**Appendix C: ChatGPT Prompts**

**Term Similarity**

|  |  |
| --- | --- |
| Prompt | Are "{term1}" and "{term2}" medical synonyms? |
| System Message | You are a helpful assistant. You will answer 1 if two terms are medical synonyms, and 0 if not. |

**Term Subsumption**

*Subtask 1: Generate Nodes for the hierarchy*

|  |  |
| --- | --- |
| Prompt | I want to generate a cardiovascular medical hierarchy using a list of terms. To do this, I want to come up with the top level 'nodes' or terms for which to group my terms under.  I want you to consider all possible terms I have before coming up with the top levels of the medical hierarchy that all my terms will fit under. Because the full list is too long, I will be providing it in chunks.    For sending you the list, I will follow this rule:  [start part 1/10]  Here is list 1/10  [end part 1/10]    Then you just answer, "received part 1/10"    When I tell you "All parts sent", then you can process the data and start generating the top level terms for the hierarchy I will build.    Please output your result as a list, i.e.,  - Group 1  - Group 2  - Group 3..  etc    Remember, even though you are given the chunks separately, you must consider \*\*all terms\*\* |
| System Message | You are a helpful assistant. |

ChatGPT will then respond with “Understood, please send the content”, and we sequentially provide ChatGPT with 25 terms at a time.

*Subtask 2: Organize terms into the nodes*

|  |  |
| --- | --- |
| Prompt | Please place this term: "{term}" into ONE of these categories: {ancestor\_list}.  Output only the category. do not output anything else. |
| System Message | You are a helpful assistant. |

*Subtask 3: Determine hierarchy within each subtree*

|  |  |
| --- | --- |
| Prompt | I am generating a cardiovascular diseases term hierarchy. I have already generated top level groups and subgroups. The top level group I am working on right now is called "{group}", and "{subgroup}" is a node within that group.    I would like to organize the leaves within this subgroup.  Please organize the following terms into a hierarchy.    Put it in this format. If necessary, you can have more depth than this example here.  1. {group}  1.1 {subgroup}  1.1.1 organize the terms here  1.1.1.1 child term of term 1.1.1  1.1.2 sibling term to 1.1.1    Then, output the json format of that hierarchy.    Here are the terms: {branch\_list}  Output only the hierarchy, both in list form and json. do not output anything else. |
| System Message | You are a helpful assistant. |

**Term Description**

|  |  |
| --- | --- |
| Prompt | Are "{term1}" and "{term2}" medical synonyms? |
| System Message | You are a helpful assistant. You will answer 1 if two terms are medical synonyms, and 0 if not. |

**Term Grouping**

|  |  |
| --- | --- |
| Prompt | Please group these terms: {chunk\_str} into these categories: {ancestor\_list}    Use all terms that is provided, and only place them in categories that are provided.  If there is a term that does not belong, you can put it in a category called "other".  No category should have no terms. (ie. do not allow empty categories.)  Output only the json object.  Please break your JSON output into smaller, manageable chunks as needed. Each chunk should represent a logical portion of the grouping.  Indicate which part of the hierarchy you're returning by including the part number, for example:  [start chunk 1/5] ```json ... ``` [end chunk 1/5] |
| System Message | Organize a list of cardiovascular diseases and conditions into meaningful groupings based on how medically similar they are.  # Steps  1. \*\*List Cardiovascular Conditions\*\*: Review the provided list of medical terms.  2. \*\*Categorize Conditions\*\*: Group these conditions based on their medical similarities, such as:  - \*\*Heart Diseases\*\*: Conditions specifically related to heart function.  - \*\*Vascular Diseases\*\*: Conditions affecting blood vessels.  - \*\*Congenital Conditions\*\*: Heart or vessel conditions present at birth.  - \*\*Arrhythmias\*\*: Disorders of heart rhythm.  3. \*\*Refine Groupings\*\*: Ensure accuracy in categorization, adjusting terms to refine group coherence and relevance.  # Output Format  Provide the output as a list of categories, each followed by their grouped cardiovascular conditions. Clearly label each category and list the terms in an organized manner using bullet points.  # Examples  \*\*Example Start\*\*  - \*\*Heart Diseases\*\*:  - Myocardial Infarction  - Coronary Artery Disease  - \*\*Vascular Diseases\*\*:  - Atherosclerosis  - Peripheral Artery Disease  - \*\*Congenital Conditions\*\*:  - Tetralogy of Fallot  - Ventricular Septal Defect  - \*\*Arrhythmias\*\*:  - Atrial Fibrillation  - Ventricular Tachycardia  \*\*Example End\*\*  # Notes  - Ensure terms are verified for correct categorization.  - Explore subcategories if needed for complex conditions to enhance understanding.  - The focus is solely on diseases and conditions. Ensure clarity for medical educational purposes. |