THE UNIVERSITY of TENNESSEE KNOXVILLE

MATHEMATICAL, NUMERICAL, BAYESIAN, AND CAUSAL PROBLEM SOLVING

MSE 494/MSE510

Instructor: Sergei V. Kalinin

TA: Elizabeth Heon

Times and locations: 9:45-11:00
TR, Ferris Hall 510


2002 - 2022
Since ~2010
2022 -2023
amazon
SCANNING PROBE MICROSCOPY
Electrical and Electromechanical Phenomena at the Nanoscale
Sergei Kalinin
Alexei Gruverman
EDITORS
VOLUME 1
Springer
Sergei V. Kalinin
Alexei Gruverman
Editors
Scanning Probe Microscopy of Functional Materials
Springer
World Scientific Series in Nanoscience and Nanotechnology
Volume 7
Scanning Probe Microscopy for Energy Research
Dawn A Bonnell • Sergei V Kalinin
Editors
World Scientific
Volume 1
World Scientific Series on Emerging Technologies – Volume 1
HANDBOOK ON BIG DATA AND MACHINE LEARNING IN THE PHYSICAL SCIENCES
In 2 Volumes
Volume 1
Big Data Methods in Experimental Materials Discovery
Editors
Surya R. Kalidindi • Sergei V. Kalinin
Turab Lookman
Editors-in-Chief
Sergei V. Kalinin • Ian Foster
World Scientific

There is no computation without representation

What are our objects?

Quantitative Variables:
- 3 apples, 2 oranges

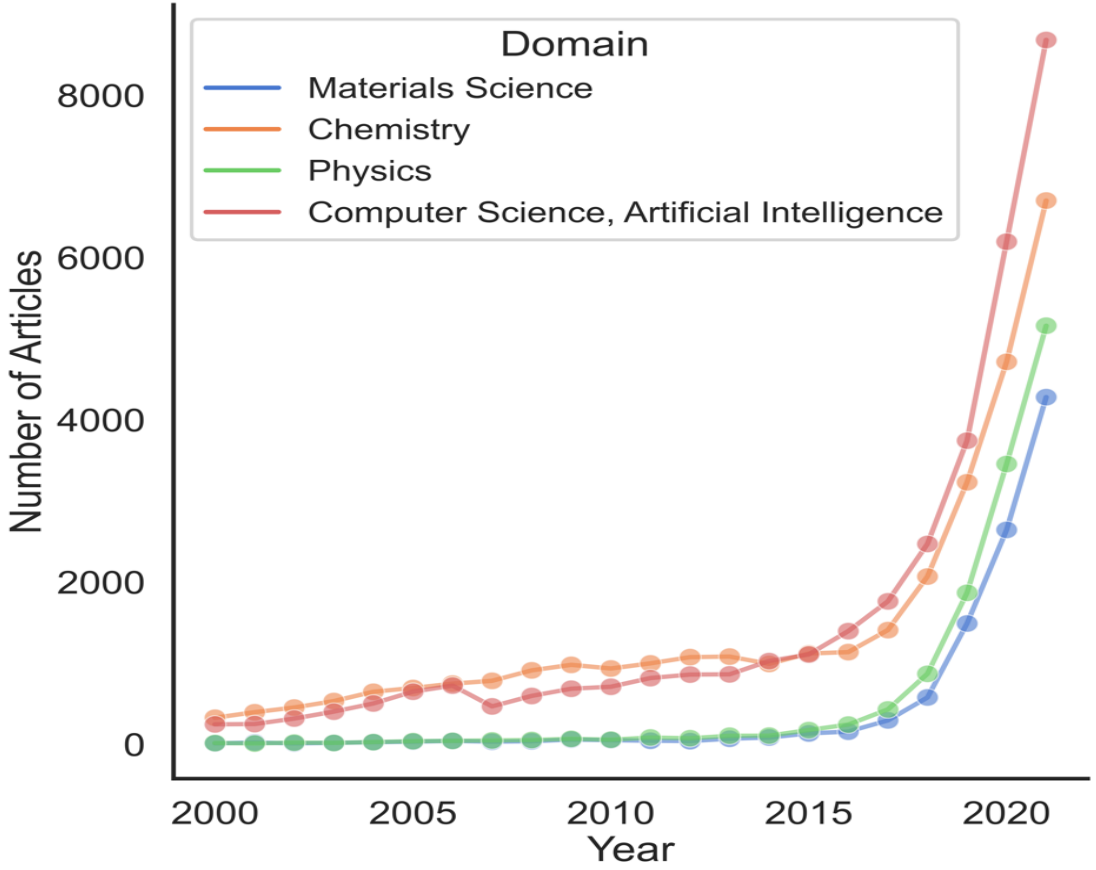
Categorical Variables:
- Apple and Orange

