

Data and Information Management in the Built Environment

Mark A. Austin

University of Maryland

austin@umd.edu

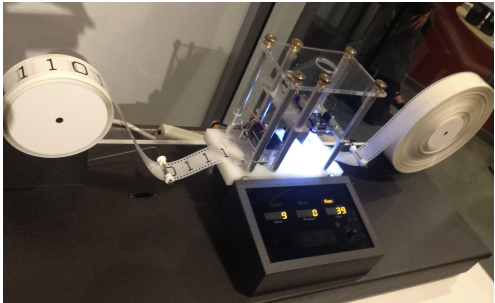
ENCE 688P, Spring Semester 2026

January 26, 2026

Early Models of Computing

Turing Machine Model: 1930s ...

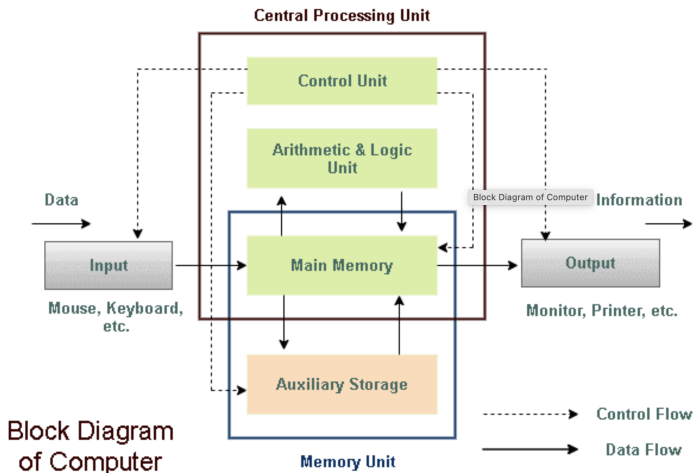
- Alan Turing (1936) created the **Turing machine** that included the **idea** of a **computer program**.



- Turing showed that you can **compute anything** using only **6 primitives**: right, left, print, scan, erase, nothing.

Early Models of Computing

Block Diagram of a Computer: 1980s ...



State-of-the-Art Computing

What does a modern computer do?

- Performs calculations – **billions** (sometimes even trillions) of **calculations per second**.
- Remembers results – **gigabytes** and terabytes of **storage**.

Modern Programming Languages

- **Modern programming languages** have a **more convenient set of primitives**.
- Can abstract methods to create new primitives (e.g., user-defined objects).
- **Anything computable in one language** is **computable** in **any other programming language**.

Man and Machine (Traditional View)

Man	Machine
<ul style="list-style-type: none">● Good at formulating solutions to problems.● Can work with incomplete data and information.● Creative.● Reasons logically, but very slow.● Performance is static.● Humans break the rules.	<ul style="list-style-type: none">● Manipulates Os and 1s.● Very specific abilities.● Requires precise descriptions of problem solving procedures.● Dumb, but very fast.● Performance doubles every 18-24 months.● Machines will follow the rules.

Sensible Problem Solving Strategy

Let engineers and computers play to their strengths:

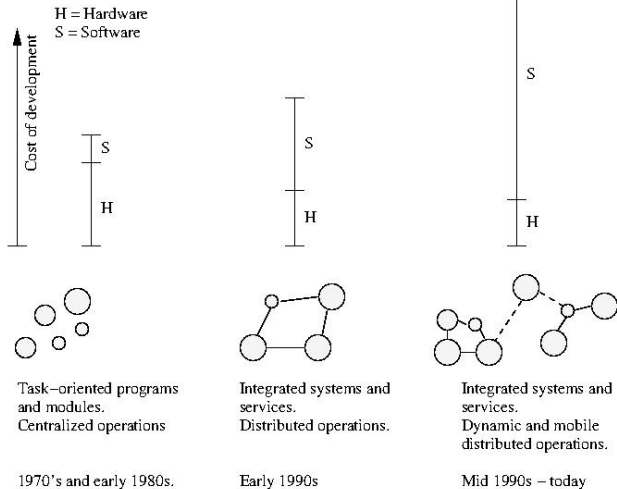
- Accelerates the solution procedure.
- Enables the analysis of problems having size and complexity beyond manual examination.
- Adds value in areas that will lead to long-term economic growth.

Getting things to work We need to:

- Describe to the computer solution procedures that are completely unambiguous.
- Look at data, organization and manipulation of data, and formal languages.

Expanding Expectations of Computing

Economics of computing and systems development



Evolution of Computer Languages

Computer Languages. Formal description – **precise grammar** – for how a problem can be solved.

Evolution. It takes about a decade for significant advances in computing to occur:

Capability	1970s	1980s	1990s
Users	Specialists	Individuals	Groups
Usage	Numerical computations	Desktop computing	E-mail, web, file transfer.
Interaction	Type at keyboard	Screen and mouse	audio/voice.
Languages	Fortran, C	MATLAB	Python, Java

Popular Computer Languages

Tend to be **designed** for a **specific set of purposes**:

- FORTRAN (1950s – today). Stands for formula translation.
- C (early 1970s – today). New operating systems.
- C++ (early 1970s – today). Object-oriented version of C.
- MATLAB (mid 1980s – today). Stands for matrix laboratory.
- Python (1990s – today). Object-oriented scripting language.
- HTML (1990s – today). Layout of web-page content.
- Java (1994 – today). Object-Oriented language for network-based computing.
- XML (late 1990s – today). Description of data on the Web.

Post- 2000 Era

Imagine: What if COVID-19 had arrived in 2000?

- No iPhone, No iPad, No iTunes.
- No Facebook, No Instagram, No WhatsApp.
- No Google Maps, No Google Streetview.
- No Dropbox, No Zoom.

Recent Advances in Technology:

- Average internet speeds: In 2000, 0.07 Mbs; In 2009, 5-7 Mbs; In 2020, 100-200 Mbs; 5G, 1000-2000 Mbs.
- Cloud-based data storage and computational services (AWS).
- New languages: [Swift](#) → App development on [iPhone/iPad](#).
- Many new types of [sensors](#) and [methods of data collection](#).

Post-2010 Era → Emergence of AI

State-of-the-Art Implementation (2020, Google, Siemens, IBM)

- AI and ML will be **deeply embedded** in new **software and algorithms**.

Artificial Intelligence:

- **Knowledge representation** and **reasoning** with ontologies and rules. Semantic graphs. Executable **event-based processing**.

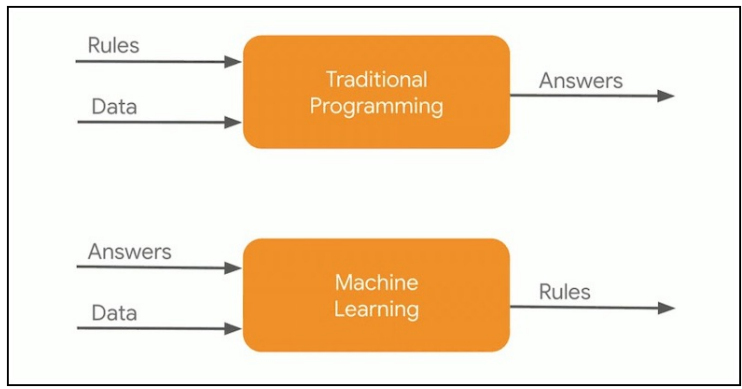
Machine Learning:

- Modern neural networks. Input-to-output prediction.
- Data mining.
- Identify **objects**, **events**, and **anomalies**.
- Learn structure and sequence. **Remember stuff**.

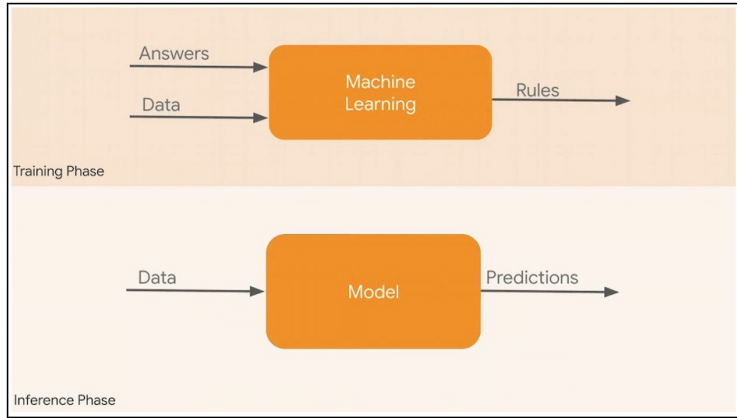
Man and Machine (AI-ML View)

Man	AI-ML Machine
<ul style="list-style-type: none"> ● Good at formulating solutions to problems. ● Can work with incomplete data and information. ● Creative. ● Reasons logically, but very slow. Forgetful. ● Performance is static. ● Humans make the rules, then they break them. 	<ul style="list-style-type: none"> ● Manipulates 0s and 1s. ● Can work with incomplete data and information. ● Creative. ● Fast logical reasoning. ● Performance doubles every 18-24 months. ● Data mining can discover the rules.

