

Claims

1. A computer-implemented method, comprising:

identifying textual information related to user information from user service information;

collecting the textual information based on selective configurations according to service demands;

performing a layered matching on the textual information based on preset background identification information in a preset list, wherein the layered matching includes a sequence of different matching methods based upon a descending order of matching accuracies, wherein the preset list includes a plurality of entries storing different preset background identification information related to the user information and associated with a particular user or groups of users, and wherein the preset list establishes mapping relationships between original and current data values, the layered matching comprising:

performing a first matching method based upon a first match accuracy between a first subset of the textual information and the plurality of entries of the preset list to produce a saved first matching method result;

in response to the saved first match method result from the first matching method, determining a lack of exact match between the first subset of the textual information and the plurality of entries of the preset list;

in response to determining the lack of exact match, performing a second matching method based upon a second matching accuracy between the first subset of the textual information and the plurality of entries of the preset list to produce a saved second matching method result;

in response to the saved second matching method result from the second matching method, determining a unique match between the first subset of the textual information and a unique one of the plurality of entries of the preset list, the unique match indicating that the first subset of the textual information contains or is contained in the unique one of the plurality of entries of the preset list;

in response to determining the unique match, updating the preset list to provide an association between the first subset of the textual information and the unique one of the plurality of entries of the preset list; and

performing a third matching method based upon a preset similarity formula having a preset similarity threshold between the textual information and the plurality of entries of the preset list to produce a similarity match result; and determining the user information based on the layered matching.

2. The computer-implemented method of claim 1, wherein performing the third matching method based on preset similarity formula further comprises:

calculating a similarity value between the textual information and each entry in the preset list using the preset similarity formula;

determining a largest similarity value; and

determining whether the largest similarity value is greater than or equal to a preset threshold.

3. The computer-implemented method of claim 2, wherein calculating the similarity value between the textual information and each entry in the preset list using the preset similarity formula further comprises:

calculating a longest common sub-string length between the textual information and that particular entry in the preset list; and

dividing the longest common sub-string length by a string length of that particular entry.

4. The computer-implemented method of claim 1, wherein updating the preset list is based on the saved second matching method result in response to success of the second matching method.

5. The computer-implemented method of claim 1, wherein determining the user information based on the layered matching further comprises determining the user information based on the first matching method, the second matching method, or the third matching method.

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

performing data cleaning on the textual information to produce cleaned textual information; and

performing the layered matching on the cleaned textual information.

7. A non-transitory, computer-readable medium storing one or more instructions executable by a computer system to perform operations comprising:

- identifying textual information related to user information from user service information;

- collecting the textual information based on selective configurations according to service demands;

- performing a layered matching on the textual information based on preset background identification information in a preset list, wherein the layered matching includes a sequence of different matching methods based upon a descending order of matching accuracies, wherein the preset list includes a plurality of entries storing different preset background identification information related to the user information and associated with a particular user or groups of users, and wherein the preset list establishes mapping relationships between original and current data values, the layered matching comprising:

 - performing a first matching method based upon a first match accuracy between a first subset of the textual information and the plurality of entries of the preset list to produce a saved first matching method result;

 - in response to the saved first match method result from the first matching method, determining a lack of exact match between the first subset of the textual information and the plurality of entries of the preset list;

 - in response to determining the lack of exact match, performing a second matching method based upon a second matching accuracy between the first subset of the textual information and the plurality of entries of the preset list to produce a saved second matching method result;

 - in response to the saved second matching method result from the second matching method, determining a unique match between the first subset of the textual information and a unique one of the plurality of entries of the preset list, the unique match indicating that the first subset of the textual information contains or is contained in the unique one of the plurality of entries of the preset list;

 - in response to determining the unique match, updating the preset list to provide an association between the first subset of the textual information and the unique one of the plurality of entries of the preset list; and

performing a third matching method based upon a preset similarity formula having a preset similarity threshold between the textual information and the plurality of entries of the preset list to produce a similarity match result; and determining the user information based on the layered matching.

