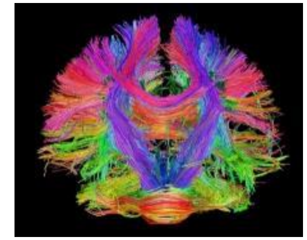



CS-570
Statistical Signal Processing

Lecture 1: Introduction to CS-570

Spring Semester 2019

Grigorios Tsagkatakis

Today's Objectives

- CS-570 Overview

- Introduction to Statistical Signal Processing

About CS-570

- Lectures

➢Monday 12:00-14:00, H208

➢Wednesday 14:00-16:00, H204

- Office Hours: 1 Hour after each class

- Prerequisites: Digital Signal Processing (CS-370), Applied Mathematics for Engineers (CS-215), Probabilities (CS-217)

About CS-570

Practical Information

- 2 individual homeworks on the material taught (30% of your final grade)
  - Exercises on MATLAB/python
  - 1st assignment will be handed out at the beginning of March
  - 2 weeks time to complete each assignment (hard deadline).

- Standalone project (max for 2 students) (50% of your final grade)
  - Research topic
  - Experimental work / analysis on experimental data
  - Submission of a project report in a technical paper form (motivation, related work, problem formulation, adopted methodology, results, conclusions & outlook)
  - Duration: mid of April - End of semester (~mid of June)

- Written Exam (20% of your final grade)

All above are compulsory for getting a grade at the end of the exam

Topics

Week 1: Introduction to Statistical Signal Processing & Review

Week 2: Introduction to optimization

Week 3: Signal sensing and reconstruction

Week 4: Computational imaging

Week 5: Deterministic signal processing

Week 6: High and low dimensional signal processing

Week 7: Statistical signal models

Week 8: Time-series modeling

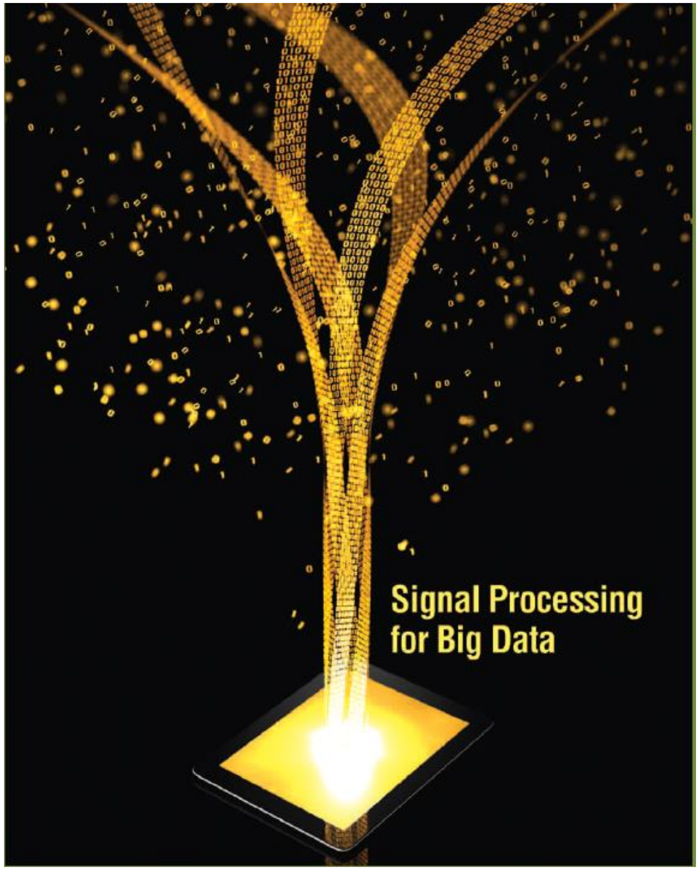
Week 9: Distributed signal processing

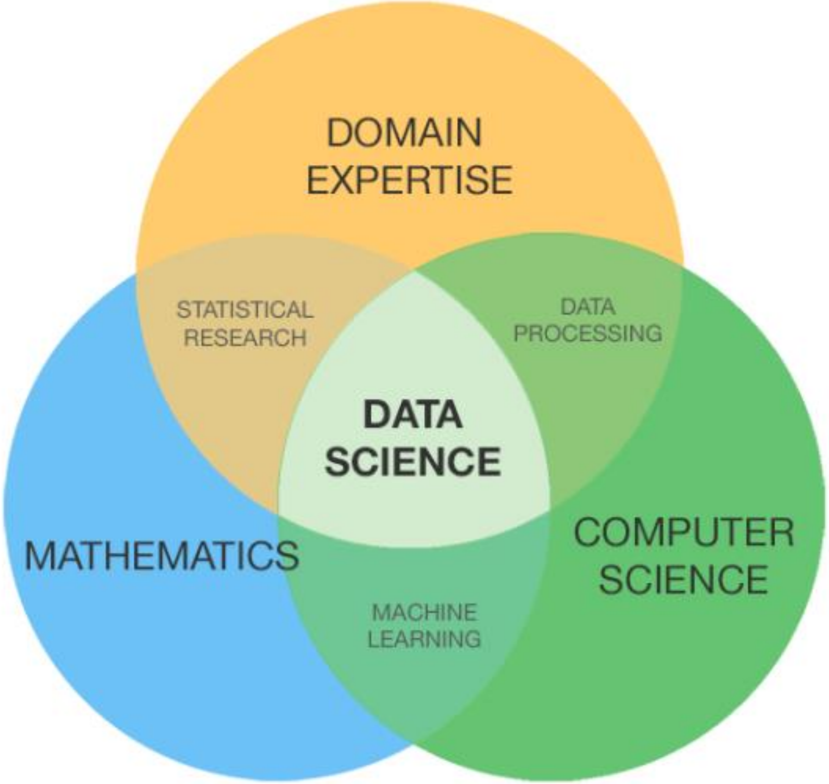
Week 10: Machine learning for signal processing

Week 11: Applications in remote sensing

Week 12: Applications in Internet-of-Things

Week 13: Review and presentations


Signal Processing
for Big Data

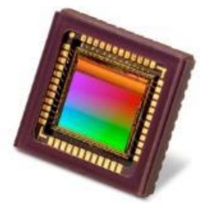

DOMAIN
EXPERTISE
STATISTICAL
RESEARCH
DATA
PROCESSING
DATA
SCIENCE
MATHEMATICS
COMPUTER
SCIENCE
MACHINE
LEARNING

Signal Processing fundamentals

- **Acquisition**
- Processing
- Analysis
- Storage
- Transmission

Temperature & Humidity

Image

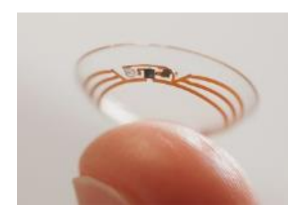


Sound

Pressure

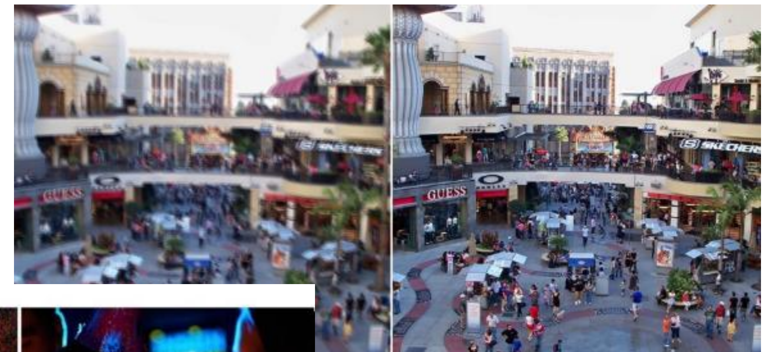
Vibration, Motion

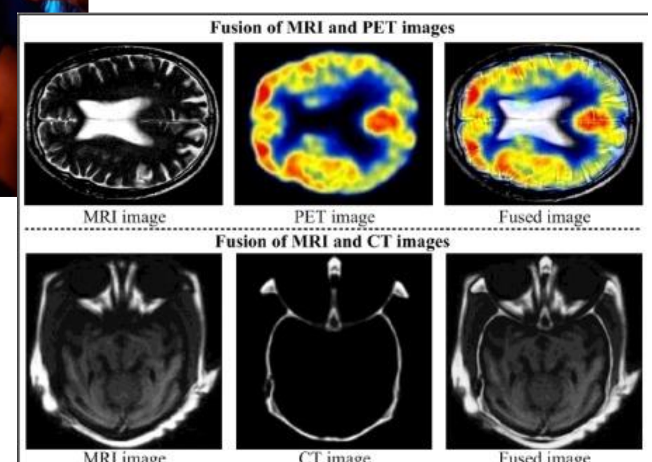
Glucose (&biometrics)



Signal Processing fundamentals

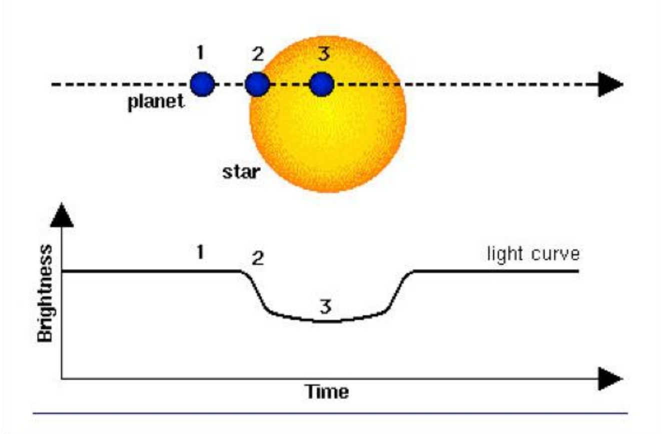
- Acquisition
- **Processing**
- Analysis
- Storage
- Transmission