Traditional Programming vs AI-ML Workflow

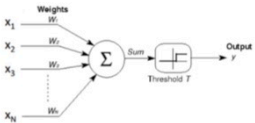


Traditional Programming vs AI-ML Workflow



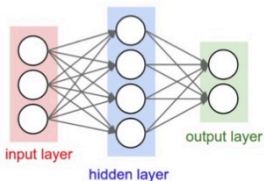
Machine Learning Capabilities (1980-1990)

Expressive Power of a Neural Network

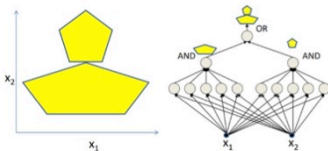
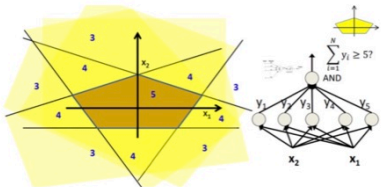


$$y = \begin{cases} 1 & \text{if } \sum_{i=1}^d w_i x_i \geq T \\ 0 & \text{else} \end{cases}$$

Neural Network with Single Hidden Layer

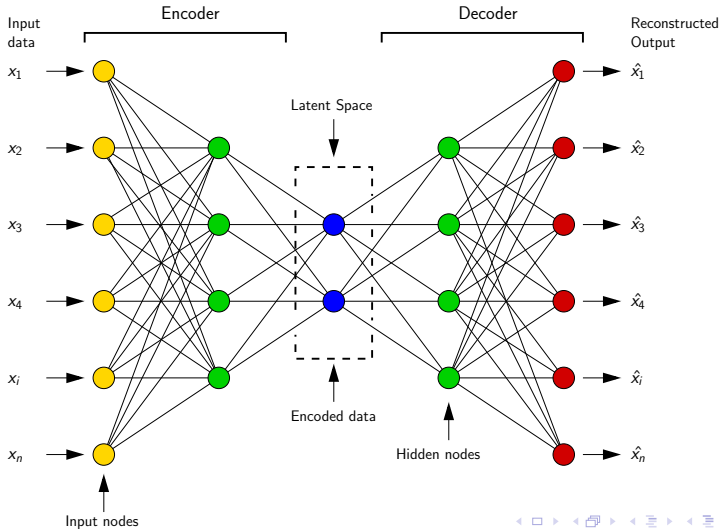


Approximation of Functions / Boolean Logic



Classification of Machine Learning Problems

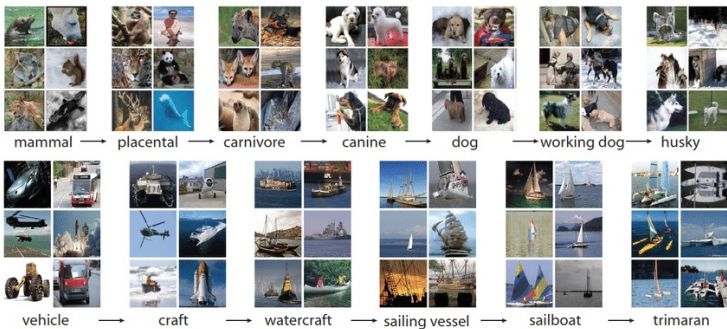
AutoEncoder (Encoder-Decoder-Reconstruction)



Classification of Machine Learning Problems

ImageNet and Deep Learning Capabilities:

- Identify objects in an image.
- 27 high-level categories; 21,800 sub-categories.



ImageNet and Deep Learning

Capabilities (2018):

- Identify relationship among multiple objects in a image.

Example. Dog riding skateboard



ImageNet and Deep Learning

Captions generated by a neural network:

The screenshot shows a web browser window with the address bar at `cs.stanford.edu/people/karpathy/deepimagesent/`. The browser tabs include "40 maps that explain...", "Amazon Web Services", "Primer | Math n Pro...", "deeplearning.net/tu...", "Deep Learning Tutor...", "deep learning", "PHILIPS - Golden Ears", "Language Technology", "MyDCare - Dashbo...", and "Other bookmarks".

The main content area displays eight images arranged in a 2x4 grid, each with a caption below it:

- Image 1: A man in a black t-shirt playing an electric guitar. Caption: "man in black shirt is playing guitar."
- Image 2: A construction worker in an orange safety vest working on a road. Caption: "construction worker in orange safety vest is working on road."
- Image 3: Two young girls playing with colorful lego toys. Caption: "two young girls are playing with lego toy."
- Image 4: A boy performing a backflip on a wakeboard in a lake. Caption: "boy is doing backflip on wakeboard."
- Image 5: A young girl in a pink dress jumping in the air. Caption: "girl in pink dress is jumping in air."
- Image 6: A black and white dog jumping over a blue and white striped bar. Caption: "black and white dog jumps over bar."
- Image 7: A young girl in a pink shirt swinging on a swing. Caption: "young girl in pink shirt is swinging on swing."
- Image 8: A man in a blue wetsuit surfing on a wave. Caption: "man in blue wetsuit is surfing on wave."