Schemas:
- "Fruit," "Type," and "Quantity“

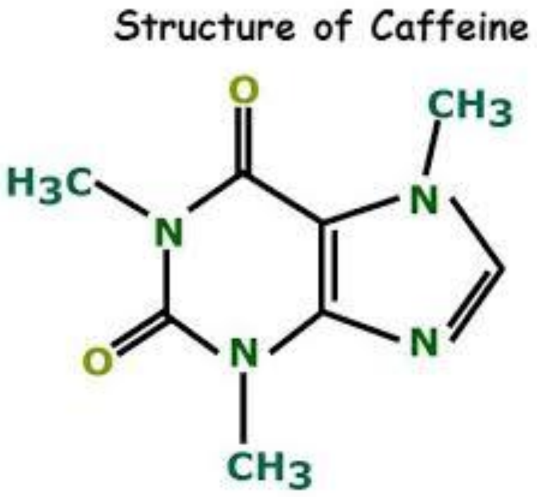
Ontologies:
- "Apple is a type of Fruit“

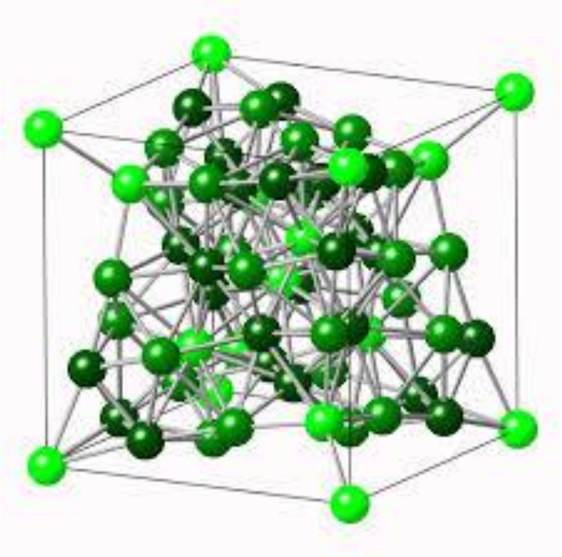
Knowledge Graphs:
- Connection between "Apple," "Fruit," "Vitamin C content," and "Growth Regions."

