STEGANOGRAPHY

Text Steganography

Cybersecurity

Period 5

Cindy Zheng

Yi Ling Wu →

History of Steganography

The Word Steganography

- "Steganographia"
  - 3-book series written by a German abbot
- About magic, but contained secret information about cryptography and steganography

Earliest Example

- 440 BC
- Herodotus wrote about how Histiaeus would shave the heads of his slaves and tattoo secret messages on them
- Once the hair grew back, the message was hidden and could be delivered

World War II

- Spies used steganography to hide microdots of information in the period of a sentence.
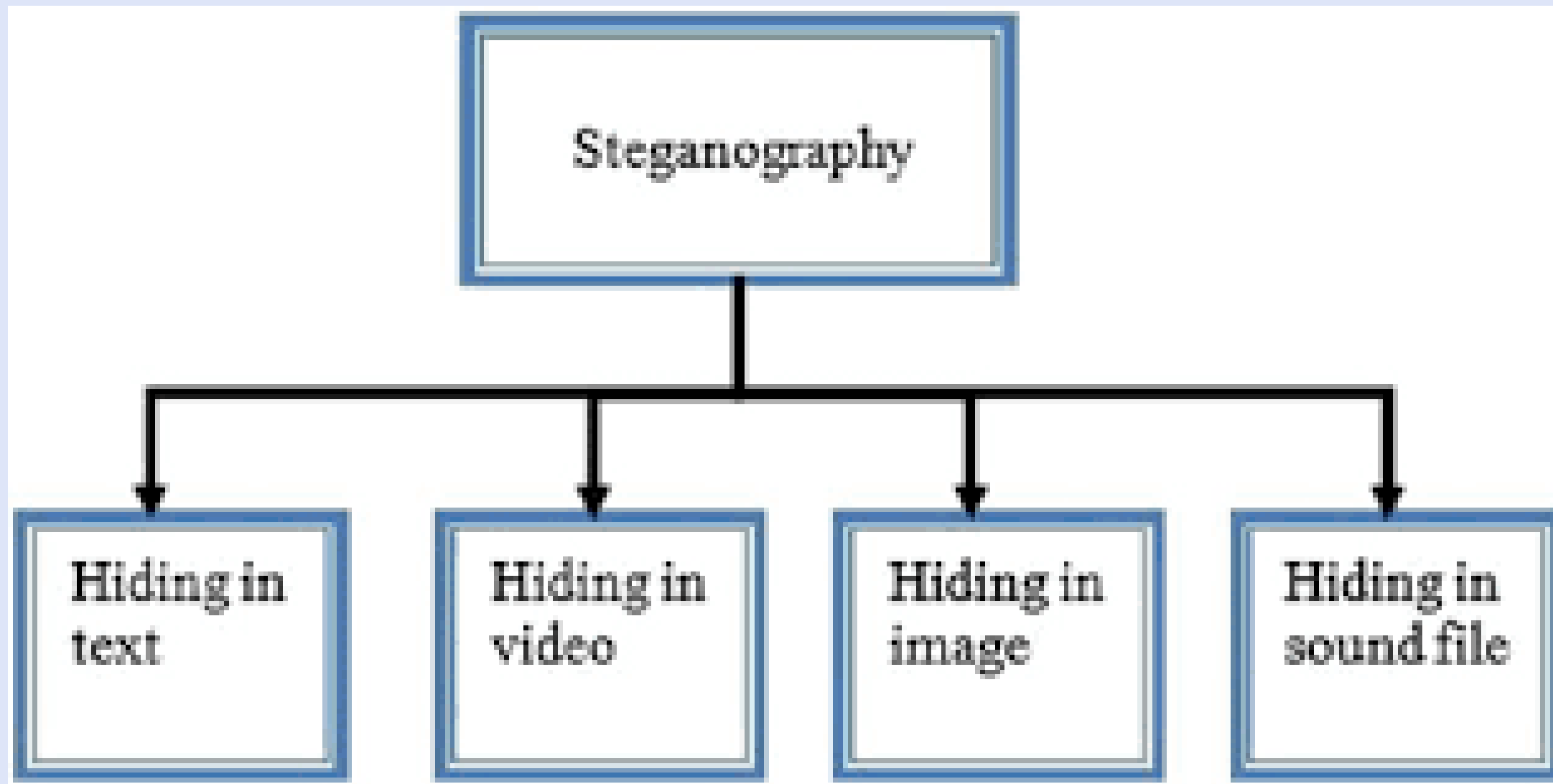
What is steganography?

