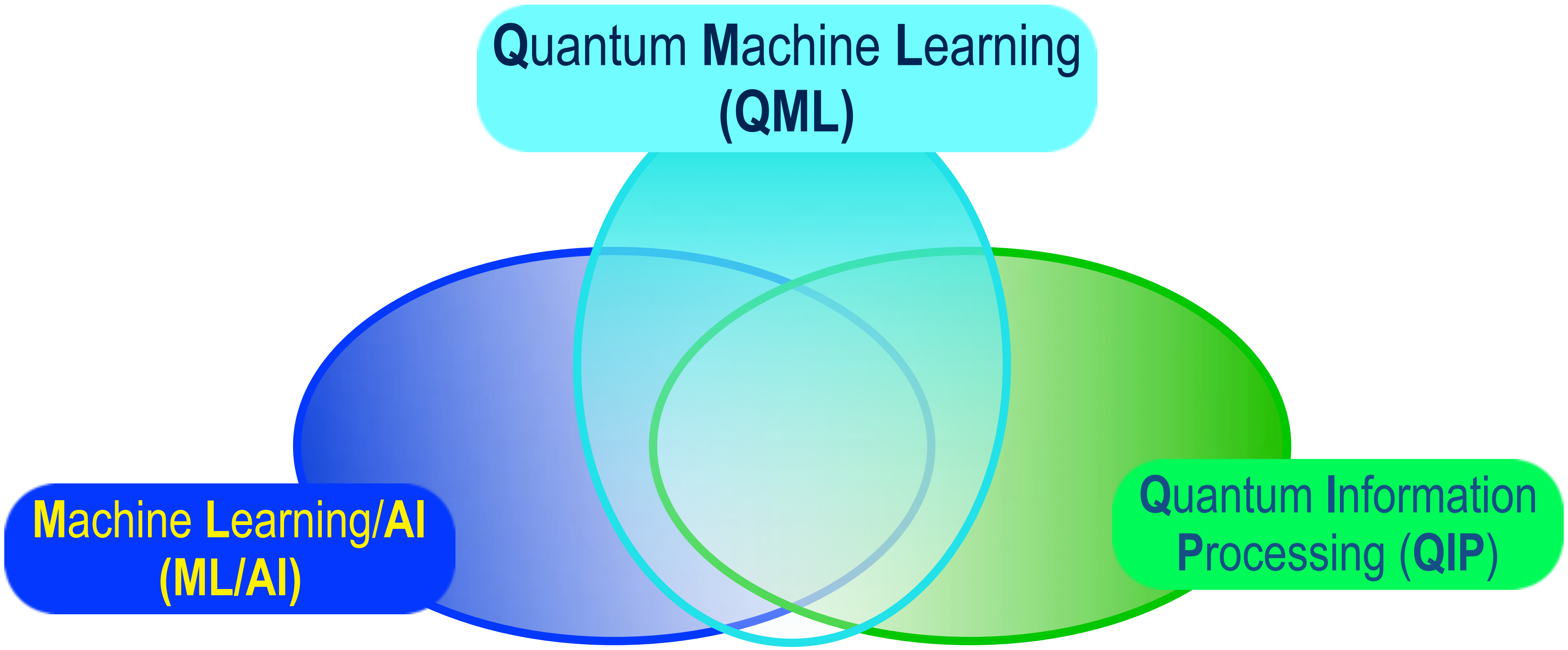
Quantum Software Consortium

Universiteit Leiden

Topics in Quantum Machine Learning

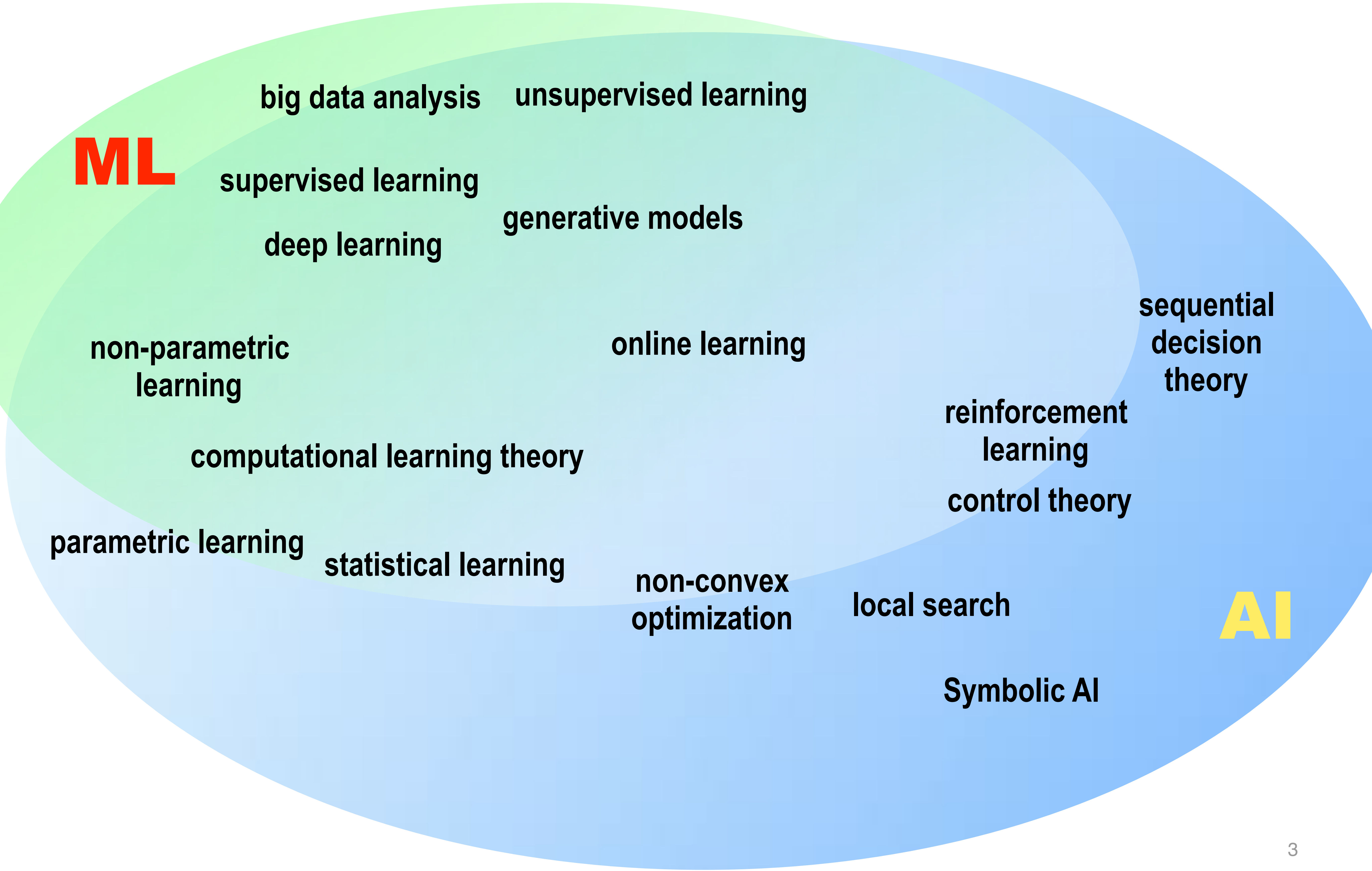
Vedran Dunjko
v.dunjko@liacs.leidenuniv.nl



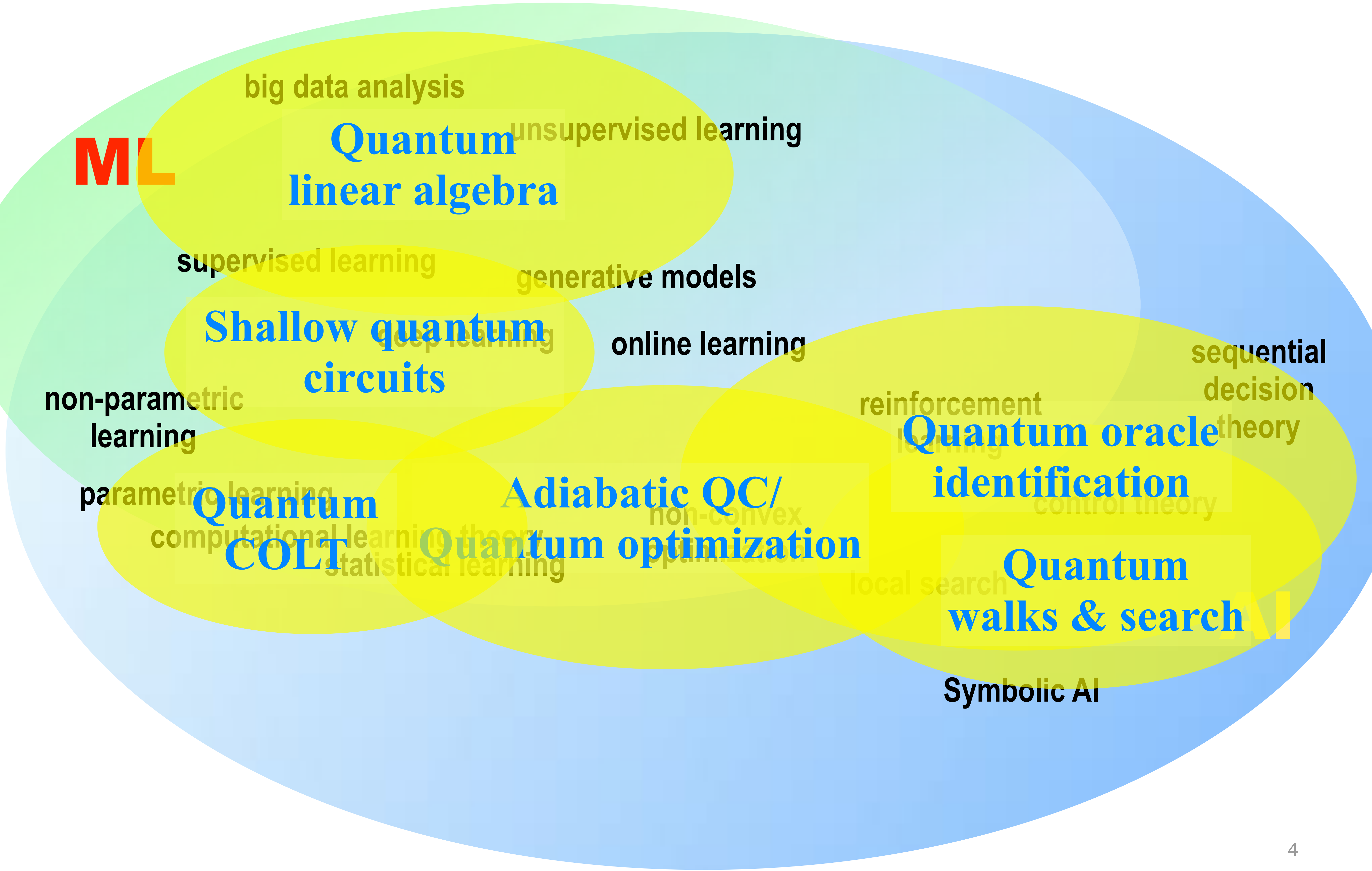
- **ML→QIP** (quantum-applied ML) ['74]
- **QIP→ML** (quantum-enhanced ML) [‘94]

- **QIP↭ML** (quantum-generalized learning) [‘00]
  - ML-inspired QM/QIP
  - Physics inspired ML/AI

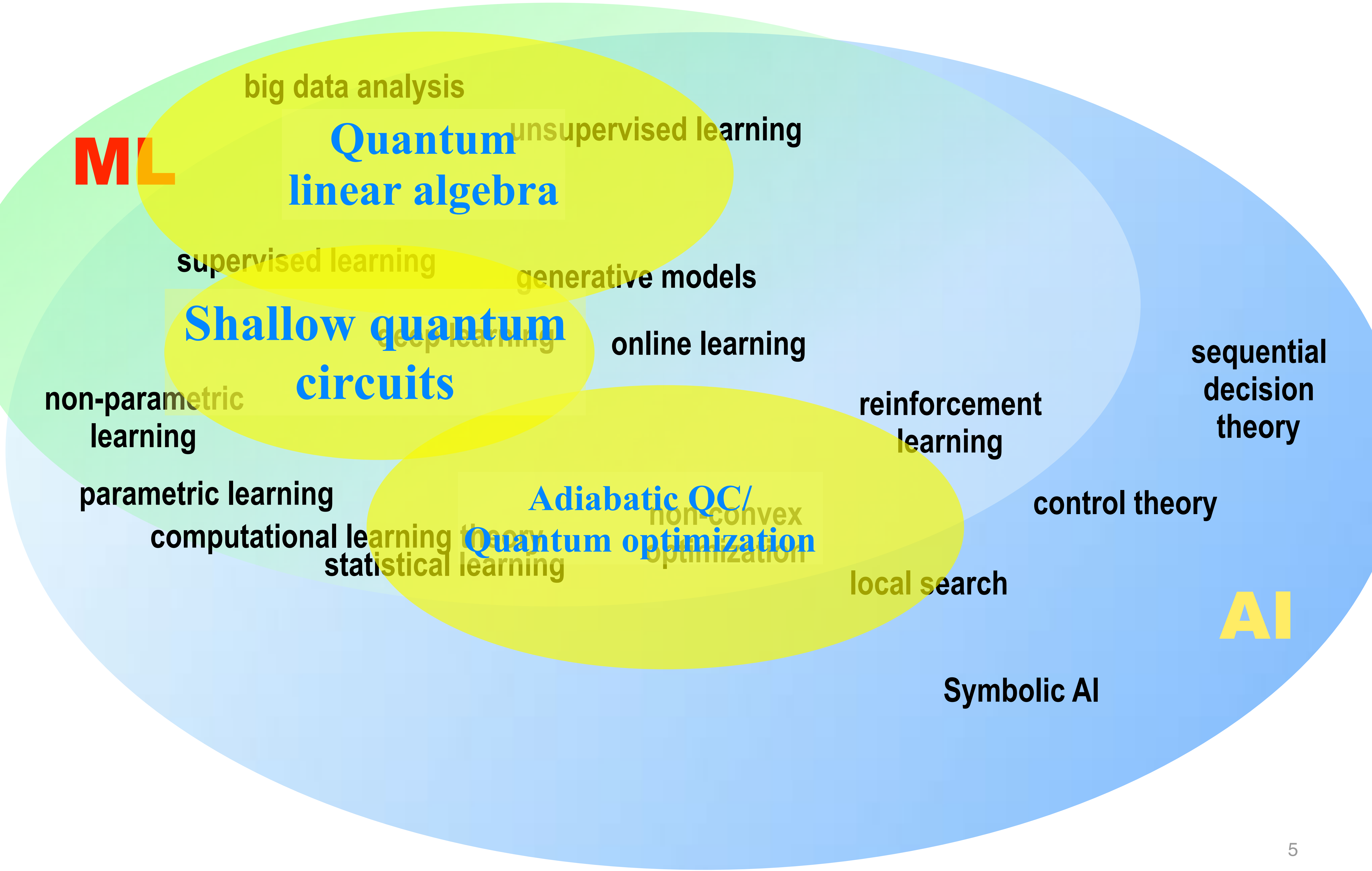
Machine learning is not one thing.
AI is not even a few things.



