

Claims

1. A computer-implemented method, comprising:

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

collecting the identified 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 different matching methods, 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;

based upon results of the layered matching, updating the preset list to include updated preset background identification information related to the user information; and

determining the user information based on the layered matching.

2. The computer-implemented method of claim 1, wherein performing the layered matching on the textual information based on the preset background identification information in the preset list further comprises:

performing a first matching on the textual information by determining whether the textual information is the same as one entry in the preset list;

in response to an unsuccessful first matching, performing a second matching on the textual information by determining whether the textual information contains or is contained in one and only one entry of the preset list; and

in response to an unsuccessful second matching, performing a third matching on the textual information based on the preset background identification information in the preset list and a preset similarity formula.

3. The computer-implemented method of claim 2, wherein performing the third matching on the textual information based on the preset background identification information in the preset list and the 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.

4. The computer-implemented method of claim 3, 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.

5. The computer-implemented method of claim 2, wherein updating the preset list to include the updated preset background identification information related to the user information is based on the second matching in response to success of the second matching.

6. The computer-implemented method of claim 2, wherein determining the user information based on the layered matching further comprises determining the user information based on the first matching, the second matching, or the third matching.

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

performing data cleaning on the textual information; and

performing the layered matching on the cleaned textual information.

8. 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 identified 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 different matching methods, 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;

based upon results of the layered matching, updating the preset list to include updated preset background identification information related to the user information; and

determining the user information based on the layered matching.

9. The non-transitory, computer-readable medium of claim 8, wherein performing the layered matching on the textual information based on the preset background identification information in the preset list further comprises one or more instructions to:

perform a first matching on the textual information by determining whether the textual information is the same as one entry in the preset list;

in response to an unsuccessful first matching, perform a second matching on the textual information by determining whether the textual information contains or is contained in one and only one entry of the preset list; and

in response to an unsuccessful second matching, perform a third matching on the textual information based on the preset background identification information in the preset list and a preset similarity formula.

10. The non-transitory, computer-readable medium of claim 9, wherein performing the third matching on the textual information based on the preset background identification information in the preset list and 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.

11. The non-transitory, computer-readable medium of claim 10, 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.

12. The non-transitory, computer-readable medium of claim 9, wherein updating the preset list to include the updated preset background identification information related to the user information includes operations comprising one or more instructions to update the preset list to include the updated preset background identification information related to the user information based upon the second matching in response to success of the second matching.

13. The non-transitory, computer-readable medium of claim 9, 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, the second matching, or the third matching.

14. The non-transitory, computer-readable medium of claim 8, further comprising one or more instructions to:

- perform data cleaning on the textual information; and

- perform the layered matching on the cleaned textual information.

15. 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 identified 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 different matching methods, 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;

based upon results of the layered matching, updating the preset list to include updated preset background identification information related to the user information; and

determining the user information based on the layered matching.

16. The computer-implemented system of claim 15, wherein performing the layered matching on the textual information based on the preset background identification information in the preset list is further configured to:

perform a first matching on the textual information by determining whether the textual information is the same as one entry in the preset list;

in response to an unsuccessful first matching, perform a second matching on the textual information by determining whether the textual information contains or is contained in one and only one entry of the preset list; and

in response to an unsuccessful second matching, perform a third matching on the textual information based on the preset background identification information in the preset list and a preset similarity formula.

17. The computer-implemented system of claim 16, wherein performing the third matching on the textual information based on the preset background identification information in the preset list and 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.

18. The computer-implemented system of claim 17, 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.

19. The computer-implemented system of claim 16, wherein updating the preset list to include the updated preset background identification information related to the user information is further configured to update the preset list to include the updated preset background identification information related to the user information based on the second matching in response to success of the second matching, and wherein determining the user information based on the layered matching is further configured to determine the user information based on the first matching, the second matching, or the third matching.

20. The computer-implemented system of claim 15, further configured to:

perform data cleaning on the textual information; and

perform the layered matching on the cleaned textual information.