

# Suffix-Based Text Encoder & Decoder

**Author:** Karol Uciecha G00436758

**Project version:** 1.0

**Java version:** Java 22

**Project Type:** Text Processing / Data Compression / Java Console Application

**Dependencies:** Pure Java (no external libraries required)

**Mapping CSV Format:** Each line should be "token,code" (e.g. "testing,1234" or "@@ing,5678")

---

## Description

This program encodes natural language text into numeric codes using a dictionary-based mapping system that includes support for suffix encoding. Each word or part of a word is mapped to an integer, allowing for compressed representations. The system preserves punctuation and casing optionally and is capable of decoding numeric sequences back into readable text.

A custom mapping is loaded from a CSV file where each line defines a mapping between a word or suffix and an integer code. Encoding uses a greedy suffix-matching algorithm to minimize code output per word. Decoding reverses this process while reapplying optional formatting markers for upper-case, capitalized, or punctuated words.

---

## Features

- **Greedy Suffix Encoding:**  
Splits words into base tokens + suffixes (e.g. "testing" -> "test" + "@@ing").
  - **Case Sensitivity:**  
Preserves all-caps and capitalized words using dedicated numeric markers.
  - **Punctuation Support:**  
Encodes and decodes symbols such as commas, periods, and exclamation marks.
  - **Interactive Menu:**  
Users can specify input/output files and configure encoding preferences.
  - **Progress Meter:**  
Displays a live visual bar during prolonged actions - CSV loading, encoding, and decoding.
  - **Performance Stats:**  
Prints timing in milliseconds for each major operation.
- 

## How to Use

1. Run the program in any IDE with "Runner.java" or open dsa.jar by typing "java -cp ./dsa.jar ie.atu.sw.Runner".
2. Choose the first menu option to load a mapping CSV file.
3. Specify input/output files, and configure preferences. (Optional)
4. Use the Encode or Decode options to transform text accordingly.
5. Default file names:
  - text.txt -> input for encoding
  - encrypted.txt -> output of encoding & input for decoding
  - decrypted.txt -> output of decoding
6. Choose the last menu option to exit the program.