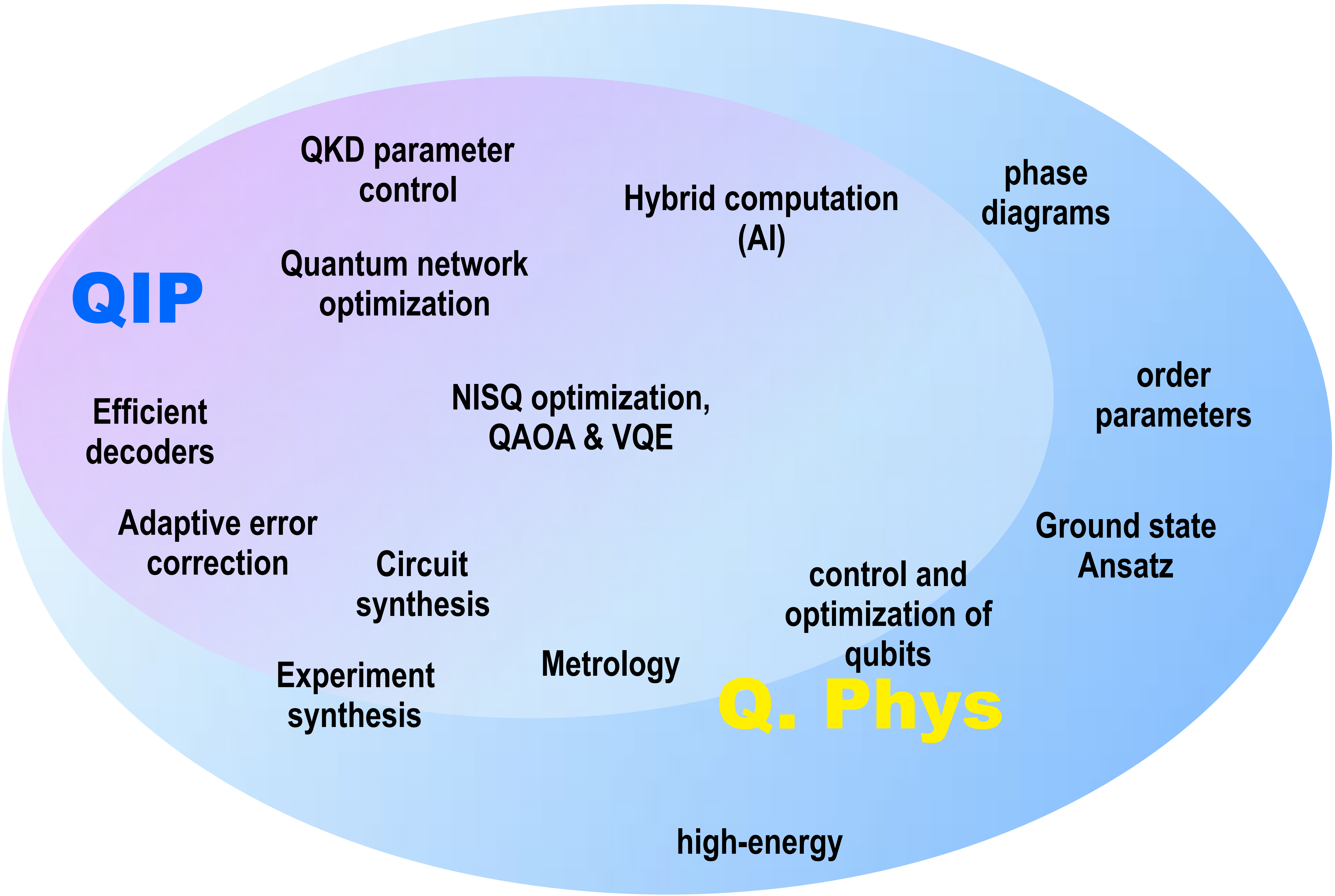
Quantum-enhanced ML is even more things

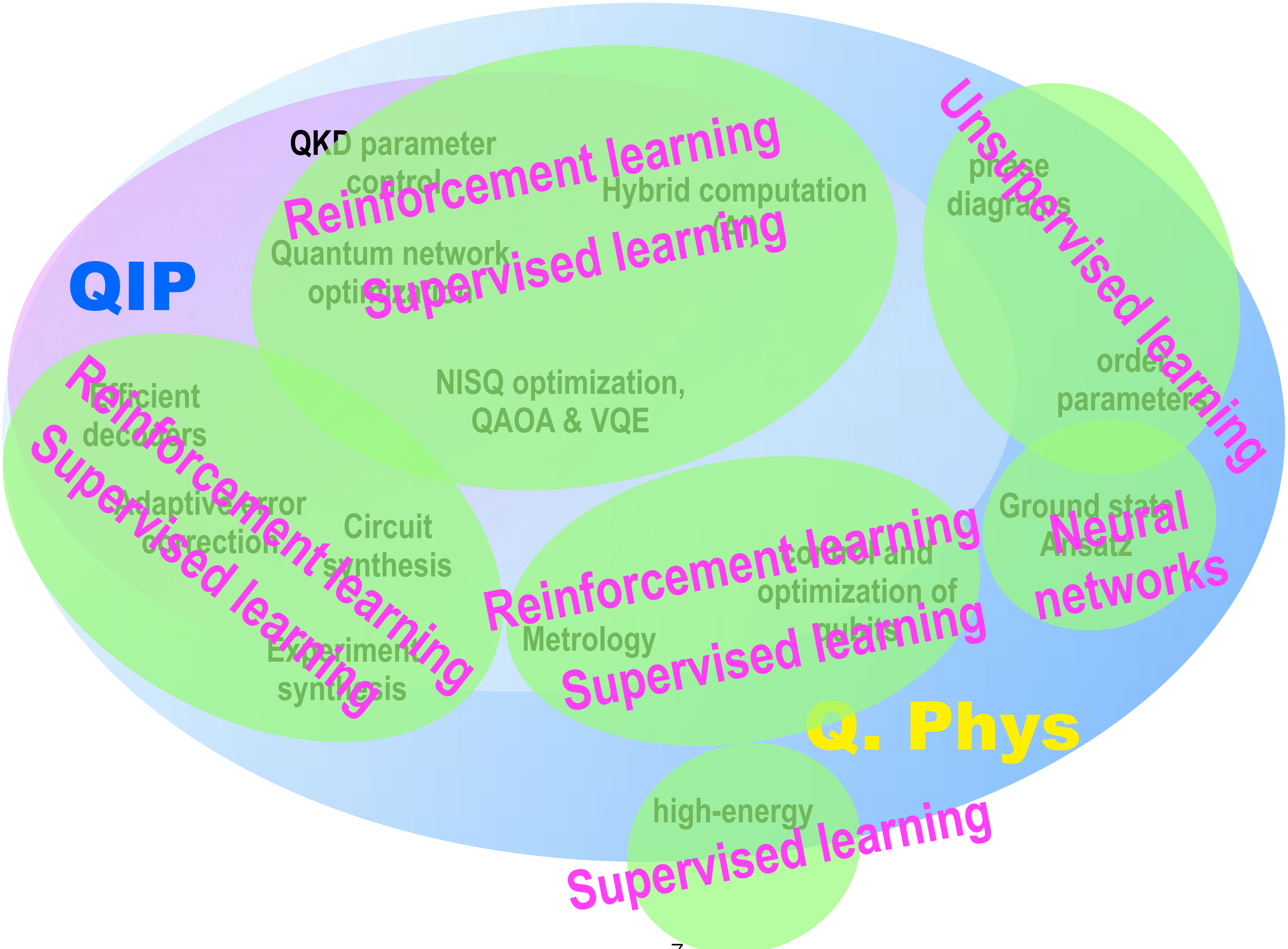


Quantum-enhanced ML is even more things



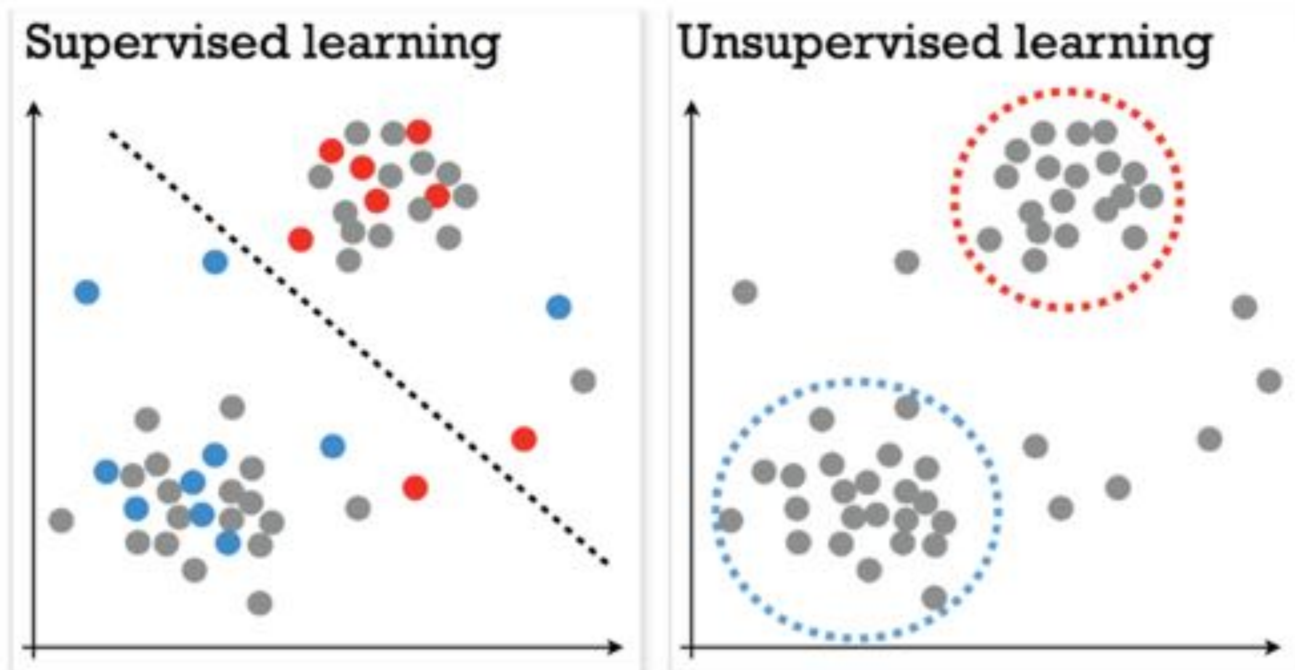
And then there's Quantum-applied ML!




QIP
Q. Phys
QKD parameter control
Reinforcement learning
Hybrid computation (AI)
Supervised learning
Quantum network optimization
NISQ optimization, QAOA & VQE
Unsupervised learning
phase diagrams
order parameters
Efficient decoders
Reinforcement learning
Supervised learning
Adaptive error correction
Circuit synthesis
Experiment synthesis
Reinforcement learning
Control and optimization of qubits
Metrology
Supervised learning
Ground state Ansatz
Neural networks
high-energy
Supervised learning

What is machine learning

Machine Learning: the **WHAT**



or

Learning *P(labels|data)* given samples from *P(data,labels)* *(also regression)*

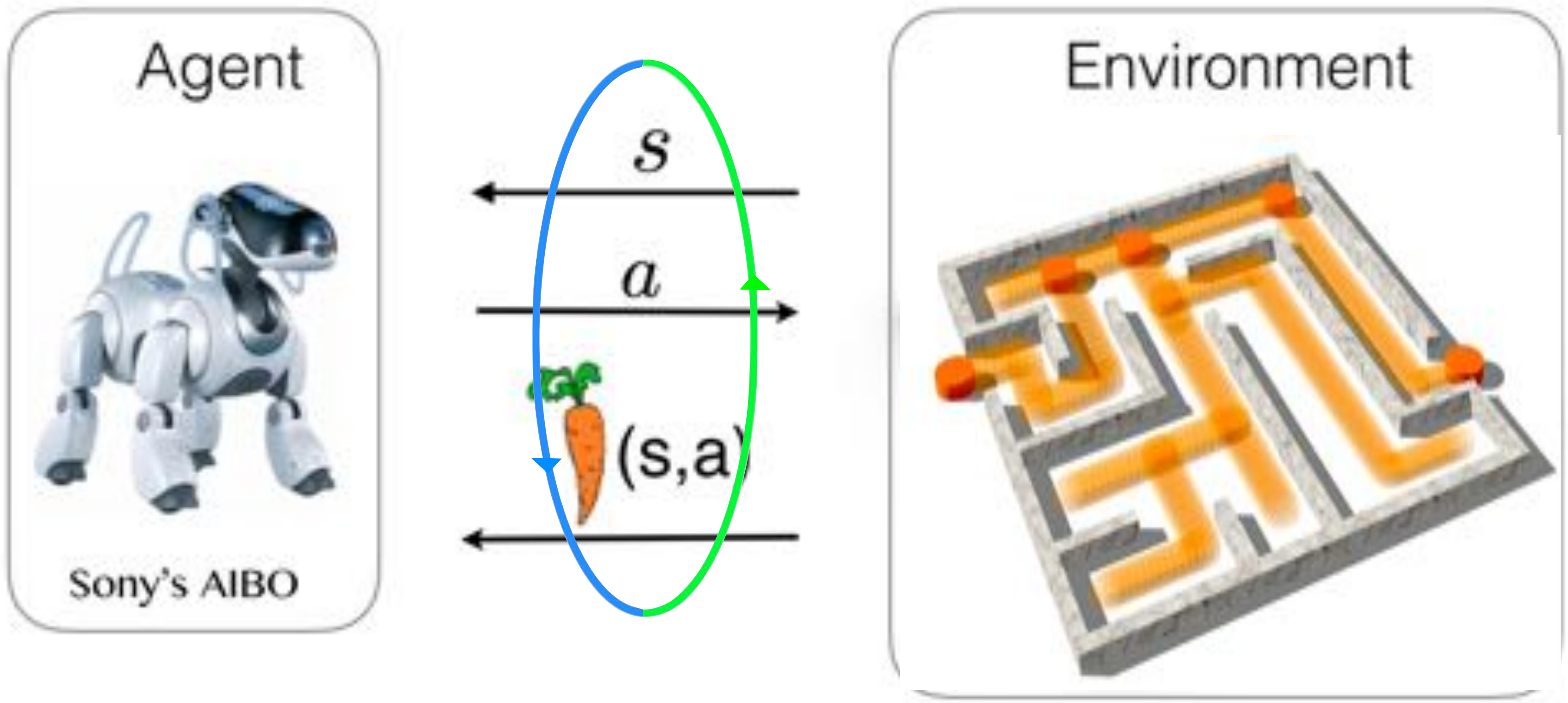
?

-generative models
-clustering (discriminative)
-feature extraction

Learning structure in *P(data)* give samples from *P(data)*

Machine Learning: the **WHAT**

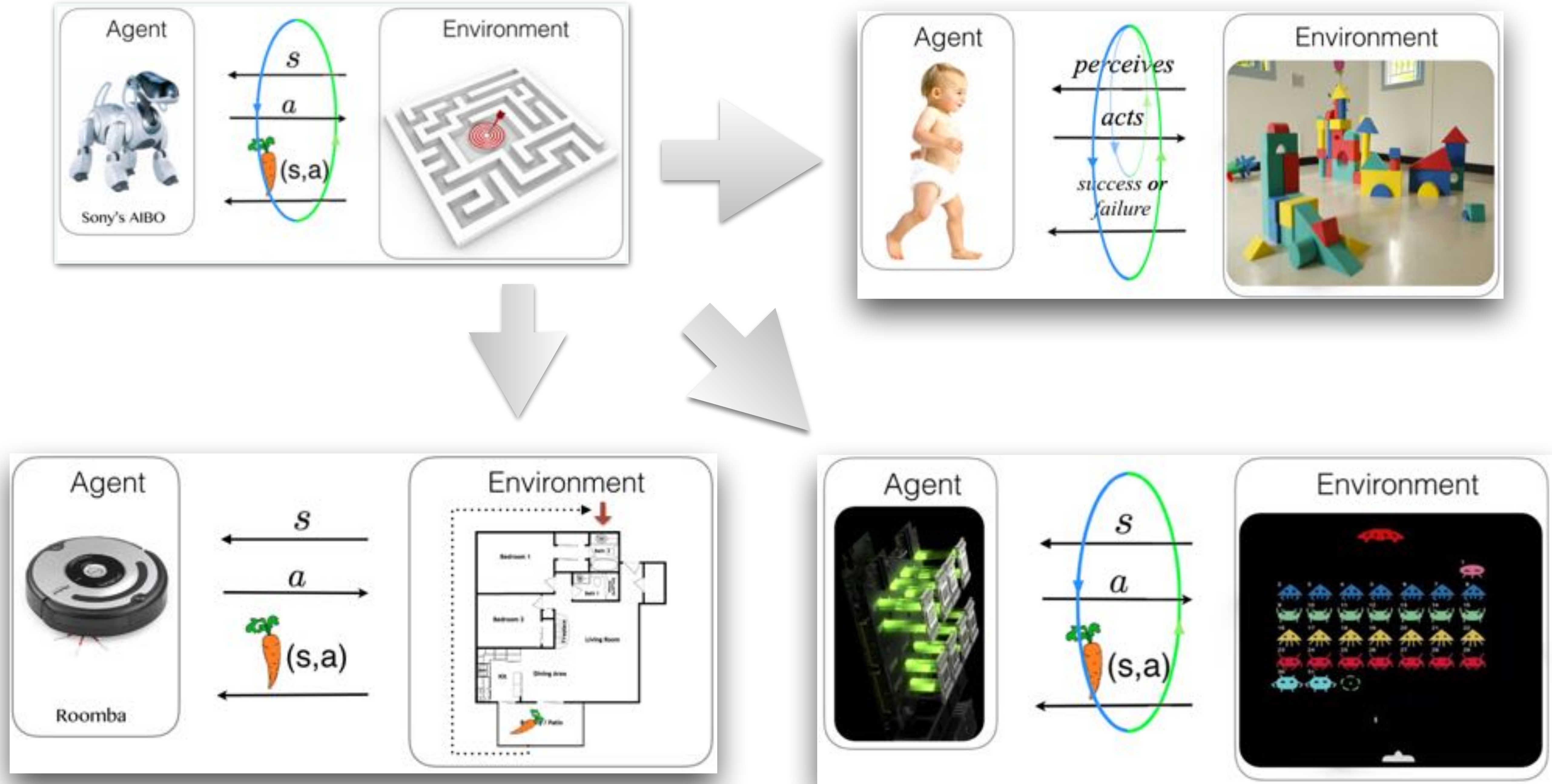
Beyond data: *reinforcement learning*



$$\mathcal{S} = \{s_1, s_2, \ldots\}$$
$$\mathcal{A} = \{a_1, a_2, \ldots\}$$

Closer to AI.
There is a body.
Interaction.
Learning.

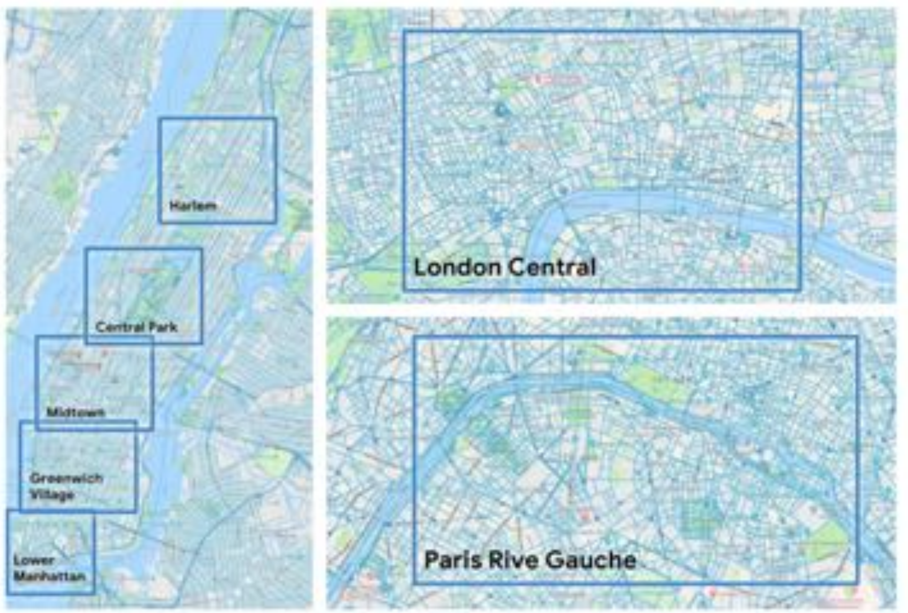
$$\pi(a|s) \qquad\qquad T(s|s', a)$$

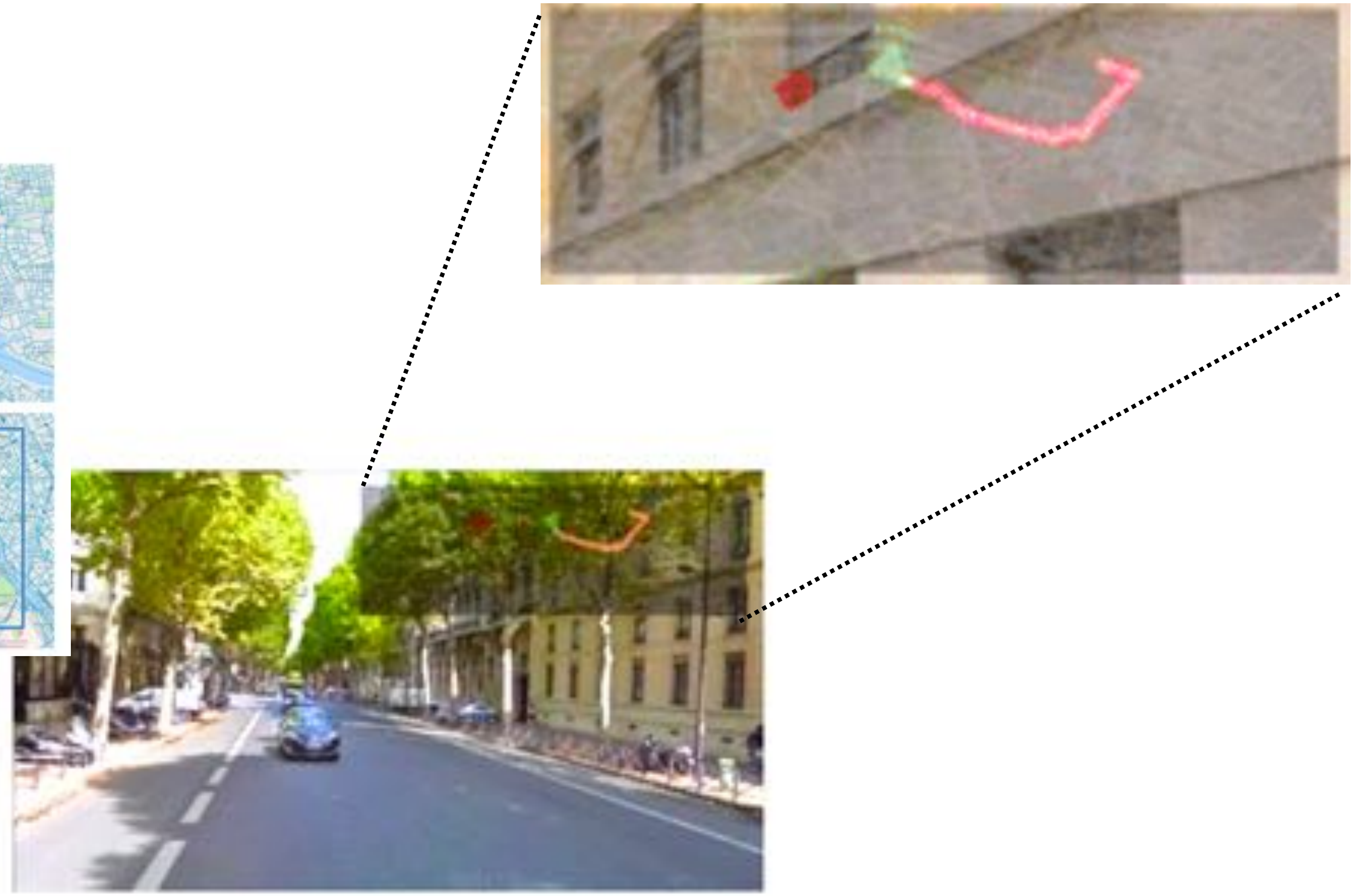


Also: MIT technology review breakthrough technology of 2017
[AlphaGo anyone?]

Using RL in Real Life

Navigating a city…



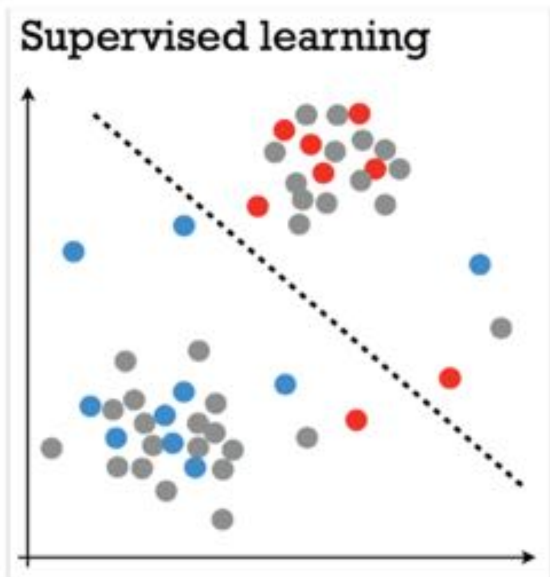


Stop-motion films of agent trained in Paris. The images are superposed with a map of the city, showing the goal location (in red) and the agent location and field of view (in green). Note that the agent does not see the map, only the lat/lon coordinates of the goal location.

https://sites.google.com/view/streetlearn

P. Mirowski et. al, *Learning to Navigate in Cities Without a Map*, arXiv:1804.00168

Machine Learning: the **HOW**

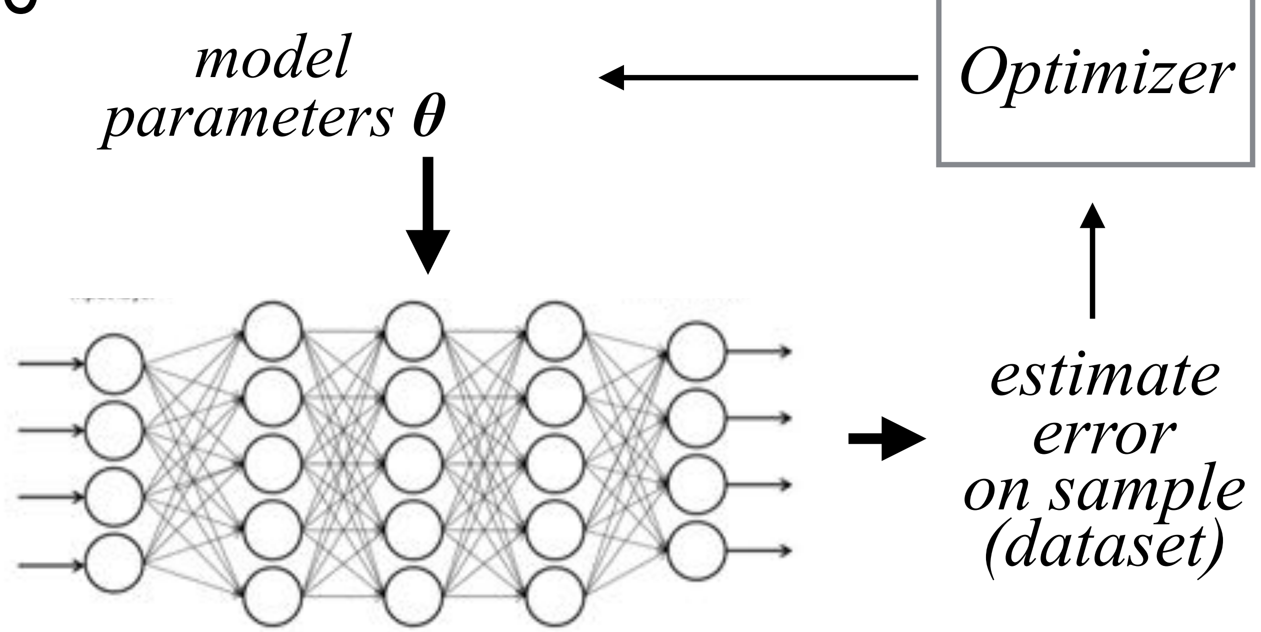


$$\text{parametrized family } \{h_\theta\}_\theta$$

$$\text{argmin}_\theta \ \text{Err\_training\_set}(\theta) + \text{Reg}(\theta)$$

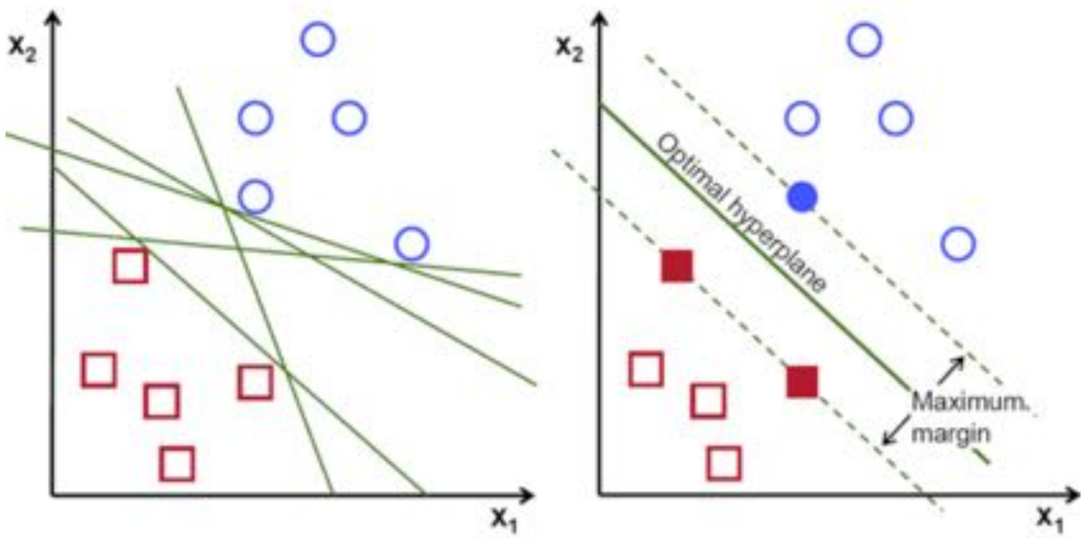
output hypothesis *h* on *Data* x *Labels*
approximating P(labels|data)

In practice



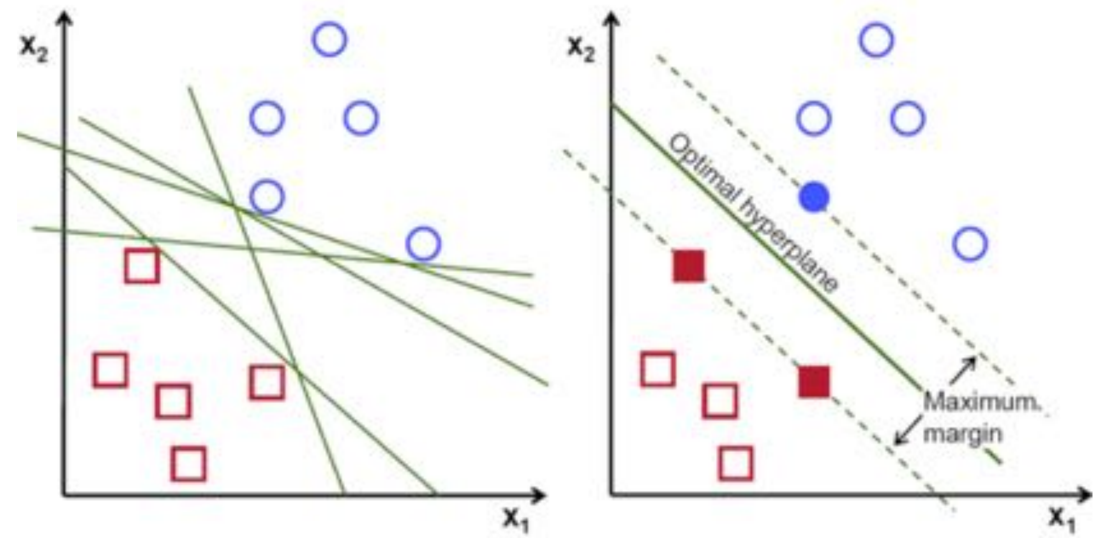
Support vector machines

separating hyperplane..

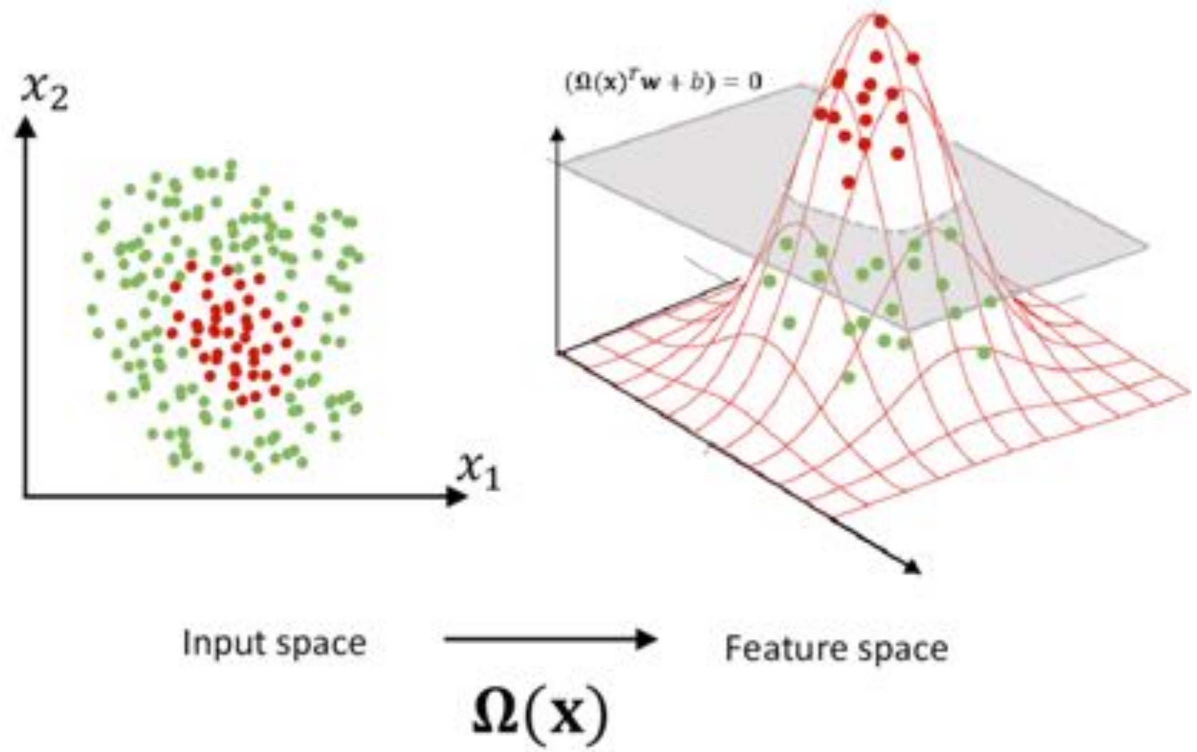


Support vector machines

separating hyperplane..



…in higher-dimensional feature space



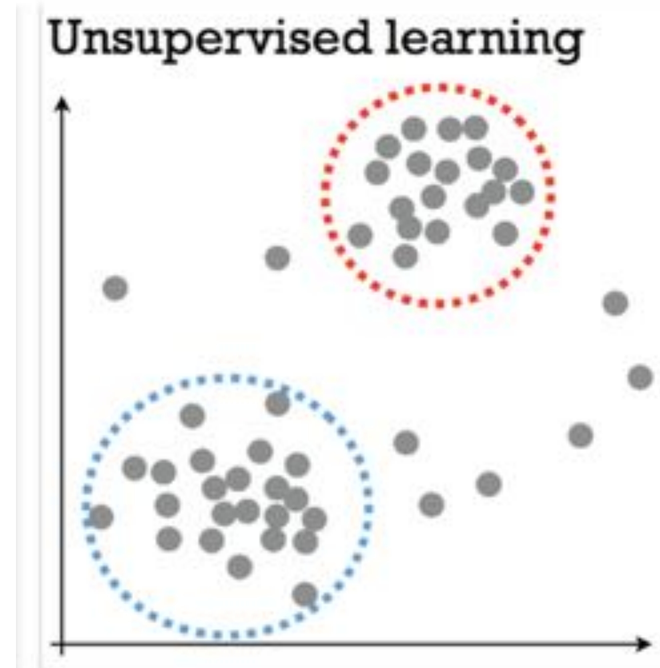
Still (algebraic) optimization over hyperplane and feature function parameters….

Machine Learning: the **HOW**

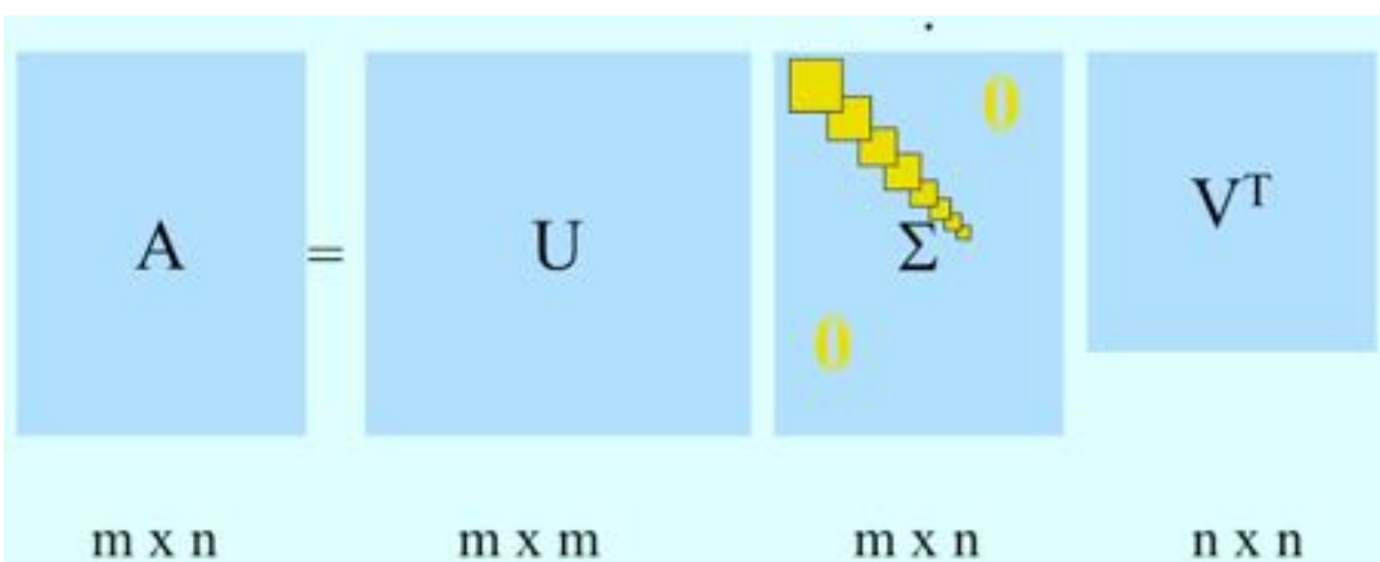
parametrized family $\{h_\theta\}_\theta$

