Papers

The following are the latest papers (sorted by release date) on prompt engineering for large language models (LLMs). We update the list of papers on a daily/weekly basis.

Overviews

- The Prompt Report: A Systematic Survey of Prompting Techniques (June 2024)
- Prompt Design and Engineering: Introduction and Advanced Methods (January 2024)
- A Survey on Hallucination in Large Language Models: Principles,Taxonomy, Challenges, and Open Questions (November 2023)
- An RL Perspective on RLHF, Prompting, and Beyond (October 2023)
- Few-shot Fine-tuning vs. In-context Learning: A Fair Comparison and Evaluation (May 2023)
- Jailbreaking ChatGPT via Prompt Engineering: An Empirical Study (May 2023)
- Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond (April 2023)
- Tool Learning with Foundation Models (April 2023)
- One Small Step for Generative AI, One Giant Leap for AGI: A Complete Survey on ChatGPT in AIGC Era (April 2023)
- A Bibliometric Review of Large Language Models Research from 2017 to 2023 (April 2023)
- A Survey of Large Language Models (April 2023)
- Nature Language Reasoning, A Survey (March 2023)
- Augmented Language Models: a Survey (February 2023)
- A Survey for In-context Learning (December 2022)
- Towards Reasoning in Large Language Models: A Survey (December 2022)
- Reasoning with Language Model Prompting: A Survey (December 2022)
- Emergent Abilities of Large Language Models (June 2022)
- A Taxonomy of Prompt Modifiers for Text-To-Image Generation (April 2022)
- Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing (July 2021)

Approaches

- Enhancing Zero-Shot Chain-of-Thought Reasoning in Large Language Models through Logic (February 2024)
- Principled Instructions Are All You Need for Questioning LLaMA-1/2, GPT-3.5/4 (December 2023)
- Walking Down the Memory Maze: Beyond Context Limit through Interactive Reading (October 2023)
- Large Language Models as Analogical Reasoners (October 2023)
- LLMLingua: Compressing Prompts for Accelerated Inference of Large Language Models (October 2023)
- Query-Dependent Prompt Evaluation and Optimization with Offline Inverse RL (September 2023)
- Chain-of-Verification Reduces Hallucination in Large Language Models (September 2023)
- Connecting Large Language Models with Evolutionary Algorithms Yields Powerful Prompt Optimizers (September 2023)
- From Sparse to Dense: GPT-4 Summarization with Chain of Density Prompting (September 2023)
- Re-Reading Improves Reasoning in Language Models (September 2023)
- Graph of Thoughts: Solving Elaborate Problems with Large Language Models (August 2023)
- Skeleton-of-Thought: Large Language Models Can Do Parallel Decoding (July 2023)
- Focused Prefix Tuning for Controllable Text Generation (June 2023)
- Exploring Lottery Prompts for Pre-trained Language Models (May 2023)
- Less Likely Brainstorming: Using Language Models to Generate Alternative Hypotheses (May 2023)
- Let's Verify Step by Step (May 2023)
- Universality and Limitations of Prompt Tuning (May 2023)
- MultiTool-CoT: GPT-3 Can Use Multiple External Tools with Chain of Thought Prompting (May 2023)
- PEARL: Prompting Large Language Models to Plan and Execute Actions Over Long Documents (May 2023)
- Reasoning with Language Model is Planning with World Model (May 2023)
- Self-Critique Prompting with Large Language Models for Inductive Instructions (May 2023)
- Better Zero-Shot Reasoning with Self-Adaptive Prompting (May 2023)

- Hierarchical Prompting Assists Large Language Model on Web Navigation (May 2023)
- Interactive Natural Language Processing (May 2023)
- Can We Edit Factual Knowledge by In-Context Learning? (May 2023)
- In-Context Learning of Large Language Models Explained as Kernel Regression (May 2023)
- Plan-and-Solve Prompting: Improving Zero-Shot Chain-of-Thought Reasoning by Large Language Models (May 2023)
- Meta-in-context learning in large language models (May 2023)
- Let's Sample Step by Step: Adaptive-Consistency for Efficient Reasoning with LLMs (May 2023)
- Post Hoc Explanations of Language Models Can Improve Language Models (May 2023)
- Compress, Then Prompt: Improving Accuracy-Efficiency Trade-off of LLM Inference with Transferable Prompt (May 2023)
- TreePrompt: Learning to Compose Tree Prompts for Explainable Visual Grounding (May 2023)
- TELeR: A General Taxonomy of LLM Prompts for Benchmarking Complex Tasks (May 2023)
- Efficient Prompting via Dynamic In-Context Learning (May 2023)
- The Web Can Be Your Oyster for Improving Large Language Models (May 2023)
- Flatness-Aware Prompt Selection Improves Accuracy and Sample Efficiency (May 2023)
- Tree of Thoughts: Deliberate Problem Solving with Large Language Models (May 2023)
- ZeroPrompt: Streaming Acoustic Encoders are Zero-Shot Masked LMs (May 2023)
- Chain-of-Symbol Prompting Elicits Planning in Large Langauge Models (May 2023)
- CooK: Empowering General-Purpose Language Models with Modular and Collaborative Knowledge (May 2023)
- What In-Context Learning "Learns" In-Context: Disentangling Task Recognition and Task Learning (May 2023)
- Reprompting: Automated Chain-of-Thought Prompt Inference Through Gibbs Sampling (May 2023)
- Satisfiability-Aided Language Models Using Declarative Prompting (May 2023)
- Pre-Training to Learn in Context (May 2023)
- Boosted Prompt Ensembles for Large Language Models (April 2023)
- Global Prompt Cell: A Portable Control Module for Effective Prompt (April 2023)

