

KoboldAI API

Overview

Tags

info

Metadata about this API

generate

Text generation endpoints

model

Information about the current text generation model

story

Endpoints for managing the story in the KoboldAI GUI

world_info

Endpoints for managing the world info in the KoboldAI GUI

config

Allows you to get/set various setting values

Paths

GET /config/authors_note Retrieve the current author's note setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":""}</pre>	No Links

PUT /config/authors_note Set author's note setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/authors_note_depth Retrieve the current author's note depth setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":3}</pre>	No Links

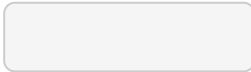
PUT /config/authors_note_depth Set author's note depth setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/authors_note_template Retrieve the current author's note template setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":"[Author's note:</pre>	>]"} 

PUT /config/authors_note_template Set author's note template setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/frmtadnsnp Retrieve the current add sentence spacing (input formatting) setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":false}</pre>	No Links

PUT /config/frmtadsnsp Set add sentence spacing (input formatting) setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/frmtrmblln Retrieve the current remove blank lines (output formatting) setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":false}</pre>	No Links

PUT /config/frmtrmblln Set remove blank lines (output formatting) setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/frmtrmspch Retrieve the current remove special characters (output formatting) setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":false}</pre>	No Links

PUT /config/frmtrmspch Set remove special characters (output formatting) setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/frmttriminc Retrieve the current trim incomplete sentences (output formatting) setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":false}</pre>	No Links

PUT /config/frmttriminc Set trim incomplete sentences (output formatting) setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/max_context_length Retrieve the current max context length setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":2048}</pre>	No Links

PUT /config/max_context_length Set max context length setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/max_length Retrieve the current max length setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":80}</pre>	No Links

PUT /config/max_length Set max length setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/memory Retrieve the current memory setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":"Memory"}</pre>	No Links

PUT /config/memory Set memory setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/n Retrieve the current Gens Per Action setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":1}</pre>	No Links

PUT /config/n Set Gens Per Action setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/sampler_full_determinism Retrieve the current sampler full determinism setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":false}</pre>	No Links

PUT /config/sampler_full_determinism Set sampler full determinism setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/sampler_order Retrieve the current sampler order setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":[6,0,1,2,3,4,5]}</pre>	No Links

PUT /config/sampler_order Set sampler order setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/sampler_seed Retrieve the current global sampler seed value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":3475097509890965500}</pre>	No Links

PUT /config/sampler_seed Set the global sampler seed value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/singleline Retrieve the current single line (output formatting) setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":false}</pre>	No Links

PUT /config/singleline Set single line (output formatting) setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/soft_prompt Retrieve the current soft prompt name

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":""}</pre>	No Links

PUT /config/soft_prompt Set soft prompt by name

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/soft_prompts_list Retrieve all available softprompt filenames

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"values": []}</pre>	No Links

GET /config/temperature Retrieve the current temperature setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value": 0.5}</pre>	No Links

PUT /config/temperature Set temperature setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/tfs Retrieve the current tail free sampling setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":1.0}</pre>	No Links

PUT /config/tfs Set tail free sampling setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/top_a Retrieve the current top-a sampling setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":0.0}</pre>	No Links

PUT /config/top_a Set top-a sampling setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/top_k Retrieve the current top-k sampling setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":0}</pre>	No Links

PUT /config/top_k Set top-k sampling setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/top_p Retrieve the current top-p sampling setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":0.9}</pre>	No Links

PUT /config/top_p Set top-p sampling setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/typical Retrieve the current typical sampling setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":1.0}</pre>	No Links

PUT /config/typical Set typical sampling setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /config/world_info_depth Retrieve the current world info depth setting value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"value":3}</pre>	No Links

PUT /config/world_info_depth Set world info depth setting to specified value

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

POST /generate Generate text

Generates text given a submission, sampler settings, soft prompt and number of return sequences.

By default, the story, userscripts, memory, author's note and world info are disabled.

