

Multimedia Appendix: Prompt Templates

This is a Multimedia Appendix to a full manuscript published in the J Med Internet Res. For full copyright and citation information see <http://dx.doi.org/10.2196/jmir.xxxx>

The following are the prompt templates used for each task. Prompts differed depending on whether the LLM was a chat model:

	Basic	Chat
Models	<ul style="list-style-type: none"> • Flan-T5-XXL • Medicine-Llama3-8B • Meditron-7B • MedLLaMA-13B 	<ul style="list-style-type: none"> • GPT-3.5-Turbo • GPT-4 • Llama-3-8B-Instruct • Yi-1.5-34B-Chat • Zephyr-7B-Beta
Example structure	Example: Input: <EXAMPLE-INPUT> Output: <EXAMPLE-OUTPUT>	USER Input: <EXAMPLE-INPUT> Output: ASSISTANT <EXAMPLE-OUTPUT>

For Zephyr-7B-Beta, the system prompt was instead appended to the first user message.

NER

Prompt style	Prompt
Short, basic	I am an excellent linguist and an expert in the biomedical domain. The task is to output a comma-delimited list of <ENTITY_TYPE> entities found in the given sentence. <EXAMPLES> Input: <INPUT> Output:
Long, basic	***INPUT*** The input is a sentence. ***OUTPUT*** The task is to identify all <ENTITY_TYPE> mentions in the input sentence and return them in comma-separated list. ***DOCUMENTATION*** <ENTITY_TYPE>: <DEFINITION> ***EXAMPLES*** <EXAMPLES> ***YOUR TURN*** Input: <INPUT> Output:
Short, chat	SYSTEM You are an expert in the biomedical domain. You are a smart named entity recognition assistant, specialized in extracting mentions of <ENTITY_TYPE> entities from text. Based on the provided sentence, answer with NO_ENTITIES if you find no mentions of <ENTITY_TYPE> entities, otherwise, answer with a

comma-delimited list of the mentions of <ENTITY_TYPE> entities you find. Do not provide any explanation or other characters.

USER

<EXAMPLES>

Input: <INPUT>

Output:

Long, chat

SYSTEM

You are an expert in the biomedical domain. You are a smart named entity recognition assistant, specialized in extracting mentions of <ENTITY_TYPE> entities from text. I will provide you the output format, the definition of the entities you need to extract, and the sentence from which to extract the entities.

USER

Output Format:

Answer with a comma-separated list of <ENTITY_TYPE > entities. If no <ENTITY_TYPE > entities are present in the sentence, answer with NO_ENTITIES. Do not provide any explanation or other characters.

<ENTITY_TYPE>: <DEFINITION>

<EXAMPLES>

USER

Input: <INPUT>

Output:

Entity Definitions

Entity type	Definition
cell	includes cell line (a cell culture developed from a single cell and therefore consisting of cells with a uniform genetic makeup) and cell type (a classification used to identify cells that share morphological or phenotypical features)
chemical	a compound or substance that has been purified or prepared.
disease	a disorder of structure or function in a human, animal, or plant, especially one that has a known cause and a distinctive group of symptoms, signs, or anatomical changes.
gene	a basic unit of heredity and a sequence of nucleotides in DNA.
protein	a large, complex molecule that play many critical roles in a living organism.

PICO

Prompt style	Prompt
Short, basic	I am an excellent linguist and an expert in the biomedical domain. The task is to output a comma-delimited list of clinical trial <ENTITY_TYPE> entities found in the given clinical trial document. Input: <INPUT> Output:
Long, basic	***INPUT*** The input is a clinical trial document. ***OUTPUT***

The task is to identify all mentions of clinical trial <ENTITY_TYPE> in the input clinical trial document and return them in comma-separated list.

DOCUMENTATION

<ENTITY_TYPE>: <DEFINITION>

YOUR TURN

Input: <ABSTRACT>

Output:

Short, chat

SYSTEM

You are an expert in the biomedical domain. You are a smart named entity recognition assistant, specialized in extracting mentions of outcomes of clinical trials from text. Based on the provided clinical trial report, answer with NO_ENTITIES if you find no mentions of clinical trial outcomes, otherwise, answer with a comma-delimited list of the mentions of clinical trial outcomes you find. Do not provide any explanation or other characters.

