

【DESCRIPTION】

【Title】

SYSTEM AND METHOD FOR PROVIDING PERSONAL DICTIONARY
BUILDING SERVICE

【Technical Field】

[1] The present invention relates to technology for providing a service for building a personal dictionary through character string-tag matching personalized for each user.

【Background Art】

[2] People may differ in literary aptitude when writing, but everyone has likely struggled at some point with choosing appropriate phrases and constructing sentences to clearly convey intended meaning to others.

[3] In particular, everyone has probably experienced the inconvenience of having to rewrite text from scratch since it was difficult to recall or find good expressions written or seen in the past, and the frustration of feeling dissatisfied with the newly written text despite recalling phrases that would be better suited to the text.

[4] In addition, when drafting documents collaboratively in small teams, many may have found it difficult to ensure consistency and efficiency in the work

since different individuals use different expressions.

[5] Recently, with the development of artificial intelligence (AI) algorithms, services have been introduced that allow users to easily and conveniently create desired sentences simply by entering keywords. However, these services provide only generic phrases and sentences as outcomes of machine learning based on big data, and cannot provide personalized expressions tailored to each user.

[6] Furthermore, such big data-based document creation services inevitably have disadvantages in that the services may be exposed to security issues, copyright issues, etc.

[7] Accordingly, there is growing need for a document creation service capable of solving security and copyright issues and allowing users to prepare documents more conveniently and efficiently based on individual experience.

[8] (Patent Document 1) US 11093557 B2 (August 17, 2021)

[9] (Patent Document 2) US 11010371 B2 (May 18, 2021)

【Disclosure】

【Technical Problem】

[10] It is an object of the present invention to provide a system and method for providing a service that may build a personal dictionary having a character string-tag matching structure based on small-scale data experienced

individually or by teams.

【Technical Solution】

[11] A system for building a personal dictionary according to the present invention for solving the above-mentioned problems is a system for providing a service for building a personal dictionary through character string-tag matching personalized for each user, the system including a user terminal including a character string selection unit configured to select a character string for which a user desires to create a tag, a character string transmission unit configured to transmit the character string selected by the character string selection unit to the service provider server, and a tag selection and input unit configured to allow the user to select one tag from among a plurality of tags received from the service provider server or manually input a tag, and a service provider server including a character string reception unit configured to receive a character string from the character string transmission unit, a first character count determination unit configured to count a number of characters in the character string received by the character string reception unit and compare the number with a first reference value, a summary extraction unit configured to generate a summary obtained by summarizing the character

string when the first character count determination unit determines that the number of characters is equal to or greater than the first reference value, a second character count determination unit configured to count the number of characters in the character string received by the character string reception unit or the number of characters in the summary generated by the summary extraction unit and compare the number with a second reference value, a word extraction unit configured to extract all words included in the character string or the summary when the second character count determination unit determines that the number of characters is less than the second reference value, a keyword extraction unit configured to extract a topic term included in the character string or the summary when the second character count determination unit determines that the number of characters is equal to or greater than the second reference value, a tag recommendation unit configured to extract, from a user database (DB) or a provider DB, a tag corresponding to the words extracted by the word extraction unit or the topic term extracted by the keyword extraction unit and transmit the tag to the tag selection and input unit of the user terminal, a dictionary storage unit configured to match the tag transmitted from the tag selection and input unit with the character string and store the tag matched with the character string in the user DB,

and a score calculation unit configured to calculate a score of the tag stored by the dictionary storage unit and store the score in the user, wherein the tag recommendation unit extracts top tags having high scores based on the score stored in the user DB, and transmit the top tags to the tag selection and input unit of the user terminal, and the score calculation unit calculates the score of the tag using, as variables, an accumulated selection count of the tag, whether the tag is manually input, a recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.

[12] Preferably, when the score calculation unit calculates the score, among the score variables, the accumulated selection count of the tag may become a default value, and the remaining score variables may function as weights applied to the default value based on a priority.

[13] Preferably, the priority may be assigned in order of whether the tag is manually input, the recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.