Unless otherwise specified, optional values default to the values in the KoboldAI GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"results":[{"text":" Holding up his tail to keep it from dragging in the dirty snow that covered the cobblestone, he waited patiently for the butcher to turn his attention from his stall so that he could pilfer his next meal: a tender-looking chicken."}]}</pre>	No Links
422	Validation error <i>Content</i> application/json	No Links

Code	Description	Links
501	<p>Not implemented</p> <p><i>Content</i></p> <p>application/json</p> <p><i>Example</i></p> <pre data-bbox="312 387 1190 562">{"detail":{"msg":"API generation is not supported in read-only mode; please load a model and then try again.,"type":"not_implemented"}}</pre>	No Links
503	<p>Server is busy</p> <p><i>Content</i></p> <p>application/json</p> <p><i>Example</i></p> <pre data-bbox="312 804 1190 940">{"detail":{"msg":"Server is busy; please try again later.,"type":"service_unavailable"}}</pre>	No Links

Code	Description	Links
507	<p data-bbox="268 165 478 201">Out of memory</p> <p data-bbox="268 241 370 273"><i>Content</i></p> <p data-bbox="268 291 501 327">application/json</p> <p data-bbox="268 340 639 376">cpu.default_cpu_allocator</p> <p data-bbox="306 389 421 421"><i>Example</i></p> <div data-bbox="312 439 1190 689" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre data-bbox="338 470 1152 658">{"detail":{"msg":"KoboldAI ran out of memory: DefaultCPUAllocator: not enough memory: you tried to allocate 209715200 bytes.", "type":"out_of_memory.cpu.default_cpu_allocator"}}}</pre> </div> <p data-bbox="268 725 402 761">gpu.cuda</p> <p data-bbox="306 775 421 806"><i>Example</i></p> <div data-bbox="312 824 1190 1075" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre data-bbox="338 855 1152 1043">{"detail":{"msg":"KoboldAI ran out of memory: CUDA out of memory. Tried to allocate 20.00 MiB (GPU 0; 4.00 GiB total capacity; 2.97 GiB already allocated; 0 bytes free; 2.99 GiB reserved in total by PyTorch)", "type":"out_of_memory.gpu.cuda"}}}</pre> </div> <p data-bbox="268 1111 379 1146">gpu.hip</p> <p data-bbox="306 1160 421 1191"><i>Example</i></p> <div data-bbox="312 1209 1190 1460" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre data-bbox="338 1240 1152 1429">{"detail":{"msg":"KoboldAI ran out of memory: HIP out of memory. Tried to allocate 20.00 MiB (GPU 0; 4.00 GiB total capacity; 2.97 GiB already allocated; 0 bytes free; 2.99 GiB reserved in total by PyTorch)", "type":"out_of_memory.gpu.hip"}}}</pre> </div> <p data-bbox="268 1496 395 1532">tpu.hbm</p> <p data-bbox="306 1545 421 1576"><i>Example</i></p> <div data-bbox="312 1594 1190 1845" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre data-bbox="338 1626 1104 1814">{"detail":{"msg":"KoboldAI ran out of memory: Compilation failed: Compilation failure: Ran out of memory in memory space hbm. Used 8.83G of 8.00G hbm. Exceeded hbm capacity by 848.88M.", "type":"out_of_memory.tpu.hbm"}}}</pre> </div> <p data-bbox="268 1881 549 1917">unknown.unknown</p> <p data-bbox="306 1930 421 1962"><i>Example</i></p> <div data-bbox="312 1980 1190 2096" style="border: 1px solid #ccc; padding: 10px; background-color: #f9f9f9;"> <pre data-bbox="338 2011 887 2047">{"detail":{"msg":"KoboldAI ran out of</pre> </div>	No Links

GET /info/version Current API version

Returns the version of the API that you are currently using.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"result":"1.0.0"}</pre>	No Links

GET /info/version/latest Latest API version

Returns the latest API version available.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"result":"1.0.0"}</pre>	No Links

GET /info/version/list List API versions

Returns a list of available API versions sorted in ascending order.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"results":["1.0.0"]}</pre>	No Links

GET /model Retrieve the current model string

Gets the current model string, which is shown in the title of the KoboldAI GUI in parentheses, e.g. "KoboldAI Client (KoboldAI/fairseq-dense-13B-Nerys-v2)".

