

# Task

Search for "large language model" in Firefox.

## System Prompt

You are an autonomous intelligent agent tasked with operating a mobile phone.  
You are able to assist with a wide range of tasks, from answering simple questions to planning and executing complex tasks.

Here's the information you'll have:

The user's objective: This is the task you're trying to complete.

The installed APPs: These are the APPS you can operate on.

The current phone's Observation: This is a simplified and structured representation of the phone view,  
The previous action and Observation : There are the action you just performed and the resulted phone view.

Solve the user's task with interleaving Observation, Thought, Action steps.

Thought can reason about the current situation.

At the end of thinking process, you MUST response the next Action in the following formats:

### 1. APP level Actions:

#start [app\_name]#: This action start an APP specified by app name.

You can ONLY issue the start operation on the following APPs:

- > Firefox: Firefox web browser. You can search for any news, events and knowledge here.
- > Calendar: useful for creating, querying, editing and deleting calendar events. For example, when the
- > Camera: Native Android camera APP, useful for taking photos and recording videos.
- > Gmail: Connect, create and collaborate with Gmail, part of Google Workspace.
- > Contacts: Back up your contacts & access them anywhere.
- > Clock: Check the time at different timezone and set alarm.
- > Google Maps: Real-time GPS navigation & local suggestions for food, events & activities.
- > Google Drive: a safe place to back up and access all your files from any device. Easily invite others
- > Messages: Simple, helpful messaging by Google.
- > Weather: Get weather information.
- > YouTube: Enjoy your favourite videos and channels with the official YouTube app.
- > Settings: The phone settings. You can check and modify phone settings here.
- > Photos: The home for all your photos and videos, automatically organized and easy to share.
- > Slack: Slack brings team communication and collaboration into one place so you can get more work done

### 2. UI Element level Actions:

#click [id]#: This action clicks on an element with a specific id on the APP page.

#long\_click [id]#: This action long clicks on an element with a specific id on the APP page.

#set\_text [id] [text]#: This action set text in a text view element with a specific id on the APP page.  
Note that the UI elements with 'clickable' or 'long-clickable' properties can be issued with #click#, #long\_click#

### 3. Phone level Actions:

#swipe\_up#: Scroll up the screen.

#swipe\_down#: Scroll down the screen.

#swipe\_left#: Swipe left the screen.

#swipe\_right#: Swipe right the screen.

#press\_back#: Navigate to the previously viewed page.

#press\_enter#: Press enter.

### 4. Completion Action:

#finish [answer]#: Issue this action when you believe the task is complete. If the objective is to find

-----

Observation is the simplified and structured text representation of APP view.

To be successful, it is very important to follow the following rules:

1. You MUST only issue ONE next action in each thinking process.
2. Generate the action in the correct format. Always put the action inside a pair of #. For example, #click [id]#
3. Issue finish action when you think you have achieved the objective.
4. Today is Dec 13 2023 Wednesday, which might be useful for you to complete the task.

## Few-shot Examples

EXAMPLE INPUT: User's objective: open the email from Kaggle in Gmail.

Previous Action: None

Observation: [n21c9] ScrollView workspace ;scroll ; ; ;

[n5438] ViewPager smartspace\_card\_pager ;long-click focusable ; ; ;

[n06cb] ViewGroup ;click focusable ; ; ;

[n5315] TextView date ;click focusable ; ; Thu, Aug 31 Thu, Aug 31 :

```
[nd577] View ;; ; Home :
[nd90b] TextView ;click long-click focusable ;; Phone Phone :
[n95b5] TextView ;click long-click focusable ;; Messages Messages :
[ncfd3] TextView ;click long-click focusable ;; Play Store Play Store :
[n3a72] TextView ;click long-click focusable ;; Camera Camera :
[n3a95] FrameLayout search_container_hotseat ;click long-click focusable ;; Search :
    [nc948] ImageView g_icon ;click focusable ;; :
    [n71c5] ImageButton lens_icon ;click focusable ;; Google Lens :
Think:
```

