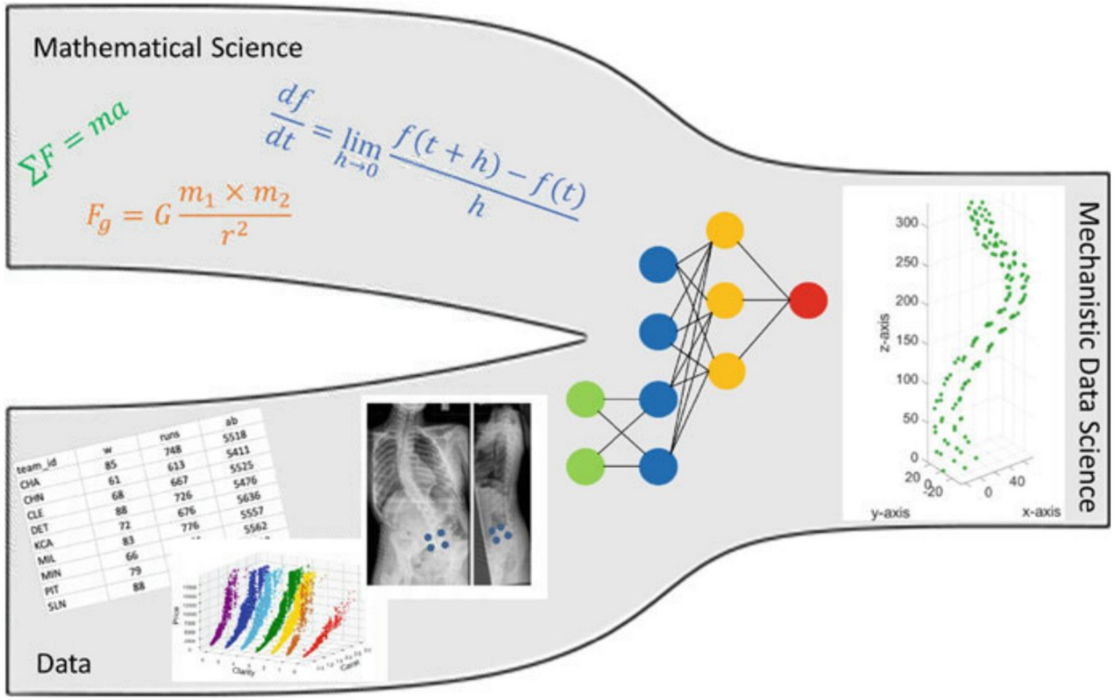
Contents

- Mechanistic Data Science for Engineering
  - Definition, goal, advantage
- History of Science
- Science, Technology, Engineering and Mathematics(STEM)
- Data Science Revolution
- Data Science for Materials Design
- Outline of Mechanistic Data Science Methodology
- Examples Describing the Three Types of MDS Problems

Mechanistic Data Science (MDS)

– Structured methodology for coupling data with mathematics and scientific principles to solve intractable problems
– Structured use of data combined with the core understanding of physical phenomena to analyze and solve problems, with the end goal of decision-making
– Hidden link between data and science
– New concept to solve intractable **S**cientific and **E**ngineering (**S&E**) problems



Mechanistic Data Science (MDS)

- Mechanistic
  - Theories which explain a phenomenon in purely physical or deterministic terms

- Goal
  - Intelligent data mining to extract scientific principles and construct useful models
  - Informed decision-making processes that combine data, reduced order models (ROMs), and physical mechanisms
- Advantages
  - can open new avenues to tackle S&E problems that are currently hard to solve today (with existing technology)
  - can solve challenging S&E problems orders of magnitude faster/better with similar accuracy compared to existing technology

Mechanistic Data Science (MDS) [ChatGPT]

- Interdisciplinary field that integrates **mechanistic modeling** (based on physical, chemical, and biological laws) with **data-driven approaches** (such as machine learning, statistical inference, and artificial intelligence). It aims to extract meaningful insights by leveraging both **first-principles models** and **empirical data** to improve predictive accuracy, interpretability, and generalizability across scientific and engineering applications
- Unlike purely data-driven machine learning, which often lacks interpretability, mechanistic data science explicitly incorporates domain knowledge through differential equations, thermodynamic laws, structural mechanics, or other governing principles, allowing for explainable and physically consistent predictions.

Mechanistic Data Science (MDS) [Perplexity]

- Emerging field that combines traditional data science techniques with scientific and mathematical models to extract information from data. Here's an overview of its definition, scope, and contents
- Structured methodology that integrates existing mathematical and scientific models with data science tools to solve complex problems
- It aims to uncover underlying mechanisms and physical phenomena in data by leveraging domain knowledge and scientific principles

