INTRODUCTION TO CRYPTOGRAPHY

Part 1 - Classical Cryptography

TABLE OF CONTENTS

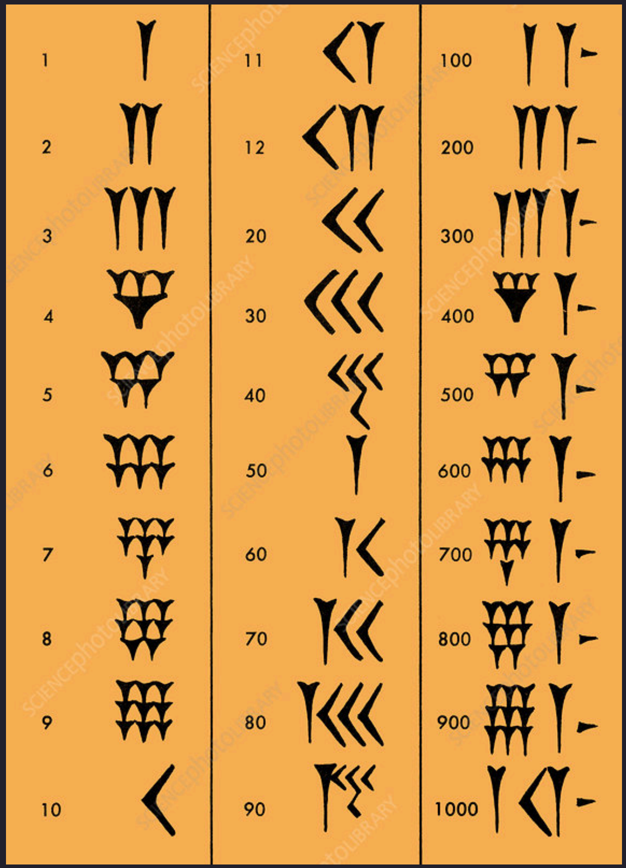
WHY CRYPTOGRAPHY?

Human societies throughout the years have developed **informational systems** ever more complex.

Numbers, for example, were introduced around 6.000 years ago within the first societies, such as the **Sumerian**.

Their objective?

Bureaucracy. To keep track of various types of items such as food, people, weapons, etc.



In certain contexts having access to information can signify the difference between **life** and **death**.

THE TRIAL OF MARY STUART

The following example has been taken from the book

The Code Book

The Science of Secrecy from Ancient Egypt to Quantum Cryptography

The Process of Mary Stuart (1/10)

15th of october, 1586.

Fotheringhay's castle, central england.

Mary Stuart, Queen of Scots, is under process for treason against Elizabeth I, Queen of England and Ireland.

The Process of Mary Stuart (2/10)

Sir Francis Walsingham, Secretary of State of the Kingdom of England, was searching for evidence of treason.

The Process of Mary Stuart (3/10)

Elizabeth wanted to make sure that Mary Stuart commited treason since

- Mary was Queen of Scotland
- It could set a dangerous precedent
- Mary is the cousin of Elizabeth

The Process of Mary Stuart (4/10)

Mary is not afraid, as she knows that all messages sent between her and her accomplices were **encrypted**.

The Process of Mary Stuart (5/10)

Francis Walsingham, already aware of this, called **Thomas Phelippes**, a linguist and one of the best decipherer of England.

