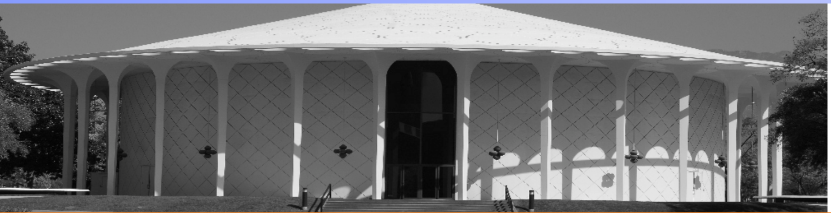
CALIFORNIA INSTITUTE OF TECHNOLOGY



universitätbonn

Data-Driven Mechanics: Constitutive Model-Free Approach

$$\inf_{y\in D}\inf_{z\in E}\|y-z\| = \inf_{z\in E}\inf_{y\in D}\|y-z\|$$

Michael Ortiz – Lecture 1
California Institute of Technology and
Rheinische Friedrich-Wilhelms Universität Bonn

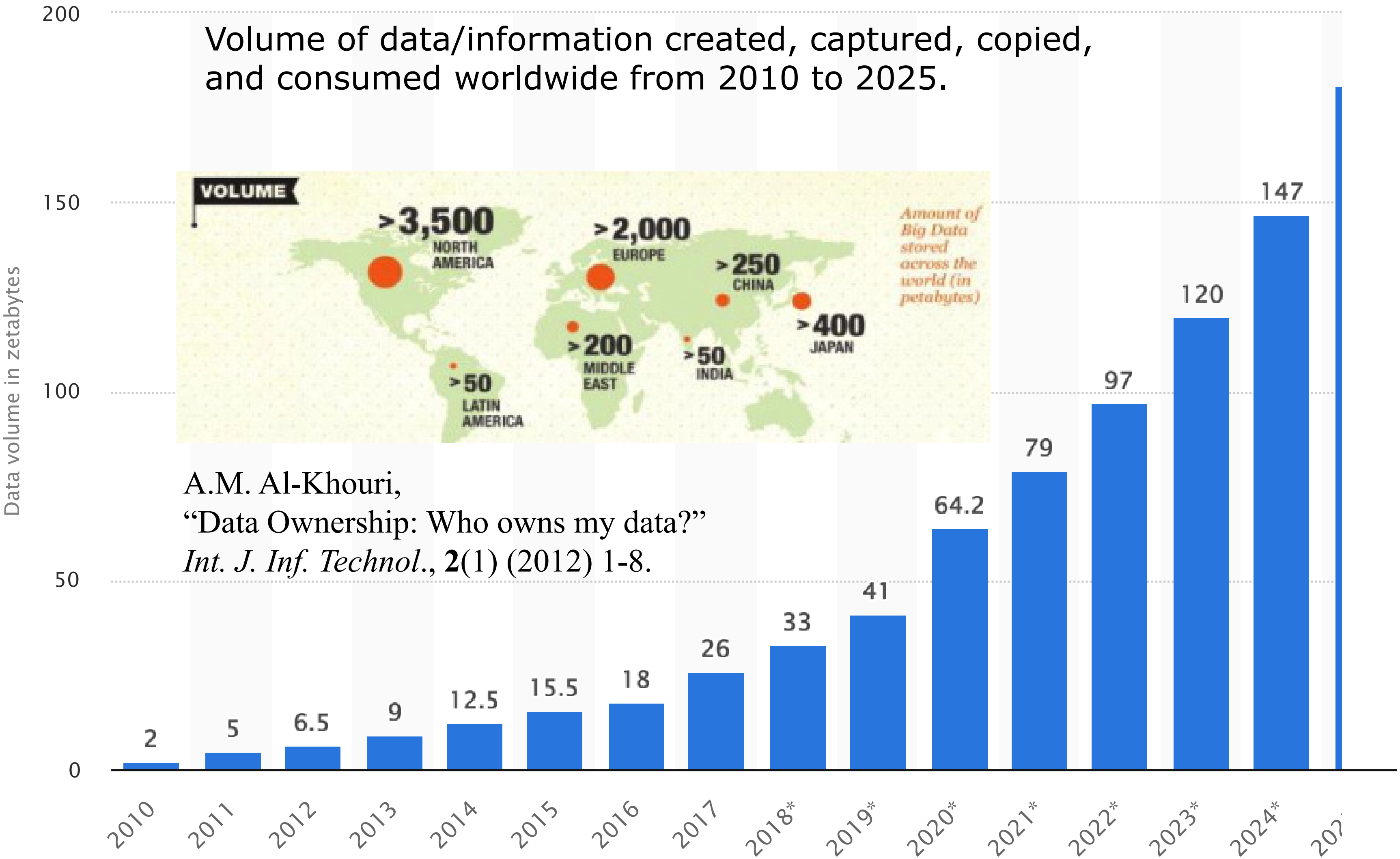
Centre International des Sciences Mécaniques (CSIM)
Udine (Italy), October 10-14, 2022

Why Data-Driven Computing Now?

- The world's data volume has increased dramatically over the past $\sim$20 years.

Why Data-Driven Computing Now?

Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2025.



A.M. Al-Khouri,
"Data Ownership: Who owns my data?"
Int. J. Inf. Technol., **2**(1) (2012) 1-8.