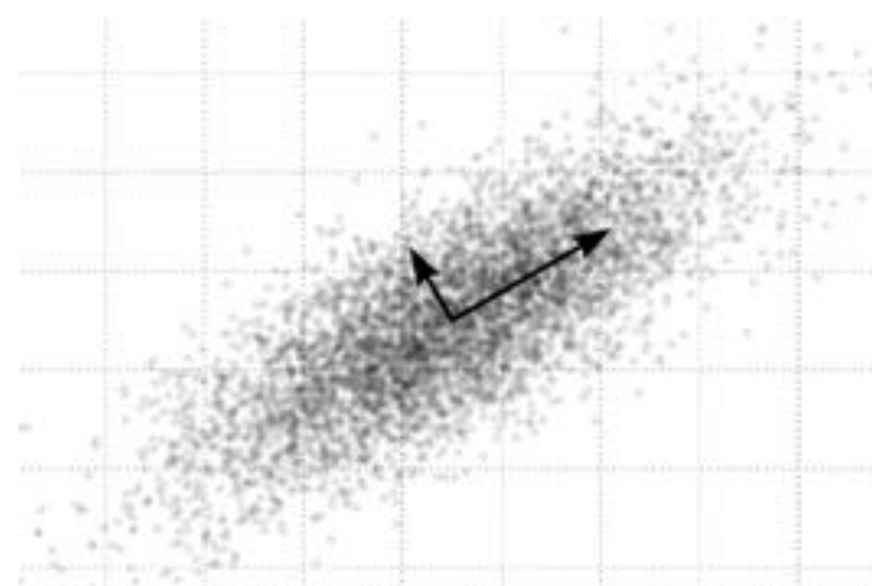
$\mathrm{argmin}_\theta \ \mathrm{Err\_training\_set}(\theta) + \mathrm{Reg}(\theta)$

$$H(\sigma) = -\sum_{\langle i\ j\rangle} J_{ij}\sigma_i\sigma_j - \mu\sum_j h_j\sigma_j$$



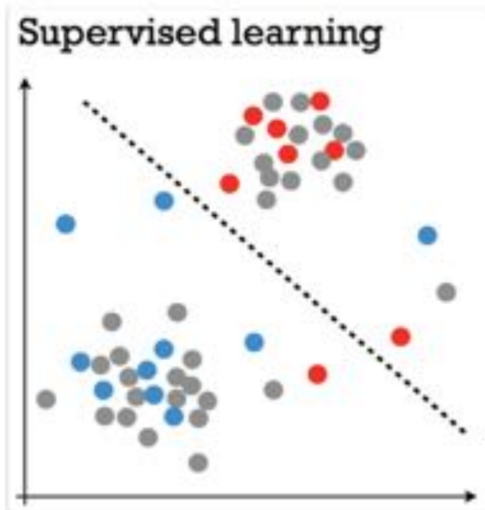
Learning structure in *P(data)*
give samples from *P(data)*



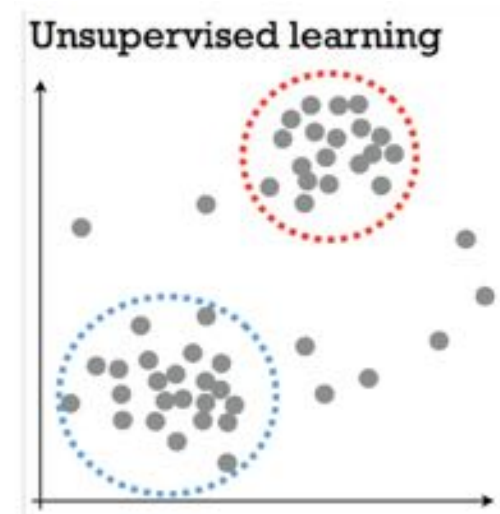


Machine Learning: the **HOW**

$$\text{argmin}_\theta \ \text{Err\_training\_set}(\theta) + \text{Reg}(\theta)$$

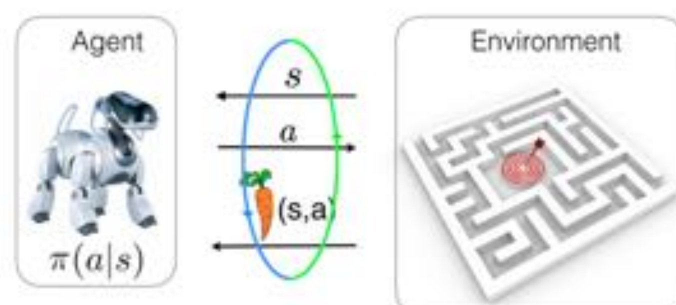


output:
hypothesis h on *Data* x *Labels*
approximating P(labels|data)



output:
hypothesis h on *Data*
"approximating" P(data)

Reinforcement learning



output:
policy π on *Actions* x *States*

Supervised learning
(learning *how to label datapoints, learning how to approximate a function, how to classify*)

Unsupervised learning
(learning *a distribution, generate. properties from samples, feature extraction & dim. reduction*)

Reinforcement learning
(learning *behavior*, *policy*, or *optimal control*)

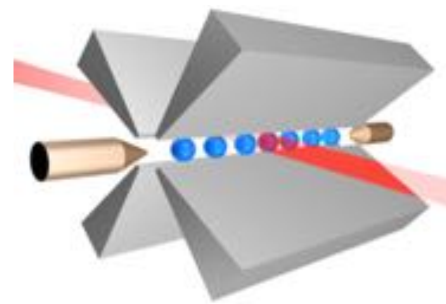
That is all ML we need for now

What about quantum computers?

Quantum computers…

…and physics

-*manipulate* registers of 2-level systems (qubits)



-full description:

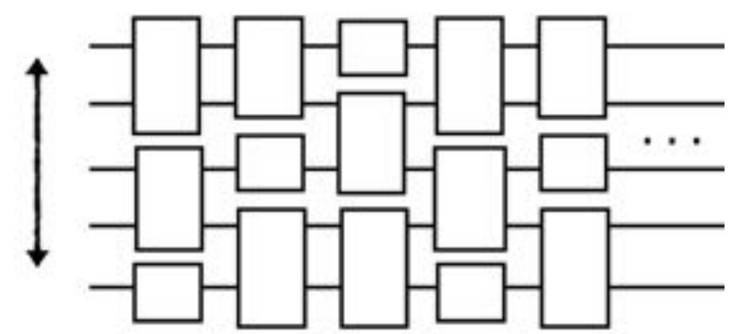
n qubits* $\rightarrow$ *2^n dimensional vector

-**manipulation**: acting locally (**gates**)

…and computer science

-likely can *efficiently* compute more things than classical computers (factoring)
e.g. factor numbers, or generate complex distributions

-even if QC is "shallow"



…and reality

special-purpose
quantum annealers

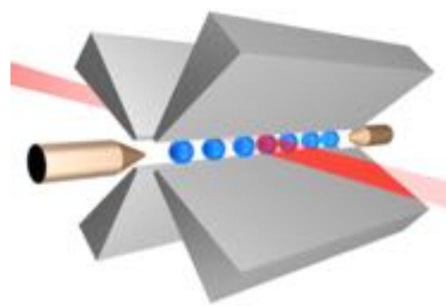
Banana
for scale

cca 50 qubit
all-purpose
noisy

Quantum computers…

…and physics

-*manipulate* registers of 2-level systems (qubits)



-full description:

n qubits* $\rightarrow$ *2^n dimensional vector

…and computer science

-can compute things likely beyond **BPP** (factoring)

-can produce distributions which are hard-to-simulate for classical computers (unless **PH collapses**)

-even if QC is "shallow"

…and reality

special-purpose
quantum annealers

Banana
for scale

cca 50 qubit
all-purpose
noisy

Quantum-enhanced supervised learning: the quantum pipeline

a) The optimization bottleneck
b) Big data & comp. complexity
c) Machine learning Models

Quantum-enhanced supervised learning: the quantum pipeline

a) The optimization bottleneck — **quantum annealers**

b) Big data & comp. complexity — **universal QC and Q. databases**

c) Machine learning Models — **restricted (shallow) architectures**

Quantum-enhanced supervised learning: the quantum pipeline

a) The optimization bottleneck — **quantum annealers**

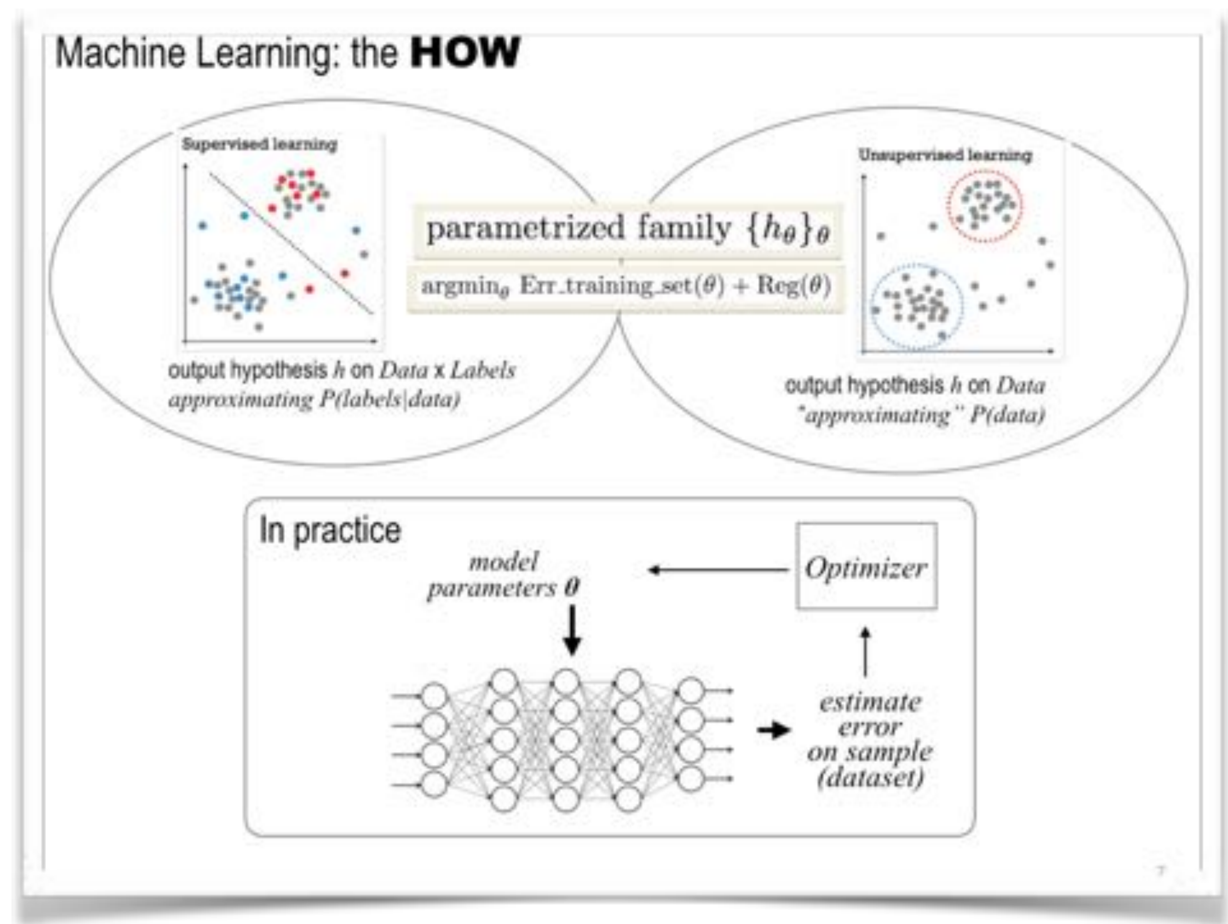
b) Big data & comp. complexity — **universal QC and Q. databases**

c) Machine learning Models — **restricted (shallow) architectures**

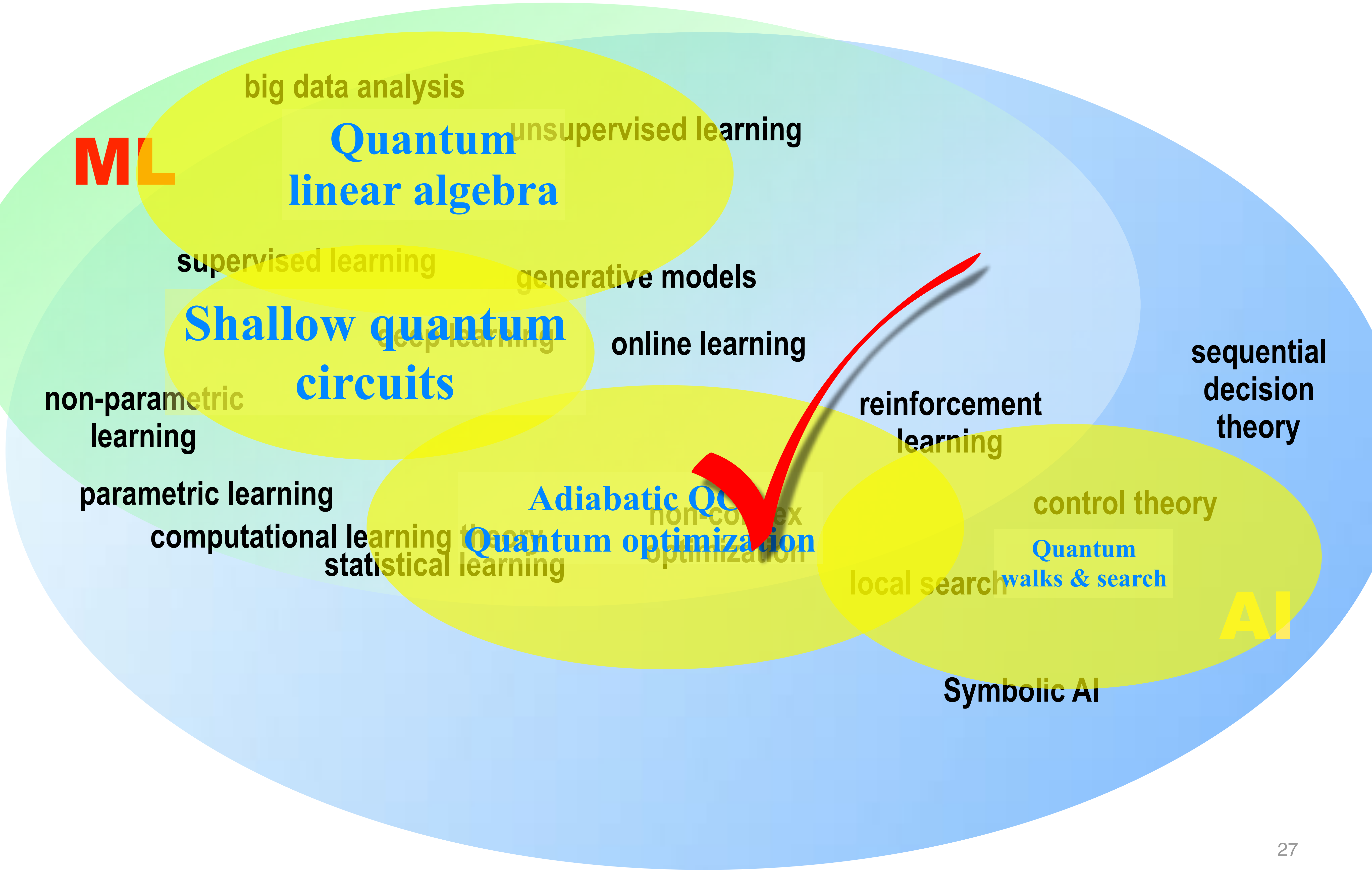
The optimization bottleneck

$$H(s) = sH_{initial} + (1-s)H_{target}; \quad s(time)$$

- Finding ground states of Hamiltonians via adiabatic evolution
- Very generic optimization problem: $\mathrm{argmin}_{|\psi\rangle} \langle\psi|H|\psi\rangle$



QeML is even more things



Quantum-enhanced supervised learning: the quantum pipeline

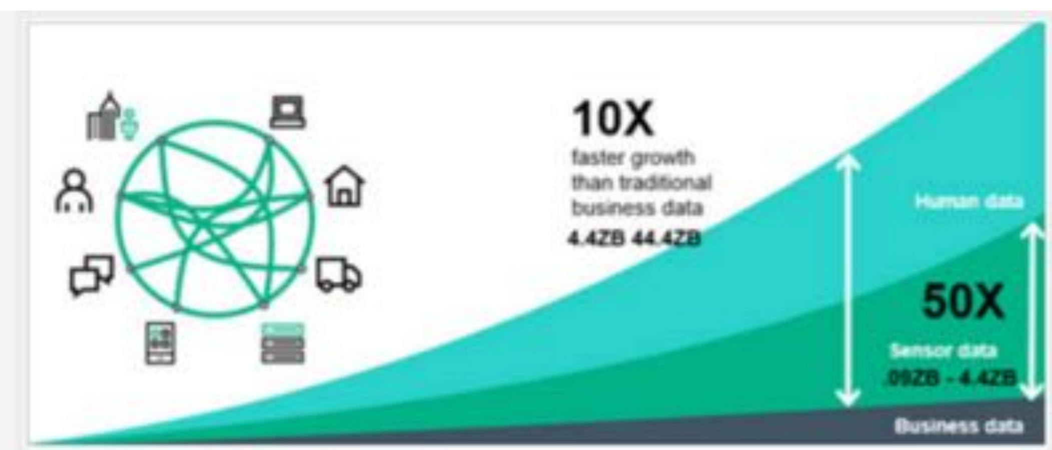
a) The optimization bottleneck — **quantum annealers**

b) Big data & comp. complexity — **universal QC and Q. databases**

c) Machine learning Models — **restricted (shallow) architectures**

Precursors of Quantum Big Data

<u>*Exponential data?*</u>



There are many sources that predict exponential data growth toward 2020 and beyond.

