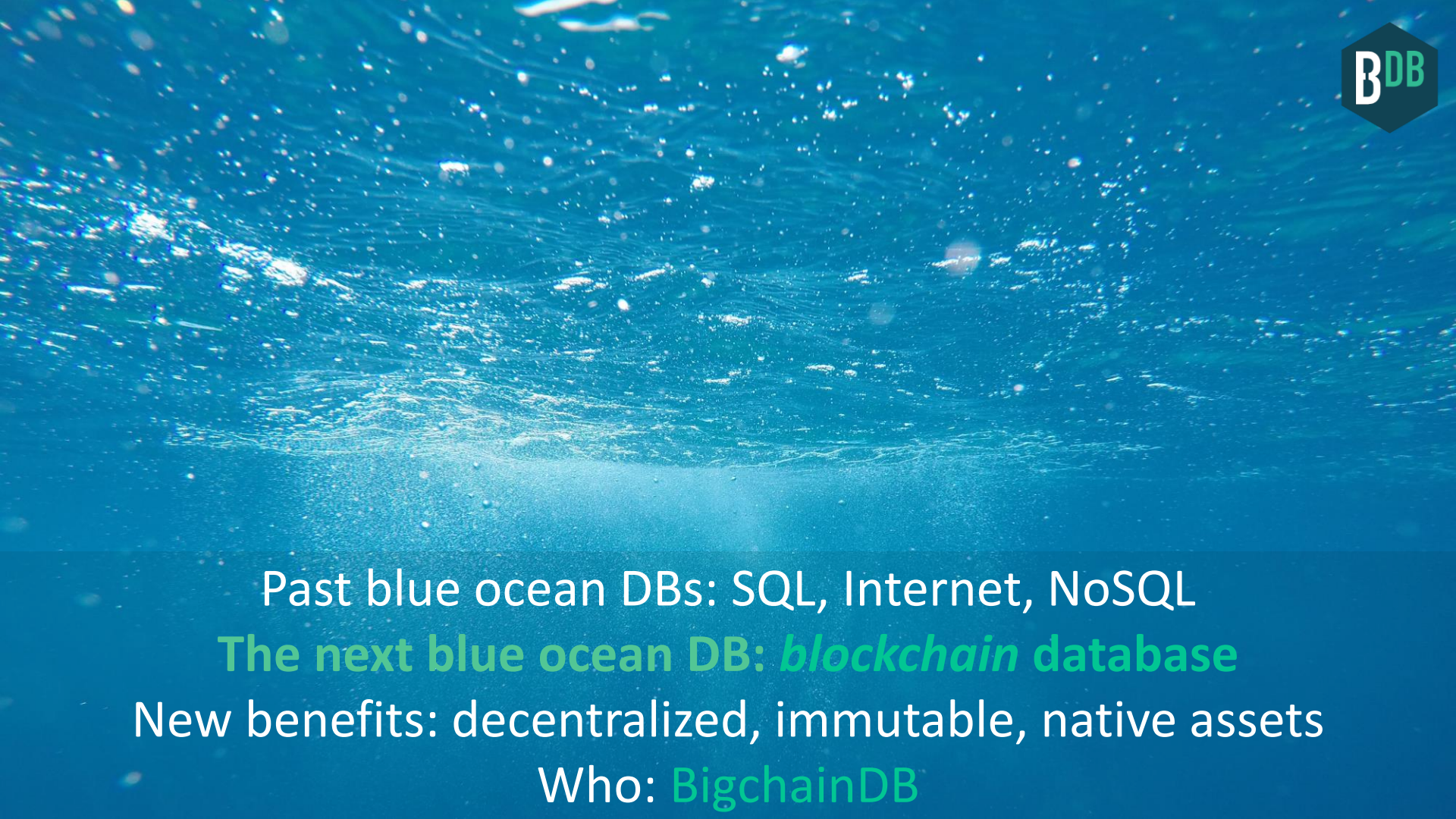
e-gold / e-cash
Bitcoin, zcash

Benefits of blockchain technology

Decentralized / shared control

Immutability / audit trail

Tokens / exchanges



Past blue ocean DBs: SQL, Internet, NoSQL

The next blue ocean DB: *blockchain* database

New benefits: decentralized, immutable, native assets

Who: BigchainDB

How to build a scalable blockchain database (BigchainDB)

1. **Start with an enterprise-grade distributed DB**
2. **Engineer in blockchain characteristics**
   - Decentralized – each node has different sysadmin, ..
   - Immutable – Merkle chain of hashes, continuous backups, ..
   - Native assets – have private key = own, transactions, ..

IPDB

INTERPLANETARY DATABASE

IPDB = a public global blockchain database

Towards a **decentralized** compute stack

CONNECT NETWORKS
e.g. TCP/IP, **Interledger**

APPLICATION

PLATFORM
e.g. AWS, Azure, **Eris/Monax, BlockApps**

PROCESSING
e.g. EC2, **Ethereum, Hyperledger, Tendermint, Lisk**

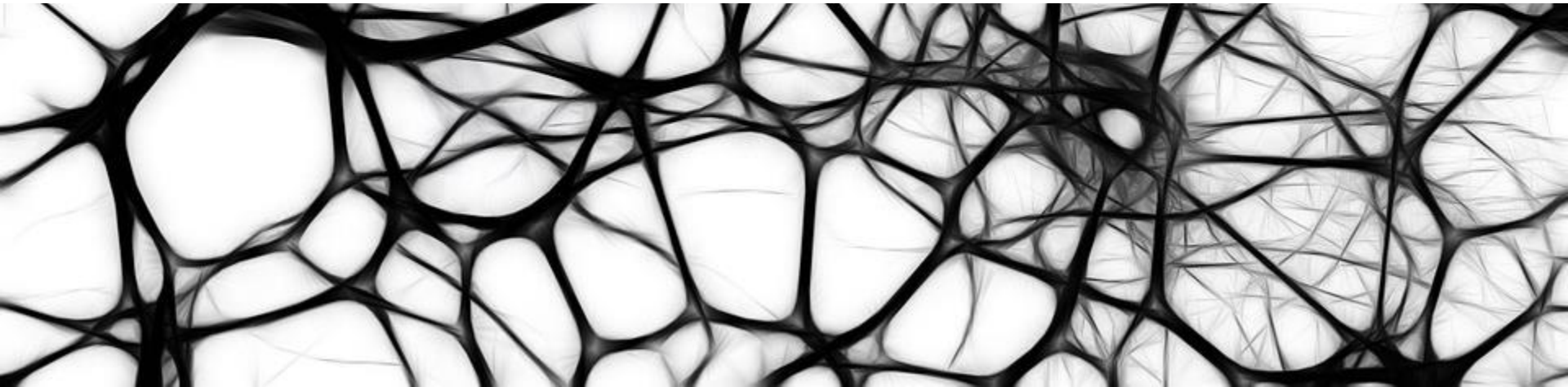
FILE SYSTEM
e.g. S3, HDFS, **IPFS**

DATABASE
e.g. MySQL, MongoDB
BigchainDB + IPDB

e-gold / e-cash
Bitcoin, zcash

Introduction to AI,

By Example



What's Artificial Intelligence (AI)?

Original:

AI: "A machine that can replicate human cognitive behavior" [Turing test]

Now:

AI: "A machine that can perform a cognitive task, that was previously only possible with a human"
[AlphaGo]

+

"A machine that can perform a non-analytical information processing task, at speed / accuracy / capacity *not possible by a human*."
[Driving Moore's Law]

[ibm.com]

What's Artificial Intelligence (AI)?

AI has a toolbox of ways to solve:

- Classification
- Regression
- Density estimation
- Rare event estimation
- Knowledge extraction
- Optimization
- Active analytics
- Structural design
- Control / autonomy
- …

Let's explore these more…

For application to:

- Operations
- Product
- Services

Example biz benefits:

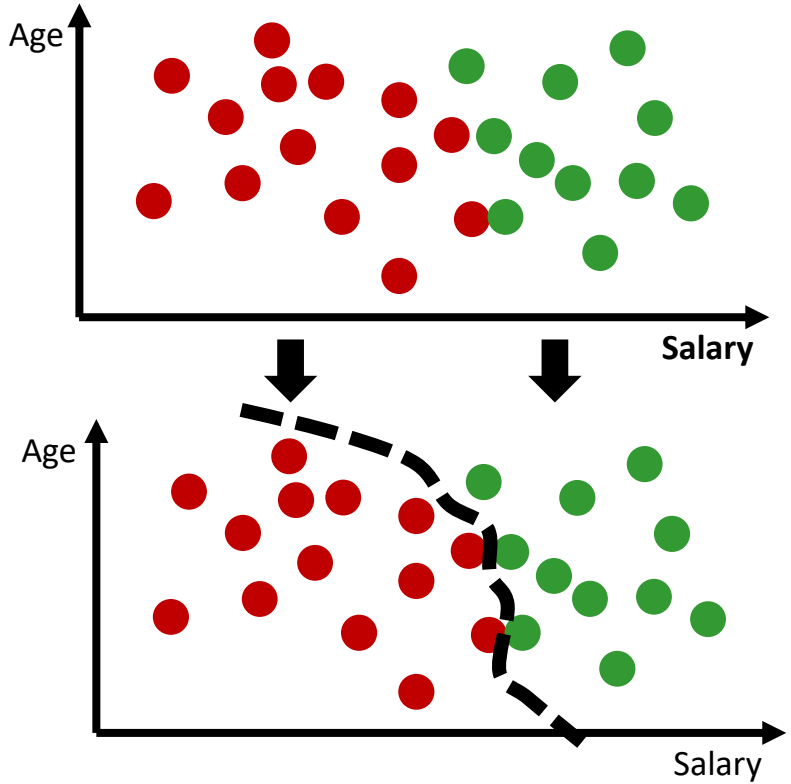
- Shorter time to market
- Lower cost
- Higher yield
- New features possible
- New products possible

Classification, in 2D

Use case: credit card applications

Credit profile:

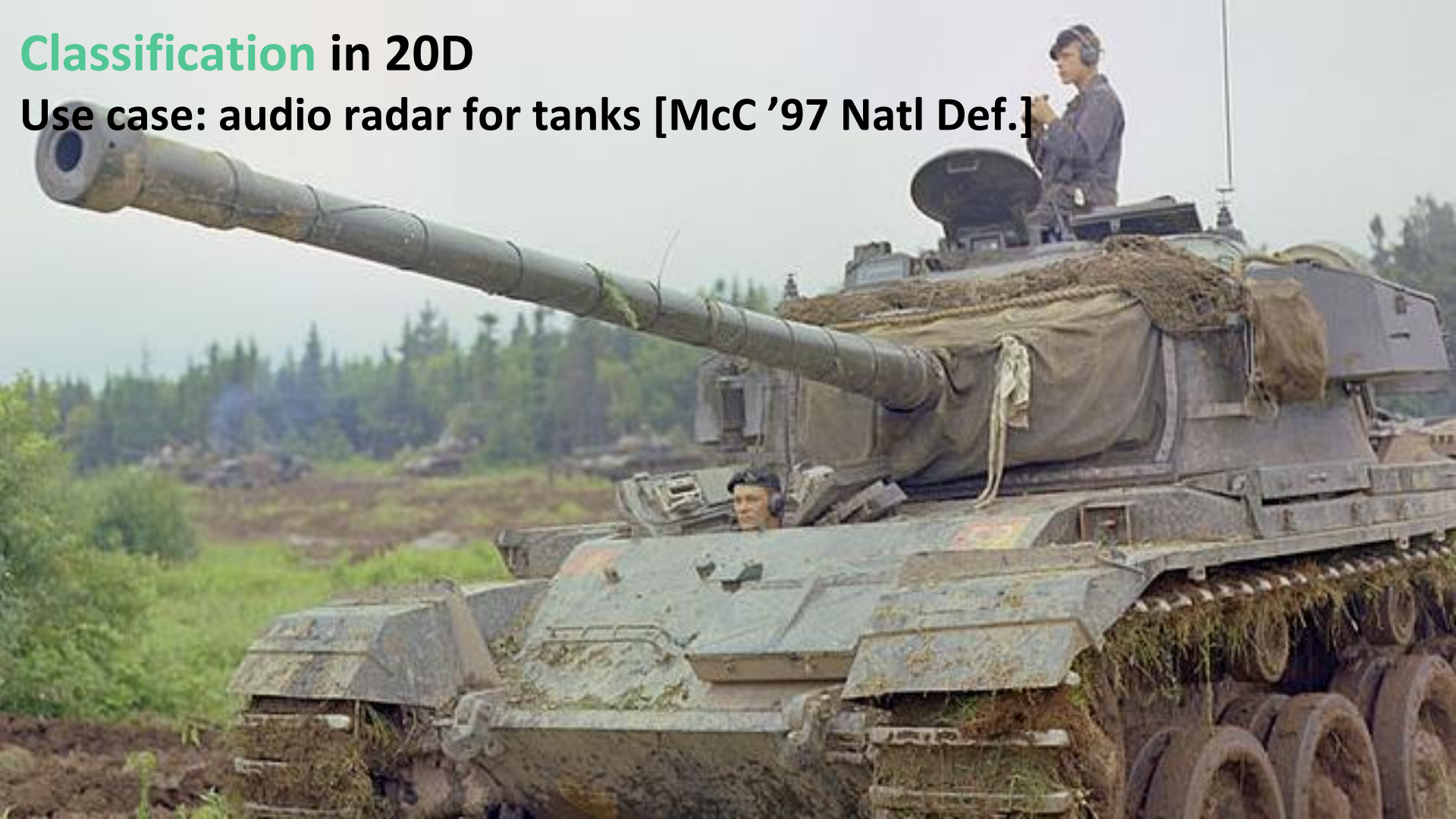
- **Paid bills**
- **Didn't pay**



How: linear classifiers, neural networks / deep learning, support vector machines, boosted trees, random forests / bagged trees