https://www.statista.com/statistics/871513/worldwide-data-created/

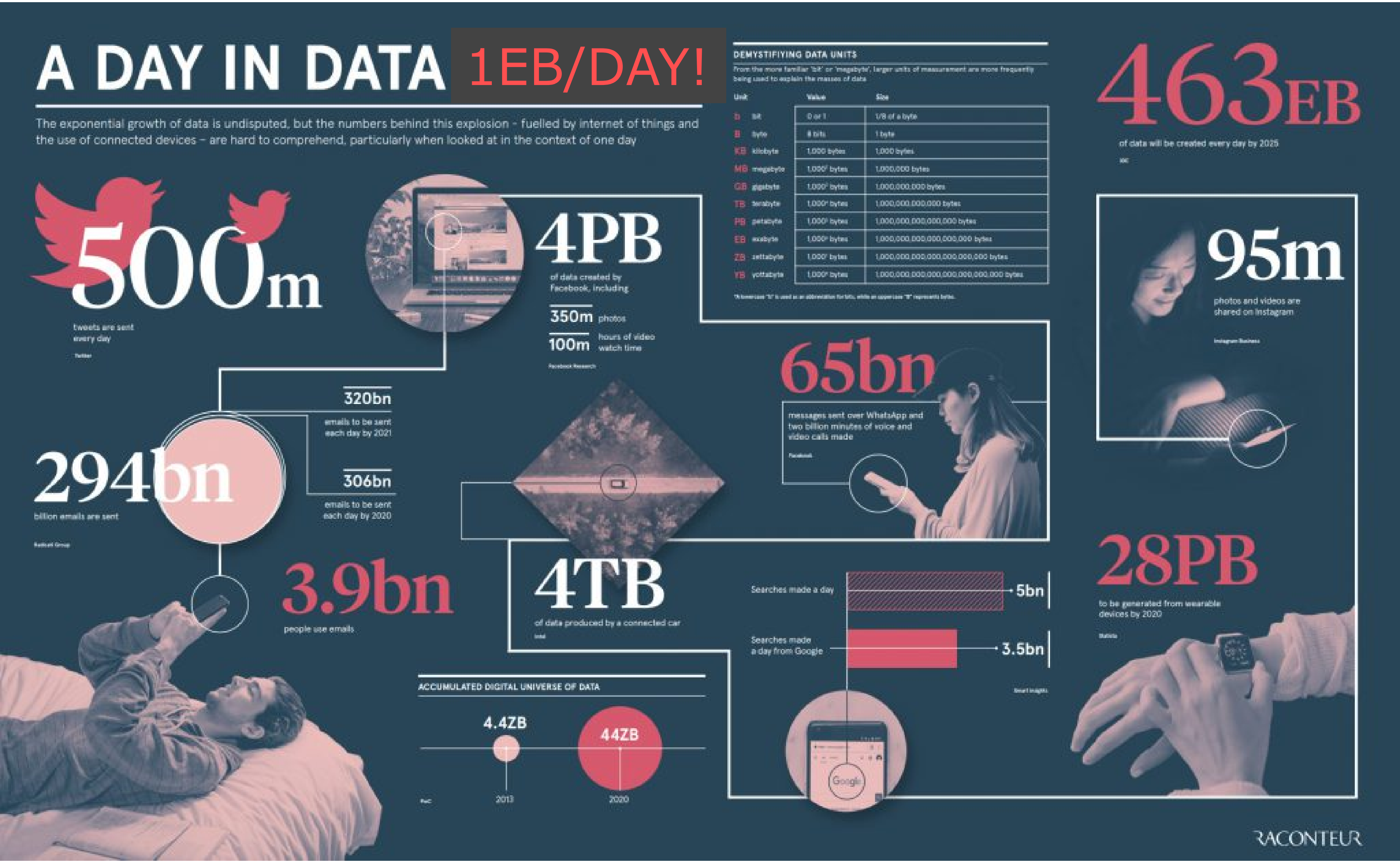
Why Data-Driven Computing Now?

• The world's data volume has increased dramatically over the past ∼20 years.

• With the advent of *cloud infrastructure* *circa* 2000, previous limits on storage became obsolete.

• As billions of new users gained *internet access* across the globe, data generation increased enormously.

Why Data-Driven Computing Now?



Why Data-Driven Computing Now?

• The world's data volume has increased dramatically over the past $\sim$20 years.

• With the advent of *cloud infrastructure* *circa* 2000, previous limits on storage became obsolete.

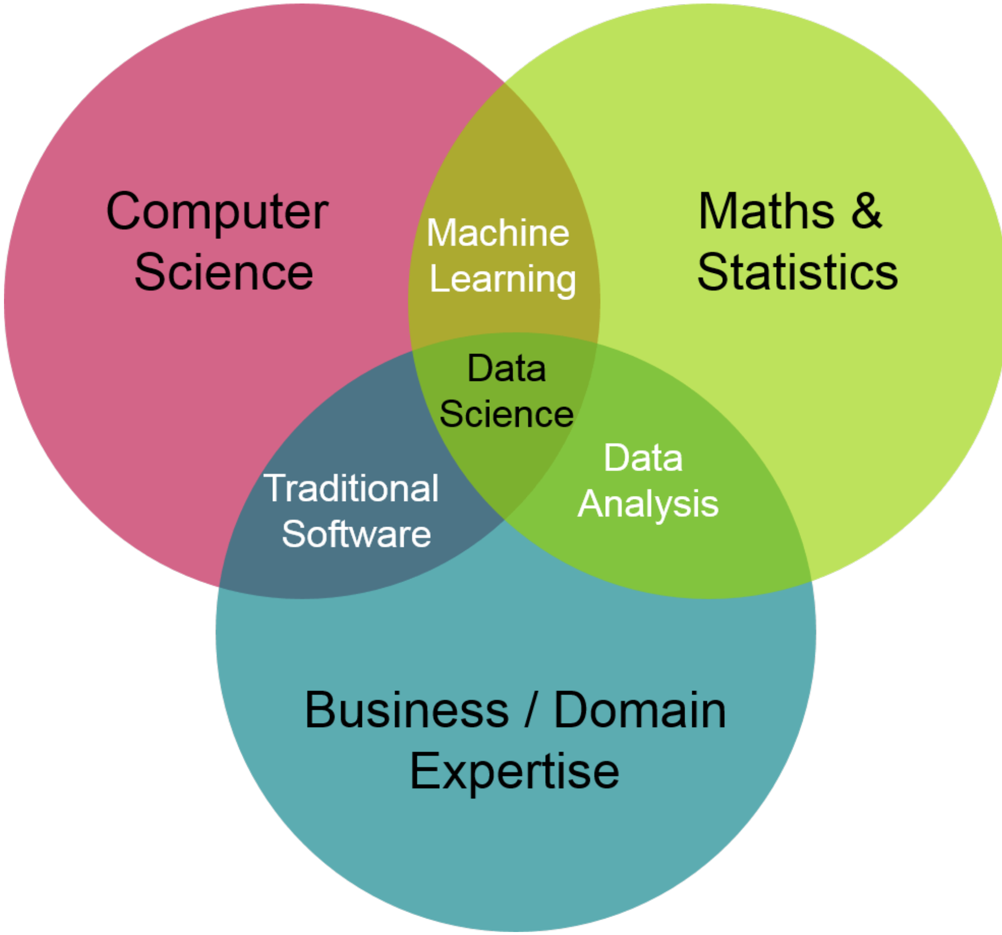
• As billions of new users gained *internet access* across the globe, data generation increased enormously.

• Soon, companies around the world adopted *Big-Data* strategies, developed tools.

• Tools, algorithms, for learning—and making predictions, decisions—from big-data sets are now generically referred to as *Machine Learning*.

• Machine Learning is a modern version of *statistical analysis* (regression, testing, inference...) and *data mining* (discovering patterns in data).

Why Data-Driven Computing Now?



Great Learning Team,
What is Data Science and How Does it Work:
A Complete Beginner's Guide, Jan 11, 2022.

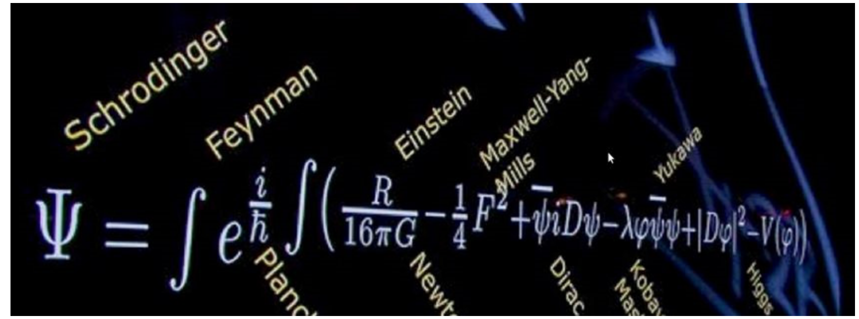
Why Data-Driven Computing Now?