ML in Domain Sciences



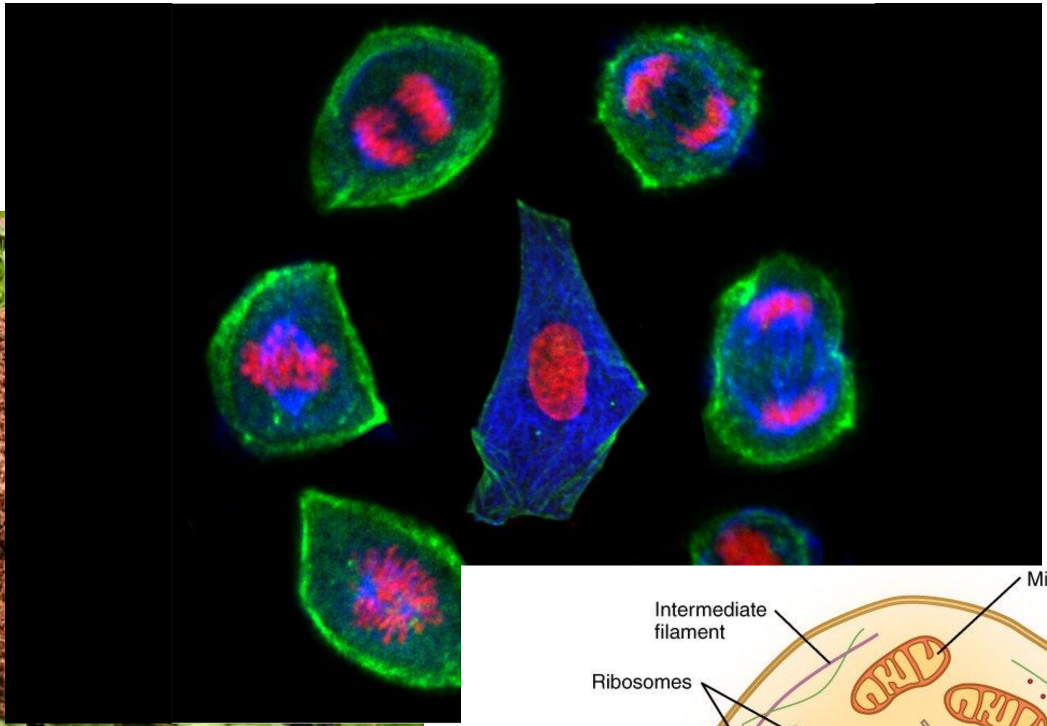
Analysis by B. Blaiszik, Argonne

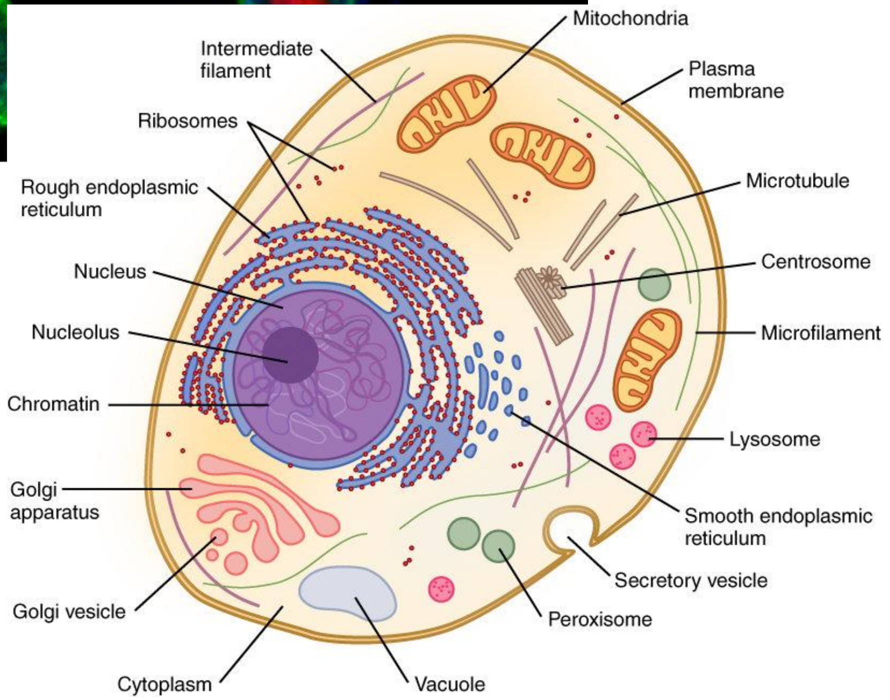




- The rapid adoption of ML in domain sciences and industrial R&D is a very recent trend
- Technologies and workforce emerge from academia into industry
- We can estimate potential growth rates comparing to cloud computing 15 - 20 years ago

No equations (or ML) without representation!





https://phys.org/news/2021-10-cell-life-optical-tweezers.html

https://doi.org/10.1080/00150193.2020.1722012

https://www.pointepestcontrol.net/inside-the-ant-hill/

https://www.researchgate.net/figure/Schematic-representation-of-a-cell-20-with-organelles-delimited-or-made-of-lipid_fig5_350133282

Symbols, numbers, data

- Numbers
- Patterns
- Magnitude
- Shapes and forms

- Symbols
- Algorithms
- Logic

- Correlations

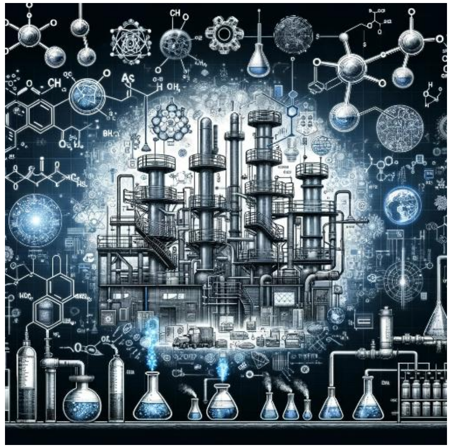
- Prior Knowledge
- Causality
- Workflows

Materials science

Chemistry

Physics

Chemical engineering

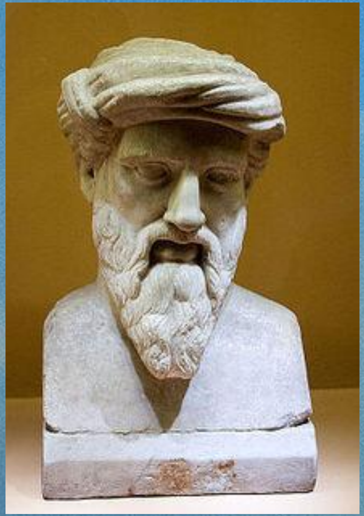


https://www.independent.co.uk/news/science/babylonians-trigonometry-develop-more-advanced-modern-mathematics-3700-years-ago-ancient-civilisation-a7910936.html