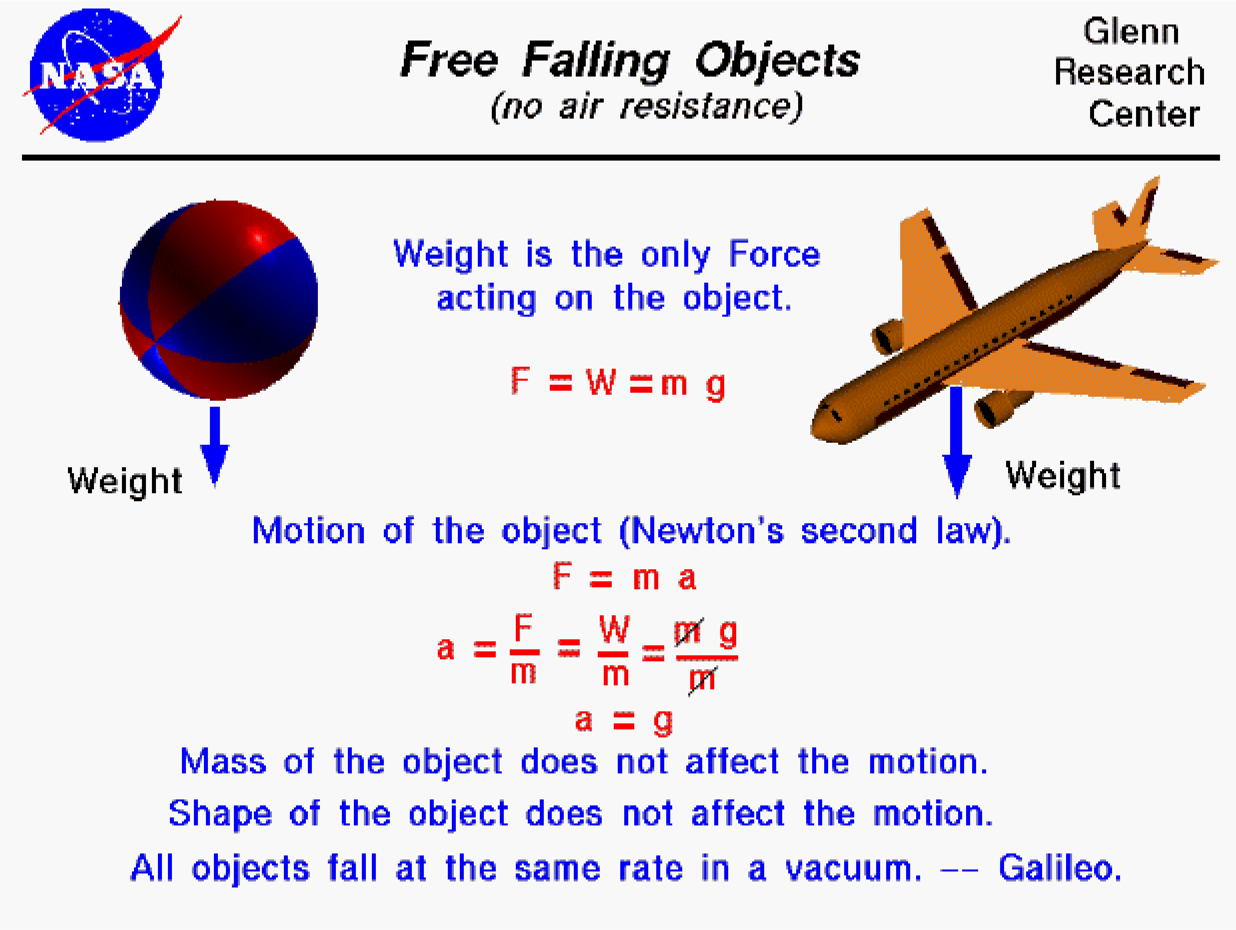
MDS vs. DS

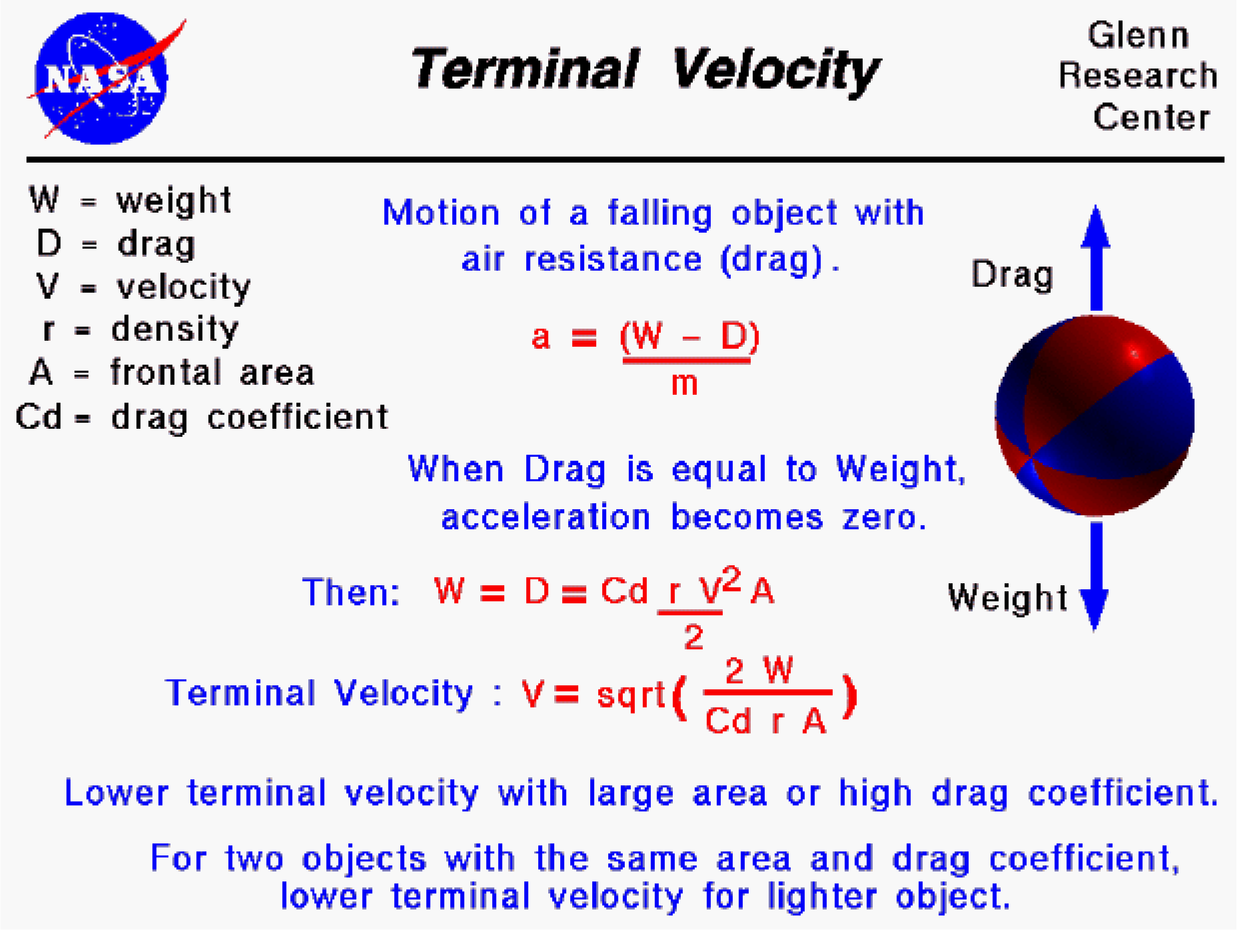
Aspect	Mechanistic Data Science (MDS)	Traditional Data Science
Approach	Integrates mathematical and scientific models with data tools	Primarily focuses on statistical techniques and machine learning
Data Requirements	Works effectively with smaller datasets due to prior knowledge	Requires large datasets for reliable pattern recognition
Focus	Establishes mechanistic relationships and causal mechanisms	Establishes statistical relationships and correlations
Application Areas	Engineering, scientific modeling, epidemic tracking, etc.	Broad applications across various industries like finance, marketing, etc.
Modular Structure	Follows a structured framework with specific modules	Flexible methodology without a fixed structure
Predictive Capability	Provides insights into underlying mechanisms; can predict behaviors not present in the data	Predicts based on patterns within the data; limited to existing data trends
Integration of Domain Knowledge	Strong emphasis on integrating domain-specific knowledge	Less emphasis on domain knowledge; more focus on data-driven insights
Validation Process	Involves hypothesis generation and model validation through experimental observations	Relies on data validation and model performance metrics

1.1 History of Science

- Ancient philosophers: Aristotle (384–322 BC), geocentric
- Scientists in the Renaissance: heliocentric
- Nicolaus Copernicus (1473–1543 AD) → Tycho Brahe (1546–1601)
- Johannes Kepler (1571–1630)
- Galileo Galilei (1564–1642)
  - Galileo's Study of Falling Objects
- Isaac Newton (1642–1726)
  - Newton's Laws of Motion: Law of inertia, force balance, reaction forces
  - Albert Einstein's theory of relativity
- Joseph Fourier (1768–1830): heat conduction and dynamics