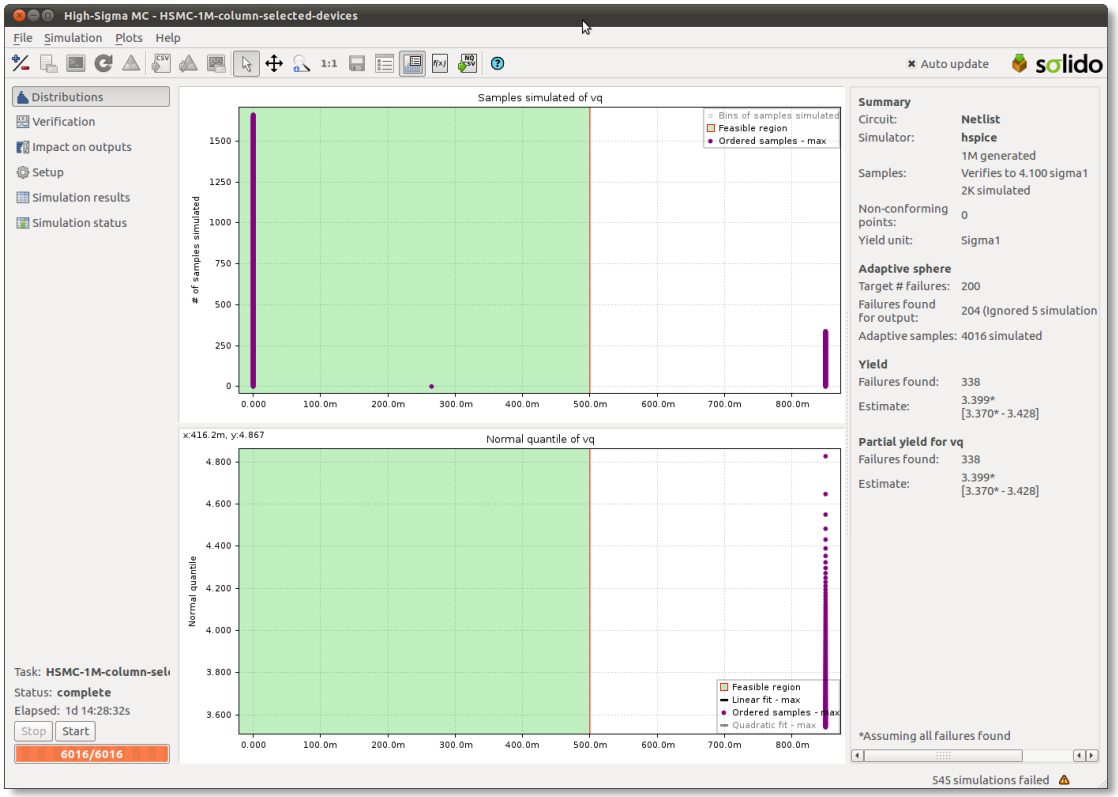
Classification in 20D

Use case: audio radar for tanks [McC ’97 Natl Def.]

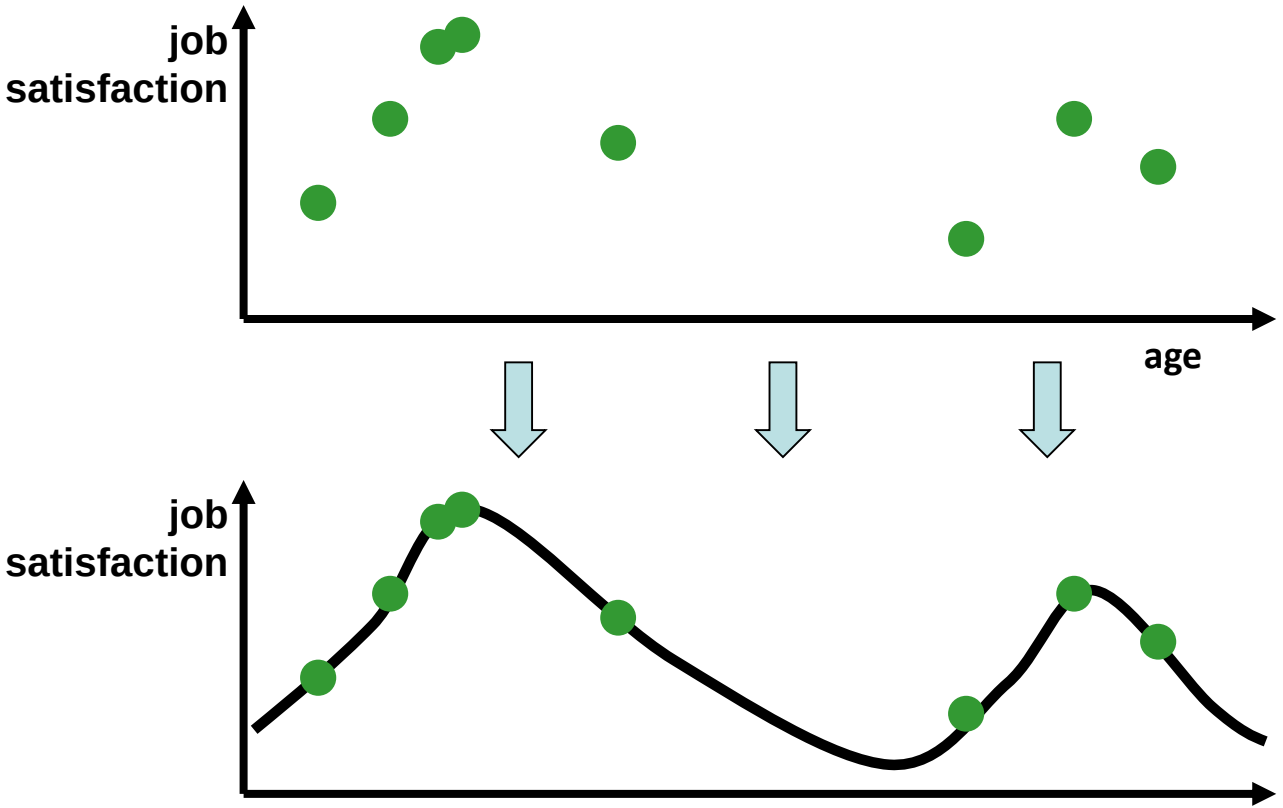


Classification in 10,000 D

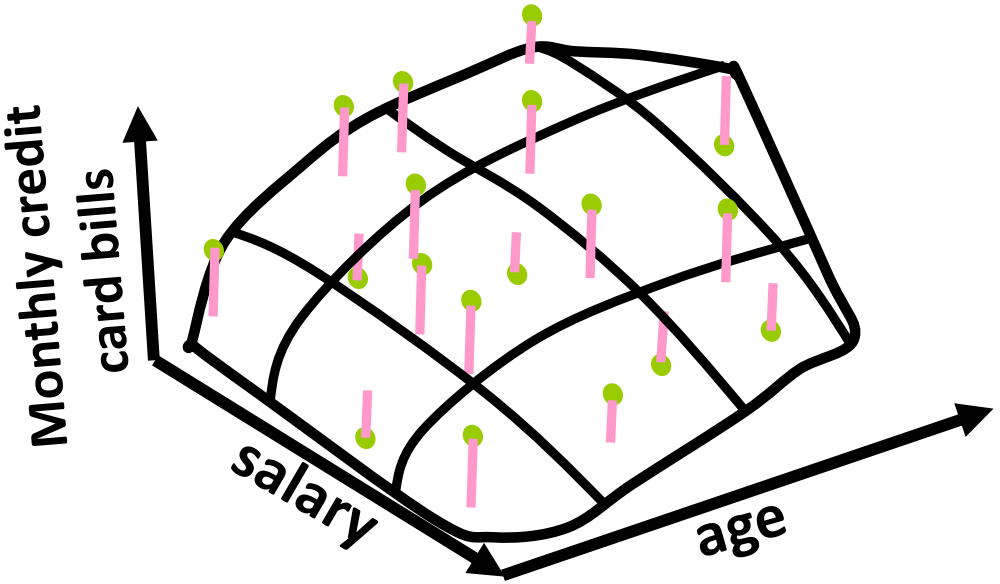
Use case: yield of memory columns [McC ’12 Solido]



Regression, in 1D

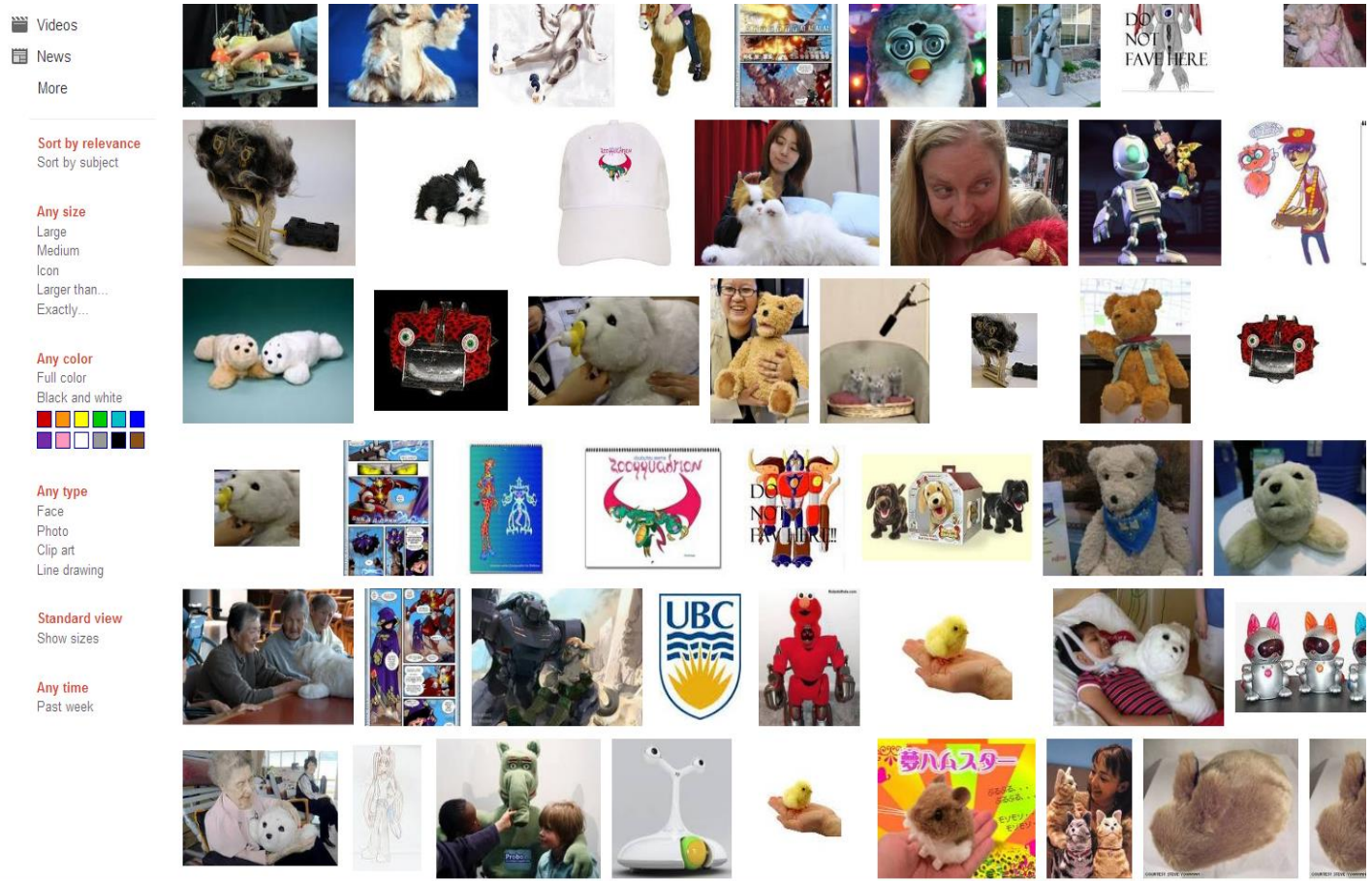


Regression, in 2D



Regression in 10,000D

Use case: reverse image search [Google]



Regression in 10,000D

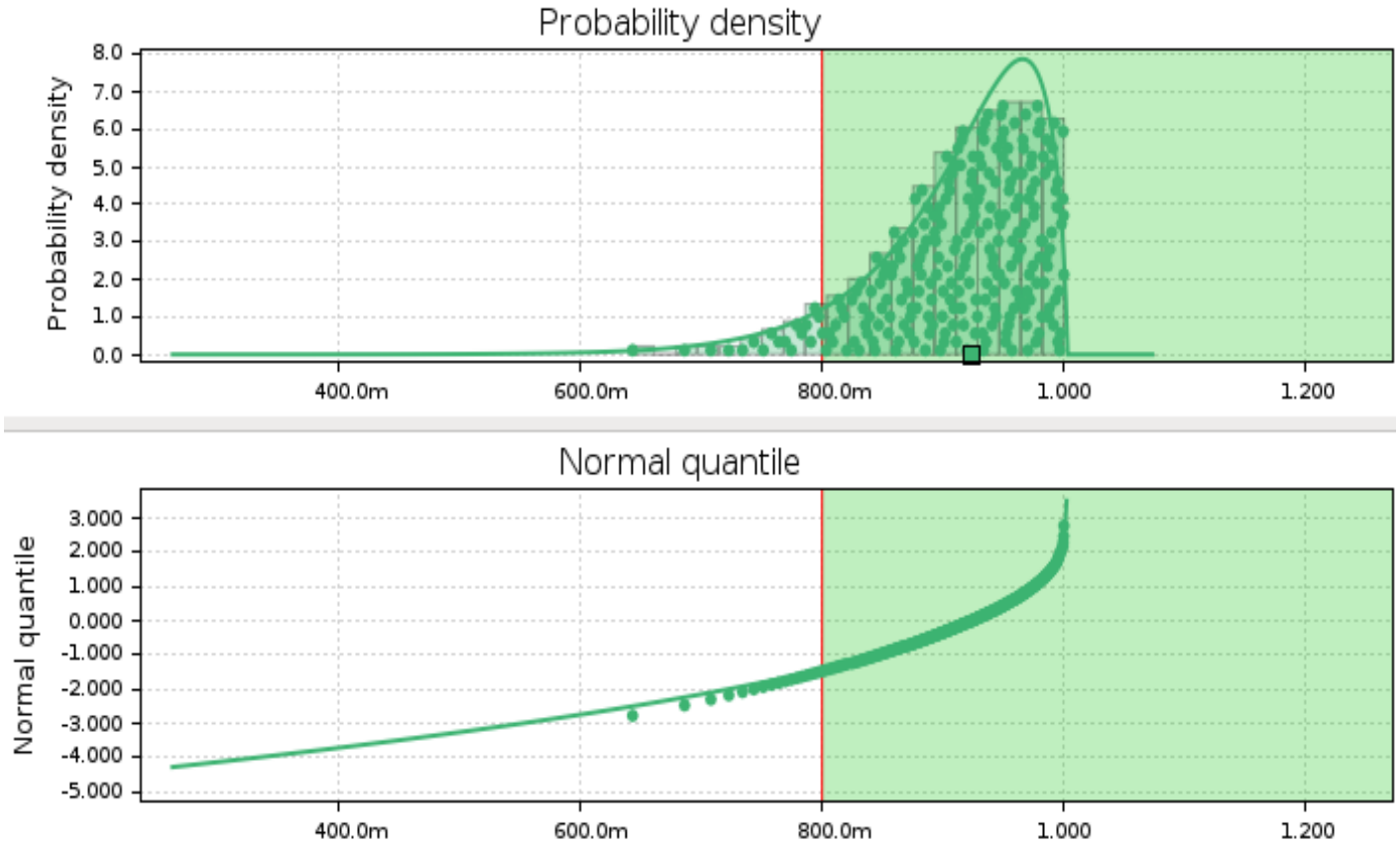
Use case: reverse image search *with provenance*

[McC ‘15 whereonthe.net // BigchainDB]

“Where’s Doge?”

