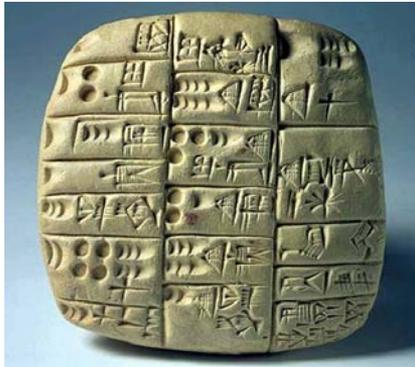


Data

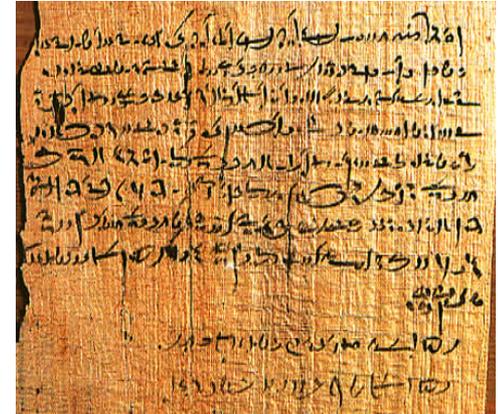
Data are representations (encodings) of information using symbols or signals (eg letters, numeric or alphanumeric codes etc) on a medium that they can be kept (eg stone, papyrus, paper, magnetic medium).



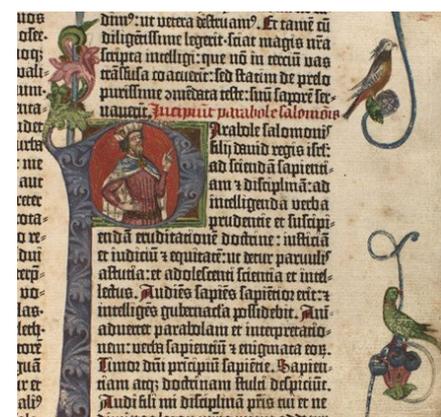
cuneiform writing



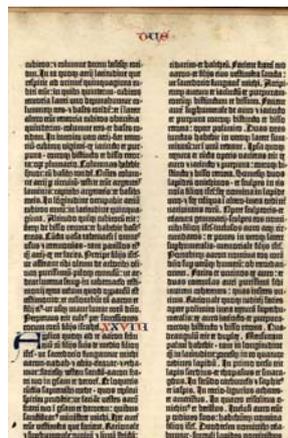
hieroglyphs writing



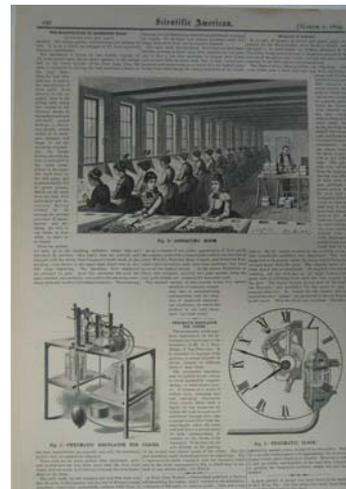
demotic writing



incunabula



Gutenberg print



Old scientific journal



digital scientific journal

Information and knowledge

- One may define **information** as a form of communication (as a service that puts someone informed with a situation).
- Information is a form of communication likely to generate **knowledge**.