- A method of hiding secret information in a non-secret medium

- Like using invisible ink on paper to hide secret messages!

- Main functions: embed and extract

- Can be used to conceal almost any type of digital content such as text, image, video, and audio within any of the said types.

What are the different kinds of steganography?

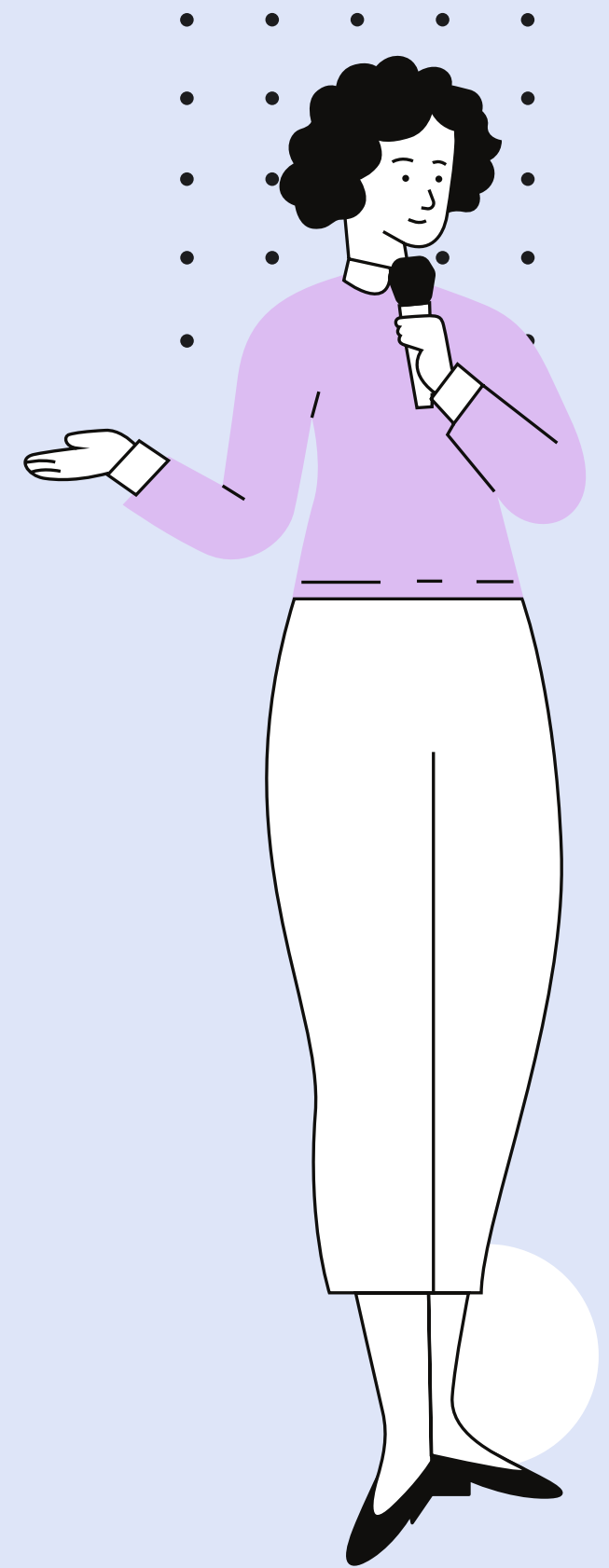Each has its advantages and disadvantages as well as its own uses