- The world's data volume has increased dramatically over the past $\sim$20 years.
- With the advent of *cloud infrastructure* *circa* 2000, previous limits on storage became obsolete.
- As billions of new users gained *internet access* across the globe, data generation increased enormously.
- Soon, companies around the world adopted *Big-Data* strategies, developed tools.
- Tools, algorithms, for learning—and making predictions, decisions—from big-data sets are now generically referred to as *Machine Learning*.
- Machine Learning is a modern version of *statistical analysis* (regression, testing, inference...) and *data mining* (discovering patterns in data).
- Key enablers (what is old? what is new?):
  - i) Advent of *high-performance and distributed computing*.
  - ii) *Cost of storing* and managing big data sets dramatically lowered.
  - iii) *Commercial data sets* available (weather, social media, medical).
  - iv) Tools available through *open-source communities* with large user bases.
- *Sea-change* in the way data is generated, managed and utilized! Here to stay, cannot be ignored by the scientific computing community.

The anatomy of a field theory

- How does Data Science intersect with *scientific computing*?
- Scientific computing deals with the *field theories* of physics.

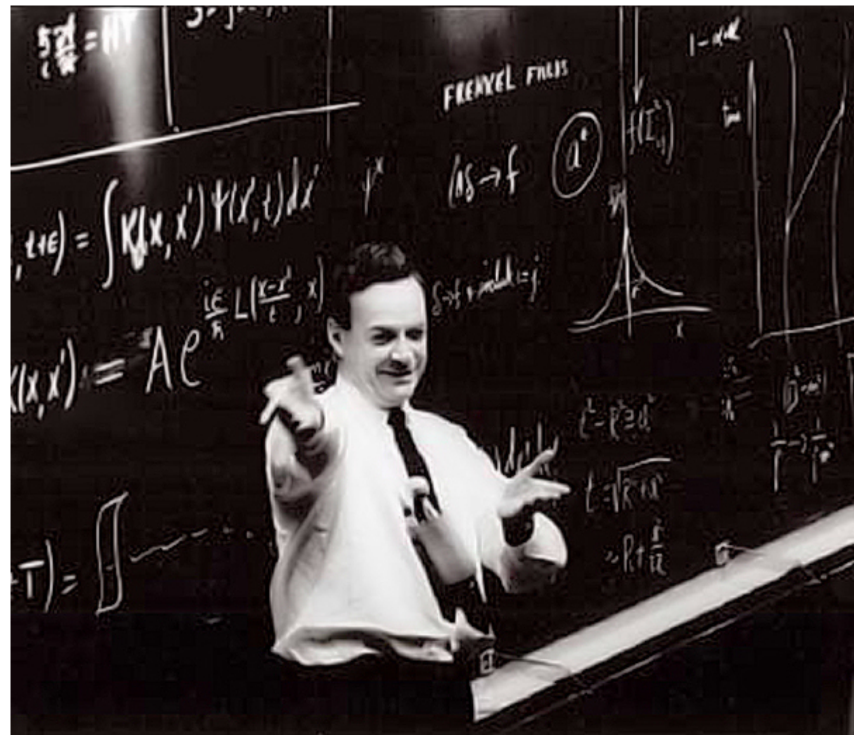
The anatomy of a field theory



James Clerk Maxwell
June 13, 1831 (Edinburgh)
November 5, 1879 (Cambridge)

$$\mathcal{L} = -\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu} - A_\mu J^\mu$$

NB: Duality structure!



Richard Feynman
May 11, 1918 (New York City)
February 15, 1988 (Los Angeles)

The anatomy of a field theory

- How does Data Science intersect with *scientific computing*?
- Scientific computing deals with the *field theories* of physics.
- Field theories describe how a *field* evolves in time or depends on other variables.
- Field theories often constructed by writing a *Lagrangian* or a *Hamiltonian* of the field and treating it as a classical or quantum system with a finite or infinite number of degrees of freedom.
- The resulting field theories are referred to as classical or quantum field theories and have a particular mathematical structure expressed in terms of *field equations*:

Field	Potential	Conservation	Material law
Gravitation	$g = -\nabla\phi$	$\nabla \cdot f + 4\pi\rho = 0$	$f = g/G$ (Newton)
Electrostatics	$E = -\nabla V$	$\nabla \cdot D = 4\pi\rho$	$D = \epsilon\, E$
Electromagnetics	$B = \nabla \times A$	$\nabla \times H = J$	$H = B/\mu$
Diffusion	$g = -\nabla c$	$\nabla \cdot J + s = 0$	$J = D\, g$ (Fick)
Heat transfer	$g = -\nabla T$	$\nabla \cdot J + s = 0$	$J = \kappa\, g$ (Fourier)
Elasticity	$\epsilon = \mathrm{sym}\nabla u$	$\nabla \cdot \sigma + f = 0$	$\sigma = \mathbb{C}\,\epsilon$ (Hooke)
General	$\epsilon = \delta u$	$\partial\sigma + f = 0$	??

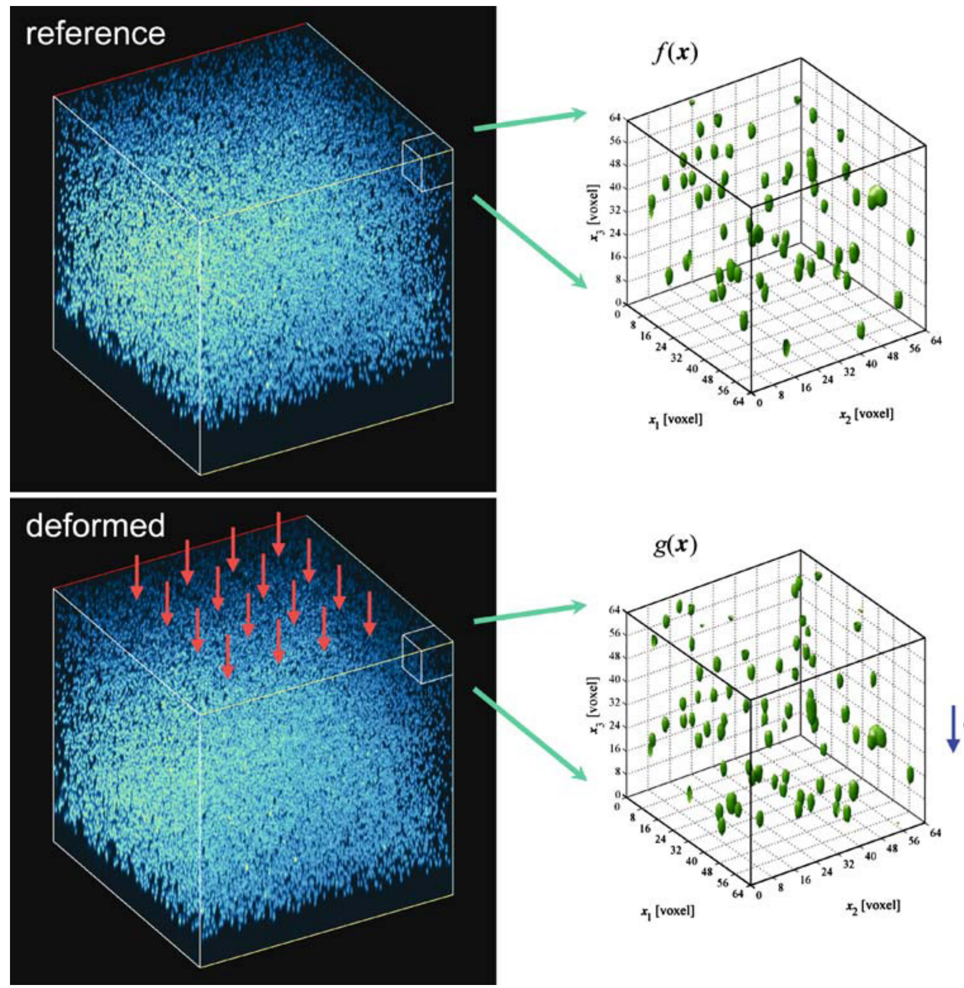
- Field equations are exactly known, *only material law is determined from data*!