- Why think step-by-step? Reasoning emerges from the locality of experience (April 2023)
- Revisiting Automated Prompting: Are We Actually Doing Better? (April 2023)
- REFINER: Reasoning Feedback on Intermediate Representations (April 2023)
- Reflexion: an autonomous agent with dynamic memory and self-reflection (March 2023)
- CAMEL: Communicative Agents for "Mind" Exploration of Large Scale Language Model Society (March 2023)
- Self-Refine: Iterative Refinement with Self-Feedback (March 2023)
- kNN Prompting: Beyond-Context Learning with Calibration-Free Nearest Neighbor Inference (March 2023)
- Visual-Language Prompt Tuning with Knowledge-guided Context Optimization (March 2023)
- Fairness-guided Few-shot Prompting for Large Language Models (March 2023)
- Context-faithful Prompting for Large Language Models (March 2023)
- Is Prompt All You Need? No. A Comprehensive and Broader View of Instruction Learning (March 2023)
- UPRISE: Universal Prompt Retrieval for Improving Zero-Shot Evaluation (March 2023)
- Model-tuning Via Prompts Makes NLP Models Adversarially Robust (March 2023)
- Structure Pretraining and Prompt Tuning for Knowledge Graph Transfer (March 2023)
- CoTEVer: Chain of Thought Prompting Annotation Toolkit for Explanation Verification (March 2023)
- Larger language models do in-context learning differently (March 2023)
- OpenICL: An Open-Source Framework for In-context Learning (March 2023)
- Dynamic Prompting: A Unified Framework for Prompt Tuning (March 2023)
- ART: Automatic multi-step reasoning and tool-use for large language models (March 2023)
- Multitask Prompt Tuning Enables Parameter-Efficient Transfer Learning (March 2023)
- Effectiveness of Data Augmentation for Prefix Tuning with Limited Data (March 2023)
- Mixture of Soft Prompts for Controllable Data Generation (March 2023)
- Prompt, Generate, then Cache: Cascade of Foundation Models makes Strong Few-shot Learners (March 2023)
- How Robust is GPT-3.5 to Predecessors? A Comprehensive Study on Language Understanding Tasks (March 2023)

- Can ChatGPT Understand Too? A Comparative Study on ChatGPT and Fine-tuned BERT (February 2023)
- EvoPrompting: Language Models for Code-Level Neural Architecture Search (February 2023)
- In-Context Instruction Learning (February 2023)
- Chain of Hindsight Aligns Language Models with Feedback (February 2023)
- Language Is Not All You Need: Aligning Perception with Language Models (February 2023)
- Automatic Prompt Augmentation and Selection with Chain-of-Thought from Labeled Data (February 2023)
- Active Prompting with Chain-of-Thought for Large Language Models (February 2023)
- More than you've asked for: A Comprehensive Analysis of Novel Prompt Injection Threats to Application-Integrated Large Language Models (February 2023)
- A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT (February 2023)
- Guiding Large Language Models via Directional Stimulus Prompting (February 2023)
- How Does In-Context Learning Help Prompt Tuning? (February 2023)
- Scalable Prompt Generation for Semi-supervised Learning with Language Models (February 2023)
- Bounding the Capabilities of Large Language Models in Open Text Generation with Prompt Constraints (February 2023)
- À-la-carte Prompt Tuning (APT): Combining Distinct Data Via Composable Prompting (February 2023)
- GraphPrompt: Unifying Pre-Training and Downstream Tasks for Graph Neural Networks (February 2023)
- The Capacity for Moral Self-Correction in Large Language Models (February 2023)
- SwitchPrompt: Learning Domain-Specific Gated Soft Prompts for Classification in Low-Resource Domains (February 2023)
- Evaluating the Robustness of Discrete Prompts (February 2023)
- Compositional Exemplars for In-context Learning (February 2023)
- Hard Prompts Made Easy: Gradient-Based Discrete Optimization for Prompt Tuning and Discovery (February 2023)
- Multimodal Chain-of-Thought Reasoning in Language Models (February 2023)
- Large Language Models Can Be Easily Distracted by Irrelevant Context (February 2023)
- Synthetic Prompting: Generating Chain-of-Thought Demonstrations for Large Language Models (February 2023)
- Progressive Prompts: Continual Learning for Language Models (January 2023)