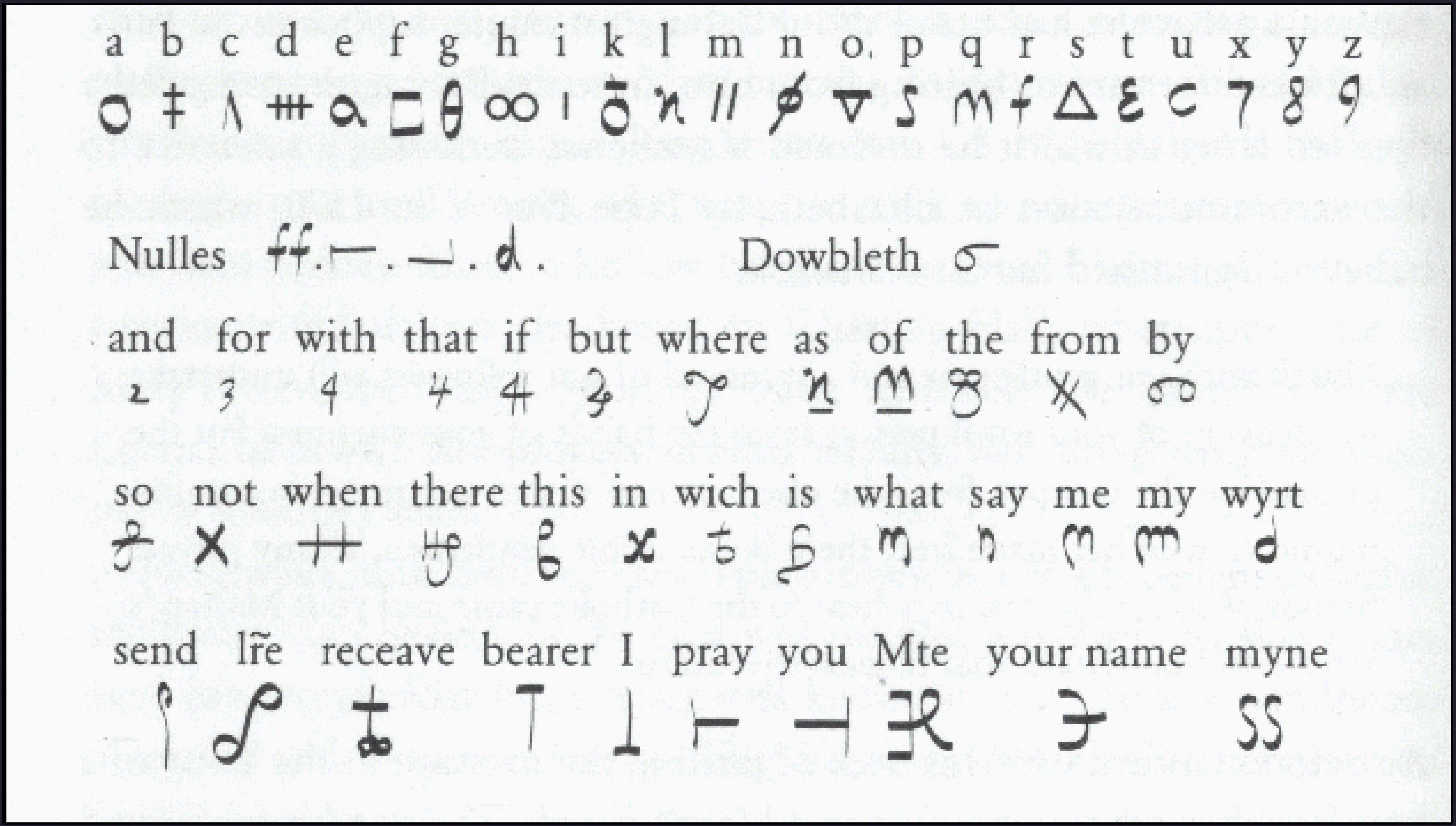
The Process of Mary Stuart (6/10)

The encryption method adopted by Mary Stuart is known as **nomenclator**.

The Process of Mary Stuart (7/10)

The idea was to use 23 new custom symbols to substitute to the normal letters of the alphabet, and 35 other symbols that represented entire words or phrases.

The Process of Mary Stuart (8/10)



The Process of Mary Stuart (9/10)

Without her knowing about it, all messages sent and received by Mary Stuart were captured and deciphered by **Walsingham**.

At the end he managed to trick Mary into writing a list with the names of all her accomplices.

The Process of Mary Stuart (10/10)

Mary Stuart was beheaded on the 8th of feburary 1587.

WHY DO WE NEED CRYPTOGRAPHY?

We need cryptography because **information** have a direct and irreversible effect on reality.

Cryptography offers tools, techniques and technologies that allow us to have more control in the way in which information can influence our life.

Many companies of today offer services related to cryptography

Let us then understand how cryptography has evolved over time.

CLASSICAL CRYPTOGRAPHY

Let's start with some **etymology** from greek

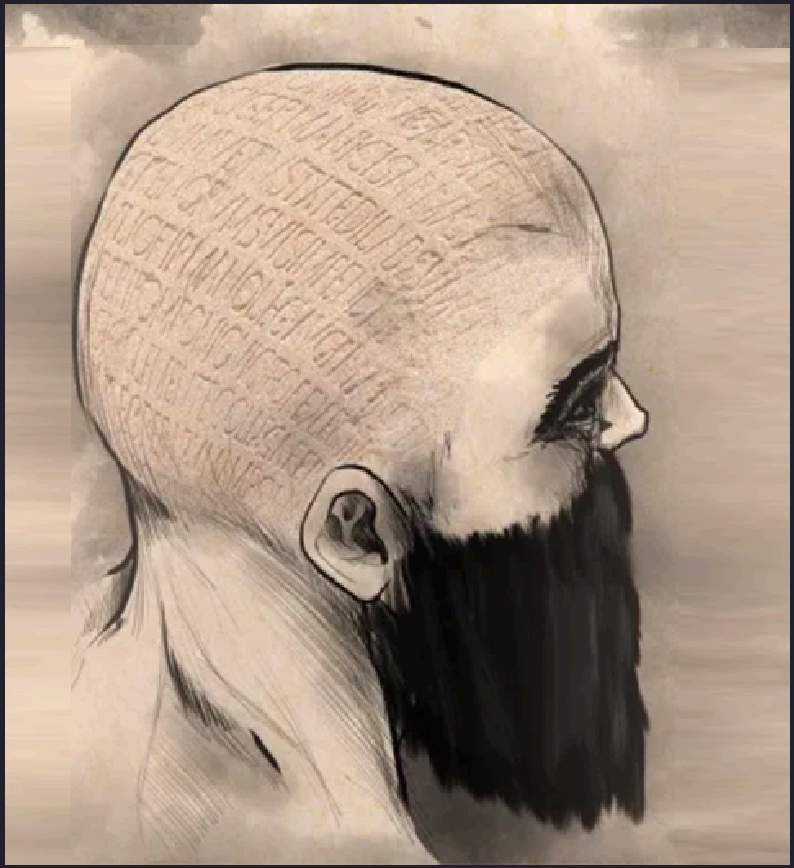
- **steganography**:
  - steganós → "covered"
  - graphía → "writing"
- **cryptography**:
  - kryptós → "secret"
  - graphía → "writing"

More technically,

The objective of **steganography** is to **hide the presence of the message**.
The objective of **cryptography** instead is to **hide the meaning of the message**.

STEGANOGRAPHY

Herodotus, one of the first writer of History, tells the practice used during the **Persian Wars** (+2500 years ago), of cutting the head of the couriers in order to write messages on their head and waiting for the hair to grow back to hide the messages during transportation.



CRYPTOGRAPHY

Caesar Cipher (1/5)

The idea is to hide the meaning of the message by **shifting** the letters of the alphabet by a fixed quantity $c = 3$.

Caesar Cipher (2/5)

We start from a **plaintext alphabet**

```
ABCDEFGHIJKLMNOPQRSTUVWXYZ
```

Caesar Cipher (2/5)

By applying the **shift**, we obtain a **ciphertext alphabet**

```
ABCDEFGHIJKLMNOPQRSTUVWXYZ
↓
DEFGHIJKLMNOPQRSTUVWXYZABC
```

Caesar Cipher (3/5)

Given a single letter, we obtan the associated ciphertext letter using the ciphertext alphabet.

$$\mathtt{A} \longrightarrow \mathtt{A} + 3 = \mathtt{D}$$

Caesar Cipher (4/5)

If we have many letters, we can encrypt them one at a time.

```
HELLO WORLD
     ↓
KHOOR ZRUOG
```

Caesar Cipher (5/5)

```
#!/usr/bin/env python3

def main():
    shift_value = 3
    cipher = Caesar(shift=shift_value)
    plaintext = "HELLO WORLD"
    ciphertext = cipher.encrypt(plaintext)
    print(f"[c={shift_value}] '{plaintext}' -> '{ciphertext}'")
```

./code/caesar.py

TRANSPOSITION AND SUBSTITUTION

Caesar cipher is a **mono-alphabetic** cipher based on **substitution**.

Classical ciphers that work on natural languages can use two main techniques

Transposition: the letters of a message swap place
Substituion: the letters of a message are substituted by other letters

VIGENÈRE CIPHER

Vigenère Cipher (1586) is a generalization of **Caesar Cipher**.

Instead of having one ciphertext alphabet, we have many ciphertext alphabets, which are used in an alternative fashion.

Esempio (1/4)

Suppose we have three ciphertext alphabets

```
ABCDEFGHIJKLMNOPQRSTUVWXYZ
            ↓
ABCDEFGHIJKLMNOPQRSTUVWXYZ
DEFGHIJKLMNOPQRSTUVWXYZABC
CDEFGHIJKLMNOPQRSTUVWXYZAB
```

Esempio (2/4)

To encrypt a sequence of letters, we select in a sequential way the various ciphertext alphabets.

After we have finished all the alphabets, we start again with the first one.

Esempio (3/4)

```
HELLO WORLD
```

$\downarrow$

```
HHNLR WRTLG
```

Esempio (4/4)

Instead of describing the ciphertext alphabets in full, we can simply write the first letter of each alphabet.

$$\texttt{ABCDEFGHIJKLMNOPQRSTUVWXYZ} \to \texttt{A}$$
$$\texttt{DEFGHIJKLMNOPQRSTUVWXYZABC} \to \texttt{D}$$
$$\texttt{CDEFGHIJKLMNOPQRSTUVWXYZAB} \to \texttt{C}$$

The **encryption key** is $\texttt{ADC}$.

```python
def main():
    key = "ADC"
    cipher = Vigenere(key)
    plaintext = "HELLO WORLD"
    ciphertext = cipher.encrypt(plaintext)
    print(f"[key='{key}'] '{plaintext}' -> '{ciphertext}'")
```

./code/vigenere.py

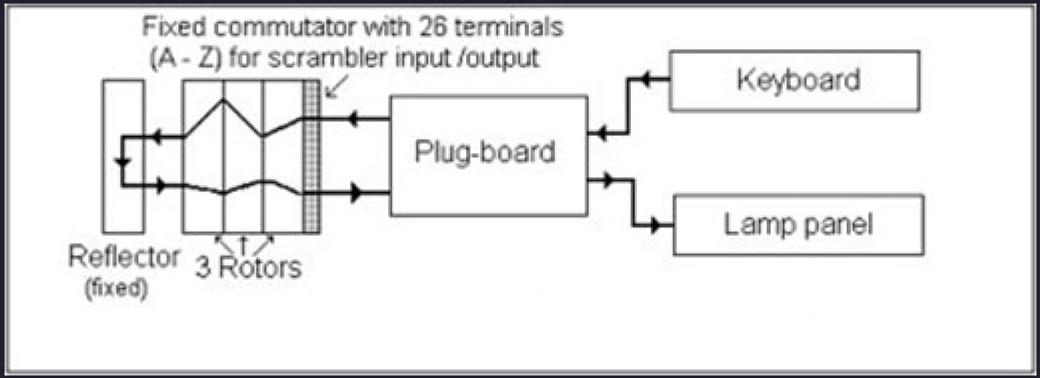
ENIGMA MACHINE

The Enigma machine (1923) is an **electro-mechanical** device that implements an extremely complex substitution cipher.

It was acquired by the german army (1926), modified and used during Second World War to protect war communications.

The key idea of enigma is that anytime you press a key, the mechanical part of the machine activates, its rotors move, then the electrical circuit closes, and from the key pressed a specific lightbulb lights up.

That lightbulb represent the encrypted letter.



- key $\longrightarrow$ plaintext letter
- lightbulb $\longrightarrow$ ciphertext letter

For those interested, I have developed a simple emulator of the machine using the **C** programming language.