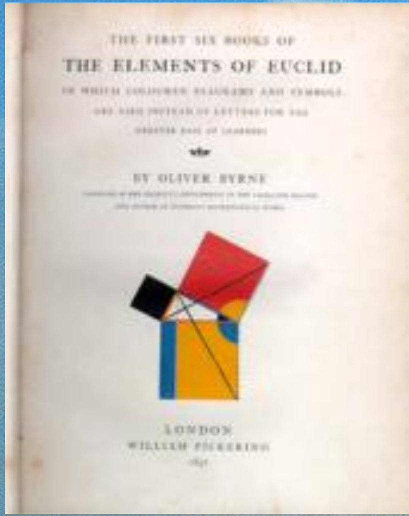
- Taxes:
  - Grain allotments
  - Workers
  - Weights of silver
- Geometry
- Pythagoras theorem
- Linear, quadratic, and cubic equations
- Hexadecimal system

- Composite and prime numbers
- Arithmetic, geometric and harmonic means
- Simple number theory
- Power of correlations

https://www.pyramid-of-giza.com/plan-your-visit-pyramids-of-giza/



Pythagoras



Euclides

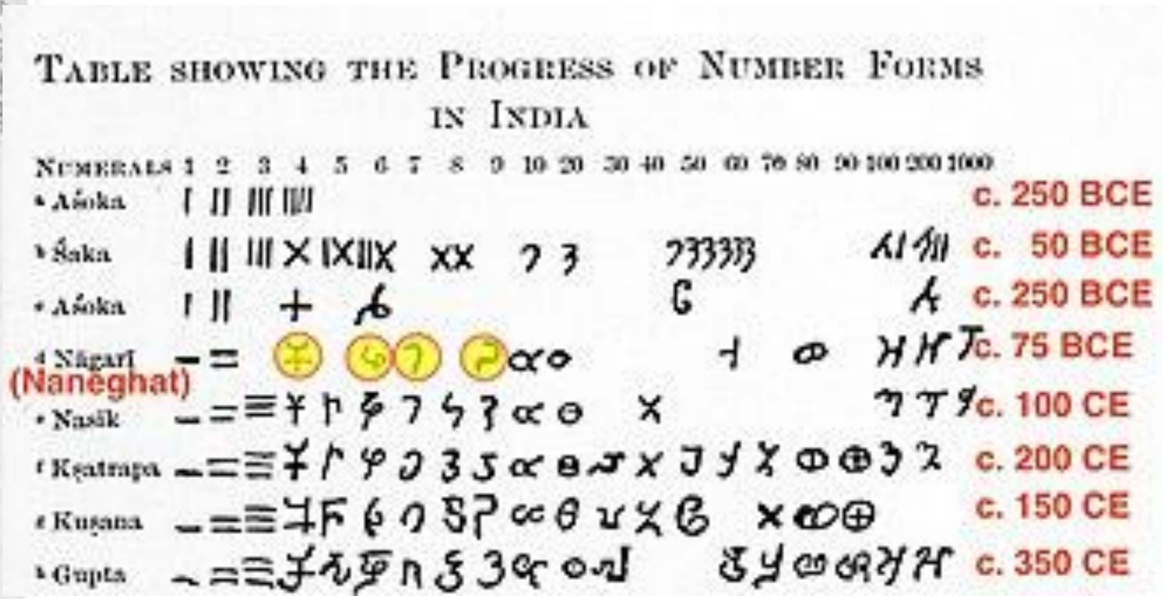
Fictional portrait of Hypatia of Alexandria by Jules Maurice Gaspard

https://assets.editorial.aetnd.com/uploads/2010/03/ancient-greece-gettyimages-475594807.jpg

India and China

The Tsinghua Bamboo Slips, containing the world's earliest decimal multiplicati on table, 305 BC during the Warring States period

The Nine Chapters on the Mathematical Art, one of the earliest surviving mathematical texts from China (2nd century AD)



https://en.wikipedia.org/wiki/History_of_mathematics

- Algebra
- Algorithms
- Early calculus
- Trigonometry
- Cryptography
- Frequency analysis