- Batch Prompting: Efficient Inference with LLM APIs (January 2023)
- Demonstrate-Search-Predict: Composing retrieval and language models for knowledge-intensive NLP (December 2022)
- On Second Thought, Let's Not Think Step by Step! Bias and Toxicity in Zero-Shot Reasoning (December 2022)
- Constitutional AI: Harmlessness from AI Feedback (December 2022)
- Successive Prompting for Decomposing Complex Questions (December 2022)
- Large Language Models are reasoners with Self-Verification (December 2022)
- Discovering Language Model Behaviors with Model-Written Evaluations (December 2022)
- Structured Prompting: Scaling In-Context Learning to 1,000 Examples (December 2022)
- PAL: Program-aided Language Models (November 2022)
- Large Language Models Are Human-Level Prompt Engineers (November 2022)
- Ignore Previous Prompt: Attack Techniques For Language Models (November 2022)
- Machine Generated Text: A Comprehensive Survey of Threat Models and Detection Methods (November 2022)
- Teaching Algorithmic Reasoning via In-context Learning (November 2022)
- Enhancing Self-Consistency and Performance of Pre-Trained Language Models through Natural Language Inference (November 2022)
- Ask Me Anything: A simple strategy for prompting language models (October 2022)
- Recitation-Augmented Language Models (October 2022)
- ReAct: Synergizing Reasoning and Acting in Language Models (October 2022)
- Prompting GPT-3 To Be Reliable (October 2022)
- Decomposed Prompting: A Modular Approach for Solving Complex Tasks (October 2022)
- Automatic Chain of Thought Prompting in Large Language Models (October 2022)
- Language Models Are Greedy Reasoners: A Systematic Formal Analysis of Chain-of-Thought (October 2022)
- Evaluating the Susceptibility of Pre-Trained Language Models via Handcrafted Adversarial Examples (September 2022)
- Dynamic Prompt Learning via Policy Gradient for Semi-structured Mathematical Reasoning (September 2022)
- Promptagator: Few-shot Dense Retrieval From 8 Examples (September 2022)
- Atlas: Few-shot Learning with Retrieval Augmented Language Models (November 2022)
- DocPrompting: Generating Code by Retrieving the Docs (July 2022)
- On the Advance of Making Language Models Better Reasoners (June 2022)

- Large Language Models are Zero-Shot Reasoners (May 2022)
- Maieutic Prompting: Logically Consistent Reasoning with Recursive Explanations (May 2022)
- MRKL Systems: A modular, neuro-symbolic architecture that combines large language models, external knowledge sources and discrete reasoning (May 2022)
- PPT: Pre-trained Prompt Tuning for Few-shot Learning (Mqy 2022)
- Toxicity Detection with Generative Prompt-based Inference (May 2022)
- Learning to Transfer Prompts for Text Generation (May 2022)
- The Unreliability of Explanations in Few-shot Prompting for Textual Reasoning (May 2022)
- A Taxonomy of Prompt Modifiers for Text-To-Image Generation (April 2022)
- PromptChainer: Chaining Large Language Model Prompts through Visual Programming (March 2022)
- Self-Consistency Improves Chain of Thought Reasoning in Language Models (March 2022)
- Training language models to follow instructions with human feedback
- Rethinking the Role of Demonstrations: What Makes In-Context Learning Work? (February 2022)
- Chain of Thought Prompting Elicits Reasoning in Large Language Models (January 2022)
- Show Your Work: Scratchpads for Intermediate Computation with Language Models (November 2021)
- AI Chains: Transparent and Controllable Human-AI Interaction by Chaining Large Language Model Prompts (October 2021)
- Generated Knowledge Prompting for Commonsense Reasoning (October 2021)
- Multitask Prompted Training Enables Zero-Shot Task Generalization (October 2021)
- Reframing Instructional Prompts to GPTk's Language (September 2021)
- Design Guidelines for Prompt Engineering Text-to-Image Generative Models (September 2021)
- Making Pre-trained Language Models Better Few-shot Learners (August 2021)
- Fantastically Ordered Prompts and Where to Find Them: Overcoming Few-Shot Prompt Order Sensitivity (April 2021)
- BERTese: Learning to Speak to BERT (April 2021)
- The Power of Scale for Parameter-Efficient Prompt Tuning (April 2021)
- Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm (February 2021)
- Calibrate Before Use: Improving Few-Shot Performance of Language Models (February 2021)

- Prefix-Tuning: Optimizing Continuous Prompts for Generation (January 2021)
- Learning to Generate Task-Specific Adapters from Task Description (January 2021)
- Making Pre-trained Language Models Better Few-shot Learners (December 2020)
- Learning from Task Descriptions (November 2020)
- AutoPrompt: Eliciting Knowledge from Language Models with Automatically Generated Prompts (October 2020)
- Language Models are Few-Shot Learners (May 2020)
- How Can We Know What Language Models Know? (July 2020)
- Scaling Laws for Neural Language Models (January 2020)