+

Much of data analysis is linear-algebra:

regression = Moore-Penrose
PCA = SVD…

Enter *quantum linear algebra*

interpret QM as linear algebra ***verbatim***

state vector	$\leftrightarrow$ (data) vector
density matrices Hamiltonians unitaries	$\leftrightarrow$ linear maps
projective measurements (swap tests)	$\leftrightarrow$ inner products

prepare states expressible as linear-algebraic manipulations of data-vectors in ***polylog(N)***

(when other quantities are well behaved)

amplitude encoding

$$\mathbf{R}^N \ni \mathbf{x} = (x_i)_i$$

$$\downarrow$$

$$|\psi\rangle \propto \sum_{i=1}^N x_i |i\rangle$$

exp(*n*) amplitudes in *n* qubits

block encoding

$$U|0\rangle|\psi\rangle = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \psi \\ 0 \end{bmatrix} = \begin{bmatrix} A\psi \\ C\psi \end{bmatrix} = |0\rangle A|\psi\rangle + |1\rangle C|\psi\rangle$$

functions of operators

$$f(A)|\psi\rangle = \alpha_0|\psi\rangle + \alpha_1 A|\psi\rangle + \alpha_0 A^2|\psi\rangle \cdots$$

$$\approx A^{-1}|\psi\rangle$$

inner products

$$P(0)_\psi = |\langle 0|\psi\rangle|^2$$

Phys. Rev. Lett. **15**,. 103, 250502 (2009)
arXiv:1806.01838

If this worked literally…this would make us *INFORMATION GODS.*

Prediction: *44 zettabytes by 2020.*

If all data is floats, this is 5.5×10^{21} *float values*

If this worked literally…this would make us *INFORMATION GODS.*

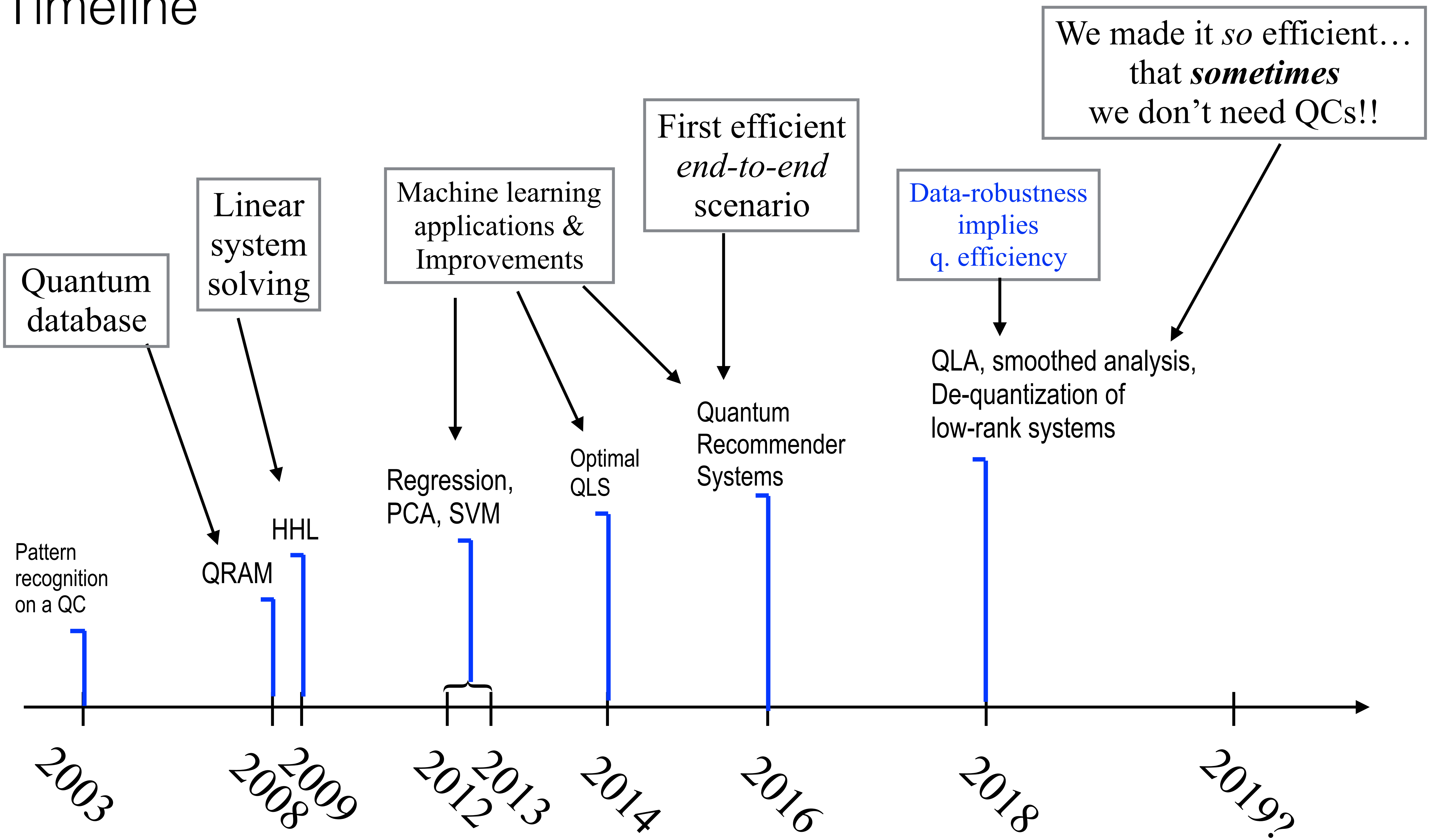
Prediction: *44 zettabytes by 2020.*

If all data is floats, this is 5.5×10^{21} float values

… can be stored in state of 73 qubits (ions, photons….)

Clearly there is a catch.
Many of them.

Timeline



We made it *so* efficient…
that ***sometimes***
we don't need QCs!!
First efficient
end-to-end
scenario
Machine learning
applications &
Improvements
Linear
system
solving
Quantum
database
Data-robustness
implies
q. efficiency
QLA, smoothed analysis,
De-quantization of
low-rank systems
Quantum
Recommender
Systems
Optimal
QLS
Regression,
PCA, SVM
HHL
QRAM
Pattern
recognition
on a QC
2003
2008
2009
2012
2013
2014
2016
2018
2019?

Summary of quantum (inspired) "big data"

interpret QM as linear algebra ***verbatim***

manipulate exponentially-sized data-vectors in system (qubit) number

HOWEVER

need full blown ideal QC
need pre-filled database (QRAM)
need appropriate condition numbers
need robustness to linear error
need right preprocessing applied
can sometimes be done classically

Summary of quantum (inspired) "big data"

interpret QM as linear algebra ***verbatim***

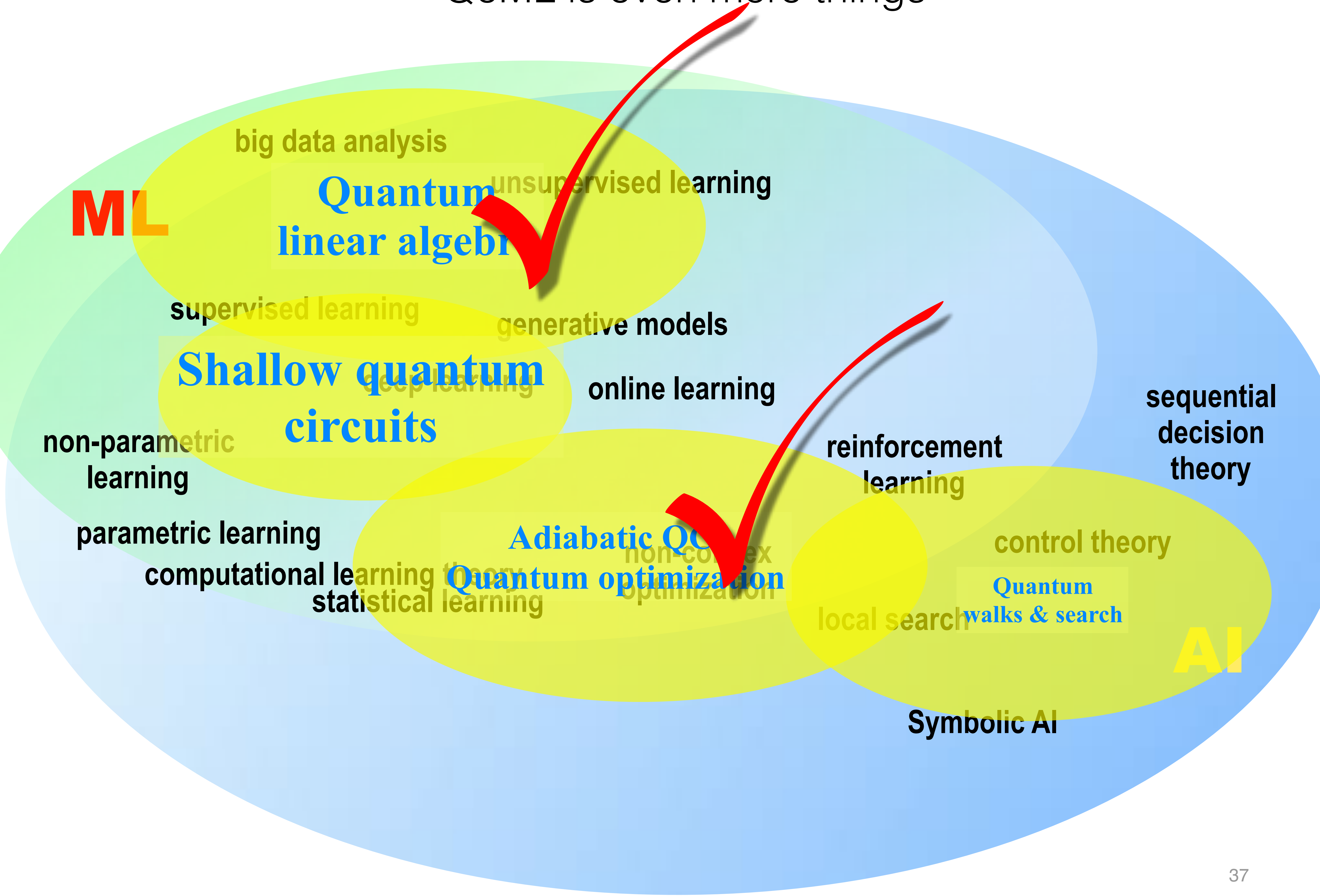
manipulate exponentially-sized data-vectors in system (qubit) number

HOWEVER

need full blown ideal QC
need pre-filled database (QRAM)
need appropriate condition numbers
need robustness to linear error
need right preprocessing applied
can sometimes be done classically…

STILL A GREAT IDEA!!

QeML is even more things



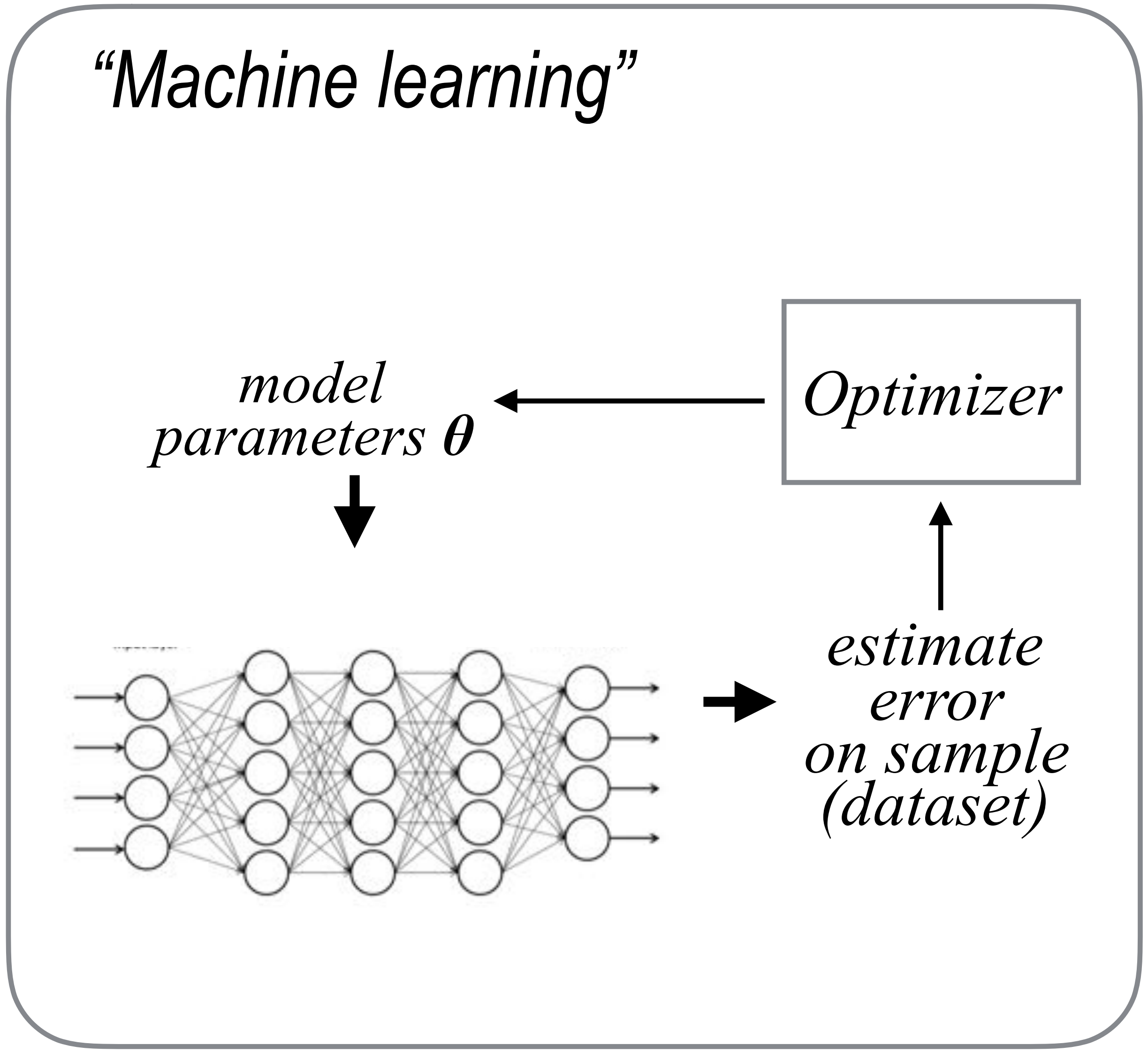
Quantum-enhanced supervised learning: the quantum pipeline

a) The optimization bottleneck — **quantum annealers**

b) Big data & comp. complexity — **universal QC and Q. databases**

c) Machine learning Models — **restricted (shallow) architectures**

(Quantum) Machine learning Models



Improving ML == speeding up algorithms… *or is it?*

Machine learning Models

A lot of machine learning:

-Take my (training) dataset *{(point, label)}*

-Take a ***model*** (tensorflow tutorials will suggest), e.g. *this-that-structure* neural network $\mathcal{N}$

-Train the ***model*** (tweak parameters of $\mathcal{N}$, until it predicts the training set well)

The math behind

"cost function"

$parametrized\ family\ \{f_\theta\}$

$\text{argmin}_\theta\ \text{Err\_training\_set}(\theta)$

What is this picture missing?

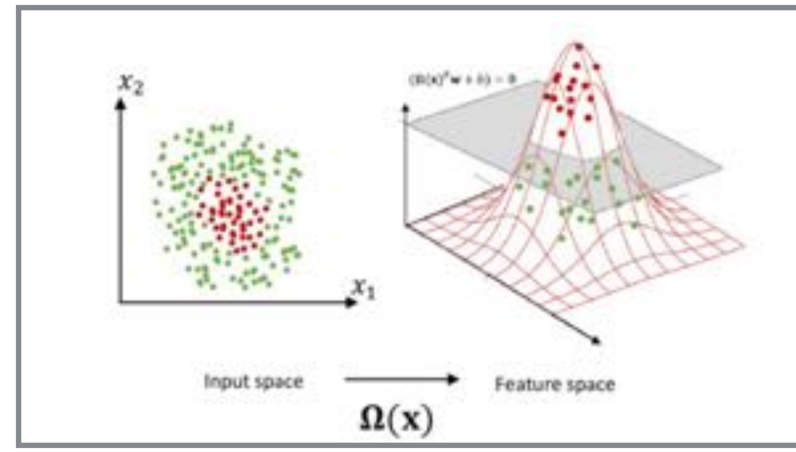
Optimization is *a part of the method, not the objective*

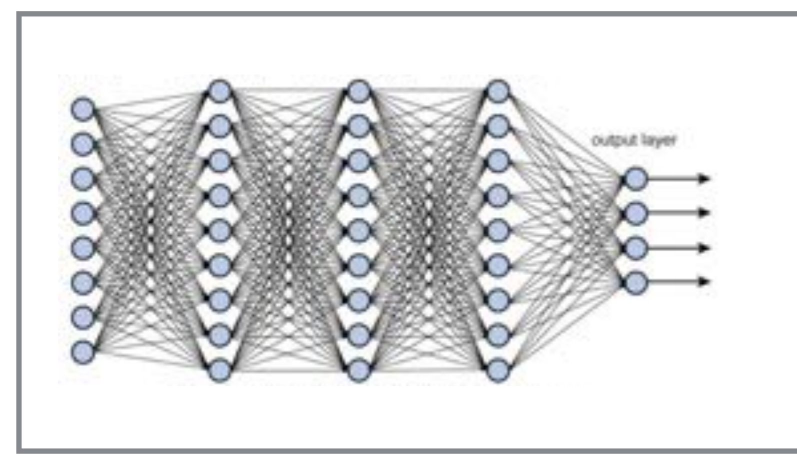
best fit v.s. "generalization performance" ***or*** *classifying well* ***beyond the training set***

Data:

Challenge:

Models:

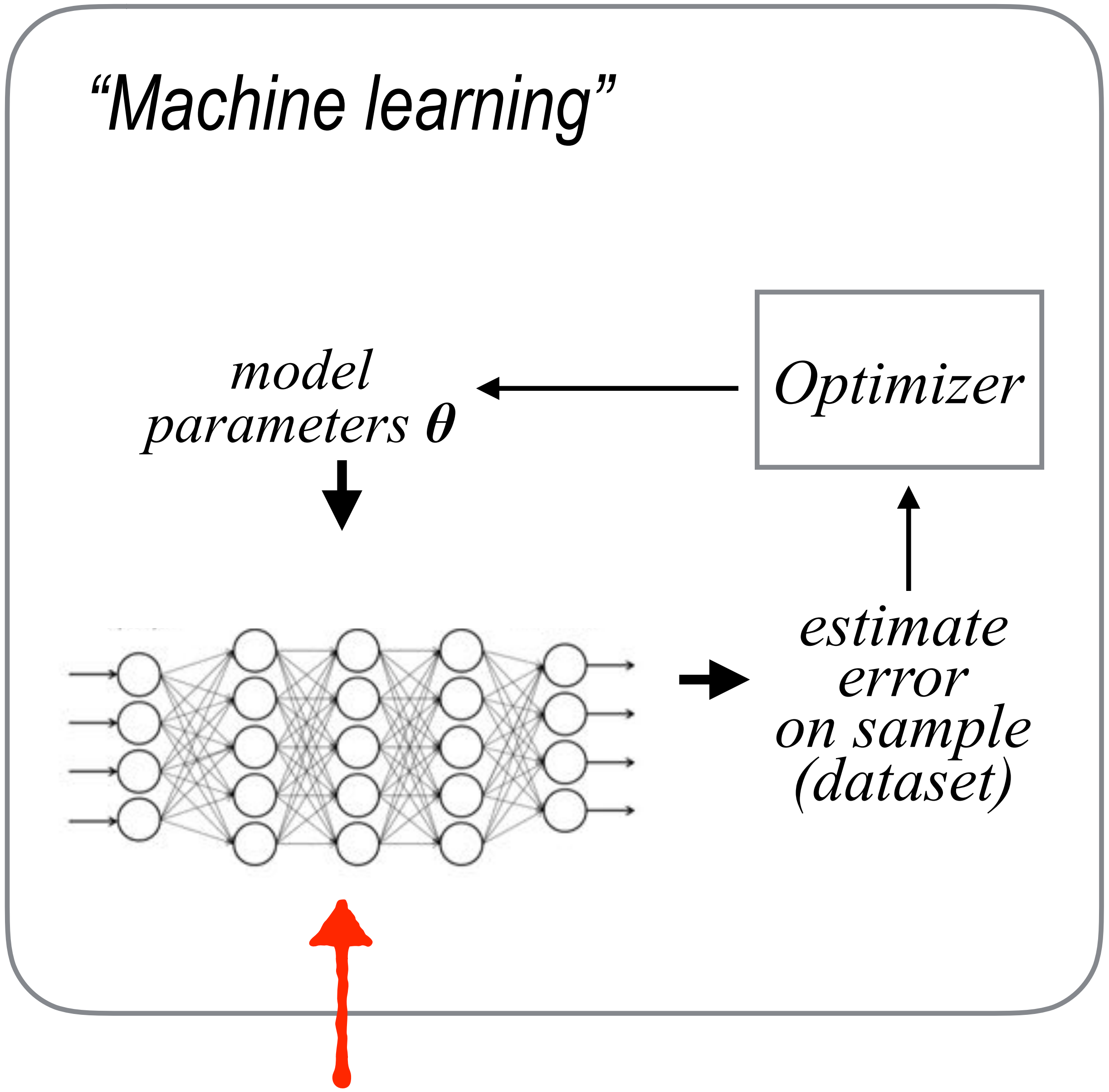




Not all models (+training algo) are born equal (for *real* datasets)…

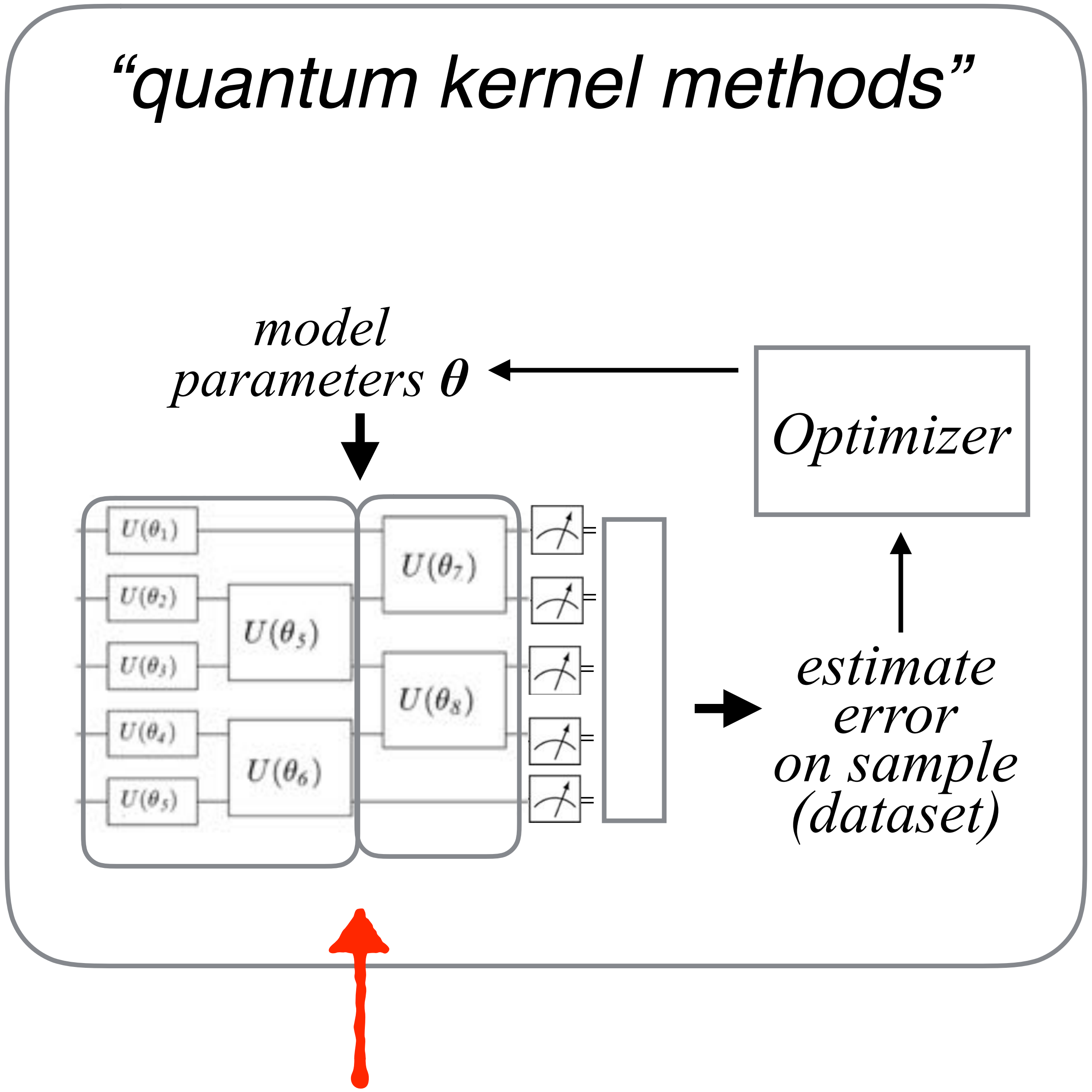
Image: 10.1016/j.compstruct.2018.03.007

Machine learning Models



family of functions.
if it's "good", we can generalize well

Quantum Machine learning Models

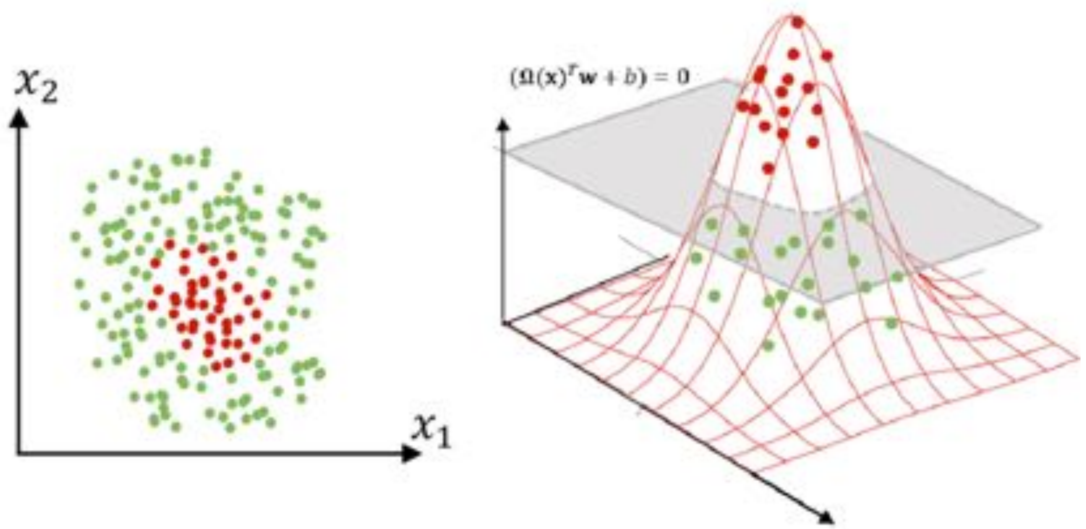


How about "shallow quantum circuits"?
-instead neural network, train a QC!
-related to ideas from q. condensed-matter physics (VQE)

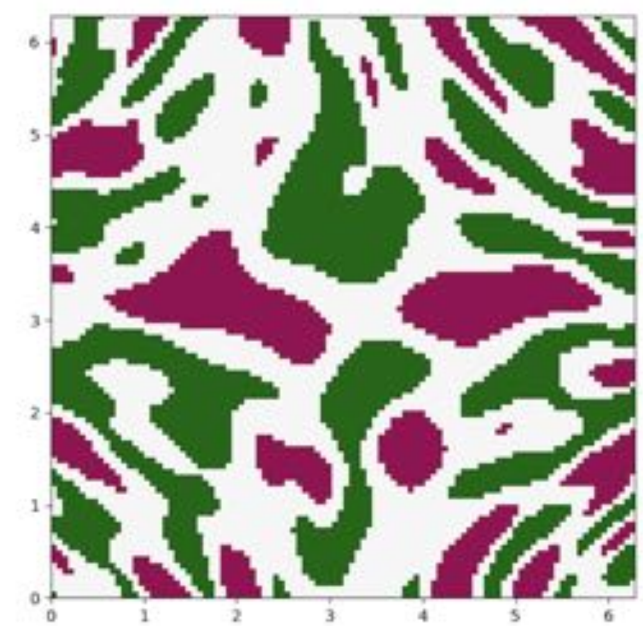
Phys. Rev. Lett. 122, 040504 2019
Nature 567, 209–212 (2019)
(c.f. Elizabeth Behrman in '90s)

The quantum feature space

- relationship between NNs, SVMs and shallow circuits for supervised learning
  (embedding - rotation - measurement = feature function - hyperplane - class)



Simple classical kernels



A weird quantum kernel

Quantum Machine learning Models
“quantum kernel methods”

<u>The good</u>

- near term architectures
- **seems to be *robust***
  <u>(noise not inherently critical!)</u>
- *possibly* ***very expressive***

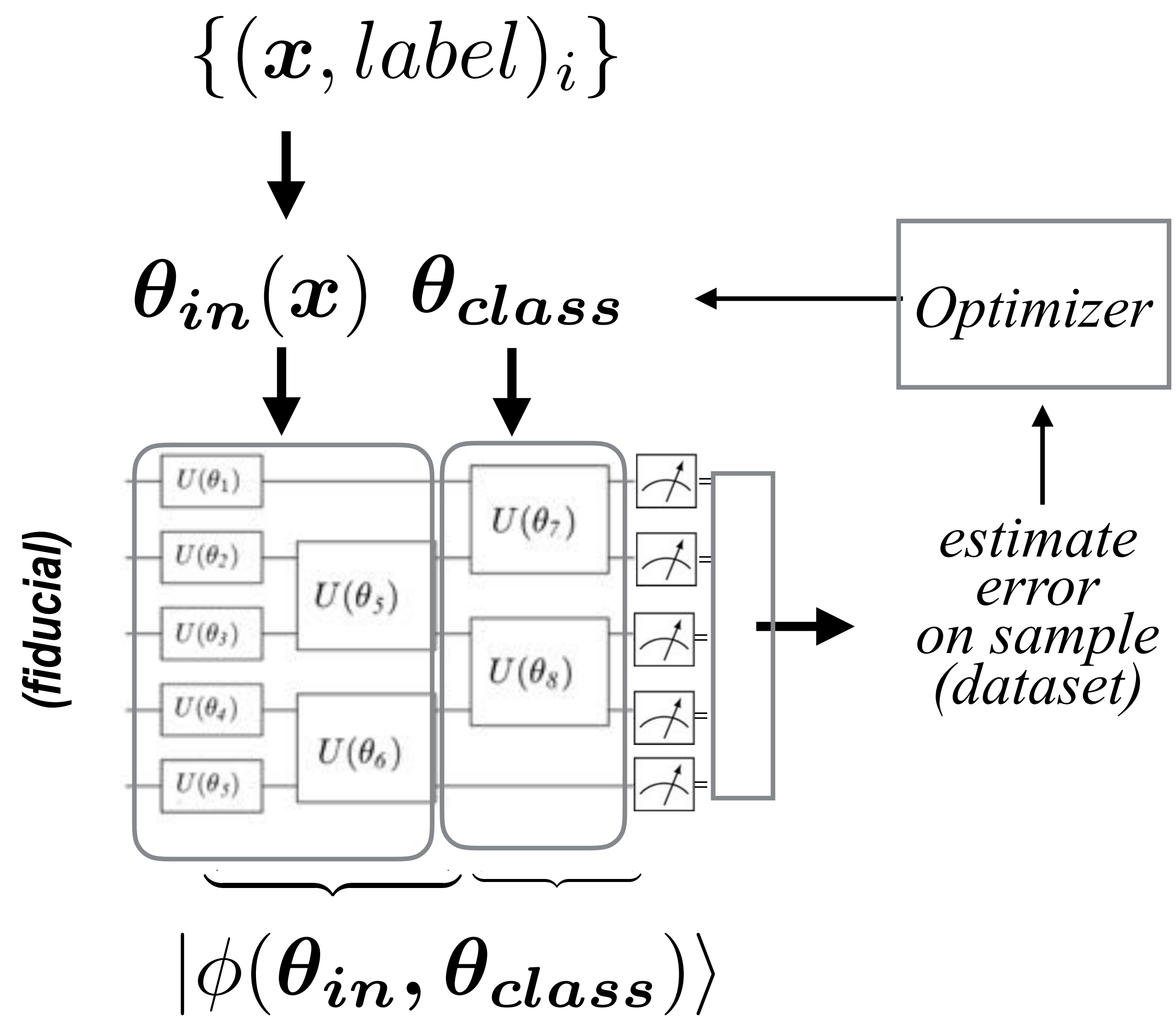
<u>The neutral</u>

- ***many parameters***
- ***model advantages*** less clear

(*contrast to variational methods!*)

<u>The bad</u>

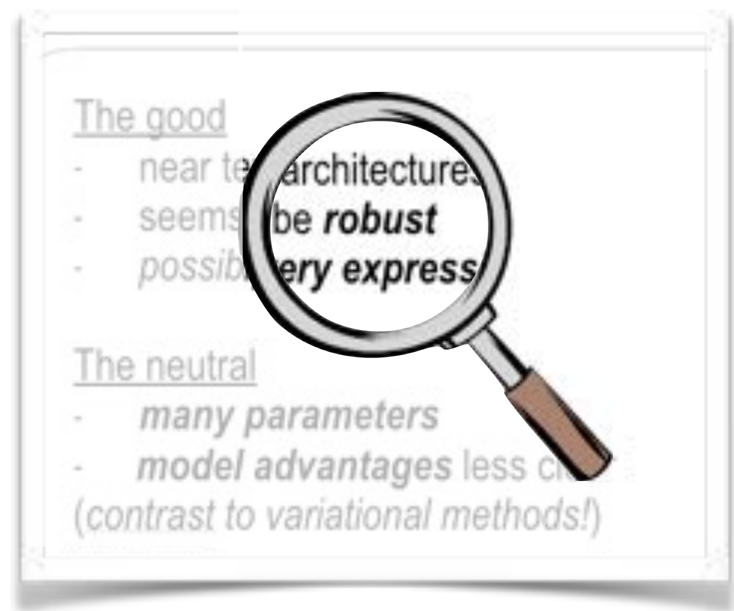
- ***barren plateaus*** *(also in DNN)*



CAVEAT: IS IT CLASSICALLY COMPUTATIONALLY HARD?!

Phys. Rev. Lett. 122, 040504 2019
Nature 567, 209–212 (2019)

A hope… *killer app for noisy QCs?*

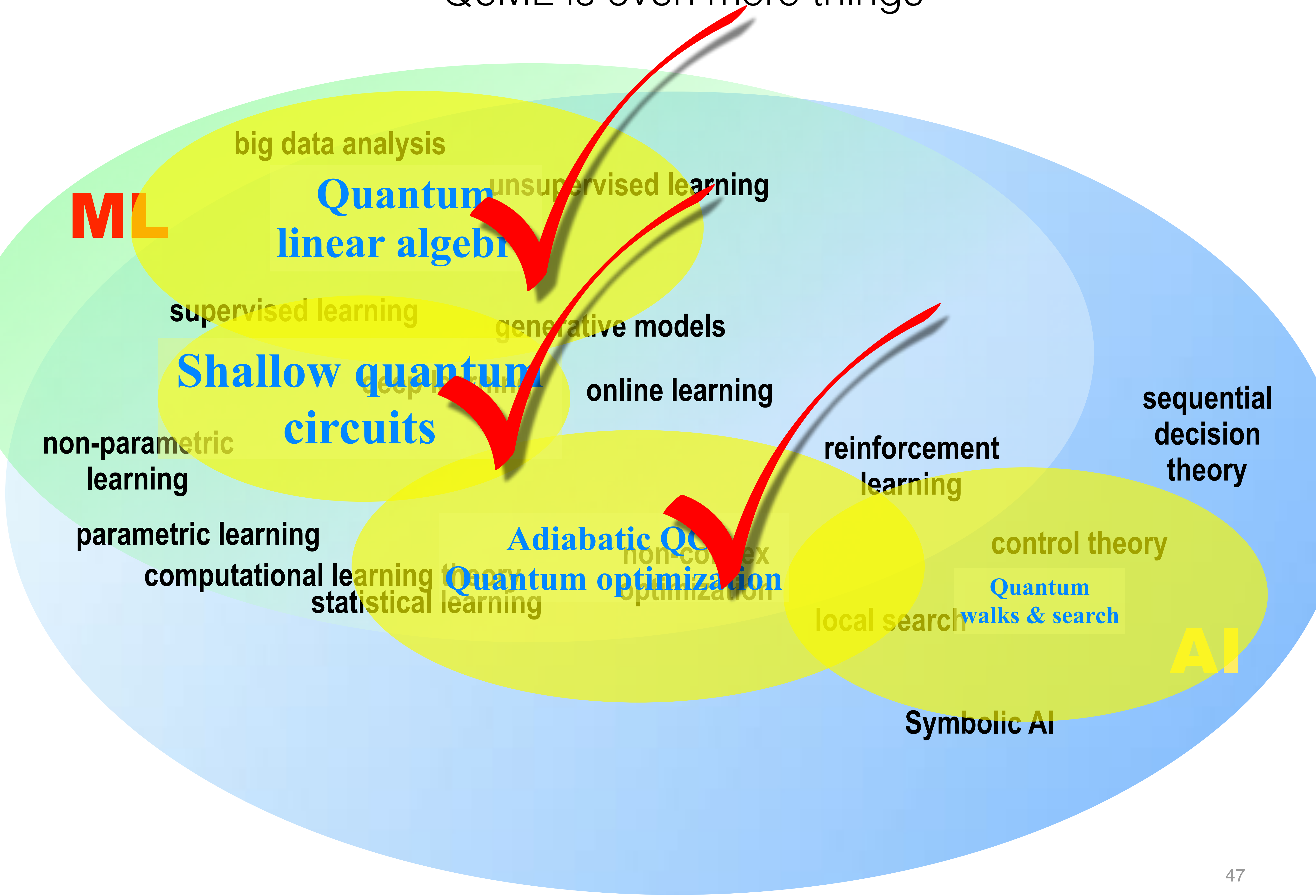


ML can be run on small QCs

BUT MORE THAN THAT

ML good for dealing with noise (in *data*)…
Can QML deal with *its own* noise (in *process*)?

QeML is even more things



Application, match, … conspiracy?

- Nice analogy Hilbert spaces - big data spaces
- Hard optimization (needed) - hard optimization (delivered)
- New learning models (needed) - shallow QC (delivered)

Application, **match**, … conspiracy?