- Thomas Edison (1847–1931): light bulb
- John Nash (1928–2015): game theory

1.2 Galileo's Study of Falling Objects





1.4 Science, Technology, Engineering and Mathematics (STEM)

- Science
  - a set of fundamental laws that describe nature and natural phenomena
- Mathematics
  - unifying language of the physical science
- Engineering
  - application of scientific principles for design and problem solving
- Technology
  - implementation of products and capabilities developed through science and engineering

1.4 Science, Technology, Engineering and Mathematics (STEM)

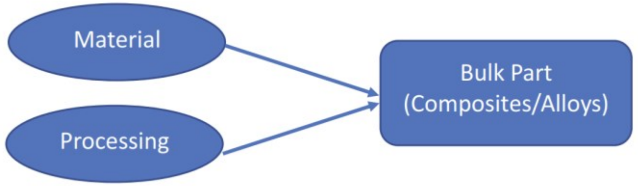
- Continuous cycles in mathematical science and technology
  - Data/observation → empirical scientific discovery → new technology and product
  - Galileo’s study of beams in bending → strength of materials
- Mathematics
  - Continuous: Fourier transform for analyzing the frequency characteristics of a signal
  - Discrete: regression/curve fitting and graphical analysis for analyzing data gathered from experiments
- Scientific method
  - Observe
  - Hypothesize
  - Test
  - Theory or Law

1.5 Data Science Revolution

- Large amounts of data have been collected on a vast array of topics
  - Smart phone: grocery store purchase, travel routes, search history
  - Platform: collect and utilize data posted on their site for various marketing purpose
  - Heavily used for product development, from the concept stage to engineering and manufacturing: increased use of sensors → IoT
- Fatigue fracture analysis and design
  - Fatigue: the initiation and slow propagation of small cracks into larger cracks under repeated cyclic loading
  - Stress life method: S-N curve (stress amplitude vs. number of cycles to failure) → estimate the fatigue life
  - Factors: material strength, microscopic impurities and voids, and surface roughness