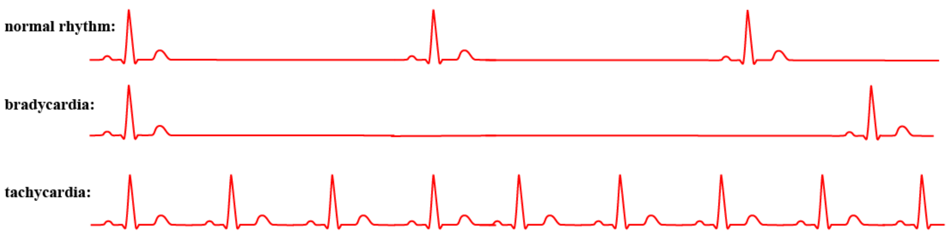


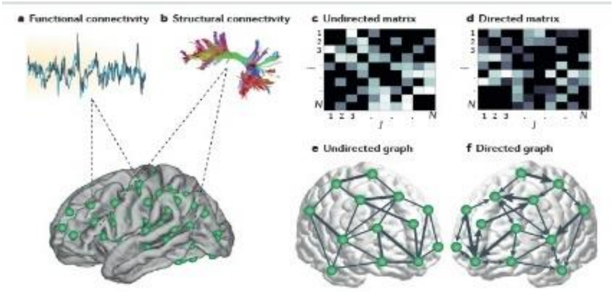


Signal Processing fundamentals

- Acquisition
- Processing
- **Analysis**
- Storage
- Transmission







Signal Processing fundamentals

- Acquisition
- Processing
- Analysis
- **Storage**
- Transmission

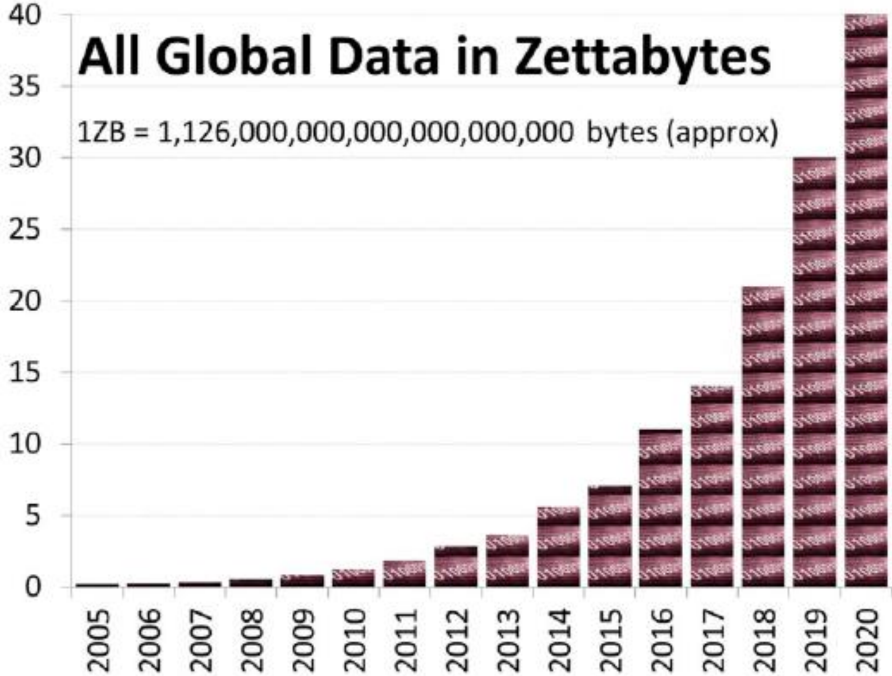
Signal Processing fundamentals

- Acquisition
- Processing
- Analysis
- Storage
- **Transmission**

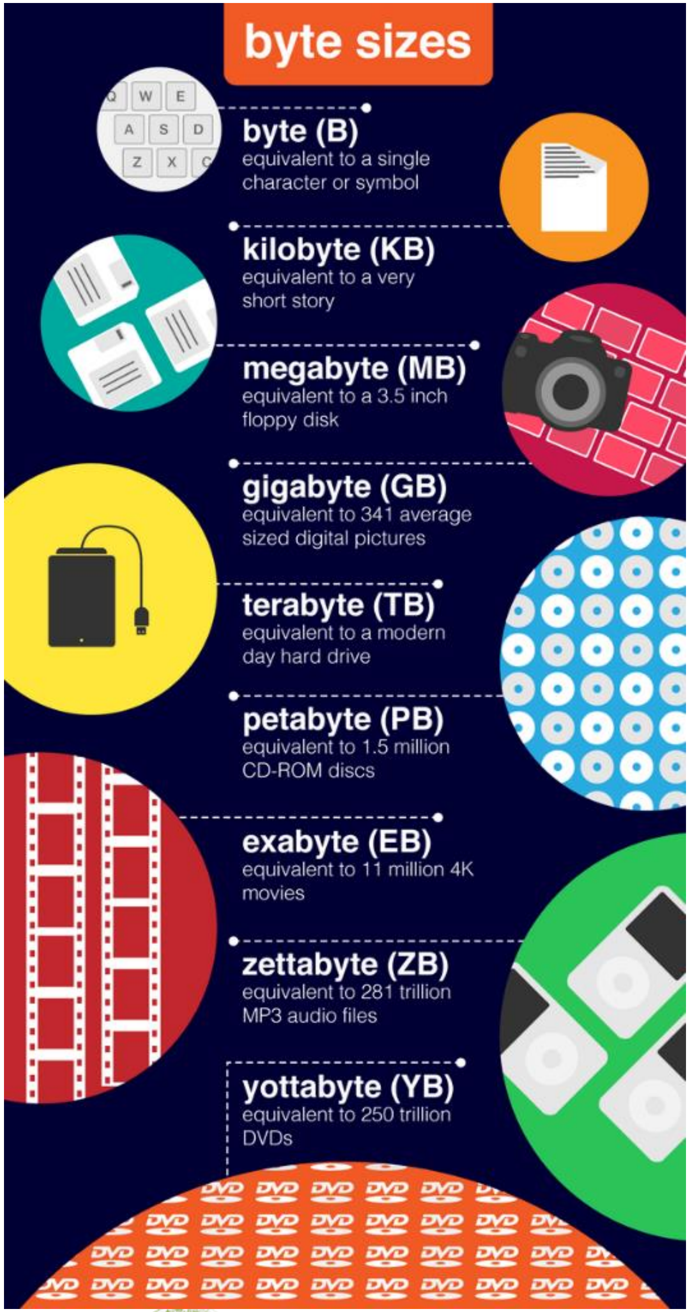
Big Data

The 5Vs

- **Volume**



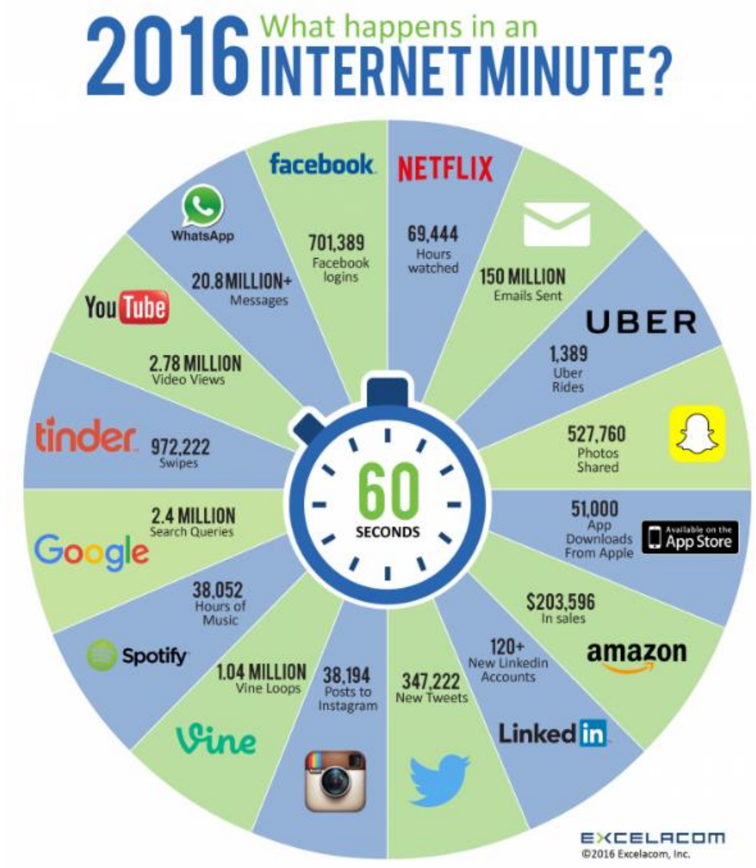
The growth in data as seen by United Nations Economic Commission for Europe.



Big Data

The 5Vs

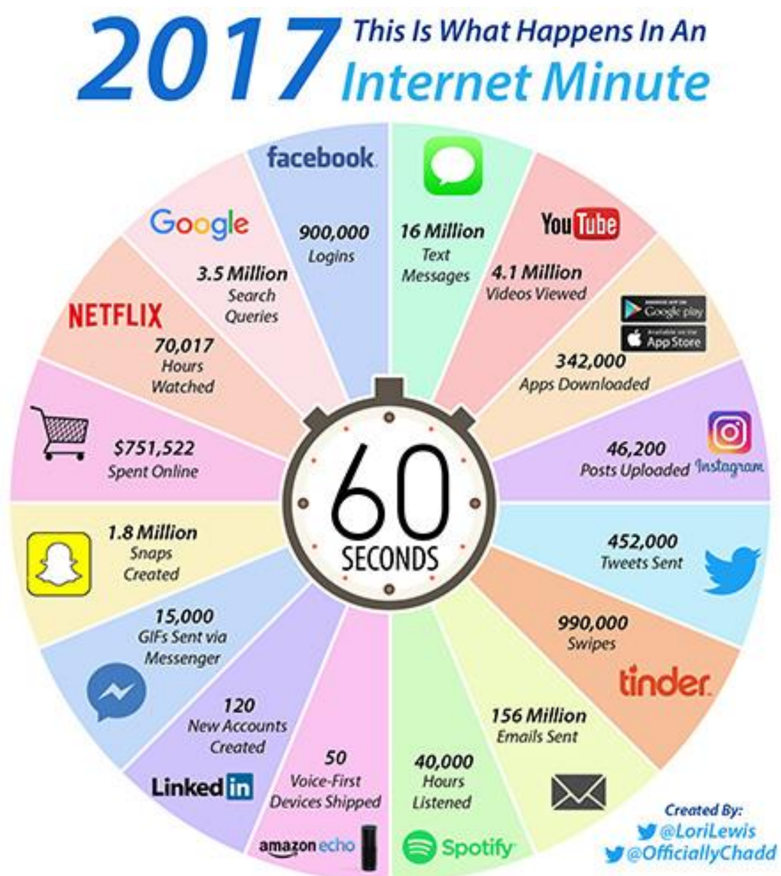
- Volume
- **Velocity**



Big Data

The 5Vs

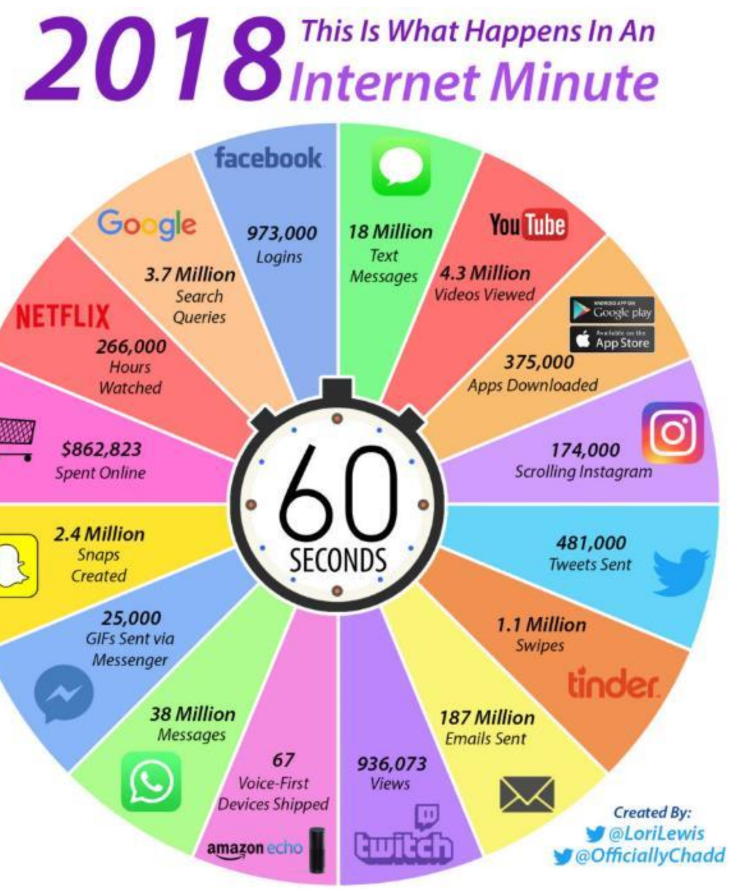
- ➢Volume
- ➢**Velocity**



Big Data

The 5Vs

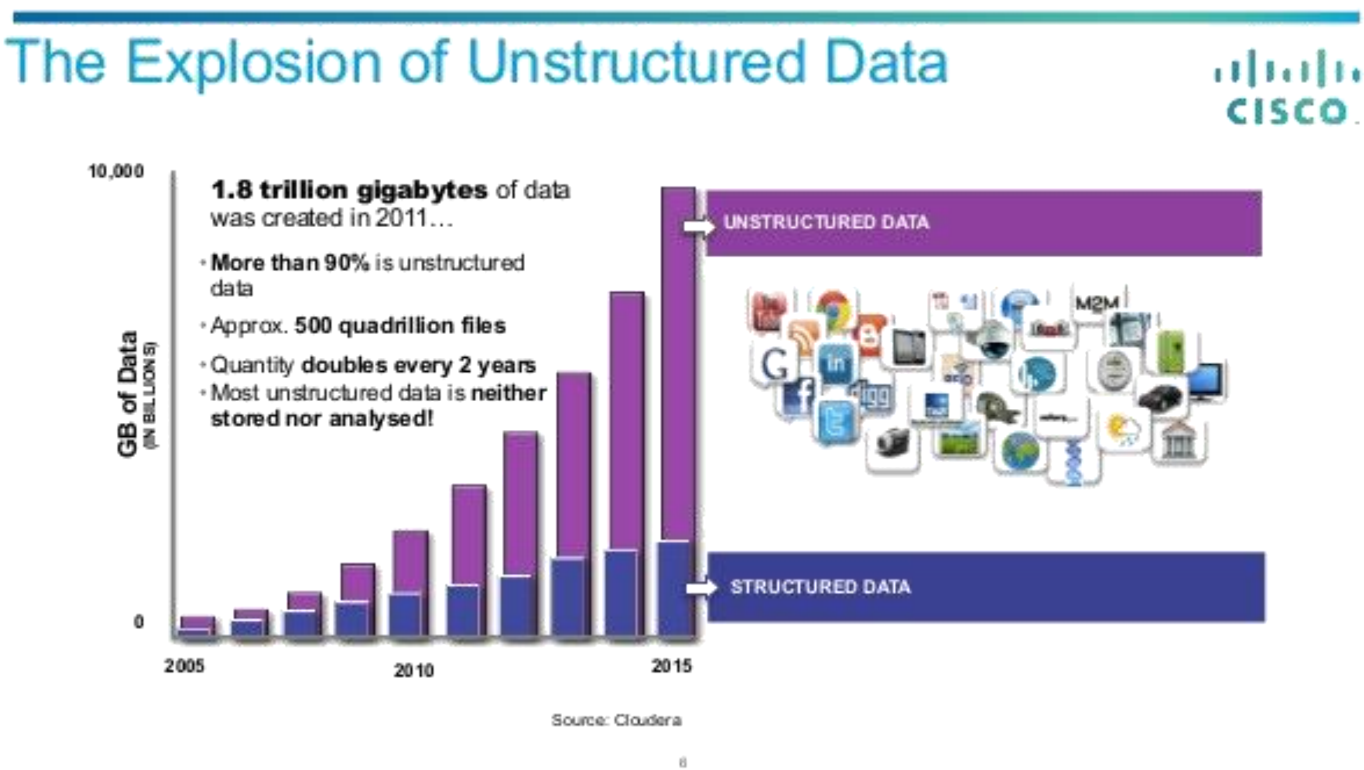
- ➢Volume
- ➢**Velocity**



Big Data

The 5Vs

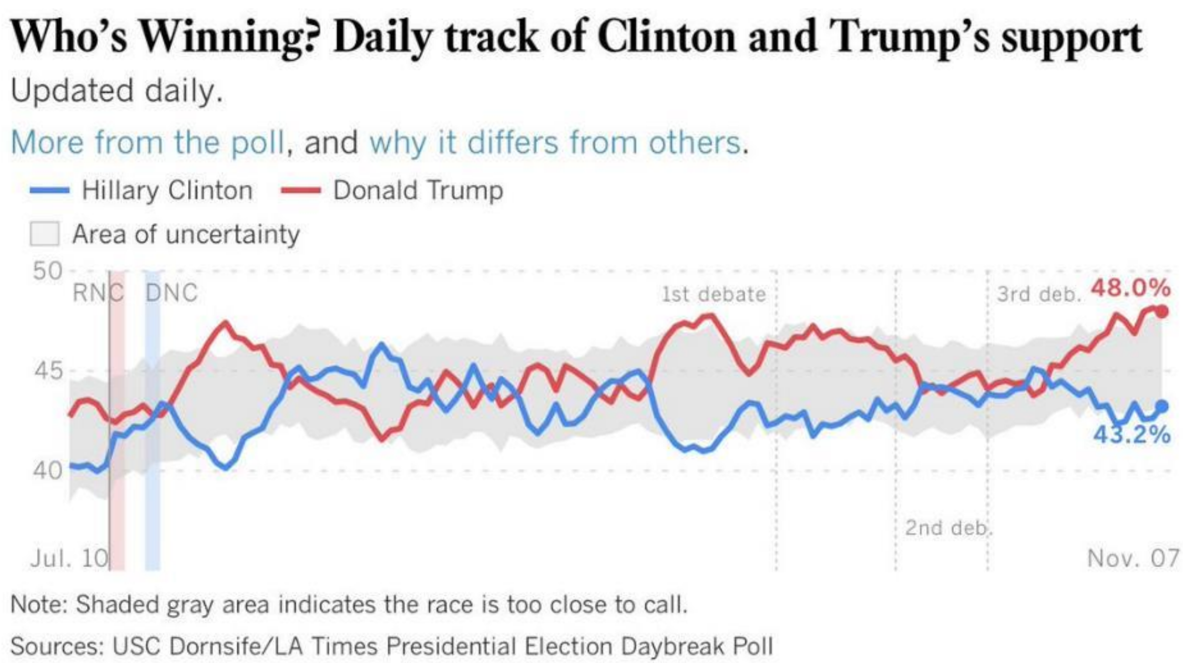
- Volume
- Velocity
- **Variety**



Big Data

The 5Vs

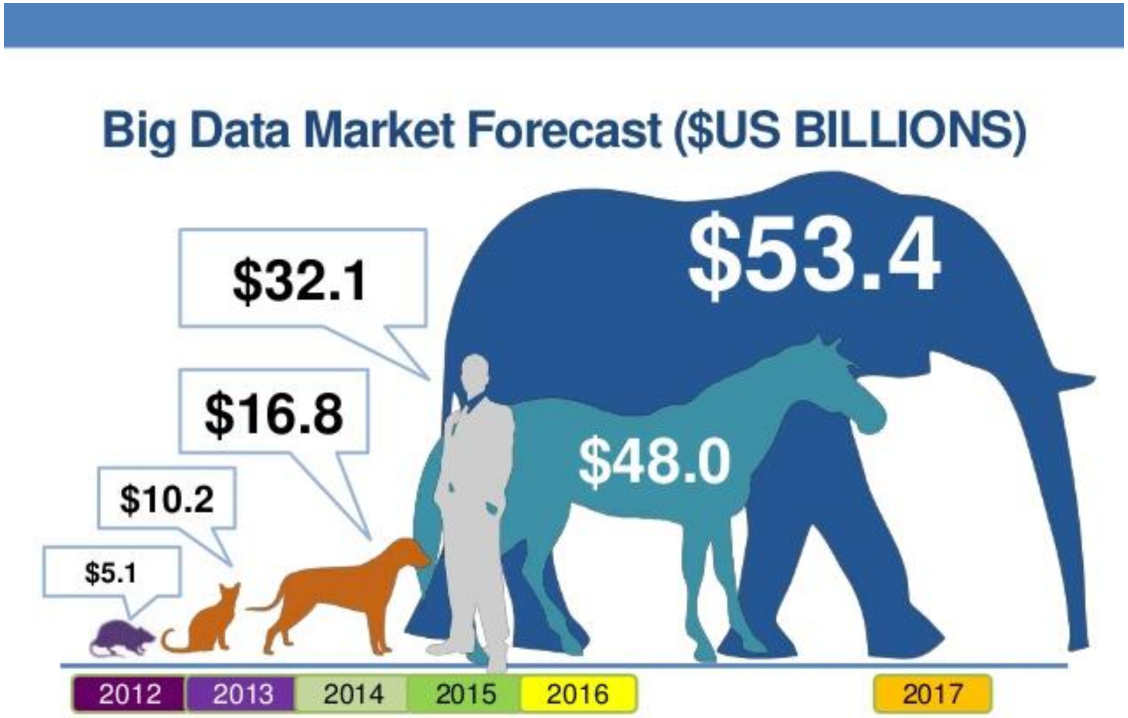
- ➢Volume
- ➢Velocity
- ➢Variety
- ➢**Veracity**



Big Data

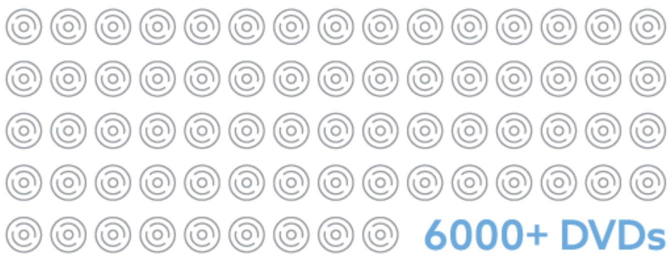
The 5Vs

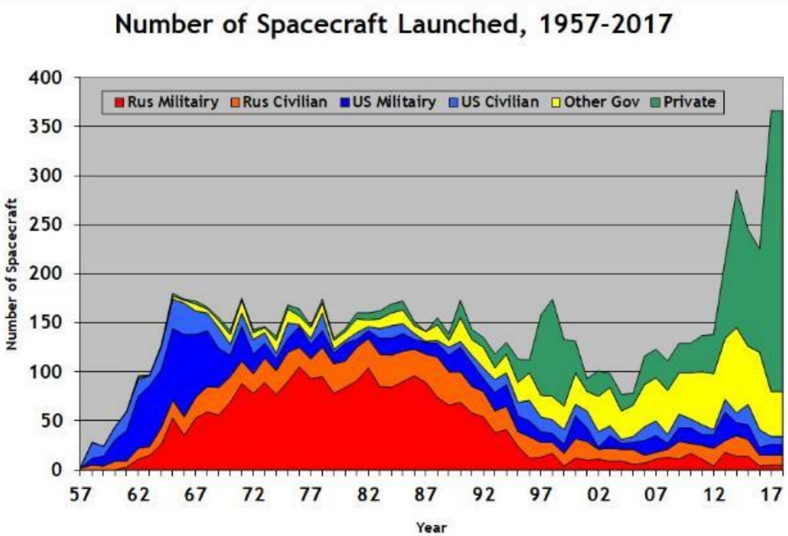
- ➢Volume
- ➢Velocity
- ➢Variety
- ➢Veracity
- ➢**Value**



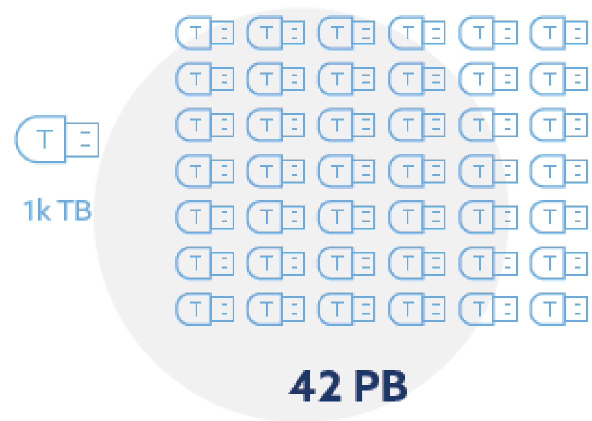
Earth Observation

The 6 Copernicus Sentinels are producing **over 12 TB** of high-quality full, free and open Earth Observation data **every day**, the equivalent of 6,000+ DVDs



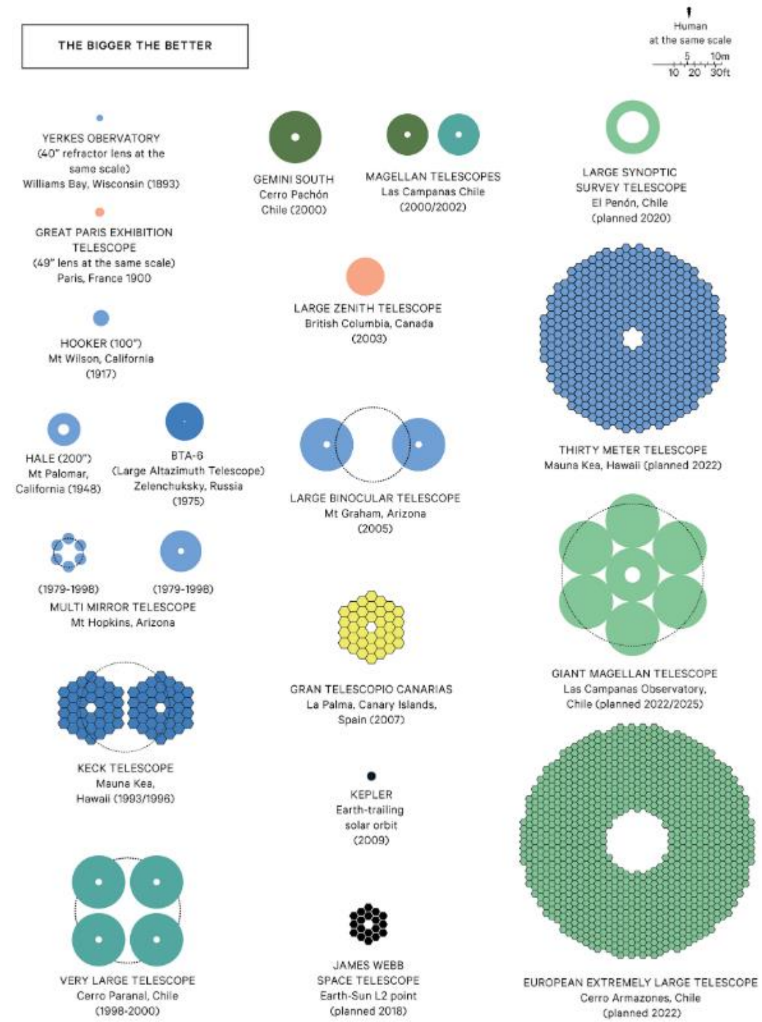


Volume of data distributed: **42 PB**

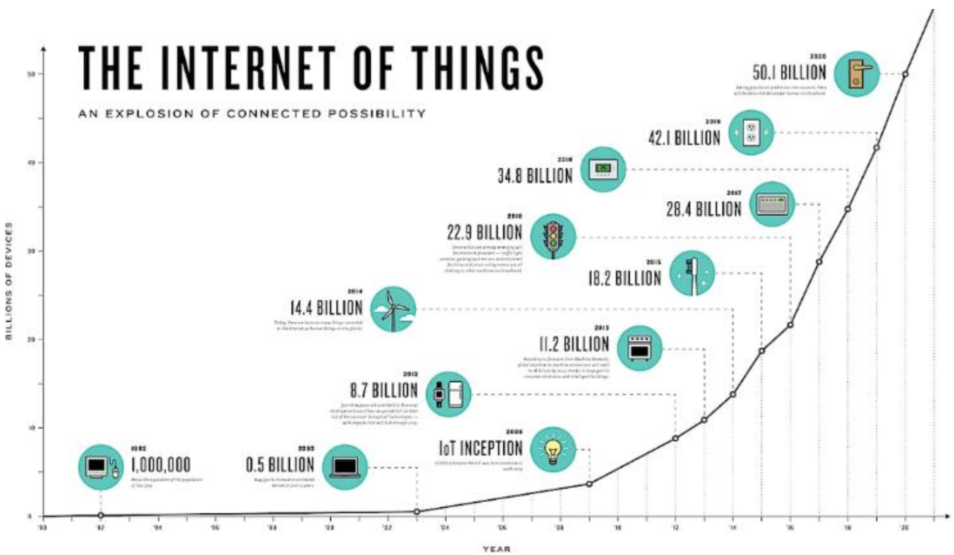


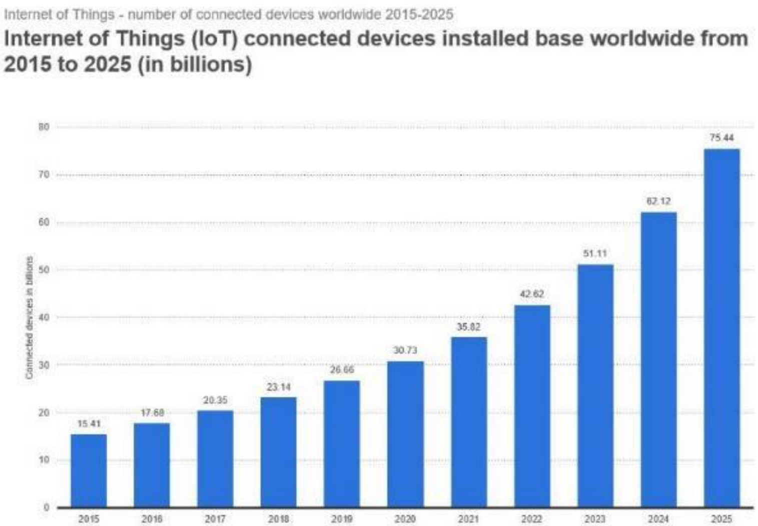
Astrophysics

Sky Survey Project	Volume	Velocity
Sloan Digital Sky Survey (SDSS)	50 TB	200 GB per day
Large Synoptic Survey Telescope (LSST)	~ 200 PB	10 TB per day
Square Kilometer Array (SKA)	~ 4.6 EB	150 TB per day