hieroglyphs writing



Rosetta Stone (the first translation service 😊)
+ Jean-Francois Champollion



Khufu (Kheops) cartouche

Information theory

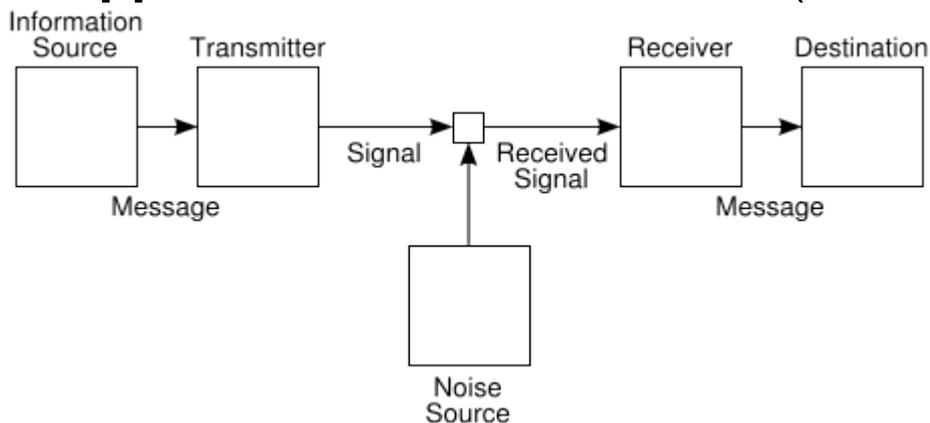
- Claude E. Shannon , "A Mathematical Theory of Communication", Bell System Technical Journal, vol. 27, pp. 379-423, 623-656, July, October, 1948

Information theory is a branch of applied mathematics and electrical engineering involving **the quantification of information**

Shannon's article laid out the basic elements of communication:

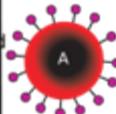
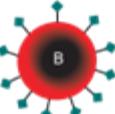
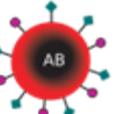
- An **information source** that produces a message
- A **transmitter** that operates on the message to create a signal which can be sent through a channel
- A **channel**, which is the medium over which the signal, carrying the information that composes the message, is sent
- A **receiver**, which transforms the signal back into the message intended for delivery
- A **destination**, which can be a person or a machine, for whom or which the message is intended (and “understand” it)

Application: communications (GSM, Wi-Fi etc), various sciences



Shannon's definition of information

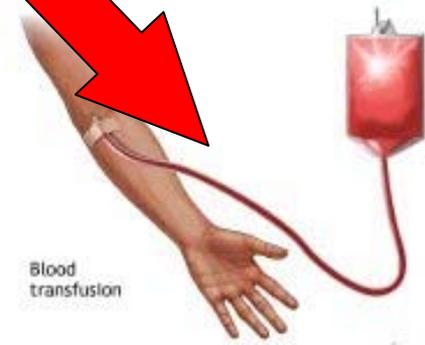
- Information is equivalent to raising the indeterminacy

	Group A	Group B	Group AB	Group O
Red blood cell type				
Antibodies present	 Anti-B	 Anti-A	None	 Anti-A and Anti-B
Antigens present	A antigen	B antigen	A and B antigens	None

e.g. Blood types

Recipient's blood			Reactions with donor's red blood cells			
ABO antigens	ABO antibodies	ABO blood type	Donor type O cells	Donor type A cells	Donor type B cells	Donor type AB cells
None	Anti-A Anti-B	O				
A	Anti-B	A				
B	Anti-A	B				
A & B	None	AB				

 Compatible
  Not compatible



D(ata) → Information → Knowledge

Data



Information



Knowledge!!!

Univ. Hosp. Tg. Mures
John Smith, 20 years

ESR

50 (mm/hr)

Dr. abcd

Univ. Hosp. Tg. Mures
John Smith, 20 years

ESR

50 (mm/hr)

Dr. abcd

Univ. Hosp. Tg. Mures
John Smith, 20 years

ESR

50 (mm/hr)

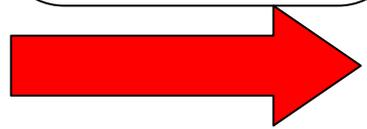
Dr. abcd



YES!!! I found the Smith's laboratory data that I lost !!!
This paper present ESR = erythrocyte sedimentation rate
It will be **useful** for dr. B to give a diagnosis and some treatment



Poor Mr. Smith, he surely has a **chronic inflammation**, the value of ESR is very high !
I shall prescribe antiinflammatory drugs !!!



D(ata) → Information → Knowledge

Data

computer

Information

Humans !!!

Knowledge!!!