Applications

- PromptRE: Weakly-Supervised Document-Level Relation Extraction via Prompting-Based Data Programming (October 2023)
- Prompting Large Language Models with Chain-of-Thought for Few-Shot Knowledge Base Question Generation (October 2023)
- Who Wrote it and Why? Prompting Large-Language Models for Authorship Verification (October 2023)
- Promptor: A Conversational and Autonomous Prompt Generation Agent for Intelligent Text Entry Techniques (October 2023)
- Thought Propagation: An Analogical Approach to Complex Reasoning with Large Language Models (October 2023)
- From Sparse to Dense: GPT-4 Summarization with Chain of Density Prompting (September 2023)
- Self-Taught Optimizer (STOP): Recursively Self-Improving Code Generation (October 2023)
- Think before you speak: Training Language Models With Pause Tokens (October 2023)
- (Dynamic) Prompting might be all you need to repair Compressed LLMs (October 2023)
- In-Context Learning in Large Language Models: A Neuroscience-inspired Analysis of Representations (September 2023)
- Understanding In-Context Learning from Repetitions (September 2023)
- Investigating the Efficacy of Large Language Models in Reflective Assessment Methods through Chain of Thoughts Prompting (September 2023)
- Automatic Prompt Rewriting for Personalized Text Generation (September 2023)
- Efficient Streaming Language Models with Attention Sinks (September 2023)
- The Dawn of LMMs: Preliminary Explorations with GPT-4V(ision) (September 2023)

- Graph Neural Prompting with Large Language Models (September 2023)
- Large Language Model Alignment: A Survey (September 2023)
- Enhancing Zero-Shot Chain-of-Thought Reasoning in Large Language Models through Logic (September 2023)
- A Practical Survey on Zero-shot Prompt Design for In-context Learning (September 2023)
- EchoPrompt: Instructing the Model to Rephrase Queries for Improved In-context Learning (September 2023)
- Prompt, Condition, and Generate: Classification of Unsupported Claims with In-Context Learning (September 2023)
- PolicyGPT: Automated Analysis of Privacy Policies with Large Language Models (September 2023)
- LLM4Jobs: Unsupervised occupation extraction and standardization leveraging Large Language Models (September 2023)
- Summarization is (Almost) Dead (September 2023)
- Investigating Zero- and Few-shot Generalization in Fact Verification (September 2023)
- Performance of the Pre-Trained Large Language Model GPT-4 on Automated Short Answer Grading (September 2023)
- Contrastive Decoding Improves Reasoning in Large Language Models (September 2023)
- Struc-Bench: Are Large Language Models Really Good at Generating Complex Structured Data? (September 2023)
- Neural Machine Translation Models Can Learn to be Few-shot Learners (September 2023)
- Chain-of-Thought Reasoning is a Policy Improvement Operator (September 2023)
- ICLEF: In-Context Learning with Expert Feedback for Explainable Style Transfer (September 2023)
- When do Generative Query and Document Expansions Fail? A Comprehensive Study Across Methods, Retrievers, and Datasets (September 2023)
- Using Large Language Models for Knowledge Engineering (LLMKE): A Case Study on Wikidata (September 2023)
- Self-Consistent Narrative Prompts on Abductive Natural Language Inference (September 2023)
- Investigating Answerability of LLMs for Long-Form Question Answering (September 2023)
- PromptTTS++: Controlling Speaker Identity in Prompt-Based Text-to-Speech Using Natural Language Descriptions (September 2023)

- An Empirical Evaluation of Prompting Strategies for Large Language Models in Zero-Shot Clinical Natural Language Processing (September 2023)
- Leveraging Contextual Information for Effective Entity Salience Detection (September 2023)
- Prompting4Debugging: Red-Teaming Text-to-Image Diffusion Models by Finding Problematic Prompts (September 2023)
- PACE: Prompting and Augmentation for Calibrated Confidence Estimation with GPT-4 in Cloud Incident Root Cause Analysis (September 2023)
- From Sparse to Dense: GPT-4 Summarization with Chain of Density Prompting (September 2023)
- Measuring and Improving Chain-of-Thought Reasoning in Vision-Language Models (September 2023)
- Zero-Resource Hallucination Prevention for Large Language Models (September 2023)
- Certifying LLM Safety against Adversarial Prompting (September 2023)
- Improving Code Generation by Dynamic Temperature Sampling (September 2023)
- Prompting a Large Language Model to Generate Diverse Motivational Messages: A Comparison with Human-Written Messages (August 2023)
- Financial News Analytics Using Fine-Tuned Llama 2 GPT Model (August 2023)
- A Study on Robustness and Reliability of Large Language Model Code Generation (August 2023)
- Large Language Models Vote: Prompting for Rare Disease Identification (August 2023)
- WizardMath: Empowering Mathematical Reasoning for Large Language Models via Reinforced Evol-Instruct (August 2023)
- Tree-of-Mixed-Thought: Combining Fast and Slow Thinking for Multi-hop Visual Reasoning (August 2023)
- Graph of Thoughts: Solving Elaborate Problems with Large Language Models (August 2023)
- Red-Teaming Large Language Models using Chain of Utterances for Safety-Alignment (August 2023)
- Boosting Logical Reasoning in Large Language Models through a New Framework: The Graph of Thought (August 2023)
- You Only Prompt Once: On the Capabilities of Prompt Learning on Large Language Models to Tackle Toxic Content (August 2023)
- LLM As DBA (August 2023)
- Interpretable Math Word Problem Solution Generation Via Step-by-step Planning (June 2023)
- In-Context Learning User Simulators for Task-Oriented Dialog Systems (June 2023)