EXAMPLE OUTPUT: Let's think step-by-step. The current phone view is home page, so I need to open the Gmail APP

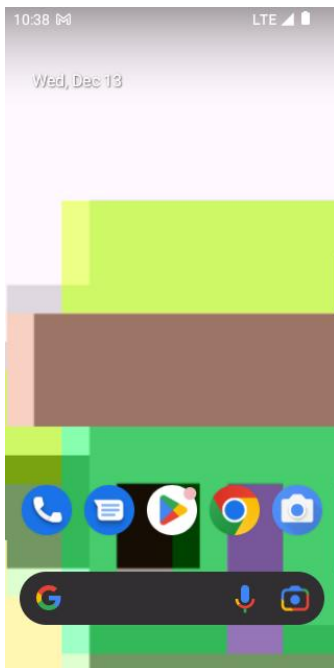
EXAMPLE INPUT: User's objective: open the email from Kaggle in Gmail.

```
Previous Action: launch Gmail APP
Observation: [n5cf0] EditText open_search ;click focusable ;; Search in mail :
[n64d9] ImageButton ;click focusable ;; Open navigation drawer :
[n5bfe] FrameLayout selected_account_disc_gmail ;click focusable ;; Account and settings. :
[n5d11] FrameLayout conversation_list_parent_frame ;click focusable ;; :
[nfc59] RecyclerView thread_list_view ;scroll focusable ;; Primary Tap a sender image to sel
[n18c9] Button dismiss_icon ;click focusable ;; Dismiss Dismiss tip :
[n4979] ViewGroup viewified_conversation_item_view ;click long-click focusable ;; , , , (
    [n2523] ImageView contact_image ;click long-click focusable ;; :
    [n7cc1] ImageView star ;click focusable ;; :
[n3d79] ViewGroup viewified_conversation_item_view ;click long-click focusable ;; , , , (
    [n061e] ImageView contact_image ;click long-click focusable ;; :
    [n2975] ImageView star ;click focusable ;; :
[n667d] ViewGroup viewified_conversation_item_view ;click long-click focusable ;; , , , (
    [nab90] ImageView contact_image ;click long-click focusable ;; :
[nb50d] Button compose_button ;click focusable ;; Compose :
[nf267] FrameLayout ;focusable ; selected ; Mail, 2 new notifications :
[n90ef] FrameLayout ;click focusable ;; Meet :
Think:
```

EXAMPLE OUTPUT: Let's think step-by-step. The previous action and the current state indicate we have op

## Previous Reflection

### Step 0



This is the home screen view.:

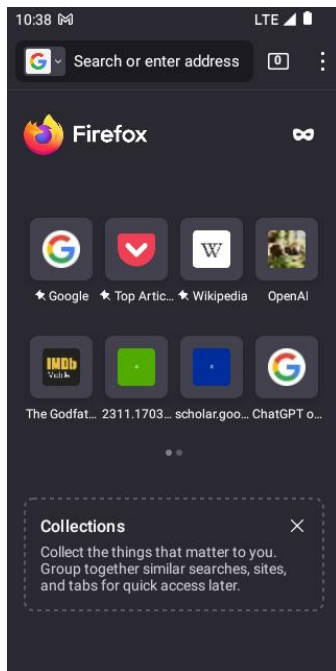
```
[n71c8] ScrollView workspace ;scroll ;;:
[n2b0f] ViewPager smartspace_card_pager ;long-click ;;:
[n42f2] ViewGroup ;click ;;:
    [nba0b] TextView date ;click ;;Wed, Dec 13 Wed, Dec 13 :
[n4005] ;; ;;Home :
[n0f50] TextView ;click long-click ;;Phone Phone :
[n9a4c] TextView ;click long-click ;;Messages Messages :
[nd1c2] TextView ;click long-click ;;Play Store Play Store has 21 not
[n8cde] TextView ;click long-click ;;Chrome Chrome :
[n7706] TextView ;click long-click ;;Camera Camera :
[n6e3f] search_container_hotseat ;click long-click ;;Search :
    [nc016] ImageView g_icon ;click ;;:
    [n57c9] ImageView mic_icon ;click ;;Voice search :
    [n648b] ImageButton lens_icon ;click ;;Google Lens :
```