Pro:
Video and audio uses for copyright and for hiding information about the owner (name, signature, etc.)
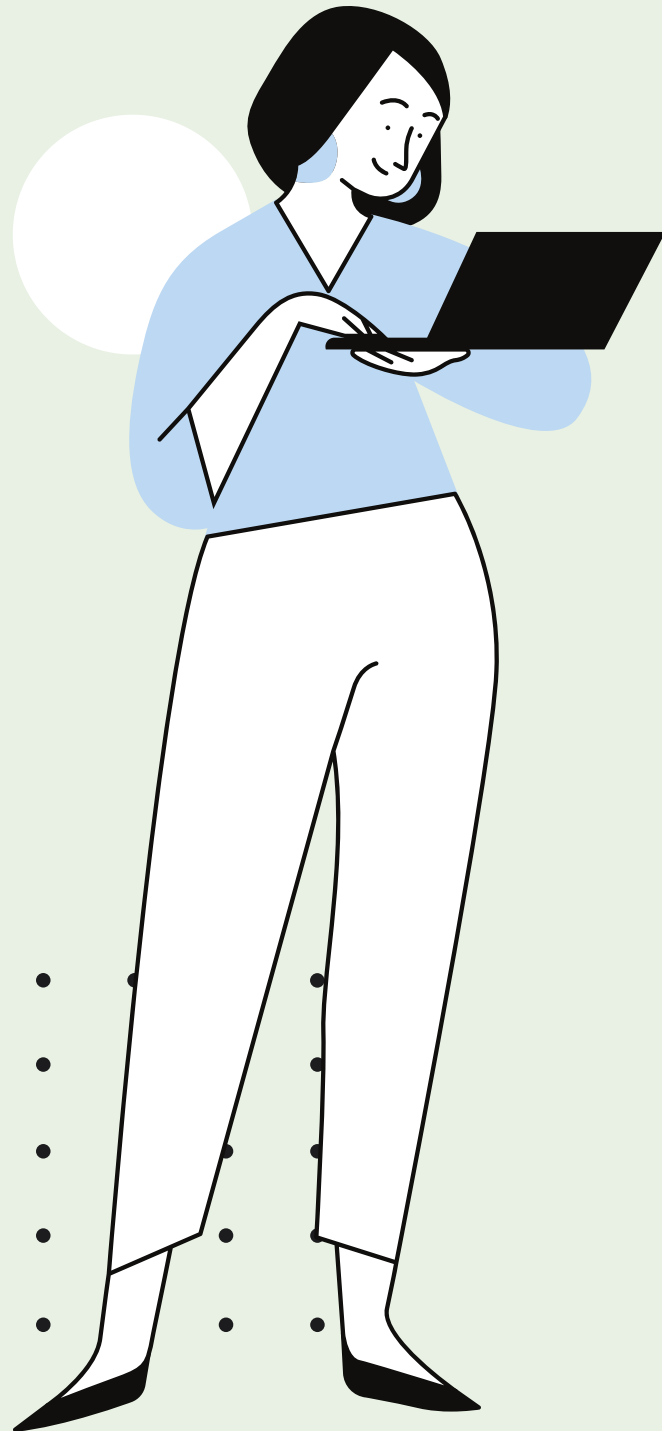
Con:
Quality of the image and audio can change

Video, audio, and image, are susceptible to compression

What is Text Steganography?

- Hides secret messages within other text that act as a cover message

- Often believed to be the trickiest because there isn't really redundant information

- It is also easier to spot differences in text

An example of White-Space Steganography!

What are the advantages of text steganography?