- SQL-PaLM: Improved Large Language ModelAdaptation for Text-to-SQL (June 2023)
- Effective Structured Prompting by Meta-Learning and Representative Verbalizer (June 2023)
- Layout and Task Aware Instruction Prompt for Zero-shot Document Image Question Answering (June 2023)
- Chain-Of-Thought Prompting Under Streaming Batch: A Case Study (June 2023)
- Red Teaming Language Model Detectors with Language Models (May 2023)
- Gorilla: Large Language Model Connected with Massive APIs (May 2023)
- Deliberate then Generate: Enhanced Prompting Framework for Text Generation (May 2023)
- What does the Failure to Reason with "Respectively" in Zero/Few-Shot Settings Tell Us about Language Models? (May 2023)
- ScoNe: Benchmarking Negation Reasoning in Language Models With Fine-Tuning and In-Context Learning (May 2023)
- SheetCopilot: Bringing Software Productivity to the Next Level through Large Language Models (May 2023)
- Grammar Prompting for Domain-Specific Language Generation with Large Language Models (May 2023)
- Mitigating Label Biases for In-context Learning (May 2023)
- Short Answer Grading Using One-shot Prompting and Text Similarity Scoring Model (May 2023)
- Strategic Reasoning with Language Models (May 2023)
- Dissecting Chain-of-Thought: A Study on Compositional In-Context Learning of MLPs (May 2023)
- Marked Personas: Using Natural Language Prompts to Measure Stereotypes in Language Models (May 2023)
- Leveraging Training Data in Few-Shot Prompting for Numerical Reasoning (May 2023)
- Exploring Effectiveness of GPT-3 in Grammatical Error Correction: A Study on Performance and Controllability in Prompt-Based Methods (May 2023)
- NOTABLE: Transferable Backdoor Attacks Against Prompt-based NLP Models (May 2023)
- Tab-CoT: Zero-shot Tabular Chain of Thought (May 2023)
- Evaluating GPT-3 Generated Explanations for Hateful Content Moderation (May 2023)
- Prompt-Guided Retrieval Augmentation for Non-Knowledge-Intensive Tasks (May 2023)
- [Zero- and Few-Shot Event Detection via Prompt-Based Meta Learning]https://arxiv.org/abs/2305.17373) (May 2023)

- Chain-of-Thought Hub: A Continuous Effort to Measure Large Language Models' Reasoning Performance (May 2023)
- Large Language Models Can be Lazy Learners: Analyze Shortcuts in In-Context Learning (May 2023)
- Heterogeneous Value Evaluation for Large Language Models (May 2023)
- PromptNER: Prompt Locating and Typing for Named Entity Recognition (May 2023)
- Small Language Models Improve Giants by Rewriting Their Outputs (May 2023)
- On the Planning Abilities of Large Language Models -- A Critical Investigation (May 2023)
- Beyond Chain-of-Thought, Effective Graph-of-Thought Reasoning in Large Language Models (May 2023)
- PRODIGY: Enabling In-context Learning Over Graphs (May 2023)
- Large Language Models are Few-Shot Health Learners (May 2023)
- Role-Play with Large Language Models (May 2023)
- Measuring Inductive Biases of In-Context Learning with Underspecified Demonstrations (May 2023)
- Fact-Checking Complex Claims with Program-Guided Reasoning (May 2023)
- Large Language Models as Tool Makers (May 2023)
- Iterative Forward Tuning Boosts In-context Learning in Language Models (May 2023)
- SwiftSage: A Generative Agent with Fast and Slow Thinking for Complex Interactive Tasks (May 2023)
- Interactive Natural Language Processing (May 2023)
- An automatically discovered chain-of-thought prompt generalizes to novel models and datasets (May 2023)
- Large Language Model Guided Tree-of-Thought (May 2023)
- Active Retrieval Augmented Generation (May 2023)
- A PhD Student's Perspective on Research in NLP in the Era of Very Large Language Models (May 2023)
- Visual Chain of Thought: Bridging Logical Gaps with Multimodal Infillings (May 2023)
- Mirages: On Anthropomorphism in Dialogue Systems (May 2023)
- Model evaluation for extreme risks (May 2023)
- Language Models Don't Always Say What They Think: Unfaithful Explanations in Chain-of-Thought Prompting (May 2023)
- Cognitive Reframing of Negative Thoughts through Human-Language Model Interaction (May 2023)
- PromptClass: Weakly-Supervised Text Classification with Prompting Enhanced Noise-Robust Self-Training (May 2023)

- Augmented Large Language Models with Parametric Knowledge Guiding (May 2023)
- Aligning Large Language Models through Synthetic Feedback (May 2023)
- Concept-aware Training Improves In-context Learning Ability of Language Models (May 2023)
- FrugalGPT: How to Use Large Language Models While Reducing Cost and Improving Performance (May 2023)
- Enhancing Black-Box Few-Shot Text Classification with Prompt-Based Data Augmentation (May 2023)
- Detecting automatically the layout of clinical documents to enhance the performances of downstream natural language processing (May 2023)
- "Is the Pope Catholic?" Applying Chain-of-Thought Reasoning to Understanding Conversational Implicatures (May 2023)
- Let's Think Frame by Frame: Evaluating Video Chain of Thought with Video Infilling and Prediction (May 2023)
- Generating Data for Symbolic Language with Large Language Models (May 2023)
- Make a Choice! Knowledge Base Question Answering with In-Context Learning (May 2023)
- Improving Language Models via Plug-and-Play Retrieval Feedback (May 2023)
- Multi-Granularity Prompts for Topic Shift Detection in Dialogue (May 2023)
- The CoT Collection: Improving Zero-shot and Few-shot Learning of Language Models via Chain-of-Thought Fine-Tuning (May 2023)
- Can Language Models Understand Physical Concepts? (May 2023)
- Evaluating Factual Consistency of Summaries with Large Language Models (May 2023)
- Dr.ICL: Demonstration-Retrieved In-context Learning (May 2023)
- Probing in Context: Toward Building Robust Classifiers via Probing Large Language Models (May 2023)
- Skill-Based Few-Shot Selection for In-Context Learning (May 2023)
- Exploring Chain-of-Thought Style Prompting for Text-to-SQL (May 2023)
- Enhancing Chat Language Models by Scaling High-quality Instructional Conversations (May 2023)
- On Learning to Summarize with Large Language Models as References (May 2023)
- Learning to Generate Novel Scientific Directions with Contextualized Literature-based Discovery (May 2023)
- Active Learning Principles for In-Context Learning with Large Language Models (May 2023)
- Two Failures of Self-Consistency in the Multi-Step Reasoning of LLMs (May 2023)

- Improving Factuality and Reasoning in Language Models through Multiagent Debate (May 2023)
- ChatCoT: Tool-Augmented Chain-of-Thought Reasoning on\ Chat-based Large Language Models (May 2023)
- WikiChat: A Few-Shot LLM-Based Chatbot Grounded with Wikipedia (May 2023)
- Query Rewriting for Retrieval-Augmented Large Language Models (May 2023)
- Discrete Prompt Optimization via Constrained Generation for Zero-shot Re-ranker (May 2023)
- Element-aware Summarization with Large Language Models: Expert-aligned Evaluation and Chain-of-Thought Method (May 2023)
- Small Language Models Improve Giants by Rewriting Their Outputs (May 2023)
- Prompting and Evaluating Large Language Models for Proactive Dialogues: Clarification, Target-guided, and Non-collaboration (May 2023)
- Prompt-Based Monte-Carlo Tree Search for Goal-Oriented Dialogue Policy Planning (May 2023)
- Mitigating Language Model Hallucination with Interactive Question-Knowledge Alignment (May 2023)
- Making Language Models Better Tool Learners with Execution Feedback (May 2023)
- Text-to-SQL Error Correction with Language Models of Code (May 2023)
- Decomposed Prompting for Machine Translation Between Related Languages using Large Language Models (May 2023)
- SPARSEFIT: Few-shot Prompting with Sparse Fine-tuning for Jointly Generating Predictions and Natural Language Explanations (May 2023)
- "According to ..." Prompting Language Models Improves Quoting from Pre-Training Data (May 2023)
- Prompt-based methods may underestimate large language models' linguistic generalizations (May 2023)
- Chain of Knowledge: A Framework for Grounding Large Language Models with Structured Knowledge Bases (May 2023)
- Measuring Inductive Biases of In-Context Learning with Underspecified Demonstrations (May 2023)
- Automated Few-shot Classification with Instruction-Finetuned Language Models (May 2023)
- Enhancing Few-shot Text-to-SQL Capabilities of Large Language Models: A Study on Prompt Design Strategies (May 2023)
- MvP: Multi-view Prompting Improves Aspect Sentiment Tuple Prediction (May 2023)
- Learning Interpretable Style Embeddings via Prompting LLMs (May 2023)
- Enhancing Small Medical Learners with Privacy-preserving Contextual Prompting (May 2023)