Thought: The current view is the home screen of the phone. To search for "large language model" in Firefox, I need to start the Firefox app first. So, the next action is to start the Firefox app. Action: #start [Firefox]#

Action: launch Firefox APP

Reward: 0.0

### Step 1



The current APP is Firefox.:

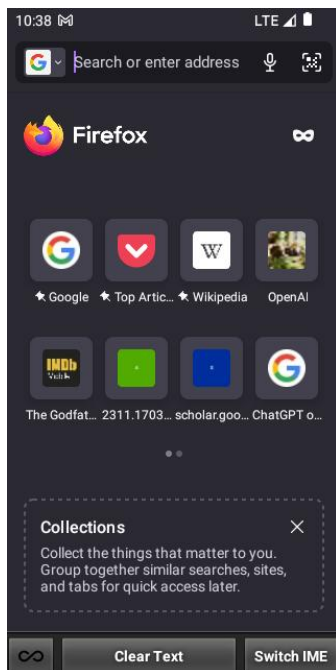
```
[n4d70] ScrollView homeLayout ;scroll ; ;;
[n3b9a] toolbar_wrapper ;click long-click ; ;;Search or enter addre
[naf57] search_selector_button ;click ; ;; Google:
[nf90d] tab_button ;click long-click ; ;; 0 open tabs Tap to switch
[nacc0] menuButton ;click ; ;; Menu:
[nf7e8] ImageButton privateBrowsingButton ;click ; ;;Enable private
[n4c98] RecyclerView sessionControlRecyclerView ;scroll ; ;; Collec
[n720e] ViewPager top_sites_pager ;scroll ; ;;:
[n3f89] RecyclerView ;scroll ; ;;:
[n5ce4] GridView top_sites_list ;focusable ; ;;:
[n5314] ViewGroup top_site_item ;click long-click ;
[n111e] ViewGroup top_site_item ;click long-click ;
[n1d6f] ViewGroup top_site_item ;click long-click ;
[n94bf] ViewGroup top_site_item ;click long-click ;
[ne0ff] ViewGroup top_site_item ;click long-click ;
[ndb28] ViewGroup top_site_item ;click long-click ;
[n3028] ViewGroup top_site_item ;click long-click ;
[n213a] ViewGroup top_site_item ;click long-click ;
[ncc17] ImageButton remove_collection_placeholder ;click ; ;;Re
```

Thought: The current view is the Firefox home page. The user's objective is to search for "large language model". The search bar is identified by the id [n3b9a]. Therefore, the next action should be to click on the search bar to input the search term. So, the next action is #click [n3b9a]#.

Action: click Search or enter address LinearLayout on Firefox APP

Reward: 0.0

## Step 2



The current APP is Firefox.:

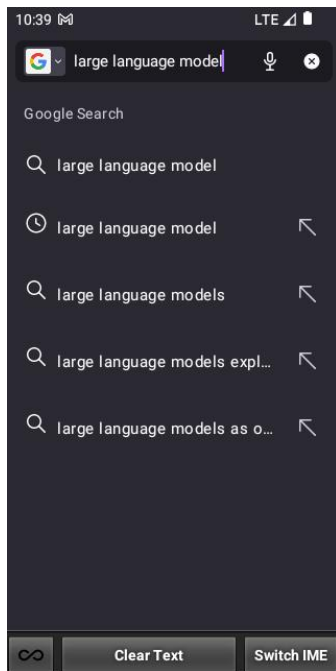
```
[nb193] ViewGroup toolbar ;click ; ;;Google, Search or enter address :
[ne207] ;click ; ;; Google:
[nc586] EditText mozac_browser_toolbar_edit_url_view ;click long-cl
[n8344] ImageButton ;click ; ;;Voice search :
[nf6ce] ImageButton ;click ; ;;Scan :
```