Machine Learning at Scale (Transformers)

Input

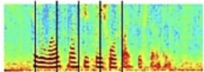
Output

Pixels:



"leopard"

Audio:



"How cold is it outside?"

"Hello, how are you?"



"Bonjour, comment allez-vous?"

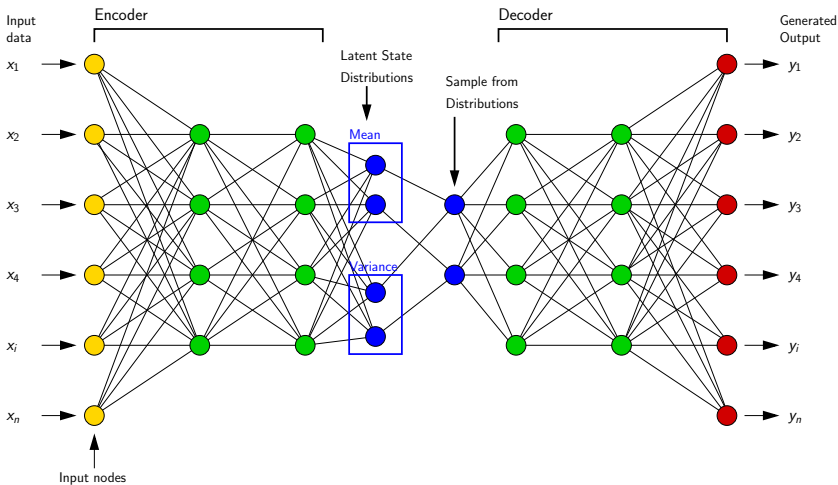
Pixels:



"A cheetah lying on top of a car"

Post-2020 Era → Explosion of Generative AI

Variational AutoEncoders (Generative Models)

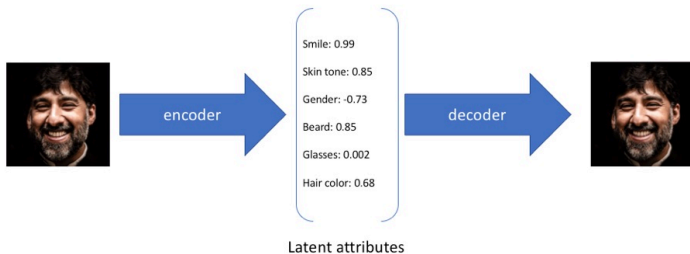


Post-2020 Era → Explosion of Generative AI

Standard Autoencoders vs. Variational Autoencoders:

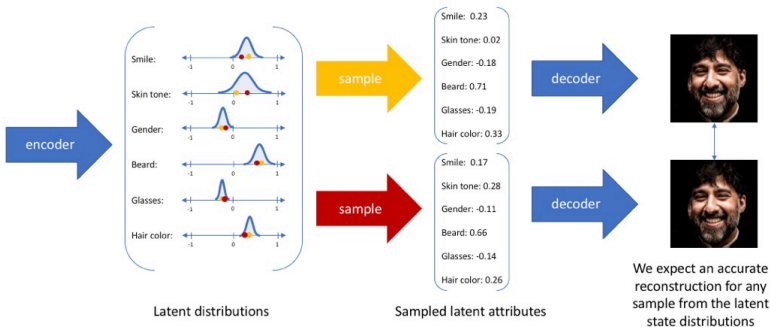
- A **standard autoencoder** outputs a **single value** for each **encoding dimension**.
- **Variational autoencoders** provide a **probability distribution** for each latent attribute.

Example: Single value representations for latent attributes:



Post-2020 Era → Explosion of Generative AI

Image Reconstruction: sampled from latent distributions ...



Source: Jordan J., Variational Autoencoders, Data Science, March 2018.

Post-2020 Era → ChatGPT

ChatGPT: Generative Pre-Trained Transformer

AU

write paragraph on civil engineering



Civil engineering is a fundamental discipline that shapes and transforms the physical infrastructure of our modern world. It encompasses the design, construction, and maintenance of a vast array of structures and systems, ranging from buildings, bridges, and roads to water supply networks, sewage systems, and environmental protection facilities. Civil engineers are responsible for ensuring the safety, functionality, and sustainability of these essential structures that support and enhance our daily lives. Their work involves a deep understanding of materials, mechanics, geotechnical engineering, structural analysis, and urban planning. By integrating scientific principles with innovative technologies, civil engineers address societal needs, address challenges posed by urbanization, and contribute to the creation of resilient and efficient urban environments. Whether it's developing eco-friendly infrastructure, earthquake-resistant buildings, or efficient transportation systems, civil engineering remains at the forefront of shaping a better and more connected world.