Over other steganography

- Lack of redundant information causes it to be harder to remove and compress
- Text is one of the biggest forms of communication so it is the most accessible and easiest
- Storing text files require less memory and is faster
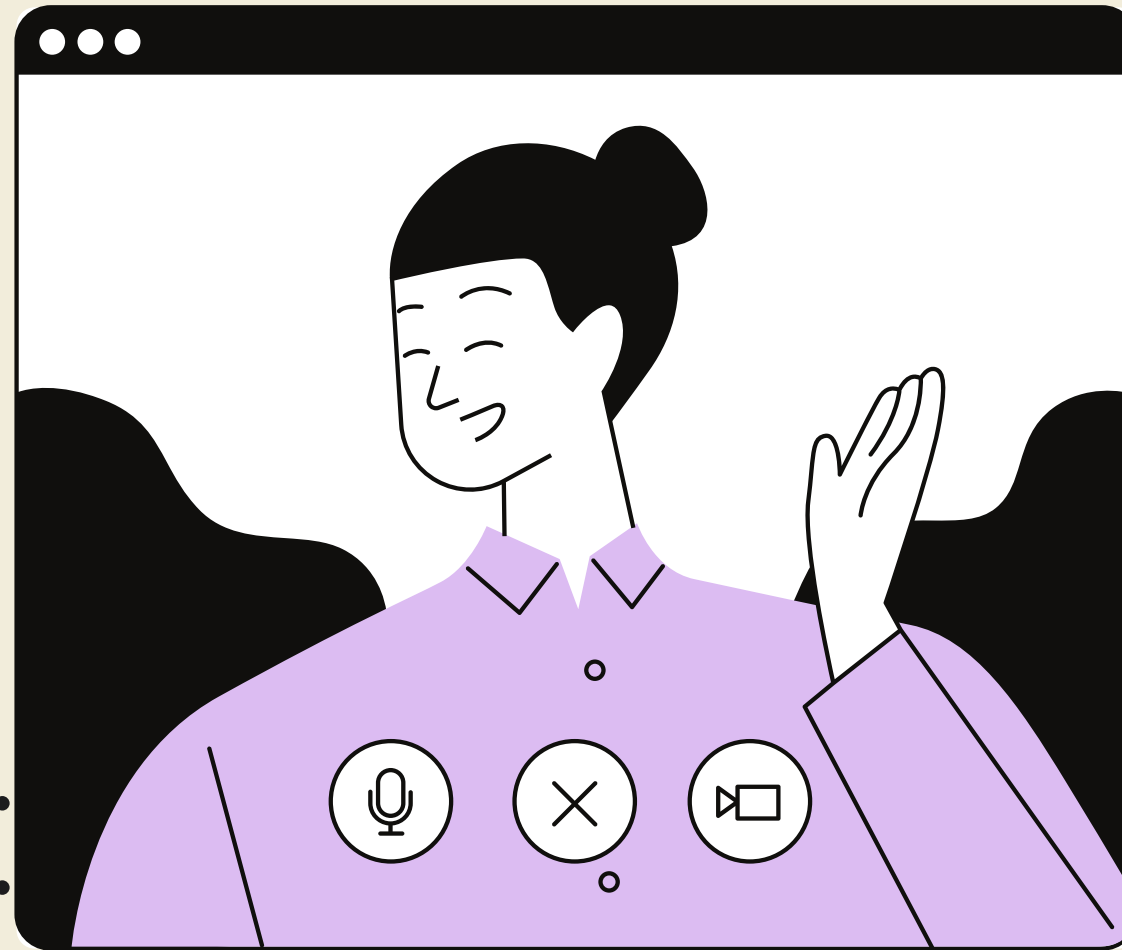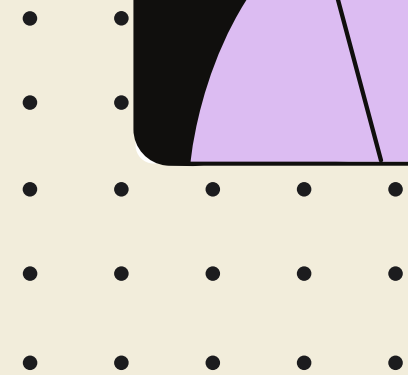
Over Cryptography

- Cryptography and steganography are often used hand in hand
- Steganography disguises the information under other media so other people don't realize that there is a hidden message at all.
- Cryptography makes the message unreadable, while steganography hides its entire existence

What are the three main types of text steganography?