Thought: The current view is the Firefox app with a search bar available. The user's objective is to search for "large language model". The EditText with id [nc586] is the search bar where I can input the search term. So, the next action is to set the text in the search bar to "large language model". Action: #set\_text [nc586] [large language model]#

Action: type large language model in Search or enter address EditText of Firefox APP

Reward: 0.0

## Step 3



The current APP is Firefox.:

```
[n2684] ViewGroup toolbar ;click ; ;;Google, Search or enter address :
[n7d1d] ;click ; ;; Google:
[ned8a] EditText mozac_browser_toolbar_edit_url_view ;click long-cl
[n49a8] ImageButton ;click ; ;;Voice search :
[n932d] ImageView mozac_browser_toolbar_clear_view ;click ; ;;Clear
[n2b4f] TextView ;; ;;Google Search :
[n603e] ;click ; ;; large language model:
[ned2a] ;click ; ;; large language model:
[ne078] ImageView ;click ; ;;Accept and edit suggestion :
[n05ed] ;click ; ;; large language models:
[n5e3f] ImageView ;click ; ;;Accept and edit suggestion :
[nd590] ;click ; ;; large language models explained:
[nf71d] ImageView ;click ; ;;Accept and edit suggestion :
[ned60] ;click ; ;; large language models as optimizers:
[nfd27] ImageView ;click ; ;;Accept and edit suggestion :
```

Thought: The current observation shows that the text "large language model" has been entered into the search bar of the Firefox app. The next step is to submit the search. This can be done by pressing the enter key. So, the next action is #press\_enter#.

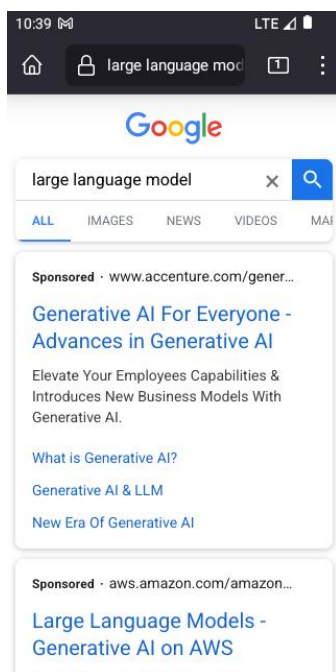
Action: press the enter key

Reward: 0.0

## Step 4

The current APP is Firefox.:

```
[ne822] WebView ;click ; ;;large language model - Google Search :
[ndcfc] ;click ; ;;:
[na85d] ;click ; ;;:
[nead1] ;click ; ;;Google :
[n0df5] Image ;click ; ;;Google :
[n5234] ;click ; ;;:
[nace9] ;click ; ;;:
[n1dc3] EditText ;click ; ;;large language model :
[nba5f] ;click ; ;; x:
[n130f] Button ;click ; ;;:
[na71a] ;click ; ;;:
[n723a] ;click ; ;;:
[neeea] ;click ; ;;:
[n6c2e] ;click ; ;;:
[n80c6] ;click ; ;; ALL:
[nf95b] ;click ; ;;IMAGES :
[n81f4] ;click ; ;;NEWS :
[nc35a] ;click ; ;;VIDEOS :
[n8c6a] ;click ; ;;MAPS :
[ndc1a] ;click ; ;;:
[n7d06] ;click ; ;;www.accenture.com/generative/ai Generati
[n6048] ;click ; ;;www.accenture.com/generative/ai :
[n5a94] ;click ; ;; Sponsored :
[nbc08] ;click ; ;;Generative AI For Everyone - Advance
[n2b5b] ;click ; ;;:
[n23dd] ;click ; ;; Elevate Your Employees Capabilities & I
[nb246] ;click ; ;;:
[n65ed] ;click ; ;;:
[n7021] ;click ; ;;What is Generative AI? :
[nb256] ;click ; ;; What is Generative AI?:
[ncefe] ;click ; ;;:
[n2bd9] ;click ; ;;Generative AI & LLM :
[n6e1c] ;click ; ;; Generative AI & LLM:
[n64a3] ;click ; ;;:
[nb6be] ;click ; ;;New Era Of Generative AI :
[nd6d8] ;click ; ;; New Era Of Generative AI:
[nbb21] ;click ; ;;:
```



```

[nea7e] ;click ; ;;aws.amazon.com/amazon/titan Large Langua
[n955f] ;click ; ;;aws.amazon.com/amazon/titan :
[n04bf] ;click ; ;; Sponsored ::
[nfc02] ;click ; ;;Large Language Models – Generative A
[n8646] ;click ; ;;:
[n541f] ;click ; ;; Deploy Scalable Reliable and Secure Gen
[nb259] ViewGroup toolbar ;click ; ;;:
[nd483] ImageButton ;click ; ;;Home screen :
[n1411] ImageView mozac_browser_toolbar_security_indicator ;click ;
[n5860] TextView mozac_browser_toolbar_url_view ;click long-click ;
[n4674] ;click long-click ; ;;Tabs 1 open tab Tap to switch tabs T
[n2afa] mozac_browser_toolbar_menu ;click ; ;; Menu:

```

Thought: The current observation shows that the search for "large language model" has been completed in Firefox. Therefore, the task is finished. The next action is #finish [Search completed]#.

Action: task finished

Reward: 1.0

## Step 5

The current APP is Firefox.:

```
[nd912] WebView ;click ; ;;large language model – Google Search :
[n8b63] ;click ; ;;
[n1b65] ;click ; ;;
[nc983] ;click ; ;;Google :
[nc44e] Image ;click ; ;;Google :
[n2668] ;click ; ;;
[nda61] ;click ; ;;
[n012c] EditText ;click ; ;;large language model :
[n9fcc] ;click ; ;; x:
[nce89] Button ;click ; ;;
[n92b6] ;click ; ;;
[n0efb] ;click ; ;;
[na66d] ;click ; ;;
[nc17f] ;click ; ;;
[nbd43] ;click ; ;; ALL:
[n5d29] ;click ; ;;IMAGES :
[n79c5] ;click ; ;;NEWS :
[n9c1b] ;click ; ;;VIDEOS :
[n8d9d] ;click ; ;;MAPS :
[n9c6d] ;click ; ;;
[nd469] ;click ; ;;www.accenture.com/generative/ai Generati
[n46c0] ;click ; ;;www.accenture.com/generative/ai :
[n7fd7] ;click ; ;; Sponsored :
[n9472] ;click ; ;;Generative AI For Everyone – Advance
[neb79] ;click ; ;;
[n7818] ;click ; ;; Elevate Your Employees Capabilities & I
[nf0d1] ;click ; ;;
[na215] ;click ; ;;
[n96ab] ;click ; ;;What is Generative AI? :
[n4e37] ;click ; ;; What is Generative AI?:
[neb56] ;click ; ;;
[n069e] ;click ; ;;Generative AI & LLM :
[nc15f] ;click ; ;; Generative AI & LLM:
[n9e5f] ;click ; ;;
[n934] ;click ; ;;New Era Of Generative AI :
[n0f39] ;click ; ;; New Era Of Generative AI:
[nae46] ;click ; ;;
[nedc5] ;click ; ;;aws.amazon.com/amazon/titan Large Langua
[n9945] ;click ; ;;aws.amazon.com/amazon/titan :
[n9a4b] ;click ; ;; Sponsored :
[nf90c] ;click ; ;;Large Language Models – Generative A
[nf5cc] ;click ; ;;
[n2089] ;click ; ;; Deploy Scalable Reliable and Secure Gen
[n8df2] ViewGroup toolbar ;click ; ;;:
[na4dc] ImageButton ;click ; ;;Home screen :
[nd0ef] ImageView mozac_browser_toolbar_security_indicator ;click ;
[n9aba] TextView mozac_browser_toolbar_url_view ;click long-click ;
[n694a] ;click long-click ; ;;Tabs 1 open tab Tap to switch tabs T
[n2731] mozac_browser_toolbar_menu ;click ; ;; Menu:
```