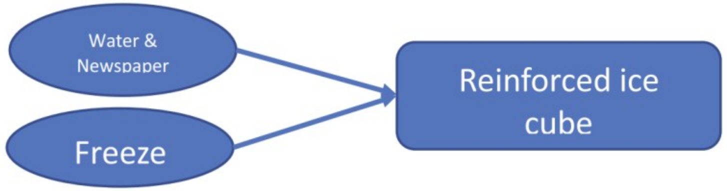
1.7 Data Science for Materials Design

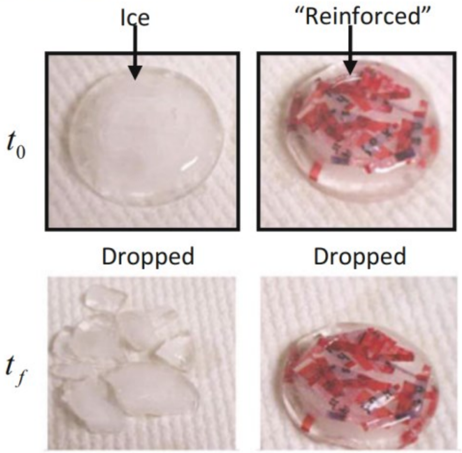
- Overall macrostructural performance of a bulk part is controlled in large part by the microstructure of the material used to make the part
- What’s in the cake mix?
  - flavor, texture, and crumble of the cake are controlled by the ingredients, the mixing, and the baking time and temperature
- Engineered components
  - strength, stiffness, and fracture resistance



“Microstructural” ingredients

“Macrostructure”

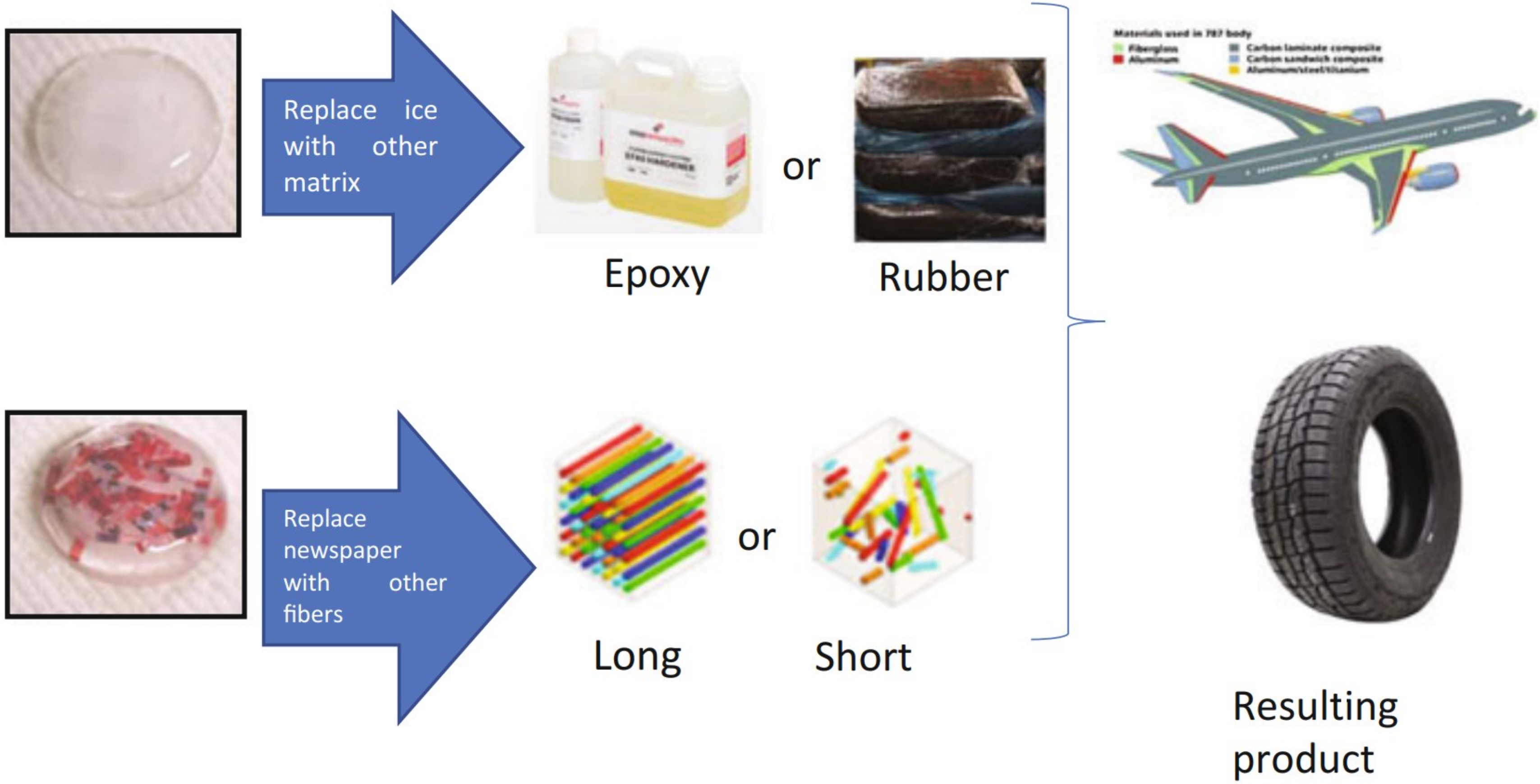




Filler material	Result
Control	Fractured, big chunks
Salt water	Small chunks
Egg	Fractured, big chunks
Sawdust	Fractured, small chunks
Wood chips	Very little fracturing
Coffee grounds	Fractured, big chunks
Blueberry	Fractured, big chunks