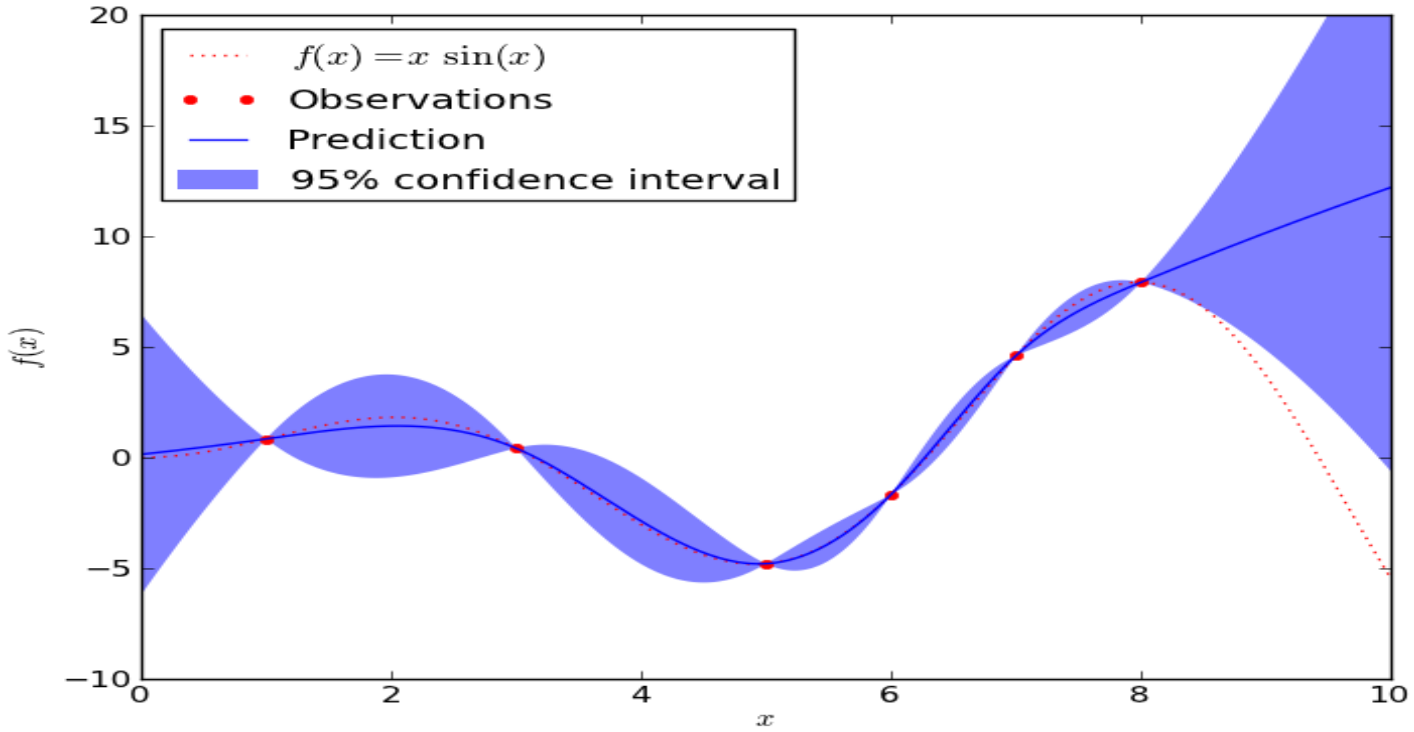
Density Estimation // Fitting a PDF

Use case: 3-sigma circuit analysis [McC '08 Solido]



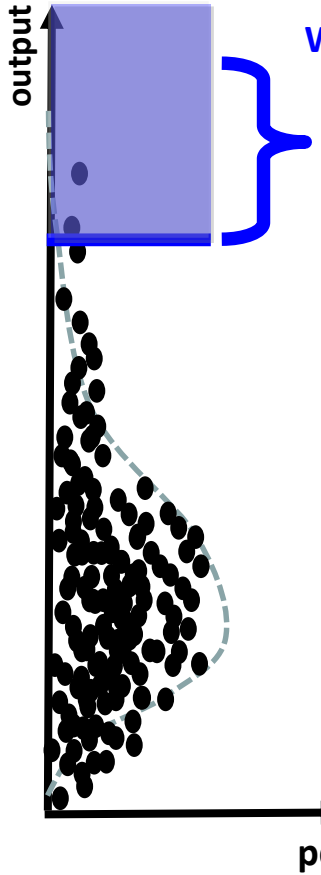
Density estimation // Fitting a PDF

Use case: regression with Confidence Intervals, for predictive circuit design [McC ’11 Solido]

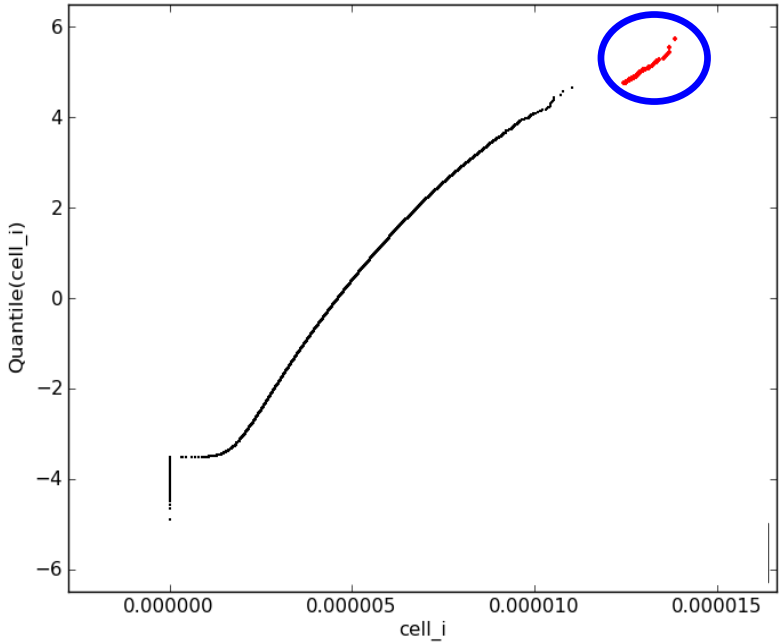


Rare event estimation / “Black swan” simulation

Use Case: 6-sigma analysis of bitcells [McC ‘09 Solido]

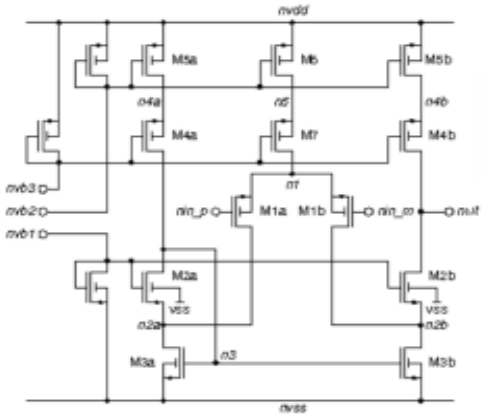


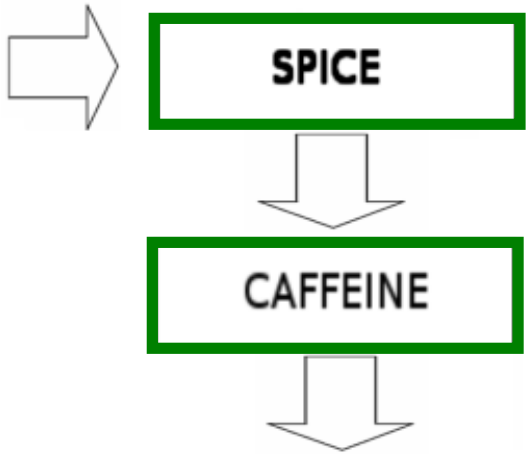
What is probability of this (rare) event happening?

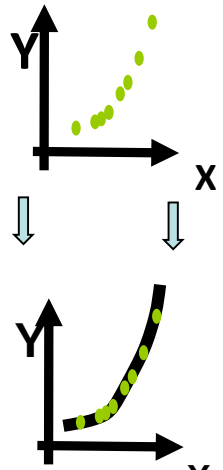


Knowledge extraction

Use case: scientific modeling of analog ccts [McC ’05 KUL]



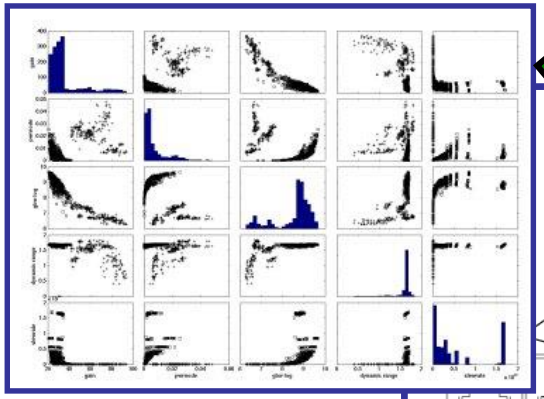


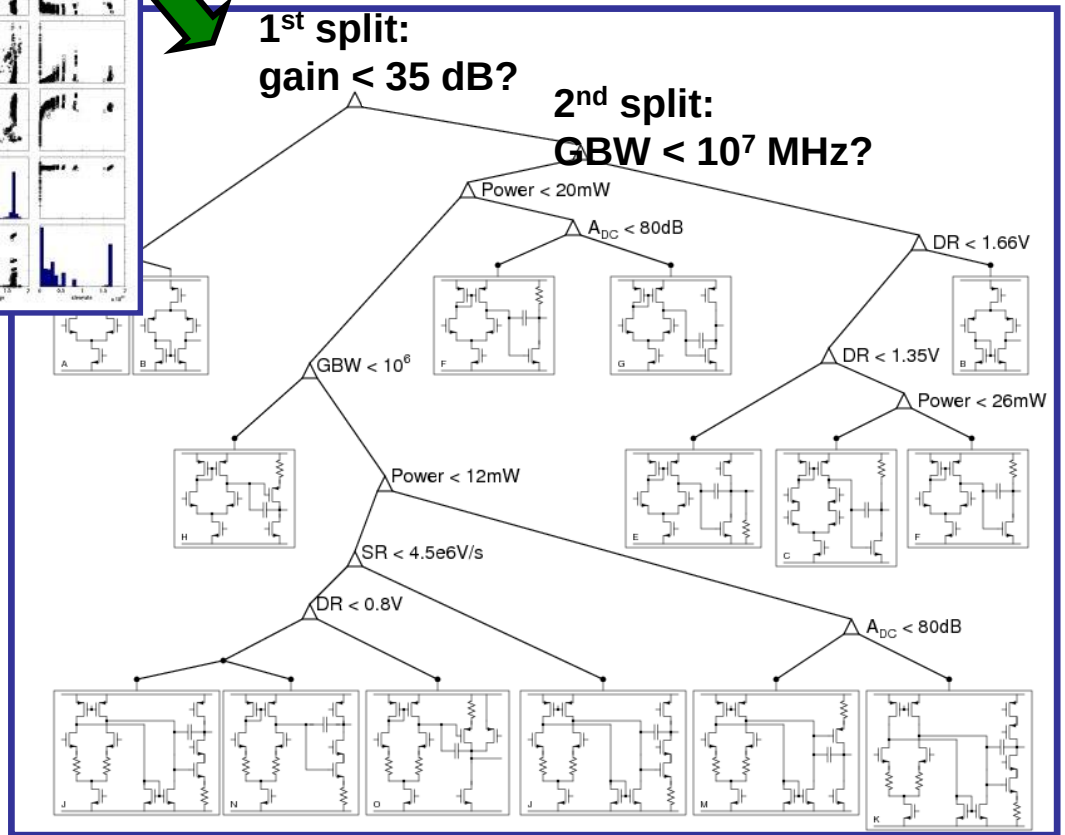


Perf.	Expression
A_{LF}	-10.3 + 7.08e-5 / id1 + 1.87 * ln(-1.95e+9 + 1.00e+10 / (vsg1*vsg3)+ 1.42e+9 *(vds2*vsd5) / (vsg1*vgs2*vsg5*id2))
f_u	10^(5.68 - 0.03 * vsg1 / vds2 - 55.43 * id1+ 5.63e-6 / id1)
PM	90.5 + 190.6 * id1 / vsg1 + 22.2 * id2 / vds2
V_{offset}	- 2.00e-3
SR_p	2.36e+7 + 1.95e+4 * id2 / id1 - 104.69 / id2 + 2.15e+9 * id2 + 4.63e+8 * id1
SR_n	- 5.72e+7 - 2.50e+11 * (id1*id2) / vgs2 + 5.53e+6 * vds2 / vgs2 + 109.72 / id1

