

# **Clinical Trial Query App**

Group member: Zhongyu Cai, Bowen Zhao, Xinyu Zhang

## **Motivation:**

Faced with the overwhelming amount of data in medical research, it can be a complicated task to manually search and select relevant clinical trial information, which could be both time-consuming and labor-intensive. The main motivation behind crafting this app is to efficiently present clinical studies data, thereby aiding individuals, researchers, and healthcare professionals in swiftly pinpointing the clinical trial information they seek. Through this app, we aim to provide easy access to thorough information regarding ongoing clinical trials on various diseases and enable users to promptly filter out content that intrigues them.

## **Overview:**

The Clinical Trial Query App is a comprehensive tool designed to offer users a real-time overview and detailed insights into ongoing clinical trials worldwide. By integrating data from diverse sources, this app provides an interactive and user-friendly interface, allowing users to explore, analyze, and comprehend the ongoing research on various diseases. With a range of features including visual representations and detailed filtering options, this app aspires to be a comprehensive solution for researchers, healthcare professionals, and anyone curious to delve deeper into this field.

## **Feature 1. Data Range Selection:**

**By Bowen Zhao**

### **Motivation:**

There are numerous critical milestones associated with clinical trials, including the start date, end date, and milestones associated with each phase of the trial. The ability to filter clinical trial data based on “date ranges” is an indispensable feature for researchers, clinicians, regulatory bodies, and participants interested in understanding the dynamics and progress of clinical research.

### **Description**

The “Date Range Selection” feature in the Clinical Trials Query is a user interface component that provides the capability to filter clinical trials based on specified start and completion dates. Concerning the “User Interface” part, A `dateRangeInput` UI element allows users to select a date range for the beginning of clinical trials. By default, it spans from 3000 days before the current system date to the current date, ensuring comprehensive coverage of recent and past trials, as well as the Completion Date Range. Concerning the “server” part, the `query_kwds` function is enhanced to incorporate date filtering. When called, it accepts optional `start_date_range` and `completion_date_range` parameters. Overall, the server filters the studies based on the provided keywords and further refines the data set according to the selected start and completion dates.

### **Use of Feature:**

Users begin by typing in relevant keywords into a text input field: "Brief title keywords". These keywords are the terms associated with the titles or conditions of clinical trials they are interested in exploring. After entering keywords, users can specify the exact time periods they are

interested in by setting the start and end dates for both the start date range and the completion date range. Once the keywords and date ranges are input, the app presents the user with a subset of clinical trials that are not only relevant to the entered keywords but also occur within the specified date ranges.

## **Feature 2: Word Cloud**

**By Zhongyu Cai**

### **Motivation:**

Amidst the wealth of information on clinical trials, it's essential to quickly pinpoint which diseases or research areas are garnering focus in a specific timeframe or sponsor type. The Word Cloud feature visually depicts high-frequency words, enabling users to instantaneously identify diseases or research areas that are currently in the spotlight. This feature can also shed light on which diseases are more likely to attract sponsorship.

### **Description:**

The Word Cloud feature, both visually appealing and functional, simplifies the understanding of data. Initially, the "getTermMatrix" function processes the filtered clinical trial data to extract a list denoting the frequency of each word. Subsequently, the Wordcloud functionality crafts a graphic depiction of word frequency. In the generated word cloud, words that are frequently mentioned in the study data are prominently displayed in a larger font, while less common words are prominently displayed in a smaller font. In this way, Word Cloud features not only provide

snapshots of popular research topics, but also help explore datasets efficiently and interactively, saving users from the tedious task of manually sifting through large amounts of textual data.

### **Use of Feature:**

When users visit the app for the first time, they see a word cloud that represents high-frequency words in keywords from the screened studies. Users can interact with various controls located in the left sidebar. Users can enter specific keywords in the "Brief Title Keywords" field, select a specific type of sponsor in the "Sponsor Type" drop-down box, and adjust the "Date range" to focus on trials conducted during a specific period. **It is important for users to click the "Change" button after making a selection to ensure that the word cloud updates accordingly.** In addition, users can fine-tune the word cloud by adjusting the "minimum frequency" and "maximum word count" Settings. This level of customization ensures that the word cloud can be customized to accurately meet the specific needs and preferences of users

## **Feature 3: World Map**

**By Xinyu Zhang**

### **Motivation:**

Given the abundant information, the World Map feature visualizes the countries leading in clinical studies, offering insights into the geographical spread of these trials. This information is invaluable for users wanting to decipher global trends.

**Description:**

The feature generates a map where each country is shaded based on the number of clinical trials conducted. Countries with a high number of studies are represented by darker shadows, allowing users to quickly understand the geographical distribution of research efforts at a glance. Use the “plot\_countries” function to create a map highlighting the amount of research conducted in each country. Based on the country name, the function uses joinCountryData2Map to merge the data set with the Spatial Data Frame (spdf). Finally, the map is drawn using the mapCountryData function. Study quantities are drawn using a gradient color scheme from white to black.

**Use of Feature:**

Users first encounter a world map representing the global distribution of clinical trials. Adjustments can be made via filters on the left sidebar. Hitting the "Change" button is vital to ensure that the world map updates based on the modifications. Users can also refine the display by tweaking the "Minimum Frequency" and "Maximum Number of Words". In addition, users can effectively customize the world map, allowing them to hone the most relevant data and facilitate a more streamlined exploration of the clinical trial landscape.

**Feature 4. American Map**

**By Bowen Zhao**

**motivation:**

Given America’s significant role in global clinical research, a map exclusive to America allows for a more obvious analysis of state-level data. This level of detail is crucial for understanding

regional variations and trends within the U.S. that a world map cannot adequately capture due to scale limitations.

## **Description**

It provides a visual representation of which states in America have higher concentrations of trials. The colour intensity represents the trial frequency—states with more trials are displayed in a darker shade, while those with fewer trials appear lighter. We first group the data by state and count the number of trials per state, creating a summary frame suitable for visualization. The “plot\_states\_us” function within the clinical trials data platform is designed to generate a colour-coded map visualizing the distribution of clinical trials across the various states of the United States. It acts as an analytical tool for professionals who need to understand the geographic distribution of clinical trials within the U.S. quickly.

## **Use of feature:**

Users can navigate to the America Map tab after filtering clinical trials through the sidebar filters, which include Sponsor Type, Start Date Range, and Completion Date Range. After applying their desired filters and clicking the "Change" button, the America Map updates to reflect the distribution of clinical trials based on the adjusted parameters. Just as with the world map, users have the option to adjust the "Minimum Frequency" for what is displayed, allowing for a customized view that emphasizes the states with a trial frequency above a certain threshold.

## **Feature 5. Top 10 Investigated Diseases Pie Chart**

**By Xinyu Zhang**

### **Motivation:**

By distilling the database down to the top 10 conditions, the pie chart allows users to focus on the diseases that have the most impact or prevalence in ongoing research. This can guide stakeholders in identifying key areas for investment, potential gaps in research, and opportunities for new study developments.

### **Description:**

It is a graphical representation that highlights the diseases most frequently studied in clinical trials. It categorizes and sums the occurrences of each disease or medical condition within the database and then displays the ten most prevalent as segments of a pie chart. By ranking diseases according to the number of associated studies and showcasing the top 10, the chart offers a clear, immediate depiction of the current research emphasis within the clinical trial landscape. Each pie segment is colour-coded and labelled with a percentage, and a corresponding legend provides a quick reference, making the chart an efficient tool for identifying predominant areas of medical research.

### **Use of feature:**

Users can navigate to the "Top 10 Investigated Diseases" tab after filtering clinical trials through the sidebar filters, which include Sponsor Type, Start Date Range, and Completion Date Range. After applying their desired filters and clicking the "Change" button, the pie chart updates to

reflect the most researched diseases within the specified criteria. This feature allows users to visually grasp the distribution of research efforts across various medical conditions.

## **Feature 6. Link to Website**

**By Zhongyu Cai**

### **Motivation:**

When users filter through a wealth of information to identify studies of interest, it can be cumbersome to search for relevant research separately in a web browser. The motivation behind implementing the link to website feature is to enhance user convenience. By incorporating this functionality, users can directly navigate to the specific page of the research study, saving time as they delve deeper into the details of the trials that capture their interest.

### **Description:**

In this function, a new column link is created by concatenating the URL with the NCT ID(National Clinical Trial identifier) to form a hyperlink. **Then further modify the link column so that only the NCT ID is displayed if the URL is not available.** The function then selects and renames the specific column to be displayed in the final table: the newly created hyperlink under the NCT ID, along with the brief title, start date, and finish date. This results in a table where NCT IDs are clickable links that direct users to their respective study pages, providing an easy and direct route to more detailed information about each clinical trial.



### **Use of Feature:**

First make adjustments in the left column, such as entering keywords in the "Brief title", selecting "Sponsor type", modifying the "date range", etc. Subsequently, the selected clinical trial information will be displayed in the table at the bottom right of the app. In this table, users can identify studies of interest. If the NCT ID is shown in blue, it means you can link directly to the study's web page. On the other hand, if the NCT ID appears black, it means that there is no direct link and the user needs to manually search for the relevant information using a web browser. **Because many studies do not have a url, please first rank the nct\_id, then you can find those studies with a hyperlink. To open the link, please first open the web page in the browser.** As a result, this feature ensures quick and seamless access to comprehensive data from clinical trials that are of interest to users.