SUMMARY

According to one embodiment of the present invention, a method includes receiving, from a requesting device, a first electronic document specifying natural language text describing an issue with a system. The method also includes synthesizing the first electronic document to create a synthesized second electronic document. The synthesizing comprises identifying one or more portions of the first electronic document for filtering from the synthesized second electronic document, based on determining that the one or more portions match one or more predefined filtering rules; performing an auto-refinement operation on the first electronic document, comprising identifying a first set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document, using a data dictionary structure specifying a plurality of previously identified terms relating to identifying solutions for issues with systems; and identifying a second set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document by determining that terms in the second set of terms match a predefined pattern matching rule relating to identifying solutions for issues with systems. The method also includes identifying a first set of electronic documents by searching within a data repository using the synthesized second electronic document comprising the first set of terms identified using the data dictionary structure; inputting the first electronic document to a first machine learning model trained to identify documents in one or more data repositories, and in response identifying a second set of electronic documents in the data repository using the first machine learning model; and selecting at least one document from the first set of electronic documents, relating to identifying the solution to the issue with the system, by correlating the first set of electronic documents and the second set of electronic documents. Additionally, the method includes returning the selected at least one document to the requesting device.

In one aspect, in combination with any example above, the method further includes assigning a weight to at least one term in the first set of terms, the weight indicating the relevancy of the term relative to the other terms in the first set of terms, where identifying the first set of electronic documents within the data repository is based at least in part on the weights assigned to the terms in the first set of terms.

In one aspect, in combination with any example above, the complex system further comprises an aircraft, and the method further includes identifying metadata specifying an area of interest in the aircraft that the electronic document relates to, where the assigned weights are determined based at least in part on the area of interest.

In one aspect, in combination with any example above, identifying the first set of electronic documents within the data repository relating to the synthesized electronic documents further includes accessing an index for a data repository using the synthesized electronic document to identify the first set of electronic documents, where the index is generated based on

In one aspect, in combination with any example above, returning at least the portion of the first set of electronic documents to the requesting device further includes determining a relevancy metric for at least one electronic document in the first set of electronic documents, indicating a measure of relevancy of the respective electronic document from the first set of electronic documents to the received electronic document, determining one or more solutions to the issue with the complex system described by the natural language text of the received electronic document, based on solutions corresponding to the electronic documents in the first set of electronic documents and the determined relevancy metrics for the at least one document in the first set of electronic documents, and returning the one or more solutions to the issue with the complex system to the requesting device.

In one aspect, in combination with any example above, the complex system further comprises an aircraft, and wherein the data repository includes at least a first electronic document that comprises a repair request for an aircraft system, an update to an aircraft system, a replacement of an aircraft system, or a combination thereof, and wherein the requesting device further comprises an operator device for a fleet of aircrafts, inclusive of the aircraft.

Another embodiment described herein provides a system that includes one or more computer processors, and a memory containing computer program code that, when executed by operation of the one or more computer processors, performs an operation. The operation includes receiving, from a requesting device, a first electronic document specifying natural language text describing a customer issue with an aircraft. The operation further includes synthesizing the first electronic document to create a synthesized second electronic document. The synthesizing includes identifying one or more portions of the first electronic document for filtering from the synthesized second electronic document, based on determining that the one or more portions match at least one predefined filtering rule; performing an auto-refinement operation on the first electronic document, comprising identifying a first set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document, using a data dictionary structure specifying a plurality of previously identified terms relating to identifying solutions for issues with aircraft; and identifying a second set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document by determining that terms in the second set of terms match a predefined pattern matching rule relating to identifying solutions for issues with aircraft. The operation also includes identifying a first set of electronic documents by searching within a data repository

using the synthesized second electronic document comprising the first set of terms identified using the data dictionary structure; inputting the first electronic document to a first machine learning model trained to identify documents in one or more data repositories, and in response identifying a second set of electronic documents in the data repository using the first machine learning model; and selecting at least one document from the first set of electronic documents, relating to identifying the solution to the issue with the system, by correlating the first set of electronic documents and the second set of electronic documents. The

operation further includes returning at least one document to the requesting device.