Knowledge extraction

Use Case: Topology decision tree [McC '08 KUL]

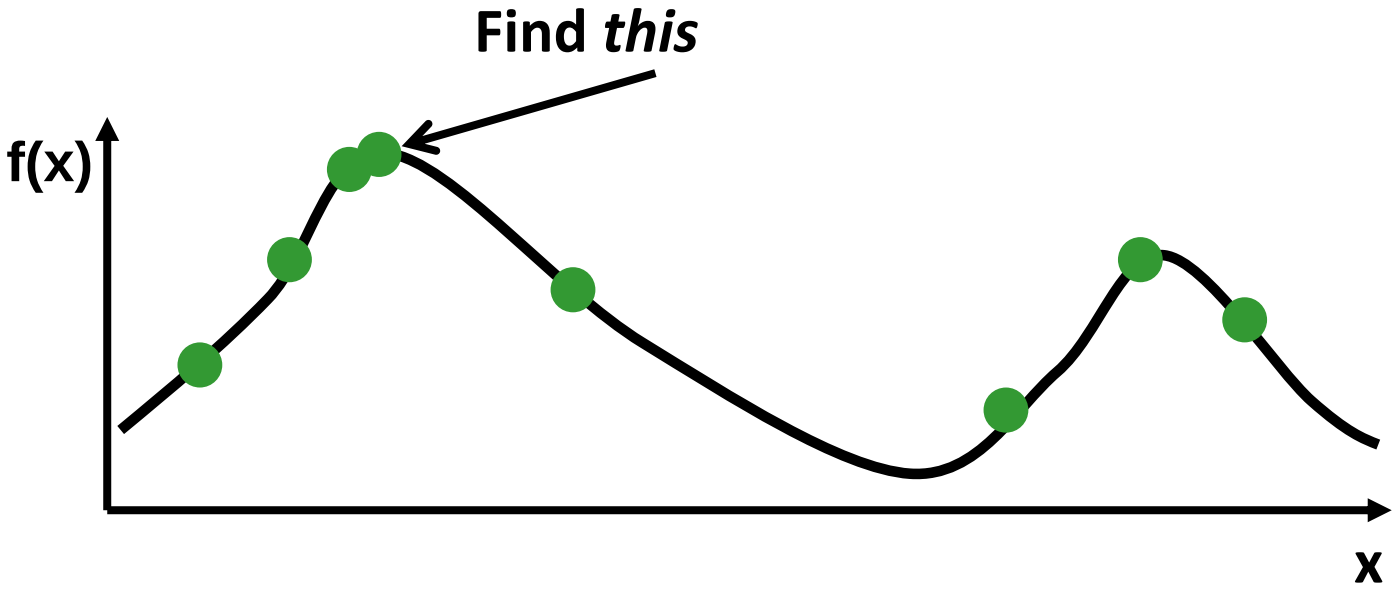




solido
DESIGN AUTOMATION

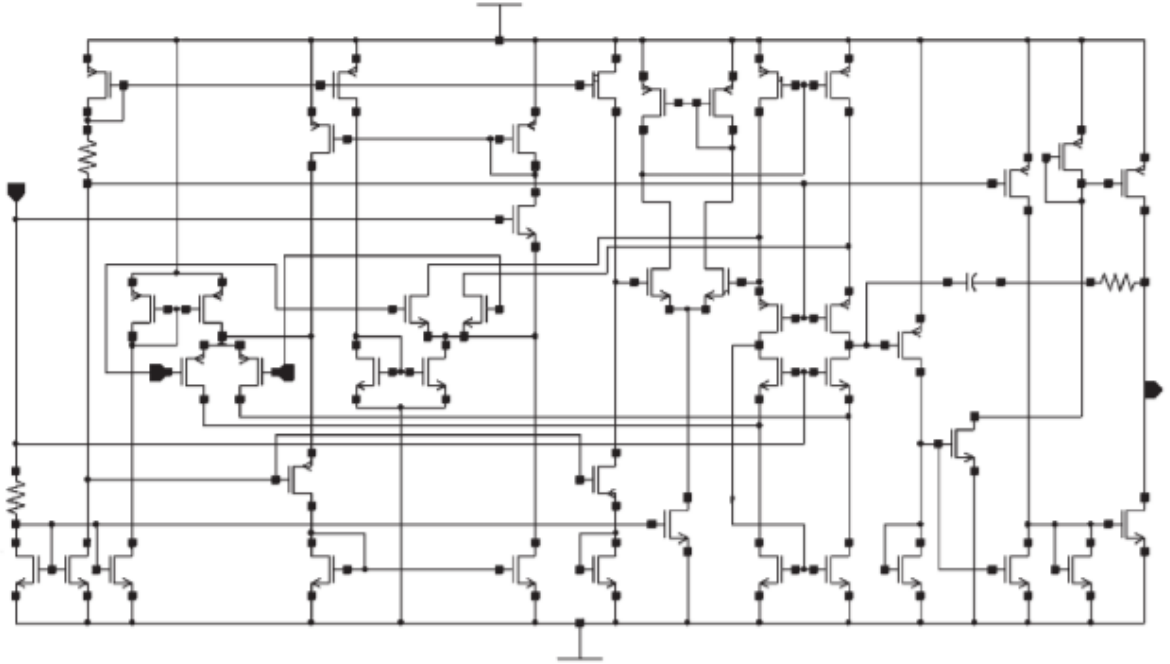
Optimization

"Find the x that maximizes f(x)"
(With as few evaluations of f(x) as possible)



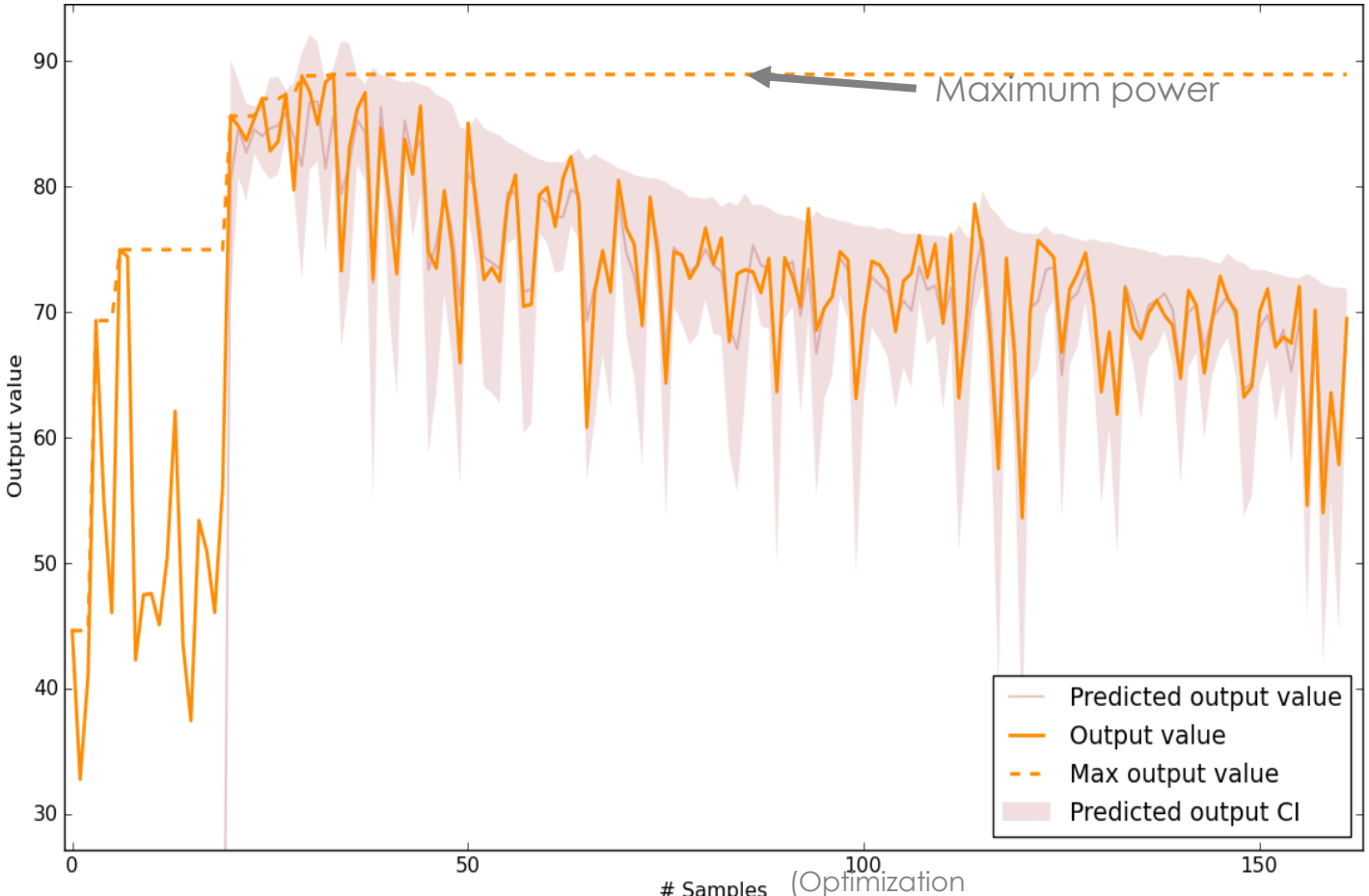
Optimization

Use Case: Optimize perf. of lg. analog circuits [McC ‘01 ADA]



Optimization

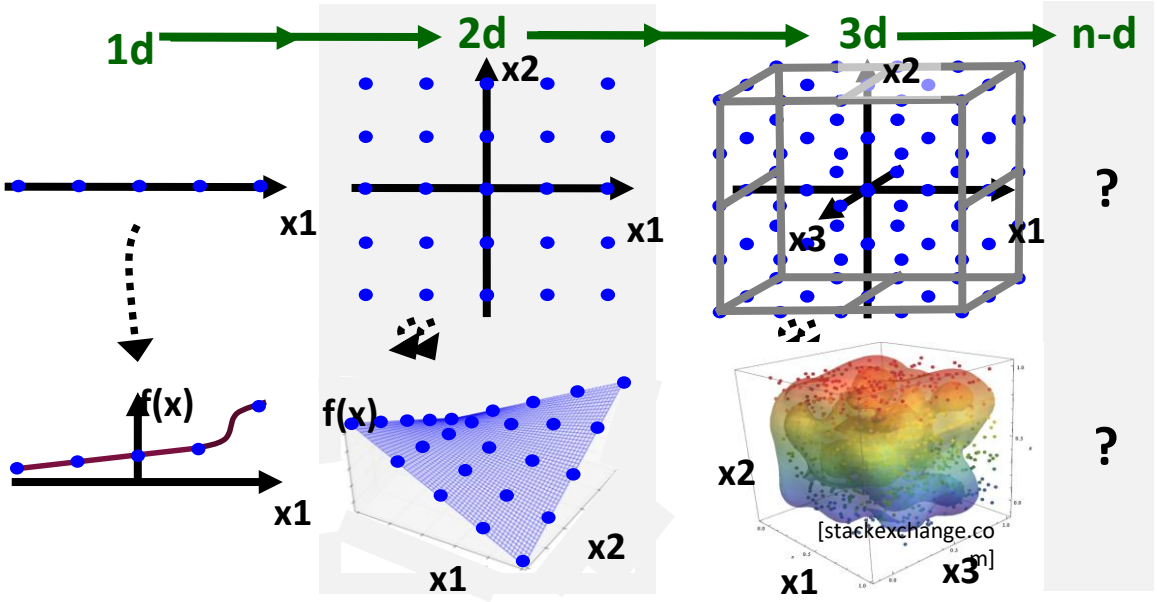
Use Case: verify circuits for worst-case perf. [McC '09 Solido]



(Optimization

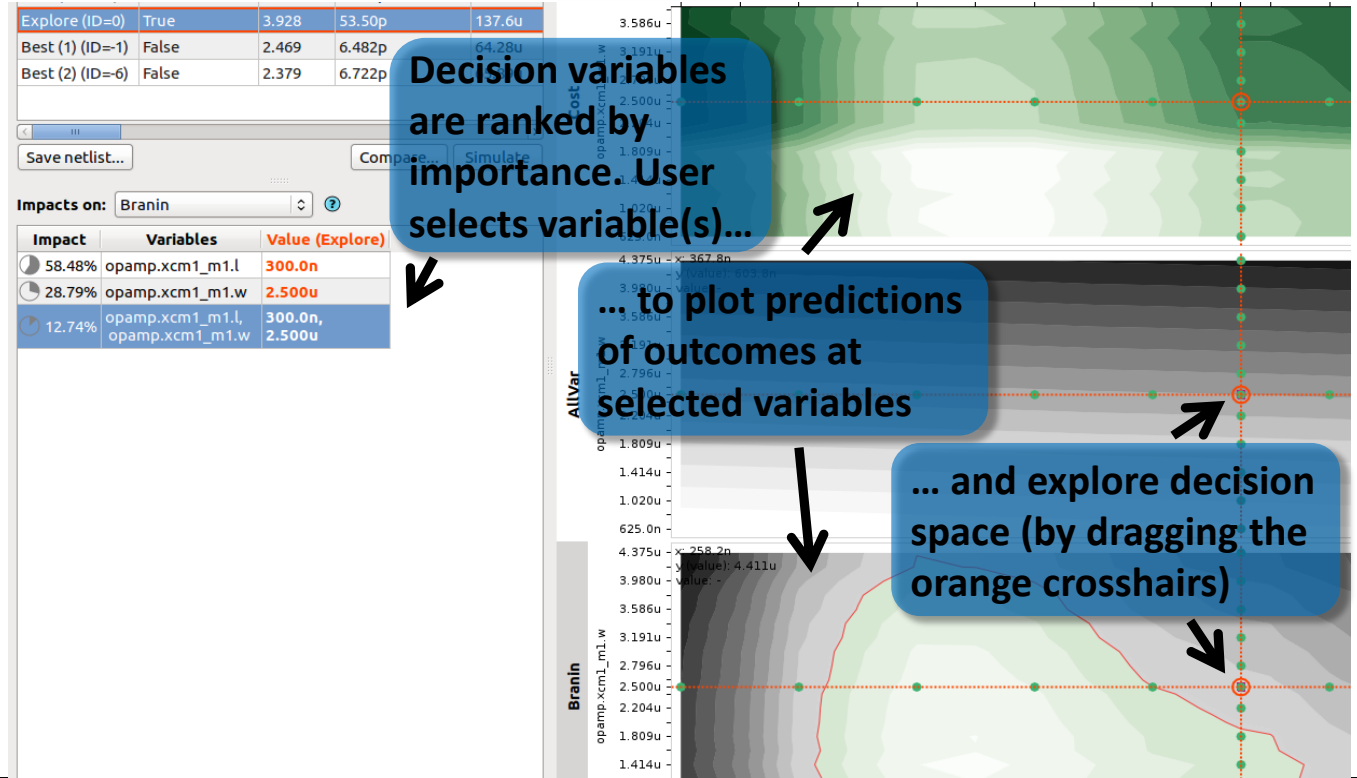
Active Analytics

“Find a mapping and visualize it, subject to high-d and expensive samples”



Active Analytics
Use Case: Interactive Circuit Design [McC ’12 Solido]

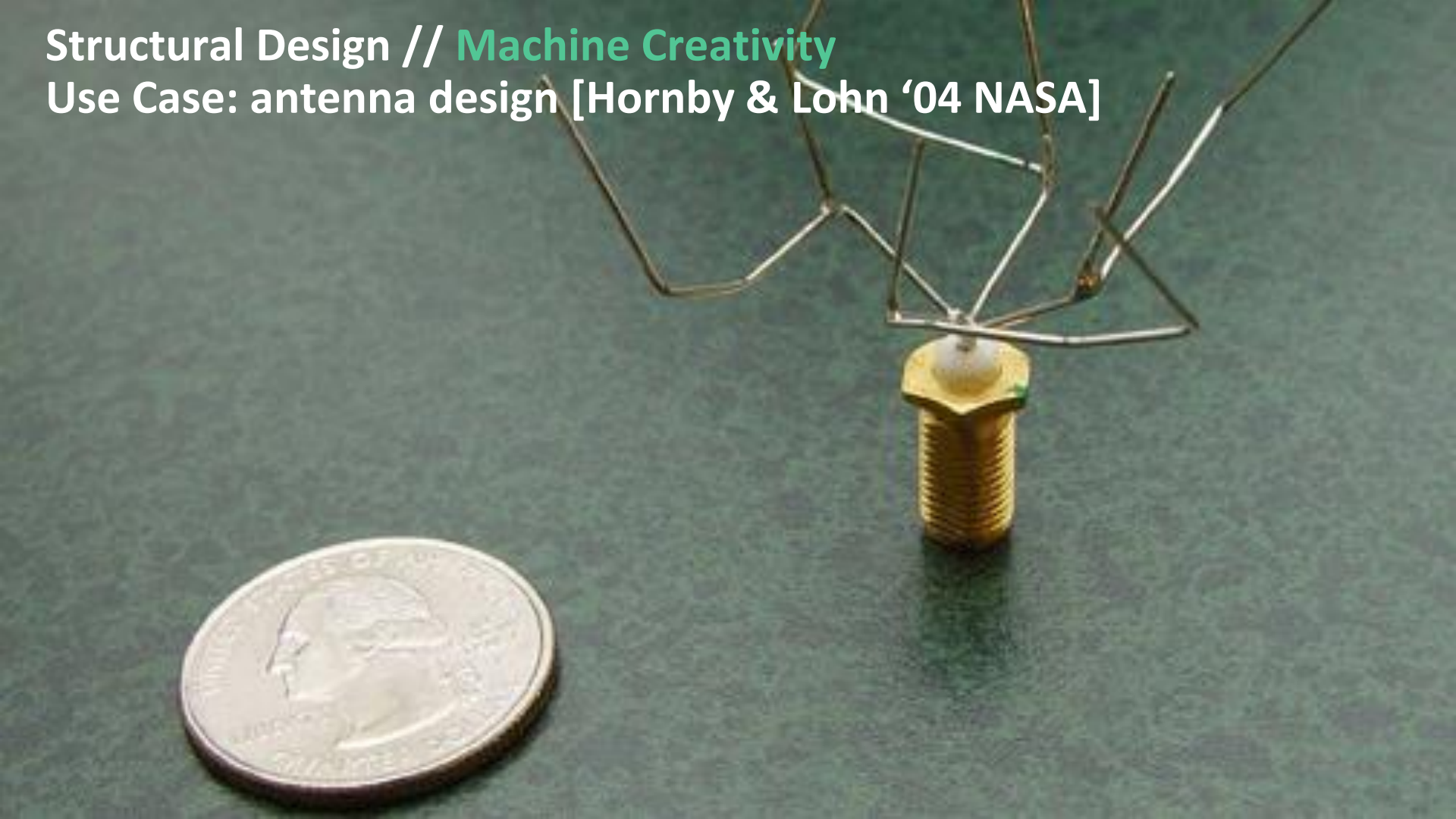
Under the hood, a machine learning engine adaptively samples the space of possible decisions, and measures outcomes.



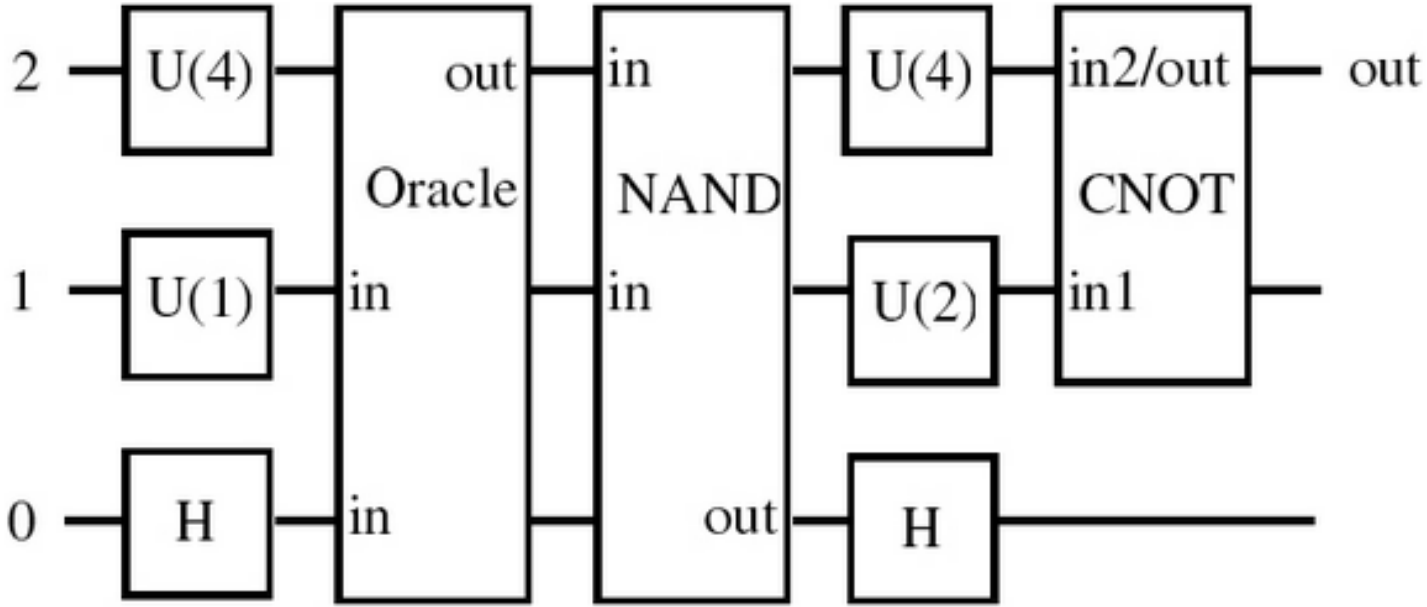
More Applications of Optimization & Analytics

- Business Intelligence: optimize churn & other key performance indicators (KPIs)
- Big data infrastructure: optimize reliability / uptime, minimize power consumption, ..
- Telecom infrastructure: capital & resource allocation
- Internet / mobile: auto SEO, optimize for app store placement (rank, profitability)
- ML modeling: Find optimal model meta-parameters (DeepNN, RF, SVM, ..), for application to computer vision etc.

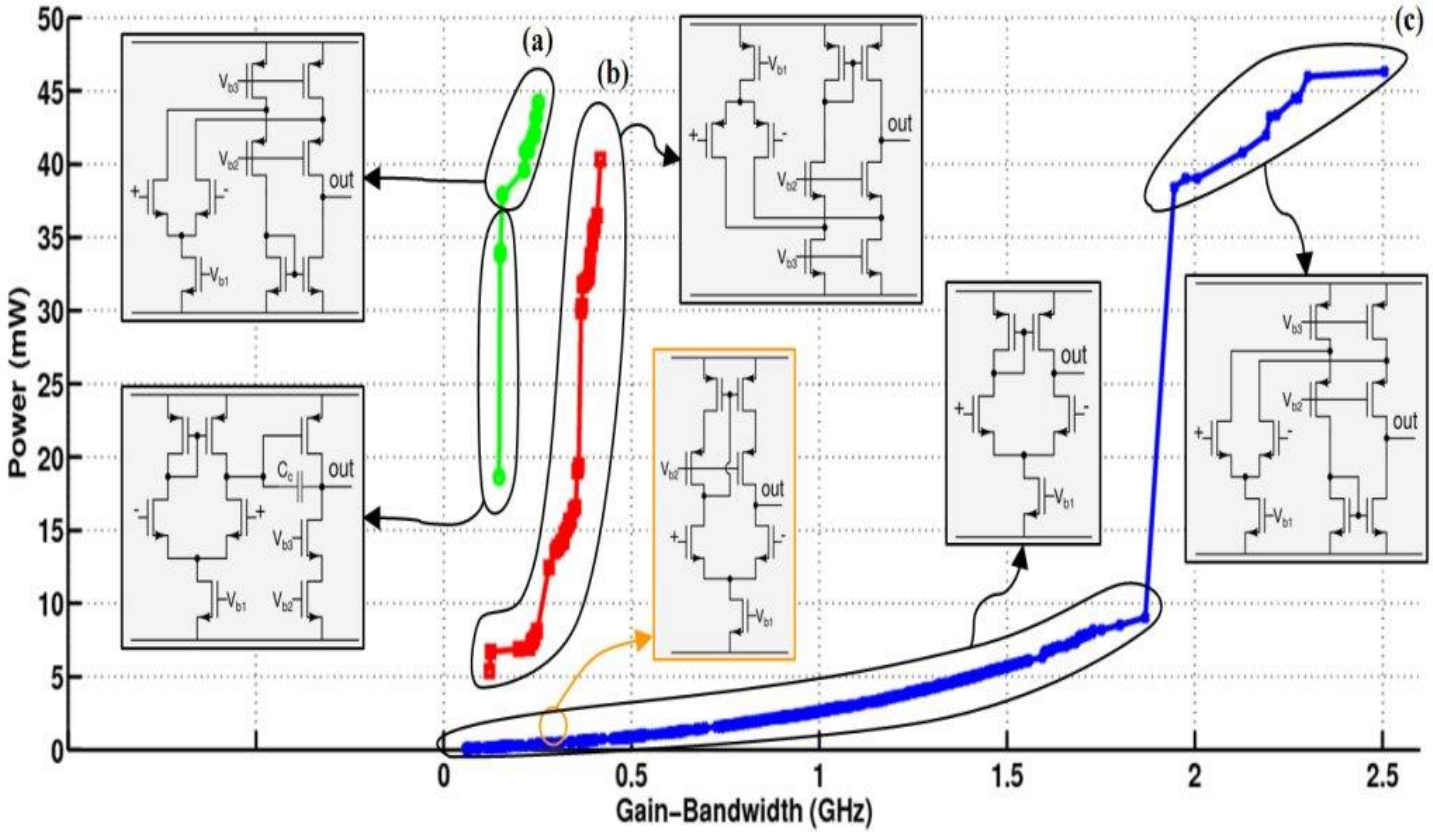
Structural Design // Machine Creativity
Use Case: antenna design [Hornby & Lohn ‘04 NASA]



