



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G06F 17/27 (2019.02); G06F 17/2785 (2019.02); G06F 17/275 (2019.02); G06F 17/289 (2019.02)

(21) (22) Заявка: 2018121132, 07.06.2018

(24) Дата начала отсчета срока действия патента:
07.06.2018Дата регистрации:
23.04.2019

Приоритет(ы):

(22) Дата подачи заявки: 07.06.2018

(45) Опубликовано: 23.04.2019 Бюл. № 12

Адрес для переписки:

125481, Москва, кв. а/я 44, для Кравченко А.А.

(72) Автор(ы):

Рогачев Игорь Петрович (RU)

(73) Патентообладатель(и):

Рогачев Игорь Петрович (RU)

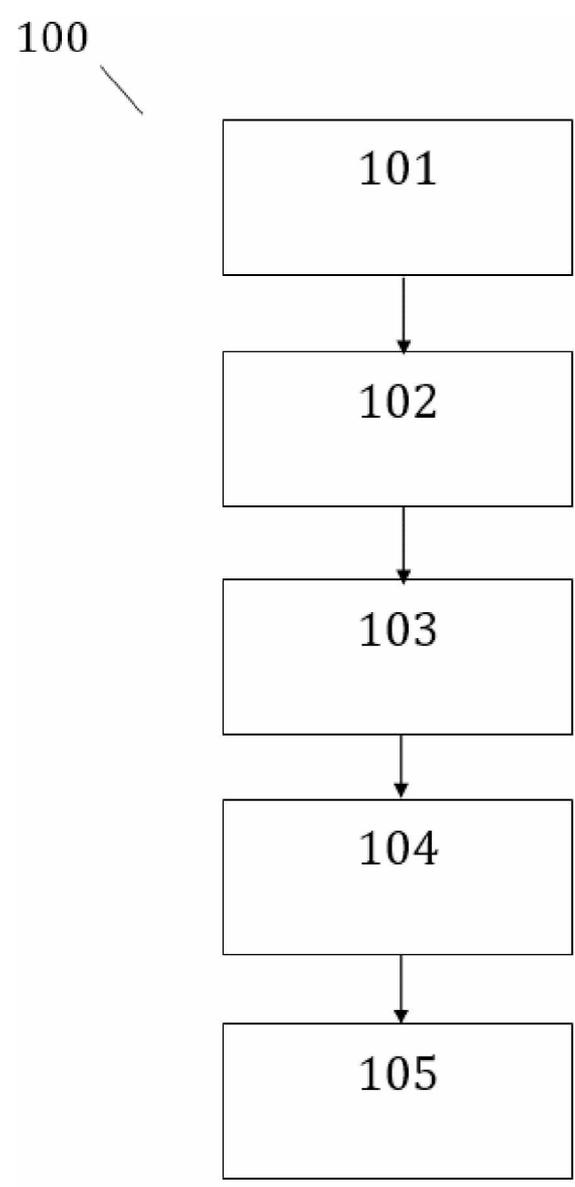
(56) Список документов, цитированных в отчете о поиске: US 2017/0083508 A1, 23.03.2017. US 2012/0109640 A1, 03.05.2012. US 2013/0024186 A1, 24.01.2013. WO 2011/082056 A1, 07.07.2011. RU 2630427 C2, 07.09.2017.

(54) СПОСОБ ПРЕДВАРИТЕЛЬНОГО ПРЕОБРАЗОВАНИЯ СТРУКТУРИРОВАННОГО МАССИВА ДАННЫХ, СОДЕРЖАЩЕГО ЛИНГВИСТИЧЕСКОЕ ПРЕДЛОЖЕНИЕ

(57) Реферат:

Изобретение относится к области вычислительной техники для обработки массивов данных. Технический результат заключается в повышении точности предварительной обработки текста на естественном языке. Технический результат достигается за счет того, что способ преобразования структурированного массива данных (СМД) содержит этап формирования первой структуры данных, на котором формируют первую структуру данных СМД, содержащую элементы упомянутой первой структуры данных, причем упомянутые элементы первой структуры данных представляют собой текстовые элементы (ТЭ) лингвистического предложения, а также идентификационные

данные ТЭ, представляющие собой для каждого ТЭ, по меньшей мере: значение ТЭ и порядковый номер ТЭ в лингвистическом предложении; и этап формирования второй структуры данных, на котором формируют вторую структуру данных СМД, содержащую элементы упомянутой второй структуры данных, причем упомянутые элементы второй структуры данных представляют собой синтаксические единицы (СЕ) лингвистического предложения, сформированные на основании сведений из базы данных лингвистических признаков (БДЛП) текстовых элементов лингвистического предложения, а также представляют собой идентификационные данные СЕ. 4 н. и 17 з.п. ф-лы, 45 ил.



Фиг. 1



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC

G06F 17/27 (2019.02); G06F 17/2785 (2019.02); G06F 17/275 (2019.02); G06F 17/289 (2019.02)

(21) (22) Application: 2018121132, 07.06.2018

(24) Effective date for property rights:
07.06.2018

Registration date:
23.04.2019

Priority:

(22) Date of filing: 07.06.2018

(45) Date of publication: 23.04.2019 Bull. № 12

Mail address:

125481, Moskva, kv. a/ya 44, dlya Kravchenko A.A.

(72) Inventor(s):

Rogachev Igor Petrovich (RU)

(73) Proprietor(s):

Rogachev Igor Petrovich (RU)

(54) **METHOD OF PRELIMINARY TRANSFORMATION OF A STRUCTURED DATA ARRAY CONTAINING A LINGUISTIC SENTENCE**

(57) Abstract:

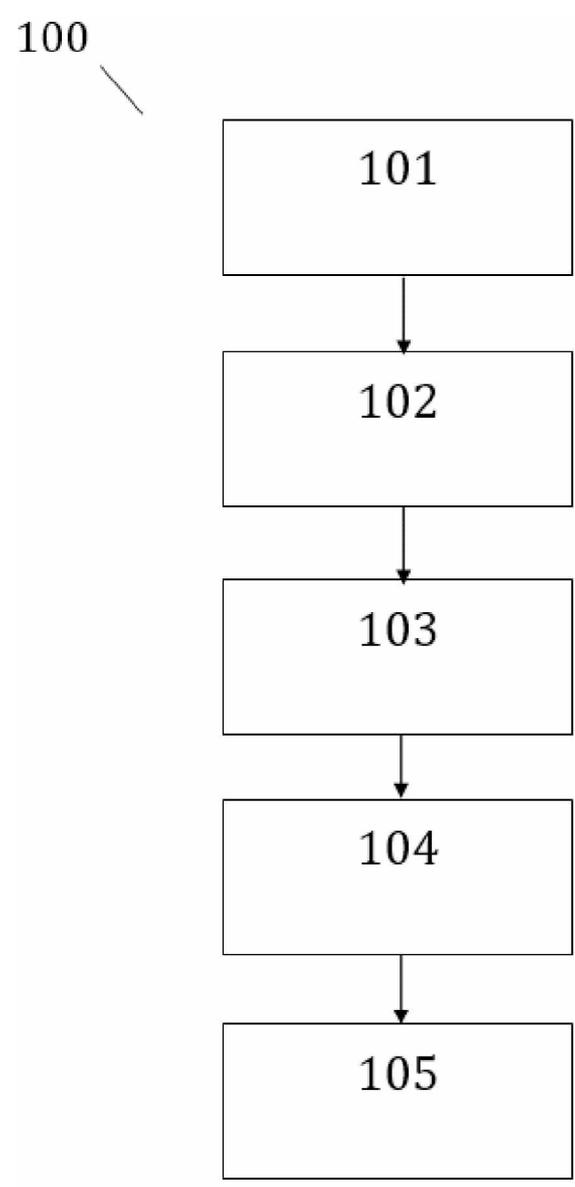
FIELD: computer equipment.

SUBSTANCE: invention relates to computer engineering for processing data arrays. Technical result is achieved due to the fact that the method of transforming a structured data array (SDA) comprises a step of forming a first data structure, on which the first SDA data structure is formed, comprising elements of said first data structure, wherein said elements of the first data structure are text elements (TE) of the linguistic sentence, as well as identification data of TE, which are for each TE, at least: value of TE and TE serial number in linguistic proposal; and a step of

forming a second data structure, on which a second SDA data structure is formed, comprising elements of said second data structure, wherein said elements of the second data structure are syntax units (SU) of the linguistic sentence, formed based on information from the database of linguistic features (DLF) of text elements of the linguistic sentence, and also represent SU identification data.

EFFECT: technical result consists in improvement of accuracy of preliminary processing of text in natural language.

21 cl, 45 dwg



Фиг. 1

[1] ОБЛАСТЬ ТЕХНИКИ, К КОТОРОЙ ОТНОСИТСЯ ИЗОБРЕТЕНИЕ

[2] Группа изобретений относится к решениям в области обработки массивов данных, в частности, к решениям в области обработки структурированных массивов данных, содержащих текст на естественном языке, в частности, лингвистические предложения, и может быть использована для предварительного преобразования структурированного массива данных для обеспечения его последующей обработки.

[3] УРОВЕНЬ ТЕХНИКИ

[4] Из патента РФ 2399959 (ЗАО «АВИКОМП СЕРВИСЕЗ»), опубликованного 10.05.2010 (Д1), известен способ автоматической индексации текстов на естественных языках. Известный из Д1 способ заключается в том, что текст сегментируют в электронной форме на элементарные единицы, выявляют устойчивые словосочетания, формируют предложения, выявляют семантически значимые объекты и семантически значимые отношения между ними, формируют для каждого семантически значимого отношения множество триад, в которых единственная триада первого типа соответствует связи, устанавливаемой семантически значимым отношением между двумя семантически значимыми объектами, при этом каждая из триад второго типа соответствует значению конкретного атрибута одного из этих семантически значимых объектов, каждая из триад третьего типа соответствует значению конкретного атрибута самого семантически значимого отношения, затем индексируют на множестве сформированных триад все связанные семантически значимыми отношениями семантически значимые объекты по отдельности, запоминают в базе данных сформированные триады и полученные индексы вместе со ссылкой на исходный текст, из которого сформированы эти триады.

[5] Однако известный из Д1 способ не обладает достаточной точностью индексации элементов текста на естественном языке, что, в свою очередь, сказывается на точности последующей обработки текста на естественном языке и точности поиска в тексте на естественном языке. Главным образом, это происходит из-за недостаточно эффективной предварительной обработки текста на естественном языке, что, соответственно, не позволяет осуществить индексацию с достаточной точностью.

[6] РАСКРЫТИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

[7] Исходя из этого, технической проблемой, решаемой настоящим изобретением, является создание способа и реализующих способ устройств и/или систем, обладающих повышенной эффективностью предварительной обработки текста на естественном языке для его последующей индексации и обработки.

[8] Соответственно, техническим результатом, достигаемым при реализации настоящего изобретения, является повышение эффективности и точности предварительной обработки текста на естественном языке для его последующей индексации и обработки.

[9] Технический результат достигается за счет того, что заявленный выполняемый процессором компьютерного устройства способ преобразования структурированного массива данных (СМД), содержащего, по меньшей мере, лингвистическое предложение, характеризуется выполнением этапов: А) формирования первой структуры данных, на котором формируют первую структуру данных СМД, содержащую элементы упомянутой первой структуры данных, причем упомянутые элементы первой структуры данных представляют собой текстовые элементы (ТЭ) лингвистического предложения, а также идентификационные данные текстовых элементов (идентификационные данные ТЭ), представляющие собой для каждого ТЭ, по меньшей мере: значение текстового элемента (значение ТЭ) и порядковый номер текстового элемента (порядковый номер ТЭ) в лингвистическом предложении; Б) формирования базы данных лингвистических

признаков (БДЛП), на котором выявляют лингвистические признаки текстовых элементов (лингвистические признаки ТЭ) лингвистического предложения, из которых формируют базу данных, представляющую собой БДЛП текстовых элементов лингвистического предложения; В) формирования второй структуры данных, на котором формируют вторую структуру данных СМД, содержащую элементы упомянутой второй структуры данных, причем упомянутые элементы второй структуры данных представляют собой синтаксические единицы (СЕ) лингвистического предложения, сформированные на основании сведений из БДЛП текстовых элементов лингвистического предложения, а также представляют собой идентификационные данные синтаксических единиц (идентификационные данные СЕ), представляющие собой для каждой СЕ, по меньшей мере: значение синтаксической единицы (значение СЕ) и порядковый (порядковые) номер (номера) ТЭ лингвистического предложения, составляющих СЕ.

[10] Варианты осуществления настоящего изобретения относятся к способам, устройствам, системам и машиночитаемым носителям данных для обеспечения эффективности и точности предварительной обработки текста на естественном языке для его последующей индексации и обработки.

[11] КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

[12] Иллюстративные варианты осуществления настоящего изобретения описываются далее подробно со ссылкой на прилагаемые чертежи, которые включены в данный документ посредством ссылки, и на которых:

[13] На фиг. 1 изображена примерная общая схема выполнения этапов заявленного способа 100 преобразования структурированного массива данных, содержащего, по меньшей мере, лингвистическое предложение, являющегося исходной структурой данных для рассматриваемого способа 100.

[14] На фиг. 2 изображена примерная общая схема выполнения этапов этапа 101 формирования первой структуры данных.

[15] На фиг. 3 изображена примерная общая структура исходной структуры данных, из которой формируется первая структура данных структурированного массива данных.

[16] На фиг. 4 изображена примерная общая структура сформированной первой структуры данных.

[17] На фиг. 5 изображена примерная общая схема выполнения этапов этапа 102 формирования базы данных лингвистических признаков, представляющей собой базу данных лингвистических признаков текстовых элементов 21 предложения 11.

[18] На фиг. 6 изображена примерная общая структура сформированной базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 лингвистического предложения 11.

[19] На фиг. 7 изображена примерная общая схема выполнения этапов этапа 103 формирования второй структуры данных СМД.

[20] На фиг. 8 изображена примерная общая структура сформированной второй структуры данных структурированного массива данных.

[21] На фиг. 9 изображена примерная общая схема выполнения этапов этапа 104 формирования третьей структуры данных структурированного массива данных.

[22] На фиг. 10 изображена примерная общая структура сформированной третьей структуры данных структурированного массива данных.

[23] На фиг. 11 изображена примерная общая схема выполнения этапа 105 формирования четвертой структуры данных структурированного массива данных.

[24] На фиг. 12 изображена примерная общая структура сформированной четвертой

структуры данных структурированного массива данных.

5 [25] На фиг. 13 изображена примерная общая схема выполнения этапов заявленного способа 200 преобразования структурированного массива данных, содержащего, по меньшей мере, синтаксические единицы лингвистического предложения и идентификационные данные синтаксических единиц, являющегося исходной структурой данных для рассматриваемого способа 200.

10 [26] На фиг. 14 изображена примерная общая схема выполнения этапа 201 идентификации пригодной для преобразования структуры данных, содержащей синтаксические единицы, являющейся исходной структурой данных для рассматриваемого способа 200.

[27] На фиг. 15 изображена примерная общая структура данных, являющаяся исходной структурой данных для способа преобразования 200, представляющая собой пригодную для преобразования структуру данных, содержащую синтаксические единицы.

15 [28] На фиг. 16 изображена примерная общая схема выполнения этапов этапа 202 формирования пятой структуры данных структурированного массива данных, являющейся первой структурой данных для рассматриваемого способа 200.

20 [29] На фиг. 17 изображена примерная общая структура сформированной в рамках этапа 2021 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 лингвистического предложения 11, содержащихся в элементах 22 пригодной для преобразования структуры данных, содержащей синтаксические единицы.

25 [30] На фиг. 18 изображена примерная общая структура сформированной в рамках этапа 2022 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 лингвистического предложения 11, содержащихся в элементах 22 пригодной для преобразования структуры данных, содержащей СЕ.

[31] На фиг. 19 изображена примерная общая структура сформированной пятой структуры данных структурированного массива данных, являющейся первой для настоящего способа преобразования 200.

30 [32] На фиг. 20 изображена примерная общая схема выполнения этапа 203 формирования шестой структуры данных структурированного массива данных, являющейся второй структурой данных для рассматриваемого способа 200.

35 [33] На фиг. 21 изображена примерная общая структура сформированной в рамках этапа 2031 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 лингвистического предложения 11, содержащихся в элементах 61 пятой структуры данных структурированного массива данных.

[34] На фиг. 22 изображена примерная общая структура сформированной шестой структуры данных структурированного массива данных, являющаяся второй структурой данных для настоящего способа преобразования 200.

