# Chapter 1: Introduction

## 1.1 Overview

Natural Language Processing (NLP) stands at the intersection between computer science and linguistics, providing a framework through which machines can understand, interpret, and generate human language. The genesis of NLP can be traced back to the mid-20th century, with early efforts focused on rule-based methods to translate text between languages [4]. These foundational efforts laid the groundwork for the field, emphasizing the goal of bridging human-computer interaction through the medium of language [5, 6, 7].

As computational power surged, methodologies evolved, and massive datasets were compiled, there was a marked transition towards statistical models in NLP. This period saw the introduction of machine learning algorithms that use vast amounts of text data to learn language patterns, significantly improving language understanding and generation tasks. These statistical models paved the way for the development of topic modeling techniques, which emerged as powerful tools for uncovering hidden thematic structures in large text corpora [8, 9]. Traditional algorithms like Latent Dirichlet Allocation (LDA) [1] became staples in the NLP toolkit [10, 11, 12], enabling researchers and practitioners to extract meaningful insights from text data across various domains.

The landscape of NLP underwent a significant transformation with the shift from traditional probabilistic models to neural-based models. Probabilistic models were grounded in the assumption that documents are admixtures of topics, where a topic is characterized by a distribution over words. And while such models offer a mathematically elegant way to discover latent topics within large corpora of text through statistical inference and are effective for a range of applications, probabilistic models often struggle with the complexities of language [13, 14]. We present our work that improve these weaknesses in Section 3.4

The advent of deep learning provided a new set of tools to address these challenges, ushering in an era of neural-based topic models. Leveraging the representational power of neural networks, these models introduced the ability to learn dense, continuous word embeddings, capturing nuanced semantic relationships that were difficult for their probabilistic counterparts to model. This transition was marked by innovative approaches that integrated the strengths of deep learning with topic modeling objectives. Early attemps included shallow graphical models [15] and neural autoregressive models [16]. As neural network architectures improved, so did the topic models. A popular method being the suite of variational-based topic models [17, 18, 19]. These advances not only enhanced the granularity and coherence of the topics extracted but also improved the models' adaptability and performance across diverse datasets.

The integration of Large Language Models (LLM) into the domain of topic modeling represents a convergence, uniting the nuanced understanding enabled by deep learning with the interpretable insight provided by topic modeling. This fusion has opened new avenues for exploring text data, allowing for more interactive, dynamic, and nuanced analysis of language. In particular, the application of these technologies to understand specialized language use cases, such as the developmental language of children and the psychological dispositions reflected in text, underscores the versatility and potential of modern NLP methods.

In this thesis, we explore how modern neural and language models can be made more interpretable, interactive, and adaptable for human-in-the-loop and human-centered applications. Our central thesis is that the next generation of topic models and predictive systems must not only achieve high performance, but also be designed around human usability, semantic coherence, and contextual understanding. To this end, we pursue two complementary directions: (1) improving topic modeling architectures through interaction and large language models, and (2) improving inference of psychological and developmental signals through the lens of large language models.

## 1.2 Enhancing Topic Models Through Interaction

Interactive topic modeling represents a pivotal shift in the development of topic modeling techniques, aiming to reconcile the often opaqueness of algorithmically generated topics with the nuanced understanding of human analysts. Traditional topic models, while powerful, frequently yield topics that, although mathematically coherent, lack interpretability or relevance from a user's perspective. The inception of interactive topic modeling emerged from the recognition of this gap, proposing a paradigm where user feedback becomes an integral part of the model refinement process

The foundational works in interactive topic modeling sought to address the static nature of traditional models by introducing mechanisms for dynamic feedback. Early efforts in this domain were characterized by the integration of user inputs into the iterative process of topic model training, thereby allowing for the direct influence of human expertise on the resulting topics [20]. Further advancements were made through the introduction of anchor words as a mechanism for interactive topic modeling. This approach, which allows users to specify words that are emblematic of the topics they wish to explore, significantly enhances the model's ability to align with user expectations and analytical goals [21, 22]. We will discuss this more in Section 2.3

The methodological innovations in interactive topic modeling have been diverse, ranging from graphical user interfaces designed for topic refinement to algorithmic adjustments that accommodate user feedback in real-time [23, 24, 25]. These developments have not only made topic models more accessible to non-expert users but have also opened new avenues for applying topic modeling in domains requiring bespoke analytical perspectives. The development of graphical interfaces that visualize topic distributions and their evolution over time has been a key innovation. Such interfaces facilitate a more intuitive interaction between the user and the model, enabling the iterative refinement of topics based on visual feedback. The flexibility of interactive topic models has led to their application across a wide range of domains, such as public health monitoring on social media [26].