In one aspect, in combination with any example above, the operation further includes retrieving metadata specifying an area of the aircraft that the electronic document relates to, and assigning a respective weight to at least one of the terms in the first set of terms and the second set of terms, based on the specified area of the aircraft that the electronic document relates to.

In one aspect, in combination with any example above, synthesizing the electronic document to create the synthesized electronic document further comprises adding one or more synonym words to the electronic document that correspond to a first word within the electronic document.

In one aspect, in combination with any example above, synthesizing the electronic document to create the synthesized electronic document further comprises replacing a first word within the electronic document with a corresponding stem word.

In one aspect, in combination with any example above, at least one word in the second set of terms is assigned a respective weight based on the corresponding at least one predefined pattern matching rule, wherein the weight is further determined based on a section of the aircraft that the electronic document relates to, wherein at least one section of the aircraft is assigned respective weights for the at least one predefined pattern matching rules.

In one aspect, in combination with any example above, identifying the second set of terms within the electronic document to include in the synthesized electronic document further comprises determining when a plurality of words, inclusive of one of the second set of terms, satisfies one or more predefined pattern matching rules.

In one aspect, in combination with any example above, synthesizing the electronic document to create the synthesized electronic document, further includes identifying a third set of terms corresponding to the electronic document to include in the synthesized electronic document, by processing the electronic document as an input to a machine learning model trained to recognize relevant terms within the electronic document.

In one aspect, in combination with any example above, the operation further includes refining the machine learning model trained to recognize relevant items within the data repository, and storing the electronic document and an associated optimum solution within the data repository, wherein the data repository includes at least a first electronic document that comprises a repair request for an aircraft system, an update to an aircraft system, a replacement of an aircraft system, or a combination thereof, and wherein the requesting device further comprises an operator device for a fleet of aircrafts, inclusive of the aircraft.

Another embodiment described herein provides a a non-transitory computer program product. The computer program product includes a non-transitory computer-readable storage medium having computer-readable program code embodied therewith, the computer-readable program code executable by one or more computer processors to perform an operation. The operation includes receiving, from a requesting device, a first electronic document specifying natural language text describing an issue with a system; and synthesizing the first electronic document to create a synthesized second electronic document. The synthesizing includes identifying one or more portions of the first electronic document for filtering from the synthesized second electronic document, based on determining that the one or more portions match one or more predefined filtering rules; performing an auto-refinement operation on the first electronic document, comprising identifying a first set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document, using a data dictionary structure specifying a plurality of previously identified terms relating to identifying solutions for issues with systems; and identifying a second set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document by determining that terms in the second set of terms match a predefined pattern matching rule relating to identifying solutions for issues with systems. The operation also includes identifying a first set of electronic documents by searching within a data repository using the synthesized second electronic document comprising the first set of terms identified using the data dictionary structure; inputting the first electronic document to a first machine learning model trained to identify documents in one or more data repositories, and in response identifying a second set of electronic documents in the data repository using the first machine learning model; selecting at least one document from the first set of electronic documents, relating to identifying the solution to the issue with the system, by correlating the first set of electronic documents and the second

set of electronic documents; and returning the selected at least one document to the requesting device.

BRIEF DESCRIPTION OF THE DRAWINGS

So that the manner in which the above recited features of the present disclosure can be understood in detail, a more particular description of the disclosure, briefly summarized above, may be had by reference to aspects, some of which are illustrated in the appended drawings.