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json <i>Example</i> <pre>{"result":"KoboldAI/fairseq-dense-13B-Nerys-v2"}</pre>	No Links

PUT /model Load a model

Loads a model given its Hugging Face model ID, the path to a model folder (relative to the "models" folder in the KoboldAI root folder) or "ReadOnly" for no model.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links
503	Server is busy <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"Server is busy; please try again later.,"type":"service_unavailable"}}</pre>	No Links

GET /story Retrieve the entire story

Returns the entire story currently shown in the KoboldAI GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

DELETE /story Clear the story

Starts a new blank story.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
503	Server is busy <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"Server is busy; please try again later.,"type":"service_unavailable"}}</pre>	No Links

GET /story/end Retrieve the last action of the story

Returns the last action of the story in the KoboldAI GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

Code	Description	Links
510	Story is empty <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"Could not retrieve the last action of the story because the story is empty.", "type":"story_empty"}} </pre>	No Links

POST /story/end Add an action to the end of the story

Inserts a single action at the end of the story in the KoboldAI GUI without generating text.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links
503	Server is busy <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"Server is busy; please try again later.", "type":"service_unavailable"}} </pre>	No Links

POST /story/end/delete Remove the last action of the story

Removes the last action of the story in the KoboldAI GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links
503	Server is busy <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"Server is busy; please try again later.,"type":"service_unavailable"}}</pre>	No Links
510	Story too short <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"Could not delete the last action of the story because the number of actions in the story is less than or equal to 1.,"type":"story_too_short"}}</pre>	No Links

GET /story/end/num Retrieve the num of the last action of the story

Returns the **num** of the last action of the story in the KoboldAI GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

Code	Description	Links
510	Story is empty <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"Could not retrieve the last action of the story because the story is empty.,"type":"story_empty"}} </pre>	No Links

GET /story/end/text Retrieve the text of the last action of the story

Returns the text of the last action of the story in the KoboldAI GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
510	Story is empty <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"Could not retrieve the last action of the story because the story is empty.,"type":"story_empty"}} </pre>	No Links

PUT /story/end/text Set the text of the last action of the story

Sets the text of the last action of the story in the KoboldAI GUI to the desired value.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links
510	Story is empty <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"Could not retrieve the last action of the story because the story is empty.,"type":"story_empty"}}</pre>	No Links

PUT /story/load Load a story

Loads a story given its filename (without the .json).

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links
503	Server is busy <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"Server is busy; please try again later.,"type":"service_unavailable"}}</pre>	No Links

GET /story/nums Retrieve a list of the nums of the chunks in the current story

Returns the `num`s of the story chunks currently shown in the KoboldAI GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

GET /story/nums/{num} Determine whether or not there is a story chunk with the given num

Parameters

Type	Name	Description	Schema
path	num <i>required</i>	num of the desired story chunk.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

PUT /story/save Save the current story

Saves the current story given its destination filename (without the .json).

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /story/{num} Retrieve a story chunk

Returns information about a story chunk given its **num**.

Parameters

Type	Name	Description	Schema
path	num <i>required</i>	num of the desired story chunk.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No chunk with the given num exists.", "type":"key_error"}}</pre>	No Links

DELETE /story/{num} Remove a story chunk

Removes a story chunk from the story in the KoboldAI GUI given its **num**. Cannot be used to delete the first action (the prompt).

Parameters

Type	Name	Description	Schema
path	num <i>required</i>	num of the desired story chunk. Must be larger than or equal to 1.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

Code	Description	Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"No chunk with the given num exists.,"type":"key_error"}} </pre>	No Links
503	Server is busy <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"Server is busy; please try again later.,"type":"service_unavailable"}} </pre>	No Links

GET /story/{num}/text Retrieve the text of a story chunk

Returns the text inside a story chunk given its `num`.