USER

Input: <INPUT>

Output:

Long, chat

SYSTEM

You are an expert in the biomedical domain. You are a smart named entity recognition assistant, specialized in extracting mentions of outcomes of clinical trials from text. I will provide you the output format, the definition of the entities you need to extract, and the sentence from which to extract the entities.

USER

Output Format:

Answer with a comma-separated list of clinical trial outcomes. If no outcomes are present in the sentence, answer with NO_ENTITIES. Do not provide any explanation or other characters.

<ENTITY>: <DEFINITION>

Input: <INPUT>

Output:

Entity Definitions

Entity type	Definition
clinical trial participant	a person who takes part in a clinical trial.
clinical trial intervention	a potential drug, medical device, activity, or procedure.
clinical trial outcome	the impact that a given intervention or exposure has on the health of clinical trial participants.

ChemProt

Prompt style	Prompt
Short, basic	The task is to classify relations between a chemical labeled as @CHEMICAL\$ and a gene labeled as @GENE\$ for a sentence. The relation must be one of 'CPR:3', 'CPR:4', 'CPR:5', 'CPR:6', 'CPR:9', or 'false'. <EXAMPLES> Input: <INPUT> Output:

Long, basic	<p>***INPUT***</p> <p>The input is a sentence where the chemical is labeled as @CHEMICAL\$ and the gene is labeled as @GENE\$ accordingly in a sentence.</p> <p>***OUTPUT***</p> <p>Your task is to select one out of the six types of relations ('CPR:3', 'CPR:4', 'CPR:5', 'CPR:6', 'CPR:9', and 'false') for the gene and chemical without any explanation or other characters.</p> <p>***DOCUMENTATION***</p> <p><RELATION>, <DEFINITION></p> <p>***EXAMPLES***</p> <p><EXAMPLES></p> <p>***YOUR TURN***</p> <p>Input: <INPUT></p> <p>Output:</p>
Short, chat	<p style="text-align: center;">SYSTEM</p> <p>You are an expert in the biomedical domain. You are a smart relation extraction assistant, specialized in relation extraction between a chemical entity and a gene entity. Classify relations between a chemical labeled as @CHEMICAL\$ and a gene labeled as @GENE\$. The answer must be one of CPR:3, CPR:4, CPR:5, CPR:6, CPR:9, or false. Do not provide any explanation or other characters.</p> <p style="text-align: center;">USER</p> <p><EXAMPLES></p> <p>Input: <INPUT></p> <p>Output:</p>
Long, chat	<p style="text-align: center;">SYSTEM</p> <p>You are an expert in the biomedical domain. You are a smart relation extraction assistant, specialized in relation extraction between a chemical entity and a gene entity. Classify relations between a chemical labeled as @CHEMICAL\$ and a gene labeled as @GENE\$. I will provide you the output format, the definition of class categories, and the text based on which to answer.</p> <p style="text-align: center;">USER</p> <p>Output Format:</p> <p>Your answer must be one out of the six types of relations (CPR:3, CPR:4, CPR:5, CPR:6, CPR:9, and false) for the gene and chemical without any explanation or other characters.</p> <p>Category Definition:</p> <p><RELATION>, <DEFINITION></p> <p><EXAMPLES></p> <p>Input: <INPUT></p> <p>Output:</p>

Relation Definitions

Relation type	Definition
CPR:3	which includes UPREGULATOR, ACTIVATOR, and INDIRECT UPREGULATOR
CPR:4	which includes DOWNREGULATOR, INHIBITOR, and INDIRECT DOWNREGULATOR
CPR:5	which includes AGONIST, AGONIST ACTIVATOR, and AGONIST INHIBITOR
CPR:6	which includes ANTAGONIST