Free Research Preview. ChatGPT may produce inaccurate information about people, places, or facts. [ChatGPT August 3 Version](#)

Post-2020 Era → ChatGPT

AU integrate $\sin(3x) \cdot \sin(2x)$ dx



To integrate the expression $\int \sin(3x) \cdot \sin(2x) dx$, we can use the product-to-sum trigonometric identity, which states that:



$$\sin(A) \cdot \sin(B) = (1/2) \cdot [\cos(A - B) - \cos(A + B)].$$

Using this identity, we can rewrite the integral as:

$$\begin{aligned} \int \sin(3x) \cdot \sin(2x) dx &= (1/2) \int [\cos(3x - 2x) - \cos(3x + 2x)] dx \\ &= (1/2) \int [\cos(x) - \cos(5x)] dx. \end{aligned}$$

Now, we can integrate each term separately:

$$\int \cos(x) dx = \sin(x) + C_1,$$

$$\int \cos(5x) dx = (1/5)\sin(5x) + C_2.$$

Putting it all together:

$$\begin{aligned} \int \sin(3x) \cdot \sin(2x) dx &= (1/2) \cdot [\sin(x) + C_1 - (1/5)\sin(5x) - C_2] \\ &= (1/2) \cdot [\sin(x) - (1/5)\sin(5x)] + (C_1/2 - C_2). \end{aligned}$$

So, the integral of $\sin(3x) \cdot \sin(2x)$ with respect to x is:

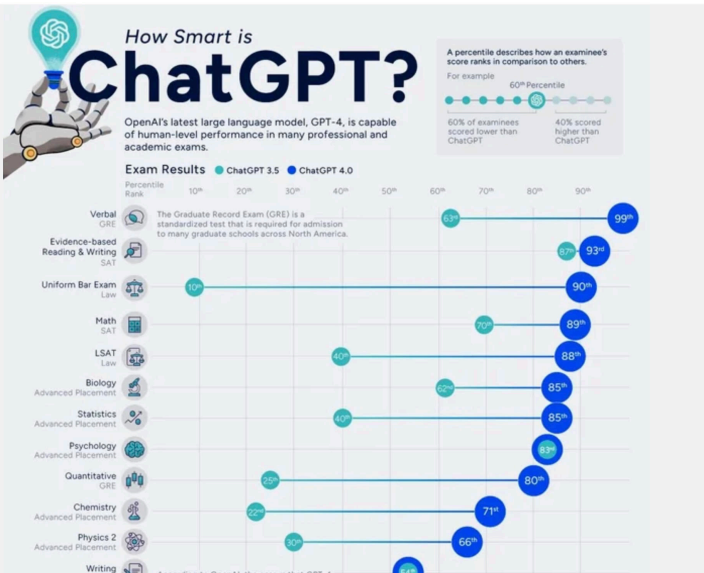
$$\int \sin(3x) \cdot \sin(2x) dx = (1/2) \cdot [\sin(x) - (1/5)\sin(5x)] + C,$$

where $C = C_1/2 - C_2$ is the constant of integration.

Regenerate



Post-2020 Era → ChatGPT



Post-2020 Era → AI Generated Architecture

Convergence: Engineering-Architecture-AI

AI-generated art ...



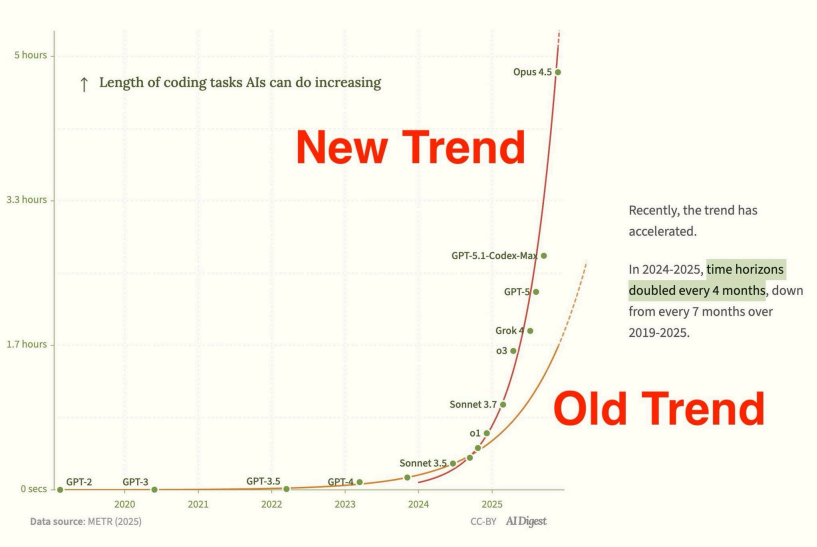
AI-generated building architecture



Post-2020 Era → AI Video Presentation



Post-2020 Era → AI Code Generation



Machine Learning in CEE

Opportunities for Machine Learning in CEE:

- Predicting system response and performance.
- Interpreting data and formulating models to predict component and subsystem-level properties.
- Information retrieval from images and text.
- Recognizing patterns in streams of sensed data.

Economic Considerations (CS Perspective):

- **Urban infrastructure** is **permanent/semi-permanent** and **very expensive** to build and maintain.
- Prioritize **improvements** to **efficiency** by identifying and **removing bottlenecks** in performance.
- Use **AI-ML** to **identify events**, **cause-and-effect relationships**, and **design of actions** that **enhance system performance**.

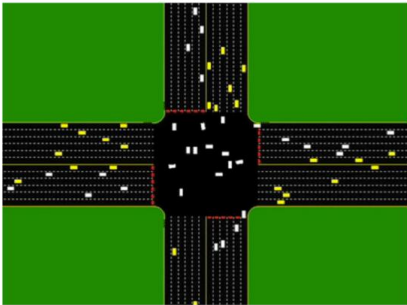
AI-ML Enabled Decision Making

AI Traffic System: Reduced congestion and accident rates:



AI-ML Enabled Decision Making (Self-Driving Cars)

Goal. Improve performance by removing bottlenecks → **no human driver**; **no traffic lights**.



Remark: 95% of the requirements are for the system software.

Source: ISR visitor from GM Research.

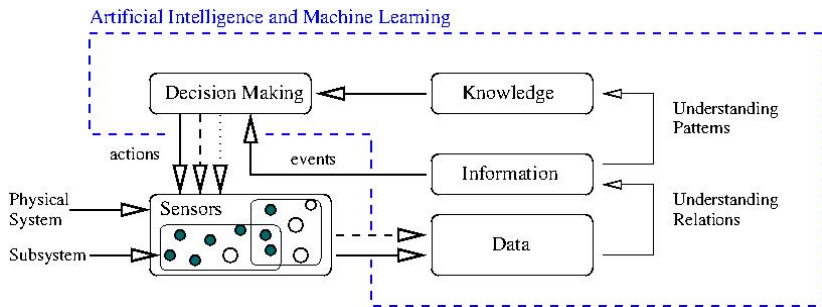
Remark: Tesla will produce self-driving cars by 2016.

Source: Elon Musk.

Stop signs and **traffic lights** are replaced by **mechanisms for vehicle-to-vehicle communication** (Adapted from <http://citylab.com>).

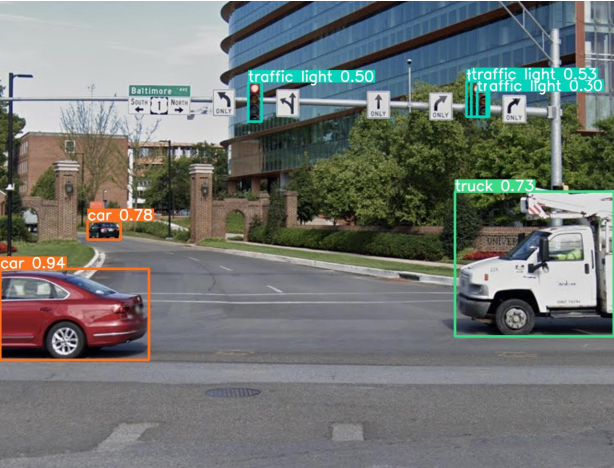
AI-ML Enabled Decision Making (Self-Driving Cars)

Solution Procedure. Pathway from **sensing** and **data collection** to ... action ... improved performance, now **enabled** by **AI** and **ML** capabilities:



AI-ML Enabled Decision Making (Self-Driving Cars)

Ainur's Experiments with Computer Vision (OpenCV):



Google DeepMind (2018-2020)

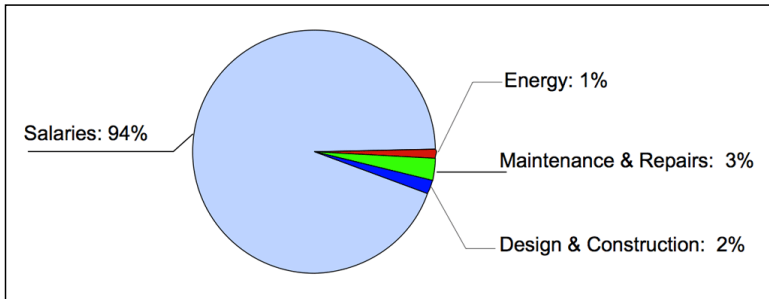
Teach Self-Driving Cars to Navigate a City without a Map



Test Cities: London, Paris, New York.