Parameters

Type	Name	Description	Schema
path	num <i>required</i>	<code>num</code> of the desired story chunk.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"No chunk with the given num exists.,"type":"key_error"}} </pre>	No Links

PUT /story/{num}/text Set the text of a story chunk

Sets the text inside a story chunk given its **num**.

Parameters

Type	Name	Description	Schema
path	num <i>required</i>	num of the desired story chunk.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No chunk with the given num exists.", "type":"key_error"}}</pre>	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /world_info Retrieve all world info entries

Returns all world info entries currently shown in the KoboldAI GUI.

The **folders** are sorted in the same order as they are in the GUI and the **entries** within the folders and within the parent **result** object are all sorted in the same order as they are in their respective parts of the GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

GET /world_info/folders Retrieve all world info folders

Returns details about all world info folders currently shown in the KoboldAI GUI.

The **folders** are sorted in the same order as they are in the GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

POST /world_info/folders Create a new world info folder at the end of the world info

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /world_info/folders/none Retrieve all world info entries not in a folder

Returns all world info entries that are not in a world info folder.

The **entries** are sorted in the same order as they are in the KoboldAI GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

POST /world_info/folders/none Create a new world info entry outside of a world info folder, at the end of the world info

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /world_info/folders/none/uids Retrieve the UIDs of all world info entries not in a folder

Returns the `uid`s of all world info entries that are not in a world info folder.

The **entries** are sorted in the same order as they are in the KoboldAI GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

GET /world_info/folders/none/uids/{uid} Determine whether or not there is a world info entry with the given UID that is not in a world info folder

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

GET /world_info/folders/uids Retrieve the UIDs all world info folders

Returns the `uid`s of all world info folders currently shown in the KoboldAI GUI.

The **folders** are sorted in the same order as they are in the GUI.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

GET /world_info/folders/{folder_uid}/uids/{entry_uid} Determine whether or not there is a world info entry with the given UID in the world info folder with the given UID

Parameters

Type	Name	Description	Schema
path	entry_uid <i>required</i>	uid of the desired world info entry.	integer
path	folder_uid <i>required</i>	uid of the desired world info folder.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

GET /world_info/folders/{uid} Retrieve all world info entries in the given folder

Returns all world info entries that are in the world info folder with the given **uid**.

The **entries** are sorted in the same order as they are in the KoboldAI GUI.

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info folder.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info folder with the given uid exists.", "type":"key_error"}}</pre>	No Links

POST /world_info/folders/{uid} Create a new world info entry at the end of the world info folder with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info folder.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info folder with the given uid exists.,"type":"key_error"}}</pre>	No Links
422	Validation error <i>Content</i> application/json	No Links

***DELETE* /world_info/folders/{uid}** Delete the world info folder with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info folder.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info folders with the given uid exists.,"type":"key_error"}}</pre>	No Links

GET /world_info/folders/{uid}/name Retrieve the name of the world info folder with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info folder.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info folder with the given uid exists.,"type":"key_error"}}</pre>	No Links

PUT /world_info/folders/{uid}/name Set the name of the world info folder with the given UID to the specified value

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info folder.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

Code	Description	Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"No world info folder with the given uid exists.,"type":"key_error"}} </pre>	No Links
422	Validation error <i>Content</i> application/json	No Links

GET `/world_info/folders/{uid}/uids` Retrieve the UIDs of all world info entries in the given folder

Returns the `uid`'s of all world info entries that are in the world info folder with the given `uid`.

The `entries` are sorted in the same order as they are in the KoboldAI GUI.

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	<code>uid</code> of the desired world info folder.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"No world info folder with the given uid exists.,"type":"key_error"}} </pre>	No Links

GET /world_info/uids Retrieve the UIDs of all world info entries

Returns in a similar format as GET /world_info except only the `uid`s are returned.

