CLAIMS:

1. A method of keyword determination, comprising:

presenting an interface comprising a plurality of information fields associated with an object;

in response to detecting completion of input in a first information field among the plurality of information fields, determining at least one candidate keyword for the object based on first information input in the first information field, the first information comprising at least one of: text or visual information; and

presenting the at least one candidate keyword for selection into at least one of the plurality of information fields.

2. The method of claim 1, further comprising:

in response to detecting completion of input in a second information field among the plurality of information fields, updating the at least one candidate keyword based on second information input in the second information field and the first information; and

replacing the presenting of the at least one candidate keyword with presenting of the at least one updated candidate keyword.

3. The method of claim 1, further comprising:

in response to detecting an update of the first information in the first information field, updating the at least one candidate keyword based on the updated first information; and

replacing the presenting of the at least one candidate keyword with presenting of the at least one updated candidate keyword.

4. The method of any of claims 1 to 3, wherein the first information comprises both the text and the visual information, and wherein determining at least one candidate keyword comprises:

extracting a text feature representation of the text and a visual feature representation of the visual information, respectively;

determining a fused feature representation for the object by fusing the text feature information and the visual feature information; and

determining the at least one candidate keyword based on the fused feature representation.

5. The method of claim 4, wherein determining the fused feature representation comprises:

performing an attention mechanism on the text feature representation and the visual feature representation, to obtain an interaction matrix between the text feature representation and the visual feature representation;

determining a text weight corresponding to the text feature representation and a visual weight corresponding to the visual feature representation based on the interaction matrix; and

performing a weighted fusion on the text feature representation and the visual feature representation based on the text weight and the visual weight, to obtain the fused feature.

6. The method of claim 4, wherein determining the at least one candidate keyword based on the fused feature representation comprises:

determining respective similarities between the object and respective keywords in a set of keywords based on the fused feature representation and respective text feature representations of the respective keywords; and

determining the at least one candidate keyword from the respective keywords based on a ranking result of the respective similarities.

7. The method of claim 6, wherein at least part of the set of keywords are generated by a content generative model.

8. The method of any of claims 1 to 3, further comprising:

in response to detecting an attribute value input in an attribute filed among the plurality of information fields, supplementing, using a language model, the attribute value; and

determining the at least one candidate keyword comprises:

determining the at least one candidate keyword based on the first information and the supplemented attribute value.

9. The method of any of claims 1 to 3, further comprising:

in response to the selection, inputting the selected keyword into the at least one of the plurality of information fields.

10. The method of any of claims 1 to 3, wherein the plurality of information fields comprise at least one of: a title field for the object, or a description field for the object.

11. An electronic device, comprising:

at least one processing unit; and

at least one memory coupled to the at least one processing unit and storing instructions executable by the at least one processing unit, the instructions, upon execution by the at least one processing unit, causing the electronic device to perform the method of any of claims 1 to 10.

12. A non-transitory computer readable storage medium having computer executable instructions stored thereon, the computer executable instructions, when executed by an electronic device, causing the electronic device perform operations comprising:

presenting an interface comprising a plurality of information fields associated with an object,

in response to detecting completion of input in a first information field among the plurality of information fields, determining at least one candidate keyword for the object based on first information input in the first information field, the first information comprising at least one of: text or visual information; and

presenting the at least one candidate keyword for selection into at least one of the plurality of information fields.