8. The non-transitory, computer-readable medium of claim 7, wherein performing the third matching method based on the preset similarity formula further comprises one or more instructions to:

calculate a similarity value between the textual information and each entry in the preset list using the preset similarity formula;

determine a largest similarity value; and

determine whether the largest similarity value is greater than or equal to a preset threshold.

9. The non-transitory, computer-readable medium of claim 8, wherein calculating the similarity value between the textual information and each entry in the preset list using the preset similarity formula further comprises one or more instructions to:

calculate a longest common sub-string length between the textual information and that particular entry in the preset list; and

divide the longest common sub-string length by a string length of that particular entry.

10. The non-transitory, computer-readable medium of claim 7, wherein updating the preset list includes operations comprising one or more instructions to update the preset list based upon the saved second matching method result in response to success of the second matching method.

11. The non-transitory, computer-readable medium of claim 7, wherein determining the user information based on the layered matching further comprises one or more instructions to determine the user information based on the first matching method, the second matching method, or the third matching method.

12. The non-transitory, computer-readable medium of claim 7, further comprising one or more instructions to:

perform data cleaning on the textual information to produce cleaned textual information; and

perform the layered matching on the cleaned textual information.

13. A computer-implemented system, comprising:

a computer memory; and

a hardware processor interoperably coupled with the computer memory and configured to perform operations comprising:

identifying textual information related to user information from user service information;

collecting the textual information based on selective configurations according to service demands;

performing a layered matching on the textual information based on preset background identification information in a preset list, wherein the layered matching includes a sequence of different matching methods based upon a descending order of matching accuracies, wherein the preset list includes a plurality of entries storing different preset background identification information related to the user information and associated with a particular user or groups of users, and wherein the preset list establishes mapping relationships between original and current data values, the layered matching comprising:

performing a first matching method based upon a first match accuracy between a first subset of the textual information and the plurality of entries of the preset list to produce a saved first matching method result;

in response to the saved first match method result from the first matching method, determining a lack of exact match between the first subset of the textual information and the plurality of entries of the preset list;

in response to determining the lack of exact match, performing a second matching method based upon a second matching accuracy between the first subset of the textual information and the plurality of entries of the preset list to produce a saved second matching method result;

in response to the saved second matching method result from the second matching method, determining a unique match between the first subset of the textual information and a unique one of the plurality of entries of the

preset list, the unique match indicating that the first subset of the textual information contains or is contained in the unique one of the plurality of entries of the preset list;

in response to determining the unique match, updating the preset list to provide an association between the first subset of the textual information and the unique one of the plurality of entries of the preset list; and

performing a third matching method based upon a preset similarity formula having a preset similarity threshold between the textual information and the plurality of entries of the preset list to produce a similarity match result; and

determining the user information based on the layered matching.

14. The computer-implemented system of claim 13, wherein performing the third matching method based on the preset similarity formula is further configured to:

calculate a similarity value between the textual information and each entry in the preset list using the preset similarity formula;

determine a largest similarity value; and

determine whether the largest similarity value is greater than or equal to a preset threshold.

15. The computer-implemented system of claim 14, wherein calculating the similarity value between the textual information and each entry in the preset list using the preset similarity formula further is further configured to:

calculate a longest common sub-string length between the textual information and that particular entry in the preset list; and

divide the longest common sub-string length by a string length of that particular entry.

16. The computer-implemented system of claim 13, wherein updating the preset list is based on the saved second matching method result in response to success of the second matching method, and wherein determining the user information based on the layered matching is further configured to determine the user information based on the first matching method, the second matching method, or the third matching method.

17. The computer-implemented system of claim 13, further configured to:
- perform data cleaning on the textual information to produce cleaned textual information; and
 - perform the layered matching on the cleaned textual information.