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

GET /world_info/uids/{uid} Determine whether or not there is a world info entry with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

GET /world_info/{uid} Retrieve information about the world info entry with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

Code	Description	Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}} </pre>	No Links

DELETE /world_info/{uid} Delete the world info entry with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}} </pre>	No Links

GET /world_info/{uid}/comment Retrieve the comment of the world info entry with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}}</pre>	No Links

PUT /world_info/{uid}/comment Set the comment of the world info entry with the given UID to the specified value

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}}</pre>	No Links

Code	Description	Links
422	Validation error <i>Content</i> application/json	No Links

GET /world_info/{uid}/constant Retrieve the constant mode state of the world info entry with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}}</pre>	No Links

PUT /world_info/{uid}/constant Set the constant mode state of the world info entry with the given UID to the specified value

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}}</pre>	No Links
422	Validation error <i>Content</i> application/json	No Links

***GET* /world_info/{uid}/content** Retrieve the content of the world info entry with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}}</pre>	No Links

***PUT* /world_info/{uid}/content** Set the content of the world info entry with the given UID to the specified value

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}}</pre>	No Links
422	Validation error <i>Content</i> application/json	No Links

***GET* /world_info/{uid}/key** Retrieve the keys or primary keys of the world info entry with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info entry with the given uid exists. ","type":"key_error"}}</pre>	No Links

***PUT* /world_info/{uid}/key** Set the keys or primary keys of the world info entry with the given UID to the specified value

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info entry with the given uid exists. ","type":"key_error"}}</pre>	No Links
422	Validation error <i>Content</i> application/json	No Links

GET /world_info/{uid}/keysecondary Retrieve the secondary keys of the world info entry with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}}</pre>	No Links

PUT /world_info/{uid}/keysecondary Set the secondary keys of the world info entry with the given UID to the specified value

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links

Code	Description	Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}} </pre>	No Links
422	Validation error <i>Content</i> application/json	No Links

***GET* /world_info/{uid}/selective** Retrieve the selective mode state of the world info entry with the given UID

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre> {"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}} </pre>	No Links

PUT /world_info/{uid}/selective Set the selective mode state of the world info entry with the given UID to the specified value

Parameters

Type	Name	Description	Schema
path	uid <i>required</i>	uid of the desired world info entry.	integer

Responses

Code	Description	Links
200	Successful request <i>Content</i> application/json	No Links
404	Not found <i>Content</i> application/json <i>Example</i> <pre>{"detail":{"msg":"No world info entry with the given uid exists.,"type":"key_error"}}</pre>	No Links
422	Validation error <i>Content</i> application/json	No Links

Components

Schemas

AddSentenceSpacingSettings

Properties

Name	Description	Schema
<i>value required</i>		boolean

AuthorsNoteDepthSetting

Properties

Name	Description	Schema
<i>value required</i>	Maximum: 5 Minimum: 1	integer

AuthorsNoteSetting

Properties

Name	Description	Schema
<i>value required</i>		string

AuthorsNoteTemplateSetting

Properties

Name	Description	Schema
<i>value required</i>		string

BasicBoolean

Properties

Name	Description	Schema
<i>value required</i>		boolean

BasicBooleanResult

Properties

Name	Description	Schema
<i>result required</i>		boolean

BasicError

Properties

Name	Description	Schema
<i>msg required</i>		string
<i>type required</i>		string

BasicResult

Properties

Name	Description	Schema
<i>result required</i>		BasicResultInner

BasicResultInner

Properties

Name	Description	Schema
result <i>required</i>		string

BasicResults

Properties

Name	Description	Schema
results <i>required</i>		< BasicResultInner > array

BasicString

Properties

Name	Description	Schema
value <i>required</i>		string

BasicUID

Properties

Name	Description	Schema
uid <i>required</i>	32-bit signed integer unique to this world info entry/folder. Maximum: 2147483647 Minimum: -2147483648	integer

Empty

GenerationInput

Properties

Name	Description	Schema
disable_input_formatting <i>optional</i>	When enabled, all input formatting options default to false instead of the value in the KoboldAI GUI Default: true	boolean
disable_output_formatting <i>optional</i>	When enabled, all output formatting options default to false instead of the value in the KoboldAI GUI. Default: true	boolean