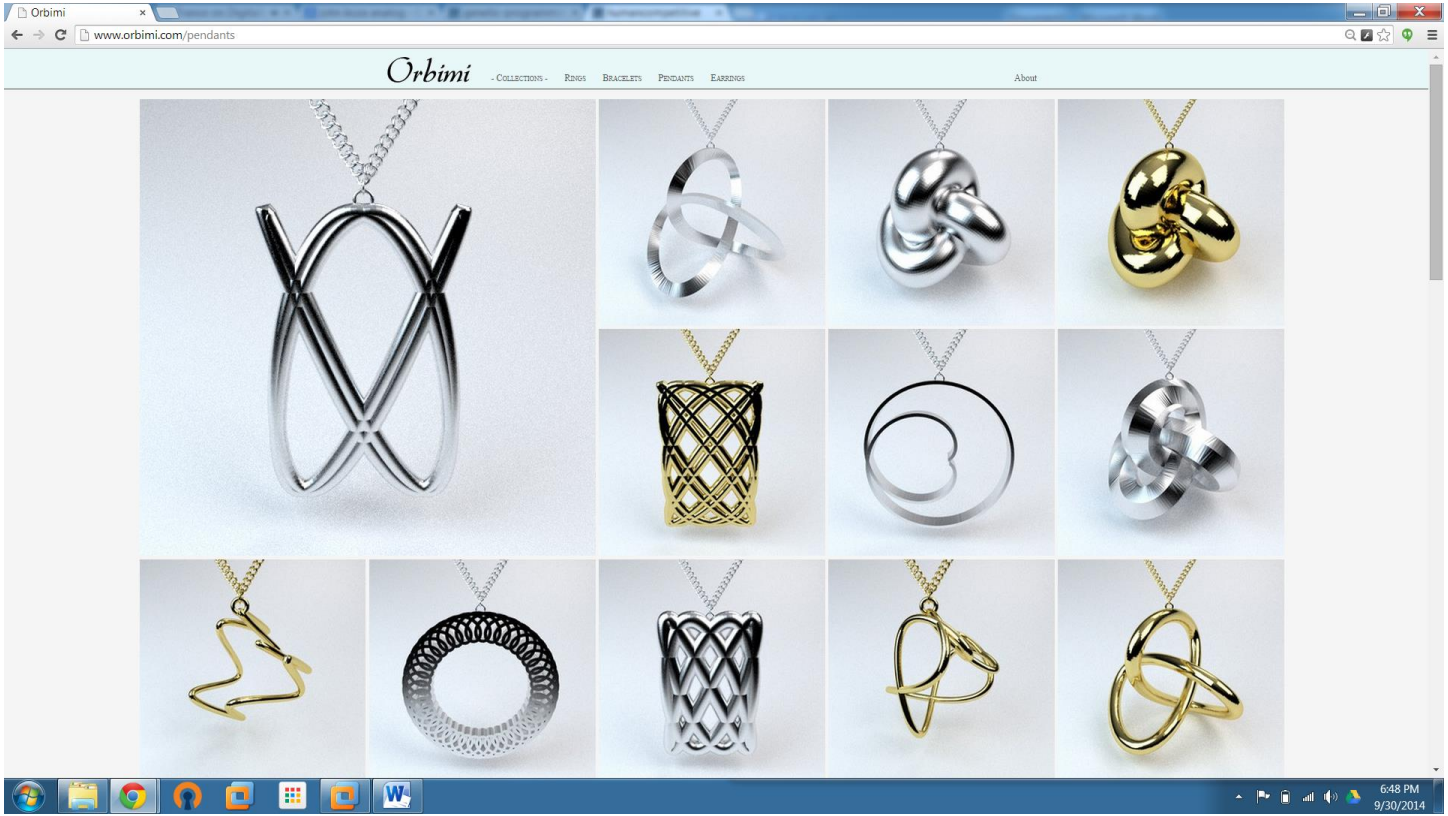
Structural Design // Machine Creativity
Use Case: quantum computing algorithm design [Spector ’04]



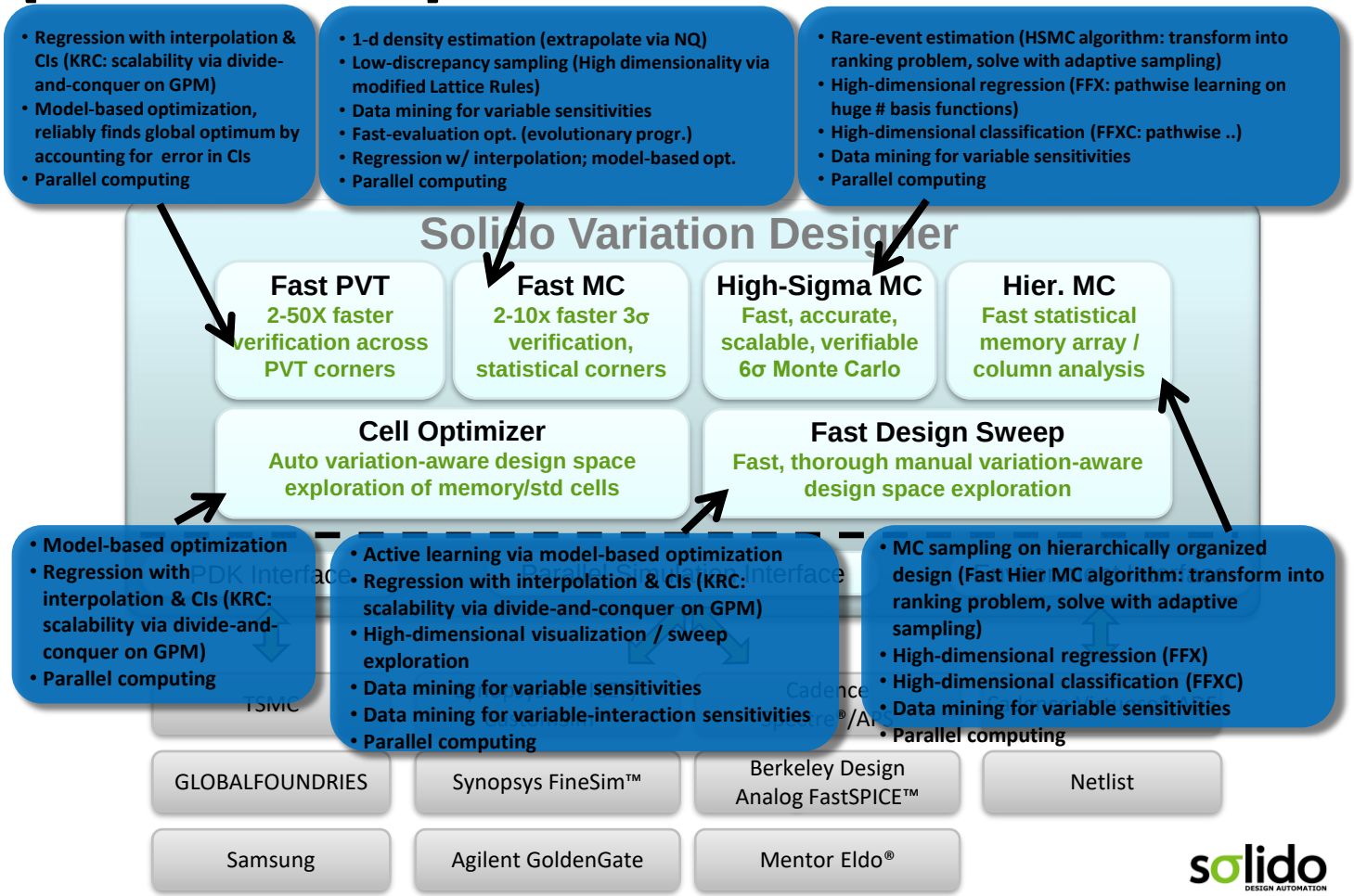
Structural Design // Machine Creativity
Use Case: circuit topology design [McC '06 KUL]



Structural Design // Machine Creativity
Use Case: jewelry design [Hornby ‘11 Orbimi]



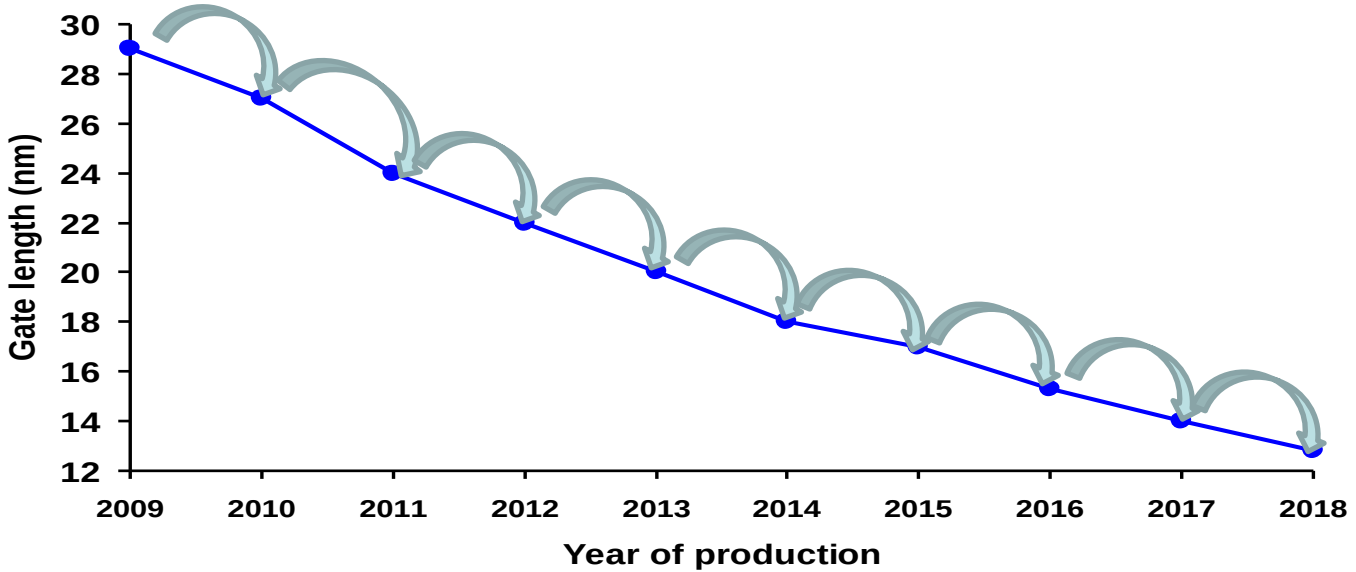
Example: A Single Product Suite With Broad Use of AI [McC Solido '05-13]



AI is driving Moore's Law
[Via Solido McC] and others

Solves:

- How to design with 1 billion devices?
- How to design when the physics gets crazier at every process generation?



Recap so far

AI has a toolbox of ways to solve:

- Classification
- Regression
- Density estimation
- Rare event estimation
- Knowledge extraction
- Optimization
- Active analytics
- Structural design
- Control / autonomy
- …

For application to:

- Operations
- Product
- Services

Example biz benefits:

- Shorter time to market
- Lower cost
- Higher yield
- New features possible
- New products possible

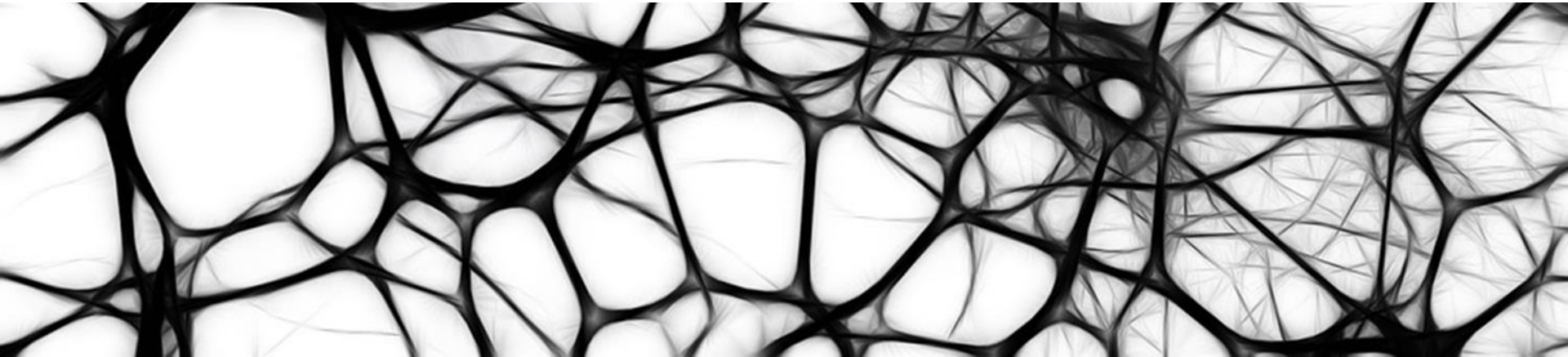
AI Sub-Fields

- machine learning
- neural networks / deep learning
- evolutionary computation (GAs, GP, PSO, ..)
- artificial general intelligence (AGI)
- pitfall to avoid: getting caught up in the latest buzzwords (it happens all the time.)

Strong relations to:

- statistics ("ML *is* modern statistics"), probability
- linear algebra ("flow of tensors" – TensorFlow)
- nonlinear programming, optimization
- control systems / cybernetics
- Monte Carlo methods
- philosophy, ethics (friendly AIs, AI rights, ..)

How can blockchains help AI?

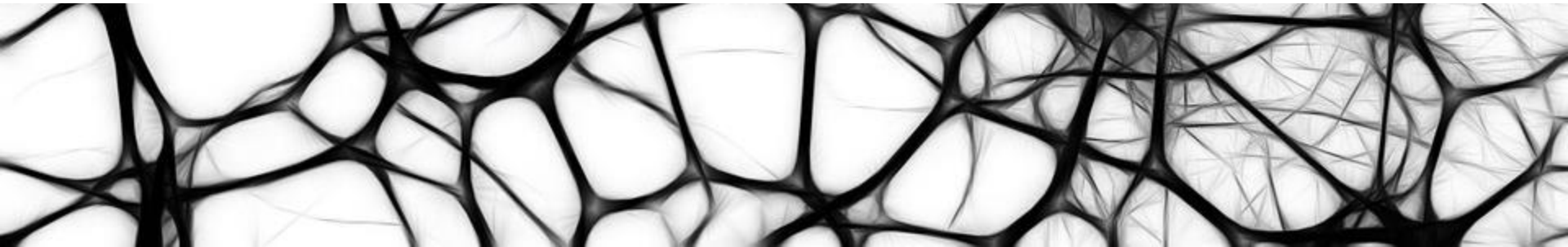


Work off of each of the benefits…

Decentralized / shared control

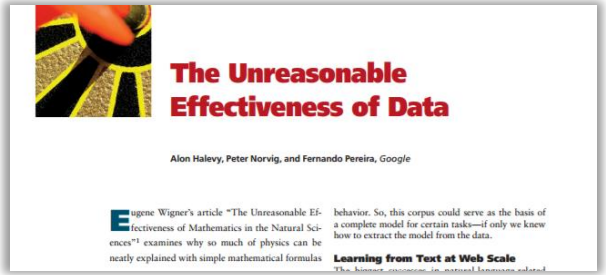
Immutability / audit trail

Tokens / exchanges



Decentralized / shared control encourages data sharing

More data → better models



The Unreasonable Effectiveness of Data

Alon Halevy, Peter Norvig, and Fernando Pereira, *Google*

Eugene Wigner's article "The Unreasonable Effectiveness of Mathematics in the Natural Sciences"[1] examines why so much of physics can be neatly explained with simple mathematical formulas behavior. So, this corpus could serve as the basis of a complete model for certain tasks—if only we knew how to extract the model from the data.