[14] Preferably, the user DB may include a tag recommendation DB configured to store a word or a topic term, a tag matching the word or the topic term, and a score of the tag, and a dictionary DB configured to store a tag and a character string matching the tag.

[15] Meanwhile, a method of building a personal dictionary according to the present invention is a method of providing a service for building a personal dictionary through character string-tag matching personalized for each user, the method including a character string selection step in which a user selects a character string for which the user intends to create a tag on a user terminal, a character string transmission step in which the character string selected in the character string selection step is transmitted from the user terminal to the service provider server, a character string reception step in which the character string transmitted in the character string transmission step is received by the service provider server, a first character count determination step in which a number of characters in the character string received in the character string reception step is counted and compared with the first reference value on the service provider server, a summary extraction step in which a summary obtained by summarizing the character string is generated when it is determined in the first character count determination step that the number of characters is equal to or greater than the first reference value on the service provider server, a second character count determination step in which the number of characters in the character string received in the character string reception step or a number of characters in

the summary generated in the summary extraction step is counted and compared with a second reference value on the service provider server, a word extraction step in which all words included in the character string or the summary are extracted when it is determined in the second character count determination step that the number of characters is less than the second reference value on the service provider server, a keyword extraction step in which a topic term included in the character string or the summary is extracted when it is determined in the second character count determination step that the number of characters is equal to or greater than the second reference value on the service provider server, a tag recommendation step in which a tag corresponding to the words extracted in the word extraction step or the topic term extracted in the keyword extraction step is extracted from a user DB or a provider DB on the service provider server and transmitted to the user terminal, a tag selection and input step in which the user selects one tag from among a plurality of tags transmitted from the service provider server or manually inputs a tag on the user terminal, a dictionary storage step in which the tag selected or manually input in the tag selection and input step is matched with the character string and stored in the user DB on the service provider server, and a score calculation step in which the score of the stored tag is calculated and stored in the user

DB on the service provider server, wherein the tag recommendation step includes extracting top tags having high scores based on the score stored in the user DB, and transmitting the top tags to the user terminal, and the score calculation step includes calculating the score of the tag using, as variables, an accumulated selection count of the tag, whether the tag is manually input, a recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.

[16] Preferably, when the score is calculated in the score calculation step, among the score variables, the accumulated selection count of the tag may become a default value, and the remaining score variables may function as weights applied to the default value based on a priority.

[17] Preferably, the priority may be assigned in order of whether the tag is manually input, the recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.

【Advantageous effects】

[18] A system and method for building a personal dictionary according to the present invention may allow a personal dictionary having a character string-tag matching structure to be built based on small-scale data experienced individually or by teams, so that individual users may

receive personalized phrases and sentences, and thus be enabled to prepare documents based on individual experience more conveniently and efficiently.

[19] In addition, in recommending a tag for building a personal dictionary, top tags are recommended in the order of scores determined on the basis of variables such as the accumulated selection count of the tag, whether the tag is manually input, recent selection history, whether a user is a dictionary creator, and whether the tag is included in a provider database, so that personalized characteristics for each user may be well reflected.

[20] Furthermore, it is possible to solve security issues, copyright issues, etc. that may arise when providing document creation services based on building of the personal dictionary.

【Description of Drawings】

[21] FIG. 1 and FIG. 2 are diagrams schematically illustrating a configuration of a personal dictionary building system according to an embodiment of the present invention;

[22] FIG. 3 is a diagram schematically illustrating a data structure of a user database in the personal dictionary building system according to an embodiment of the present invention;

[23] FIG. 4 is a flowchart schematically illustrating a method of building a personal dictionary according to an embodiment of the present invention; and

[24] FIG. 5 is a flowchart specifically illustrating the method of building the personal dictionary according to the embodiment of the present invention.