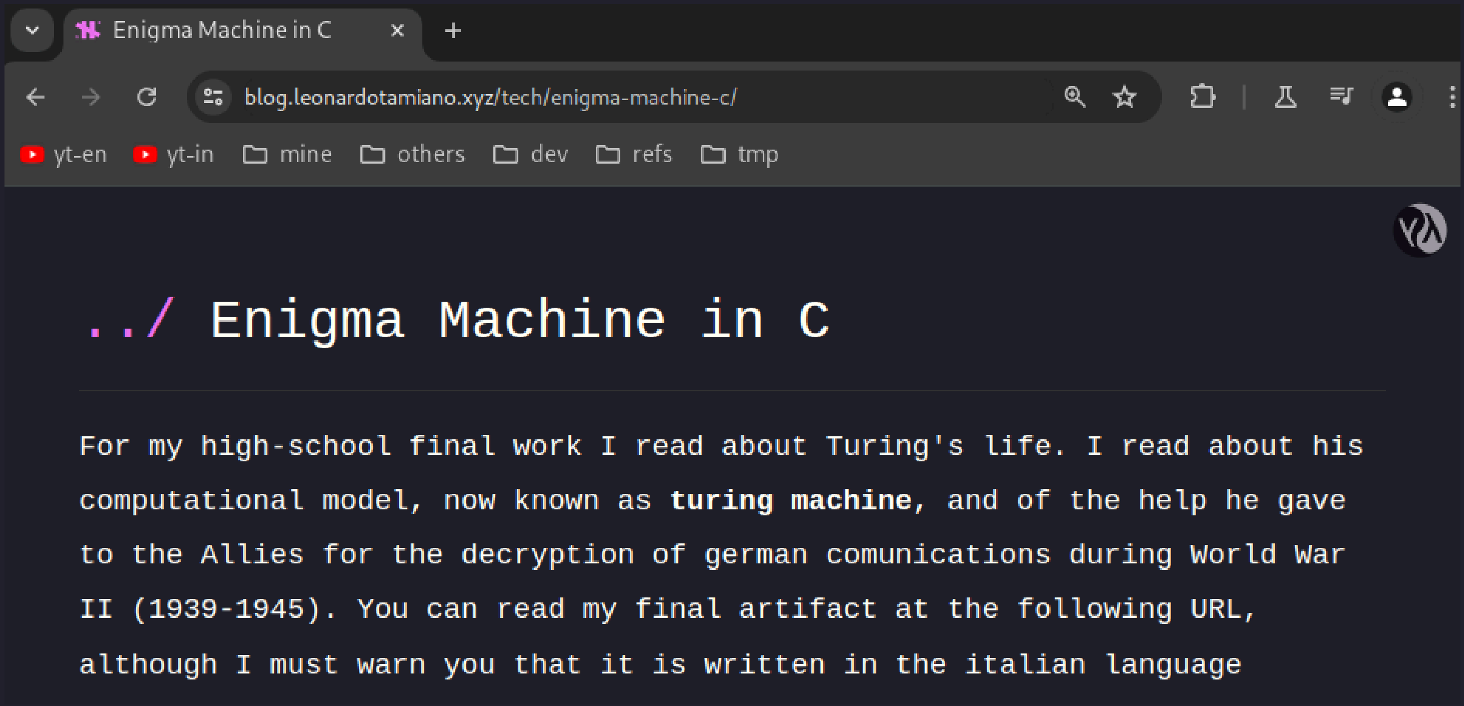
The project is available on github

https://github.com/LeonardoE95/enigma-machine

```
Enigma> info
Enigma> Current configuration...
        Rotors (from left to right): M3-II, M3-I, M3-III
               Position: 0, 0, 0
                   Ring: 0, 0, 0
        Reflector: M3-B
        Plugboard: 6 plugs
                   (A, M)
                   (F, I)
                   (N, V)
                   (P, S)
                   (T, U)
                   (W, Z)
Enigma> encrypt HELLO
MIJEN
```

https://github.com/LeonardoE95/enigma-machine

I've also written a blog post about it.



https://blog.leonardotamiano.xyz/tech/enigma-machine-c/

THE PROBLEMS OF CLASSICAL CRYPTOGRAPHY

The first ciphers, among which we find Caesar and Vigenère ciphers, suffer from a problem linked to the **key space**.

That is, the **key space** of these ciphers is, simply, too small. A modern computer can bruteforce all the possible keys in a short window of time.

In **Caesar's cipher**, we have 26 possible keys.

In **Vigeneré's cipher**, we have 26^n possible keys when used with a key of n characters.

The **Enigma Machine** has a much bigger key space

$$\approx 158.962.555.217.826.360.000$$

Still, it suffered from a different problem.