Why Data-Driven Computing Now?

- How does Data-Science intersect with *computational mechanics?*
- *Material data is currently plentiful* due to dramatic advances in experimental science (DIC, EBSD, microscopy, tomography…) and multiscale computing (DFT → MD → DDD → SM → Hom)

Digital Volume Correlation (DVC): Two confocal volume images of an agarose gel with randomly dispersed fluorescent particles before and after mechanical loading. The full displacement vector field is measured using 3D volume correlation methods.

C. Franck, S. Hong, S.A. Maskarinec, D.A. Tirrell and G. Ravichandran, *Experimental Mechanics* (2007) **47**:427–438.



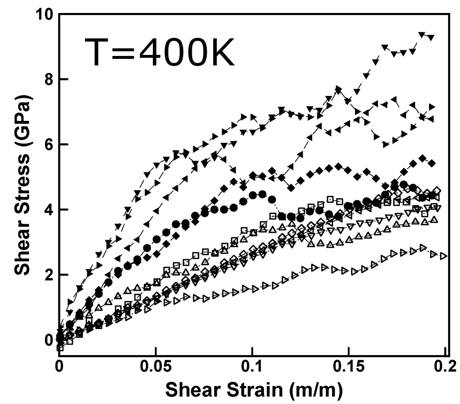
Why Data-Driven Computing Now?

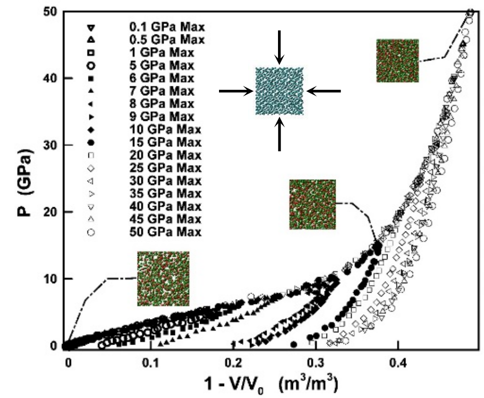
- Material data can also be generated in large volumes from high-fidelity *micromechanical calculations* (DFT, MD, DD…)
- New role for *multiscale analysis*: Data generation

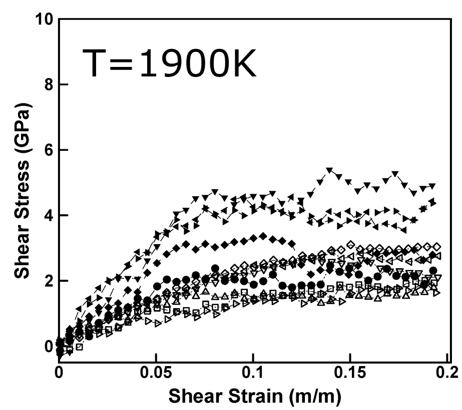
Amorphous SiO_2 glass: LAMMPS MD calculations of amorphous silica glass under *pressure-shear* loading over a range of *temperatures* and *strain rates*. RVEs are quenched from the melt, then analyzed using the BKS potential with Ewald summation.

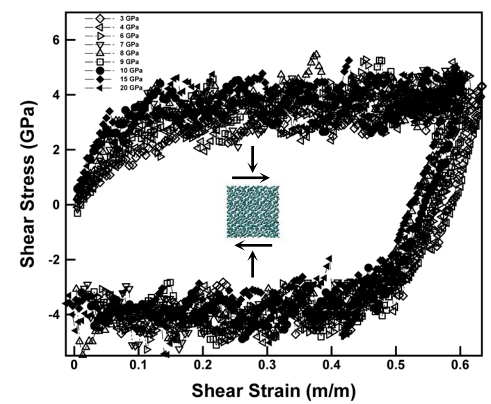
Schill, W., Heyden, S., Conti, S. & MO, *JMPS*, **113** (2018) 105-125.

Schill, W., Mendez, J.P., Stainier, L. & MO, *JMPS*, **140** (2020) 103940.









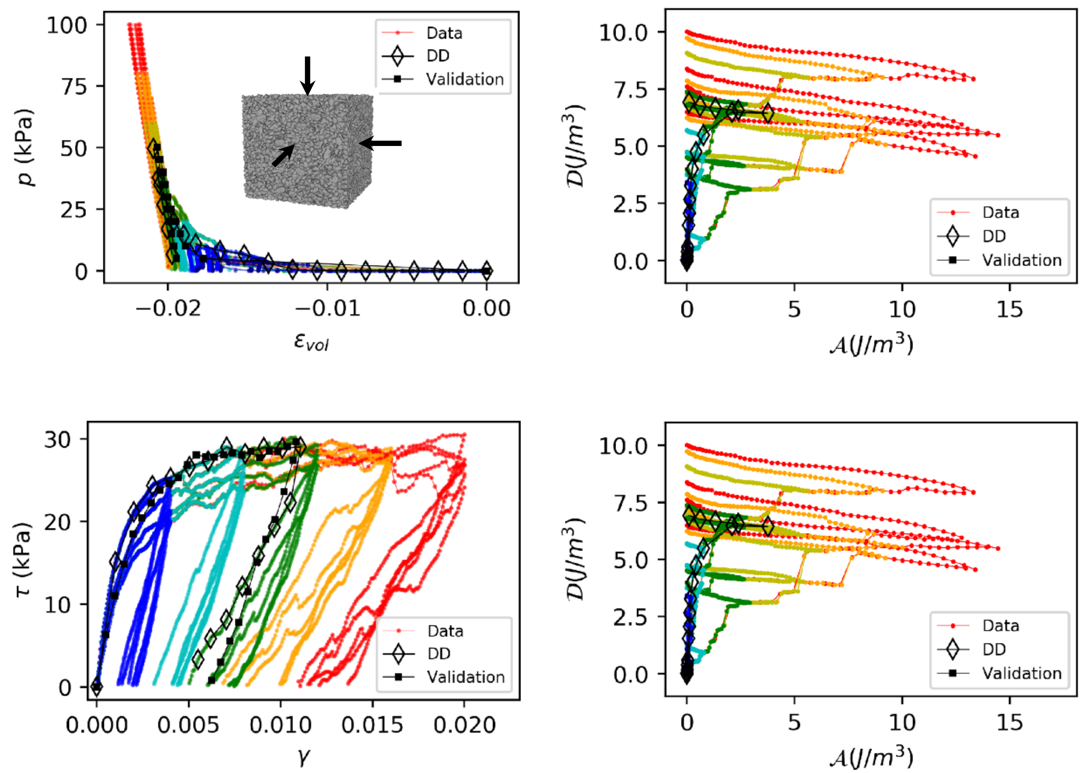
Why Data-Driven Computing Now?