Modern Buildings (Economics)

Lifecycle costs in office buildings over a 30-Year period:



Energy systems have a huge impact on building occupant comfort and indoor air quality which, in turn, affects salary performance.

Source: United Technologies Research Center, 2009.

Modern Buildings (Integrated Energy Systems)

Trend toward Integrated Energy Systems:

- Commercial and residential buildings consume 1/3 of the world's energy.
- And by 2025, buildings will consume more energy than the transportation and industrial sectors combined.
- **Standard models** of building operation rely on **centrally produced power** as a source of high-grade energy.
- Advances in technology allow for consideration of alternatives, such as **local production of power**.

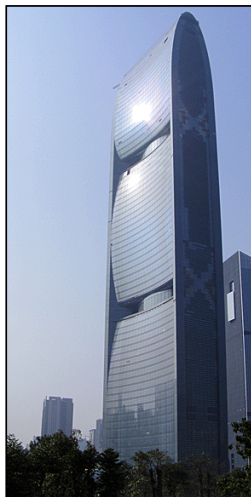
Examples:

- Solar power; small-scale combined heat and power systems.
- **Electricity production** through use of **ducted wind turbines**.

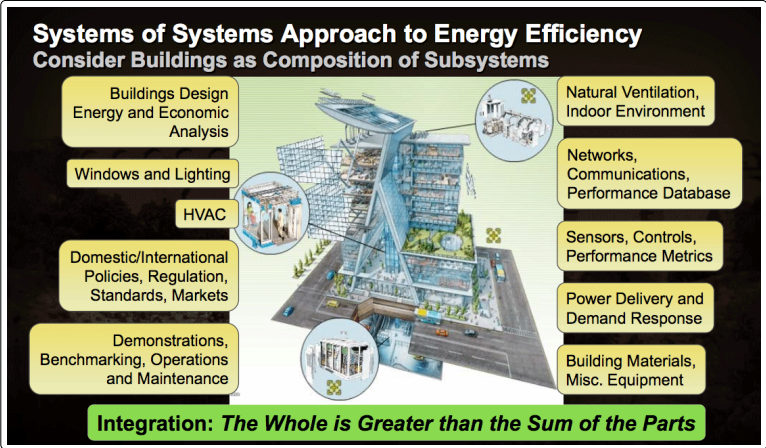
Modern Buildings (Integrated Energy Systems)

Pearl River Tower (2010):

- High performance structure designed to produce as much energy as it consumes.
- Guides wind to a pair of openings at its mechanical floors.
- Wind drives turbines that generate energy for the heating, ventilation and air conditioning systems.
- Openings provide structural relief, by allowing wind to pass through the building.

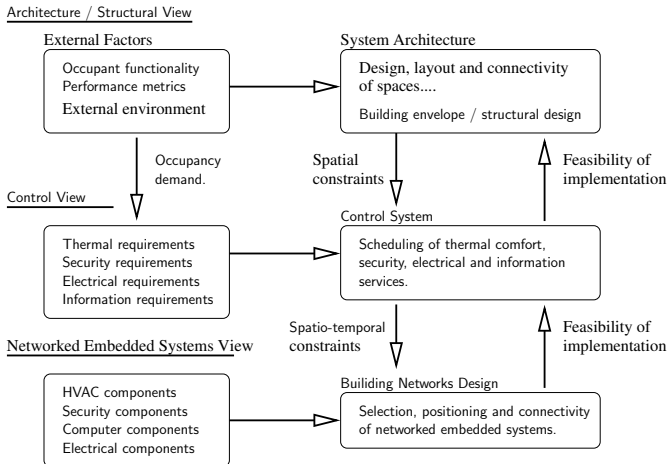


Modern Buildings (Automation Systems Design)



Modern Buildings (Traditional Approach to Design)

Interaction of Multiple-Domains:



Modern Buildings (Platform-Based Design)

Factors Driving Design

Architectural requirements.
Occupancy requirements.
External loads (gravity, thermal, ...)