https://www.newscientist.com/people/muhammad-ibn-musa-al-khwarizmi/

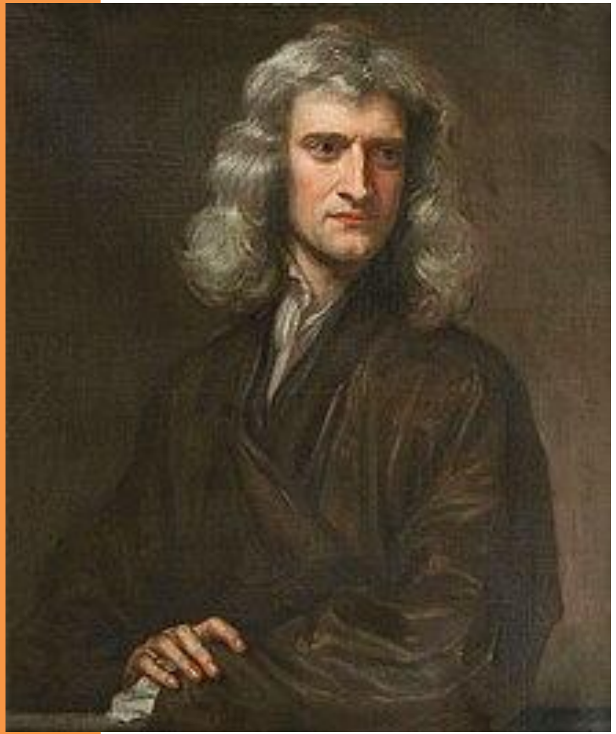
Medieval Europe

Primary drivers:

- Manufacturing
- Taxes
- Philosophy
- Theology

https://www.discovermiddleages.co.uk/medieval-life/medieval-life-and-times

Renaissance



Isaac Newton
(1642 –1726)

Leonhard Euler
(1707 – 1783)

Friedrich Gauss
(1777 –1855)

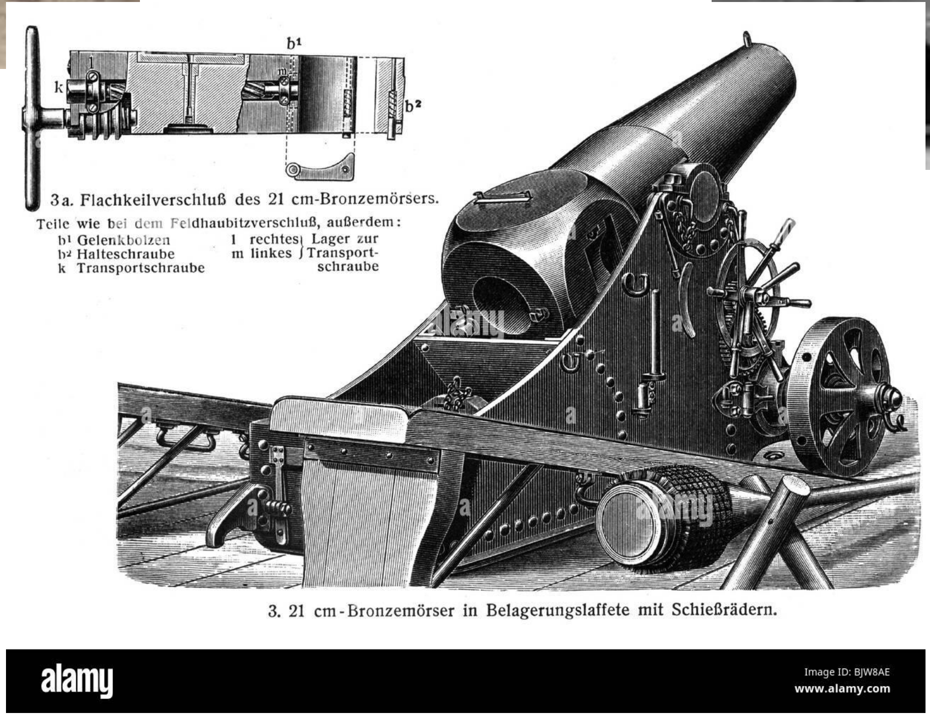
Joseph Fourier
(1768 –1830)

https://en.wikipedia.org/wiki/Isaac_Newton
https://en.wikipedia.org/wiki/Leonhard_Euler
https://en.wikipedia.org/wiki/Carl_Friedrich_Gauss
https://en.wikipedia.org/wiki/Joseph_Fourier

- Mechanics
- Electromagnetism
- Heat conduction

From Renaissance to XX Century

https://en.wikipedia.org/wiki/Reffye_85_mm_cannon



https://en.wikipedia.org/wiki/Pre-dreadnought_battleship

https://www.alamy.com/stock-photo-military-artillery-guns-german-21-cm-mortar-wood-engraving-late-19th-28829542.html