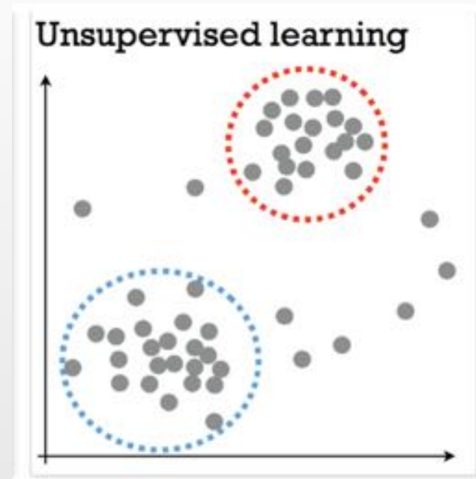
- Nice analogy Hilbert spaces - big data spaces
  Problem: preparations can offset speed-up;
  ML: not here! processing must be robust -> low cost

- Hard optimization (needed) - hard optimization (delivered)
  Problem: optimization just heuristic, quality unknown
  ML: well all we do is domain-specific! If it works, it works!

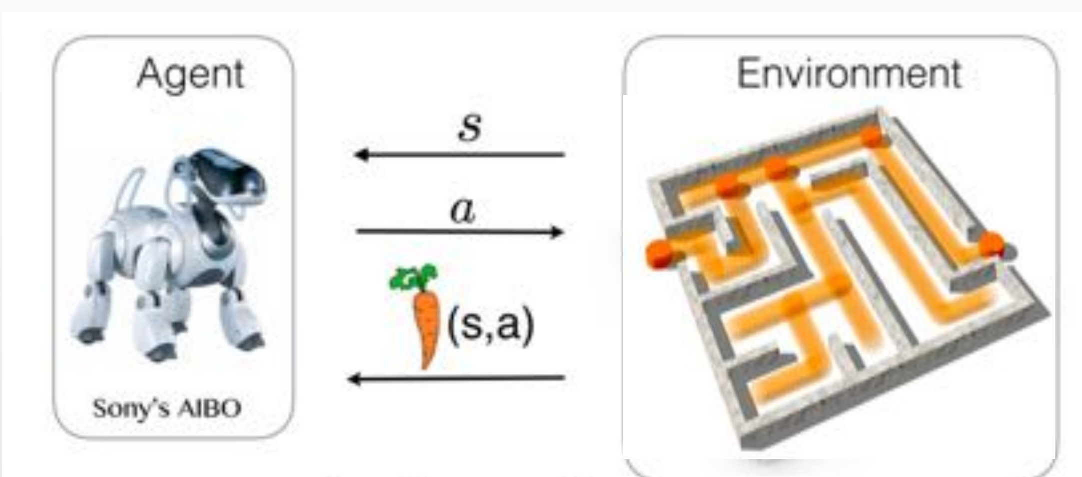
- New learning models (needed) - shallow QC (delivered)
  Problem: noisy models, bad estimates (in VQE)
  ML: not estimating! Train model, could be even better than exact (elements of regularization)

Application, **match**, … conspiracy?

Application, **match**, … conspiracy?



Quantum-enhanced unsupervised learning

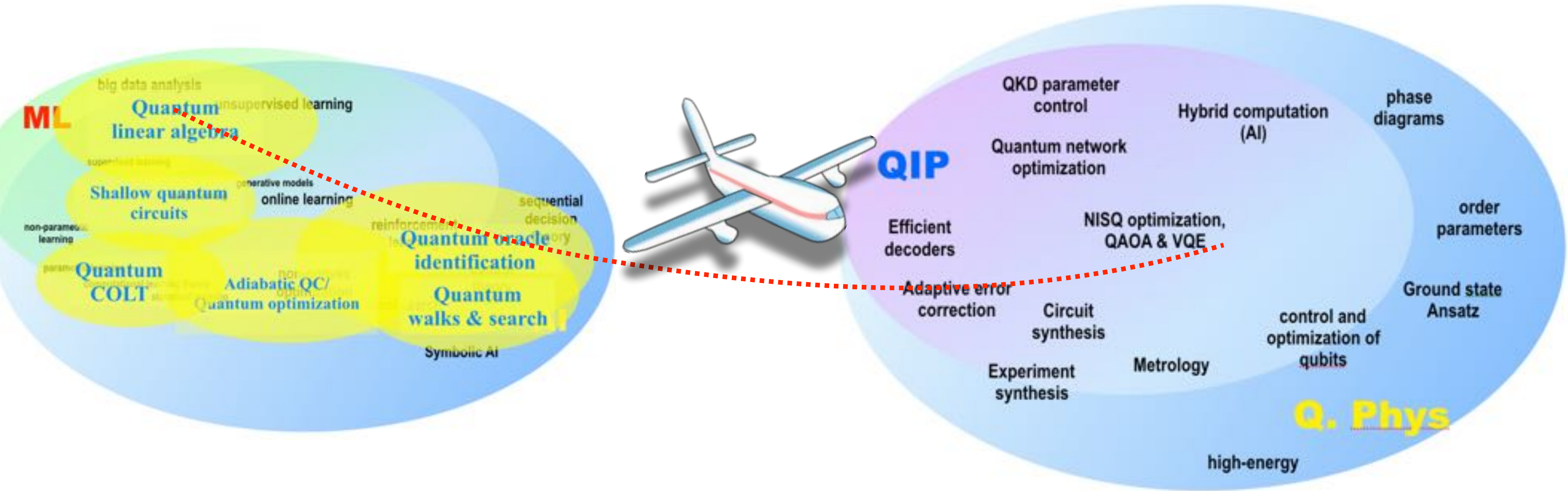


Quantum-enhanced reinforcement learning

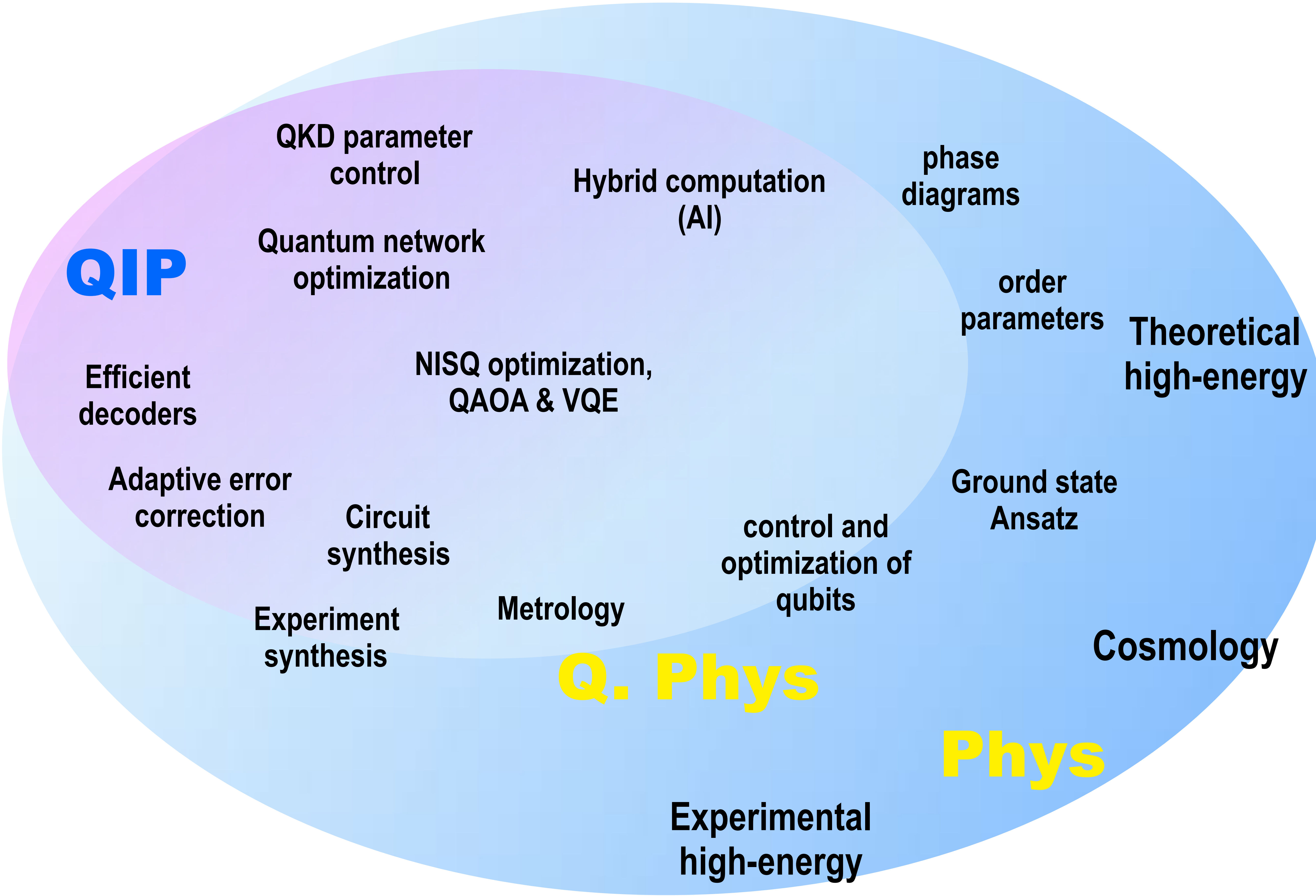
Towards quantum AI

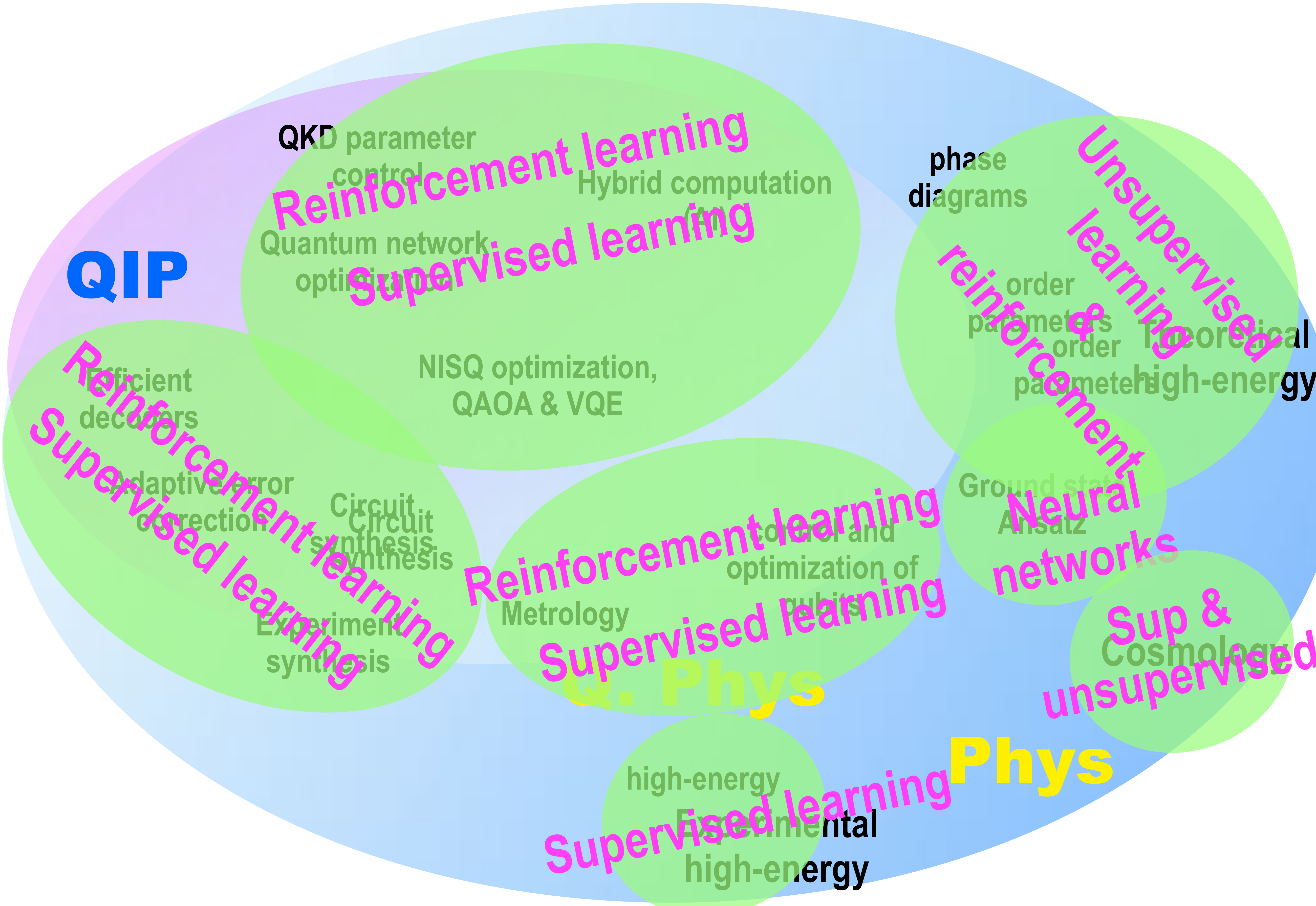
Application, **match**, … conspiracy?

Application, **match**, … conspiracy?



Machine learning
in the physics domain


QKD parameter control
Hybrid computation (AI)
phase diagrams
Quantum network optimization
QIP
order parameters
Theoretical high-energy
NISQ optimization, QAOA & VQE
Efficient decoders
Ground state Ansatz
Adaptive error correction
Circuit synthesis
control and optimization of qubits
Metrology
Experiment synthesis
Cosmology
Q. Phys
Phys
Experimental high-energy


QKD parameter control
Reinforcement learning
Hybrid computation (AI)
Supervised learning
Quantum network optimization
QIP
NISQ optimization,
QAOA & VQE
Efficient decoders
Reinforcement learning
Supervised learning
Adaptive error correction
Circuit synthesis
Circuit synthesis
Experiment synthesis
Reinforcement learning
Control and optimization of qubits
Metrology
Supervised learning
Q. Phys
phase diagrams
Unsupervised learning & reinforcement
order parameters
order parameters
Theoretical high-energy
Ground state Ansatz
Neural networks
Sup & unsupervised
Cosmology
Phys
high-energy
Experimental high-energy
Supervised learning

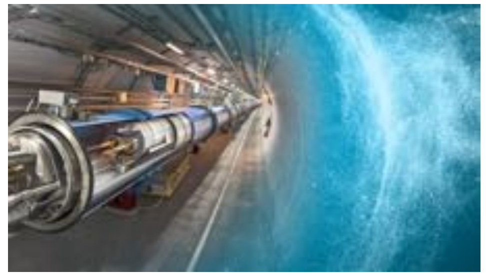
<u>Supervised learning</u>
(learning *how to label datapoints, learning how to approximate a function, how to classify*)

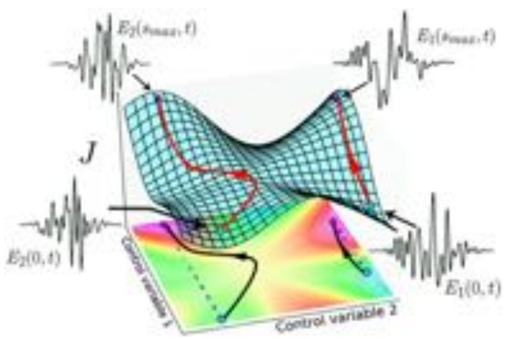
<u>Unsupervised learning</u>
(learning *a distribution, generate. properties from samples, feature extraction & dim. reduction*)

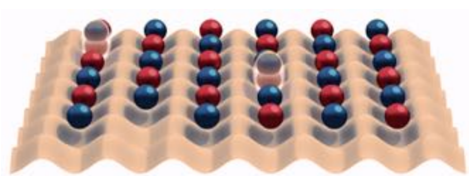
<u>Reinforcement learning</u>
(learning *behavior*, *policy*, or *optimal control*)

Big picture

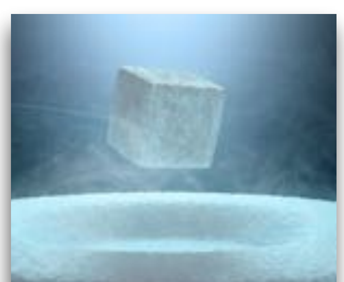
200-petabyte
(2017!)