However, these methods have yet to be adapted for neural-based topic models. We seek to fill this lacunae. Our work introduces a novel avenue for interaction with neural topic models, through a user-friendly interface and a flexible framework for making neural topic models interactive.

### 1.2.1 Large Language Model Enhanced Topic Models

The integration of LLM into topic modeling represents a significant advancement in the field NLP. This evolution from using simple embeddings to the full-scale application of LLM for topic modeling has dramatically enhanced the ability to capture complex semantic relationships within large text corpora. The initial phase of incorporating LLM into topic modeling focused on leveraging word and document embeddings. These embeddings, dense vector representations of words and documents, captured nuanced semantic and syntactic relationships that were not easily accessible with traditional bag-of-words models [27, 28, 29, 30].

Very recent work has led to the full utilization of LLM in topic modeling, leveraging the entire architecture of models like BERT and GPT for both extracting and understanding topics. Leveraging the capabilities of models like GPT-3 for zero-shot or few-shot learning allowed for the generation of topics without extensive retraining. This approach utilizes the model's inherent understanding of language to identify themes and topics within texts based on minimal prompts or examples [31].

Our work seeks to fill the gap between LLM and probabilistic models. Leveraging the world-knowledge of LLM and the interpretibility of probabilistic models, we develop a combined model that takes the best-of-both-worlds to develop coherent and interpretable topics.

## 1.3 Leveraging Large Language Models for Improved Precision in Predictive Linguistics

Topic modeling and LLM, have significantly broadened the scope of language understanding in human-centered applications. This section delves into two pivotal projects: the utilization of topic models to assess the language depth of children and the application of LLM to infer psychological dispositions. It is a continuing problem in computational linguistics to effectively model language, with many efforts applied to the acquisition of language in children [32, 33, 34, 35], however these models follow heuristic-based models that are first, computational linguistic models.

Unsupervised learning has been used to generate grammars of children data to understand language acquisition [36]. However, no work yet exists seeking to use the full capability of unsupervised learning to model the language depth and breadth of children language. This is where our work comes in. We aim to build robust models that can model the language acquisition of children across languages by modeling the breadth and depth of their vocabulary with topic models.

On a separate front, the inference of psychological dispositions from textual data represents another frontier where NLP intersects with human-centric concerns. There is a long line of research seeking to predict the psychological and personality state of users based on their textual data [37, 38].

As the field of NLP advanced, researchers revisited this problem, now with stronger computational models [39, 40, 41, 42, 43].

Now, armed with the world knowledge and wide-ranging capabilities of LLM, researchers are starting to test these kinds of models in predicting psychological disposition [44, 45, 46]. We join this group of researchers, pushing the possibilities of using LLM to predict psychological traits of users from textual data.

## 1.4 Thesis Contribution

This dissertation advances the field of topic modeling by proposing novel interactive neural topic models that integrate user feedback and large language models to improve both topic interpretability and downstream performance. Traditional topic models often suffer from limited semantic coherence and lack of adaptability, particularly when applied to real-world tasks. In response, this work develops and evaluates interactive methods that allow users to iteratively shape topic distributions, leading to more focused and relevant topic spaces. Furthermore, by leveraging LLM such as BERT within the topic modeling framework, as priors, the models bridge the gap between probabilistic structure and rich contextual knowledge, enabling more adaptable representations. These models are shown to outperform prior approaches in both interpretability and classification tasks.

Beyond modeling advances, this dissertation applies LLM to the task of predicting psychological dispositions from text, revealing both their promise and pitfalls. LLM consistently outperform classical machine learning models on this task, especially in low-data settings. However, their performance exposes a deeper issue: standard evaluation methods and datasets often fail to capture the nuanced, latent nature of psychological traits. LLMs can pick up on superficial linguistic cues, leading to overfitting to annotation biases or stylistic artifacts. This dissertation argues that while LLMs show clear potential for high-level inference tasks, they require structured human feedback and iterative refinement to be reliable in sensitive domains. These findings underscore the importance of incorporating human-in-the-loop approaches and designing evaluations that reflect the complexity of real-world applications

## 1.5 Proposal Outline

Chapter 2 reviews relevant previous work. We introduce the modern neural networks in NLP. We then dive into the two research directions this proposal will take.

Chapter 3 will introduce I-NTM, an interactive framework for neural topic modeling

Chapter 4 will dive into the intricacies of using large language models for predicting personality traits of people from textual data

In Chapter 5 we explore using open-source LLM to predict psycological dispositions and the challenges of user-grounded evaluations