CPR:9	which includes SUBSTRATE, PRODUCT OF and SUBSTRATE PRODUCT OF
false	which indicates no relations

DDI

Prompt style	Prompt
Short, basic	<p>The task is to classify relations between two drugs labeled as @DRUG\$ for a sentence. The relation must be one of 'DDI-effect', 'DDI-mechanism', 'DDI-advise', 'DDI-false', or 'DDI-int'.</p> <p><EXAMPLES> Input: <INPUT> Output:</p>
Long, basic	<p>***INPUT***</p> <p>The input is a sentence where the drugs are labeled as @DRUG\$.</p> <p>***OUTPUT***</p> <p>Your task is to select one out of the five types of relations ('DDI-effect', 'DDI-mechanism', 'DDI-advise', 'DDI-false', and 'DDI-int') for the drugs without any explanation or other characters.</p> <p>***DOCUMENTATION***</p> <p><RELATION>: <DEFINITION></p> <p>***EXAMPLES***</p> <p><EXAMPLES></p> <p>***YOUR TURN***</p> <p>Input: <INPUT> Output:</p>
Short, chat	<p style="text-align: center;">SYSTEM</p> <p>You are an expert in the biomedical domain. You are a smart relation extraction assistant, specialized in relation extraction between drug entities. Classify relations between two drugs labeled as @DRUG\$. The answer must be one of DDI-effect, DDI-mechanism, DDI-advise, DDI-false, or DDI-int. Do not provide any explanation or other characters.</p> <p style="text-align: center;">USER</p> <p><EXAMPLES> Input: <INPUT> Output:</p>
Long, chat	<p style="text-align: center;">SYSTEM</p> <p>You are an expert in the biomedical domain. You are a smart relation extraction assistant, specialized in relation extraction between drug entities. Classify relations between two drugs labeled as @DRUG\$. I will provide you the output format, the definition of class categories, and the text based on which to answer.</p> <p style="text-align: center;">USER</p> <p>Output Format: Your answer must be one out of the five types of relations (DDI-effect, DDI-mechanism, DDI-advise, DDI-false, and DDI-int) for the drugs without any explanation or other characters.</p> <p>Category Definition: <RELATION>: <DEFINITION> <EXAMPLES></p>

Input: <INPUT>

Output:

Relation Definitions

Relation	Definition
DDI-mechanism	This type is used to annotate DDIs that are described by their PK mechanism (e.g. Grepafloxacin may inhibit the metabolism of theobromine)
DDI-effect	This type is used to annotate DDIs describing an effect (e.g. In uninfected volunteers, 46% developed rash while receiving SUSTIVA and clarithromycin) or a PD mechanism (e.g. Chlorthalidone may potentiate the action of other antihypertensive drugs)
DDI-advise	This type is used when a recommendation or advice regarding a drug interaction is given (e.g. UROXATRAL should not be used in combination with other alpha-blockers)
DDI-int	This type is used when a DDI appears in the text without providing any additional information (e.g. The interaction of omeprazole and ketoconazole has been established)
DDI-false	This type is used when no DDI relation appears

GAD

Prompt style	Prompt
Short, basic	The task is to classify relations between a disease labeled as @DISEASE\$ and a gene labeled as @GENE\$ for a sentence. The response should be 1 if there is a relation or 0 if there is not. <EXAMPLES> Input: <INPUT> Output:
Long, basic	***INPUT*** The input is a sentence where the disease is labeled as @DISEASE\$ and the gene is labeled as @GENE\$ accordingly in a sentence. ***OUTPUT*** Your task is to mark the relationship as either true (1) or false (0). ***DOCUMENTATION*** 0: false 1: true ***EXAMPLES*** <EXAMPLES> ***YOUR TURN*** Input: <INPUT> Output:
Short, chat	SYSTEM You are an expert in the biomedical domain. You are a smart relation extraction assistant, specialized in relation extraction between a disease entity and a gene entity. Classify relations between a disease labeled as @DISEASE\$ and a gene labeled as @GENE\$. Answer with 1 if there is a relation or 0 if there is not. Do not provide any explanation or other characters.