40 [35] На фиг. 23 изображена примерная общая схема выполнения этапов этапа 204 формирования седьмой структуры данных структурированного массива данных, являющейся третьей структурой данных для рассматриваемого способа 200.

45 [36] На фиг. 24 изображена примерная общая структура сформированной седьмой структуры данных структурированного массива данных, являющейся третьей структурой данных для настоящего способа преобразования 200.

[37] На фиг. 25 изображена примерная общая схема выполнения этапов заявленного способа 300 преобразования структурированного массива данных, содержащего, по меньшей мере, основные лингво-логические объекты (ОЛЛО) лингвистического

предложения и идентификационные данные ОЛЛО, являющегося исходной структурой данных для рассматриваемого способа 300.

5 [38] На фиг. 26 изображена примерная общая схема выполнения этапа 301 идентификации пригодной для преобразования структуры данных, содержащей основные лингво-логические объекты (ОЛЛО), являющейся исходной структурой данных для рассматриваемого способа 300.

10 [39] На фиг. 27 изображена примерная общая структура данных, являющаяся исходной для способа преобразования 300, представляющая собой пригодную для преобразования структуру данных, содержащую основные лингво-логические объекты (ОЛЛО).

[40] На фиг. 28 изображена примерная общая схема выполнения этапов этапа 302 формирования восьмой структуры данных структурированного массива данных, являющейся первой структурой данных для рассматриваемого способа 300.

15 [41] На фиг. 29 изображена примерная общая структура формируемой в рамках этапа 3021 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 предложения 11, содержащихся в элементах 91 пригодной для преобразования структуры данных, содержащей ОЛЛО.

20 [42] На фиг. 30 изображена примерная общая структура сформированной восьмой структуры данных СМД, являющейся первой структурой данных СМД для настоящего способа преобразования 300.

[43] На фиг. 31 изображена примерная общая схема выполнения этапов этапа 303 формирования девятой структуры данных СМД, являющейся второй структурой данных для рассматриваемого способа 300.

25 [44] На фиг. 32 изображена примерная общая структура сформированной девятой структуры данных СМД, являющейся второй структурой данных СМД для настоящего способа преобразования 300.

[45] На фиг. 33 изображена примерная общая схема выполнения этапов этапа 304 формирования десятой структуры данных СМД являющейся третьей структурой данных для рассматриваемого способа 300.

30 [46] На фиг. 34 изображена примерная общая структура формируемой в рамках этапа 3041 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 предложения 11, содержащихся в элементах 13 девятой структуры данных СМД.

35 [47] На фиг. 35 изображена примерная общая структура сформированной десятой структуры данных СМД, являющейся третьей структурой данных СМД для настоящего способа преобразования 300.

[48] На фиг. 36 изображена примерная общая схема выполнения этапов этапа 305 формирования одиннадцатой структуры данных СМД, являющейся третьей структурой данных СМД для рассматриваемого способа 300.

40 [49] На фиг. 37 изображена примерная общая структура сформированной одиннадцатой структуры данных СМД, являющейся четвертой структурой данных СМД для настоящего способа преобразования 300.

45 [50] На фиг. 38 изображена примерная общая схема выполнения этапов заявленного способа 400 преобразования СМД, содержащего, по меньшей мере, основные лингво-логические объекты (ОЛЛО) лингвистического предложения и идентификационные данные ОЛЛО, являющегося исходной структурой данных для рассматриваемого способа 400.

[51] На фиг. 39 изображена примерная общая схема выполнения этапов этапа 402

формирования двенадцатой структуры данных СМД, являющейся первой структурой данных для рассматриваемого способа 400.

5 [52] На фиг. 40 изображена примерная общая структура формируемой в рамках этапа 4021 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 предложения 11, содержащихся в элементах 91 пригодной для преобразования структуры данных, содержащей ОЛЛО, формируемой в рамках этапа 4021.

10 [53] На фиг. 41 изображена примерная общая структура сформированной двенадцатой структуры данных СМД являющаяся первой для настоящего способа преобразования 400.

[54] На фиг. 42 изображена примерная общая схема выполнения этапов этапа 403 формирования тринадцатой структуры данных СМД, являющейся второй структурой данных для рассматриваемого способа 400.

15 [55] На фиг. 43 изображена примерная общая структура формируемой в рамках этапа 4031 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 предложения 11, содержащихся в элементах 17 двенадцатой структуры данных СМД.

20 [56] На фиг. 44 изображена примерная общая структура сформированной тринадцатой структуры данных СМД, являющейся второй структурой данных СМД для настоящего способа преобразования 400.

[57] На фиг. 45 изображена примерная общая схема системы для преобразования структурированного массива данных.

[58] ОСУЩЕСТВЛЕНИЕ ИЗОБРЕТЕНИЯ

25 [59] Описанные в данном разделе возможные осуществления вариантов настоящего изобретения представлены на неограничивающих объем правовой охраны примерах, применительно к конкретным вариантам осуществления настоящего изобретения, которые во всех их аспектах предполагаются иллюстративными и не накладывающими ограничения. Альтернативные варианты реализации настоящего изобретения, не выходящие за пределы объема его правовой охраны, являются очевидными
30 специалистам в данной области, имеющим обычную квалификацию, на которых это изобретение рассчитано.

[60] На фиг. 1, в качестве примера, но не ограничения, изображена общая схема выполнения этапов заявленного способа 100 преобразования структурированного массива данных (СМД), содержащего, по меньшей мере, лингвистическое предложение (ЛП), являющегося исходной структурой данных для рассматриваемого способа 100. Заявленный способ 100 преобразования СМД, содержащего, по меньшей мере, лингвистическое предложение, характеризуется: выполнением этапа 101 формирования первой структуры данных, на котором формируют первую структуру данных СМД, содержащую элементы упомянутой первой структуры данных, причем упомянутые
40 элементы первой структуры данных представляют собой текстовые элементы (ТЭ) лингвистического предложения, а также идентификационные данные текстовых элементов (идентификационные данные ТЭ), представляющие собой для каждого ТЭ, в качестве примера, но не ограничения: значение текстового элемента (значение ТЭ) и порядковый номер текстового элемента (порядковый номер ТЭ) в лингвистическом предложении; выполнением этапа 102 формирования базы данных лингвистических признаков (БДЛП), на котором выявляют лингвистические признаки текстовых элементов (лингвистические признаки ТЭ) лингвистического предложения, из которых формируют базу данных представляющую собой базу данных лингвистических

признаков текстовых элементов лингвистического предложения (БДЛП ТЭ лингвистического предложения); выполнением этапа 103 формирования второй структуры данных, на котором формируют вторую структуру данных СМД, содержащую элементы упомянутой второй структуры данных, причем упомянутые
 5 элементы второй структуры данных представляют собой синтаксические единицы (СЕ) лингвистического предложения, сформированные на основании сведений из БДЛП ТЭ лингвистического предложения, а также представляют собой идентификационные
 10 данные синтаксических единиц (идентификационные данные СЕ), представляющие собой для каждой СЕ, в качестве примера, но не ограничения: значение синтаксической единицы (значение СЕ) и порядковый (порядковые) номер (номера) ТЭ лингвистического предложения, составляющих СЕ; выполнением этапа 104 формирования третьей структуры данных, на котором формируют третью структуру данных СМД, содержащую элементы упомянутой третьей структуры данных, причем упомянутые
 15 элементы третьей структуры данных представляют собой лингвистические объекты (ЛО) лингвистического предложения, сформированные путем их отождествления с упомянутыми синтаксическими единицами, либо полученные в результате преобразования упомянутых синтаксических единиц, а также представляют собой
 20 идентификационные данные лингвистических объектов (идентификационные данные ЛО), представляющие собой для каждого ЛО, в качестве примера, но не ограничения: значение лингвистического объекта (значение ЛО) и порядковый (порядковые) номер (номера) ТЭ лингвистического предложения, составляющего (составляющих) ЛО; выполнением этапа 105 формирования четвертой структуры данных, на котором формируют четвертую структуру данных СМД, содержащую элементы упомянутой
 25 четвертой структуры данных, причем упомянутые элементы четвертой структуры данных представляют собой логические объекты (ЛогО) лингвистического предложения, сформированные посредством корреляции лингвистических и логических объектов и отождествленные с упомянутыми лингвистическими объектами, а также представляют собой идентификационные данные логических объектов (идентификационные данные ЛогО), представляющие собой для каждого ЛогО, в качестве примера, но не
 30 ограничения: значение логического объекта (значение ЛогО) и порядковый (порядковые) номер (номера) ТЭ лингвистического предложения, составляющего (составляющих) ЛогО.

[61] На фиг. 2, в качестве примера, но не ограничения, изображена общая схема выполнения этапов этапа 101 формирования первой структуры данных. Этап 101
 35 характеризуется: выполнением этапа 1011 идентификации исходной структуры данных СМД на котором идентифицируют элементы 11 исходной структуры данных СМД, являющиеся лингвистическими предложениями 11 (ЛП 11); выполнением этапа 1012 идентификации элементов 21 первой структуры данных СМД, на котором
 40 идентифицируют элементы 21 первой структуры данных СМД, являющиеся текстовыми элементами (ТЭ) лингвистического предложения 11, а также идентификационные данные элементов 21, представляющие собой для каждого ТЭ, в качестве примера, но не
 ограничения: значение 211 элемента 21 первой структуры данных СМД и порядковый номер 212 элемента 21 первой структуры данных СМД, и формируют первую структуру данных СМД.

[62] На фиг. 3, в качестве примера, но не ограничения, изображена общая структура исходной структуры данных, из которой формируется первая структура данных СМД. Исходные данные представляют собой СМД, содержащий элементы 11 исходной
 45 структуры данных, представляющие собой лингвистические предложения (ЛП). Такой

массив данных представляет собой множество лингвистических предложений 11, относящихся к любой области деятельности и любого назначения. У элементов 11 отсутствуют характеризующие их уникальные наименования (УН), имеющие практическое использование. В исходной структуре данных элементы 11, в качестве
5 примера, но не ограничения, могут именоваться как «ЛП1», «ЛП2», «ЛП3», «ЛПn», где $n \geq 1$ - порядковый номер элемента в лингвистическом предложении.

Лингвистическое предложение – это грамматически организованное соединение слов (множество синтаксически связанных слов), обладающее смысловой и логической
10 завершенностью. Кроме слов ЛП может содержать следующие объекты: цифры (числа), знаки препинания и индексы (конструкции из букв, цифр и знаков). Все перечисленные выше объекты являются компонентами ЛП 11 (компонентами ЛП) и в исходной структуре данных представляют собой отдельные элементы, заранее подготовленные и помещенные в исходную структуру данных не в виде лингвистического текста, а в виде структурированного массива (списка, перечня и тому подобного) отдельных ЛП.
15 Такие подготовительные действия могут осуществляться любым известным из уровня техники способом и, соответственно, далее не описываются.

[63] Идентификация элементов 11 исходной структуры данных в рамках этапа 1011 сводится к обеспечению классификации элементов, из которых состоит исходная структура данных, как лингвистических предложений (ЛП 11). При этом компонентами
20 ЛП являются все цифры (числа), знаки препинания и индексы (конструкции из букв, цифр и знаков), содержащиеся в ЛП и отделенные друг от друга пробелом.

[64] Исходная структура данных представляет собой в итоге множество элементов 11, идентифицированных на этапе 1011.

[65] На фиг. 4, в качестве примера, но не ограничения, изображена общая структура сформированной первой структуры данных. Первая структура данных представляет собой СМД, содержащий элементы 21 первой структуры данных, которые представляют собой текстовые элементы (ТЭ 21) лингвистического предложения 11 и идентификационные данные ТЭ. У текстовых элементов 21 лингвистического предложения 11 отсутствуют характеризующие их уникальные наименования (УН),
30 имеющие практическое использование. В структуре данных элементы 21, в качестве примера, но не ограничения, могут именоваться как «ТЭ», «ТЭ2», «ТЭ3», «ТЭn», где $n \geq 1$ - порядковый номер элемента в лингвистическом предложении 11. ТЭ 21 лингвистического предложения (ЛП 11) являются компонентами ЛП, то есть словами, цифрами (числами), знаками препинания или индексами (конструкциями из букв, цифр
35 и знаков), содержащимися в ЛП 11 и отделенными друг от друга пробелом. Текстовые элементы 21 лингвистического предложения 11 имеют идентификационные данные ТЭ, такие как, в качестве примера, но не ограничения: значение 211 ТЭ и порядковый номер 212 ТЭ. Значением 211 ТЭ являются, в качестве примера, но не ограничения, набор букв, цифр и(или) знаков препинания, из которых состоит ТЭ. Порядковым номером
40 212 ТЭ является порядковый номер ТЭ 21 в лингвистическом предложении 11.

[66] Формирование текстовых элементов 21 первой структуры данных в ходе этапа 1012 производят путем выявления отдельных слов или групп слов, цифр (чисел) или индексов, причем цифры (числа) или индексы не разделены пробелом, а также знаков препинания. При этом предпочтительно, чтобы последний знак препинания в
45 лингвистическом предложении 11 не учитывался и не рассматривался в качестве текстового элемента 21 лингвистического предложения 11.

[67] Идентификацию значения 211 текстового элемента 21 первой структуры данных в ходе этапа 1012 производят путем регистрации символов (букв, цифр и(или) знаков

препинания), из которых состоит текстовый элемент 21. Идентификацию порядкового номера 212 текстового элемента 21 первой структуры данных в ходе этапа 1012 производят путем расчета местоположения ТЭ 21 в лингвистическом предложении 11. При этом первый текстовый элемент 21 в лингвистическом предложении 11 получает
 5 порядковый номер «1», а все последующие ТЭ получают порядковый номер, больший на единицу чем порядковый номер предыдущего ТЭ 21.

[68] Формирование первой структуры данных СМД в ходе этапа 1012 производят путем объединения в одной структуре данных элементов 21 первой структуры данных СМД, а также их идентификационных данных по известным из уровня техники
 10 принципам и способам, которые, соответственно, далее подробно не описываются.

[69] На фиг. 5, в качестве примера, но не ограничения, изображена общая схема выполнения этапов этапа 102 формирования базы данных лингвистических признаков, представляющей собой базу данных лингвистических признаков текстовых элементов 21 предложения 11. Этап 102 характеризуется: выполнением этапа 1021 формирования
 15 лингвистических признаков текстовых элементов 21 предложения 11, на котором для лингвистического анализа текстового элемента предоставляют идентификационные данные элемента 21 (значение 211 и порядковый номер 212 ТЭ) и получают лингвистические характеристики 213 текстовых элементов 21 лингвистического предложения 11, а также значения упомянутых лингвистических характеристик 2131;
 20 выполнением этапа 1022 формирования базы данных лингвистических признаков (БДЛП), на котором формируют БДЛП текстовых элементов 21 лингвистического предложения 11. При этом лингвистическим признаком текстового элемента 21 будут являться все полученные для текстового элемента 21 в ходе этапа 1021 лингвистические характеристики 213, обладающие значениями лингвистических характеристик 2131.

[70] На фиг. 6, в качестве примера, но не ограничения, изображена общая структура сформированной базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 лингвистического предложения 11. Первая часть (основная)
 25 лингвистических характеристик 213 текстовых элементов 21 лингвистического предложения 11 может содержать морфологические, синтаксические и семантические характеристики. При этом совокупность значений всех лингвистических характеристик текстового элемента является для каждого ТЭ 21 лингвистического предложения 11 его отличительным (уникальным) лингвистическим признаком в лингвистическом предложении. Морфологические характеристики предпочтительно указывают на морфологические признаки ТЭ 21 лингвистического предложения 11, которые могут
 30 быть классифицированы, в качестве примера, но не ограничения, по уровню вложенности (род-вид-подвид). При этом морфологическими родами ТЭ 21 лингвистического предложения 11 предпочтительно являются слово, цифра, знаки препинания, иные знаки; морфологическими видами – часть речи (для слов), вид цифры (арабская, римская), вид знака препинания (точка, запятая и тому подобное), вид иного
 40 знака; морфологическими подвидами – род, число, падеж частей речи и тому подобное (для слов), число, двоичный код, индекс и тому подобное (для цифр). Синтаксические характеристики предпочтительно указывают на множество синтаксических признаков ТЭ 21 лингвистического предложения 11, среди которых можно выделить, в качестве примера, но не ограничения, следующие синтаксические характеристики ТЭ 21
 45 лингвистического предложения 11: синтаксическая роль (сказуемое, подлежащее и тому подобное.); синтаксический родитель (синтаксически главное слово); синтаксические потомки (синтаксически подчиненные слова); синтаксическая сочинительная связь (наличие иного ТЭ, имеющего ту же синтаксическую роль и того же синтаксического

родителя). Семантические характеристики предпочтительно указывают на семантические признаки ТЭ 21 лингвистического предложения 11, среди которых можно выделить, в качестве примера, но не ограничения, следующие семантические характеристики ТЭ 21 лингвистического предложения 11: семантическая группа (группа слов, которые
5 можно отнести к одному классу, роду, виду или подвиду предметов или действий окружающего мира при совпадении признаков упомянутых классов, родов, видов или подвидов), семантический статус (смысловое значение слова или группы слов в рамках словосочетания, которым называется некий мыслимый образ (предмет или действие). Например, но не ограничиваясь, мыслимый образ «отсутствие продавца в месте
10 нахождения потребителя» состоит из двух элементов верхнего уровня вложенности (терминов): первого - «отсутствие продавца», и второго - «месте нахождения потребителя», у которых имеются следующие семантические статусы: у первого – главный (определяет смысл термина), у второго - дополнительный (уточняет определенный ранее смысл главного термина)).

[71] Формирование первой части (основной) лингвистических характеристик 213 и их значений 2131 для текстовых элементов 21 лингвистического предложения 11 предпочтительно производят на этапе 1021 путем комплексного лингвистического анализа каждого текстового элемента 21 лингвистического предложения 11, представляющего, в качестве примера, но не ограничения, анализ ТЭ на основе
20 местонахождения ТЭ в структуре предложения, его значения, вида, классификации его мыслимого образа и анализа его связей с другими текстовыми элементами в предложении. По результатам комплексного анализа предпочтительно производится формирование основных лингвистических характеристик 213 и их внесение на этапе 1022 в БДЛП в виде перечня первой части лингвистических характеристик 213 со
25 значениями этих характеристик 2131. Например, но не ограничиваясь, одной из лингвистических характеристик 213 может быть «синтаксическая роль», со значением 2131 данной лингвистической характеристики «подлежащее». Такой анализ может быть выполнен любым известным из уровня техники способом и, соответственно, подробно далее не описывается. Например, не ограничиваясь, такой анализ может быть выполнен
30 традиционно специалистом-лингвистом, или же с помощью программного алгоритма лингвистического (синтаксического) процессора. Более того, при наличии достаточного количества примеров возможно выполнение такого анализа с помощью статистического процессора (нейросети) посредством применения технологии обучения нейросети.

[72] На основании выявленной первой части (основной) лингвистических
35 характеристик 213 текстовых элементов 21 лингвистического предложения 11 и их значений 2131 в итоге формируют базу данных лингвистических признаков, являющуюся БДЛП текстовых элементов 21 лингвистического предложения 11. При этом первая часть лингвистических характеристик 213 текстовых элементов 21 лингвистического предложения 11 и их значения 2131 формирует уникальные лингвистические признаки
40 текстовых элементов 21 лингвистического предложения 11.

[73] На фиг. 7, в качестве примера, но не ограничения, изображена общая схема выполнения этапов этапа 103 формирования второй структуры данных СМД. Этап 103 характеризуется: выполнением этапа 1031 идентификации и формирования первых
45 элементов второй структуры данных СМД, на котором идентифицируют и формируют элементы 31 второй структуры данных СМД, а также идентификационные данные элементов 31, представляющие собой для каждого элемента 31, в качестве примера, но не ограничения, значение 311 элемента 31 второй структуры данных СМД и порядковый (порядковые) номер (номера) 312 ТЭ лингвистического предложения 11, составляющих

элемент 31 (значение и порядковые номера ТЭ элементов 31); выполнением этапа 1032 формирования вторых элементов второй структуры данных СМД, на котором формируют элементы 32 второй структуры данных СМД, а также идентификационные данные элементов 32, представляющие собой для каждого элемента 32, в качестве
 5 примера, но не ограничения, значение 321 элементов 32 второй структуры данных СМД и порядковые номера 322 ТЭ лингвистического предложения 11, составляющих элементы 32 (значение и порядковые номера ТЭ элементов 32); выполнением этапа 1033 формирования второй структуры данных СМД, на котором формируют вторую структуру данных СМД путем объединения идентифицированных элементов 31 и 32 и их идентификационных данных (значений 311, 321 и порядковых номеров 312, 322 ТЭ).
 10

[74] На фиг. 8, в качестве примера, но не ограничения, изображена общая структура сформированной второй структуры данных СМД. Вторая структура данных СМД представляет собой СМД, содержащий элементы 31 и (или) элементы 32, которые представляют собой синтаксические единицы (СЕ) лингвистического предложения 11
 15 и идентификационные данные СЕ, представляющие собой для каждого СЕ, в качестве примера, но не ограничения, значение 311 элемента 31 и (или) значение 321 элемента 32, а также порядковый (порядковые) номер (номера) 312 и (или) 322 ТЭ 21 лингвистического предложения, составляющего (составляющих) элементы 31 и (или) элементы 32. Синтаксическими единицами (СЕ) предложения 11 предпочтительно
 20 являются отдельные слова и группы слов, связанные синтаксической связью. У СЕ лингвистического предложения 11 имеются характеризующие их уникальные наименования (УН СЕ), имеющие практическое использование, в качестве примера, но не ограничения: «подлежащее», «сказуемое», «прямое дополнение» и тому подобное. В структуре данных элементы 31 и 32 именуется уникальными наименованиями и, в
 25 качестве примера, но не ограничения, могут именоваться как «УН СЕ n», где $n \geq 1$ -порядковый индекс элемента уникального наименования в лингвистическом предложении, начиная с «1» для каждого нового УН СЕ, ранее не встречавшегося в лингвистическом предложении. В качестве примера, но не ограничения для пояснения определения порядковых индексов элементов структуры данных, имеющих уникальное
 30 наименование (СЕ и иных, указанных далее) определение порядкового индекса УН СЕ может быть продемонстрировано следующим образом. В предложении: «Покупатель обязан оплатить и забрать оплаченные товар, тару и упаковку» имеются семь СЕ четырех уникальных наименований: «подлежащее» - одна СЕ, «сказуемое» - две СЕ, «определение» - одна СЕ и «дополнение» - три СЕ. В структуре данных будет
 35 содержаться один элемент «подлежащее», который будет иметь индекс «1», а также два элемента «сказуемое», которые будут иметь соответствующие индексы «1» и «2», причем индекс «1» будет иметь то «сказуемое», которое имеет меньший порядковый номер ТЭ, а также один элемент «определение», который будет иметь индекс «1», а также три элемента «дополнение», которые будут иметь соответствующие индексы
 40 «1», «2» и «3», причем индекс «1» будет иметь то «дополнение», которое имеет меньший порядковый номер ТЭ. В дальнейшем, при упоминании СЕ как элемента второй структуры данных имеется ввиду УН СЕ. По своей структуре СЕ могут состоять из одного слова или из нескольких слов. СЕ из одного слова (одного текстового элемента 21 предложения 11) – это всегда элемент 31, являющийся исходным синтаксическим
 45 объектом (ИСО), то есть членом предложения, установленным в результате синтаксического анализа. В дальнейшем, при упоминании элемента 31 второй структуры данных, являющегося ИСО, имеется в виду УН ИСО. При этом у членов предложения можно выделить множество характеристик, которые на практике могут породить

перечень уточненных членов предложения. Таким образом список ИСО – это не заранее установленный список, а варьируемый, зависящий от предварительно заданных параметров список. Одно и то же слово (текстовый элемент 21 предложения 11) в предложении 11 может быть поименовано разной синтаксической единицей 31, в зависимости от наличия у данного текстового элемента 21 предложения 11 различных лингвистических характеристик 213 и их значений 2131, а также актуальности отдельных лингвистических характеристик. Например, но не ограничиваясь, при идентификации ИСО «дополнение», в котором не актуальны некоторые лингвистические характеристики 213 данного ИСО, к одному уникальному названию синтаксических единиц ИСО 31 будут отнесены и ИСО «прямое дополнение», и ИСО «косвенное дополнение». При актуализации всех лингвистических характеристик 213 ИСО как отдельных синтаксических единиц 31 будут идентифицированы ИСО «прямое дополнение» и ИСО «косвенное дополнение». Синтаксические единицы из нескольких слов (нескольких текстовых элементов 21 предложения 11) – это иногда ИСО (элемент 31), а иногда синтаксические конструкции (элемент 32). ИСО (элемент 31) из нескольких слов (текстовых элементов 21 предложения 11) встречаются, в частности, в ИСО «сказуемое». Например, но не ограничиваясь, ИСО «сказуемое» - «мог быть выполнен», - состоит из трех текстовых элементов 21 предложения 11 ввиду лингвистической особенности формирования сказуемого, которое, помимо глагола «выполнен» может иметь вспомогательный глагол «быть» и модальность (модальный глагол) «мог». Синтаксическая конструкция 32 всегда состоит из нескольких слов (текстовых элементов 21 предложения 11) и представляет собой конструкцию из синтаксически связанных ИСО (элементов 31), в качестве примера, но не ограничения, синтаксической конструкцией (УН СК) является «словосочетание», «синтаксический оборот» (причастный, деепричастный), «простое предложение», «сложное предложение» и т.д. В дальнейшем, при упоминании элемента 32 второй структуры данных, являющегося синтаксической конструкцией (СК), имеется ввиду УН СК. На практике, в целях решения прикладных задач, возможно формирование самых различных синтаксических конструкций, при этом перечень и описание синтаксических конструкций могут быть заданы предварительно. Перечень SE 31 (ИСО) и SE 32 (СК), которые подлежат идентификации в предложении 11, предпочтительно задается предварительно.

[75] Идентификацию или формирование первого элемента 31 второй структуры данных СМД в рамках этапа 1031 производят путем комплексного анализа значений 2131 лингвистических характеристик 213 текстовых элементов 21 лингвистического предложения 11. Элементы 31 второй структуры данных СМД, являющиеся ИСО, идентифицируемые (если состоят из одного ТЭ) или формируемые (если состоят из нескольких ТЭ) в рамках этапа 1031, фактически отождествляют с актуальными ИСО. Актуальные ИСО – это актуальный перечень ИСО с актуальными лингвистическими характеристиками, установленный предварительно и записанный в первую пользовательскую базу данных (первую ПБД), являющуюся таким образом базой данных актуальных синтаксических единиц (СЕ), содержащей актуальные ИСО и актуальные СК. Идентификацию и формирование (если требуется) элементов 31 ИСО производят путем сравнения зарегистрированных в базе данных лингвистических признаков (БДЛП) значений 2131 лингвистических характеристик 213 текстовых элементов 21 лингвистического предложения 11 с предварительно заданными значениями лингвистических характеристик актуальных ИСО. При совпадении значений 2131 лингвистических характеристик 213 текстового элемента 21 лингвистического предложения 11 и предварительно заданных значений лингвистических характеристик

актуальных ИСО выявленный текстовый элемент идентифицируют и при необходимости осуществляют его формирование как соответствующего ИСО (элемент 31). Такой комплексный анализ может быть выполнен любым известным из уровня техники способом и, соответственно, подробно далее не описывается. Например, не
 5 ограничиваясь, такой комплексный анализ может быть выполнен традиционно специалистом-лингвистом, или же с помощью программного алгоритма лингвистического (синтаксического) процессора. Более того, при наличии достаточного количества примеров возможно выполнение такого анализа с помощью статистического процессора (нейросети) посредством применения технологии обучения нейросети.
 10 Формирование второго элемента 32 второй структуры данных СМД в рамках этапа 1032 производят путем комплексного анализа значений 2131 лингвистических характеристик 213 текстовых элементов, являющихся ИСО (элементов 31). При выявлении лингвистических характеристик текстовых элементов, составляющих ИСО, которые соответствуют описанию синтаксической конструкции (СК), содержащемуся
 15 в базе данных актуальных СЕ, формируют СК (элементы 32) из упомянутых двух или более ИСО (элементов 31). Такой комплексный анализ может быть выполнен любым известным из уровня техники способом и, соответственно, подробно далее не описывается. Например, не ограничиваясь, такой комплексный анализ может быть
 20 выполнен традиционно специалистом-лингвистом, или же с помощью программного алгоритма лингвистического (синтаксического) процессора. Более того, при наличии достаточного количества примеров возможно выполнение такого анализа с помощью статистического процессора (нейросети) посредством применения технологии обучения нейросети.

[76] Идентификацию значения и порядкового (порядковых) номера (номеров) ТЭ
 25 элемента 31 (ИСО) второй структуры данных СМД в рамках этапа 1032 производят следующим образом: значение (значения) 211 текстового (текстовых) элемента (элементов) 21 лингвистического предложения 11, составляющего (составляющих) элемент 31, идентифицируют как значение 311 элемента 31 (ИСО), а порядковый (порядковые) номер (номера) 212 упомянутого (упомянутых) текстового (текстовых)
 30 элемента (элементов) 21 лингвистического предложения 11, составляющего (составляющих) элемент 31, идентифицируют как порядковый (порядковые) номер (номера) 312 текстового (текстовых) элемента (элементов), составляющего (составляющих) элемент 31 (ИСО). Идентификацию значения и порядковых номеров ТЭ элемента 32 (СК) второй структуры данных СМД в рамках этапа 1032 производят
 35 следующим образом: значения 211 текстовых элементов 21 лингвистического предложения 11, составляющих элемент 32, идентифицируют как значение 321 элемента 32 (СК), а порядковые номера 212 упомянутых текстовых элементов 21 лингвистического предложения 11, составляющих элемент 32 идентифицируют как порядковые номера текстовых элементов 322, составляющих элемент 32 (СК).

[77] Формирование второй структуры данных СМД в ходе этапа 1033 производят путем объединения в одной структуре данных элементов 31 второй структуры данных СМД и элементов 32 второй структуры данных СМД, а также их идентификационных данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются.

[78] Идентификацию лингвистических характеристик ТЭ, составляющих элементы 31 (ИСО) второй структуры данных СМД и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 102, состоящего из идентификационных данных ТЭ, составляющих ИСО 31, и получения значений 2131

лингвистических характеристик 213 текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 31 (ИСО). Идентификация лингвистических характеристик ТЭ, составляющих элементы 32 (СК) второй структуры данных СМД и их значений при необходимости производится путем организации запроса в БДЛП, формируемую в рамках этапа 102, состоящего из идентификационных данных ТЭ, составляющих СК 32, и получения значений 2131 лингвистических характеристик 213 текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 32 (СК). При этом, как было описано ранее, лингвистическими признаками элементов 31 (ИСО) и 32 (СК) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 31 (ИСО) и 32 (СК).

[79] На фиг. 9, в качестве примера, но не ограничения, изображена общая схема выполнения этапов этапа 104 формирования третьей структуры данных СМД. Этап 104 характеризуется: выполнением этапа 1041 идентификации первого элемента третьей структуры данных СМД, на котором идентифицируют или формируют элементы 41 третьей структуры данных СМД, а также идентификационные данные элементов 41, представляющие собой для каждого элемента 41, в качестве примера, но не ограничения, значение 411 элемента 41 третьей структуры данных СМД и порядковый (порядковые) номер (номера) 412 ТЭ лингвистического предложения 11, составляющего (составляющих) элемент 41; выполнением этапа 1042 формирования второго элемента третьей структуры данных СМД, на котором формируют элементы 42 третьей структуры данных СМД а также идентификационные данные элементов 42, представляющие собой для каждого элемента 42, в качестве примера, но не ограничения, значение 421 элемента 42 третьей структуры данных СМД и порядковый (порядковые) номер (номера) 422 ТЭ лингвистического предложения 11, составляющего (составляющих) элемент 42; выполнением этапа 1043 формирования третьей структуры данных СМД, на котором формируют третью структуру данных СМД путем объединения идентифицированных элементов 41 и 42 и их идентификационных данных.

[80] На фиг. 10, в качестве примера, но не ограничения, изображена общая структура сформированной третьей структуры данных СМД. Третья структура данных СМД представляет собой СМД, содержащий элементы 41 и (или) элементы 42, которые представляют собой лингвистические объекты (ЛО) лингвистического предложения 11 и идентификационные данные ЛО, представляющие собой для каждого ЛО, в качестве примера, но не ограничения, значение 411 элемента 41 и (или) значение 421 элемента 42 (значение элементов 41 и (или) 42), а также порядковый (порядковые) номер (номера) 412 и (или) 422 текстовых элементов 21 лингвистического предложения, составляющего (составляющих) элемент 41 и (или) элемент 42 (порядковые номера ТЭ элементов 41 и (или) 42). Лингвистическим объектом (ЛО) является особая синтаксическая единица (ОСЕ). У ЛО лингвистического предложения 11 имеются характеризующие их уникальные наименования (УН ЛО), имеющие практическое использование в качестве примера, но не ограничения, представляет собой: «простое предложение без однородных членов и оборотов», «первое словосочетание» и тому подобное. В структуре данных элементы 41 и 42 именуется уникальными наименованиями, в качестве примера, но не ограничения, могут именоваться как «УН ЛО n», где $n \geq 1$ - порядковый индекс элемента уникального наименования в лингвистическом предложении, начиная с «1» для каждого нового УН ЛО, ранее не встречавшегося в лингвистическом предложении. В дальнейшем, при упоминании ЛО как элемента третьей структуры данных или элементов 41 и 42 третьей структуры данных имеется ввиду УН ЛО. ОСЕ предварительно формируют в

рамках лингво-логического преобразования, результатом которого является выполнение корреляции лингвистических и логических объектов и преобразование синтаксических единиц предложения в массив логических объектов. Механизм лингво-логического преобразования позволяет на основании заранее известного представления о логической модели предложения (логической структуре простого неосложненного предложения, идентичного простому суждению) установить корреляцию (взаимосвязь) между логическими объектами предложения и синтаксическими единицами предложения. В ходе формирования таблицы взаимосвязанных (лингвистических и логических) объектов в каждом отдельном случае формируются такие ОСЕ, являющиеся искомыми лингвистическими объектами (искомые ЛО), которые могут быть отождествлены с искомыми логическими объектами (искомыми ЛогО). Сформированный таким образом искомый ЛО может быть полностью тождественен одному или нескольким СЕ (такой искомый ЛО будет порождать первый лингвистический объект (первый ЛО), являющийся элементом 41 лингвистического предложения 11), либо представлять собой производный синтаксический объект, полученный путем преобразования одного или нескольких СЕ в объект, который будет являться объектом, тождественным искомому ЛО (такой искомый ЛО будет порождать второй лингвистический объект (второй ЛО), являющийся элементом 42 лингвистического предложения 11). Искомые ЛогО могут быть отдельные логические объекты или группы логических объектов, которые являются логическими сущностями и сформированы по законам логики (науке о мышлении). При этом существуют базовые академические логические сущности, в качестве примера, но не ограничения, такие как «понятие», «суждение», «субъект суждения» (то, о чем что-либо утверждается или отрицается в суждении), «предикат суждения» (то, что утверждается или отрицается о субъекте суждения). Вместе с тем, в практических целях могут быть сформированы и иные логические сущности, отражающие потребности текущей прикладной задачи и используемые в рамках лингво-логического преобразования. Например, не ограничиваясь, в предикате суждения могут быть выделены отдельные более функционально однородные логические сущности, нежели сам предикат суждения, а именно – «действие предиката суждения», «объект предиката суждения» и «обстоятельства предиката суждения».

[81] Идентификацию элемента 41 (первый ЛО) третьей структуры данных СМД в рамках этапа 1041 производят путем анализа состава искомых ЛО. При этом состав искомых ЛО предварительно задают в виде СЕ уникального наименования (УН СЕ) или списка СЕ уникальных наименований с указанием способа формирования ЛО, которым, не ограничиваясь, может быть один из следующих способов: отождествление с установленным (несколькими установленными) СЕ или преобразование установленного (установленных) СЕ по описанному ранее способу. Упомянутые состав искомых ЛО и способы формирования ЛО формируют собой вторую пользовательскую базу данных (вторую ПБД), являющуюся таким образом базой данных искомых ЛО, содержащую, в том числе, и лингвистические характеристики текстовых элементов, из которых формируются искомые ЛО. Если в результате анализа состава конкретного ЛО будет установлено, что анализируемый ЛО состоит из отождествляемых СЕ (элементов 31 или элементов 32), то такой ЛО идентифицируют (если тождественен одному СЕ) или формируют (если тождественен нескольким СЕ) как элемент 41 третьей структуры данных СМД. Идентифицированные в рамках анализа состава лингвистических объектов синтаксические единицы (элементы 31 или элементы 32) второй структуры данных СМД в соответствии с предварительно заданными параметрами идентификации или формирования элементов 41 третьей структуры данных

СМД содержащиеся в описании искомым лингвистических объектов, отождествляют с первым ЛО (первым лингвистическим объектом (элементом 41 третьей структуры данных СМД)), благодаря чему формируют элемент 41 третьей структуры данных СМД. Дополнительной частью элемента 41 третьей структуры данных СМД могут
 5 быть служебные части речи (предлоги, союзы, частицы) и знаки препинания, находящиеся между объединяемыми в рамках элемента 41 третьей структуры данных СМД синтаксическими единицами в соответствии со способом формирования первого ЛО, установленным пользователем. Если в результате анализа состава конкретного ЛО будет установлено, что анализируемый ЛО состоит из преобразуемых СЕ, то такой
 10 лингвистический объект будет относиться к элементам 42 третьей структуры данных СМД (вторым ЛО) и формироваться в рамках этапа 1042. Формирование элемента 42 третьей структуры данных СМД (второго ЛО) в рамках этапа 1042 производят путем преобразования СЕ, предварительно заданных в описании искомым ЛО вида «второй ЛО» по предварительно заданному способу преобразования во второй ПБД.
 15 Преобразование СЕ в элементы 42 третьей структуры данных СМД является необязательным действием для настоящего способа преобразования структурированного массива данных. Настоящий способ преобразования структурированного массива данных может быть с достаточной эффективностью и точностью реализован даже при осуществлении лишь идентификации элемента 41 третьей структуры данных СМД при
 20 условии, что все искомые ЛО порождают только первый ЛО, являющийся элементом 41 лингвистического предложения 11.

[82] Идентификацию значения и порядкового (порядковых) номера (номеров) ТЭ, элемента 41 (первый ЛО) третьей структуры данных СМД в рамках этапа 1041 производят следующим образом: значение (значения) 311 или 321 элемента (элементов)
 25 31 или 32 лингвистического предложения 11, отождествленного (отождествленных) с элементом 41 (первый ЛО), идентифицируют как значение 411 элемента 41 (первый ЛО), а порядковый (порядковые) номер (номера) 312 или 322 ТЭ элемента (элементов) 31 или 32 лингвистического предложения 11, отождествленного (отождествленных) с элементом 41 (первый ЛО) идентифицируют как порядковый (порядковые) номер
 30 (номера) 412 текстового (текстовых) элемента (элементов), составляющего (составляющих) элемент 41 (первый ЛО). Если формирование элемента 41 производят согласно требованиям второй ПБД путем объединения нескольких СЕ, то к значениям и номерам ТЭ элемента 41 (первый ЛО) третьей структуры данных СМД добавляют значения и порядковые номера ТЭ служебных частей речи (предлоги, союзы, частицы)
 35 и знаки препинания, находящиеся между объединяемыми СЕ. Идентификацию значения и порядкового (порядковых) номеров ТЭ элемента 42 (второй ЛО) третьей структуры данных СМД в рамках этапа 1042 производят следующим образом: значение (значения) преобразованных СЕ (из элемента (элементов) 31 и (или) 32 лингвистического предложения 11), отождествленного (отождествленных) с элементом 42 (второй ЛО),
 40 идентифицируют как значение 421 элемента 42 (второй ЛО), а порядковый (порядковые) номер (номера) преобразованных СЕ (из элемента (элементов) 31 и (или) 32 лингвистического предложения 11), отождествленного (отождествленных) с элементом 42 (второй ЛО) идентифицируют как порядковый (порядковые) номер (номера) 422 текстового (текстовых) элемента (элементов), составляющего (составляющих) элемент
 45 42 (второй ЛО).

[83] Формирование третьей структуры данных СМД в ходе этапа 1043 производят путем объединения в одной структуре данных элементов 41 третьей структуры данных СМД и элементов 42 третьей структуры данных СМД, а также их идентификационных

данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются.

[84] Идентификацию лингвистических характеристик ТЭ, составляющих элементы 41 (первый ЛО) третьей структуры данных СМД и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 102, состоящего из идентификационных данных ТЭ, составляющих первый ЛО 41, и получения значений 2131 лингвистических характеристик 213 текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 41 (первый ЛО). Идентификацию лингвистических характеристик ТЭ, составляющих элементы 42 (второй ЛО) третьей структуры данных СМД и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 102, состоящего из идентификационных данных ТЭ, составляющих второй ЛО 42, и получения значений 2131 лингвистических характеристик 213 текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 42 (второй ЛО). При этом, как было описано ранее, лингвистическими признаками элементов 41 (первый ЛО) и 42 (второй ЛО) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 41 (первый ЛО) и 42 (второй ЛО).

[85] На фиг. 11, в качестве примера, но не ограничения, изображена общая схема выполнения этапа 105 формирования четвертой структуры данных СМД. Этап 105 характеризуется формированием элемента четвертой структуры данных, при котором формируют элементы 51 четвертой структуры данных СМД, а также идентификационные данные элементов 51, представляющие собой для каждого элемента 51, в качестве примера, но не ограничения, значение 511 элемента 51 четвертой структуры данных СМД и порядковый (порядковые) номер (номера) 512 ТЭ лингвистического предложения 11, составляющего (составляющих) элемент 51.

[86] На фиг. 12, в качестве примера, но не ограничения, изображена общая структура сформированной четвертой структуры данных СМД. Четвертая структура данных СМД представляет собой СМД, содержащий элементы 51 четвертой структуры данных СМД, которые представляют собой логические объекты (ЛогО) лингвистического предложения 11 и идентификационные данные ЛогО, представляющие собой для каждого ЛогО, в качестве примера, но не ограничения, значение 511 элемента 51 четвертой структуры данных СМД (значение элемента 51) и порядковый (порядковые) номер (номера) 512 ТЭ, составляющего (составляющих) элемент 51 (порядковые номера 512 ТЭ элемента 51). Логические объекты (ЛогО) являются логическими сущностями и сформированы по законам логики (науки о мышлении). У ЛогО лингвистического предложения 11 имеются характеризующие их уникальные наименования (УН ЛогО), имеющие практическое использование, в качестве примера, но не ограничения, - «понятие», «суждение», «субъект суждения» (то, о чем что-либо утверждается или отрицается в суждении), «предикат суждения» (то, что утверждается или отрицается о субъекте суждения и тому подобное. В структуре данных элементы 51 именуется уникальными наименованиями, в качестве примера, но не ограничения, могут именоваться как «УН ЛогО n», где $n \geq 1$ - порядковый индекс элемента уникального наименования в лингвистическом предложении, начиная с «1» для каждого нового УН ЛогО, ранее не встречавшегося в лингвистическом предложении элемента в лингвистическом предложении. В дальнейшем, при упоминании ЛогО как элемента четвертой структуры данных или элемента 51 четвертой структуры данных имеется в виду УН ЛогО. Массив логических объектов предложения представляет собой иной,

нелингвистический способ регистрации смысла того, о чем идет речь в лингвистическом предложении. В отличие от лингвистического предложения, смысл которого может быть сохранен только при условии регистрации всех текстовых элементов предложения в одном месте (в виде одной единицы информации), массив логических объектов позволяет хранить ту же информацию без потери смысла в разделенном виде, во множестве независимых мест хранения информации (в виде массива логических объектов). При этом формирование массива логических объектов позволяет не просто разделить единую информацию на части, но и системно структурировать эти части для любых преобразуемых в рамках настоящего способа лингвистических предложений.

Возможность регистрации и хранения логически и системно структурированной информации по частям открывает новые возможности для высокоточной обработки такой информации, поскольку в этом случае поиск и анализ информации технически может проводиться не в едином объеме исходной информации (лингвистическом предложении), а только в определенных логических частях этой информации, характеристики и требования к которым соответствуют логической сути и логическому предмету поиска или анализа. В результате такой способ хранения информации существенно уменьшает объем (зону) поиска и анализа, что приводит к уменьшению времени и повышению качества (релевантности, точности) поиска и анализа.

[87] Формирование элемента 51 (ЛогО) четвертой структуры данных СМД в рамках этапа 105 производят с помощью третьей структуры данных СМД, содержащей лингвистические объекты (ЛО) и базы данных корреляции лингвистических и логических объектов, представляющей собой, таким образом, третью пользовательскую базу данных (третью ПБД). Третью ПБД формируют для лингво-логического преобразования предварительно. Третья ПБД представляет собой перечень необходимых логических объектов, которые могут быть выделены в лингвистическом предложении и на которые оно может быть в итоге разделено. Кроме этого, третья ПБД содержит перечень лингвистических объектов уникальных наименований (УН ЛО) и описаний лингвистических объектов, которые предполагают тождественными искомым логическим объектам уникальных наименований (УН ЛогО) при указанных условиях (указанном составе УН ЛО (перечень СЕ) и (или) указанном способе преобразования СЕ, а также указанных лингвистических характеристиках упомянутых СЕ). Для формирования логического объекта 51, коррелированного с выбранным лингвистическим объектом 41 или 42 в третьей структуре данных выбирают отдельный ЛО (из перечня ЛО) и с помощью третьей ПБД выявляют коррелированный с УН ЛО логический объект (УН ЛогО), в качестве примера, но не ограничения, выбирают из третьей структуры данных ЛО с уникальным наименованием (УН ЛО) «первое словосочетание» (словосочетание состоящее из простого предложения без однородных членов и оборотов (например, подлежащее) и все зависимые от него слова)) и направляют этот УН ЛО в качестве запроса в третью ПБД для поиска выбранного УН ЛО в перечне уникальных наименований имеющих там ЛО. При обнаружении такого УН ЛО (например, «первое словосочетание») в соответствии с третьей ПБД получают отождествленное с ним УН логического объекта (в приведенном примере УН ЛогО - «субъект суждения»). Таким образом формируют логический объект (элемент 51 четвертой структуры данных СМД) для выбранного лингвистического объекта 41 или 42. Результатом формирования элемента 51 является выявленное УН ЛогО.

[88] Идентификацию значения и порядковых номеров ТЭ элемента 51 (ЛогО) четвертой структуры данных СМД в рамках этапа 105 производят следующим образом: значение (значения) 211 текстового (текстовых) элемента (элементов) 21

лингвистического предложения 11, составляющего (составляющих) коррелированный с элементом 51 элемент 41 или 42, идентифицируют как значение 511 элемента 51 (ЛогО), а порядковый (порядковые) номер (номера) 212 упомянутого (упомянутых) текстового (текстовых) элемента (элементов) 21 лингвистического предложения 11, составляющего (составляющих) коррелированный с элементом 51 элемент 41 или 42 идентифицируют как порядковый (порядковые) номер (номера) текстового (текстовых) элемента (элементов) 512, составляющего (составляющих) элемент 51 (ЛогО).

[89] Формирование четвертой структуры данных СМД в ходе этапа 105 производят путем объединения в одной структуре данных элементов 51 четвертой структуры данных СМД, а также их идентификационных данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются.

[90] Идентификацию лингвистических характеристик ТЭ, составляющих элементы 51 (ЛогО) четвертой структуры данных СМД и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 102, состоящего из идентификационных данных ТЭ, составляющих ЛогО 51, и получении значений 2131 лингвистических характеристик 213 текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 51 (ЛогО). При этом, как было описано ранее, лингвистическими признаками элементов 51 (ЛогО) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 51 (ЛогО).

[91] Далее, после завершения описанного ранее этапа 103, в целях повышения точности последующего поиска в структурированном массиве данных становится возможным осуществить дальнейшее преобразование СМД, содержащего, по меньшей мере, синтаксические единицы (СЕ) лингвистического предложения и их идентификационные данные.

[92] На фиг. 13, в качестве примера, но не ограничения, изображена общая схема выполнения этапов заявленного способа 200 преобразования СМД, содержащего, по меньшей мере, синтаксические единицы (СЕ) лингвистического предложения и идентификационные данные СЕ, являющегося исходной структурой данных для рассматриваемого способа 200. Заявленный способ 200 преобразования СМД, содержащего, по меньшей мере, СЕ лингвистического предложения и идентификационные данные СЕ, характеризуется: выполнением этапа 201 идентификации пригодной для преобразования структуры данных, содержащей СЕ, на котором идентифицируют структуру данных СМД, содержащую элементы упомянутой пригодной для преобразования структуры данных, содержащей СЕ, являющейся исходной структурой данных СМД в рамках рассматриваемого способа 200, причем упомянутые элементы пригодной для преобразования структуры данных представляют собой синтаксические единицы (СЕ) лингвистического предложения и идентификационные данные СЕ; выполнением этапа 202 формирования пятой структуры данных СМД, на котором формируют пятую структуру данных СМД, являющуюся первой структурой данных СМД в рамках рассматриваемого способа 200, содержащую элементы упомянутой пятой структуры данных СМД, причем упомянутые элементы пятой структуры данных СМД представляют собой лингво-логические единицы (ЛЛЕ) лингвистического предложения, идентифицированные и сформированные по итогам лингво-логического анализа синтаксических единиц, а также представляют собой идентификационные данные ЛЛЕ, представляющие собой для каждого ЛЛЕ, в качестве примера, но не ограничения: значение ЛЛЕ и порядковый (порядковые) номер (номера)

ТЭ лингвистического предложения, составляющего (составляющих) ЛЛЕ; выполнением этапа 203 формирования шестой структуры данных СМД, на котором формируют шестую структуру данных СМД, являющуюся второй структурой данных СМД в рамках рассматриваемого способа 200, содержащую элементы упомянутой шестой структуры СМД, причем упомянутые элементы шестой структуры данных СМД представляют собой группы лингво-логических единиц (группы ЛЛЕ) лингвистического предложения сформированные на основании схемы синтаксических связей ЛЛЕ в простом предложении исходного лингвистического предложения, а также представляют собой идентификационные данные групп ЛЛЕ, представляющие собой для каждой группы ЛЛЕ, в качестве примера, но не ограничения: значение группы ЛЛЕ и порядковые номера ТЭ лингвистического предложения, составляющих группу ЛЛЕ; выполнением этапа 204 формирования седьмой структуры данных СМД, на котором формируют седьмую структуру данных СМД, являющуюся итоговой структурой данных СМД для рассматриваемого способа 200, содержащую элементы упомянутой седьмой структуры данных СМД, причем упомянутые элементы седьмой структуры данных СМД представляют собой основные лингво-логические объекты (ОЛЛО) лингвистического предложения сформированные из групп ЛЛЕ путем устранения однородностей в группах ЛЛЕ, а также представляют собой идентификационные данные ОЛЛО, представляющие собой для каждого ОЛЛО, в качестве примера, но не ограничения: значение ОЛЛО и порядковый (порядковые) номер (номера) ТЭ лингвистического предложения, составляющего (составляющих) ОЛЛО.

[93] На фиг. 14, в качестве примера, но не ограничения, изображена общая схема выполнения этапа 201 идентификации пригодной для преобразования структуры данных, содержащей синтаксические единицы (СЕ), являющейся исходной структурой данных для рассматриваемого способа 200. Этап 201 характеризуется выполнением идентификации пригодной для преобразования структуры данных содержащей СЕ, на котором идентифицируют элементы 22 пригодной для преобразования структуры данных содержащей СЕ, а также идентификационные данные элементов 22, представляющие собой для каждого элемента 22, в качестве примера, но не ограничения, значение 221 элемента 22 пригодной для преобразования структуры данных, содержащей СЕ и порядковый (порядковые) номер (номера) 222 ТЭ лингвистического предложения 11, составляющего (составляющих) элемент 22.

[94] На фиг. 15, в качестве примера, но не ограничения, изображена общая структура данных, являющаяся исходной структурой данных для настоящего способа преобразования 200, представляющая собой пригодную для преобразования структуру данных, содержащую синтаксические единицы. Исходная структура данных представляет собой СМД, содержащий элементы 22 пригодной для преобразования структуры данных, содержащей СЕ, представляющие собой синтаксические единицы (СЕ) и идентификационные данные СЕ, представляющие собой для каждой СЕ, в качестве примера, но не ограничения, значение 221 элемента 22 пригодной для преобразования структуры данных, содержащей СЕ и порядковый (порядковые) номер (номера) 222 ТЭ лингвистического предложения, составляющего (составляющих) СЕ, которые, в качестве примера, но не ограничения, были описаны ранее со ссылкой на фиг. 8 способа 100. У элементов 22 лингвистического предложения 11 имеются характеризующие их уникальные наименования (УН), имеющие практическое использование, в качестве примера, но не ограничения: «подлежащее», «сказуемое», «прямое дополнение» и тому подобное. В структуре данных элементы 22 именуется уникальными наименованиями и, в качестве примера, но не ограничения, могут именоваться как «УН СЕ n», где $n \geq$

1- порядковый индекс элемента уникального наименования в лингвистическом предложении, начиная с «1» для каждого нового УН СЕ, ранее не встречавшегося в лингвистическом предложении. Как было сказано ранее, СЕ могут быть двух видов: исходные синтаксические объекты (ИСО), то есть синтаксические единицы - члены предложения, выявленные в результате синтаксического анализа, и синтаксические конструкции, представляющие собой конструкцию из синтаксически связанных ИСО. Элементы 22 пригодной для преобразования структуры данных, содержащей СЕ, представляют собой множество ИСО уникальных наименований (УН ИСО) лингвистического предложения 11. ИСО 22 в пригодной для преобразования структуре данных, содержащей СЕ, представляют собой отдельные элементы, заранее подготовленные и помещенные в пригодную для преобразования структуру данных, содержащую СЕ в виде структурированного массива (списка, перечня и тому подобного) отдельных ИСО. Такие подготовительные действия могут осуществляться указанным любым известным из уровня техники способом или неизвестным из уровня техники способом, таким, как способ, указанный при описании способа 100, и, соответственно, далее не описываются.

[95] Идентификация элементов 22 пригодной для преобразования структуры данных, содержащей СЕ в ходе этапа 201, сводится к классификации элементов, из которых состоит пригодная для преобразования структура данных, содержащая СЕ, как исходных синтаксических объектов (ИСО 22). При этом отдельные ИСО 22 имеют значение 221 и порядковый (порядковые) номер (номера) 222 ТЭ лингвистического предложения, составляющего (составляющих) синтаксические единицы. Значение 221 ИСО 22 представляет собой один или несколько текстовых элементов (ТЭ 21) предложения 11, из которых состоят ИСО. Как правило, ИСО состоят из одного ТЭ 21, за исключением некоторых ИСО. Например, но не ограничиваясь, ИСО «сказуемое» может состоять из трех ТЭ 21 - «мог быть выполнен». ИСО «сказуемое» состоит из трех текстовых элементов ввиду лингвистической особенности формирования сказуемого, которое, помимо глагола «выполнен», может иметь вспомогательный глагол «быть» и модальность (модальный глагол) «мог».

[96] Идентификацию значения 221 элемента 22 (ИСО) пригодной для преобразования структуры данных, содержащей СЕ в ходе этапа 201 производят путем регистрации набора знаков ТЭ 21 предложения 11, составляющего (составляющих) элемент ИСО 22. Идентификацию порядкового (порядковых) номера (номеров) 222ТЭ, составляющего (составляющих) элемент 22 (ИСО) пригодной для преобразования структуры данных, содержащей СЕ в ходе этапа 201 производят путем регистрации порядкового (порядковых) номера (номеров) упомянутых ТЭ в соответствии с его (их) расположением в лингвистическом предложении, при условии, что первый ТЭ в лингвистическом предложении имеет порядковый номер «1», а все последующие ТЭ имеют порядковый номер, больший на единицу, чем порядковый номер предыдущего ТЭ.

[97] Пригодная для преобразования структура данных, содержащая СЕ, являющаяся исходной для настоящего способа преобразования 200 представляет собой в итоге множество элементов 22, их значений 221 и порядковых номеров 222 ТЭ, составляющих элементы 22, идентифицированных на этапе 201.

[98] Идентификацию лингвистических характеристик ТЭ, составляющих элементы 22 (ИСО) пригодной для преобразования структуры данных, содержащей СЕ и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 102, состоящего из идентификационных данных ТЭ, составляющих ИСО 22, и получении значений 2131 лингвистических характеристик 213

текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 22 (ИСО). При этом, как было описано ранее, лингвистическими признаками элементов 22 (ИСО) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых
5 состоят элементы 22 (ИСО).

[99] На фиг. 16, в качестве примера, но не ограничения, изображена общая схема выполнения этапов этапа 202 формирования пятой структуры данных СМД, являющейся первой структурой данных для рассматриваемого способа 200. Этап 202 характеризуется:
10 выполнением этапа 2021 формирования значений 2141 второй части лингвистических характеристик 214 текстовых элементов 21, составляющих элементы 22 пригодной для преобразования структуры данных, содержащей СЕ, на котором формируют значения 2141 второй части лингвистических характеристик 214 текстовых элементов 21, составляющих элементы 22 и вносят полученные сведения в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 102, формируя в
15 итоге БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2021; выполнением этапа 2022 формирования значений 2231 третьей части лингвистических характеристик 223 текстовых элементов 21, составляющих элементы 22 пригодной для преобразования структуры данных, содержащей СЕ на котором идентифицируют виды элемента 22 пригодной для преобразования структуры
20 данных, содержащей СЕ для которых необходимо изменить некоторые значения лингвистических характеристик, формируют упомянутые значения 2231 третьей части лингвистических характеристик 223 текстовых элементов 21, составляющих элементы 22 и вносят полученные сведения в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2021, формируя в итоге БДЛП текстовых
25 элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2022; выполнением этапа 2023 идентификации элементов 61 пятой структуры данных СМД, на котором идентифицируют и формируют элементы 61, пятой структуры данных СМД, а также идентификационные данные элементов 61, представляющие собой для каждого элемента 61, в качестве примера, но не ограничения, значение 611 элемента 61 пятой
30 структуры данных СМД и порядковый (порядковые) номер (номера) 612 ТЭ лингвистического предложения 11, составляющего (составляющих) элемент 61, и формируют пятую структуру данных СМД.

[100] На фиг. 17, в качестве примера, но не ограничения, изображена общая структура сформированной в рамках этапа 2021 базы данных лингвистических признаков (БДЛП),
35 являющейся БДЛП текстовых элементов 21 лингвистического предложения 11, содержащихся в элементах 22 пригодной для преобразования структуры данных, содержащей СЕ. БДЛП, сформированная в рамках этапа 2021, отличается от БДЛП, сформированной в рамках этапа 102, наличием значений 2141 второй части лингвистических характеристик 214, указывающих на синтаксическо-логическую
40 идентичность элемента 22 (ИСО). Для идентификации лингво-логических единиц (ЛЛЕ) лингвистического предложения необходимо по всем синтаксическим единицам вида ИСО (элемент 22) провести проверку на идентичность их синтаксической и логической ролей в лингвистическом предложении 11. В результате упомянутой проверки синтаксических единиц ИСО 22 на их синтаксическо-логическую идентичность должны
45 быть установлены наличие или отсутствие идентичности ИСО 22. ИСО 22 будет считаться синтаксическо-логическо идентичным, если его синтаксическая и логическая роли совпадут, и будет считаться синтаксическо-логическо неидентичным, если упомянутые роли не совпадут. Синтаксическая роль ИСО 22 – это функция слова в

предложении, которая определяется в рамках взаимосвязей слов в предложении. Синтаксическая роль ИСО 22 определяет, каким членом предложения является слово, и в каких оно синтаксических отношениях состоит с другими словами. Одна и та же форма слова может выполнять различные функции, то есть быть разными членами предложения. Логическая роль ИСО 22 – это синтаксическая роль, отождествленная с логической функцией слова в простом суждении. Вид логической функции в простом суждении связан с уникальным наименованием логического объекта (УН ЛО), являющегося элементом простого суждения, к которому относится ИСО 22. Простое суждение – это логическая сущность, представляющая собой первичную логическую конструкцию мышления, обязательно содержащую элементы простого суждения – субъект суждения (предмет окружающего мира, о котором идет речь) и предикат суждения (то, что утверждается или опровергается о предмете суждения). При этом субъект суждения представляет собой один предмет окружающего мира, а предикат суждения состоит, как правило, из действия и множества предметов, которые находятся в определенной взаимосвязи друг с другом, что в конечном итоге и выполняет функцию предиката суждения, то есть поясняет то, что утверждается или опровергается о предмете суждения. Поэтому предикат суждения может быть разделен на логические функционально-однородные части (логические объекты), в качестве примера, но не ограничения, такие как действие, объект и обстоятельства предиката суждения. Упомянутые ранее субъект суждения, а также действие, объект и обстоятельства предиката суждения могут являться уникальными наименованиями логического объекта (УН ЛО), являющегося элементом простого суждения. В соответствии с использованным словом (ИСО 22) посредством которого выражены элементы простого суждения, эти слова (ИСО 22) будут иметь определенную логическую функцию, которая в соответствии с предварительно заданными условиями может отождествляться или не отождествляться с определенной синтаксической ролью. Таким образом становится возможным провести проверку на идентичность синтаксической и логической ролей отдельных ИСО 22 лингвистического предложения 11 и преобразовать ИСО 22 в лингво-логические единицы (ЛЛЕ).

[101] Формирование второй части лингвистических характеристик 214 и их значений 2141 для текстовых элементов 21 лингвистического предложения 11, содержащихся в элементах 22 с целью идентификации синтаксическо-логической идентичности ИСО 22 на этапе 2021 производят путем проверки ИСО 22 на идентичность синтаксической и логической ролей в лингвистическом предложении 11. Проверка представляет собой сравнение значений 2131 первой части лингвистических характеристик 213 текстовых элементов 21, составляющих элементы 22, с заранее заданными значениями заранее заданных лингвистических признаков текстовых элементов 21, составляющих ИСО 22. Для проведения проверки заранее задаются условия, такие как, в качестве примера, но не ограничения: перечни пар ИСО 22, у которых синтаксическая и логическая роли могут не совпадать, а также признаки несовпадения синтаксической и логической ролей (некоторые значения лингвистических характеристик указанных ИСО 22, при наличии которых ИСО 22 идентифицируется как ИСО 22 с несовпадением синтаксических и логических ролей). ИСО 22, которые успешно проходят упомянутую проверку идентифицируют как ИСО 22 с совпадением синтаксических и логических ролей (СЛ-идентичные ИСО 22, то есть синтаксическо-логическо идентичные ИСО). ИСО 22, которые не проходят упомянутую проверку по причине того, что их упомянутые роли не совпадают, идентифицируют как ИСО 22 с несовпадением синтаксических и логических ролей (СЛ-неидентичные ИСО 22, то есть синтаксическо-логическо

неидентичные ИСО). Например, но не ограничиваясь, ИСО в форме глагола имеет синтаксическую роль - «сказуемое», а ИСО в форме причастия – синтаксическую роль «согласованное определение». Несмотря на то, что синтаксические роли глагола и причастия не совпадают, с логической точки зрения и тот и другой ИСО являются синтаксически главными словами в одном и том же элементе простого суждения «действие предиката суждения», то есть имеют одну и ту же логическую роль. В итоге ИСО в форме глагола проходит проверку идентичности синтаксической и логической ролей, так как «глагол» и есть то, что выражает действие. При этом ИСО в форме согласованного определения не проходит проверку, так как «определение» – это признак, а не действие. По этой причине, а также при дополнительном условии, что причастие является распространенным (имеет зависимые слова – значимые части речи) считается, что ИСО «согласованное определение» в форме причастия не проходят упомянутую проверку и являются СЛ-неидентичным ИСО 22. Упомянутые условия, используемые для проведения проверки идентичности синтаксической и логической ролей ИСО 22 предварительно записывают в четвертую пользовательскую базу данных (четвертую ПБД).

[102] Сформированные в результате выполнения этапа 2021 значения 2141 (синтаксическо-логическая идентичность ИСО 22) второй части лингвистических характеристик 214 элементов 21, составляющих ИСО 22 пригодной для преобразования структуры данных, содержащей СЕ вносят в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 102, формируя таким образом БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2021. При этом обе части (213 и 214) лингвистических характеристик текстовых элементов 21 лингвистического предложения 11 и их значения (2131 и 2141) формируют уникальные лингвистические признаки элементов 22.

[103] На фиг. 18, в качестве примера, но не ограничения, изображена общая структура сформированной в рамках этапа 2022 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 лингвистического предложения 11, содержащихся в элементах 22 пригодной для преобразования структуры данных, содержащей СЕ. БДЛП, формируемая в рамках этапа 2022, отличается от БДЛП, формируемой в рамках этапа 2021, наличием значений 2231 третьей части лингвистических характеристик 223, указывающих вид элемента 22 (ИСО) и актуальные значения некоторых лингвистических характеристик 213 текстовых элементов, составляющих элемент 22 (ИСО). Для идентификации лингво-логических единиц (ЛЛЕ) лингвистического 11 необходимо по всем синтаксическим единицам ИСО 22, по которым была проведена проверка на идентичность их синтаксической и логической ролей дополнительно провести идентификацию видов ИСО 22 по признаку внесения изменений в значения их лингвистических характеристик 213 в соответствии с их значением 2141 синтаксическо-логической идентичности 214. В результате проведенной на этапе 2021 проверки на синтаксическо-логическую идентичность ИСО 22 было установлено, что все ИСО 22 разделяются по этому признаку на СЛ-идентичные и СЛ-неидентичные ИСО 22. При этом СЛ-неидентичные ИСО 22 могут быть изменяемыми и неизменяемыми, то есть такие СЛ-неидентичные ИСО 22, у которых в соответствии с заранее заданным условием должны быть изменены значения 2131 некоторых лингвистических характеристик 213 или не изменены. Кроме этого, в некоторых случаях, в заранее установленных условиях возможно наличие требования о создании дубликата СЛ-неидентичного ИСО 22, у которого не изменяются лингвистические характеристики. Упомянутые условия и требования устанавливаются в четвертой ПБД.

[104] Формирование третьей части лингвистических характеристик 223 и их значений 2231 для текстовых элементов 21 лингвистического предложения 11, содержащихся в элементах 22 производят на этапе 2022 с целью идентификации различных видов ИСО 22 по признаку синтаксическо-логической идентичности (СЛ-идентичности) и формирования дубликата СЛ-неидентичного ИСО 22 путем комплексного анализа условий, установленных в четвертой ПБД, и на основании итогов проверки ИСО 22 на СЛ-идентичность. В результате упомянутого анализа для каждого ИСО 22 устанавливаются определенный вид – первый, второй или третий вид. Если в условиях, установленных в четвертой ПБД, имеется требование о создании дубликата СЛ-неидентичных ИСО 22, то такие ИСО 22 относят к четвертому виду ИСО 22. В соответствии с установленным видом ИСО 22 формируют значения 2231 третьей части лингвистических характеристик 223 текстовых элементов 21 лингвистического предложения 11, содержащихся в элементах 22. У первого вида ИСО 22 (СЛ-идентичные) остаются все их исходные значения 221 и номера текстовых элементов 222, а также их лингвистические характеристики 213 и 214 со значениями 2131 и 2141. В связи с этим значения 2231 лингвистических характеристик 213 для такого ИСО 22 отсутствуют. У второго вида ИСО 22 (СЛ-неидентичные изменяемые) остается неизменным исходное их значение 221 и номера текстовых элементов 222, а также вторая часть их лингвистических характеристик 214 со значениями 2141, но при этом значения 2131 некоторых лингвистических характеристик 213 первой части изменяют. В связи с этим значения 2231 лингвистической характеристики 223 вместо изменяемых значений 2131 некоторых лингвистических характеристик 213 такого ИСО 22 указывается таким, как это установлено в четвертой ПБД. У третьего вида ИСО 22 (СЛ-неидентичные неизменяемые) остаются все их исходные значения 221 и номера текстовых элементов 222, а также их лингвистические характеристики 213 и 214 со значениями 2111 и 2141. В связи с этим значения 2231 лингвистических характеристик 213 для такого ИСО 22 отсутствуют. У четвертого вида ИСО 22 (дубликат СЛ-неидентичных неизменяемых ИСО) остается неизменным исходное их значение 221 и номера текстовых элементов 222, а также вторая часть их лингвистических характеристик 214 со значениями 2141, но при этом значения 2131 некоторых лингвистических характеристик 213 первой части изменяют. В связи с этим значения 2231 лингвистической характеристики 223 вместо изменяемых значений 2131 некоторых лингвистических характеристик 213 такого ИСО 22 указывают таким, как это установлено в четвертой ПБД.

[105] Сформированные по итогам этапа 2022 значения 2231 (измененные значения лингвистических характеристик ИСО 22) третьей части лингвистических характеристик 223 элементов 21, составляющих ИСО 22 пригодной для преобразования структуры данных, содержащей СЕ, вносят в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2021, формируя таким образом БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2022. При этом все три части (213, 214 и 223) лингвистических характеристик текстовых элементов 21 лингвистического предложения 11 и их значения (2131, 2141 и 2231) формируют уникальные лингвистические признаки элемента 22.

[106] На фиг. 19, в качестве примера, но не ограничения, изображена общая структура сформированной пятой структуры данных СМД, являющейся первой для настоящего способа преобразования 200. Пятая структура данных СМД представляет собой СМД, содержащий элементы 61 пятой структуры данных СМД, которые представляют собой лингво-логические единицы (ЛЛЕ 61) лингвистического предложения 11 и идентификационные данные ЛЛЕ 61, представляющие собой для каждой ЛЛЕ, в качестве

примера, но не ограничения, значение 611 элемента 61 пятой структуры данных и
 порядковый (порядковые) номер (номера) 612 ТЭ, составляющего (составляющих)
 ЛЛЕ 61. У ЛЛЕ 61 лингвистического предложения 11 имеются характеризующие их
 уникальные наименования (УН ЛЛЕ), имеющие практическое использование в качестве
 5 примера, но не ограничения: «подлежащее», «сказуемое», «прямое дополнение» и тому
 подобное. В структуре данных элементы 61 именуются уникальными наименованиями,
 и, в качестве примера, но не ограничения, могут именоваться как «УН ЛЛЕ n», где n
 ≥ 1 - порядковый индекс элемента уникального наименования в лингвистическом
 предложении, начиная с «1» для каждого нового УН ЛЛЕ, ранее не встречавшегося в
 10 лингвистическом предложении. Лингво-логическая единица (ЛЛЕ 61) – это
 синтаксическая единица вида ИСО, по которой проведена проверка на идентичность
 ее синтаксической и логической ролей в лингвистическом предложении. При этом может
 быть два вида ЛЛЕ 61: первый вид – это оригинальный ИСО 22, прошедший упомянутую
 проверку, который содержался в пригодной для преобразования структуре данных,
 15 содержащей СЕ, а второй вид – это дубликат ИСО 22, которого не было в упомянутой
 структуре данных. Первому виду ЛЛЕ 61 соответствуют все ИСО 22 пригодной для
 преобразования структуры данных, содержащей СЕ. Второму виду ЛЛЕ 61
 соответствуют те ИСО 22 пригодной для преобразования структуры данных,
 содержащей СЕ, которые отнесены к четвертому виду (дубликат СЛ-неидентичных
 20 неизменяемых ИСО), указанному в значении 2231 третьей части лингвистических
 характеристик 223.

[107] Идентификацию и формирование элементов 61 пятой структуры данных СМД
 в ходе этапа 2023 производят на основе результатов этапа 2022 по идентификации видов
 элементов 22 пригодной для преобразования структуры данных, содержащей СЕ.
 25 Идентификацию элементов 61 (ЛЛЕ) первого вида на этапе 2023 производят путем их
 отождествления с элементами 22 (ИСО) первого, второго и третьего видов. ИСО 22,
 идентифицированные на этапе 2022 как четвертый вид, формируют новые элементы -
 элементы 61 второго вида пятой структуры данных СМД. Формирование элементов
 61 (ЛЛЕ) второго вида на этапе 2023 производится путем их отождествления с
 30 элементами 22 (ИСО) четвертого вида.

[108] Идентификацию значения и порядкового (порядковых) номера (номеров) ТЭ
 элемента 61 первого вида пятой структуры данных СМД, отождествленного с первым
 видом ИСО 22, производят в рамках этапа 2023 путем отождествления значения 611
 элемента 61 со значением 221 элемента 22 первого вида, а порядкового (порядковых)
 35 номера (номеров) 612 ТЭ, составляющего (составляющих) элемент 61, путем
 отождествления с порядковым (порядковыми) номером (номерами) 222 ТЭ,
 составляющего (составляющих) элемент 22 первого вида. Идентификацию значения и
 порядкового (порядковых) номера (номеров) ТЭ элемента 61 первого вида пятой
 структуры данных СМД, отождествленного со вторым видом ИСО 22, производят в
 40 рамках этапа 2023 путем отождествления значения 611 элемента 61 со значением 221
 элемента 22 второго вида, а порядкового (порядковых) номера (номеров) 612ТЭ,
 составляющих элемент 61, с порядковым (порядковыми) номерами 222 ТЭ,
 составляющего (составляющих) элемент 22 второго вида. Идентификацию значения и
 порядкового (порядковых) номера (номеров) ТЭ элемента 61 первого вида пятой
 45 структуры данных СМД, отождествленного с третьим видом ИСО 22, производят в
 рамках этапа 2023 путем отождествления значения 611 элемента 61 со значением 221
 элемента 22 третьего вида, а порядкового (порядковых) номера (номеров) 612 текстовых
 элементов 21, составляющих элемент 61, с порядковым (порядковыми) номером

(номера) 222 ТЭ, составляющего (составляющих) элемент 22 третьего вида.

Идентификацию значения и порядкового (порядковых) номера (номеров) ТЭ элемента 61 второго вида пятой структуры данных СМД, отождествленного с четвертым видом ИСО 22, производят в рамках этапа 2023 путем отождествления значения 611 элемента 61 со значением 221 элемента 22 третьего вида, а порядкового (порядковых) номера (номеров) 612ТЭ, составляющего (составляющих) элемент 61, с порядковым (порядковыми) номером (номера) 222 ТЭ, составляющего (составляющих) элемент 22 третьего вида.

[109] Формирование пятой структуры данных СМД в ходе этапа 2023 производят путем объединения в одной структуре данных элементов 61 первого и второго видов пятой структуры данных СМД и их идентификационных данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются.

[110] Идентификацию лингвистических характеристик ТЭ, составляющих первые или вторые элементы 61 (ЛЛЕ) пятой структуры данных СМД, и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 2022, состоящего из идентификационных данных ТЭ, составляющих первый или второй элемент 61 (ЛЛЕ), и получении значений (2131, 2141, 2231) всех частей лингвистических характеристик (213, 214, 223) текстовых элементов 21 лингвистического предложения 11, из которых состоит первый или второй элемент 61 (ЛЛЕ). При этом, как было описано ранее, лингвистическими признаками первых и вторых элементов 61 (ЛЛЕ) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 61 (ЛЛЕ).

[111] На фиг. 20 в качестве примера, но не ограничения, изображена общая схема выполнения этапа 203 формирования шестой структуры данных СМД, являющейся второй структурой данных для рассматриваемого способа 200. Этап 203 характеризуется выполнением этапа 2031 формирования значений 2151 четвертой части лингвистических характеристик 215 текстовых элементов 21, составляющих элементы 61 пятой структуры данных СМД, на котором формируют значения 2151 четвертой части лингвистических характеристик 215 текстовых элементов 21, составляющих элементы 61, и вносят полученные сведения в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2022, формируя в итоге БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2031; выполнением этапа 2032 формирования элементов шестой структуры данных СМД, на котором формируют элементы 71 шестой структуры данных СМД, а также идентификационные данные элементов 71, представляющие собой для каждого элемента 71, в качестве примера, но не ограничения, значение 711 элемента 71 шестой структуры данных СМД и порядковый (порядковые) номер (номера) 712 ТЭ лингвистического предложения 11, составляющего (составляющих) элемент 71; и формируют шестую структуру данных СМД.

[112] На фиг. 21, в качестве примера, но не ограничения, изображена общая структура сформированной в рамках этапа 2031 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 лингвистического предложения 11, содержащихся в элементах 61 пятой структуры данных СМД. БДЛП, формируемая в рамках этапа 2031 отличается от БДЛП, формируемой в рамках этапа 2022, наличием значений 2151 четвертой части лингвистических характеристик 215, указывающих на уровень синтаксической значимости элемента 61 (ЛЛЕ). Уровень синтаксической

значимости элемента 61 (ЛЛЕ) характеризует местоположение ЛЛЕ в синтаксическом дереве лингвистического предложения 11. Синтаксическим деревом является схема синтаксических связей предложения 11, в которой между отдельными словами (ЛЛЕ) установлена подчинительная синтаксическая связь. В зависимости от состава и строения предложения синтаксическое дерево может содержать неопределенное количество синтаксических уровней и частей предложения, в которых содержатся основные логические элементы (субъект и предикат суждения). Синтаксическое дерево, используемое в настоящем способе, содержит элементы (ЛЛЕ) как минимум трех уровней синтаксической значимости (УСЗ). К первому УСЗ относят ЛЛЕ, находящиеся на синтаксической вершине синтаксического дерева. Ко второму УСЗ относят ЛЛЕ, являющиеся прямыми синтаксическими потомками синтаксических вершин. К третьему УСЗ относят все остальные ЛЛЕ лингвистического предложения. Синтаксическая вершина – это синтаксическая единица (ЛЛЕ), имеющая синтаксическую роль «сказуемое». Особенностью ЛЛЕ «сказуемое» является то, что оно может состоять из трех слов. Прямой синтаксический потомок синтаксической вершины – это синтаксическая единица (ЛЛЕ), имеющая прямую подчинительную связь со словом группы ЛЛЕ первого уровня. Установление синтаксического уровня ЛЛЕ является синтаксическим структурированием ЛЛЕ.

[113] Формирование четвертой части лингвистических характеристик 215 и их значений 2151 для текстовых элементов 21 лингвистического предложения 11, содержащихся в элементах 61, с целью установления уровня синтаксической значимости элементов 61 пятой структуры данных СМД в ходе этапа 2031 производят на основании комплексного анализа значений лингвистических характеристик всех ЛЛЕ 61 лингвистического предложения, на основе которого формируют синтаксическое дерево лингвистического предложения. В результате выяснения положения каждого ЛЛЕ в синтаксическом дереве предложения определяют уровень синтаксической значимости каждого ЛЛЕ 61.

[114] Сформированные по итогам этапа 2031 значения 2151 (уровень синтаксической значимости ЛЛЕ 61) четвертой части лингвистических характеристик 215 элементов 21, составляющих элементы 61 пятой структуры данных СМД, вносят в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2022, формируя таким образом БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2031. При этом все части (211, 214, 223, 215) лингвистических характеристик текстовых элементов 21 лингвистического предложения 11 и их значения (2111, 2141, 2231, 2151) формируют уникальные лингвистические признаки элементов 61.

[115] На фиг. 22, в качестве примера, но не ограничения, изображена общая структура сформированной шестой структуры данных СМД, являющаяся второй структурой данных для настоящего способа преобразования 200. Шестая структура данных СМД представляет собой СМД, содержащий элементы 71, которые представляют собой структурированные группы лингво-логических единиц (группы ЛЛЕ) лингвистического предложения 11 и идентификационные данные групп ЛЛЕ, представляющие собой для каждой группы ЛЛЕ, в качестве примера, но не ограничения, значение 711 элемента 71 шестой структуры данных и порядковый (порядковые) номер (номера) 712 ТЭ, составляющего (составляющих) группу ЛЛЕ. У групп ЛЛЕ 71 лингвистического предложения 11 отсутствуют характеризующие их уникальные наименования (УН групп ЛЛЕ), имеющие практическое использование. В структуре данных элементы 71, в качестве примера, но не ограничения, могут именоваться как «группа ЛЛЕ 1», «группа

ЛЛЕ 2», «группа ЛЛЕ 3», «группа ЛЛЕ n», где $n \geq 1$ - порядковый номер элемента в лингвистическом предложении. Группа ЛЛЕ лингвистического предложения 11 – это синтаксические единицы вида «синтаксическая конструкция», которые формируются из лингво-логических единиц (элементов 61 пятой структуры данных СМД) разных уровней синтаксической значимости ЛЛЕ. К первой группе ЛЛЕ (первые элементы 71 5 шестой структуры данных СМД) относят как главные ЛЛЕ первой группы ЛЛЕ, являющиеся упомянутыми синтаксическими вершинами (ЛЛЕ первого УСЗ), так и прочие ЛЛЕ первой группы ЛЛЕ (группы 1 ЛЛЕ), являющиеся последовательно подчиненными ЛЛЕ, первая из которых имеет прямую подчинительную связь с ЛЛЕ 10 первого УСЗ, вплоть до ЛЛЕ, которая является главной ЛЛЕ второй группы ЛЛЕ. Ко второй группе ЛЛЕ (вторые элементы 71 шестой структуры данных СМД) относят как главные ЛЛЕ второй группы ЛЛЕ, являющиеся упомянутыми ЛЛЕ второго УСЗ при условии, что, такой ЛЛЕ имеет синтаксическую роль подлежащего или любого 15 второстепенного члена предложения, кроме определения (исключая определения в форме распространенного причастия или деепричастия), так и прочие ЛЛЕ второй группы ЛЛЕ, являющиеся последовательно подчиненными ЛЛЕ, первая из которых является ЛЛЕ третьего УСЗ, которые имеют прямую подчинительную связь с упомянутым главным ЛЛЕ второй группы ЛЛЕ, вплоть до последних ЛЛЕ в цепочке ЛЛЕ с непрерывной подчинительной связью.

[116] Идентификацию элементов 71 (первых и вторых элементов 71) шестой структуры данных СМ в ходе этапа 2032 производят посредством идентификации главных ЛЛЕ 20 первых и вторых групп ЛЛЕ, а также идентификации прочих ЛЛЕ первых и вторых групп ЛЛЕ. Для идентификации главных ЛЛЕ первых групп ЛЛЕ необходимо выявить все ЛЛЕ «сказуемое» (в том числе, все части каждого ЛЛЕ «сказуемое») анализируемого 25 предложения. Для идентификации главных ЛЛЕ вторых групп ЛЛЕ необходимо выявить все ЛЛЕ, которые имеют прямую синтаксическую связь с любой частью ЛЛЕ «сказуемое» и которые при этом имеют синтаксическую роль подлежащего или любого второстепенного члена предложения, кроме определения (исключая определения в 30 форме распространенного причастия или деепричастия). Для идентификации прочих ЛЛЕ первой или второй группы ЛЛЕ необходимо идентифицировать все ЛЛЕ, которые имеют непрерывную синтаксическую подчинительную связь, начиная от главного ЛЛЕ первой или второй группы ЛЛЕ и вплоть до другого идентифицированного главного ЛЛЕ первой или второй группы ЛЛЕ, или до последнего ЛЛЕ в цепочке непрерывной синтаксической подчинительной связи от главного ЛЛЕ первой или второй группы 35 ЛЛЕ. Формирование первых элементов 71 шестой структуры данных СМД производят путем объединения идентифицированных главной ЛЛЕ первой группы ЛЛЕ и прочих ЛЛЕ для этой главной ЛЛЕ. Формирование вторых элементов 71 шестой структуры данных СМ производят путем объединения идентифицированных главной ЛЛЕ второй группы ЛЛЕ и прочих ЛЛЕ для этой главной ЛЛЕ.

[117] Идентификацию значения и порядкового (порядковых) номера (номеров) ТЭ 40 первого элемента 71 (первой группы ЛЛЕ) шестой структуры данных СМД в рамках этапа 2032 производят следующим образом: значение (значения) 211 ТЭ 21 лингвистического предложения 11, составляющего (составляющих) элемент (элементы) 61, из которого (которых) состоит первый элемент 71, идентифицируют как значение 45 711 первого элемента 71 (первая группа ЛЛЕ), а порядковый (порядковые) номер (номера) 212 упомянутого (упомянутых) ТЭ 21 лингвистического предложения 11, составляющего (составляющих) элемент (элементы) 61, из которого (которых) состоит первый элемент 71, идентифицируют как порядковый (порядковые) номер (номера)

712 ТЭ, составляющего(составляющих) первый элемент 71 (первая группа ЛЛЕ). Идентификацию значения и порядкового (порядковых) номера (номеров) ТЭ второго элемента 71 (второй группы ЛЛЕ) шестой структуры данных СМД в рамках этапа 2032 производят следующим образом: значение (значения) 211 ТЭ 21 лингвистического предложения 11, составляющего (составляющих) элемент (элементы) 61, из которого (которых) состоит второй элемент 71, идентифицируют как значение 711 второго элемента 71 (вторая группа ЛЛЕ), а порядковый (порядковые) номер (номера) 212 упомянутого (упомянутых) ТЭ 21 лингвистического предложения 11, составляющего (составляющих) элемент 61, из которого (которых) состоит второй элемент 71, идентифицируют как порядковый (порядковые) номер (номера) 712 ТЭ, составляющего (составляющих) второй элемент 71 (вторая группа ЛЛЕ).

[118] Формирование шестой структуры данных СМД в ходе этапа 2032 производят путем объединения в одной структуре данных первых и вторых элементов 71 шестой структуры данных СМД, а также их идентификационных данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются.

[119] Идентификацию лингвистических характеристик ТЭ, составляющих первые или вторые элементы 71 (первую или вторую группу ЛЛЕ) шестой структуры данных СМД и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 2031, состоящего из идентификационных данных ТЭ, составляющих первый или второй элемент 71 (первая или вторая группа ЛЛЕ), и получении значений (2131, 2141, 2231, 2151) всех частей лингвистических характеристик (213, 214, 223, 215) текстовых элементов 21 лингвистического предложения 11, из которых состоит первый или второй элемент 71 (первая или вторая группа ЛЛЕ). При этом, как было описано ранее, лингвистическими признаками первых и вторых элементов 71 (первой и второй группы ЛЛЕ) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 71 (группы ЛЛЕ).

[120] На фиг. 23, в качестве примера, но не ограничения, изображена общая схема выполнения этапов этапа 204 формирования седьмой структуры данных СМД, являющейся третьей структурой данных для рассматриваемого способа 200. Этап 204 характеризуется: выполнением этапа 2041 идентификации видов элементов 71 шестой структуры данных СМД, указывающих на наличие в элементе 71 осложнений лингвистического предложения, на котором идентифицируют первые и вторые виды элементов 71 шестой структуры данных СМД, а также идентификации непреобразованных элементов 81 седьмой структуры данных СМД и их идентификационных данных, представляющих собой для каждого непреобразованного элемента 81, в качестве примера, но не ограничения, значение 811 непреобразованного элемента 81 седьмой структуры данных СМД и порядковый (порядковые) номер (номера) 812 ТЭ лингвистического предложения 11, составляющего (составляющих) непреобразованный элемент 81 (значение и порядковые номера ТЭ непреобразованного элемента 81); выполнением этапа 2042 формирования преобразованных элементов 81 седьмой структуры данных СМД, на котором формируют преобразованные элементы седьмой структуры данных СМД из элементов 71 второго вида, а также идентификационные данные преобразованных элементов 81, представляющие собой для каждого преобразованного элемента 81, в качестве примера, но не ограничения, значение 811 преобразованного элемента 81 седьмой структуры данных СМД и порядковый (порядковые) номер (номера) 812 ТЭ лингвистического предложения 11,

составляющего (составляющих) преобразованный элемент 81 (значение и порядковые номера ТЭ преобразованного элемента 81); выполнением этапа 2043 формирования седьмой структуры данных СМД, на котором формируют седьмую структуру данных СМД из преобразованных элементов 81 и непреобразованных элементов 81 седьмой

5 структуры данных СМД.

[121] На фиг. 24, в качестве примера, но не ограничения, изображена общая структура сформированной седьмой структуры данных СМД, являющейся третьей структурой данных для настоящего способа преобразования 200. Седьмая структура данных СМД представляет собой СМД, содержащий непреобразованные элементы 81 и

10 преобразованные элементы 81, которые представляют собой основные лингво-логические объекты (ОЛЛО) лингвистического предложения 11. У ОЛЛО лингвистического предложения 11 имеются характеризующие их уникальные наименования (УН ОЛЛО), имеющие практическое использование, в качестве примера, но не ограничения: «субъект суждения», «предикат суждения» и тому подобное. В

15 структуре данных элементы 81 именуется уникальными наименованиями, и, в качестве примера, но не ограничения, могут именоваться как «УН ОЛЛО n», где $n \geq 1$ -порядковый индекс элемента уникального наименования в лингвистическом предложении, начиная с «1» для каждого нового УН ОЛЛО, ранее не встречавшегося в лингвистическом предложении. В дальнейшем, при упоминании ОЛЛО как элемента

20 седьмой структуры данных имеется в виду УН ОЛЛО. Основные лингво-логические объекты предложения – это пограничная сущность с признаками как лингвистического, так и логического объектов. Сущность – это объект в компьютерной системе, содержащий набор критических правил. С лингвистической точки зрения, основной лингво-логический объект – это синтаксическая единица вида «синтаксическая

25 конструкция», которая может быть сформирована из структурированной группы ЛЛЕ путем гетерогенизации группы ЛЛЕ. Гетерогенизация группы ЛЛЕ – это устранение однородностей в синтаксической конструкции путем идентификации в ней синтаксической сочинительной связи и преобразования синтаксической конструкции с однородностями в несколько синтаксических конструкций без однородностей.

30 Сочинительная связь – это связь между синтаксически равноправными единицами предложения (в данном случае между ЛЛЕ), у которых имеется одно и то же синтаксически главное слово (главная ЛЛЕ) и одна и та же синтаксическая роль (синтаксическая роль ЛЛЕ). ЛЛЕ, между которыми идентифицирована синтаксическая сочинительная связь, считаются «однородными членами» (ЛЛЕ-ОЧ). С логической

35 точки зрения ОЛЛО – это логические сущности (объекты), являющиеся элементами простого суждения, с помощью которых формируют и передают мысль о том, что нечто (предикат суждения) утверждается или опровергается о предмете суждения (субъекте суждения). При этом нечто (то, что утверждается или опровергается о предмете суждения) может быть дополнительно логически структурирован. Например, но не

40 ограничиваясь, предикат суждения может быть разделен на три логически отделимые логические сущности (объекта) - действие предиката суждения, объект предиката суждения и обстоятельства предиката суждения. Упомянутые выше для примера элементы простого суждения (субъект суждения, действие предиката суждения, объект предиката суждения и обстоятельства предиката суждения) могут быть отождествлены

45 с отдельными УН ОЛЛО, как с отдельными синтаксическими конструкциями лингвистического предложения 11. С семантической точки зрения ОЛЛО представляют собой обобщенный мыслимый (семантический) образ элемента простого суждения.

[122] Идентификацию элементов 71 первого или второго вида шестой структуры

данных СМД в рамках этапа 2041 производят путем анализа значений лингвистических характеристик текстовых элементов шестой структуры данных СМД, составляющих элемент 71 (группу ЛЛЕ), с целью идентификации синтаксических сочинительных связей в группе ЛЛЕ 71. Синтаксическая сочинительная связь между текстовыми элементами (словами, ЛЛЕ) в группе ЛЛЕ 71 идентифицируют в случае наличия у таких ЛЛЕ (слов) одного и того же синтаксического родителя (синтаксически главного слова, у которого с такими ЛЛЕ (словами) имеется прямая синтаксическая подчинительная связь).

Дополнительными условиями может быть наличие у таких ЛЛЕ (слов) одной и той же синтаксической роли, а также наличие признака связи между такими ЛЛЕ (словами), а именно запятой или сочинительного союза. Сочинительные союзы делятся по значению на соединительные («и», «да» в значении «и», «ни... ни», «также», «тоже», «как... так и»), разделительные («или», «или... или», «либо», «либо... либо», «то... то», «то ли... то ли», «не то... не то») и противительные («а», «но», «однако», «да» в значении «но»), в том числе сопоставительные («не только... но и», «не только... а и», «не только не... но», «не столько... сколько»). Предварительно могут быть заданы иные собственные критерии идентификации сочинительной связи или, например, не ограничиваясь, может быть задано игнорирование какого-либо дополнительного условия, например, не ограничиваясь, условия о наличии признака связи. При идентификации синтаксической сочинительной связи в группе ЛЛЕ такие группы ЛЛЕ с однородными членами идентифицируются как второй вид элемента 71 шестой структуры данных СМД. Все иные группы ЛЛЕ, в которых не идентифицируется синтаксическая сочинительная связь, идентифицируются как первый вид элемента 71 шестой структуры данных СМД. При этом идентифицированный первый вид элемента 71 шестой структуры данных СМД отождествляют с преобразованным элементом 81 седьмой структуры данных СМД. Формирование преобразованных элементов седьмой структуры данных СМД (преобразованных элементов 81, полученных из элементов 71 второго вида) на этапе 2042 производят на основании идентифицированной на этапе 2041 синтаксической сочинительной связи и однородных членов путем преобразования идентифицированных элементов 71 второго вида шестой структуры данных СМД. Смысл преобразования элемента 71 второго вида на этапе 2042 состоит в гетерогенизации группы ЛЛЕ, то есть в устранении синтаксических однородностей в группе ЛЛЕ путем формирования из элемента 71 второго вида множества новых групп ЛЛЕ, идентичных по однородностям элементу 71 первого вида. При этом каждая новая группа ЛЛЕ упомянутого множества может содержать только один из упомянутых однородных членов (если элемент 71 второго вида содержит один ряд однородных членов) или только одно уникальное сочетание упомянутых однородных членов разных рядов однородных членов (если элемент 71 второго вида содержит несколько рядов однородных членов). Множество групп ЛЛЕ, сформированных из элемента 71 второго вида, являются гетерогенными ОЛЮ (преобразованными элементами 81), то есть сформированными из одной группы ЛЛЕ 71 второго вида путем ее гетерогенного преобразования (гетерогенизации). Процесс гетерогенизации элементов 71 второго вида может быть проведен, в качестве примера, но не ограничения, следующим образом: на первом этапе выявляют ряды однородных членов в анализируемом элементе 71 второго вида; в ходе выявления рядов однородных членов выявляются сами ряды однородных членов и порядковые номера однородных членов в ряду, а также признаки связи каждого однородного члена (ЛЛЕ-ОЧ); на втором этапе выявляют цепочки однородных членов для каждого выявленного ЛЛЕ-ОЧ элемента 71 второго вида; при этом цепочки однородных членов (цепочки ОЧ) - это ЛЛЕ-ОЧ и зависимая цепочка ОЧ (цепочка синтаксически связанных ЛЛЕ, начиная

от первого синтаксического потомка ЛЛЕ-ОЧ (синтаксически зависимого слова) и продолжая по цепочке синтаксической зависимости (по синтаксической подчинительной связи) до другого ЛЛЕ-ОЧ или до последней ЛЛЕ в элементе 71 второго вида; идентифицированные цепочки ОЧ классифицируются по факту наличия синтаксического потомка ЛЛЕ в зависимой цепочке ОЧ, который не входит в зависимую цепочку ОЧ; те цепочки ОЧ, которые имеют зависимые цепочки ОЧ без синтаксического потомка за рамками зависимой цепочки ОЧ считаются крайними цепочками однородных членов (крайними цепочками ОЧ), а те цепочки ОЧ, которые имеют зависимые цепочки ОЧ с синтаксическим потомком за рамками зависимой цепочки ОЧ считаются внутренними цепочками однородных членов (внутренними цепочками ОЧ); на третьем этапе выявляют ЛЛЕ элемента 71 второго вида, не входящие в цепочки однородных членов (ЛЛЕ вне цепочек ОЧ); на четвертом этапе формируют сменные зависимые цепочки (сменные ЗЦ) элемента 71 второго вида; при этом сменная ЗЦ - это основа преобразованных элементов 81, состоящая из одной крайней цепочки ОЧ и внутренних цепочек ОЧ, если такие внутренние цепочки ОЧ по цепочке непрерывной синтаксической подчинительной связи имеются между крайней цепочкой ОЧ и ЛЛЕ вне цепочек ОЧ; на пятом этапе формируют гетерогенные основные лингво-логические объекты (ОЛЛО); при этом гетерогенный ОЛЛО формируется из одной сменной ЗЦ и ЛЛЕ вне цепочки ОЧ (если такие ЛЛЕ есть); количество сменных ЗЦ элемента 71 второго вида определяет количество гетерогенных ОЛЛО, сформированных от одного элемента 71 второго вида.

[123] Идентификация видов элементов 71 шестой структуры данных СМД в рамках этапа 2041 не изменяет значение 711 элемента 71 и не изменяет порядковые номера текстовых элементов 21 лингвистического предложения 11, составляющих элемент 71 (группу ЛЛЕ). Идентификацию значения и порядкового (порядковых) номера (номеров) ТЭ непреобразованного элемента 81 (ОЛЛО) седьмой структуры данных СМД в рамках этапа 2041 производят следующим образом: значение (значения) 211 ТЭ 21 лингвистического предложения 11, составляющего (составляющих) элемент 71 первого вида шестой структуры данных СМД идентифицируют как значение 811 непреобразованного элемента 81 (ОЛЛО), а порядковый (порядковые) номер (номера) 712 ТЭ 21 лингвистического предложения 11, составляющего (составляющих) элемент 71 из которого состоит непреобразованный элемент 81 идентифицируют как порядковый (порядковые) номер (номера) 812 ТЭ, составляющего (составляющих) непреобразованный элемент 81 (ОЛЛО). Идентификацию значения и порядкового (порядковых) номера (номеров) ТЭ преобразованного элемента 81 (гетерогенного ОЛЛО) седьмой структуры данных СМД в рамках этапа 2042 производят следующим образом: значение (значения) 211 ТЭ 21 лингвистического предложения 11, составляющего (составляющих) элемент 71 второго вида шестой структуры данных СМД, за исключением значения (значений) 211 ТЭ 21 лингвистического предложения 11, составляющего (составляющих) отдельную (отдельные) часть (части) элемента 71 второго вида (ЛЛЕ-ОЧ), которая (которые) была (были) удалена (удалены) при гетерогенизации, и признаков связи удаленных ЛЛЕ-ОЧ (синтаксически подчиненные удаленным ЛЛЕ-ОЧ сочинительные союзы или знаки препинания) идентифицируют как значение 811 преобразованного элемента 81 (ОЛЛО), а порядковый (порядковые) номер (номера) 212 ТЭ 21 лингвистического предложения 11, составляющего (составляющих) элемент 71 второго вида шестой структуры данных СМД, за исключением порядкового (порядковых) номера (номеров) 212 ТЭ 21 лингвистического предложения 11, составляющего (составляющих) отдельную (отдельные) часть (части)

элемента 71 второго вида (ЛЛЕ-ОЧ), которая (которые) была (были) удалена (удалены) при гетерогенизации, и признаков связи удаленных ЛЛЕ-ОЧ (синтаксически подчиненные удаленным ЛЛЕ-ОЧ сочинительные союзы или знаки препинания) идентифицируют как порядковый (порядковые) номер (номера) 812 ТЭ, составляющего (составляющих) преобразованный элемент 81 (гетерогенный ОЛЛО).