Random and Statistics generation:
- generates cover using the secret message and a generation algorithm
- Drawbacks: extra complexity: Large amount of space and time

Linguistic:
- Manipulates the linguistic structure of the cover message
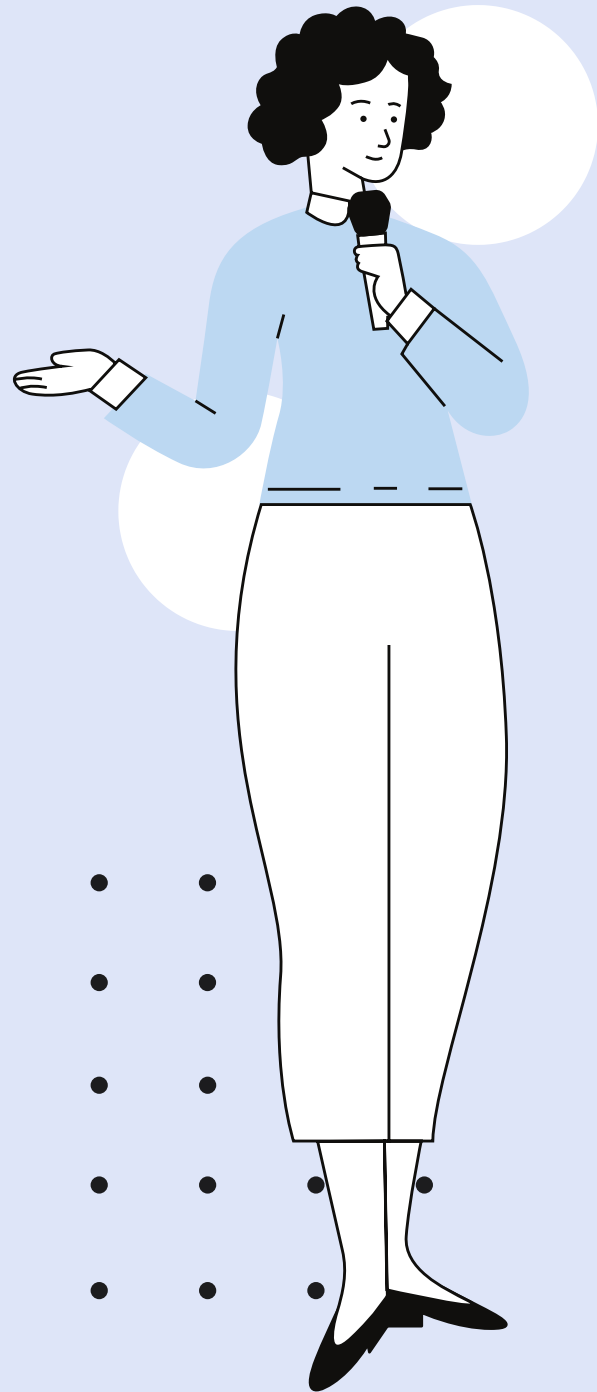
Syntactic text steganography
- structure of the text is altered without changing the meaning of the tone
- Con: Data size may get very large