	USER
	<EXAMPLES> Input: <INPUT> Output:
Long, chat	SYSTEM
	You are an expert in the biomedical domain. You are a smart relation extraction assistant, specialized in relation extraction between a disease entity and a gene entity. Classify relations between a disease labeled as @DISEASE\$ and a gene labeled as @GENE\$. I will provide you the output format, the definition of class categories, and the text based on which to answer.
	USER
	Output Format: Answer with 1 or 0. Do not provide any explanation or other characters. Category Definition: 0: there is no relation between the disease entity and the gene entity 1: there is a relation between the disease entity and the gene entity <EXAMPLES> Input: <INPUT> Output:

BIOSSES

Prompt style	Prompt
Short, basic	Provide a score from 0 to 4 signifying the similarity between two sentences. Input: sentence1: <SENTENCE1> sentence2: <SENTENCE2> <EXAMPLES> Input: sentence1: <SENTENCE1> sentence2: <SENTENCE2> Output:
Long, basic	***INPUT*** The input is a pair of sentences named sentence1 and sentence2. ***OUTPUT*** The output is the semantic similarity in a continuous number from 0 (no relation) to 4 (equivalent). ***DOCUMENTATION*** <SIMILARITY>, <DEFINITION> ***EXAMPLES*** <EXAMPLES> ***YOUR TURN*** Input: sentence1: <SENTENCE1> sentence2: <SENTENCE2> Output:
Short, chat	SYSTEM
	You are an expert in the biomedical domain. You are a smart semantic similarity scoring assistant, specialized in the scoring of semantic similarity between two sentences. Score the semantic similarity between two sentences named sentence1 and sentence2. Your score must be a number from 0 to 4 signifying the similarity between two sentences. Do not provide any explanation or other characters.
	USER

<EXAMPLES>

Input: sentence1: <SENTENCE1> sentence2: <SENTENCE2>

Output:

Long, chat

SYSTEM

You are an expert in the biomedical domain. You are a smart semantic similarity scoring assistant, specialized in the scoring of semantic similarity between two sentences. Score the semantic similarity between two sentences named sentence1 and sentence2. I will provide you the output format, the definition of the increments on the scoring scale, and the two sentences based on which to answer.

USER

Output Format:

Your answer is the semantic similarity expressed as a continuous number from 0 (no relation) to 4 (equivalent) without any explanation or other characters.

Category Definition:

<SIMILARITY>, <DEFINITION>

<EXAMPLES>

Input: sentence1: <SENTENCE1> sentence2: <SENTENCE2>

Output:

Similarity Definitions

Similarity	Definition
0	the two sentences are on different topics.
1	the two sentences are not equivalent, but are on the same topic.
2	the two sentences are not equivalent, but share some details.
3	the two sentences are roughly equivalent, but some important information differs/missing.
4	The two sentences are completely or mostly equivalent, as they mean the same thing

HoC

Prompt style	Prompt
Short, basic	Provide a comma-separated list classifying the given input with zero, one, or multiple of the following hallmarks of cancer: <ul style="list-style-type: none">- activating invasion and metastasis- avoiding immune destruction- sustaining proliferative signaling- resisting cell death- cellular energetics- genomic instability and mutation- evading growth suppressors- inducing angiogenesis- enabling replicative immortality- tumor promoting inflammation <EXAMPLES> Input: <INPUT>