- Fact-Checking Complex Claims with Program-Guided Reasoning (May 2023)
- A Benchmark on Extremely Weakly Supervised Text Classification: Reconcile Seed Matching and Prompting Approaches (May 2023)
- This Prompt is Measuring <MASK>: Evaluating Bias Evaluation in Language Models (May 2023)
- Enhancing Cross-lingual Natural Language Inference by Soft Prompting with Multilingual Verbalizer (May 2023)
- Evaluating Prompt-based Question Answering for Object Prediction in the Open Research Knowledge Graph (May 2023)
- Explaining How Transformers Use Context to Build Predictions (May 2023)
- PiVe: Prompting with Iterative Verification Improving Graph-based Generative Capability of LLMs (May 2023)
- PromptNER: A Prompting Method for Few-shot Named Entity Recognition via k Nearest Neighbor Search (May 2023)
- Logic-LM: Empowering Large Language Models with Symbolic Solvers for Faithful Logical Reasoning (May 2023)
- Enhancing Few-shot NER with Prompt Ordering based Data Augmentation (May 2023)
- Chain-of-thought prompting for responding to in-depth dialogue questions with LLM (May 2023)
- How to Prompt LLMs for Text-to-SQL: A Study in Zero-shot, Single-domain, and Cross-domain Settings (May 2023)
- Evaluation of medium-large Language Models at zero-shot closed book generative question answering (May 2023)
- Few-Shot Dialogue Summarization via Skeleton-Assisted Prompt Transfer (May 2023)
- Can NLP Models Correctly Reason Over Contexts that Break the Common Assumptions? (May 2023)
- Reasoning Implicit Sentiment with Chain-of-Thought Prompting (May 2023)
- Writing your own book: A method for going from closed to open book QA to improve robustness and performance of smaller LLMs (May 2023)
- AutoTrial: Prompting Language Models for Clinical Trial Design (May 2023)
- CRITIC: Large Language Models Can Self-Correct with Tool-Interactive Critiquing (May 2023)
- Controlling the Extraction of Memorized Data from Large Language Models via Prompt-Tuning (May 2023)
- Prompting with Pseudo-Code Instructions (May 2023)
- TrueTeacher: Learning Factual Consistency Evaluation with Large Language Models (May 2023)

- Aligning Instruction Tasks Unlocks Large Language Models as Zero-Shot Relation Extractors (May 2023)
- Exploiting Biased Models to De-bias Text: A Gender-Fair Rewriting Model (May 2023)
- Learning In-context Learning for Named Entity Recognition (May 2023)
- Take a Break in the Middle: Investigating Subgoals towards Hierarchical Script Generation (May 2023)
- TEPrompt: Task Enlightenment Prompt Learning for Implicit Discourse Relation Recognition (May 2023)
- Large Language Models can be Guided to Evade AI-Generated Text Detection (May 2023)
- Temporal Knowledge Graph Forecasting Without Knowledge Using In-Context Learning (May 2023)
- Prompting the Hidden Talent of Web-Scale Speech Models for Zero-Shot Task Generalization (May 2023)
- Think Outside the Code: Brainstorming Boosts Large Language Models in Code Generation (May 2023)
- Improving Language Model Negotiation with Self-Play and In-Context Learning from AI Feedback (May 2023)
- ConvXAI: Delivering Heterogeneous AI Explanations via Conversations to Support Human-AI Scientific Writing (May 2023)
- StructGPT: A General Framework for Large Language Model to Reason over Structured Data (May 2023)
- Towards Expert-Level Medical Question Answering with Large Language Models (May 2023)
- Large Language Models are Built-in Autoregressive Search Engines (May 2023)
- MsPrompt: Multi-step Prompt Learning for Debiasing Few-shot Event Detection (May 2023)
- Exploring the Impact of Layer Normalization for Zero-shot Neural Machine Translation (May 2023)
- SGP-TOD: Building Task Bots Effortlessly via Schema-Guided LLM Prompting (May 2023)
- Multi-modal Visual Understanding with Prompts for Semantic Information Disentanglement of Image (May 2023)
- Soft Prompt Decoding for Multilingual Dense Retrieval (May 2023)
- PaLM 2 Technical Report (May 2023)
- Are LLMs All You Need for Task-Oriented Dialogue? (April 2023)
- HiPrompt: Few-Shot Biomedical Knowledge Fusion via Hierarchy-Oriented Prompting (April 2023)
- Approximating Human Evaluation of Social Chatbots with Prompting (April 2023)

