

What is claimed is:

1. A method for topic early warning, comprising:

acquiring a self-defined keyword;

calculating similarity between the self-defined keyword and each word in a corpus, and

5 acquiring extended keywords related to the self-defined keyword from the corpus according to the similarity;

selecting one or more target keywords from the extended keywords according to a type of the extended keywords and similarity between the extended keywords and the self-defined keyword, and adding the target keywords to a target keyword list;

10 performing real-time monitoring according to the target keywords in the target keyword list; and performing topic early warning when it is monitored that the number of topics corresponding to the target keywords reaches a preset threshold.

2. The method of claim 1, wherein the selecting one or more target keywords from the extended keywords according to a type of the extended keywords and similarity between the extended keywords and the self-defined keyword, and adding the target keywords to a target keyword list comprises:

classifying the extended keywords in accordance with a preset type;

15 selecting, from each type of extended keywords, top h extended keywords having the highest similarity with the self-defined keyword as the target keywords, wherein h is a positive integer greater than 0;

20 aggregating the target keywords selected from each type, and generating the target keyword list for monitoring.

25 3. The method of claim 1, wherein after the acquiring a self-defined keyword, the method further comprises:

calculating a word vector corresponding to the self-defined keyword;

wherein the calculating similarity between the self-defined keyword and each word in a corpus, and acquiring extended keywords related to the self-defined keyword from the corpus according to the similarity comprises:

calculating similarity between the word vector of the self-defined keyword and a word vector of
5 each word in the corpus;

acquiring the extended keywords related to the self-defined keyword from the corpus according to the similarity between the word vectors.

4. The method of claim 1, wherein the calculating similarity between the self-defined keyword
10 and each word in a corpus, and acquiring extended keywords related to the self-defined keyword from the corpus according to the similarity comprises:

calculating the similarity between the self-defined keyword and each word in the corpus by using the Pearson Correlation Coefficient method;

acquiring top K words having the highest similarity with the self-defined keyword as the
15 extended keywords of the self-defined keyword, wherein K is a positive integer greater than 0.

5. The method of claim 1, wherein the performing real-time monitoring according to the target keywords in the target keyword list comprises:

performing real-time monitoring for each target keyword in the target keyword list in a form of a
20 sliding window.

6. An apparatus for topic early warning, comprising:

a self-defined keyword acquiring module, configured to acquire a self-defined keyword;

an extended keyword acquiring module, configured to calculate similarity between the
25 self-defined keyword and each word in a corpus, and acquire extended keywords related to the self-defined keyword from the corpus according to the similarity;

a target keyword selecting module, configured to select one or more target keywords from the

extended keywords according to a type of the extended keywords and similarity between the extended keywords and the self-defined keyword, and add the target keywords to a target keyword list;

a monitoring module, configured to perform real-time monitoring according to the target

5 keywords in the target keyword list; and

an early warning module, configured to perform topic early warning when it is monitored that the number of topics corresponding to the target keywords reaches a preset threshold.

7. The apparatus of claim 6, wherein the target keyword selecting module comprises:

10 a classifying module, configured to classify the extended keywords in accordance with a preset type;

a selecting module, configured to select, from each type of extended keywords, top h extended keywords having the highest similarity with the self-defined keyword as the target keywords, wherein h is a positive integer greater than 0;

15 an aggregating module, configured to aggregate the target keywords selected from each type, and generate the target keyword list for monitoring.

8. The apparatus of claim 6, wherein the apparatus further comprises:

a calculating module, configured to calculate a word vector corresponding to the self-defined

20 keyword; and

the extended keyword acquiring module is further configured to calculate similarity between the word vector of the self-defined keyword and a word vector of each word in the corpus, and acquire the extended keywords related to the self-defined keyword from the corpus according to the similarity between the word vectors.

25

9. The apparatus of claim 6, wherein the extended keyword acquiring module is further configured to calculate the similarity between the self-defined keyword and each word in the

corpus by using the Pearson Correlation Coefficient method, and acquire top K words having the highest similarity with the self-defined keyword as the extended keywords of the self-defined keyword, wherein K is a positive integer greater than 0.

5 10. The apparatus of claim 6, wherein the early warning module is further configured to perform real-time monitoring for each target keyword in the target keyword list in a form of a sliding window.

11. A computer equipment, comprising a memory and a processor, the memory storing
10 computer-readable instructions, the computer-readable instructions being executed by the processor to cause the processor to execute following steps:

acquiring a self-defined keyword;

calculating similarity between the self-defined keyword and each word in a corpus, and
acquiring extended keywords related to the self-defined keyword from the corpus according to

15 the similarity;

selecting one or more target keywords from the extended keywords according to a type of the
extended keywords and similarity between the extended keywords and the self-defined keyword,
and adding the target keywords to a target keyword list;

performing real-time monitoring according to the target keywords in the target keyword list; and

20 performing topic early warning when it is monitored that the number of topics corresponding to
the target keywords reaches a preset threshold.