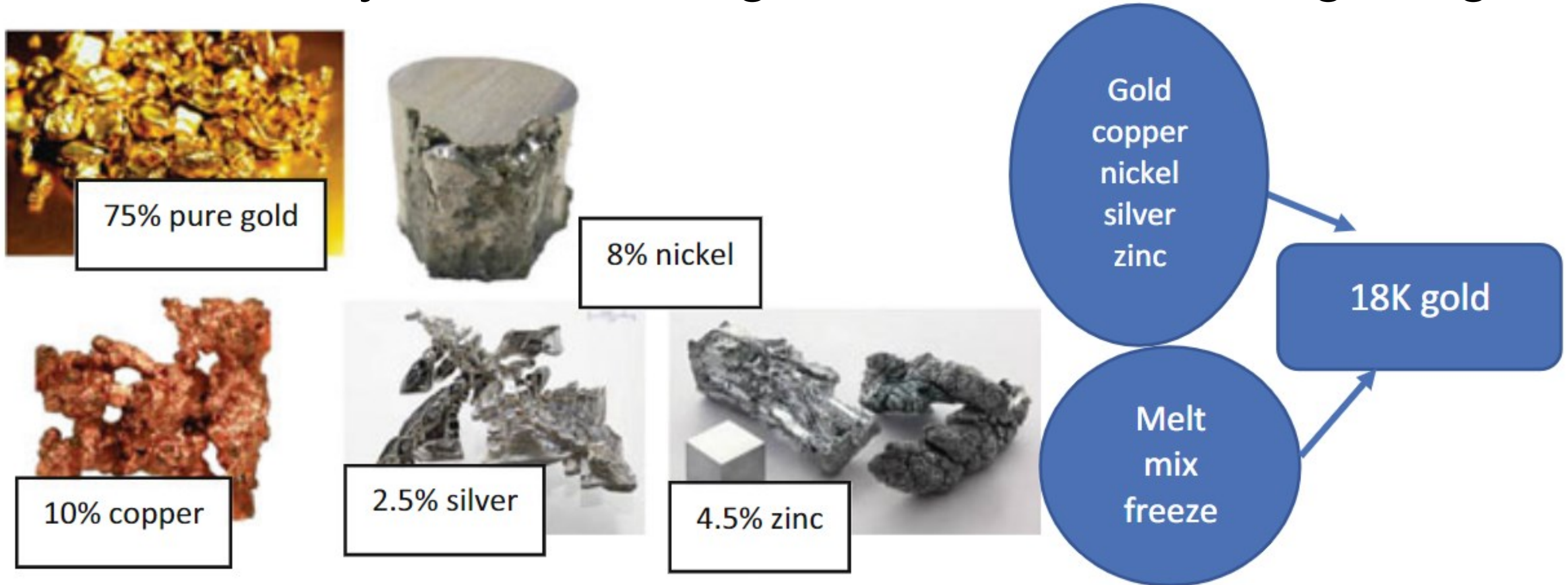
1.8 From Everyday Applications to Materials Design

- Reinforced ice cube vs. carbon fiber reinforced composite material, tire (rubber, steel or polymer cords)



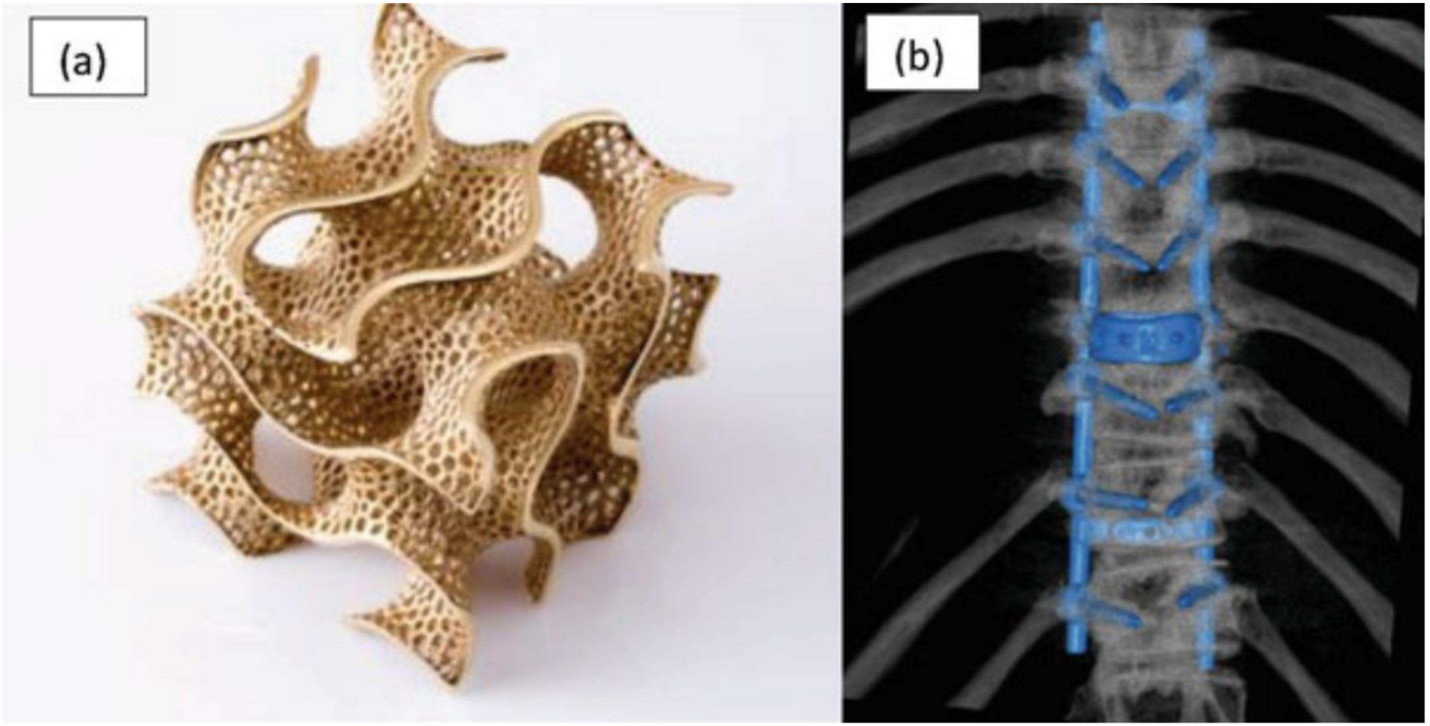
1.8 Data Science for Materials Design: Example