- Automated Reading Passage Generation with OpenAI's Large Language Model (April 2023)
- WebBrain: Learning to Generate Factually Correct Articles for Queries by Grounding on Large Web Corpus (April 2023)
- Prompt Pre-Training with Twenty-Thousand Classes for Open-Vocabulary Visual Recognition (April 2023)
- GPT detectors are biased against non-native English writers (April 2023)
- Zero-Shot Next-Item Recommendation using Large Pretrained Language Models (April 2023)
- Large Language Models as Master Key: Unlocking the Secrets of Materials Science with GPT (April 2023)
- Efficiently Aligned Cross-Lingual Transfer Learning for Conversational Tasks using Prompt-Tuning (April 2023)
- Better Language Models of Code through Self-Improvement (April 2023)
- PromptORE -- A Novel Approach Towards Fully Unsupervised Relation Extraction (April 2023)
- Assessing Language Model Deployment with Risk Cards (April 2023)
- Enhancing Large Language Models with Climate Resources (March 2023)
- BloombergGPT: A Large Language Model for Finance (March 2023)
- Medical Intervention Duration Estimation Using Language-enhanced Transformer Encoder with Medical Prompts (March 2023)
- Soft-prompt tuning to predict lung cancer using primary care free-text Dutch medical notes (March 2023)
- TaskMatrix.AI: Completing Tasks by Connecting Foundation Models with Millions of APIs (March 2023)
- Larger Probes Tell a Different Story: Extending Psycholinguistic Datasets Via In-Context Learning (March 2023)
- Linguistically Informed ChatGPT Prompts to Enhance Japanese-Chinese Machine Translation: A Case Study on Attributive Clauses (March 2023)
- Knowledge-augmented Frame Semantic Parsing with Hybrid Prompt-tuning (March 2023)
- Debiasing Scores and Prompts of 2D Diffusion for Robust Text-to-3D Generation (March 2023)
- Zero-shot Model Diagnosis (March 2023)
- Prompting Large Language Models to Generate Code-Mixed Texts: The Case of South East Asian Languages (March 2023)
- SPeC: A Soft Prompt-Based Calibration on Mitigating Performance Variability in Clinical Notes Summarization (March 2023)
- Large Language Models and Simple, Stupid Bugs (March 2023)

- Can Generative Pre-trained Transformers (GPT) Pass Assessments in Higher Education Programming Courses? (March 2023)
- SelfCheckGPT: Zero-Resource Black-Box Hallucination Detection for Generative Large Language Models (March 2023)
- Large Language Models in the Workplace: A Case Study on Prompt Engineering for Job Type Classification (March 2023)
- ICL-D3IE: In-Context Learning with Diverse Demonstrations Updating for Document Information Extraction (March 2023)
- MathPrompter: Mathematical Reasoning using Large Language Models (March 2023)
- Prompt-Based Learning for Thread Structure Prediction in Cybersecurity Forums (March 2023)
- Choice Over Control: How Users Write with Large Language Models using Diegetic and Non-Diegetic Prompting (March 2023)
- Prompting Large Language Models with Answer Heuristics for Knowledge-based Visual Question Answering (March 2023)
- Soft Prompt Guided Joint Learning for Cross-Domain Sentiment Analysis (March 2023)
- SpeechPrompt v2: Prompt Tuning for Speech Classification Tasks (March 2023)
- Goal Driven Discovery of Distributional Differences via Language Descriptions (February 2023)
- Navigating the Grey Area: Expressions of Overconfidence and Uncertainty in Language Models (February 2023)
- TabGenie: A Toolkit for Table-to-Text Generation (February 2023)
- SGL-PT: A Strong Graph Learner with Graph Prompt Tuning (February 2023)
- Few-Shot Table-to-Text Generation with Prompt-based Adapter (February 2023)
- Language Models Are Few-shot Learners for Prognostic Prediction (February 2023)
- STA: Self-controlled Text Augmentation for Improving Text Classifications (February 2023)
- Check Your Facts and Try Again: Improving Large Language Models with External Knowledge and Automated Feedback (February 2023)
- How Generative AI models such as ChatGPT can be (Mis)Used in SPC Practice, Education, and Research? An Exploratory Study (February 2023)
- Grimm in Wonderland: Prompt Engineering with Midjourney to Illustrate Fairytales (February 2023)
- LabelPrompt: Effective Prompt-based Learning for Relation Classification (February 2023)
- Language Model Crossover: Variation through Few-Shot Prompting (February 2023)

- Prompt Tuning of Deep Neural Networks for Speaker-adaptive Visual Speech Recognition (February 2023)
- The Capacity for Moral Self-Correction in Large Language Models (February 2023)
- Prompting for Multimodal Hateful Meme Classification (February 2023)
- PLACES: Prompting Language Models for Social Conversation Synthesis (February 2023)
- Toolformer: Language Models Can Teach Themselves to Use Tools (February 2023)
- Commonsense-Aware Prompting for Controllable Empathetic Dialogue Generation (February 2023)
- Crawling the Internal Knowledge-Base of Language Models (January 2023)
- Legal Prompt Engineering for Multilingual Legal Judgement Prediction (December 2022)
- Investigating Prompt Engineering in Diffusion Models (November 2022)
- Learn to Explain: Multimodal Reasoning via Thought Chains for Science Question Answering (September 2022)
- Conversing with Copilot: Exploring Prompt Engineering for Solving CS1 Problems Using Natural Language (October 2022)
- Piloting Copilot and Codex: Hot Temperature, Cold Prompts, or Black Magic? (October 2022)
- Plot Writing From Scratch Pre-Trained Language Models (July 2022)
- Survey of Hallucination in Natural Language Generation (February 2022)

Collections

- Chain-of-Thought Papers
- Papers with Code
- Prompt Papers

Last updated on September 19, 2024

Copyright © 2024 DAIR.AI