Univ. Hosp. Tg. Mures
John Smith, 20 years

ESD

50

Dr. abcd

Univ. Hosp. Tg. Mures
John Smith, 20 years

ESD

50

Dr. abcd

Univ. Hosp. Tg. Mures
John Smith, 20 years

ESD

50

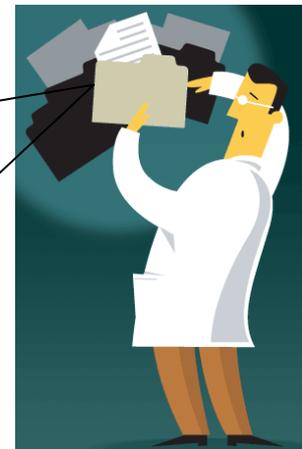
Dr. abcd

???

YES!!! I found the
Smith's laboratory data
that I lost !!!
This paper present
ESD = erythrocyte
sedimentation rate
It will be **useful** for dr.
B to give a diagnosis



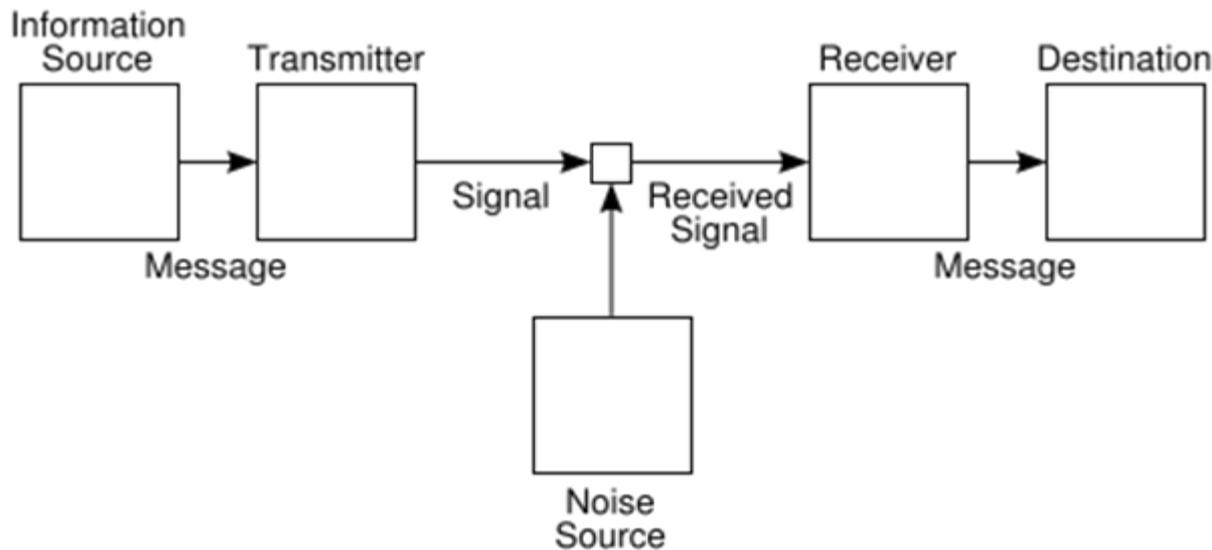
Poor Mr. Smith, he
surely has a
**chronic
inflammation**, the
value of ESD is very
high !
**I shall prescribe
antiinflammatory
drugs !!!**



Search ID: grin510

Information theory – some conclusions

- The **characteristic feature** separating **data from information** is **the usefulness**, which implies there is a receiver of data to understand his meaning, in order to generate information.
- **Transforming data into informations** is one of the most important functions of data processing task.
- Information theory is a branch of applied mathematics and electrical engineering involving **the quantification of information**.



The Quantification of Information (1).

- In the “language” of probability theory, the **amount of information** provided by an event is **even greater** as the **probability of occurrence of that event is less**.

Experiment 1: 100 eggs from the same farm, **99** eggs were broken, **1** egg no

Experiment 2: 10000 eggs from the same farm, **9999** eggs were broken, **1** egg no



This new separator cracks the shells of 3,600 eggs an hour, and holds them apart till the contents drain into the cups, where each is inspected. (Int. News)

Probability of Experiment 1:

1/100

Probability of Experiment 2:

1/10000

Based on probability, which egg is stronger ???

The Quantification of Information (2).