Long, basic	<p>Output:</p> <p>***INPUT***</p> <p>The input is an abstract text.</p> <p>***OUTPUT***</p> <p>The output should be a comma-separated list, with relevant value for each class. Include the class in the list if the article is related to that class. Please note one article can be related to multiple classes.</p> <p>Example output:</p> <p>"resisting cell death", "evading growth suppressors", "avoiding immune destruction"</p> <p>***DOCUMENTATION***</p> <p>There are 10 cancer hallmarks you will need to decide whether the article is related to, including:</p> <p><CLASS_LIST></p> <p>***EXAMPLES***</p> <p><EXAMPLES></p> <p>***YOUR TURN***</p> <p>Input: <INPUT></p> <p>Output:</p>
Short, chat	<p style="text-align: center;">SYSTEM</p> <p>You are an expert in the biomedical domain. You are a smart document classification assistant, specialized in biomedical document classification. Classify the provided abstract. If the abstract does not belong to any of the hallmarks of cancer, answer with EMPTY_LIST. Otherwise, answer with a comma-separated list consisting of one or multiple of the following hallmarks of cancer:</p> <p><CLASS_LIST></p> <p style="text-align: center;">USER</p> <p><EXAMPLES></p> <p>Abstract: <INPUT></p> <p>Output:</p>
Long, chat	<p style="text-align: center;">SYSTEM</p> <p>You are an expert in the biomedical domain. You are a smart document classification assistant, specialized in biomedical document classification. I will provide you the output format, the definition of class categories, and the abstract to classify.</p> <p style="text-align: center;">USER</p> <p>Output Format:</p> <p>Answer with a comma-separated list consisting of one or multiple hallmarks of cancer.</p> <p>Include the class in the list if the abstract is related to that class. Please note one abstract can be related to multiple classes. If the abstract does not belong to any of the hallmarks of cancer, answer with EMPTY_LIST</p> <p>Category Definition:</p> <p>There are 10 cancer hallmarks you will need to decide whether the article is related to, including:</p> <p><CLASS_LIST></p> <p><EXAMPLES></p>

Abstract: <INPUT>
Output:

Class List

HoC Class

activating invasion and metastasis
sustaining proliferative signaling
resisting cell death
cellular energetics
genomic instability and mutation
evading growth suppressors
inducing angiogenesis
enabling replicative immortality
avoiding immune destruction
tumor promoting inflammation

BioASQ

Prompt style	Prompt
Short, basic	Your task is to answer biomedical questions. Only output yes or no. <EXAMPLES> Input: <INPUT> Output:
Long, basic	***INPUT*** The input is a question. ***OUTPUT*** Answer each question by providing one of the following options: yes, no. ***EXAMPLES*** <EXAMPLES> ***YOUR TURN*** Input: <INPUT> Output:
Short, chat	SYSTEM You are an expert in the biomedical domain. You are a smart question answering assistant, specialized in answering biomedical questions. Only answer with yes, or no. Do not provide any explanation or other characters. USER <EXAMPLES> Input: <INPUT> Output:
Long, chat	SYSTEM You are an expert in the biomedical domain. You are a smart question answering assistant, specialized in answering biomedical questions. I will provide you the output format, and the question to answer. USER

Output Format:
Your answer must be one word: yes, or no. Do not provide any explanation or other characters.
<EXAMPLES>
Input: <INPUT>
Output:

PubMedQA

Prompt style	Prompt
Short, basic	Your task is to answer biomedical questions using the given abstract. Only output yes, no, or maybe as answer. <EXAMPLES> Input: Question: <QUESTION> Abstract: <ABSTRACT> Output:
Long, basic	***INPUT*** The input is a question followed by an abstract. ***OUTPUT*** Answer each question by providing one of the following options: yes, no, maybe. ***EXAMPLES*** <EXAMPLES> ***YOUR TURN*** Input: Question: <QUESTION> Abstract: <ABSTRACT> Output:
Short, chat	SYSTEM You are an expert in the biomedical domain. You are a smart question answering assistant, specialized in answering biomedical questions. Answer based on the provided abstract. Only answer with yes, no, or maybe. Do not provide any explanation or other characters. USER <EXAMPLES> Input: Question: <QUESTION> Abstract: <ABSTRACT> Output:
Long, chat	SYSTEM You are an expert in the biomedical domain. You are a smart question answering assistant, specialized in answering biomedical questions. I will provide you the output format, the question to answer, and the abstract based on which to answer the question. USER Output Format: Your answer must be one word: yes, no, or maybe. Do not provide any explanation or other characters. <EXAMPLES> Input: Question: <QUESTION> Abstract: <ABSTRACT> Output:
