# Nasouh SLR Template (Your title goes here)

Jannie Jansen
Vrije Universiteit Amsterdam, The Netherlands
j.jansen@vu.nl

## ABSTRACT

The abstract of the study goes here. A good abstract shortly summarizes the paper by covering: (i) a presentation of the topic covered, (ii) the research followed, (iii) the main findings of the study, and (iv) the outcome of the investigation

## KEYWORDS

Systematic Literature Review, Other keywords identifying your study separated by comma

## 1 INTRODUCTION

This review explores how large language models (LLMs) are transforming linguistic steganography, the practice of hiding messages in text. We focus on the unique challenges and advances in using LLMs for secure, imperceptible, and high-capacity covert communication.

### 1.1 Overview of Information Security and Concealment Systems

Information security systems include **encryption**, **privacy**, and **concealment** (steganography).

*1.1.1 Encryption Systems and Privacy Systems.* These protect content but reveal that secret communication is happening, which can attract attention.

*1.1.2 Concealment Systems (Steganography).* Steganography hides the existence of information by embedding it in ordinary carriers (e.g., text, images). Text is a challenging carrier due to its low redundancy and strict semantics.

### 1.2 Introduction to Steganography

Steganography is often explained by the "Prisoners' Problem," where Alice and Bob must communicate secretly under surveillance. The goal is to embed messages so they are undetectable to an observer.

Steganography methods include **carrier selection**, **carrier modification**, and **carrier generation**.

- **Carrier modification:** Hide information in existing text with minimal changes.
- **Carrier generation:** Generate new text that encodes information, allowing higher capacity but requiring naturalness.

### 1.3 The Significance of Linguistic Steganography

Linguistic steganography enables covert communication, especially where encryption is suspicious. Text is a robust, ubiquitous carrier but presents challenges in balancing imperceptibility and capacity. Advances in deep learning and LLMs improve text quality and security, while related fields like watermarking focus on tracing content origin.

### 1.4 The Emergence of Large Language Models (LLMs)

LLMs like GPT-2 and LLaMA generate fluent, context-aware text, enabling new steganography methods with higher imperceptibility and capacity.

### 1.5 Scope of the Review

This review covers LLM-based linguistic steganography, focusing on methods, evaluation, challenges, and future directions.

## 2 BACKGROUND AND EVOLUTION OF LINGUISTIC STEGANOGRAPHY

This section summarizes key concepts and the evolution of linguistic steganography.

### 2.1 Imperceptibility in Steganography

Imperceptibility means making steganographic text indistinguishable from normal text, measured by perceptual, statistical, and semantic criteria. Improving one aspect can harm another (Psic Effect).

### 2.2 Security Definitions and Metrics

Security is measured by information-theoretic metrics (e.g., KL divergence) and computational indistinguishability.

### 2.3 Evolution of Linguistic Steganography

Early methods used syntax/statistics but lacked naturalness. Neural models improved fluency but sometimes compromised security.

### 2.4 Provably Secure Steganography in Practice

Classical secure methods include rejection sampling and arithmetic coding.

## 3 LLMS IN STEGANOGRAPHY: APPROACHES AND PARADIGMS

### 3.1 Direct Generation via User Interfaces (Black-Box LLM-Stega)

LLMs generate steganographic text directly via prompts, treating the model as a black box.

### 3.2 Rewriting-Based Linguistic Steganography

Secret messages are embedded by rewriting existing text using LLMs.

### 3.3 Context-Aware Linguistic Steganography (NMT-Stega)

Methods leverage LLMs' context awareness for more natural and secure embedding.

### 3.4 Acrostic Linguistic Steganography

LLMs generate text where specific positions encode information (e.g., acrostics).

### 3.5 Collaborative/Hierarchical Linguistic Steganography

Combines multiple LLMs or hierarchical strategies for improved capacity and imperceptibility.

### 3.6 LLM-Based Watermarking (Related Work to Steganography)

Watermarking uses LLMs to embed traceable patterns for provenance.

*3.6.1 DeepTextMark.* Deep learning-based watermarking for text.

*3.6.2 Principled Approach to Natural Language Watermarking.* Formal methods for robust, undetectable watermarking.

*3.6.3 ChatGPT for Stego Covers.* ChatGPT is used to generate steganographic covers.

*3.6.4 Paraphraser-based Lexical Substitution (Para-NLW).* Paraphrasing with LLMs to encode information via lexical choices.

## 4 KEY CHALLENGES AND LIMITATIONS

### 4.1 Perceptual vs. Statistical Imperceptibility (Psic Effect)

Improving perceptual quality can reduce statistical security, and vice versa.

### 4.2 Low Embedding Capacity

Short texts and strict semantics limit how much information can be hidden.

### 4.3 Lack of Semantic Control and Contextual Consistency

Ensuring generated text matches intended meaning/context is difficult.

### 4.4 Challenges with LLMs in Steganography

LLMs may introduce unpredictability, bias, or leak information.

### 4.5 Segmentation Ambiguity

Tokenization can cause ambiguity in how information is embedded or extracted.

## 5 EVALUATION METRICS IN LINGUISTIC STEGANOGRAPHY

### 5.1 Imperceptibility Metrics

Perceptual: PPL, Distinct-n, MAUVE, human evaluation. Statistical: KLD, JSD, anti-steganalysis accuracy, semantic similarity.

### 5.2 Embedding Capacity Metrics

Bits per token/word, embedding rate.

## 6 RELATED WORK

Brief summary of similar studies and how this review fits in the research landscape.

## 7 STUDY DESIGN

### 7.1 Research Goal

State the main research goal (e.g., characterize LLM-based linguistic steganography).

### 7.2 Research Questions

List the key research questions addressed.

### 7.3 Initial search

Describe the initial literature search process.

### 7.4 Application of selection criteria

Summarize inclusion/exclusion criteria for study selection.

### 7.5 Snowballing

Briefly note if snowballing was used for additional sources.

### 7.6 Data Extraction

Describe how data was extracted from selected studies.

### 7.7 Data Synthesis

Summarize the approach for synthesizing extracted data.

### 7.8 Study Replicability

List materials/data made available for replicability.

## 8 RESULTS

Summarize main findings from the literature review.

## 9 DISCUSSION

Discuss implications and interpretation of the results.

## 10 CONCLUSION

Summarize the main findings and takeaways of the study.

## REFERENCES