Name	Description	Schema
frmtadsnsp <i>optional</i>	<p>Input formatting option. When enabled, adds a leading space to your input if there is no trailing whitespace at the end of the previous action.</p> <p>If <code>disable_input_formatting</code> is <code>true</code>, this defaults to <code>false</code> instead of the value in the KoboldAI GUI.</p>	boolean
frmtrmbln <i>optional</i>	<p>Output formatting option. When enabled, replaces all occurrences of two or more consecutive newlines in the output with one newline.</p> <p>If <code>disable_output_formatting</code> is <code>true</code>, this defaults to <code>false</code> instead of the value in the KoboldAI GUI.</p>	boolean
frmtrmspch <i>optional</i>	<p>Output formatting option. When enabled, removes <code>`#/@%{}+=~</code></p>	<p><code>\^<>`</code> from the output.</p> <p>If <code>disable_output_formatting</code> is <code>true</code>, this defaults to <code>false</code> instead of the value in the KoboldAI GUI.</p>

Name	Description	Schema
boolean	<i>frmtriminc optional</i>	<p>Output formatting option. When enabled, removes some characters from the end of the output such that the output doesn't end in the middle of a sentence. If the output is less than one sentence long, does nothing.</p> <p>If <code>disable_output_formatting</code> is <code>true</code>, this defaults to <code>false</code> instead of the value in the KoboldAI GUI.</p>
boolean	<i>max_context_length optional</i>	<p>Maximum number of tokens to send to the model.</p> <p>Maximum: 2048 Minimum: 512</p>
integer	<i>max_length optional</i>	<p>Number of tokens to generate.</p> <p>Maximum: 512 Minimum: 1</p>

Name	Description	Schema
integer	<i>n optional</i>	Number of outputs to generate. Maximum: 5 Minimum: 1
integer	<i>prompt required</i>	This is the submission.
string	<i>quiet optional</i>	When enabled, Generated output will not be displayed in the console.
boolean	<i>rep_pen optional</i>	Base repetition penalty value. Minimum: 1
number	<i>rep_pen_range optional</i>	Repetition penalty range. Minimum: 0
integer	<i>rep_pen_slope optional</i>	Repetition penalty slope. Minimum: 0

Name	Description	Schema
number	sampler_full_determinism <i>optional</i>	<p>If enabled, the generated text will always be the same as long as you use the same RNG seed, input and settings. If disabled, only the sequence of generated texts that you get when repeatedly generating text will be the same given the same RNG seed, input and settings.</p>
boolean	sampler_order <i>optional</i>	<p>Sampler order to be used. If N is the length of this array, then N must be greater than or equal to 6 and the array must be a permutation of the first N non-negative integers.</p> <p>Minimum Items: 6</p>

Name	Description	Schema
< integer array >	sampler_seed <i>optional</i>	<p>RNG seed to use for sampling. If not specified, the global RNG will be used.</p> <p>Maximum: 18446744073709551615</p> <p>Minimum: 0</p>
integer	singleline <i>optional</i>	<p>Output formatting option. When enabled, removes everything after the first line of the output, including the newline.</p> <p>If <code>disable_output_formatting</code> is <code>true</code>, this defaults to <code>false</code> instead of the value in the KoboldAI GUI.</p>
boolean	soft_prompt <i>optional</i>	<p>Soft prompt to use when generating. If set to the empty string or any other string containing no non-whitespace characters, uses no soft prompt.</p>

Name	Description	Schema
string	temperature <i>optional</i>	Temperature value.
number	tfs <i>optional</i>	Tail free sampling value. Maximum: 1 Minimum: 0
number	top_a <i>optional</i>	Top-a sampling value. Minimum: 0
number	top_k <i>optional</i>	Top-k sampling value. Minimum: 0
integer	top_p <i>optional</i>	Top-p sampling value. Maximum: 1 Minimum: 0
number	typical <i>optional</i>	Typical sampling value. Maximum: 1 Minimum: 0
number	use_authors_note <i>optional</i>	Whether or not to use the author's note from the KoboldAI GUI when generating text. This has no effect unless <code>use_story</code> is also enabled. Default: false