- Tire tread material design: durability
  - One of the key materials property metrics: tan(δ)
  - High tan(δ) for low temperatures (better ice and wet grip) and a low tan(δ) for high temperatures (better rolling friction)
  - Function of the matrix materials(rubber matrices and fillers), micro-structure, and the operating conditions such as temperature and frequency
- Gold and Gold Alloys for Wedding Cakes and Wedding Rings

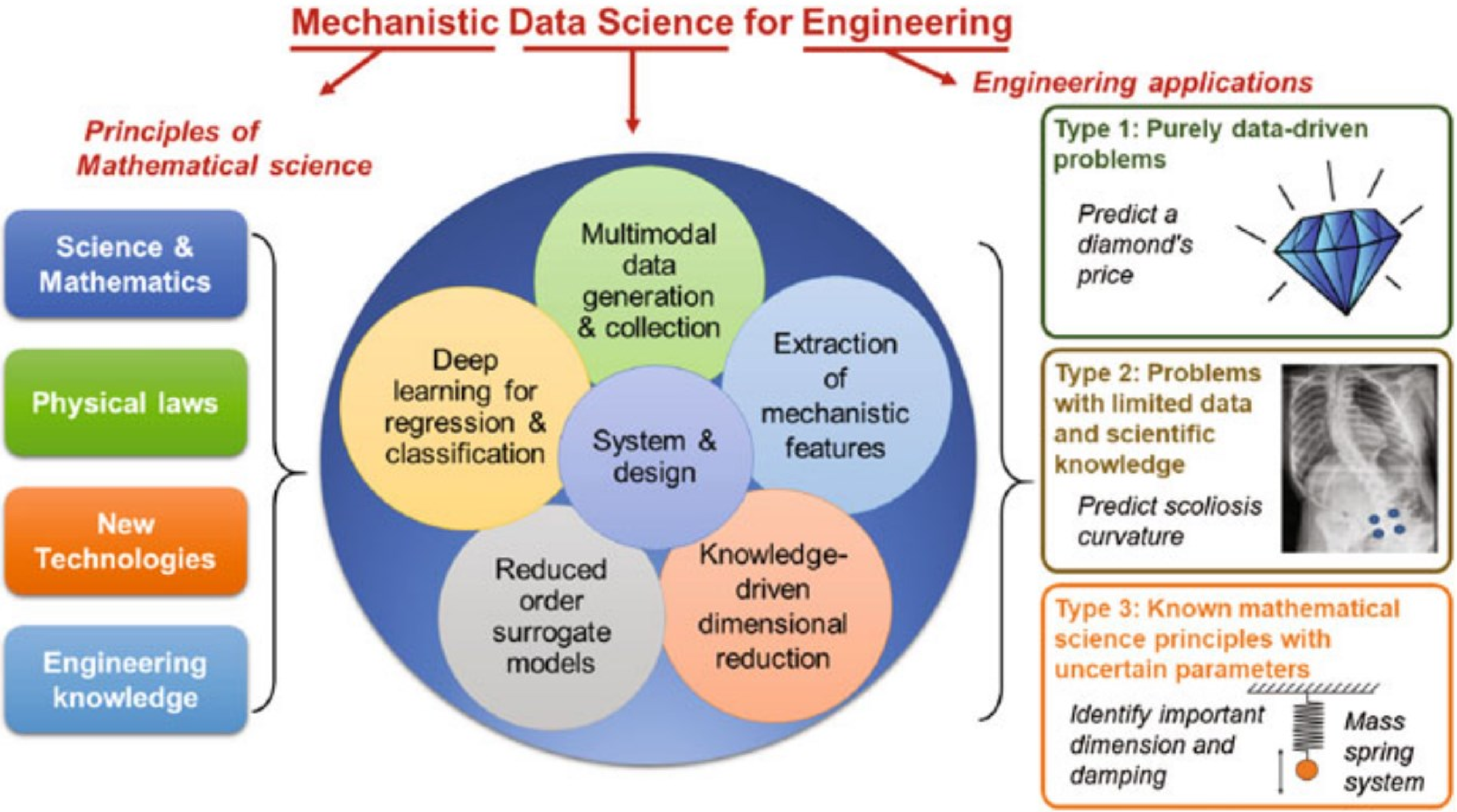


1.9 Twenty-First Century Data Science

- AlphaGo (Deep Mind, 2014)
  - in 2016, AlphaGo beat 18-time world champion Mr. Lee Sedol
- 3D Printing: From Gold Jewelry to Customized Implants