- Material data can also be generated in large volumes from high-fidelity *micromechanical calculations* (DFT, MD, DD…)
- New role for *multiscale analysis*: Data generation

Granular matls. (dry sand): Level-Set Discrete Element Method (LS-DEM) simulation of granular material samples. 3D irregular *rigid particles* interact through *frictional contact*. Particle morphology described by level-set functions. Note calculation of *dissipation and free energy*.

Karapiperis, K., Harmon, J., And, E., Viggiani, G. & Andrade, J.E., *JMPS*, **144** (2020) 104103.

Karapiperis, K., Stainier, L., Ortiz, M. & Andrade, J.E., *JMPS*, **147** (2021) 104239.

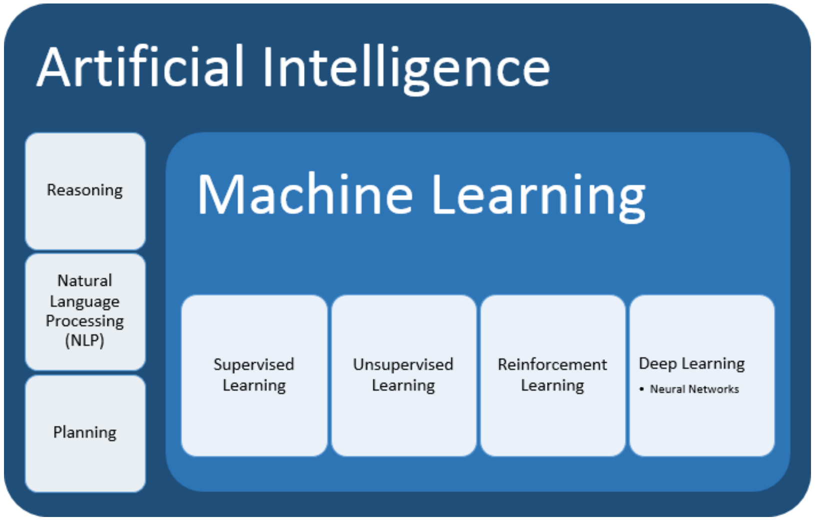


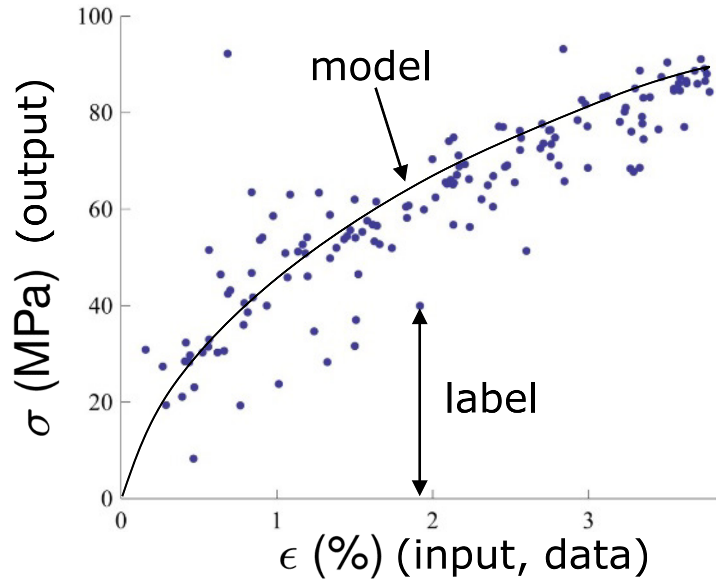
Model-Free Data-Driven Computing

- Mechanics of materials is presently *data rich*: Challenge and opportunity!
- *Data-Driven Computing*: Forge a closer connection between data and predictions.
- *Model-Based* (supervised) *vs.* *Model-Free* (unsupervised) Data-Driven computing:

Model-Based Data-Driven computing:	Data	$\rightarrow$	Model	$\rightarrow$	Prediction
Model-Free Data-Driven computing:	Data		$\longrightarrow$		Prediction

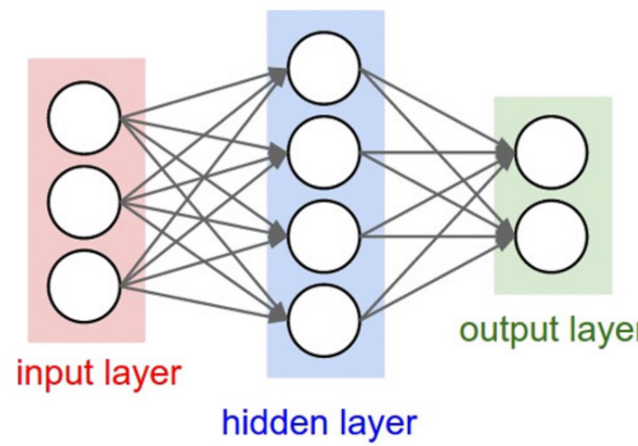
Model-Free Data-Driven Computing





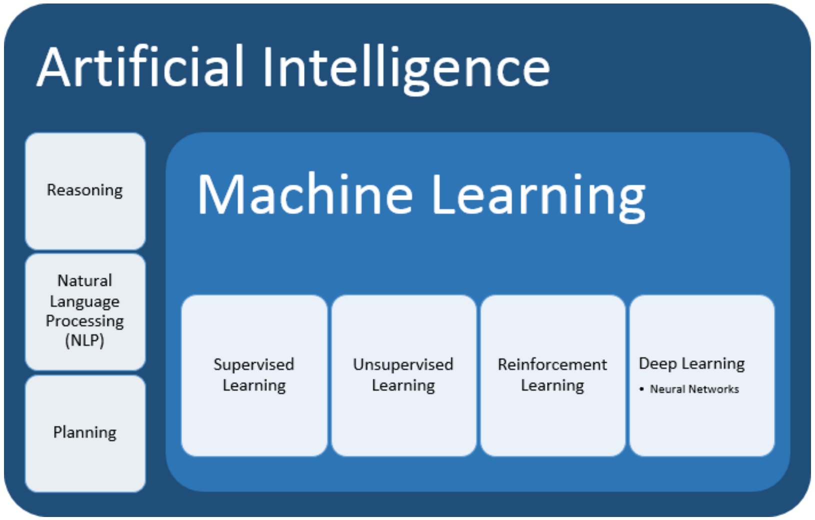
Supervised learning: Find (e.g., by *regression*) a function (e.g., *deep Neural Network*) from data containing both inputs and outputs (labels).