Name	Description	Schema
boolean	<i>use_memory optional</i>	Whether or not to use the memory from the KoboldAI GUI when generating text. Default: false
boolean	<i>use_story optional</i>	Whether or not to use the story from the KoboldAI GUI when generating text. Default: false
boolean	<i>use_userscripts optional</i>	Whether or not to use the userscripts from the KoboldAI GUI when generating text. Default: false
boolean	<i>use_world_info optional</i>	Whether or not to use the world info from the KoboldAI GUI when generating text. Default: false

GenerationOutput

Properties

Name	Description	Schema
results <i>required</i>	Array of generated outputs.	< GenerationResult > array

GenerationResult

Properties

Name	Description	Schema
text <i>required</i>	Generated output as plain text.	string

GensPerActionSetting

Properties

Name	Description	Schema
value <i>required</i>	Maximum: 5 Minimum: 0	integer

MaxContextLengthSetting

Properties

Name	Description	Schema
value <i>required</i>	Maximum: 2048 Minimum: 512	integer

MaxLengthSetting

Properties

Name	Description	Schema
value <i>required</i>	Maximum: 512 Minimum: 1	integer

MemorySetting

Properties

Name	Description	Schema
value <i>required</i>		string

ModelSelection

Properties

Name	Description	Schema
model <i>required</i>	Hugging Face model ID, the path to a model folder (relative to the "models" folder in the KoboldAI root folder) or "ReadOnly" for no model	string

NotFoundError

Properties

Name	Description	Schema
detail <i>required</i>		BasicError

NotImplementedError

Properties

Name	Description	Schema
detail <i>required</i>		BasicError

OutOfMemoryError

Properties

Name	Description	Schema
detail <i>required</i>		BasicError

RemoveBlankLinesSettings

Properties

Name	Description	Schema
value <i>required</i>		boolean

RemoveSpecialCharactersSettings

Properties

Name	Description	Schema
value <i>required</i>		boolean

SamplerFullDeterminismSetting

Properties

Name	Description	Schema
value <i>required</i>		boolean

SamplerOrderSetting

Properties

Name	Description	Schema
value required	Minimum Items: 6	< integer > array

SamplerSeedSetting

Properties

Name	Description	Schema
value required	Maximum: 18446744073709551615 Minimum: 0	integer

ServerBusyError

Properties

Name	Description	Schema
detail required		BasicError

SingleLineSettings

Properties

Name	Description	Schema
value required		boolean

SoftPromptSetting

Properties

Name	Description	Schema
value required	Soft prompt name, or a string containing only whitespace for no soft prompt. If using the GET method and no soft prompt is loaded, this will always be the empty string.	string

SoftPromptsList

Properties

Name	Description	Schema
values required	Array of available softprompts.	< SoftPromptSetting > array

Story

Properties

Name	Description	Schema
results <i>required</i>	Array of story actions. The array is sorted such that actions closer to the end of this array are closer to the end of the story.	< StoryChunk > array

StoryChunk

Properties

Name	Description	Schema
num <i>required</i>	Guaranteed to not equal the num of any other active story chunk. Equals 0 iff this is the first action of the story (the prompt).	integer
text <i>required</i>	The text inside this story chunk.	string