CLAIMS

1. A method, comprising:

receiving, from a requesting device, a first electronic document specifying natural language text describing an issue with a system;

synthesizing the first electronic document to create a synthesized second electronic document, the synthesizing comprising:

identifying one or more portions of the first electronic document for filtering from the synthesized second electronic document, based on determining that the one or more portions match one or more predefined filtering rules;

performing an auto-refinement operation on the first electronic document, comprising identifying a first set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document, using a data dictionary structure specifying a plurality of previously identified terms relating to identifying solutions for issues with systems; and

identifying a second set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document by determining that terms in the second set of terms match a predefined pattern matching rule relating to identifying solutions for issues with systems;

identifying a first set of electronic documents by searching within a data repository using the synthesized second electronic document comprising the first set of terms identified using the data dictionary structure;

inputting the first electronic document to a first machine learning model trained to identify documents in one or more data repositories, and in response identifying a second set of electronic documents in the data repository using the first machine learning model;

selecting at least one document from the first set of electronic documents, relating to identifying the solution to the issue with the system, by correlating the first set of electronic documents and the second set of electronic documents; and

returning the selected at least one document to the requesting device.

2. The method of claim 1, further comprising:

assigning a respective weight to at least one term in the first set of terms, the weight indicating a relevancy of the at least one term relative to other terms in the first set of terms,

wherein identifying the first set of electronic documents by searching within the data repository is based at least in part on the weight assigned to the at least one term in the first set of terms.

3. The method of claim 2, wherein the system further comprises an aircraft, and wherein the method further comprises:

identifying metadata specifying an area of interest in the aircraft that the first electronic document relates to,

wherein the assigned weights are determined based at least in part on the area of interest.

4. The method of claim 1, wherein identifying the first set of electronic documents by searching within the data repository using the synthesized second electronic document further comprises:

accessing an index for the data repository using the synthesized second electronic document to identify the first set of electronic documents, wherein the index is generated based on a plurality of synthesized electronic documents corresponding to a plurality of electronic documents stored within the data repository.

5. The method of claim 1, wherein synthesizing the first electronic document to create the synthesized second electronic document further comprises:

adding one or more synonym words to the second electronic document that correspond to a first word within the first electronic document; and

replacing a first word within the first electronic document with a corresponding stem word in the second electronic document.

6. The method of claim 1, further comprising:

storing the first electronic document and an associated identified solution, based on the at least one document, within the data repository; and

updating an index for the data repository to identify the stored first electronic document and the associated identified solution.

7. The method of claim 1, wherein the system further comprises an aircraft, wherein at least one word in the second set of terms is assigned a respective weight based on the corresponding at

least one predefined pattern matching rule, wherein the weight is further determined based on a section of the aircraft that the first electronic document relates to.

8. The method of claim 1, wherein synthesizing the first electronic document to create the synthesized second electronic document, further comprises:

identifying a third set of terms corresponding to the first electronic document to include in the synthesized second electronic document, by processing the first electronic document as an input to a second machine learning model trained to recognize relevant terms within electronic documents specifying natural language text describing issues with the system.

9. The method of claim 1, wherein returning the selected at least one document to the requesting device further comprises:

determining a relevancy metric for at least one electronic document in the first set of electronic documents, indicating a measure of relevancy of the at least one electronic document to the received first electronic document;

determining one or more solutions to the issue with the system described by the natural language text of the received first electronic document, based on solutions corresponding to the electronic documents in the first set of electronic documents and the determined relevancy metrics for the at least one electronic document; and

returning the one or more solutions to the issue with the system to the requesting device.

10. The method of claim 1, wherein the system further comprises an aircraft, and wherein the data repository includes at least one electronic document that comprises a repair request for an aircraft system, an update to an aircraft system, a replacement of an aircraft system, or a combination thereof, and wherein the requesting device further comprises an operator device for a fleet of aircrafts, inclusive of the aircraft.

11. The method of claim 1, wherein the synthesizing further comprises:

identifying a third set of terms corresponding to the first electronic document, for inclusion in the synthesized second electronic document, by analyzing the first electronic document using a second trained machine learning model.

12. The method of claim 1, wherein the synthesizing further comprises:

identifying a third set of terms corresponding to the first electronic document, for inclusion in the synthesized second electronic document, based on one or more terms previously identified as relating to a requesting entity associated with the requesting device.