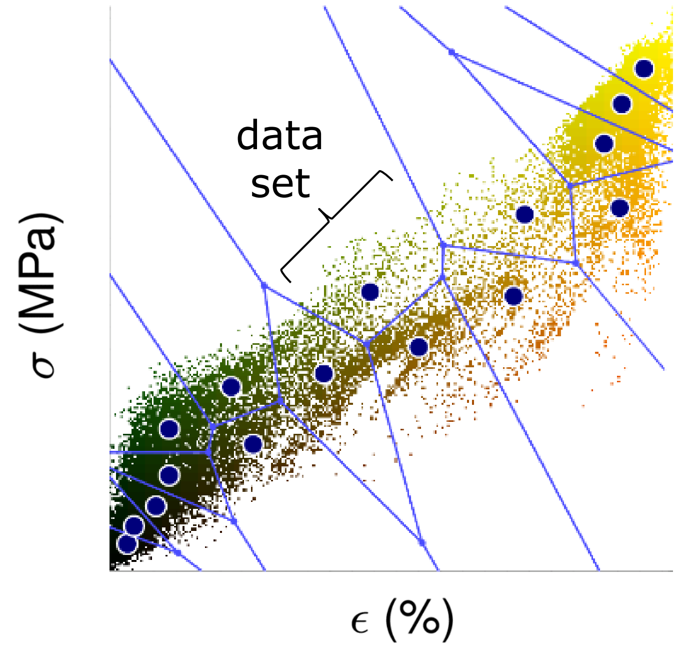
J. Hurwitz & D. Kirsch, *Machine Learning*, John Wiley & Sons, 2018.



Deep Neural Network representation, regression

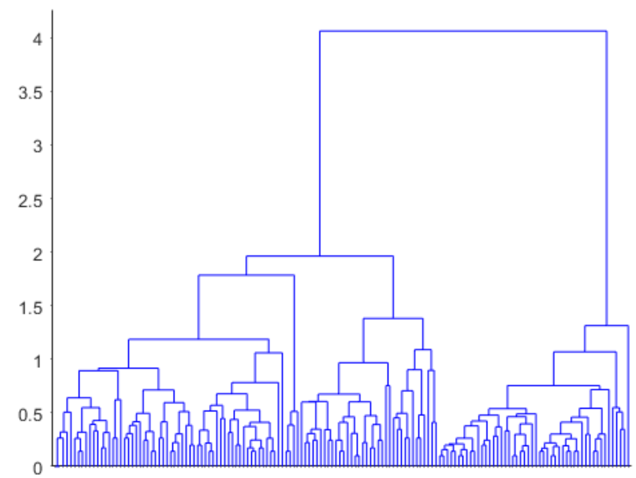
Model-Free Data-Driven Computing





Unsupervised learning: Find structure in ***unlabeled data*** sets (e.g., grouping, clustering, density), make predictions directly from ***data structures***.

J. Hurwitz & D. Kirsch, *Machine Learning*, John Wiley & Sons, 2018.



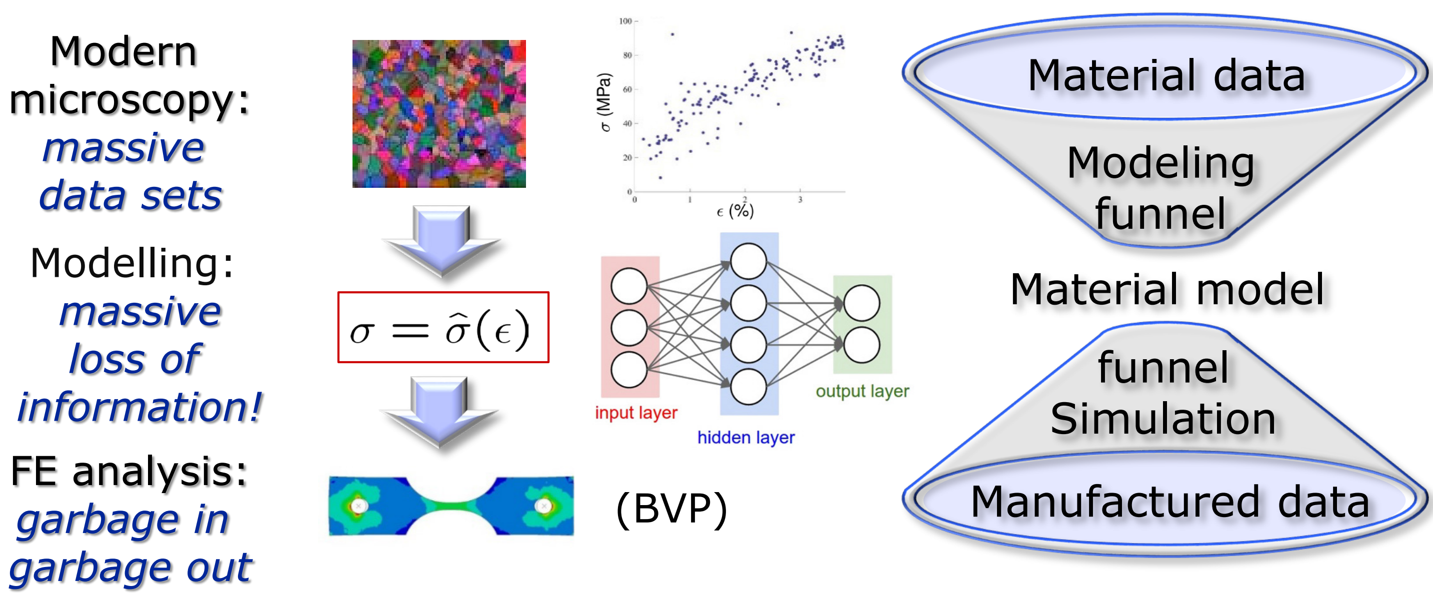
Hierarchical k-means representation, set based

Model-Free Data-Driven Computing

- Mechanics of materials is presently *data rich*: Challenge and opportunity!
- *Data-Driven Computing*: Forge a closer connection between data and predictions.
- *Model-Based* (supervised) *vs.* *Model-Free* (unsupervised) Data-Driven computing:

Model-Based Data-Driven computing:	Data	$\rightarrow$	Model	$\rightarrow$	Prediction
Model-Free Data-Driven computing:	Data		$\longrightarrow$		Prediction

- Critique of Model-Based Data-Driven computing:
  i) Modeling is *lossy* (information in model $<$ information in data).