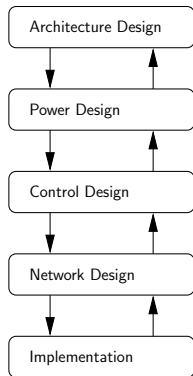
Ventilation requirements.
Energy generation requirements.

Sequence of operations.
Comfort requirements.

Control speed requirements.
Sensor and actuator requirements.

Layout requirements.

Design Flow



Performance

Maximum ventilation.
Maximum power generation.
Cost estimates.

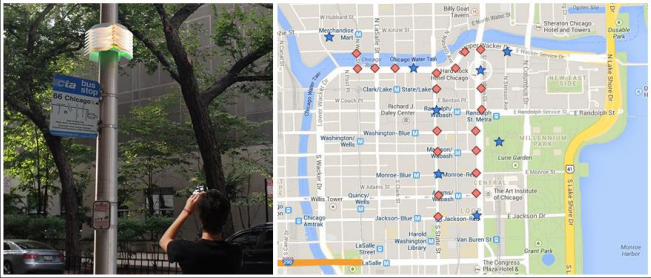
Minimum response time.
Control accuracy.

Maximum available bandwidth.
Maximum computational speed.
Maximum storage size.

Actual ventilation.
Actual power generation.
Actual network speed.
Actual layout constraints.
Actual installation cost.

Smart Cities: Urban Sensing in Chicago

Array of Things, Chicago (EOL 2022). Modular sensor boxes collect real-time data on the environment, infrastructure and activity.



Basic Questions. How is the city used? What is going on?

Source: <https://arrayofthings.github.io>

Smart Cities: Urban Sensing in Chicago

What Data is Collected?

The nodes will initially measure temperature, barometric pressure, light, vibration, carbon monoxide, nitrogen dioxide, sulfur dioxide, ozone, ambient sound intensity, pedestrian and vehicle traffic, and surface temperature. Continued research and development will help create sensors to monitor other urban factors of interest such as flooding and standing water, precipitation, wind, and pollutants.

Array of Things is interested in monitoring the city's environment and activity, not individuals. In fact, the technology and policy have been designed to specifically avoid any potential collection of data about individuals, so privacy protection is built into the design of the sensors and into the operating policies. Array of Things will not collect any personal or private information.

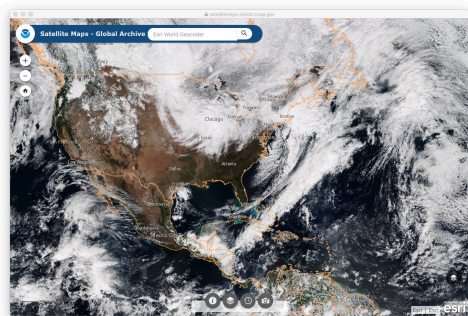


Global Applications

Answering Big Science Questions

Satellite Imagery and Measurements

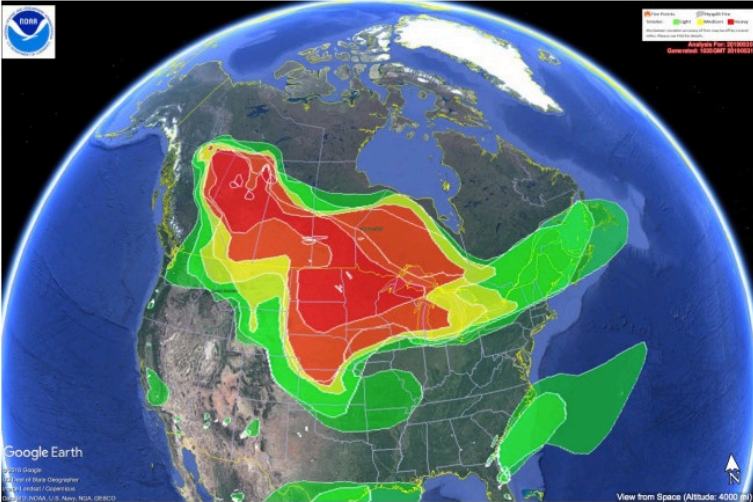
Understanding Climate Change



Example. Measure spatial and temporal extent of annual Snow Pack → Estimate water resources available for agriculture and urban consumption.

Canadian Wildfires impact US

Wildfires in Alberta: Smoke covers millions of square miles:



Recurring Themes and Key Points

Recurring Themes

- Information-age systems offer enhanced functionality and better performance, but their design is more difficult than in the past.
- Physical systems and computational systems fail in completely different ways.
- **Sensor networks** will form the **eyes and ears** of complex control and information systems.
- As system complexity increases, **more and more of the functionality** will be **managed by software!**