- The simplest case is represented by the statement that admits only two answers alleged as **equiprobable (yes or no, present or absent, head/tail, 0/1 etc..)**

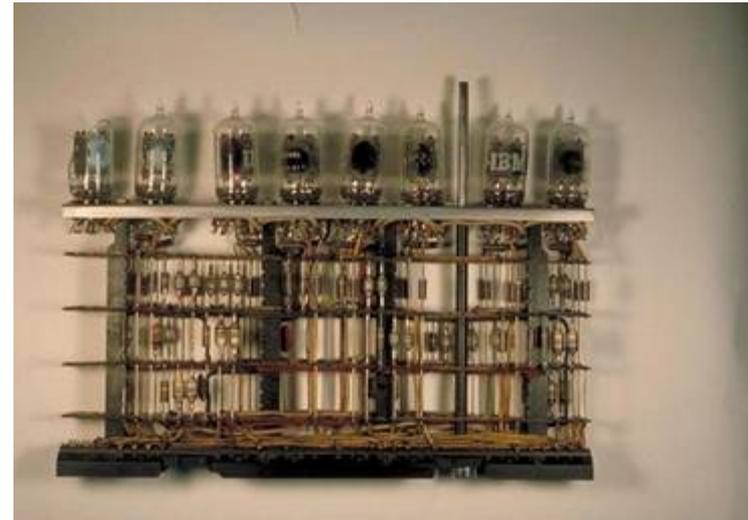
Real life experiment : throwing one new coin 100 times

**We expect the probability of experiment:
50 head/50 tail**



Informatics – IBM vacuum tubes module:

Powered vacuum tube =1
Vacuum tube without electric power=0



The Quantity of Information (1)

- **to measure something we need an unit of measurement !!!**
- But, unlike the computers (which use binary system), we think in decimal system

Unit of measurement in decimal system:

•E.g. for **mass = a measure of the amount of matter in the body** (while weight is a measure of the force on the object caused by a gravitational field).

The (old) unit : 1 gramme (g) = the absolute weight of a volume of pure water equal to the cube of the hundredth part of a metre, and at the temperature of melting ice (later 4 °C).

Bigger units :

1kg (kilogramme) = 1000 g = 10^3 g
(from greek “chilioi” means 1000)

1 tonne = 1000 kg = 10^3 kg

Unit of measurement for information (used in binary system):

From the information theory viewpoint, **the unit of quantity of information** called **"bit" (binary unit)** is defined as **the quantity of information obtained by producing an event from a experiment which may have only two outcomes equally probable (equiprobable).**

In more technical terms, a **bit** is the basic unit of information in computing and telecommunications; **it is the maximum amount of information that can be stored by a device or other physical system that can normally exist in only two distinct states.** These may be the two stable positions of an electrical switch, two distinct voltage or current levels allowed by a circuit, two distinct levels of light intensity, two directions of magnetization or polarization, etc.

The Quantity of Information (2)

- to measure something we need an unit of measurement !!!
- But, unlike the computers (which use binary system), we think in decimal system

Unit of measurement in decimal system:

•E.g. for **mass = a measure of the amount of matter in the body** (while weight is a measure of the force on the object caused by a gravitational field).

The (old) unit : 1 gramme (g)

Bigger units:

1kg (kilogramme) = 1000 g = 10^3 g
(from greek “chilioi” means 1000) =
= the base of decimal system at power of 3

1 tonne = 1000 kg = 10^3 kg

Unit of measurement for information (used in binary system):

The unit : 1 bit (b)

Bigger units :

1 Byte (B) = 2^3 bits (b)=

the base of binary system at power of 3=
= 8 bits (b)

1 Kilobyte (KB) = the value of the pure power of two closer to 1000 (kilo)= 2^{10}
B=1024 B

1MB (megabyte)= 2^{10} KB=1024 KB

1GB (gigabyte)= 2^{10} MB=1024 MB

1TB (terabyte)= 2^{10} GB=1024 GB

1PB (petabyte)= 2^{10} TB=1024 TB

Other units:

1 Kb (kilobit), Mb, Gb, Tb etc, where

1 KB (kilobyte) = 2^3 Kb (kilobits) etc.