CLAIMS

1. A computer-implemented method comprising:

maintaining a plurality of low-rank adaptation (LoRA) models that each include a set of weights configured to modify parameters of a large-language model (LLM) to cause the LLM to generate text having a corresponding property;

presenting a set of manipulable user-interface controls to allow configuration of properties of text generated using the LLM and wherein output of the LLM is to be modified based on one of more of the LoRA models of the plurality of LoRA models selected based on a state of the user-interface controls as manipulated; and

providing a preview of LLM output corresponding to a current state of the user-interface controls during presentation and manipulation thereof, wherein providing the preview includes iteratively providing a prompt to the LLM and outputting the output of the LLM responsive to the prompt for each iteration, wherein, for each iteration, the output of the LLM is modified based on one or more of the LoRA models of the plurality of LORA models selected based on the current state of the user-interface controls as manipulated.

2. The computer-implemented method of claim 1, further comprising:

identifying a first LoRA model of the plurality of LORA models based on the current state of the user-interface controls as manipulated, wherein the first LoRA model corresponds to a first property of text generated by the LLM; and

identifying a range of degrees of a second property of text generated by the LLM based on a LoRA model corresponding to a degree in the identified range being composable with the first LoRA model;

wherein presenting the set of manipulable user-interface controls comprises displaying indications of degrees of the second property in the identified range as selectable and displaying indications of degrees of the second property that are not in the identified range as not selectable.

3. The computer-implemented method of claim 1, wherein presenting the set of user-interface controls comprises:

identifying a first LoRA model of the plurality of LORA models based on the current state of the user-interface controls as manipulated, wherein the first LoRA model corresponds to a first property of text generated by the LLM;

identifying a subset of the plurality of LORA models based on the first LoRA model; and

presenting user-interface controls associated with the properties corresponding to the identified subset of LORA models that are manipulable to select one or more of the corresponding properties.

4. The computer-implemented method of claim 1, wherein presenting the set of user-interface controls comprises:

presenting a first user-interface control that is manipulable to change a degree of a first property of the text generated by the LLM;

wherein manipulation of the set of user-interface controls a selection of the degree of the first property.

5. The computer-implemented method of claim 1, wherein presenting the set of user-interface controls comprises:

presenting a first set of user-interface controls that are manipulable to select respective properties of text within a first category of properties; and

in response to receiving manipulation of a user-interface control of the first set of user-interface controls to select a property from among the first category of properties, presenting a second set of user-interface controls that are manipulable to select respective properties of text within a second category of properties.

6. The computer-implemented method of claim 5, wherein the properties within the second category of properties are determined based on the selected property from among the first category of properties.

7. The computer-implemented method of claim 1, wherein the set of user-interface controls are presented to allow configuration of properties of text generated by the LLM for output by a chat agent.

8. The computer-implemented method of claim 7, further comprising:

storing identifiers of the plurality of LoRA models in a repository or instruction set that is accessible to the chat agent during execution of the chat agent.

9. The computer-implemented method of claim 1, wherein the set of user-interface controls are output for display by a browser application, and wherein the method further comprises:

storing identifiers of the plurality of LORA models in the browser application.

10. The computer-implemented method of claim 1:

wherein providing the prompt to the LLM comprises specifying an order in which the one or more LoRA models selected based on the current state of the user-interface controls as manipulated are to be composed.

11. The computer-implemented method of claim 1, further comprising:

training the plurality of LORA models based on a set of chat conversations that have been tagged as containing the property corresponding to each of the plurality of LORA models.

12. A non-transitory computer readable storage medium storing executable instructions, execution of which by a processor causing the processor to:

maintain a plurality of low-rank adaptation (LoRA) models that each include a set of weights configured to modify parameters of a large-language model (LLM) to cause the LLM to generate text having a corresponding property;

present a set of manipulable user-interface controls to allow configuration of properties of text generated using the LLM and wherein output of the LLM is to be modified based on one of more of the LoRA models of the plurality of LoRA models selected based on a state of the user-interface controls as manipulated; and

provide a preview of LLM output corresponding to a current state of the user-interface controls during presentation and manipulation thereof, wherein providing the preview includes iteratively providing a prompt to the LLM and outputting the output of the LLM responsive to the prompt for each iteration, wherein, for each iteration, the output of the LLM is modified based on one or more of the LoRA models of the plurality of LORA models selected based on the current state of the user-interface controls as manipulated.

13. The non-transitory computer readable storage medium of claim 12, wherein execution of the instructions further causes the processor to:

identify a first LoRA model of the plurality of LORA models based on the current state of the user-interface controls as manipulated, wherein the first LoRA model corresponds to a first property of text generated by the LLM; and

identify a range of degrees of a second property of text generated by the LLM based on a LoRA model corresponding to a degree in the identified range being composable with the first LoRA model;

wherein presenting the set of manipulable user-interface controls comprises displaying indications of degrees of the second property in the identified range as selectable and displaying indications of degrees of the second property that are not in the identified range as not selectable.

14. The non-transitory computer readable storage medium of claim 12, wherein presenting the set of user-interface controls comprises:

identifying a first LoRA model of the plurality of LoRA models based on the current state of the user-interface controls as manipulated, wherein the first LoRA model corresponds to a first property of text generated by the LLM;

identifying a subset of the plurality of LORA models based on the first LoRA model; and

presenting user-interface controls associated with the properties corresponding to the identified subset of LORA models that are manipulable to select one or more of the corresponding properties.

15. The non-transitory computer readable storage medium of claim 12, wherein presenting the set of user-interface controls comprises:

presenting a first user-interface control that is manipulable to change a degree of a first property of the text generated by the LLM;

wherein manipulation of the set of user-interface controls a selection of the degree of the first property.

16. The non-transitory computer readable storage medium of claim 12, wherein presenting the set of user-interface controls comprises:

presenting a first set of user-interface controls that are manipulable to select respective properties of text within a first category of properties; and

in response to receiving manipulation of a user-interface control of the first set of user-interface controls to select a property from among the first category of properties, presenting a second set of user-interface controls that are manipulable to select respective properties of text within a second category of properties.

17. The non-transitory computer readable storage medium of claim 12, wherein the set of user-interface controls are presented to allow configuration of properties of text generated by the LLM for output by a chat agent.

18. The non-transitory computer readable storage medium of claim 17, wherein execution of the instructions further causes the processor to:

storing identifiers of the plurality of LORA models in a repository or instruction set that is accessible to the chat agent during execution of the chat agent.

19. The non-transitory computer readable storage medium of claim 12:

wherein providing the prompt to the LLM comprises specifying an order in which the one or more LoRA models selected based on the current state of the user-interface controls as manipulated are to be composed.

20. A system comprising:

at least one hardware processor; and

at least one non-transitory memory storing instructions, which, when executed by the at least one hardware processor, cause the system to:

maintain a plurality of low-rank adaptation (LoRA) models that each include a set of weights configured to modify parameters of a large-language model (LLM) to cause the LLM to generate text having a corresponding property;

present a set of manipulable user-interface controls to allow configuration of properties of text generated using the LLM and wherein output of the LLM is to be modified based on one of more of the LoRA models of the plurality of LORA models selected based on a state of the user-interface controls as manipulated; and

provide a preview of LLM output corresponding to a current state of the user-interface controls during presentation and manipulation thereof, wherein providing the preview includes iteratively providing a prompt to the LLM and outputting the output of the LLM responsive to the prompt for each iteration, wherein, for each iteration, the output of the LLM is modified based on one or more of the LoRA models of the plurality of LoRA models selected based on the current state of the user-interface controls as manipulated.