Semantic text steganography
- synonyms of words are used to hide messages
- Con: Susceptible to third party attacks as the words may be changed but the cover text could still be the same
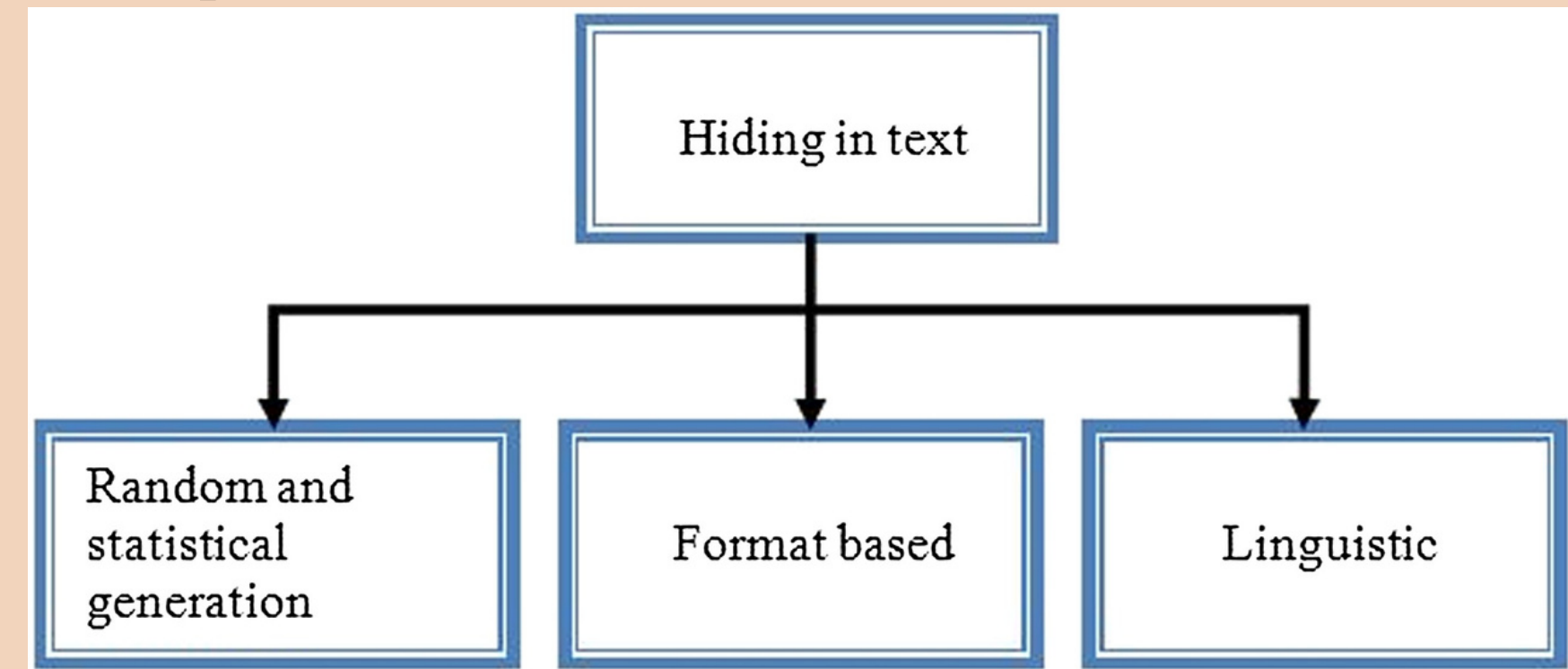
What is Format-based Text Steganography?