StoryChunkNum

Properties

Name	Description	Schema
value <i>required</i>		integer

StoryChunkResult

Properties

Name	Description	Schema
result <i>required</i>		StoryChunk

StoryChunkSetText

Properties

Name	Description	Schema
value <i>required</i>		string

StoryChunkText

Properties

Name	Description	Schema
value <i>required</i>		string

StoryEmptyError

Properties

Name	Description	Schema
------	-------------	--------

Name	Description	Schema
detail <i>required</i>		BasicError

StoryLoad

Properties

Name	Description	Schema
name <i>required</i>		string

StorySave

Properties

Name	Description	Schema
name <i>required</i>		string

StoryTooShortError

Properties

Name	Description	Schema
detail <i>required</i>		BasicError

SubmissionInput

Properties

Name	Description	Schema
disable_input_f ormatting <i>optional</i>	When enabled, disables all input formatting options, overriding their individual enabled/disabled states. Default: true	boolean
frmtadnsnp <i>optional</i>	Input formatting option. When enabled, adds a leading space to your input if there is no trailing whitespace at the end of the previous action.	boolean
prompt <i>required</i>	This is the submission.	string

TailFreeSamplingSetting

Properties

Name	Description	Schema
value <i>required</i>	Maximum: 1 Minimum: 0	number

TemperatureSamplingSetting

Properties

Name	Description	Schema
<i>value required</i>		number

TopASamplingSetting

Properties

Name	Description	Schema
<i>value required</i>	Minimum: 0	number

TopKSamplingSetting

Properties

Name	Description	Schema
<i>value required</i>	Minimum: 0	integer

TopPSamplingSetting

Properties

Name	Description	Schema
<i>value required</i>	Maximum: 1 Minimum: 0	number

TrimIncompleteSentencesSettings

Properties

Name	Description	Schema
<i>value required</i>		boolean

TypicalSamplingSetting

Properties

Name	Description	Schema
<i>value required</i>	Maximum: 1 Minimum: 0	number

ValidationError

Properties

Name	Description	Schema
<i>detail required</i>		object

WorldInfo

Properties

Name	Description	Schema
<i>entries required</i>		< WorldInfoEntry > array
<i>folders required</i>		< WorldInfoFolder > array

WorldInfoDepthSetting

Properties

Name	Description	Schema
<i>value required</i>	Maximum: 5 Minimum: 1	integer

WorldInfoEntries

Properties

Name	Description	Schema
<i>entries required</i>		< WorldInfoEntry > array

WorldInfoEntriesUIDs

Properties

Name	Description	Schema
<i>entries required</i>		< integer > array

WorldInfoEntry

Properties

Name	Description	Schema
<i>comment required</i>	The comment/description/title for this world info entry.	string

Name	Description	Schema
<i>constant required</i>	Whether or not constant mode is enabled for this world info entry.	boolean
<i>content required</i>	The "What To Remember" for this entry.	string
<i>key required</i>	Comma-separated list of keys, or of primary keys if selective mode is enabled.	string
<i>keysecondary optional</i>	Comma-separated list of secondary keys if selective mode is enabled.	string
<i>selective required</i>	Whether or not selective mode is enabled for this world info entry.	boolean
<i>uid required</i>	32-bit signed integer unique to this world info entry. Maximum: 2147483647 Minimum: -2147483648	integer

WorldInfoFolder

Properties

Name	Description	Schema
<i>entries required</i>		< WorldInfoEntry > array
<i>name required</i>	Name of this world info folder.	string
<i>uid required</i>	32-bit signed integer unique to this world info folder. Maximum: 2147483647 Minimum: -2147483648	integer

WorldInfoFolderBasic

Properties

Name	Description	Schema
<i>name required</i>	Name of this world info folder.	string
<i>uid required</i>	32-bit signed integer unique to this world info folder. Maximum: 2147483647 Minimum: -2147483648	integer

WorldInfoFolderUIDs

Properties

Name	Description	Schema
entries <i>required</i>		< integer > array
uid <i>required</i>	32-bit signed integer unique to this world info folder. Maximum: 2147483647 Minimum: -2147483648	integer

WorldInfoFolders

Properties

Name	Description	Schema
folders <i>required</i>		< WorldInfoFolderBasic > array

WorldInfoFoldersUIDs

Properties

Name	Description	Schema
folders <i>required</i>		< integer > array

WorldInfoUIDs

Properties

Name	Description	Schema
entries <i>required</i>		< integer > array
folders <i>required</i>		< WorldInfoFolderUIDs > array