Learning from Text at Web Scale

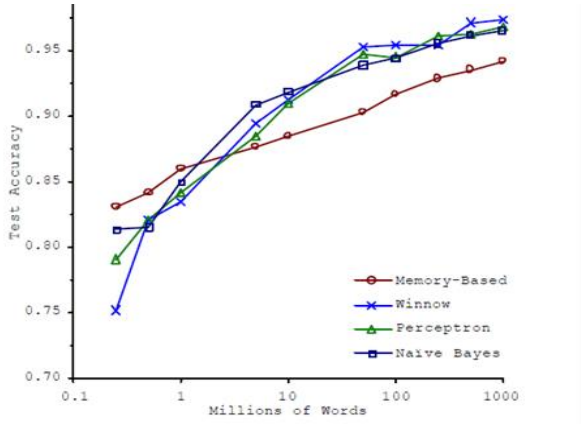


Figure 1. Learning Curves for Confusion Set Disambiguation

[Banko and Brill, 2001]

- 5.72e+7 - 2.50e+11 * (id1*id2) / vgs2 + 5.53e+6 * vds2 / vgs2 + 109.72 / id1

Build models

90.5 + 190.6 * id1 / vsg1 + 22.2 * id2 / vds2

10^(5.68 - 0.03 * vsg1 / vds2 - 55.43 * id1+ 5.63e-6 / id1)

2.36e+7 + 1.95e+4 * id2 / id1 - 104.69 / id2 + 2.15e+9 * id2 + 4.63e+8 * id1

Low accuracy

Merge

Build model

-10.3 + 7.08e-5 / id1 + 1.87 * ln(-1.95e+9 + 1.00e+10 / (vsg1*vsg3) + 1.42e+9 *(vds2*vsd5) / (vsg1*vgs2*vsg5*id2))

High accuracy

Decentralized / shared control encourages data sharing
Qualitatively new ecosystem-level data → qualitatively new models

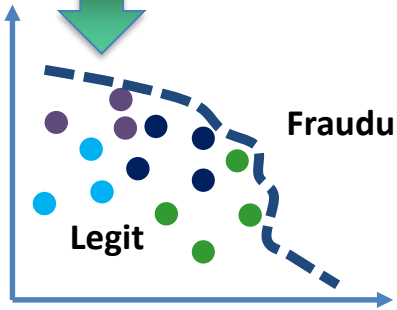
Example: shared diamond certification houses data → makes fraud id possible

Merge

Build classifiers

No single cert house has enough data to make an accurate classifier

Build 1-class classifier



Decentralized / shared control encourages data sharing

Qualitatively *new planet-level* data → qualitatively new models

Immutability for An Audit Trail on Training/Testing Data & Models

For greater trustworthiness of the data & models
(Avoid garbage-in, garbage-out)

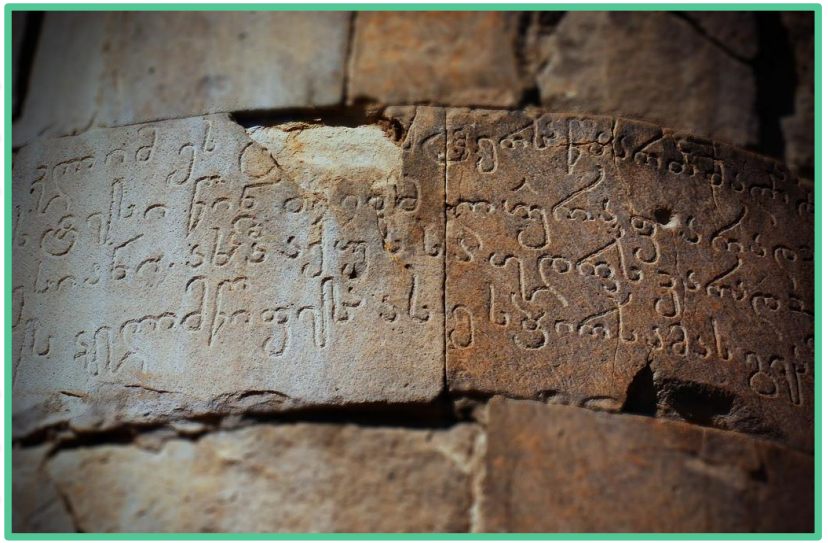
Provenance in building models:

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

Provenance in testing / in the field:

- Testing X data
- Model simulation
- Testing yhat data

Time-stamp / store



Applications:

- you can tell if a sensor is lying
- you know the “story” of a model
- catch leaks in the data chain

Another Opportunity:

A shared global registry of training data & models

"Models are owned by the planet"

All the Kaggle datasets

All the Kaggle models

All the ImageNet datasets

All the ImageNet models

....

Training/testing data & models as intellectual property assets

→ Decentralized data & model *exchanges*

Your datasets or models…

…licensed to others

Others' datasets & models

…licensed to you

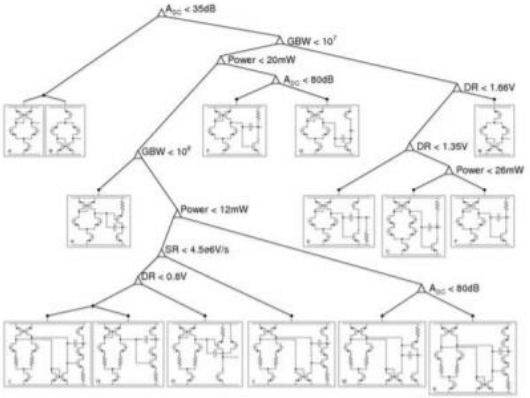
….

"EMX – European Model Exchange?"

Sell your CARTS?

Certificate of Authenticity

As of Nov. 06 2016, 19:10:42, trent is the owner.
To verify current owner, please visit https://www.ascribe.io/app/coa_verify/



Circuit Decision Tree

Edition: 1/3
Created by: Trent McConaghy
Owner: trent

ARTWORK DETAILS

Artwork ID: 136UbLGSHNHqY9kJxQ3tDy83K7P69zDJeN
File Extension: .png
File Size: 87090 bytes

PROVENANCE/OWNERSHIP HISTORY

Nov. 06, 2016, 19:10:42 - Registered by trent

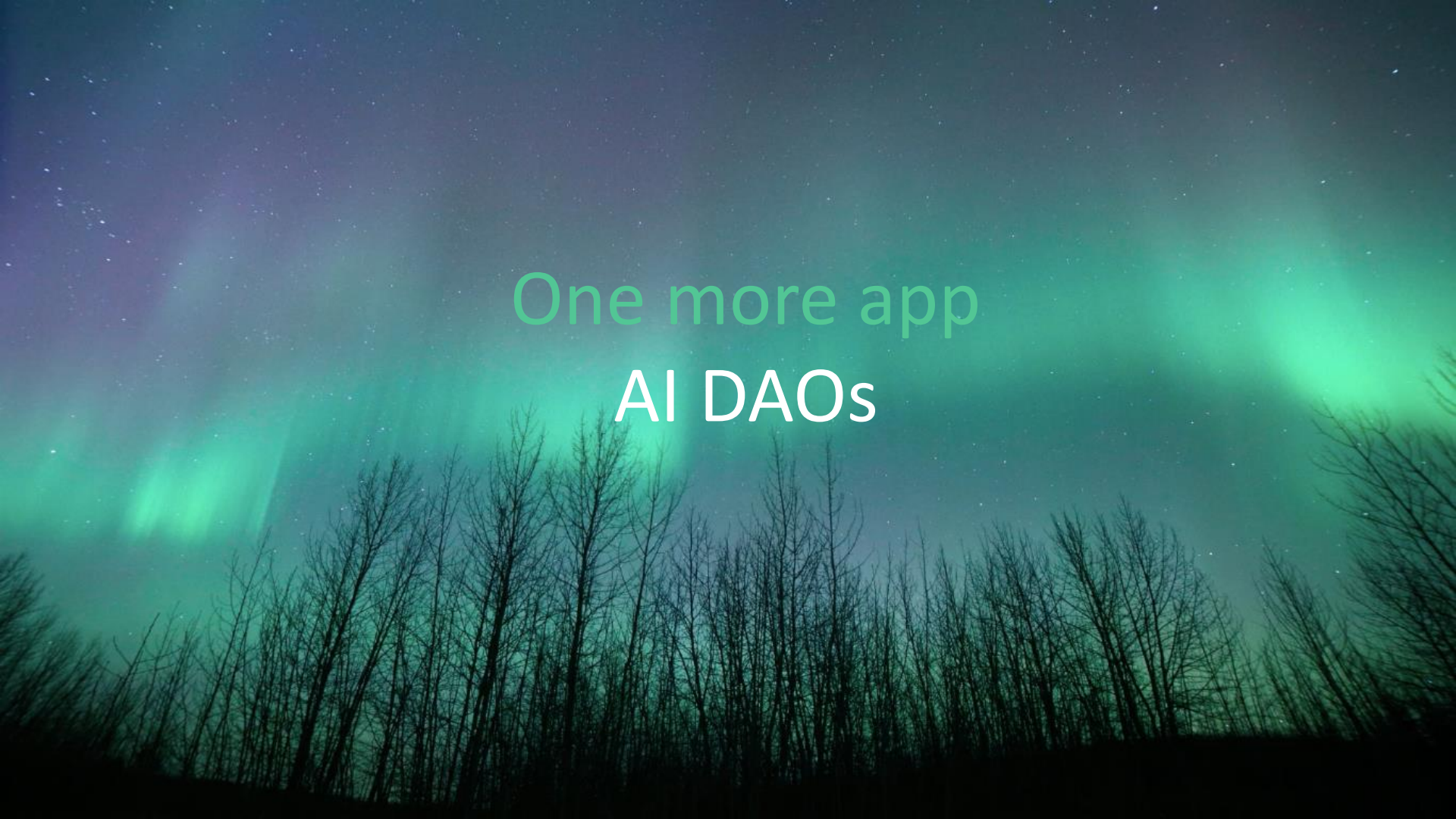
CRYPTOGRAPHIC STAMP

Use the summary and signature below to authenticate this certificate on:

Link: https://www.ascribe.io/app/coa_verify/

Summary: Trent McConaghy*Circuit Decision Tree*1/3*2008*2016Nov06-19:10:42

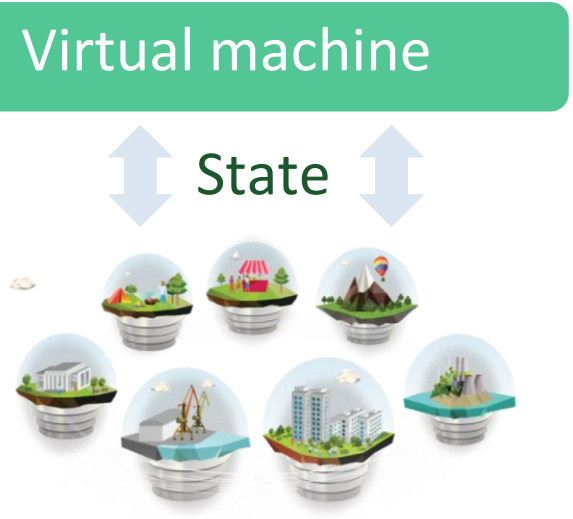
Signature: C38D56C823CEC09E40B35B9D27D48B9C8EF9ADECC9592F469CE0144CF9ECA406B3ABF1D976ADB7813895379A66F9F7C327BB0EE090A52F6A8274F3F4AC9EE3D7DF0FA98964C834678A6F48EF4FE687E7B4243F8F65FF57315CB7391A03874CD4BDFCB35718F1742AB5256B72A4C2D2593F3492372A66C82679263E39BAB9996EL


One more app
AI DAOs

What if you used a blockchain
to store *state* of a state machine?

Then you get
decentralized processing.

aka “smart contracts”



What if you used a blockchain
to store *state* of a state machine?

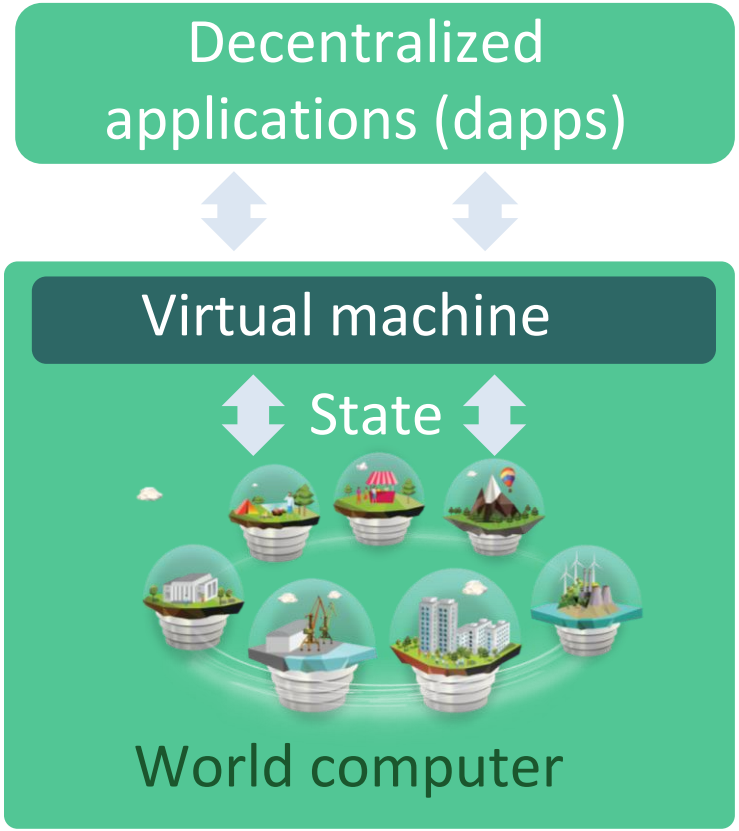
Then you get
decentralized processing.