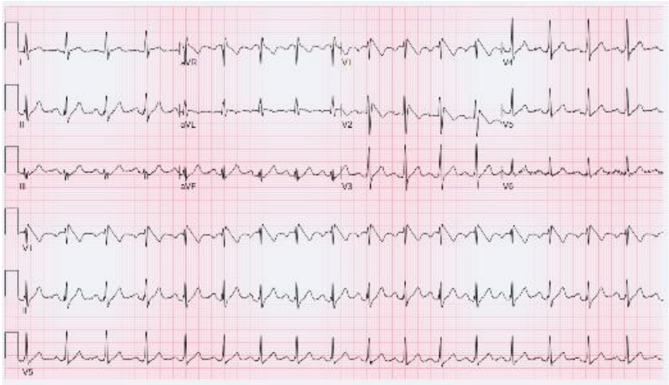


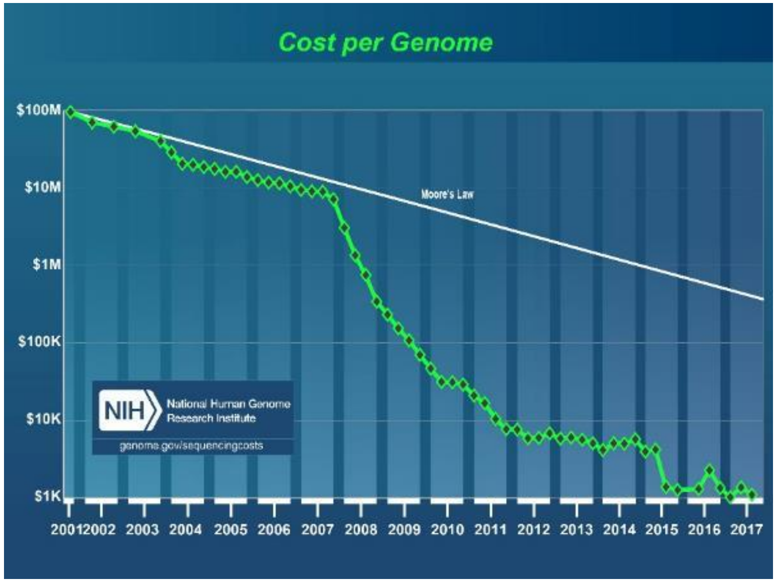
Internet-of-Things

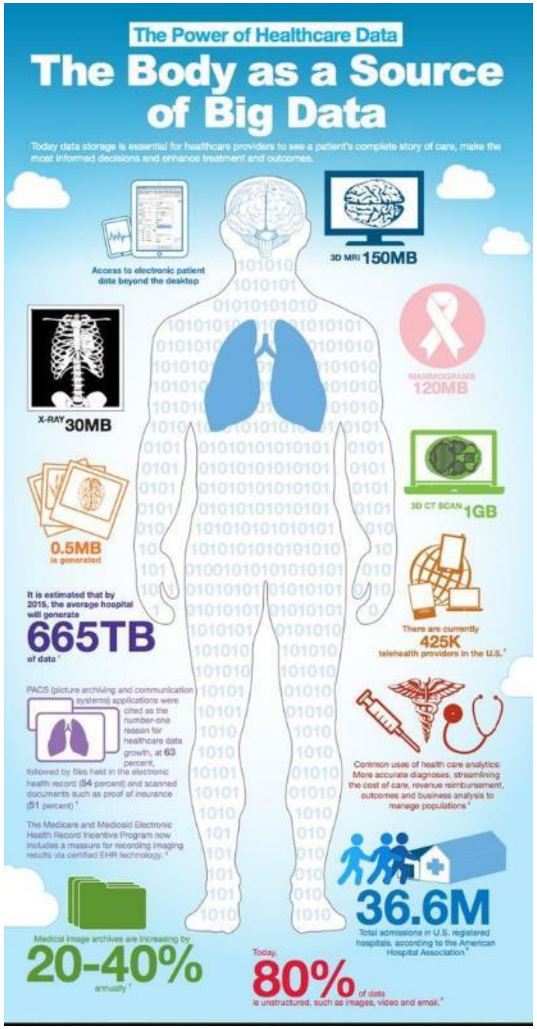




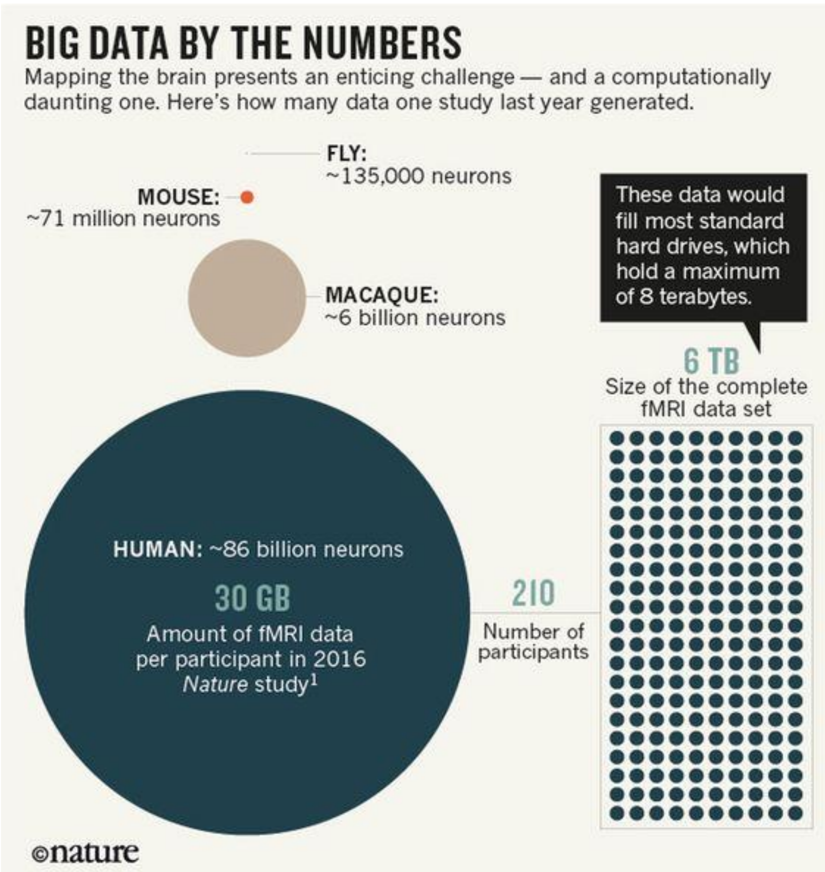
Biomedical signals



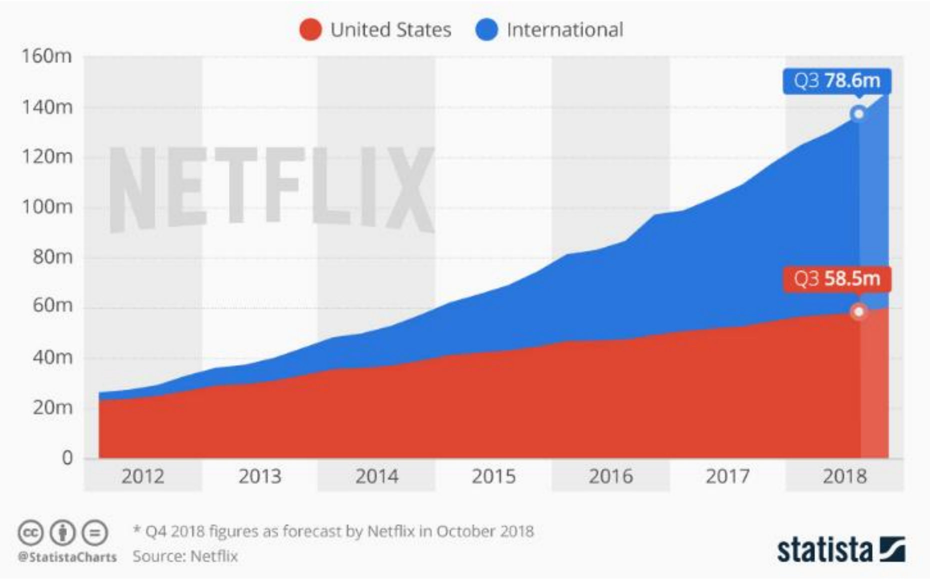


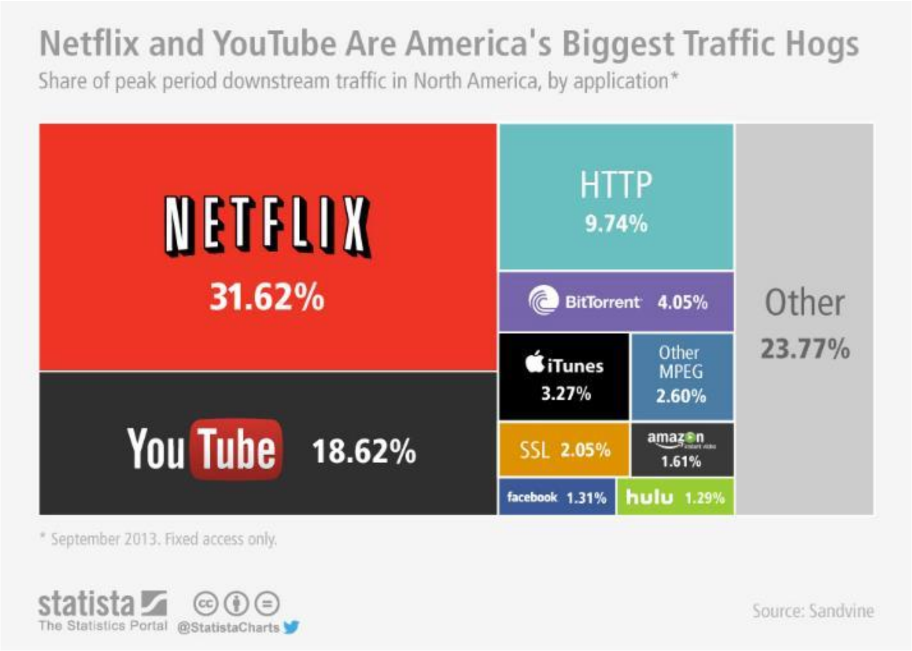


Neuroscience



Imaging technologies

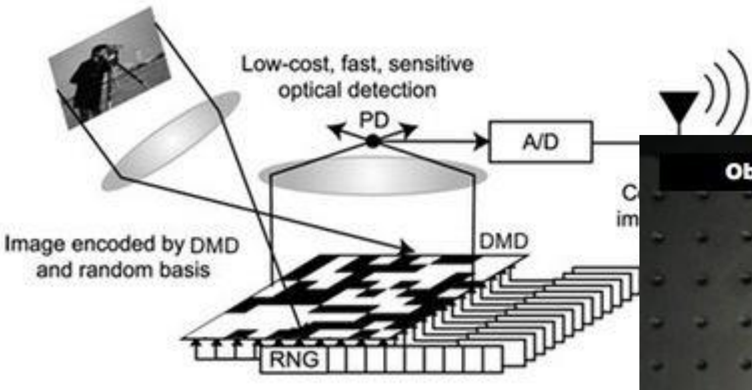


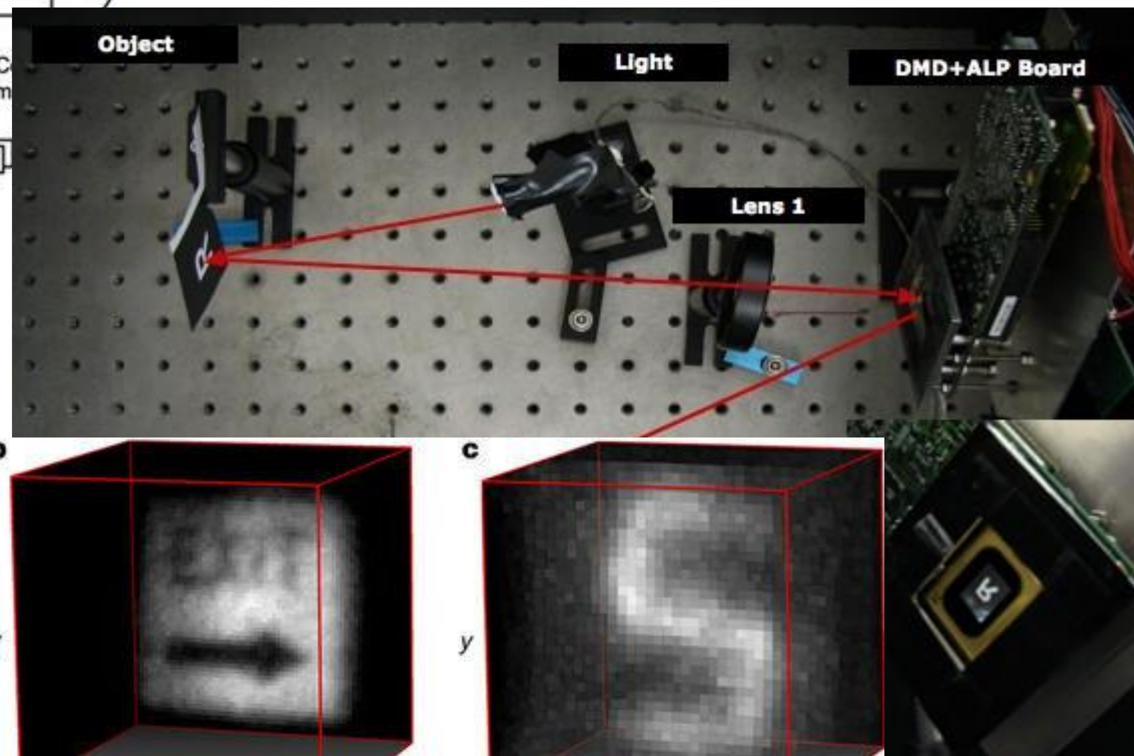


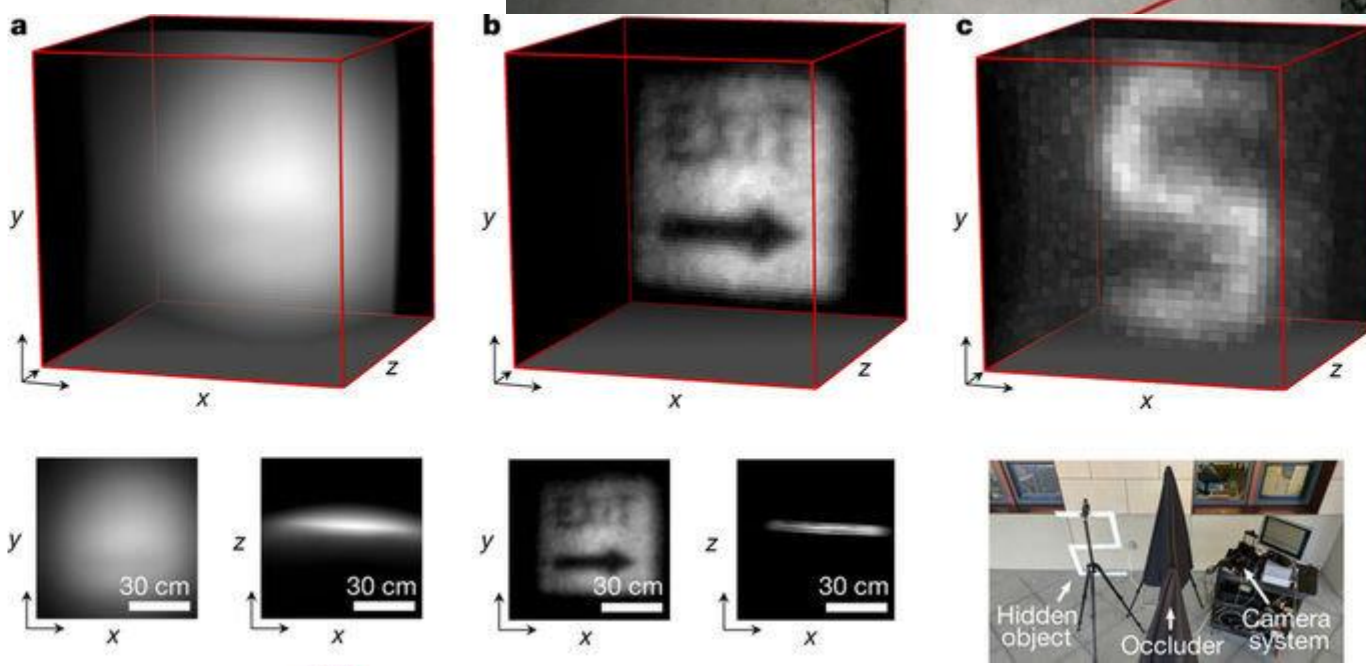

Massive Scale
Outliers, Missing Values
Challenges
Real-Time Constraints
Cloud Storage
Signal Processing and Learning for Big Data
Parallel, Decentralized
Time/Data Adaptive
Models and Optimization
Robust
Succinct, Sparse
Prediction, Forecasting
Dimensionality Reduction
Cleansing, Imputation
Tasks
Regression, Classification, Clustering

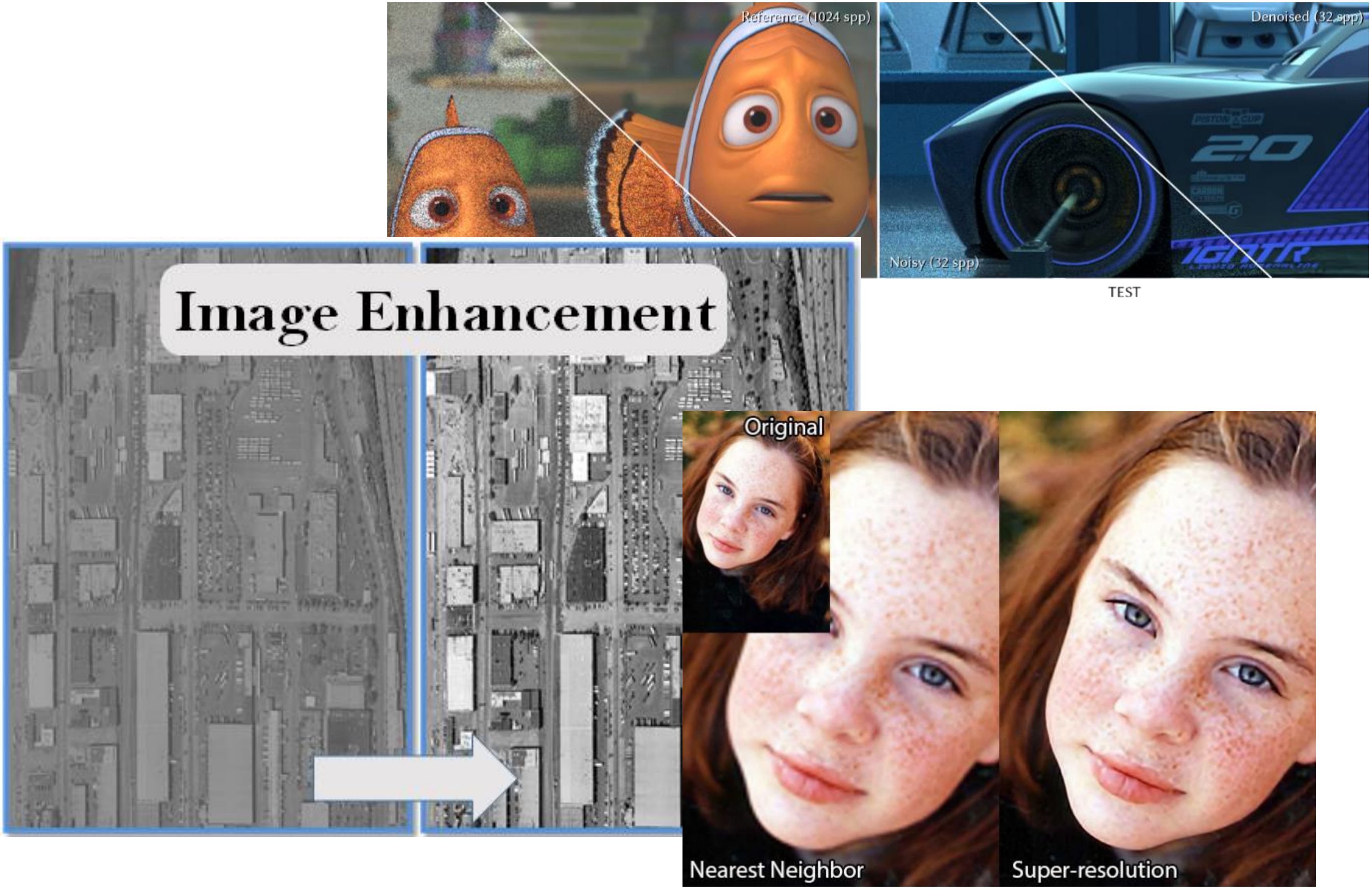
Fundamental Signal Processing

- Signal Sensing: Compressed Sensing
- Inverse problems: Signal Denoising, Enhancement
- Filtering
- Time-series modeling and prediction
- Modeling: dimensionality reduction