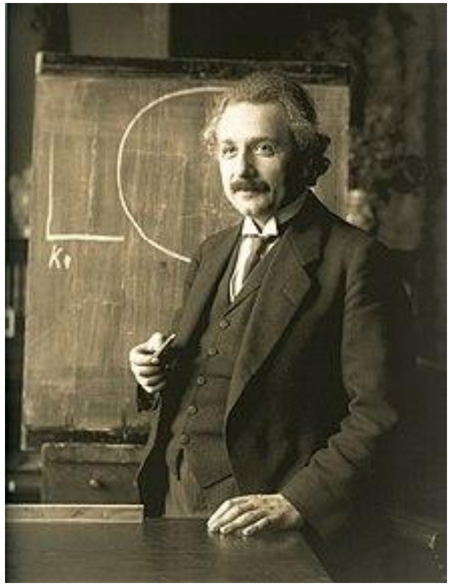
Gödel proved that the world of pure mathematics is inexhaustible; no finite set of axioms and rules of inference can ever encompass the whole of mathematics; given any finite set of axioms, we can find meaningful mathematical questions which the axioms leave unanswered. I hope that an analogous Situation exists in the physical world. If my view of the future is correct, it means that the world of physics and astronomy is also inexhaustible; no matter how far we go into the future, there will always be new things happening, new information coming in, new worlds to explore, a constantly expanding domain of life, consciousness, and memory.

— Freeman Dyson

From Lecture 1, 'Philosophy', in a series of four James Arthur Lectures, 'Lectures on Time and its Mysteries' at New York University (Autumn 1978).

Printed in 'Time Without End: Physics and Biology in an Open Universe', *Reviews of Modern Physics* (Jul 1979), **51**, 449.

The Quantum Era



Einstein

Bohr

Noether

Schwinger

Feynman

- Quantum mechanics
- Relativity
- Quantum electrodynamics
- Discovery of spin
- … and so on

For a physicist mathematics is not just a tool by means of which phenomena can be calculated, it is the main source of concepts and principles by means of which new theories can be created.

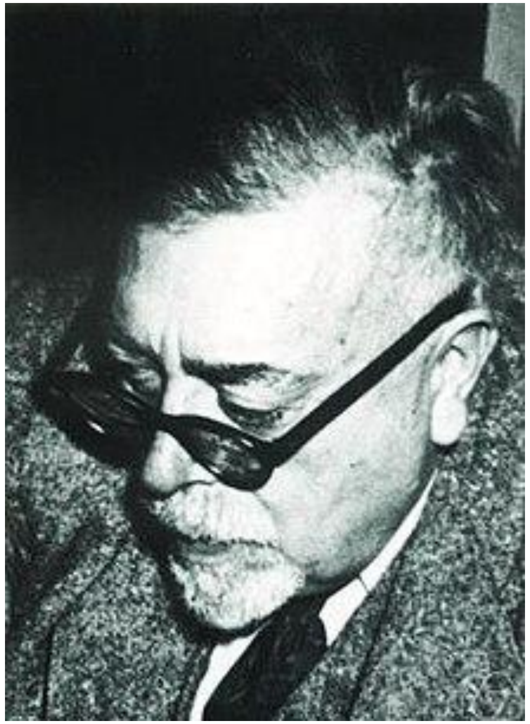
— Freeman Dyson

In 'Mathematics in the Physical Sciences', *Scientific American* (Sep 1964), **211**, No. 3, 129.

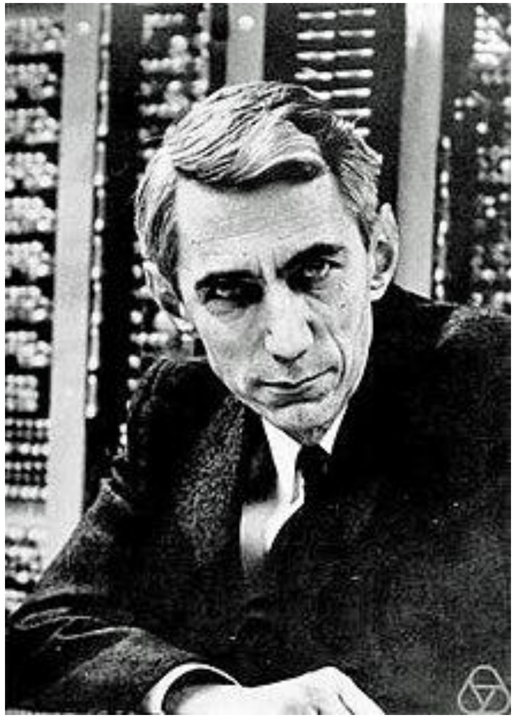
Science and Engineering of Big Numbers

John von Neumann
(1903 – 1957)

Eugene Wigner
(1902 – 1995)



Norbert Wiener
(1894 – 1964)