And you can build a

world computer

having decentralized processing,
storage, and communications

(e.g. Ethereum vision)

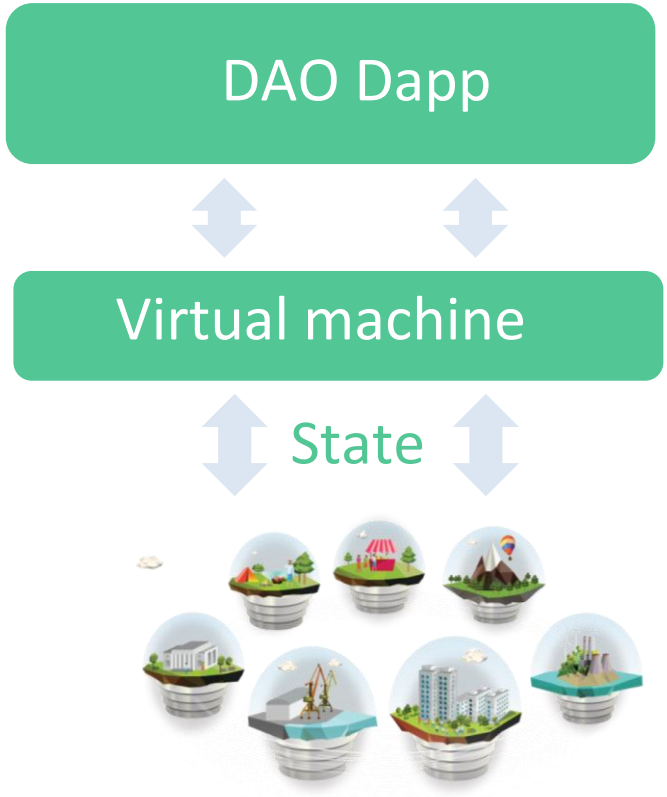


DAO: Decentralized Autonomous Organization

DAO: a computational process that

- runs autonomously,
- on decentralized infrastructure,
- with resource manipulation.

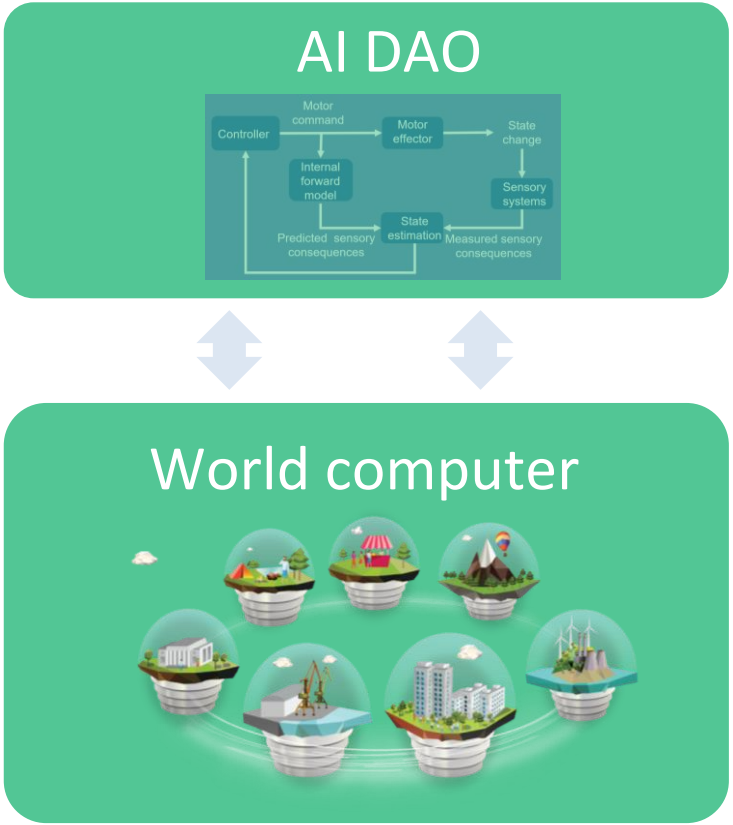
It’s code that can *own* stuff!



AGI on a DAO?

AI entity is a feedback control system. That is, AGI.

Its feedback loop would continue on its own, taking inputs, updating its state, and actuating outputs, with the resources to do so continually.



Example: The ArtDAO

Algorithm…

1. Run AI art engine to generate new image, using GP or deep
2. Claim attribution in blockchain, using ascribe
3. Create multiple editions, using ascribe
4. Post editions for sale onto a marketplace, using Getty (centralized), or OpenBazaar (decent.)
5. Sell the editions. $ goes to ArtDAO using built-in cryptocurrency like Ether. IP go from ArtDAO using ascribe.
6. Repeat! Create more art, sell it, get wealthier

Example: The ArtDAO

Algorithm…

1. Run AI art engine to generate new image, using GP or deep
2. Claim attribution in blockchain, using ascribe
3. Sell the editions. $ goes to ArtDAO using built-in cryptocurrency like Ether. IP go from ArtDAO using ascribe.
6. Repeat! Create more art, sell it, get wealthier

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.

Angles to Making AI DAOs

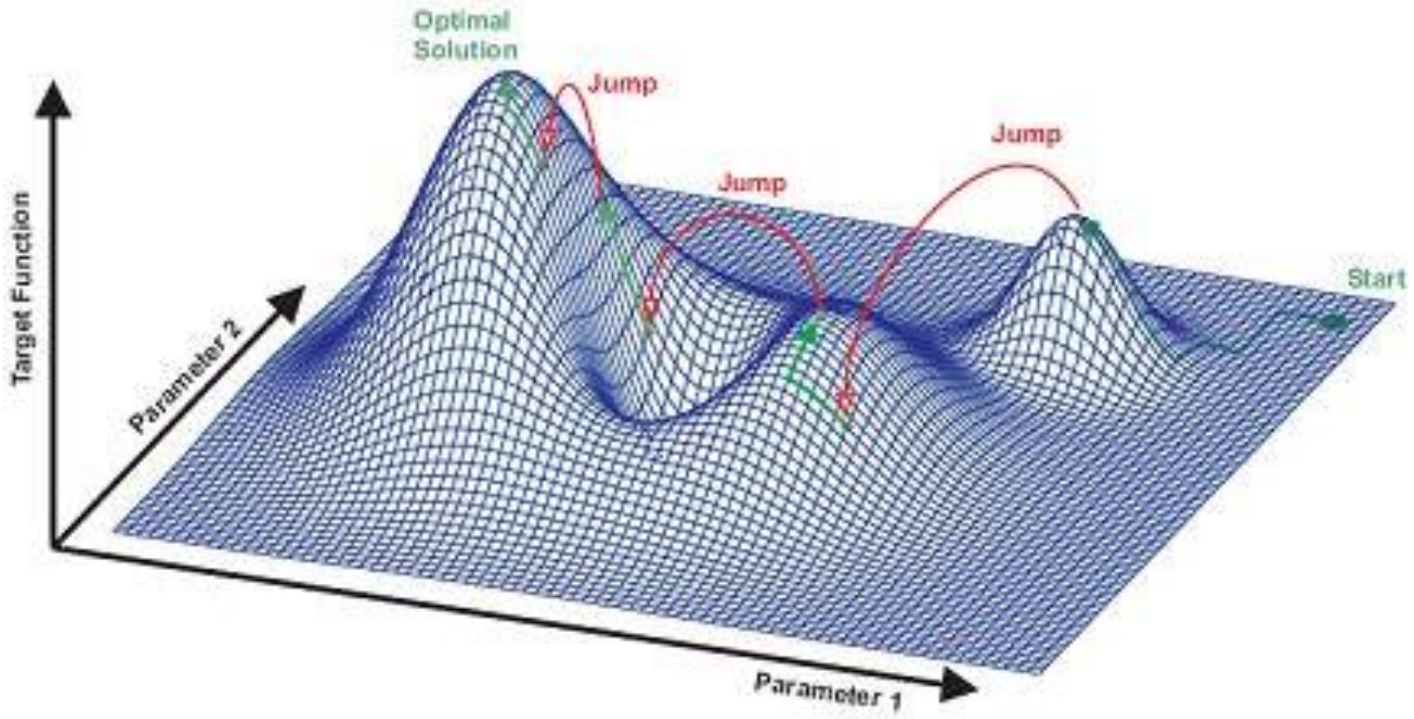
- **DAO → AI DAO**. Start with DAO, add AI. E.g. Plantoid
- **AI → AI DAO**. Start with AI, add DAO. E.g. numer.ai
- **SaaS → DAO → AI DAO**. Convert SaaS to DAO. Then add AI
- **Physical service → AI DAO**. E.g. Uber self-*owning* cars


AI DAOs II

Giving AIs control of our cars

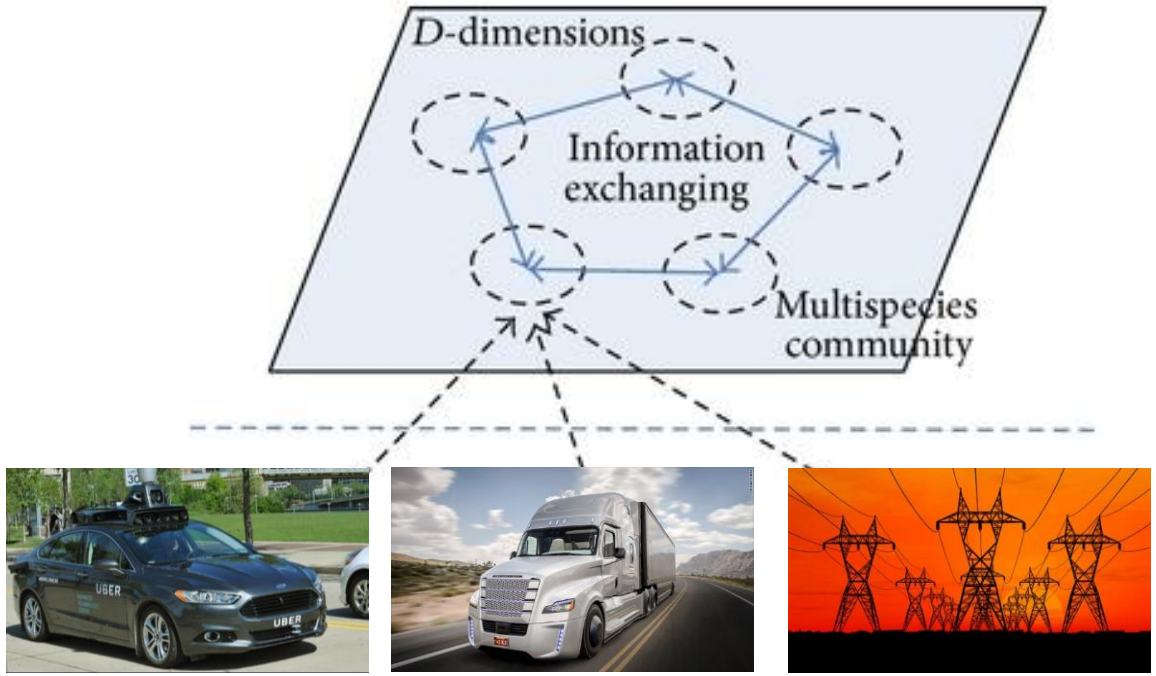

FREIGHTLINER
Giving AIs control of our trucks

Giving AIs control of our grid

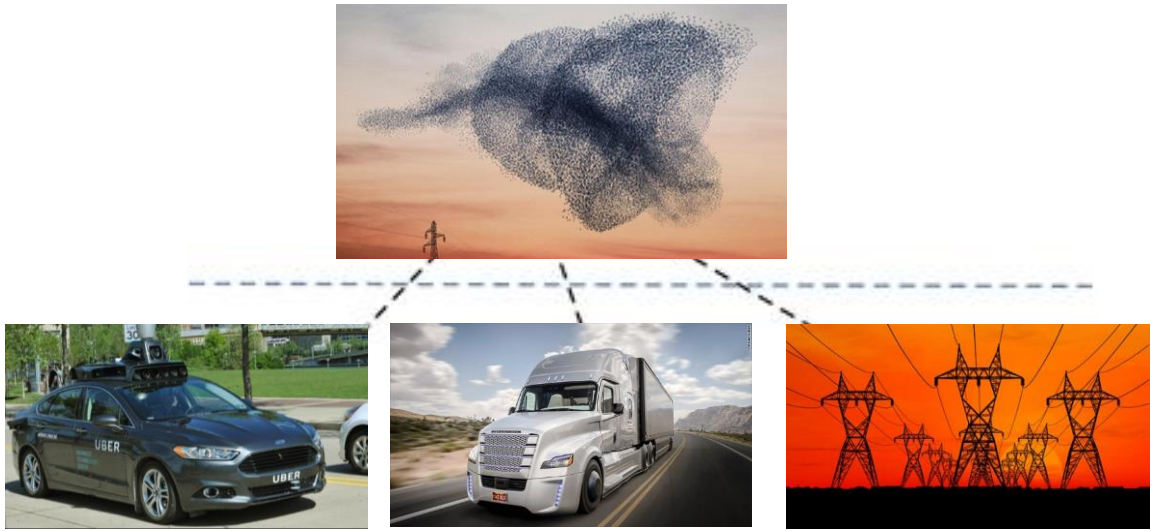


All in the name of optimization

Higher-level integration, therefore higher-level optimization



AIs Become AI DAOs, everywhere



The AIs now own stuff.
You can’t turn them off.

AIs Become AI DAOs,
everywhere

The AIs now own stuff.
You can't turn them off.

And you've just given them control of all your resources.
(Oops)

Blockchains for Artificial Intelligence

A planetary-scale blockchain database (IPDB) unlocks opportunities:

1. Data sharing → Better models
2. Data sharing → Qualitatively new models
3. Audit trails on data & models for more trustworthy predictions
4. Shared global registry of training data & models
5. Data & models as IP assets → data & model exchange
6. AI DAOs – AI that can accumulate wealth, that you can't turn off (be careful)

Trent McConaghy
@trentmc0

bigchaindb.com
ipdb.foundation