12. The computer equipment of claim 11, wherein the step of selecting one or more target
keywords from the extended keywords according to a type of the extended keywords and
25 similarity between the extended keywords and the self-defined keyword, and adding the target
keywords to a target keyword list, which is executed by the processor, comprises:

classifying the extended keywords in accordance with a preset type;

selecting top h extended keywords having the highest similarity with the self-defined keyword from each type of extended keywords as the target keywords, wherein h is a positive integer greater than 0;

aggregating the target keywords selected from each type, and generating the target keyword list
5 for monitoring.

13. The computer equipment of claim 11, wherein after the step of acquiring a self-defined keyword, the processor is further configured to execute following steps:

calculating a word vector corresponding to the self-defined keyword;

10 wherein calculating similarity between the self-defined keyword and each word in a corpus, and acquiring extended keywords related to the self-defined keyword from the corpus according to the similarity comprises:

calculating similarity between the word vector of the self-defined keyword and a word vector of each word in the corpus;

15 acquiring the extended keywords related to the self-defined keyword from the corpus according to the similarity between the word vectors.

14. The computer equipment of claim 11, wherein the step of calculating similarity between the self-defined keyword and each word in a corpus, and acquiring extended keywords related to the self-defined keyword from the corpus according to the similarity, which is executed by the
20 processor, comprises:

calculating the similarity between the self-defined keyword and each word in the corpus by using the Pearson Correlation Coefficient method;

acquiring top K words having the highest similarity with the self-defined keyword as the
25 extended keywords of the self-defined keyword, wherein K is a positive integer greater than 0.

15. The computer equipment of claim 11, wherein the step of performing real-time monitoring

according to the target keywords in the target keyword list, which is executed by the processor, comprises: performing real-time monitoring for each target keyword in the target keyword list in a form of a sliding window.

- 5 16. One or more non-volatile readable storage medium(s) storing computer-readable instructions which, when executed by one or more processors, cause the one or more processors to perform the following steps:
- acquiring a self-defined keyword;
- calculating similarity between the self-defined keyword and each word in a corpus, and
- 10 acquiring extended keywords related to the self-defined keyword from the corpus according to the similarity;
- selecting one or more target keywords from the extended keywords according to a type of the extended keywords and similarity between the extended keywords and the self-defined keyword, and adding the target keywords to a target keyword list;
- 15 performing real-time monitoring according to the target keywords in the target keyword list; and performing topic early warning when it is monitored that the number of topics corresponding to the target keywords reaches a preset threshold.

17. The non-volatile readable storage medium of claim 16, wherein the step of selecting one or
- 20 more target keywords from the extended keywords according to a type of the extended keywords and similarity between the extended keywords and the self-defined keyword, and adding the target keywords to a target keyword list, which is executed by the processor, comprises:
- classifying the extended keywords in accordance with a preset type;
- 25 selecting, from each type of extended keywords, top h extended keywords having the highest similarity with the self-defined keyword as the target keywords, wherein h is a positive integer greater than 0;

aggregating the target keywords selected from each type, and generating the target keyword list for monitoring.

18. The non-volatile readable storage medium of claim 16, wherein after the step of acquiring a self-defined keyword, the processor is further configured to execute following steps:

calculating a word vector corresponding to the self-defined keyword;

wherein calculating similarity between the self-defined keyword and each word in a corpus, and acquiring extended keywords related to the self-defined keyword from the corpus according to the similarity comprises:

calculating similarity between the word vector of the self-defined keyword and a word vector of each word in the corpus;

acquiring the extended keywords related to the self-defined keyword from the corpus according to the similarity between the word vectors.

19. The non-volatile readable storage medium of claim 16, wherein the step of calculating similarity between the self-defined keyword and each word in a corpus, and acquiring extended keywords related to the self-defined keyword from the corpus according to the similarity, which is executed by the processor, comprises:

calculating the similarity between the self-defined keyword and each word in the corpus by

using the Pearson Correlation Coefficient method;

acquiring top K words having the highest similarity with the self-defined keyword as the extended keywords of the self-defined keyword, wherein K is a positive integer greater than 0.

20. The non-volatile readable storage medium of claim 16, wherein the step of performing real-time monitoring according to the target keywords in the target keyword list, which is executed by the processor, comprises: performing real-time monitoring for each target keyword in the target keyword list in a form of a sliding window.