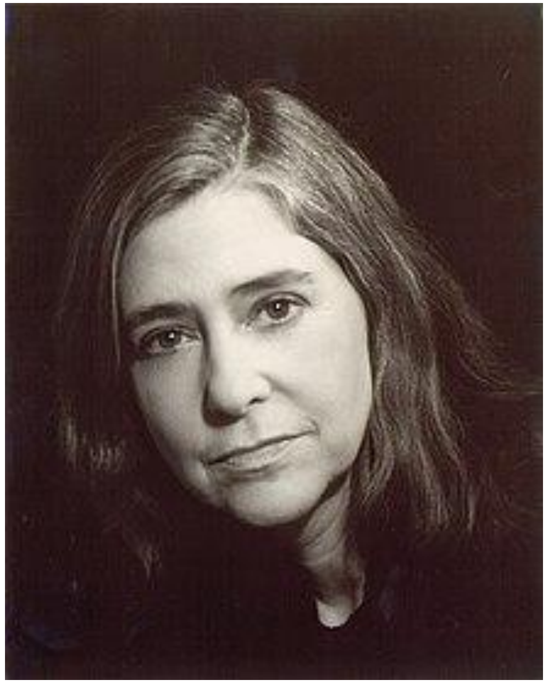


Claude Shannon
(1916 – 2001)

https://en.wikipedia.org/wiki/John_von_Neumann
https://en.wikipedia.org/wiki/Eugene_Wigner
https://en.wikipedia.org/wiki/Norbert_Wiener
https://en.wikipedia.org/wiki/Claude_Shannon

Science and Engineering of Big Numbers

Grace Hopper
(1906 –1992)



Margaret Hamilton
(1936 -)

Katherine Johnson
(1918 –2020)

https://en.wikipedia.org/wiki/Grace_Hopper

https://en.wikipedia.org/wiki/Margaret_Hamilton_(software_engineer)

https://en.wikipedia.org/wiki/Katherine_Johnson

Much like in early 80ies (PCs), mid-90ies (Internet), mid-200 (cloud), and 2012 (Deep learning), we are now observing the emergence of the disruptive technology

aws

ChatGPT

1981

1991

2001

2011

2021

Automated laboratories!

We have:

- Automated laboratories
- Cloud infrastructure, edge, HPC

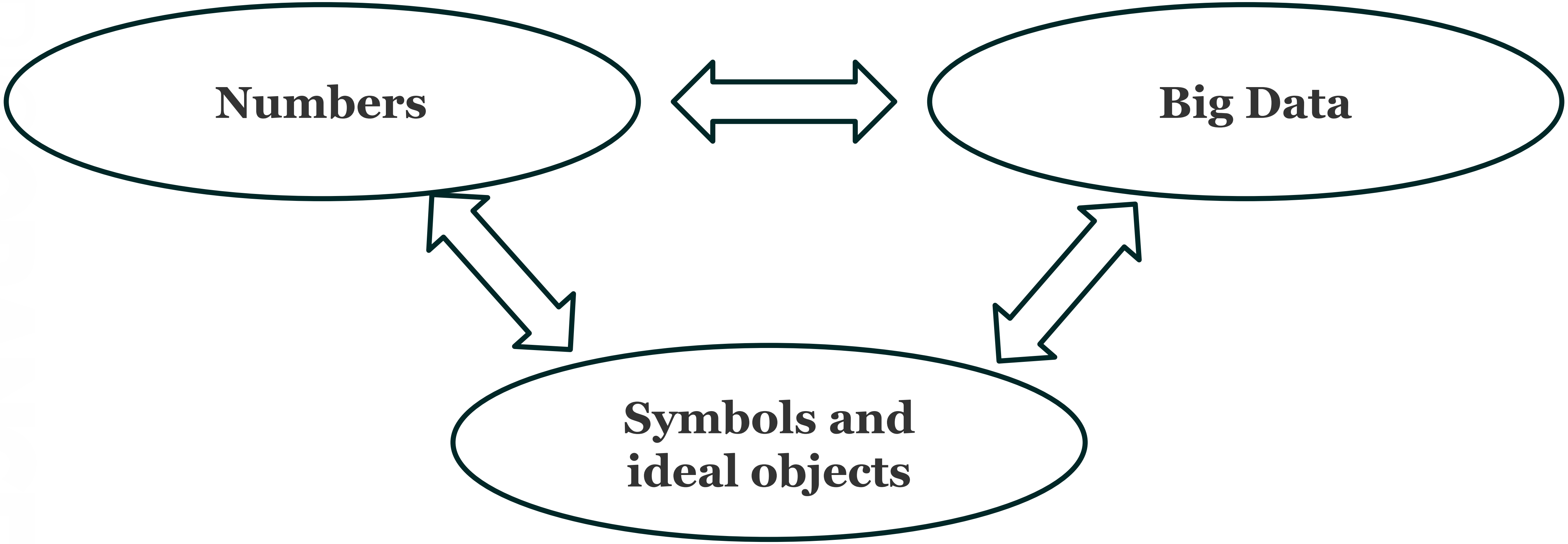
Challenge:

ML to connect it to meaningful workflows

https://www.computerhistory.org/timeline/1981/

New addition: big data

We are now entering the new era of science when in addition to classical mathematics and physics tools we have new intuition and methods of big data



Course Information

Faculty Contact Information:

Instructor:	Prof. Sergei V. Kalinin,
Office:	314 IAMM
E-mail:	sergei2@utk.edu
Teaching Assistant:	Kamyar Barakati, kbarakat@vols.utk.edu

Instructor Availability:
Please don't hesitate to email me with updates, questions, or concerns. I will typically respond within 24 hours during the week and 48 hours on the weekend. I will notify you if I will be out of town and if connection issues may delay a response.

Meeting Time: 9:45-11:00 TR, Ferris Hall 502

The lectures and materials will be posted on Canvas and at GitHub:
https://github.com/SergeiVKalinin/MSE_Spring2025

Office Hours:
Friday 1:30 - 3:00 PM are open for 1:1 meetings to discuss any course related item. Can also be made by scheduling via email.

Course Outline

1. Mathematical methods in physical science
2. Introduction into Python and SciPy ecosystems
3. Classical numerical methods with Python
4. Regression methods (linear, functional fits, symbolic)
5. Bayesian data analysis
6. Causal methods
7. Gaussian Processes, Bayesian Optimization, and active learning
8. Gaussian processes meet physics

Value Proposition

1. Achieve proficiency in scientific Python, tapping into its diverse mathematical applications.
2. Gain the ability to solve both foundational and advanced equations with confidence.
3. Deep dive into data analysis, mastering tools that empower data-driven decisions.
4. Grasp the nuances of Bayesian methods, learning how to weave together data with prior knowledge.
5. Explore the intricate landscape of causal analysis within the ML spectrum.
6. Develop a solid understanding of probabilistic techniques for decision-making in uncertain scenarios.

Prerequisites

To be successful in this course you will need a general background in materials science. Python or similar programming experience, while not essential, will be extremely useful. Students without any prior programming experience should expect to spend extra time outside of class learning basic skills.

Outcomes

1. This course aims to provide students with the skills needed to link physics, numerical methods, and big data

2. Students should learn how to combine intuition from mathematics, physics, and machine learning methods

3. The course is designed to provide students with basic knowledge of numerical methods, causal analysis, and Bayesian methods

4. Preparedness for the Future of Science and Industry: With insights into automated labs, large language models in scientific workflows, and federated tools and workflows, students will be prepared for the future of industry. These skills are increasingly important as companies automate processes and incorporate AI into their workflows. This knowledge can help students stand out in the job market and be prepared for the careers of the future.

This and that

Learning Environment:
The class will be delivered as in-person lectures. The Jupyter notebooks, code libraries, and videos provided. Weekly programming exercises will be assigned via Google Colabs and those students wishing to interact with the instructor in person should attend office hours.

Use of ChatGPT:
Strongly encouraged both for programming and written assignments. However, the students have to be aware of the limitations of the generative models.

Grading & Policies:

• Midterms (2)	30%
• Homeworks	40%
• Final Project	30%

Reference Materials

I will provide copies of lecture notes, presentations, and Colabs on GitHub and Canvas. There is no specific textbook for the course, and we will take material from a variety of sources including:

- Andrew Bird et al, Python Workshop, https://www.packtpub.com/product/the-python-workshop/9781839218859

- Oswaldo Martin, Bayesian Analysis with Python - Second Edition, https://subscription.packtpub.com/book/data/9781789341652/

- Alexander Molak, Causal Inference and Discovery in Python, https://subscription.packtpub.com/book/data/9781804612989/

Homework 1:

- Create new Colab, https://colab.google/
- Chapter 1-4 and 10, Python Workshop.

Homework, midterm, and finals format

- All homeworks, midterms, and finals will be in the Google Colab format
- Use the code for programming exercises and markdown fields for text responses
- Share in the “comment” or “editor” modes
- The Colabs should save all graph outputs
- The Colabs should be able to run from the beginning to end (e.g. if I restart the runtime and run all)
- Submit to sergei2vk@gmail.com

Homework 1:

- Create new Colab, https://colab.google/
- Chapter 1-4 and 10, Python Workshop.