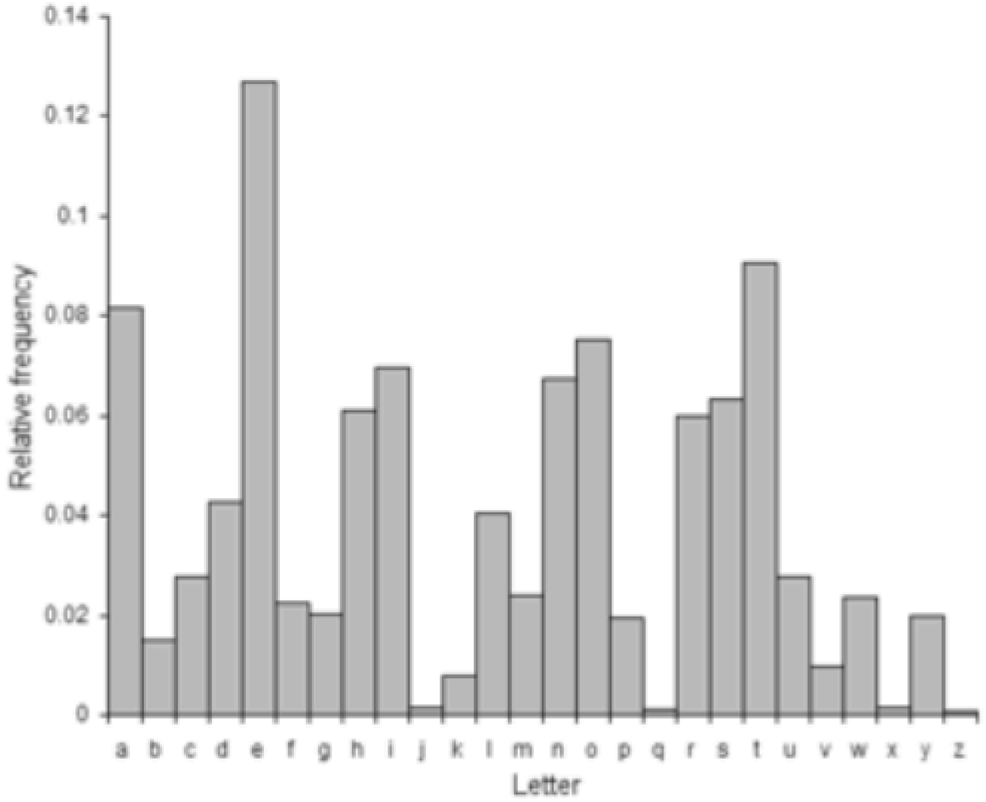
Beyond key space, classical ciphers suffer from a foundamental issue rooted in the fact that these ciphers work at the level of **single letters** of natural languages such as **italian**, **english**, etc.

The problem is that

The frequency of letters in natural languages is not uniform

Frequence of letters in the english language

E	11.1607%	56.88	M	3.0129%	15.36
A	8.4966%	43.31	H	3.0034%	15.31
R	7.5809%	38.64	G	2.4705%	12.59
I	7.5448%	38.45	B	2.0720%	10.56
O	7.1635%	36.51	F	1.8121%	9.24
T	6.9509%	35.43	Y	1.7779%	9.06
N	6.6544%	33.92	W	1.2899%	6.57
S	5.7351%	29.23	K	1.1016%	5.61
L	5.4893%	27.98	V	1.0074%	5.13
C	4.5388%	23.13	X	0.2902%	1.48
U	3.6308%	18.51	Z	0.2722%	1.39
D	3.3844%	17.25	J	0.1965%	1.00
P	3.1671%	16.14	Q	0.1962%	(1)



This observation was made by **Al-Kindi**, an Arab mathematician, around 800 A.C. and it represented the birth of **cryptanalysis**.

The objective of **cryptanalysis** is to break ciphers:

- discover the key used
- decipher text without knowing the key
- cipher text without knowing the key

Manuscript on Deciphering Cryptographic Messages

(al-Kindi)

TOWARDS MODERN CRYPTOGRAPHY

Even though the basic objectives of cryptography have not changed, throughout years what changed are the **techniques** used to achieve these objectives.

To begin our journey in understading modern cryptography, it is useful to start from a new foundamental idea

Kerckhoffs's principle

Kerckhoffs's principle

The security of a cryptographic system should not rely on the secrecy of the algorithm. Instead, it should be based on the secrecy of the cryptographic key.

A good cryptographic system should remain secure even if the algorithm used is known.

Anoher foundamental change between classical is due the introduction of the **bit** as the foundamental unit of information.

In this context, it is important to mention the bachelor thesis **Claude Shannon**, in which he showed the connection between

$$\texttt{Boolean Algebra} \leftrightarrow \texttt{Logic Circuits}$$

Claude Shannon, A symbolic analysis of relay and switching circuits, 1937

In the next video we will discuss in depth the following ideas

- The CIA Triad
- Mathematics
- Cryptographic Primitives
- Cryptographic Protocols