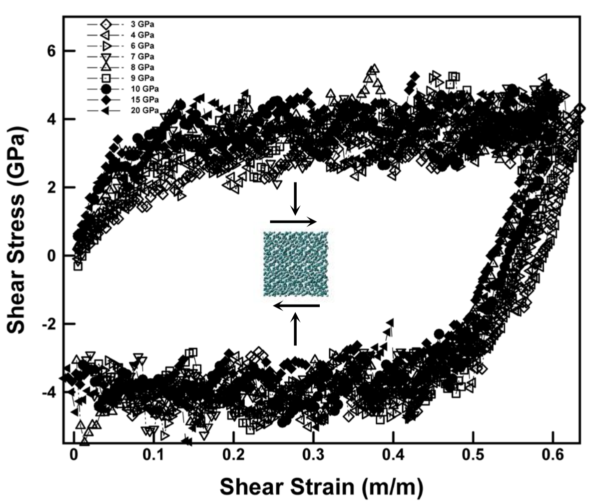


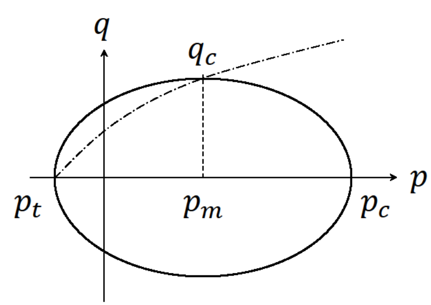
Model-Free Data-Driven Computing

- Mechanics of materials is presently *data rich*: Challenge and opportunity!
- *Data-Driven Computing*: Forge a closer connection between data and predictions.
- *Model-Based* (supervised) *vs.* *Model-Free* (unsupervised) Data-Driven computing:

Model-Based Data-Driven computing: Data $\rightarrow$ Model $\rightarrow$ Prediction
Model-Free Data-Driven computing: Data $\longrightarrow$ Prediction

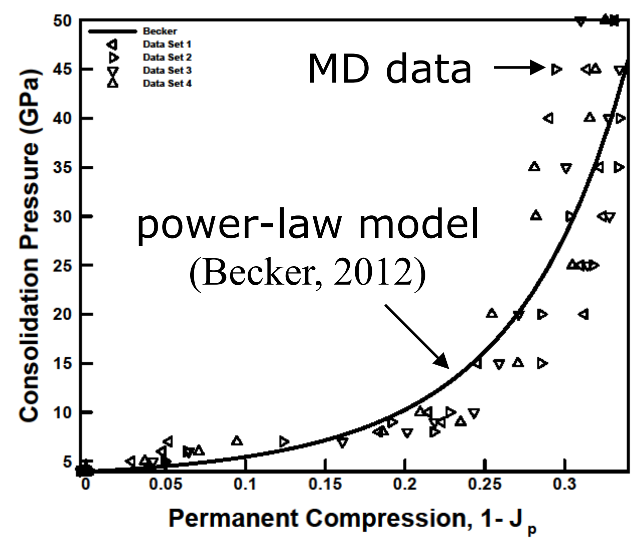
- Critique of Model-Based Data-Driven computing:
  - i) Modeling is *lossy* (information in model $<$ information in data).
  - ii) Modeling is *ad hoc* (involves arbitrary decisions, sausage making).
  - iii) Modeling introduces *biases, modeling error, epistemic uncertainty*.





Silica glass (SiO_2)
Critical state model
(Cam-Clay)

W. Schill *et al.*, *JMPS*, **113** (2018) 105-125.

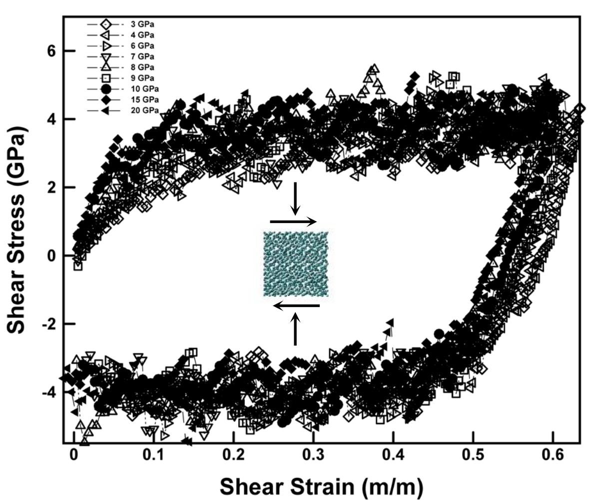


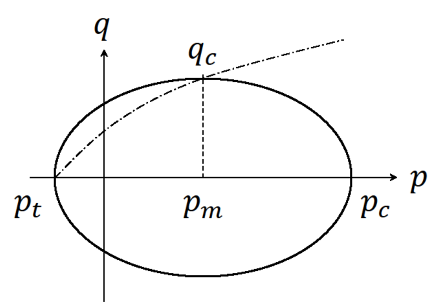
Model-Free Data-Driven Computing

- Mechanics of materials is presently *data rich*: Challenge and opportunity!
- *Data-Driven Computing*: Forge a closer connection between data and predictions.
- *Model-Based* (supervised) *vs.* *Model-Free* (unsupervised) Data-Driven computing:

Model-Based Data-Driven computing:	Data	$\rightarrow$	Model	$\rightarrow$	Prediction
Model-Free Data-Driven computing:	Data		$\longrightarrow$		Prediction

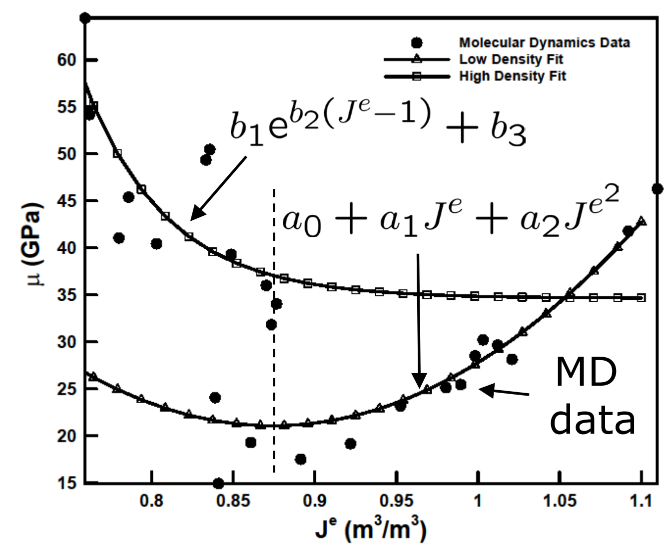
- Critique of Model-Based Data-Driven computing:
  - i) Modeling is *lossy* (information in model $<$ information in data).
  - ii) Modeling is *ad hoc* (involves arbitrary decisions, sausage making).
  - iii) Modeling introduces *biases, modeling error, epistemic uncertainty*.