Low-cost, fast, sensitive optical detection
PD
A/D
Image encoded by DMD and random basis
DMD
RNG

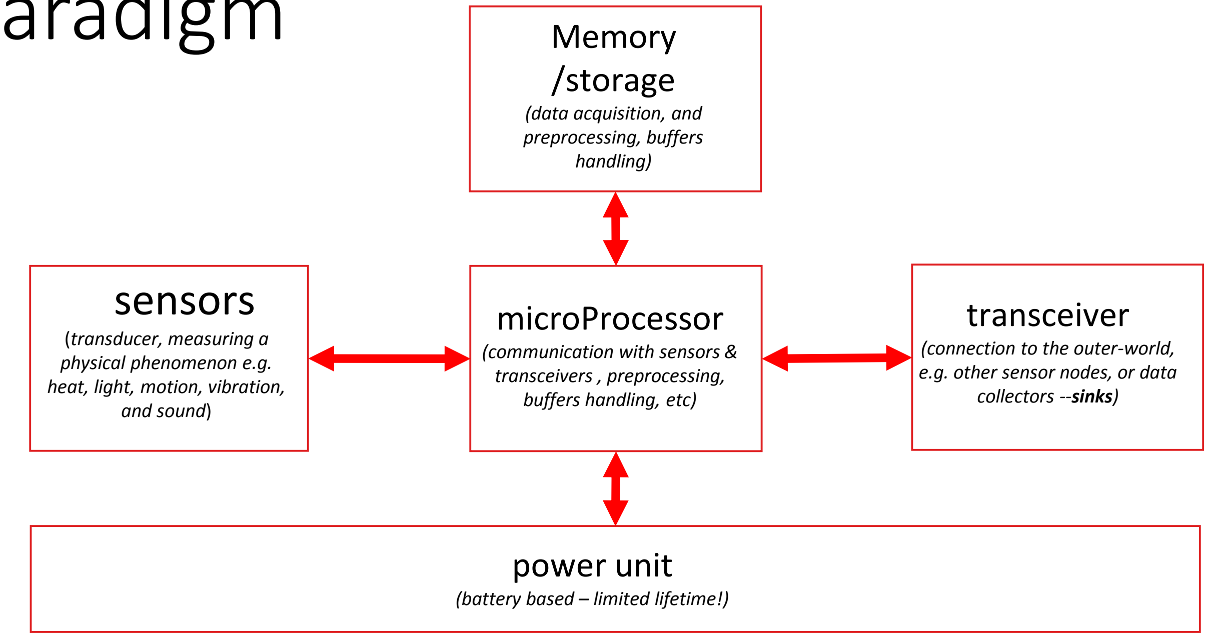

Object
Light
DMD+ALP Board
Lens 1


a
b
c
y
x
z
30 cm
Hidden object
Occluder
Camera system


Reference (1024 spp)
Denoised (32 spp)
Noisy (32 spp)
Image Enhancement
Original
Nearest Neighbor
Super-resolution

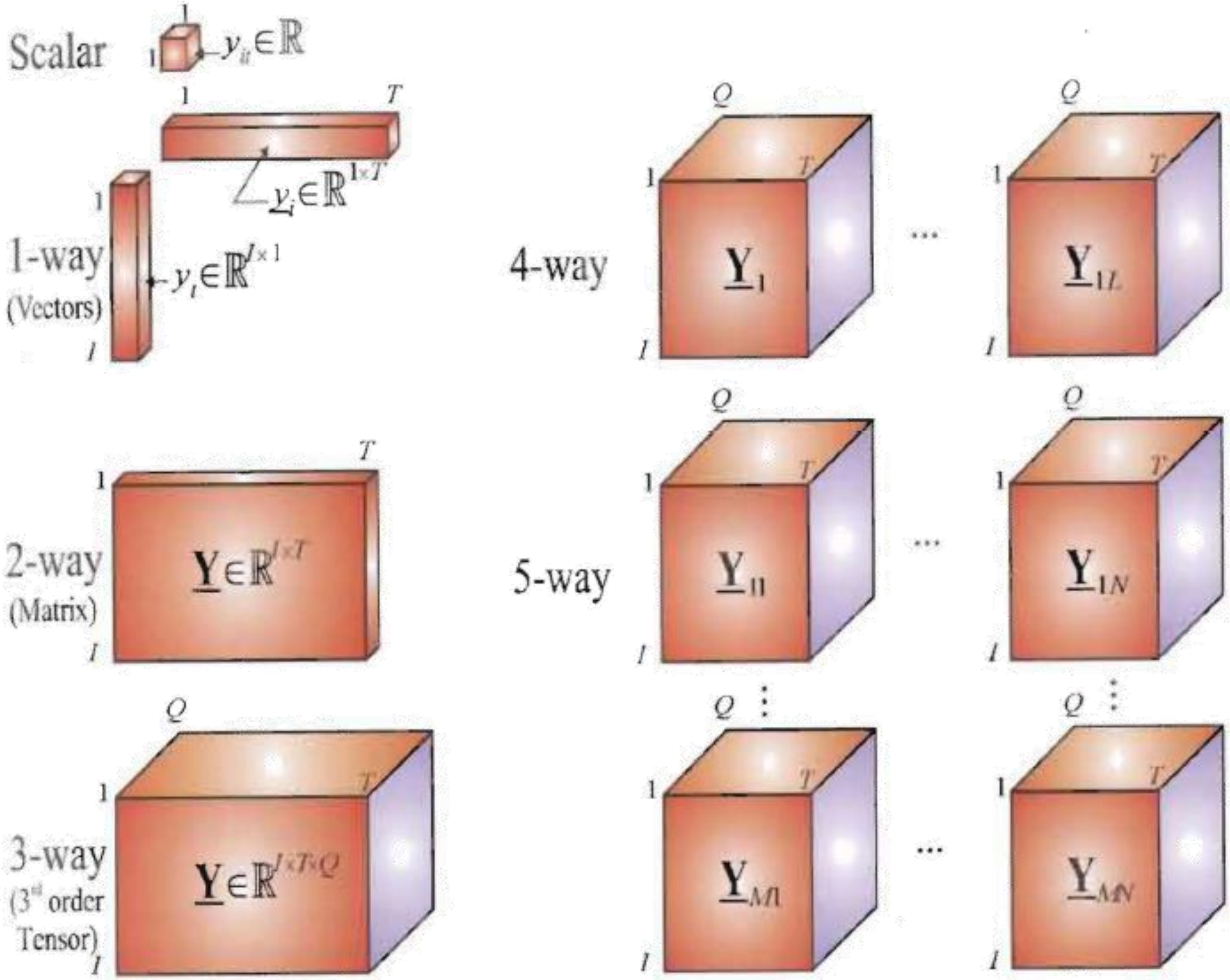
TEST

Sensing node paradigm



- Sensor Node
  - Basic unit in sensor network
  - Contains on-board sensors, processor, memory, transceiver, and power supply

Review of basic concepts



Vectors

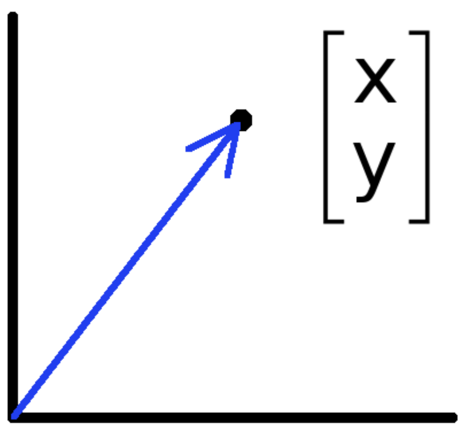
- A column vector $\mathbf{v} \in \mathbb{R}^{n \times 1}$ where $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$

- A row vector $\mathbf{v}^T \in \mathbb{R}^{1 \times n}$ where $\mathbf{v}^T = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$

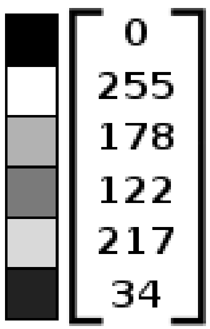
T denotes the transpose operation

Vectors

- Vectors can represent an offset in 2D or 3D space
- Points are just vectors from the origin



- Data (pixels, gradients at an image keypoint, etc) can also be treated as a vector
- Such vectors don't have a geometric interpretation, but calculations like "distance" can still have value

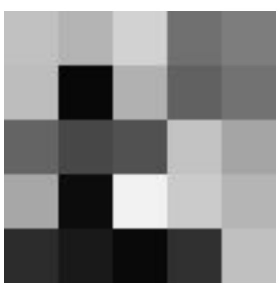


Matrix

- A matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is an array of numbers with size $m \downarrow$ by $n \rightarrow$, i.e. m rows and n columns.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & & & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

- If $m = n$, we say that $\mathbf{A}$ is square.



$$= \begin{bmatrix} 193 & 180 & 210 & 112 & 125 \\ 189 & 8 & 177 & 97 & 114 \\ 100 & 71 & 81 & 195 & 165 \\ 167 & 12 & 242 & 203 & 181 \\ 44 & 25 & 9 & 48 & 192 \end{bmatrix}$$

Basic Matrix Operations

- Addition $$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} a+1 & b+2 \\ c+3 & d+4 \end{bmatrix}$$

  - Can only add a matrix with matching dimensions, or a scalar.