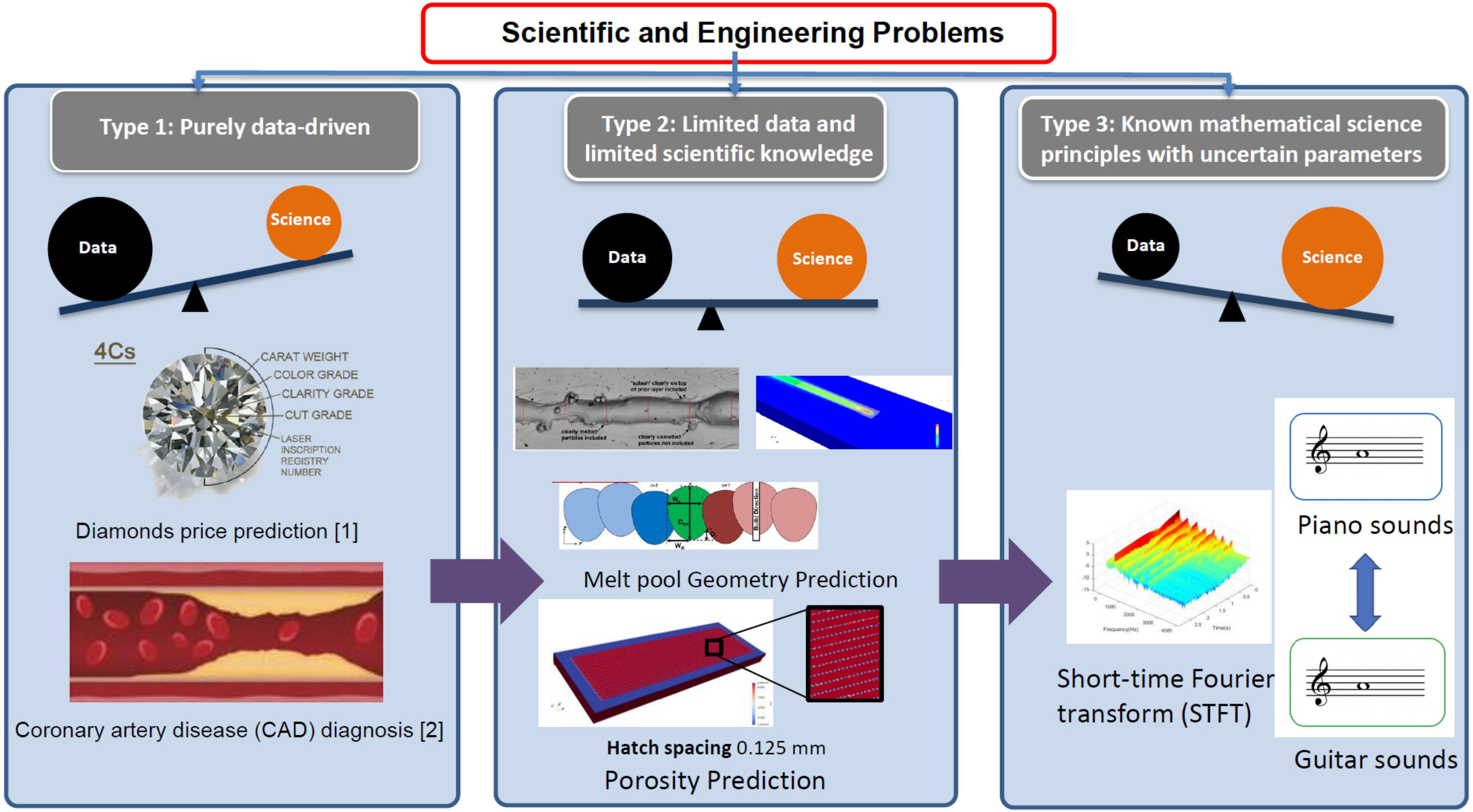
1.10 Outline of Mechanistic Data Science Methodology



Three Types of Problems addressed with MDS

- Type 1: purely data-driven problem
  - Problem with abundant data but undeveloped or unavailable fundamental principles
  - Example: marketing behavior of people based on characteristics such as age and gender
- Type 2: limited data and scientific knowledge
  - Neither the data nor the scientific principles provide a complete solution
  - Example: biomechanical problems such as scoliosis progression (fundamental scientific principles can be used to compute the direction of bone growth, but the data is required to characterized the effects of age and gender)
- Type 3: known mathematical science principles with uncertain parameters
  - Fundamental understanding of the world
  - Example: physics problems (determining the actual spring stiffness and damping properties of an actual spring mass system based on data collected from multiple cameras at different angles)