Silica glass (SiO_2)
Critical state model
(Cam-Clay)

W. Schill *et al.*, *JMPS*, **113** (2018) 105-125.

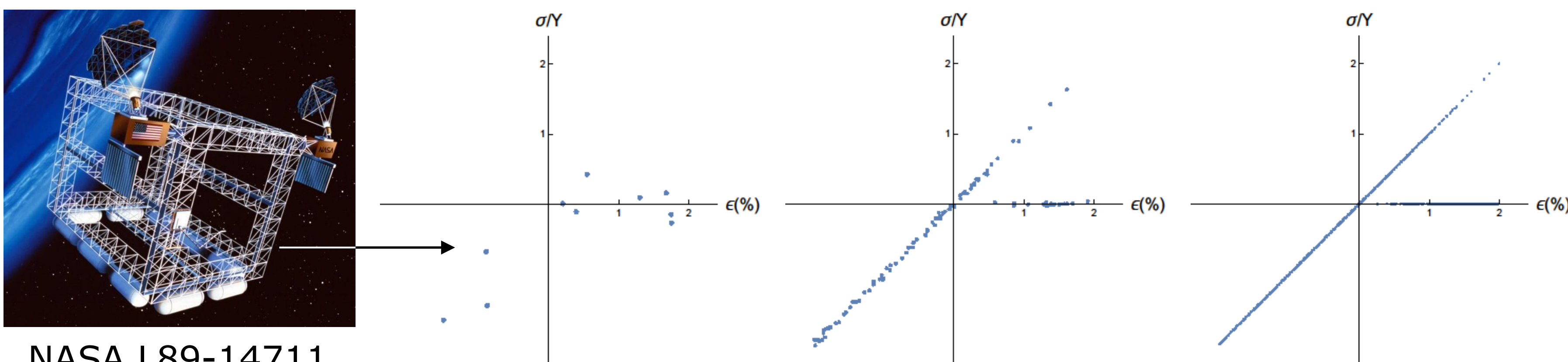


Model-Free Data-Driven Computing

- Mechanics of materials is presently *data rich*: Challenge and opportunity!
- *Data-Driven Computing*: Forge a closer connection between data and predictions.
- *Model-Based* (supervised) *vs.* *Model-Free* (unsupervised) Data-Driven computing:

Model-Based Data-Driven computing:	Data	$\rightarrow$	Model	$\rightarrow$	Prediction
Model-Free Data-Driven computing:	Data		$\longrightarrow$		Prediction

- Critique of Model-Based Data-Driven computing:
  - i) Modeling is *lossy* (information in model $<$ information in data).
  - ii) Modeling is *ad hoc* (involves arbitrary decisions, sausage making).
  - iii) Modeling introduces *biases, modeling error, epistemic uncertainty*.
  - iv) Modeling is *ill-posed* (no notion of continuity, convergence, with respect to data).



NASA L89-14711

Data set sequence: *Convergence of solutions?*

Model-Free Data-Driven Computing

- Mechanics of materials is presently *data rich*: Challenge and opportunity!
- *Data-Driven Computing*: Forge a closer connection between data and predictions.
- *Model-Based* (supervised) *vs.* *Model-Free* (unsupervised) Data-Driven computing:

Model-Based Data-Driven computing:	Data	$\rightarrow$	Model	$\rightarrow$	Prediction
Model-Free Data-Driven computing:	Data		$\longrightarrow$		Prediction

- Critique of Model-Based Data-Driven computing:
  i) Modeling is *lossy* (information in model $<$ information in data).
  ii) Modeling is *ad hoc* (involves arbitrary decisions, sausage making).
  iii) Modeling introduces *biases, modeling error, epistemic uncertainty*.
  iv) Modeling is *ill-posed* (no notion of continuity, convergence, with respect to data).

- Cut-out the middleman! Model-Free Data-Driven computing: Make predictions directly from the data.

The data, all the data, nothing but the data

How?

Why Data-Driven Computing Now?

- It is possible to formulate boundary-value problems, convergent approximation schemes, *based directly on material data*, without the intermediate step of constitutive modeling
- The *Model-Free Data-Driven paradigm* forges a direct connection between material data and solutions of boundary-value problems, in the spirit of unsupervised learning
- Results extend to *infinite-dimensional problems* (linear, finite elasticity)
- *Solvers*? Data structures, searching? Connections to Machine Learning?
- Extension to *time-dependent problems*? (e.g., dynamics)
- Extension to *inelasticity*? (viscoelasticity, viscoplasticity, plasticity)
- *Probability? Inference?* (scatter, random materials, random loads…)
- *Where does the data come from*?
  - *Multiscale* Data-Driven schemes
  - Data-Driven *Material Identification*

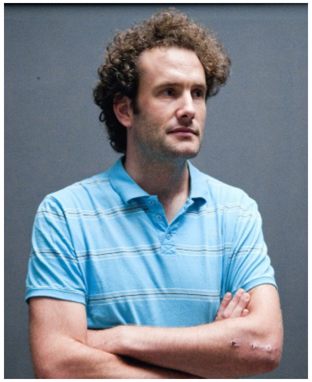
Why Data-Driven Computing Now?

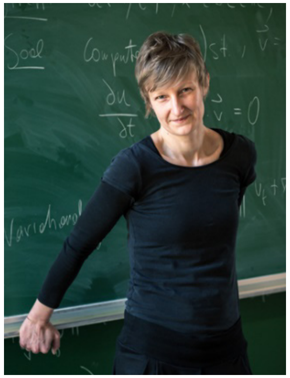
TIME TABLE

TIME	Monday	Tuesday	Wednesday	Thursday	Friday
	October 10	October 11	October 12	October 13	October 14
09.00 - 09.45	Registration	Schönlieb	Stainier	Doblaré	Doblaré
09.45 - 10.30	Ortiz	Schönlieb	Stainier	Doblaré	Doblaré
11.00 - 11.45	Ortiz	Ortiz	Schönlieb	Stainier	Doblaré
11.45 - 12.30	Ortiz	Ortiz	Schönlieb	Stainier	Ortiz
14.00 - 14.45	Réthoré	Schönlieb	Reese	Stainier	
14.45 - 15.30	Réthoré	Schönlieb	Reese	Reese	
16.00 - 16.45	Réthoré	Réthoré	Stainier (DDCM)	Reese	
16.45 - 17.30	Réthoré (DIC)	Réthoré (DDI)	Poster Session	Reese	
18.00	Welcome aperitif				

Why Data-Driven Computing Now? – Lecture plan








- MD: Physically-informed neural networks in predictive physics
- MO: Fundamentals of (model-free) Data-Driven mechanics
- JR: Data-Driven material identification, experimental methods
- SR: Data structures, solvers, algorithmic strategies, plasticity
- CBS: Imaging, mathematical approaches, hybrid modeling
- LS: Extensions to finite elasticity, multiscale analysis, design

Why Data-Driven Computing Now?

TIME TABLE

TIME	Monday	Tuesday	Wednesday	Thursday	Friday
	October 10	October 11	October 12	October 13	October 14
09.00 - 09.45	Registration	Schönlieb	Stainier	Doblaré	Doblaré
09.45 - 10.30	Ortiz	Schönlieb	Stainier	Doblaré	Doblaré
11.00 - 11.45	Ortiz	Ortiz	Schönlieb	Stainier	Doblaré
11.45 - 12.30	Ortiz	Ortiz	Schönlieb	Stainier	Ortiz
14.00 - 14.45	Réthoré	Schönlieb	Reese	Stainier	
14.45 - 15.30	Réthoré	Schönlieb	Reese	Reese	
16.00 - 16.45	Réthoré	Réthoré	Stainier (DDCM)	Reese	
16.45 - 17.30	Réthoré (DIC)	Réthoré (DDI)	Poster Session	Reese	
18.00	Welcome aperitif				

to be continued…