- Scaling $$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + 7 = \begin{bmatrix} a+7 & b+7 \\ c+7 & d+7 \end{bmatrix}$$

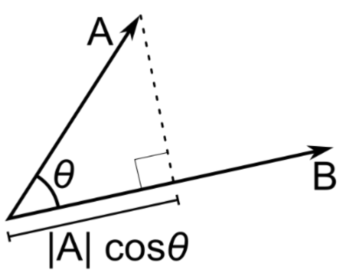
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \times 3 = \begin{bmatrix} 3a & 3b \\ 3c & 3d \end{bmatrix}$$

Matrix Operations

- Inner product (dot product) of vectors
  - Multiply corresponding entries of two vectors and add up the result
  - x·y is also |x||y|cos(*the angle between x and y*)

$$\mathbf{x}^T\mathbf{y} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{i=1}^{n} x_i y_i \qquad \text{(scalar)}$$

  - If B is a unit vector, then A·B gives the length of A which lies in the direction of B



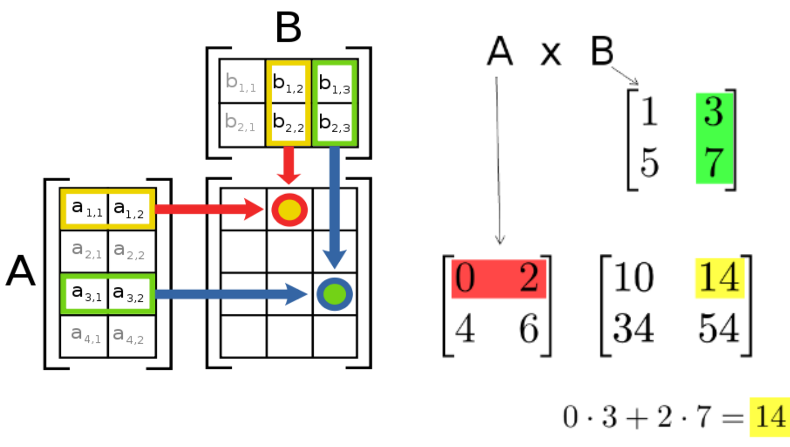
Matrix Operations

Matrix Multiplication

$$C = AB. \quad C_{i,j} = \sum_k A_{i,k} B_{k,j}.$$

- Properties

$$A(B + C) = AB + AC.$$
$$A(BC) = (AB)C.$$
$$(AB)^\top = B^\top A^\top$$
$$x^\top y = \left(x^\top y\right)^\top = y^\top x$$



- Powers
    - We can refer to the matrix product AA as A^2, and AAA as A^3, etc.
    - Only square matrices can be multiplied that way

Matrix Operations

- Transpose $\begin{bmatrix} 0 & 1 \\ 2 & 3 \\ 4 & 5 \end{bmatrix}^T = \begin{bmatrix} 0 & 2 & 4 \\ 1 & 3 & 5 \end{bmatrix}$ $(ABC)^T = C^T B^T A^T$

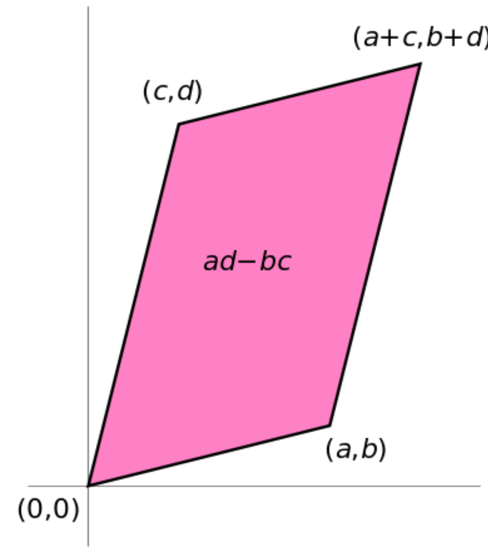
- Determinant
  - $\det(\mathbf{A})$ returns a scalar
  - Represents area of the parallelogram described by the vectors in the rows of the matrix
  - For $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\det(\mathbf{A}) = ad - bc$
  - Properties:

$$\det(\mathbf{AB}) = \det(\mathbf{BA})$$

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det(\mathbf{A})}$$

$$\det(\mathbf{A}^T) = \det(\mathbf{A})$$

$$\det(\mathbf{A}) = 0 \Leftrightarrow \mathbf{A} \text{ is singular}$$



Matrix Operations

- Trace $$\mathrm{tr}(\mathbf{A}) = \text{sum of diagonal elements}$$

$$\mathbf{tr}(\begin{bmatrix} 1 & 3 \\ 5 & 7 \end{bmatrix}) = 1 + 7 = 8$$

  - Invariant to a lot of transformations
  - Properties: $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$

$$\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$$

Special matrices

- Identity matrix **I** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

- Diagonal matrix $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 2.5 \end{bmatrix}$

- Symmetric matrix $\mathbf{A}^T = \mathbf{A}$ $\begin{bmatrix} 1 & 2 & 5 \\ 2 & 1 & 7 \\ 5 & 7 & 1 \end{bmatrix}$

- Skew-symmetric matrix $\mathbf{A}^T = -\mathbf{A}$ $\begin{bmatrix} 1 & -2 & -5 \\ 2 & 1 & -7 \\ 5 & 7 & 1 \end{bmatrix}$

Matrix Inverse

- Given a matrix **A**, its inverse $\mathbf{A}^{-1}$ is a matrix such that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

- Inverse does not always exist. If $\mathbf{A}^{-1}$ exists, **A** is *invertible* or *non-singular*. Otherwise, it's *singular*.
- For matrices that are invertible

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}$$
$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$
$$\mathbf{A}^{-T} \triangleq (\mathbf{A}^{T})^{-1} = (\mathbf{A}^{-1})^{T}$$

System of linear equations

$$Ax = b$$

$A \in \mathbb{R}^{m \times n}$ is a known matrix

$b \in \mathbb{R}^{m}$ is a known vector

$x \in \mathbb{R}^{n}$ is a vector of unknown variables

$$A_{1,:}x = b_1$$

$$A_{2,:}x = b_2$$

$$\dots$$

$$A_{m,:}x = b_m$$

$$A_{1,1}x_1 + A_{1,2}x_2 + \cdots + A_{1,n}x_n = b_1$$

Solution of system

- Inverse of a matrix $A^{-1}A = I_n$

$$Ax = b$$
$$A^{-1}Ax = A^{-1}b$$
$$I_n x = A^{-1}b$$
$$x = A^{-1}b.$$

- Solution of systems of linear equ
- Provided A^{-1} exists

- If both *x* and *y* are solutions then $z = \alpha x + (1 - \alpha)y$ is also a solution for any real α

Linear combinations

- Linear combination of some set of vectors *{v*(1)*, . . . , v*(*n*)*}* is given by multiplying each vector *v*(*i*) by a corresponding scalar coefficient and adding the results: $\sum_i c_i \boldsymbol{v}^{(i)}$

- The **span** of a set of vectors is the set of all points obtainable by linear combination of the original vectors.

Norms

$$||\boldsymbol{x}||_p = \left(\sum_i |x_i|^p\right)^{\frac{1}{p}} \quad p \in \mathbb{R}, p \geq 1$$

a norm is any function f that satisfies

- $f(\boldsymbol{x}) = 0 \Rightarrow \boldsymbol{x} = \boldsymbol{0}$
- $f(\boldsymbol{x}+\boldsymbol{y}) \leq f(\boldsymbol{x}) + f(\boldsymbol{y})$ (the **triangle inequality**)
- $\forall \alpha \in \mathbb{R}, f(\alpha \boldsymbol{x}) = |\alpha| f(\boldsymbol{x})$

Norms

- L_2 norm, also known as Euclidean norm

$$\|x\|_2 = \sqrt{\sum_i x_i^2}$$

- L_{21} norm $||\boldsymbol{x}||_1 = \sum_i |x_i|$
- Infinite norm (or max norm) $||\boldsymbol{x}||_\infty = \max_i |x_i|$
- Frobenius norm (Matrix norm)

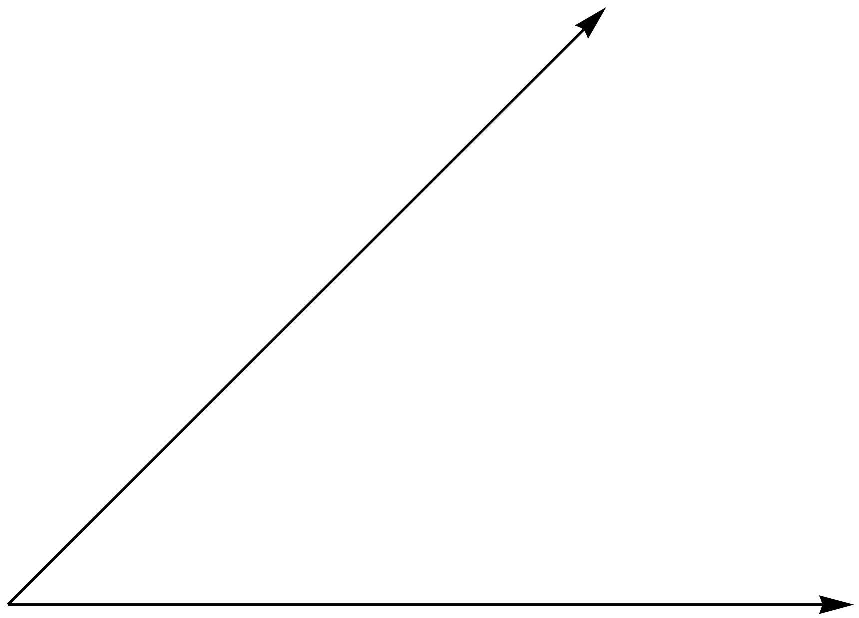
$$||A||_F = \sqrt{\sum_{i,j} A_{i,j}^2},$$

Linear independence

- Suppose we have a set of vectors $\mathbf{v_1, \ldots, v_n}$
- If we can express $\mathbf{v}_1$ as a linear combination of the other vectors $\mathbf{v}_2 \ldots \mathbf{v}_n$, then $\mathbf{v}_1$ is linearly *dependent* on the other vectors.
  - The direction $\mathbf{v}_1$ can be expressed as a combination of the directions $\mathbf{v}_2 \ldots \mathbf{v}_n$. (E.g. $\mathbf{v}_1 = .7\ \mathbf{v}_2 - .7\ \mathbf{v}_4$)
- If no vector is linearly dependent on the rest of the set, the set is linearly *independent*.
  - Common case: a set of vectors $\mathbf{v_1, \ldots, v_n}$ is always linearly independent if each vector is perpendicular to every other vector (and non-zero)

Linear independence

Linearly independent set



Not linearly independent