[25] *Description of key reference symbols in the drawing*

[26] 100: user terminal, 110: character string selection unit

[27] 120: character string transmission unit, 130: tag selection and input unit

[28] 200: service provider server, 210: character string reception unit

[29] 221: first character count determination unit, 222: second character count determination unit

[30] 223: summary extraction unit, 231: word extraction unit

[31] 232: keyword extraction unit, 240: tag recommendation unit

[32] 250: dictionary storage unit, 260: score calculation unit

[33] 271: provider DB, 272: user DB

[34] 273: provider DB generation unit, 274: tag recommendation DB

[35] 275: dictionary DB

[36] S100: character string selection step, S200: character string transmission step

[37] S300: character string reception step, S400: first character count determination step

[38] S410: summary extraction step, S500: second character count determination step

[39] S610: word extraction step, S620: keyword extraction step

[40] S700: tag recommendation step, S800: tag selection and input step

[41] S900: dictionary storage step, S1000: score calculation step

【Best Mode】

[42] Hereinafter, a system and method for building a personal dictionary according to the embodiments of the present invention will be described in detail with reference to the accompanying drawings. However, a detailed description of well-known functions and configurations that may unnecessarily obscure the gist of the present invention will be omitted.

[43] First, the personal dictionary building system according to the embodiments of the present invention will be described in detail with reference to FIG. 1 to FIG. 3.

[44] FIG. 1 and FIG. 2 are diagrams schematically

illustrating a configuration of the personal dictionary building system according to an embodiment of the present invention, and FIG. 3 is a diagram schematically illustrating a data structure of a user database (DB) in the personal dictionary building system according to an embodiment of the present invention.

[45] As shown in FIG. 1, the personal dictionary building system according to the embodiments of the present invention is a system that provides a service for building a personal dictionary through personalized character string-tag matching for each user. The system broadly includes a user terminal 100 and a service provider server 200, and a personal dictionary building service may be implemented as Software-as-a-Service (SaaS)-based software on an online cloud platform through an internet browser.

[46] Referring to FIG. 2, the user terminal 100 includes a character string selection unit 110 that selects a character string for which a user desires to create a tag, a character string transmission unit 120 that transmits the character string selected by the character string selection unit 110 to the service provider server 200, and a tag selection and input unit 130 that allows the user to select one tag from among a plurality of tags received from the service provider server 200 or manually input a tag.

[47] The character string selection unit 110 is

implemented as a web extension program of an internet browser installed on the user terminal 100, and allows the user to drag a specific phrase from newspaper articles, documents, materials, etc. accessed through the internet browser, thereby selecting a character string to be transmitted to the service provider server 200.

[48] In addition, the character string selection unit 110 allows the user to input and drag a specific phrase through an editing program implemented as a web-based cloud application accessed through the internet browser installed on the user terminal 100, thereby selecting a character string to be transmitted to the service provider server 200.

[49] Further, the character string selection unit 110 may be implemented as a voice recognition program installed on the user terminal 100, thereby converting voice input by the user into a character string and selecting a character string to be transmitted to the service provider server 200.

[50] The character string transmission unit 120 is implemented as a communication module embedded in the user terminal 100 to transmit, via wired or wireless communication, the character string selected by the character string selection unit 110 to the service provider server 200.

[51] The tag selection and input unit 130 transmits a character string to the service provider server 200, and

then allows the user to select one tag from among a plurality of tags transmitted from the service provider server 200 or manually input a tag.

[52] The tag selected or input through the tag selection and input unit 130 is transmitted to the service provider server 200 through the communication module embedded in the user terminal 100.

[53] Similarly to the character string selection unit 110, the tag selection and input unit 130 may be implemented as a web extension program of an internet browser installed on the user terminal 100, an editing program implemented as a cloud application, or a voice recognition program.

[54] Meanwhile, the service provider server 200 includes a character string reception unit 210 that receives a character string from the character string transmission unit 120, a first character count determination unit 221 that counts the number of characters in the character string received by the character string reception unit 210 and compares the number with a first reference value, a summary extraction unit 223 that generates a summary obtained by summarizing the character string when the first character count determination unit 221 determines that the number of characters is equal to or greater than the first reference value, a second character count determination unit 222 that counts the number of characters in the character string