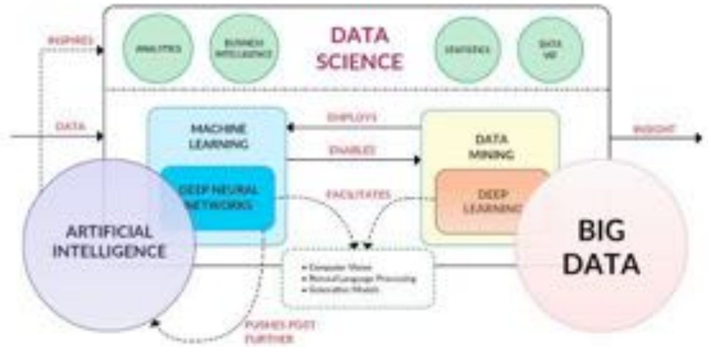


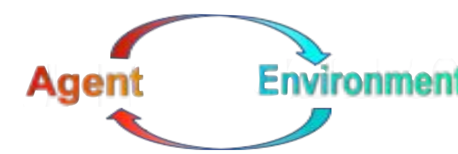


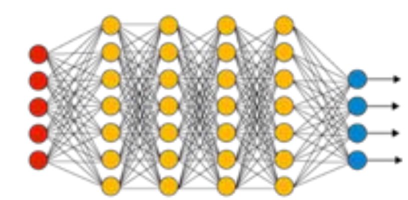
$\hat{H} = -t \sum_{\langle i,j\rangle,\sigma} \hat{c}^\dagger_{i,\sigma}\hat{c}_{j,\sigma} + U\sum_{j} \hat{n}_{j,\uparrow}\hat{n}_{j,\downarrow} + J\sum_{\langle i,j\rangle} \hat{\mathbf{S}}_i\cdot\hat{\mathbf{S}}_j$



hard computations
new theory & experiments







AI/ML assisted computation
machine-assisted research

Figure from: https://hackernoon.com/how-big-data-is-empowering-ai-and-machine-learning-4e93a1004c8f

Particle physics (and cosmology)

Many-body quantum matter

Chemistry and materials

Facilitating quantum computers

“Machine learning and the physical sciences”
Carleo et al., https://arxiv.org/pdf/1903.10563.pdf

Particle physics and cosmology

-“big data” aspects: event selection, jet tagging, triggering;
(photometric red shift, gravitational lens finding)

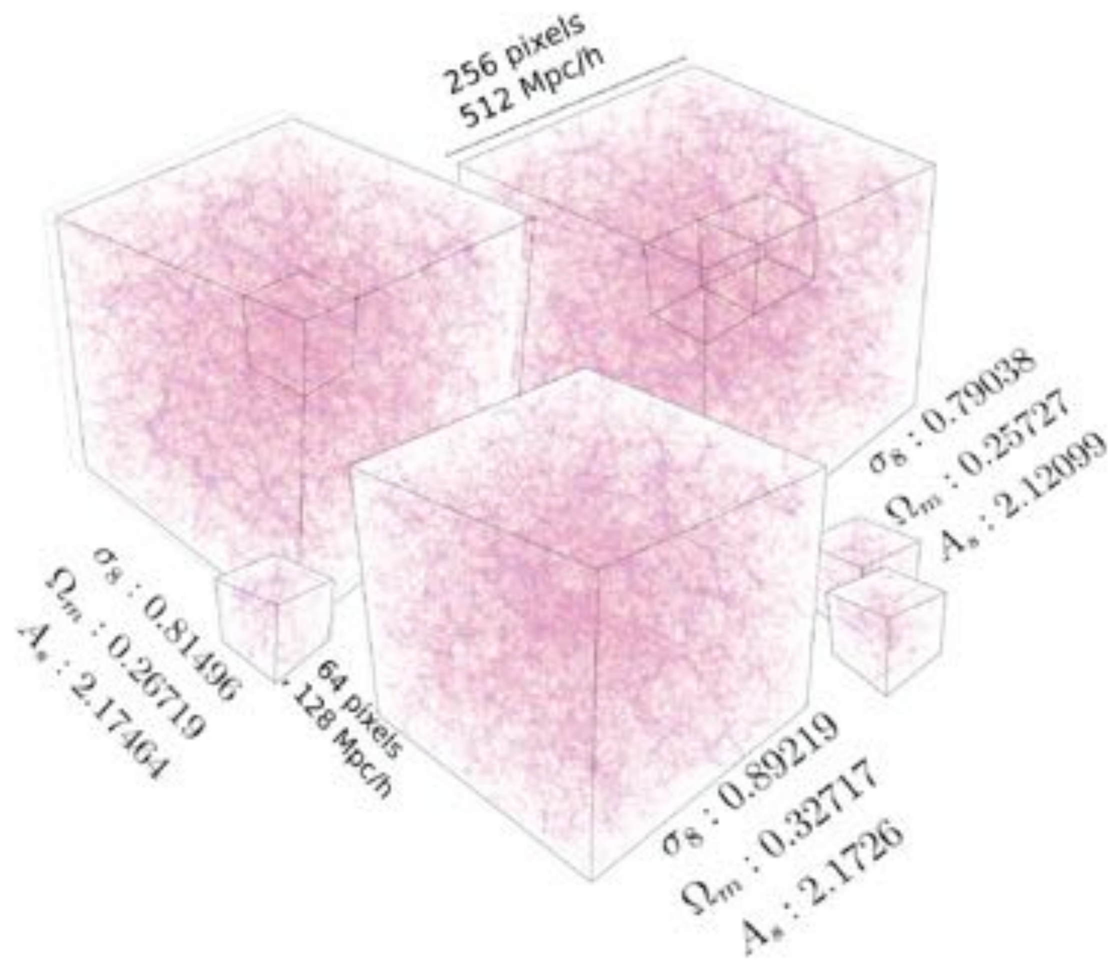
-simulation and inverse problems

-applications in theory

“Machine learning and the physical sciences”
Carleo et al., https://arxiv.org/pdf/1903.10563.pdf

Example: *Estimating Cosmological Parameters from the Dark Matter Distribution*

$$\Lambda CDM$$
$$\text{(cosm. parameters)} \longrightarrow \text{distr. of matter}$$



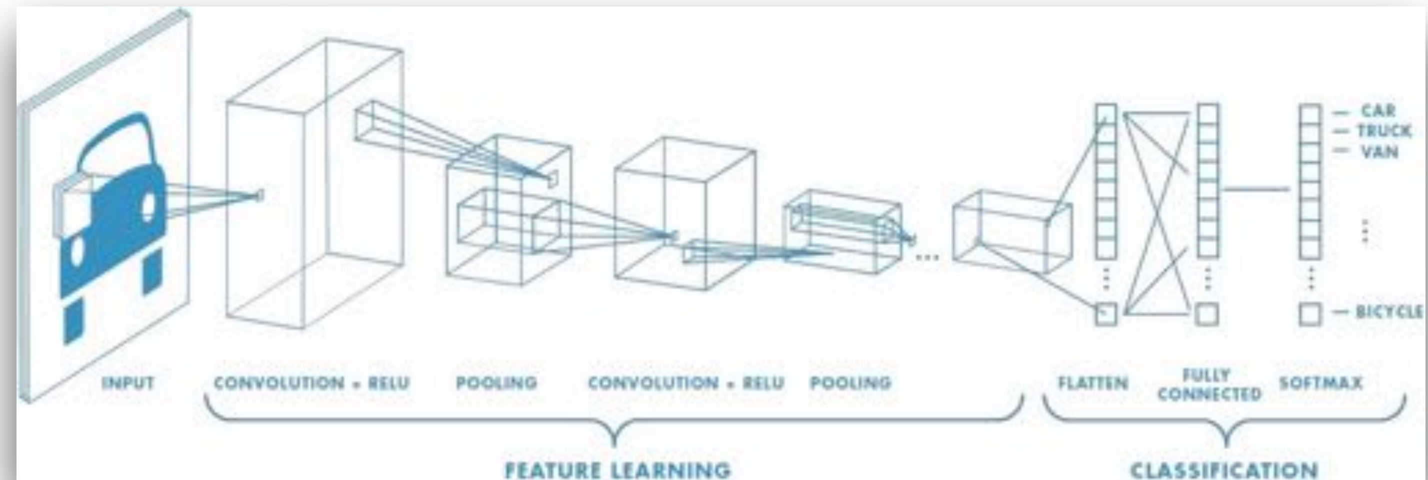
What are the cosmological parameters from observed universe?

"Inverse simulation?"

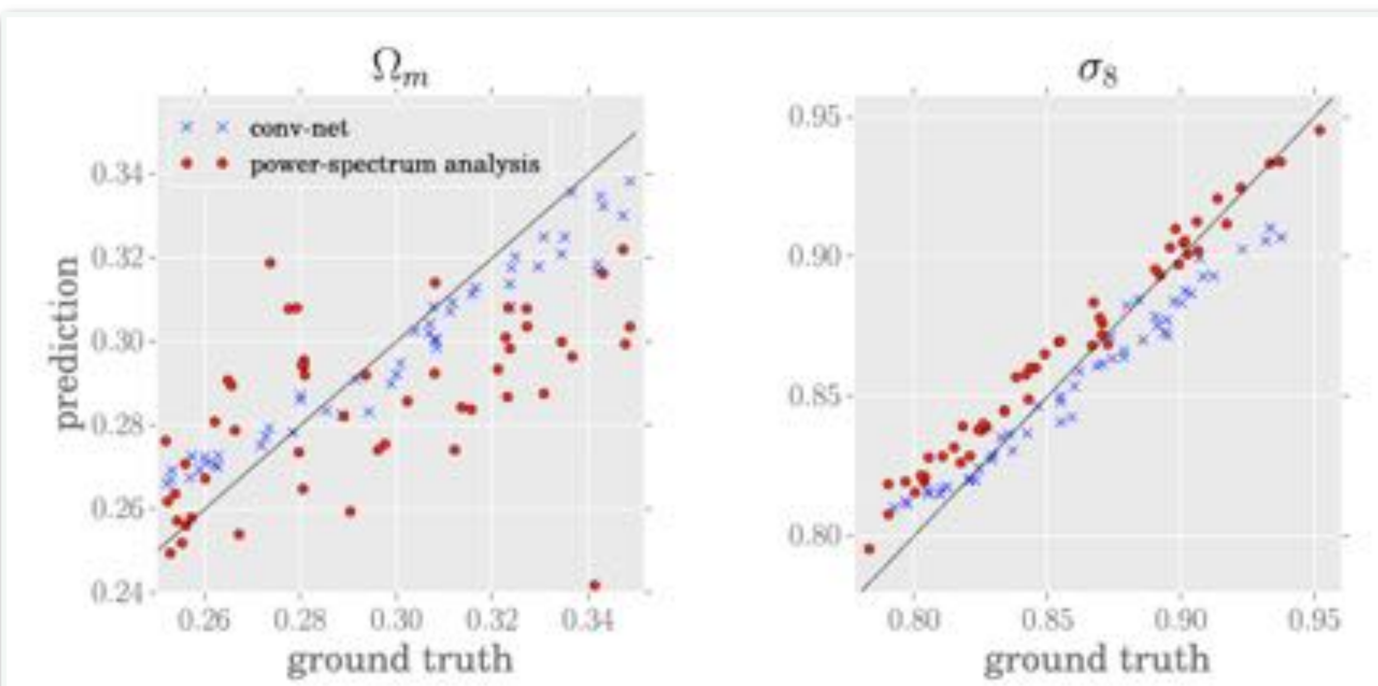
arXiv:1711.02033v1

Example: *Estimating Cosmological Parameters from the Dark Matter Distribution*

Machine learning solution:



Train NN to output *correct parameters* given <u>the universe</u>;
Training set: $(universe, parameters)$
Learning goal: $(parameters \mid universe)$



arXiv:1711.02033v1

Many-body quantum matter

- -neural quantum states (approximate *the wavefunction*)
    - -expressivity, learning from data, variational approaches
- -assisted many-body simulations
    - -learned hard sampling
- -classification of many-body phases of matter

“Machine learning and the physical sciences”
Carleo et al., https://arxiv.org/pdf/1903.10563.pdf

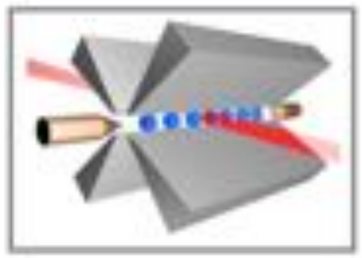
Machine learning
in quantum information processing

Enabling quantum information processing devices

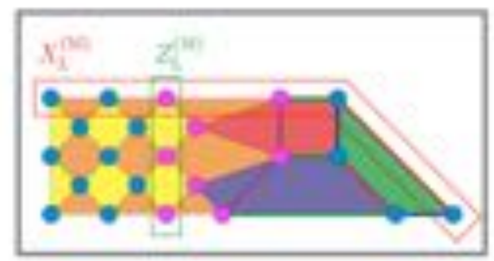
Interplay at various levels of complexity and abstraction

quantum control

choose parameters
(time-dependently)
to realize a desired
unitary operation



optimizing quantum communication & computation



Optimize local error
correction, and fight
noise processes
adaptively and
autonomously

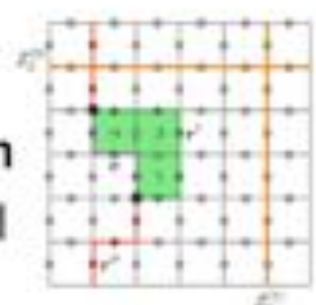
economic (quasi)-algorithmic strategies

-ML-enhanced classical pre&post-processing

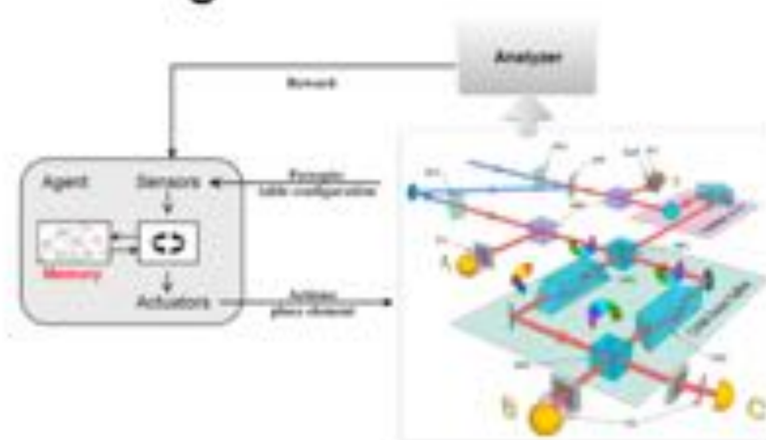
-decoders for error correction

-quantum memories (dynamical decoupling)

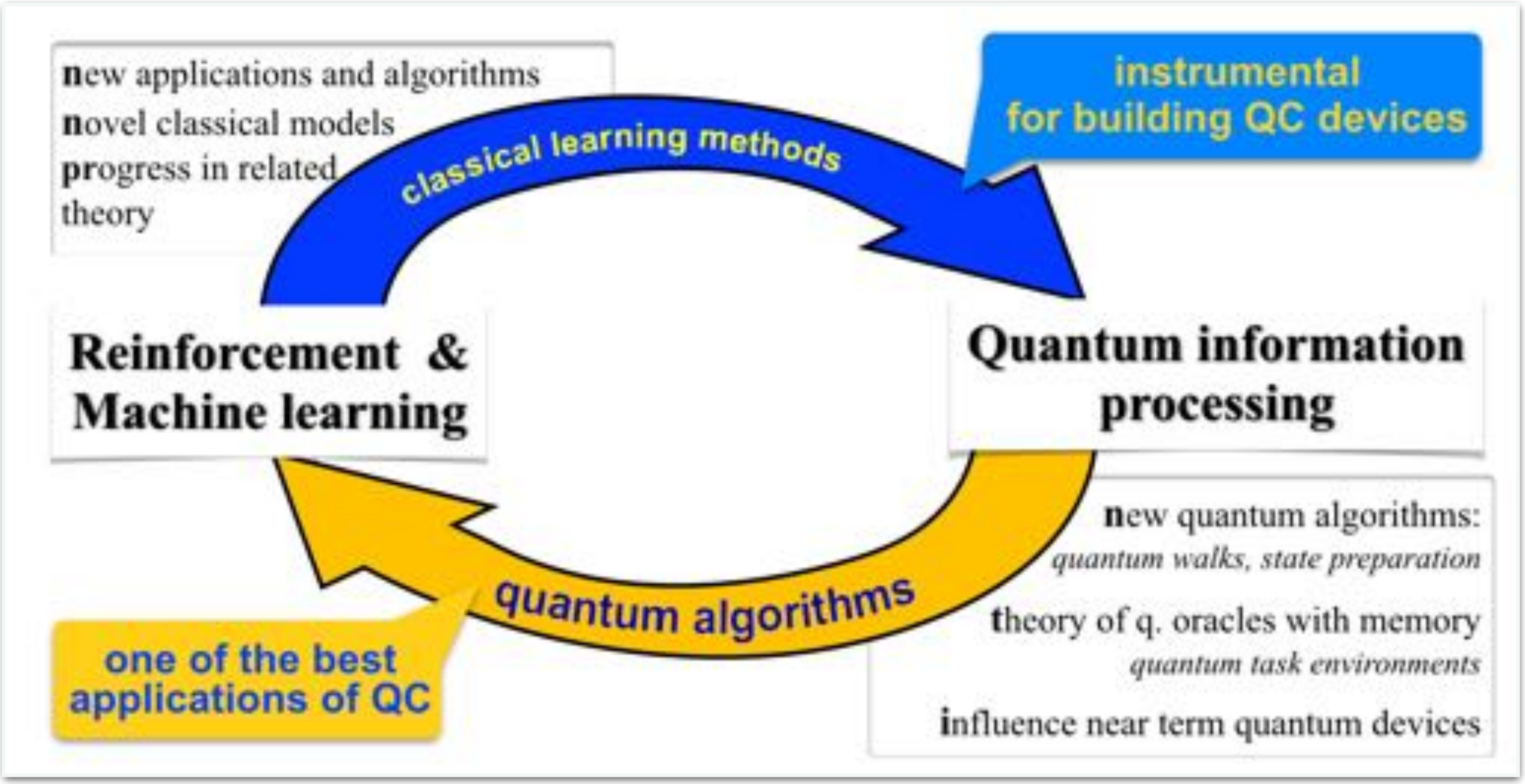
-tailored error correcting codes

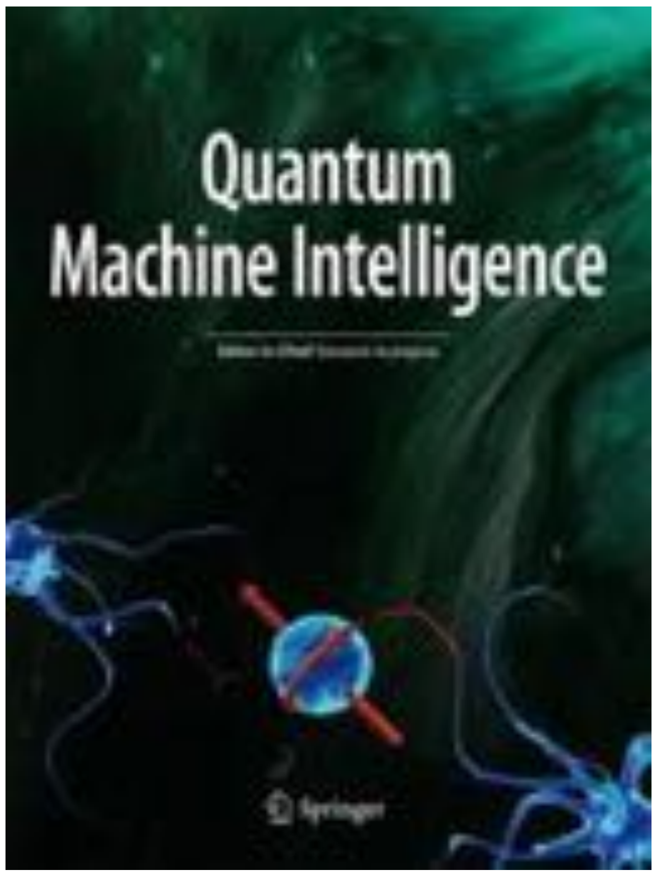


designing new experiments and doing research



Application, match, … **conspiracy**?





Editor-in-Chief

Giovanni Acampora, University of Naples Federico II, Italy

CALL FOR PAPERS

Field Editors

1) Quantum Machine Learning

Seth Lloyd (MIT), USA

2) Quantum Computing for Artificial Intelligence

Hans Jürgen Briegel, (Innsbruck, Austria)

3) Artificial Intelligence for Quantum Information Processing

Chin-Teng Lin (Sydney, Australia)

4) Quantum- and Bio-inspired Computational Intelligence

Francisco Herrera (Granada, Spain)

5) Quantum Optimization

Davide Venturelli (USRA, USA)