13. A system, comprising:

one or more computer processors; and

a memory containing computer program code that, when executed by operation of the one or more computer processors, performs an operation comprising:

receiving, from a requesting device, a first electronic document specifying natural language text describing a customer issue with an aircraft;

synthesizing the first electronic document to create a synthesized second electronic document, the synthesizing comprising:

identifying one or more portions of the first electronic document for filtering from the synthesized second electronic document, based on determining that the one or more portions match at least one predefined filtering rule;

performing an auto-refinement operation on the first electronic document, comprising identifying a first set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document, using a data dictionary structure specifying a plurality of previously identified terms relating to identifying solutions for issues with aircraft; and

identifying a second set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document by determining that terms in the second set of terms match a predefined pattern matching rule relating to identifying solutions for issues with aircraft;

identifying a first set of electronic documents by searching within a data repository using the synthesized second electronic document comprising the first set of terms identified using the data dictionary structure;

inputting the first electronic document to a first machine learning model trained to identify documents in one or more data repositories, and in response identifying a second set of electronic documents in the data repository using the first machine learning model;

selecting at least one document from the first set of electronic documents, relating to identifying the solution to the issue with the system, by correlating the first set of electronic documents and the second set of electronic documents; and

returning the selected at least one document to the requesting device.

14. The system of claim 13, the operation further comprising:

retrieving metadata specifying an area of the aircraft that the first electronic document relates to, and

assigning a respective weight to at least one term in the first set of terms, based on the area of the aircraft that the first electronic document relates to.

15. The system of claim 13, wherein synthesizing the first electronic document to create the synthesized second electronic document further comprises adding one or more synonym words to the second electronic document that correspond to a first word within the first electronic document.

16. The system of claim 13, wherein at least one word in the second set of terms is assigned a weight based on the satisfied at least one predefined pattern matching rule, and wherein the weight is further determined based on a section of the aircraft that the first electronic document relates to.

17. The system of claim 13, wherein identifying the second set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document further comprises determining when a plurality of words, inclusive of one of the second set of terms, satisfies one or more predefined pattern matching rules.

18. The system of claim 13, wherein synthesizing the first electronic document to create the synthesized second electronic document, further comprises:

identifying a third set of terms corresponding to the first electronic document to include in the synthesized second electronic document, by processing the first electronic document as input to a second machine learning model trained to recognize relevant terms within electronic documents specifying natural language text describing issues with aircraft.

19. The system of claim 13, the operation further comprising:

storing the first electronic document and an associated identified solution within the data repository.

20. A non-transitory computer program product, the computer program product comprising:

a non-transitory computer-readable storage medium having computer-readable program code embodied therewith, the computer-readable program code executable by one or more computer processors to perform an operation, the operation comprising:

receiving, from a requesting device, a first electronic document specifying natural language text describing an issue with a system;

synthesizing the first electronic document to create a synthesized second electronic document, the synthesizing comprising:

identifying one or more portions of the first electronic document for filtering from the synthesized second electronic document, based on determining that the one or more portions match one or more predefined filtering rules;

performing an auto-refinement operation on the first electronic document, comprising identifying a first set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document, using a data dictionary structure specifying a plurality of previously identified terms relating to identifying solutions for issues with systems; and

identifying a second set of terms corresponding to the first electronic document for inclusion in the synthesized second electronic document by determining that terms in the second set of terms match a predefined pattern matching rule relating to identifying solutions for issues with systems;

identifying a first set of electronic documents by searching within a data repository using the synthesized second electronic document comprising the first set of terms identified using the data dictionary structure;

inputting the first electronic document to a first machine learning model trained to identify documents in one or more data repositories, and in response identifying a second set of electronic documents in the data repository using the first machine learning model;

selecting at least one document from the first set of electronic documents, relating to identifying the solution to the issue with the system, by correlating the first set of electronic documents and the second set of electronic documents; and

returning the selected at least one document to the requesting device.