- Modifications to the physical structure of the text (spaces and special characters)
- The cover message will not look different to the naked eye
- Added special characters in the cover message hold value that can be pieced together to form a secret message
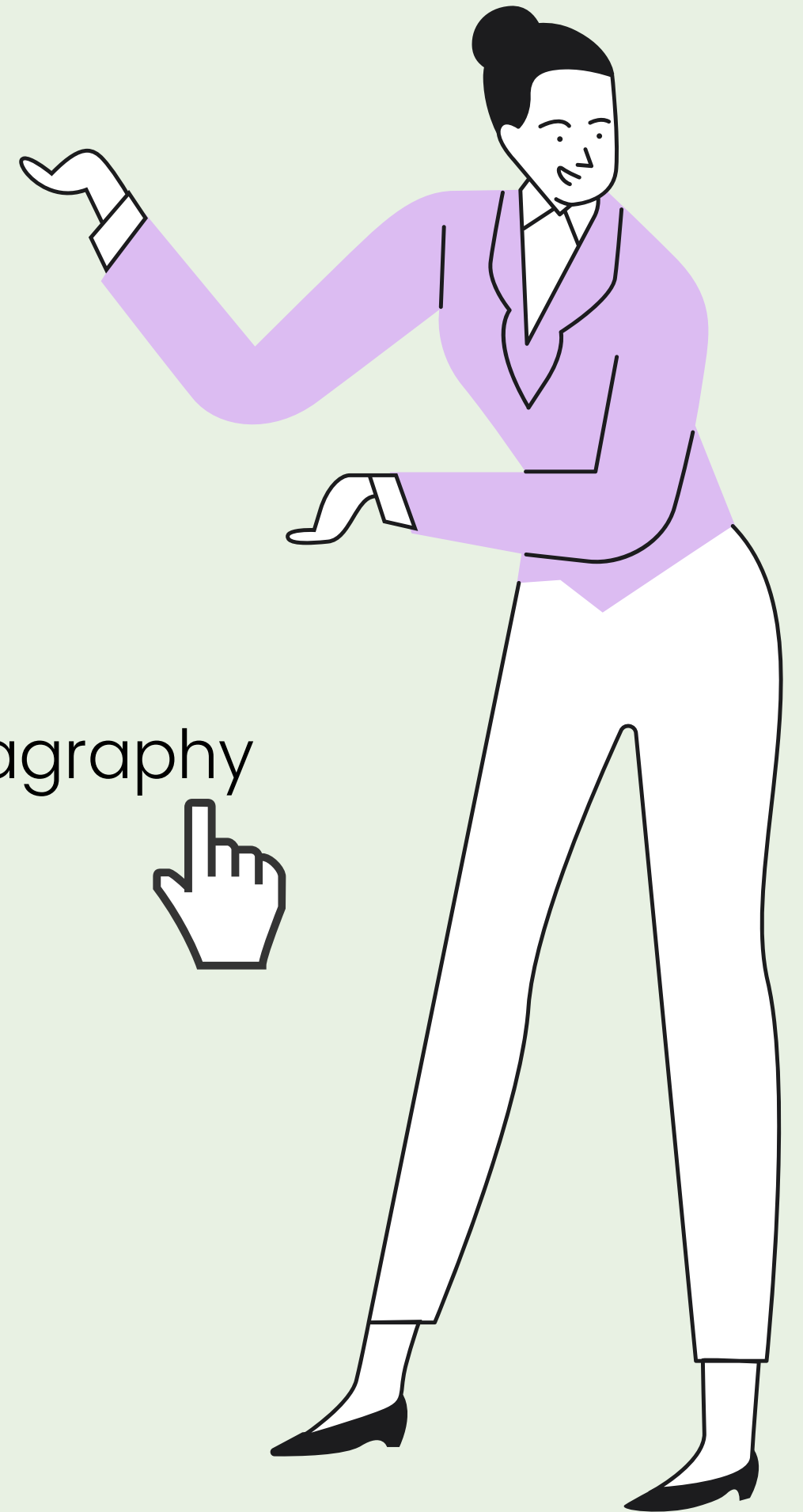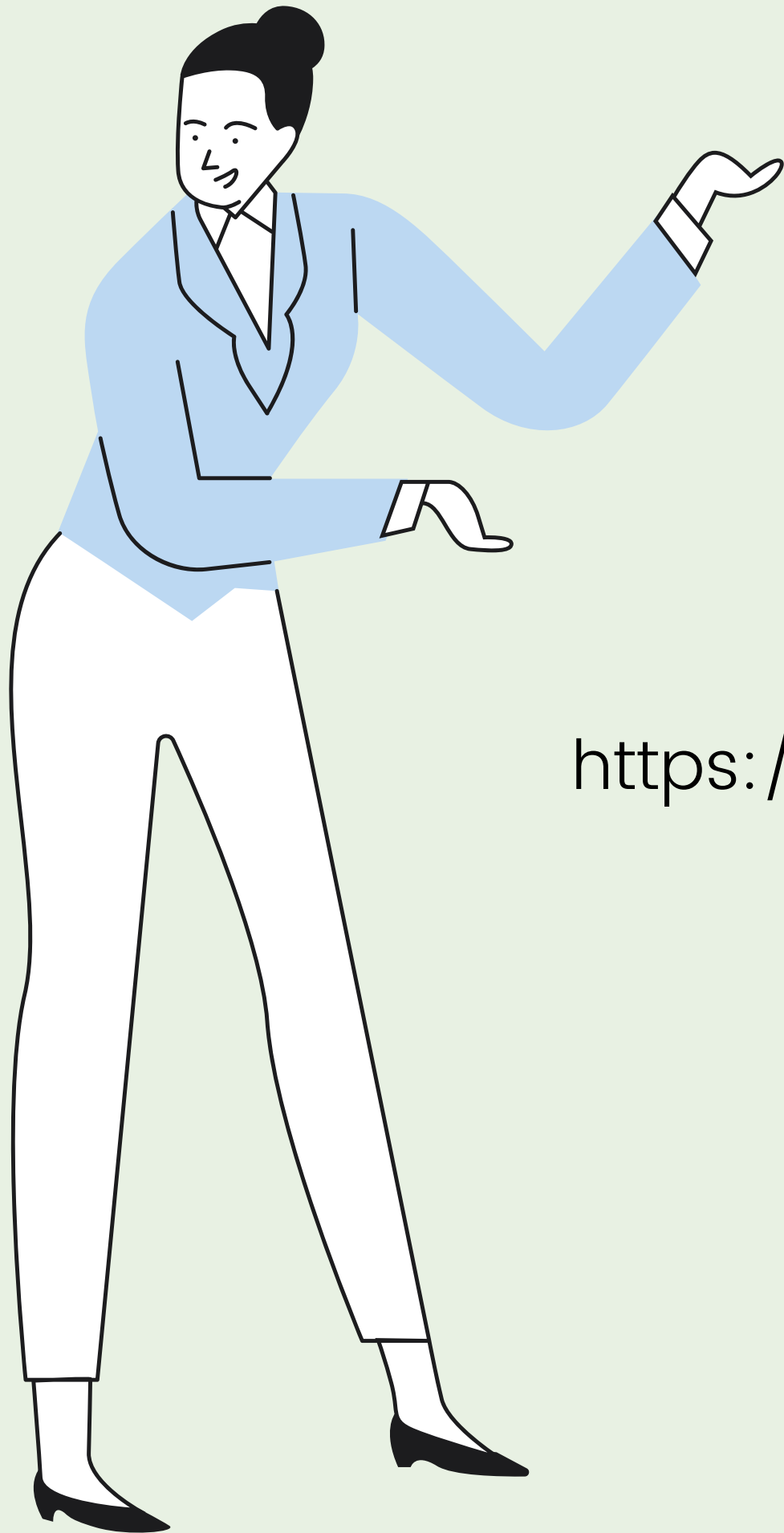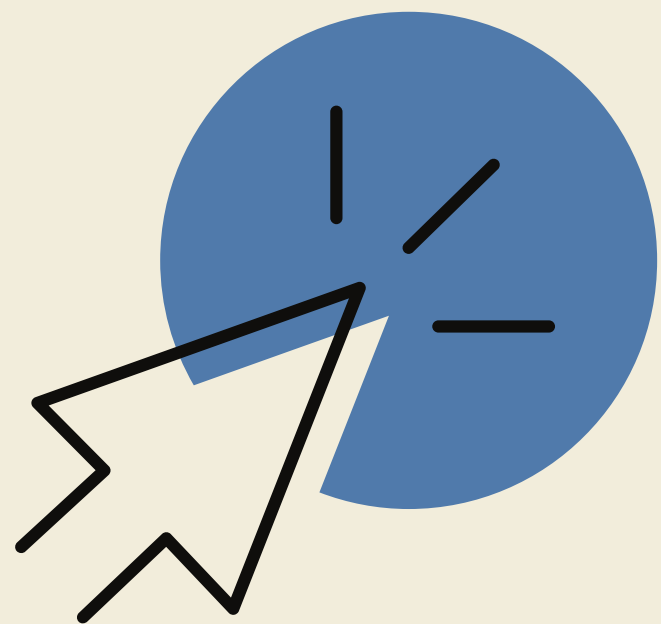
Drawbacks:
- Depending on the algorithm/method, word processors can detect misspellings or extra white spaces
- Having access to the original text file can also lead to detection through comparison
- Secret message destroyed with copying/reprinting

→

DEMO

https://github.com/YLW-105/textstegonagraphy

Sources

- https://embeddedsw.net/doc/Openpuff_thesis_StegChat_a_synonym_substitution_based_algorithm_for_text_steganography.pdf

- https://onlinelibrary.wiley.com/doi/full/10.1002/sec.1752#:~:text=Text steganography is a mechanism,the original secret message 5

- https://arxiv.org/ftp/arxiv/papers/1302/1302.2718.pdf

- https://searchsecurity.techtarget.com/definition/steganography