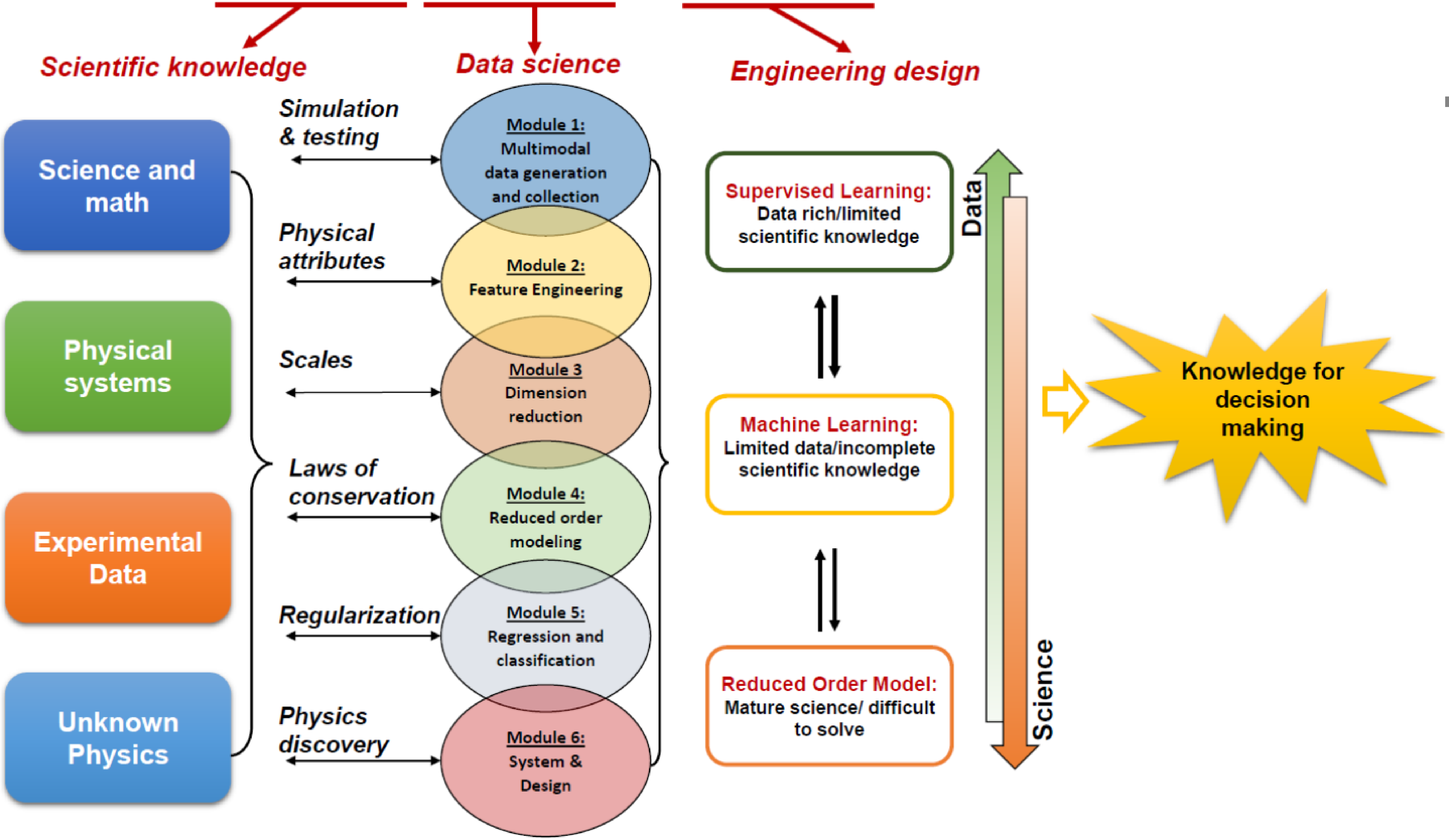
Three Types of Problems addressed with MDS



[1] https://www.cupidjewellery.com/pages/about-diamond-4c
[2]https://www.gleneagles.com.sg/healthplus/article/signs-you-need-to-check-your-heart

Z. Gan, KK. Jones, Y. Lu, WK. Liu, "Benchmark Study of Melted Track Geometries in Laser Powder Bed Fusion of Inconel 625 ", Invited Article, AFRL Benchmark Study. (2021)

Mechanistic data science for engineering



MDS Procedure

- Step 1: generation and collection of multimodal data (such as testing, simulations, or databases)
- Step 2: extract mechanistic features using basic mathematical tools, including continuous and discrete/digital analysis
- Step 3: perform knowledge-driven dimension reduction to streamline the analysis
- Step 4: reduced order surrogate models will be created to introduce the fundamental physics into the solution of the problem
  - Fourier analysis, regression, continuous and discrete mathematics, and image analysis
- Step 5: deep learning algorithms, such as neural networks, will be performed for regression and classification
- Step 6: These data and mechanistic analysis steps will be coupled for the system and design

Systems and Design

- Piano example with spring mass system (Type 3 general)
- Feature-based diamond pricing (Type 1 general)
- Additive manufacturing (Type 1 advanced)
- Composite design (Type 3 advanced)
- Indentation analysis for materials property prediction (Type 2 advanced)
- Early warning of rainfall induced landslides (Type 3 advanced)

1.11 Examples Describing the Three Types of MDS Problems

- Pure Data Science: Type 1
  - Determining Price of a Diamond Based on Features
  - Sports Analytics: “Moneyball”, Data Science for Optimizing a Baseball Team Roster
- Mixed Data Science and Surrogate: Type 2
  - Predicting Patient-Specific Scoliosis Curvature
- Type 3 Problem
  - Identifying Important Dimensions and Damping in a Mass-Spring System