received by the character string reception unit 210 or the number of characters in the summary generated by the summary extraction unit 223 and compares the number with a second reference value, a word extraction unit 231 that extracts all words included in the character string or summary when the second character count determination unit 222 determines that the number of characters is less than the second reference value, a keyword extraction unit 232 that extracts a topic term included in the character string or the summary when the second character count determination unit 222 determines that the number of characters is equal to or greater than the second reference value, a tag recommendation unit 240 that extracts, from the user DB 272 or the provider DB 271, a tag corresponding to the words extracted by the word extraction unit 231 or the topic term extracted by the keyword extraction unit 232 and transmits the tag to the tag selection and input unit 130 of the user terminal 100, a dictionary storage unit 250 that matches the tag transmitted from the tag selection and input unit 130 with the character string and stores the tag matched with the character string in the user DB 272, and a score calculation unit 260 that calculates a score of the tag stored by the dictionary storage unit 250 and stores the score in the user DB 272.

[55] The character string reception unit 210 is

implemented as a communication module equipped in the service provider server 200, and receives a character string from the character string transmission unit 120 of the user terminal 100 via wired or wireless communication.

[56] The first character count determination unit 221, the summary extraction unit 223, the second character count determination unit 222, the word extraction unit 231, and the keyword extraction unit 232 are configured to extract a word or a keyword from a character string for character string-tag matching.

[57] The first character count determination unit 221 counts the number of characters in the character string received by the character string reception unit 210 and compares the number with the first reference value.

[58] When the number of characters in the character string counted by the first character count determination unit 221 is greater than or equal to the first reference value (for example, 100), the summary extraction unit 223 generates a summary.

[59] The summary extraction unit 223 preferably uses an extractive summarization method that extracts an important or core sentence without change from a given character string, and it is possible to use various known summary extraction algorithms, such as TextRank.

[60] The second character count determination unit 222

counts the number of characters in the character string received by the character string reception unit 210 or the number of characters in the summary generated by the summary extraction unit 223 and compares the number with the second reference value.

[61] When the number of characters counted by the second character count determination unit 222 is less than the second reference value (for example, 20), the word extraction unit 231 extracts all words included in the character string or the summary.

[62] On the other hand, when the number of characters counted by the second character count determination unit 222 is greater than or equal to the second threshold (for example, 20), the keyword extraction unit 232 extracts a topic term included in the character string or the summary.

[63] The word extraction unit 231 extracts all words included in the character string or the summary by removing unnecessary characters such as particles and stop words using preprocessing algorithms such as tokenization, cleaning, and normalization.

[64] The keyword extraction unit 232 extracts an important or core word from a given character string using topic term extraction algorithms such as count-based vectorization CountVectorizer and word embedding, and it is possible to use various known keyword extraction algorithms.

[65] That is, in the present invention, when extracting a word or a topic term from a character string for character string-tag matching, accuracy and efficiency of extraction of the word or the topic term may be improved by implementing a step-by-step process of summary extraction, topic term extraction, and word extraction based on the number of characters in the character string.

[66] The tag recommendation unit 240 extracts a tag corresponding to the words extracted by the word extraction unit 231 or the topic term extracted by the keyword extraction unit 232 from the user DB 272 or provider DB 271 and transmits the tag to the tag selection and input unit 130 of the user terminal 100 via a communication module provided in the service provider server 200.

[67] To this end, referring to FIG. 3, the user DB 272 may include a tag recommendation DB 274 that stores a word or a topic term, a tag matching the word or the topic term, and a score of the tag.

[68] For example, the tag recommendation DB 274 stores tags such as "computer", "software", and "HTML" that match the word or topic term "JAVA", and each tag is assigned a score of 100, 200, 50 points, etc.

[69] The tag recommendation unit 240 extracts top tags (for example, four top tags) having high scores based on the score stored in the tag recommendation DB 274 and

transmits the top tags to the tag selection and input unit 130 of the user terminal 100.

[70] When the user DB 272 does not have a tag corresponding to the words extracted by the word extraction unit 231 or the topic term extracted by the keyword extraction unit 232, a tag is extracted from the provider DB 271.

[71] The provider DB 271 is a general-purpose DB provided by the provider so that all users may use the DB. The provider DB generation unit 273 periodically and randomly crawls Internet data to generate general-purpose tags, and updated tags are stored therein.

[72] The dictionary storage unit 250 matches tags transmitted from the tag selection and input unit 130 with the character strings and stores the tags matched with the character strings in the user DB 272.

[73] TO this end, referring to FIG. 3, the user DB 272 may include a dictionary DB 275 that stores tags and character strings matched with the tags.

[74] The dictionary DB 275 that stores tags and character strings matched with the tags becomes a personal dictionary, personalized for each user and built using the personal dictionary building system according to the present invention.

[75] That is, when the user searches the dictionary DB

275 for a tag, a character string matching the tag may be provided, and the provided character string may be utilized when creating a document.

[76] The score calculation unit 260 calculates and updates the scores of tags stored in the tag recommendation DB 274.

[77] In this instance, a score of a tag is determined using, as variables, the accumulated selection count of the tag, whether the tag is manually input, recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.

[78] The accumulated selection count of the tag refers to the accumulated number of times the tag has been selected by the user. The greater the accumulated selection count, the higher the score that may be assigned so that the tag may be positioned higher among the recommended tags.

[79] With regard to whether the tag is manually input, when the user manually inputs a tag instead of selecting a recommended tag, the manually input tag has the highest priority in creating the personal dictionary, and therefore the highest score is assigned. When the user selects one of the recommended tags, a score is assigned so that the tag has the next priority. Optionally, when the user selects a recommended tag as inappropriate, a penalty may be applied.

[80] With regard to recent selection history of tags, a tag that has been selected more recently by the user may be assigned a higher score so that the tag may be positioned higher among the recommended tags.

[81] With regard to whether a user is a dictionary creator, when several users participate in building the personal dictionary, a tag selected by a user who first created the dictionary may be assigned a higher score so that the tag may be positioned higher among the recommended tags.

[82] This is to ensure that, in team-based collaboration, a tag created by a user who first created the dictionary, that is, a team leader, is prioritized over tags created by team members.

[83] With regard to whether the tag is included in the provider DB, when a tag is included in the provider DB 271, this indicates that the tag is a general-purpose tag reflecting recent trends, and therefore a higher score may be assigned so that the tag may be positioned higher among the recommended tags.

[84] When calculating the score, as shown in the formula below, among the score variables, the accumulated selection count of the tag becomes a default value, and the remaining score variables may function as weights applied to the default value based on priority.

[85]

[86] (Score Calculation Formula)

[87] Score = Previous Score + Default Value × Weight 1 ×
Weight 2 × ... × Weight n - Penalty

[88]

[89] In the formula above, the penalty is the score applied when the user selects a recommended tag as inappropriate.

[90] In the personal dictionary building system according to an embodiment of the present invention, weights are assigned in the order of whether the tag is manually entered, recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.

[91] For example, a default value according to the accumulated selection count of the tag may be assigned by normalizing, to a value between 0 and 1, the accumulated selection count of all tags matched to one word or topic term using a sigmoid function.

[92] Depending on whether the tag is manually input, weight 1 may be assigned as 2 for a manually input tag, as 1.5 for a tag selected from among the recommended tags, and as 1.0 for a tag not selected from among the recommended tags.

[93] Weight 2 may be assigned as a value between 1.4 and

1.0 in the recent selection order according to the recent selection history.

[94] Depending on whether a user is a dictionary creator, weight 3 may be assigned as 1.3 when the user is the dictionary creator who first created the dictionary, and as 1.0 otherwise.

[95] Depending on whether the tag is included in the provider DB, weight 4 may be assigned as 1.1 when the tag is included in the provider DB 271, and as 1.0 otherwise.

[96] The penalty may be assigned as a value of 0.1 when the user selects a recommended tag as inappropriate.

[97] In addition, in the formula, the previous score may be assigned by normalizing, to a value between 0 and 1, the previous scores of all tags matched to one word or topic term using a sigmoid function. This is to prevent an old tag from continuously maintaining a high score due to accumulated scoring, which may cause an error in which the old tag remains higher than newly added latest tags.

[98] As described above, in recommending a tag for building a personal dictionary, the personal dictionary building system according to an embodiment of the present invention recommends top tags in the order of scores determined using, as variables, the accumulated selection count of the tag, whether the tag is manually input, recent selection history, whether a user is a dictionary creator,

and whether the tag is included in the provider DB, so that personalized characteristics for each user may be well reflected.

[99] Next, a method of building the personal dictionary according to an embodiment of the present invention will be described in detail with reference to FIG. 4 and FIG. 5.

[100] FIG. 4 is a flowchart schematically illustrating the method of building the personal dictionary according to the embodiment of the present invention, and FIG. 5 is a flowchart specifically illustrating the method of building the personal dictionary according to the embodiment of the present invention.

[101] As shown in FIG. 4 and FIG. 5, the method of building the personal dictionary according to the embodiment of the present invention includes a character string selection step S100, a character string transmission step S200, a character string reception step S300, a first character count determination step S400, a summary extraction step S410, a second character count determination step S500, a word extraction step S610, a keyword extraction step S620, a tag recommendation step S700, a tag selection and input step S800, a dictionary storage step S900, and a score calculation step S1000.

[102] Specifically, the method of building the personal dictionary according to the embodiment of the present

invention is a method of providing a service that builds the personal dictionary through personalized character string-tag matching for each user, and includes the character string selection step S100 in which the user selects a character string for which the user intends to create a tag on the user terminal 100, the character string transmission step S200 in which the character string selected in the character string selection step S100 is transmitted from the user terminal 100 to the service provider server 200, the character string reception step S300 in which the character string transmitted in the character string transmission step S200 is received by the service provider server 200, the first character count determination step S400 in which the number of characters in the character string received in the character string reception step S300 is counted and compared with the first reference value on the service provider server 200, the summary extraction step S410 in which a summary obtained by summarizing the character string is generated when it is determined in the first character count determination step S400 that the number of characters is equal to or greater than the first reference value on the service provider server 200, the second character count determination step S500 in which the number of characters in the character string received in the character string reception step S300

or the number of characters in the summary generated in the summary extraction step S410 is counted and compared with the second reference value on the service provider server 200, the word extraction step S610 in which all words included in the character string or the summary are extracted when it is determined in the second character count determination step S500 that the number of characters is less than the second reference value on the service provider server 200, the keyword extraction step S620 in which the topic term included in the character string or the summary is extracted when it is determined in the second character count determination step S500 that the number of characters is equal to or greater than the second reference value on the service provider server 200, the tag recommendation step S700 in which a tag corresponding to the words extracted in the word extraction step S610 or the topic term extracted in the keyword extraction step S620 is extracted from the user DB 272 or the provider DB 271 on the service provider server 200 and transmitted to the user terminal 100, the tag selection and input step S800 in which the user selects one tag from among a plurality of tags transmitted from the service provider server 200 or manually inputs a tag on the user terminal 100, the dictionary storage step S900 in which the tag selected or manually input in the tag selection and input step S800 is matched

with the character string and stored in the user DB 272 on the service provider server 200, and the score calculation step S1000 in which the score of the stored tag is calculated and stored in the user DB 272 on the service provider server 200.

[103] The character string selection step S100 is a step in which the user selects a character string for which the user intends to create a tag on the user terminal 100.

[104] Here, the user may access the personal dictionary building service, which is implemented as SaaS-based software, on the user terminal 100 through an Internet browser, log in to the service provider server 200, and select a character string through the character string selection unit 110, which includes a web extension program of the Internet browser installed on the user terminal 100, an editing program implemented as a cloud application, or a voice recognition program.

[105] When logging in, the user may import the personal dictionary built by the user, and when the user participates in a personal dictionary of another person that is being built, the personal dictionary of the other person in which the user participates may be imported.

[106] Subsequently, the character string transmission step S200 is a step in which the character string selected in the character string selection step S100 is transmitted

from the user terminal 100 to the service provider server 200.

[107] In addition, the character string reception step S300 is a step in which the character string transmitted in the character string transmission step S200 is received by the service provider server 200.

[108] Subsequently, the first character count determination step S400 is a step in which the first character count determination unit 221 counts the number of characters in the character string received in the character string reception step S300 and compares the number with the first reference value on the service provider server 200.

[109] In this instance, when the number of characters in the character string counted in the first character count determination step S400 is equal to or greater than the first reference value (for example, 100), the summary extraction step S410 is performed to create a summary obtained by summarizing the character string. The summary extraction step S410 is performed through the summary extraction unit 223.

[110] Further, the second character count determination step S500 is a step in which the second character count determination unit 222 counts the number of characters in the character string received in the character string reception step S300 or the number of characters in the

summary generated in the summary extraction step S410 and compares the number with the second reference value on the service provider server 200.

[111] In this instance, when it is determined in the second character count determination step S500 that the number is less than the second reference value, the word extraction step S610 is performed, and when it is determined in the second character count determination step S500 that the number is equal to or greater than the second reference value, the keyword extraction step S620 is performed.

[112] Here, the word extraction step S610 is a step in which all words included in the character string or the summary are extracted by the word extraction unit 231 on the service provider server 200.

[113] In addition, the keyword extraction step S620 is a step in which the keyword extraction unit 232 extracts the topic term included in the character string or the summary on the service provider server 200.

[114] Subsequently, the tag recommendation step S700 is a step in which the tag recommendation unit 240 extracts a tag corresponding to the words extracted in the word extraction step S610 or the topic term extracted in the keyword extraction step S620 from the user DB 272 or the provider DB 271 on the service provider server 200 and transmits the tag to the user terminal 100.

[115] In this instance, the tag recommendation unit 240 extracts top tags (for example, the four top tags) having high scores based on the score stored in the tag recommendation DB 274 of the user DB 272.

[116] A score of a tag is determined using, as variables, the accumulated selection count of the tag, whether the tag is manually input, recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.

[117] Subsequently, the tag selection and input step S800 is a step in which the user selects one tag from among a plurality of tags transmitted from the service provider server 200 or manually inputs a tag through the tag selection and input unit 130 on the user terminal 100.

[118] In this instance, an input window through which the user may manually input a tag and a selection window through which the user may input selection/inapplicable information for the recommended tags are displayed on the user terminal 100.

[119] Further, the dictionary storage step S900 is a step in which the dictionary storage unit 250 matches the tag selected or manually input in the tag selection and input step S800 with the character string and store the tag matched with the character string in the user DB 272 on the service provider server 200.

[120] Subsequently, the score calculation step S1000 is a step in which the score calculation unit 260 calculates the score of the stored tag and stores the score in the user DB 272 on the service provider server 200.

[121] The score may be calculated such that, among the score variables, the accumulated selection count of the tag becomes a default value, and the remaining score variables serve as weights applied to the default value according to the priorities. Optionally, when the user selects a recommended tag as inapplicable, the tag may be applied as a penalty when calculating the score.

[122] As described above, in recommending a tag for building the personal dictionary, the method of building the personal dictionary according to the embodiment of the present invention recommends top tags in the order of scores determined using, as variables, the accumulated selection count of the tag, whether the tag is manually input, recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB, so that personalized characteristics for each user may be well reflected.

[123] Even though the embodiments of the present invention have been described in detail with reference to the accompanying drawings, the embodiments disclosed in the specification and the drawings are merely used for the

purpose of facilitating understanding of the technical spirit of the invention, and are not intended to limit the scope of the invention as set forth in the claims. Therefore, it will be understood by those of ordinary skill in the art that various modifications and equivalent other embodiments are possible therefrom.

【CLAIMS】

【Claim 1】 A system for providing a service for building a personal dictionary through character string-tag matching personalized for each user, the system comprising:

a user terminal comprising a character string selection unit configured to select a character string for which a user desires to create a tag, a character string transmission unit configured to transmit the character string selected by the character string selection unit to the service provider server, and a tag selection and input unit configured to allow the user to select one tag from among a plurality of tags received from the service provider server or manually input a tag; and

a service provider server comprising a character string reception unit configured to receive a character string from the character string transmission unit, a first character count determination unit configured to count a number of characters in the character string received by the character string reception unit and compare the number with a first reference value, a summary extraction unit configured to generate a summary obtained by summarizing the character string when the first character count determination unit determines that the number of characters is equal to or greater than the first reference value, a second character

count determination unit configured to count the number of characters in the character string received by the character string reception unit or the number of characters in the summary generated by the summary extraction unit and compare the number with a second reference value, a word extraction unit configured to extract all words included in the character string or the summary when the second character count determination unit determines that the number of characters is less than the second reference value, a keyword extraction unit configured to extract a topic term included in the character string or the summary when the second character count determination unit determines that the number of characters is equal to or greater than the second reference value, a tag recommendation unit configured to extract, from a user database (DB) or a provider DB, a tag corresponding to the words extracted by the word extraction unit or the topic term extracted by the keyword extraction unit and transmit the tag to the tag selection and input unit of the user terminal, a dictionary storage unit configured to match the tag transmitted from the tag selection and input unit with the character string and store the tag matched with the character string in the user DB, and a score calculation unit configured to calculate a score of the tag stored by the dictionary storage unit and store the score in the user, wherein:

the tag recommendation unit extracts top tags having high scores based on the score stored in the user DB, and transmit the top tags to the tag selection and input unit of the user terminal, and

the score calculation unit calculates the score of the tag using, as variables, an accumulated selection count of the tag, whether the tag is manually input, a recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.

【Claim 2】 The system according to claim 1, wherein, when the score calculation unit calculates the score, among the score variables, the accumulated selection count of the tag becomes a default value, and the remaining score variables function as weights applied to the default value based on a priority.

【Claim 3】 The system according to claim 2, wherein the priority is assigned in order of whether the tag is manually input, the recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.

【Claim 4】 The system according to claim 1, wherein the user DB comprises:

a tag recommendation DB configured to store a word or a topic term, a tag matching the word or the topic term, and a score of the tag; and

a dictionary DB configured to store a tag and a character string matching the tag.

【Claim 5】 A method of providing a service for building a personal dictionary through character string-tag matching personalized for each user, the method comprising:

a character string selection step in which a user selects a character string for which the user intends to create a tag on a user terminal;

a character string transmission step in which the character string selected in the character string selection step is transmitted from the user terminal to the service provider server;

a character string reception step in which the character string transmitted in the character string transmission step is received by the service provider server;

a first character count determination step in which a number of characters in the character string received in the character string reception step is counted and compared with the first reference value on the service provider server;

a summary extraction step in which a summary obtained

by summarizing the character string is generated when it is determined in the first character count determination step that the number of characters is equal to or greater than the first reference value on the service provider server;

a second character count determination step in which the number of characters in the character string received in the character string reception step or a number of characters in the summary generated in the summary extraction step is counted and compared with a second reference value on the service provider server;

a word extraction step in which all words included in the character string or the summary are extracted when it is determined in the second character count determination step that the number of characters is less than the second reference value on the service provider server;

a keyword extraction step in which a topic term included in the character string or the summary is extracted when it is determined in the second character count determination step that the number of characters is equal to or greater than the second reference value on the service provider server;

a tag recommendation step in which a tag corresponding to the words extracted in the word extraction step or the topic term extracted in the keyword extraction step is extracted from a user DB or a provider DB on the service

provider server and transmitted to the user terminal;

a tag selection and input step in which the user selects one tag from among a plurality of tags transmitted from the service provider server or manually inputs a tag on the user terminal;

a dictionary storage step in which the tag selected or manually input in the tag selection and input step is matched with the character string and stored in the user DB on the service provider server; and

a score calculation step in which the score of the stored tag is calculated and stored in the user DB on the service provider server, wherein:

the tag recommendation step comprises extracting top tags having high scores based on the score stored in the user DB, and transmitting the top tags to the user terminal, and

the score calculation step comprises calculating the score of the tag using, as variables, an accumulated selection count of the tag, whether the tag is manually input, a recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.

【Claim 6】 The method according to claim 5, wherein, when the score is calculated in the score calculation step,

among the score variables, the accumulated selection count of the tag becomes a default value, and the remaining score variables function as weights applied to the default value based on a priority.

【Claim 7】 The method according to claim 6, wherein the priority is assigned in order of whether the tag is manually input, the recent selection history, whether a user is a dictionary creator, and whether the tag is included in the provider DB.