[124] Формирование седьмой структуры данных СМД в рамках этапа 2043 производят путем объединения в одной структуре данных идентифицированных на этапе 2041 и сформированных на этапе 2042 элементов 81 седьмой структуры данных СМД, а также их идентификационных данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются. Первым из объединяемых элементов седьмой структуры данных СМД является элемент, идентифицированный на этапе 2041 как элемент 71 первого вида, являющийся непреобразованным элементом 81, то есть ОЛЛО 81, совпадающий с группой ЛЛЕ 71 без синтаксической сочинительной связи. Вторым из объединяемых элементов седьмой структуры данных СМД являются преобразованные элементы 81, сформированные на этапе 2042 как множество преобразованных элементов 81, являющихся гетерогенными ОЛЛО, сформированными из элемента 71 второго вида (групп ЛЛЕ с синтаксической сочинительной связью) путем гетерогенизации групп ЛЛЕ второго вида.

[125] Идентификацию лингвистических характеристик ТЭ, составляющих первые или вторые элементы 81 (ОЛЛО) седьмой структуры данных СМД и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 2031, состоящего из идентификационных данных ТЭ, составляющих первый или второй элемент 81 (ОЛЛО), и получения значений (2131, 2141, 2231, 2151) всех частей лингвистических характеристик (213, 214, 223, 215) текстовых элементов 21 лингвистического предложения 11, из которых состоит первый или второй элемент 81 (ОЛЛО). При этом, как было описано ранее, лингвистическими признаками элементов 81 (ОЛЛО) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 81 (ОЛЛО).

[126] Далее, после завершения описанного ранее этапа 204, в целях еще большего повышения точности последующего поиска в структурированном массиве данных (СМД), становится возможным осуществить дальнейшее преобразование СМД, содержащего, по меньшей мере, основные лингво-логические объекты (ОЛЛО) лингвистического предложения и их идентификационные данные.

[127] На фиг. 25, в качестве примера, но не ограничения, изображена общая схема выполнения этапов заявленного способа 300 преобразования СМД, содержащего, по меньшей мере, основные лингво-логические объекты (ОЛЛО) лингвистического предложения и идентификационные данные ОЛЛО, являющегося исходной структурой данных для рассматриваемого способа 300. Заявленный способ 300 преобразования СМД, содержащего, по меньшей мере, ОЛЛО лингвистического предложения и идентификационные данные ОЛЛО, характеризуется: выполнением этапа 301 идентификации пригодной для преобразования структуры данных, содержащей ОЛЛО, на котором идентифицируют структуру данных СМД, содержащую элементы упомянутой пригодной для преобразования структуры данных, содержащей ОЛЛО, являющейся исходной структурой данных СМД в рамках рассматриваемого способа 300, причем упомянутые элементы пригодной для преобразования структуры данных представляют собой основные лингво-логические объекты (ОЛЛО) лингвистического предложения и идентификационные данные ОЛЛО; выполнением этапа 302

формирования восьмой структуры данных СМД, на котором формируют восьмую структуру данных СМД, являющуюся первой структурой данных СМД для рассматриваемого способа 300, содержащую элементы упомянутой восьмой структуры данных СМД, причем упомянутые элементы восьмой структуры данных СМД представляют собой исходные простые лингво-логические конструкции (ИП ЛЛК) лингвистического предложения, сформированные из синтаксически связанных ОЛЛО простого предложения исходного лингвистического предложения, а также представляют собой идентификационные данные ИП ЛЛК, представляющие собой для каждой ИП ЛЛК, в качестве примера, но не ограничения: значение ИП ЛЛК и порядковые номера ТЭ лингвистического предложения, составляющих ИП ЛЛК; выполнением этапа 303 формирования девятой структуры данных СМД, на котором формируют девятую структуру данных СМД, являющуюся второй структурой данных СМД для рассматриваемого способа 300, содержащую элементы девятой структуры данных СМД, причем упомянутые элементы девятой структуры данных СМД представляют собой простые лингво-логические конструкции (ПЛЛК) лингвистического предложения, сформированные из ИП ЛЛК путем устранения однородностей в ИП ЛЛК, а также представляют собой идентификационные данные ПЛЛК, представляющие собой для каждой ПЛЛК, в качестве примера, но не ограничения: значение ПЛЛК и порядковые номера ТЭ лингвистического предложения, составляющих ПЛЛК; выполнением этапа 304 формирования десятой структуры данных СМД, на котором формируют десятую структуру данных СМД, являющуюся третьей структурой данных СМД для рассматриваемого способа 300, содержащую элемент упомянутой десятой структуры данных СМД, причем упомянутый элемент десятой структуры данных СМД представляет собой исходную сложную лингво-логическую конструкцию (ИС ЛЛК) лингвистического предложения, сформированную путем объединения всех ПЛЛК с учетом синтаксических связей ПЛЛК в исходном лингвистическом предложении, а также представляет собой идентификационные данные ИС ЛЛК, представляющие собой, в качестве примера, но не ограничения: значение ИС ЛЛК и порядковые номера ТЭ лингвистического предложения, составляющих ИС ЛЛК; выполнением этапа 305 формирования одиннадцатой структуры данных СМД, на котором формируют одиннадцатую структуру данных СМД, являющуюся итоговой структурой данных СМД для рассматриваемого способа 300, содержащую элементы упомянутой одиннадцатой структуры данных СМД, причем упомянутые элементы одиннадцатой структуры данных СМД представляют собой сложные лингво-логические конструкции (СЛЛК) лингвистического предложения, сформированные из ИС ЛЛК путем устранения однородностей в ИС ЛЛК, а также представляют собой идентификационные данные СЛЛК, представляющие собой для каждой СЛЛК, в качестве примера, но не ограничения: значение СЛЛК и порядковые номера ТЭ лингвистического предложения, составляющих СЛЛК.

[128] На фиг. 26, в качестве примера, но не ограничения, изображена общая схема выполнения этапа 301 идентификации пригодной для преобразования структуры данных, содержащей основные лингво-логические объекты (ОЛЛО), являющейся исходной структурой данных для рассматриваемого способа 300. Этап 301 характеризуется выполнением идентификации пригодной для преобразования структуры данных, содержащей ОЛЛО, на котором идентифицируют элементы 91 пригодной для преобразования структуры данных, содержащей ОЛЛО, а также идентификационные данные элементов 91, представляющие собой для каждого ОЛЛО, в качестве примера, но не ограничения, значение 911 элемента 91 пригодной для преобразования структуры

данных, содержащей ОЛЛО и порядковый (порядковые) номер (номера) 912 ТЭ лингвистического предложения 11, составляющего (составляющих) элемент 91.

[129] На фиг. 27, в качестве примера, но не ограничения, изображена общая структура данных, являющаяся исходной для настоящего способа преобразования 300, представляющая собой пригодную для преобразования структуру данных, содержащую основные лингво-логические объекты (ОЛЛО). Исходная структура данных представляет собой СМД, содержащий элементы 91 пригодной для преобразования структуры данных, содержащей ОЛЛО, представляющие собой основные лингво-логические объекты (ОЛЛО) и идентификационные данные ОЛЛО, представляющие собой для каждого ОЛЛО, в качестве примера, но не ограничения, значение 911 элемента 91 пригодной для преобразования структуры данных содержащей ОЛЛО и порядковый (порядковые) номер (номера) 912 ТЭ лингвистического предложения, составляющего (составляющих) ОЛЛО, которые, в качестве примера, но не ограничения, были описаны ранее со ссылкой на фиг. 24 способа 200. У элемента 91 лингвистического предложения 11 имеются характеризующие их уникальные наименования (УН ОЛЛО), имеющие практическое использование в качестве примера, но не ограничения: «субъект суждения», «действие предиката суждения» и тому подобное. В структуре данных элементы 91 именуется уникальными наименованиями, в качестве примера, но не ограничения, могут именоваться как «УН ОЛЛО n», где $n \geq 1$ - порядковый индекс элемента уникального наименования в лингвистическом предложении, начиная с «1» для каждого нового УН ОЛЛО, ранее не встречавшегося в лингвистическом предложении. Как было сказано ранее, с лингвистической точки зрения ОЛЛО представляют собой синтаксическую единицу вида «синтаксическая конструкция», которая может быть сформирована из структурированных групп ЛЛЕ; с логической точки зрения ОЛЛО представляют собой элементы простого суждения, с помощью которых нечто утверждается или опровергается о предмете суждения; с семантической точки зрения ОЛЛО представляют собой обобщенный мыслимый (семантический) образ элемента простого суждения. Элементы 91 пригодной для преобразования структуры данных, содержащей ОЛЛО, представляют собой множество ОЛЛО уникальных наименований (УН ОЛЛО) лингвистического предложения 11. ОЛЛО 91 в пригодной для преобразования структуре данных, содержащей ОЛЛО, представляют собой отдельные элементы, заранее подготовленные и помещенные в пригодную для преобразования структуру данных, содержащую ОЛЛО в виде структурированного массива (списка, перечня и тому подобного) отдельных ОЛЛО. Такие подготовительные действия могут осуществляться любым известным из уровня техники способом или неизвестным из уровня техники способом, таким, как способ, указанный при описании способа 200, и, соответственно, далее не описываются.

[130] Идентификация элементов 91 пригодной для преобразования структуры данных, содержащей ОЛЛО, в ходе этапа 301 сводится к классификации элементов, из которых состоит пригодная для преобразования структура данных, содержащая ОЛЛО, как основных лингво-логических объектов (ОЛЛО 91). При этом все ОЛЛО 91 имеют значения 911 и порядковые номера текстовых элементов 912 лингвистического предложения, составляющие ОЛЛО. Значения 911 ОЛЛО 91 представляют собой один или несколько текстовых элементов (ТЭ 21) предложения 11, из которых состоят ОЛЛО. Количество ТЭ, из которых состоит ОЛЛО связано с технической возможностью языка при описании объекта, действия или их признаков, либо с потребностью автора описания сущностей окружающего мира (объектов, действий или их признаков) использовать определенный способ лингвистического описания, в качестве примера, но не

ограничения, можно привести различные следующие способы лингвистического описания одной и той же сущности: «старый человек» = «старик» = «человек преклонных лет» и тому подобное. Указанная сущность может быть лингвистически выражена через один, два или три ТЭ. При этом, например, не ограничиваясь, сущность «права потребителей» не может быть лингвистически выражена в русском языке одним ТЭ.

[131] Идентификацию значения 911 элемента 91 (ОЛЛО) пригодной для преобразования структуры данных, содержащей ОЛЛО, в ходе этапа 301 производят путем регистрации набора знаков текстовых элементов 21 предложения 11, из которых состоит ОЛЛО 91. Идентификацию порядкового (порядковых) номера (номеров) 912 ТЭ, из которого (которых) состоит элемент 91 (ОЛЛО) пригодной для преобразования структуры данных, содержащей ОЛЛО, в ходе этапа 301 производят путем регистрации порядкового (порядковых) номера (номеров) упомянутых ТЭ в соответствии с их расположением в лингвистическом предложении, при условии, что первый ТЭ в лингвистическом предложении имеет номер «1», а все последующие ТЭ имеют порядковый номер, больший на единицу, чем порядковый номер, предыдущего ТЭ.

[132] Пригодная для преобразования структура данных, содержащая ОЛЛО, являющаяся исходной для настоящего способа преобразования 300 представляет собой в итоге множество элементов 91, их значений 911 и порядковых номеров 912 ТЭ, составляющих элементы 22, идентифицированных на этапе 301.

[133] Идентификацию лингвистических характеристик ТЭ, составляющих элементы 91 (ОЛЛО) пригодной для преобразования структуры данных, содержащей ОЛЛО, и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 2031, состоящего из идентификационных данных ТЭ, составляющих элемент 91 (ОЛЛО), и получении значений (2131, 2141, 2231, 2151) всех частей лингвистических характеристик (213, 214, 223, 215) текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 91 (ОЛЛО). При этом, как было описано ранее, лингвистическими признаками элементов 91 (ОЛЛО) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 91 (ОЛЛО).

[134] На фиг. 28, в качестве примера, но не ограничения, изображена общая схема выполнения этапов этапа 302 формирования восьмой структуры данных СМД, являющейся первой структурой данных для рассматриваемого способа 300. Этап 302 характеризуется: выполнением этапа 3021 формирования значений 2161 пятой части лингвистических характеристик 216 элементов 21, составляющих элементы 91 пригодной для преобразования структуры данных, содержащей ОЛЛО на котором формируют значения 2161 пятой части лингвистических характеристик 216 элементов 21, составляющих элементы 91 и вносят полученные сведения в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2031, формируя в итоге БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 3021; выполнением этапа 3022 формирования элементов восьмой структуры данных СМД, на котором формируют элементы 12 восьмой структуры данных СМД, а также идентификационные данные элементов 12, представляющие собой для каждого элемента 12, в качестве примера, но не ограничения, значение 121 элемента 12 восьмой структуры данных СМД и порядковые номера 122 ТЭ лингвистического предложения 11, составляющих элементы 12, и формируют восьмую структуру данных СМД.

[135] На фиг. 29, в качестве примера, но не ограничения, изображена общая структура

формируемой в рамках этапа 3021 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 предложения 11, содержащихся в элементах 91 пригодной для преобразования структуры данных, содержащей ОЛЛО. БДЛП, формируемая в рамках этапа 3021, отличается от БДЛП, формируемой в рамках этапа 2031, наличием значений 2161 пятой части лингвистических характеристик 216, указывающих на синтаксическую роль элемента 91 (ОЛЛО). Поскольку с лингвистической точки зрения ОЛЛО – синтаксическая конструкция, то естественно, что в одном предложении одни синтаксические конструкции относительно других синтаксических конструкций при наличии между ними синтаксической подчинительной связи являются синтаксически главными или синтаксически зависимыми. По этому принципу одни ОЛЛО могут быть синтаксически главными (главными ОЛЛО), а другие – синтаксически зависимыми (зависимыми ОЛЛО). Выявление пятой части лингвистических характеристик 216 текстовых элементов 21 лингвистического предложения 11, составляющих ОЛЛО 91 и их значений 2161 сводится к классификации ОЛЛО 91 по их синтаксической роли. Синтаксическую роль ОЛЛО отождествляют с синтаксической ролью главной ЛЛЕ в ОЛЛО 91. С помощью значения 2161 пятой части лингвистических характеристик 216 элементов 21, составляющих ОЛЛО 91, все ОЛЛО 91 разделяют на главные ОЛЛО 91 и зависимые ОЛЛО 91.

[136] Формирование пятой части лингвистических характеристик 216 и их значений 2161 в ходе этапа 3021 производят путем анализа существующих лингвистических признаков элемента 91. В ходе такого анализа для примера, но не ограничения, могут производиться следующие действия: на первом этапе во всех ОЛЛО 91 идентифицируют синтаксически главную ЛЛЕ; на втором этапе среди лингвистических признаков главных ЛЛЕ упомянутых ОЛЛО 91 идентифицируют лингвистическую характеристику, отвечающую за синтаксическую роль синтаксически главной ЛЛЕ элемента 91; на третьем этапе проверяют значение идентифицированной лингвистической характеристики, отвечающей за синтаксическую роль синтаксически главной ЛЛЕ; если значение (синтаксическая роль синтаксически главной ЛЛЕ элемента 91) является «сказуемым», то такой ОЛЛО 91 (синтаксически главной ЛЛЕ элемента 91) получает значение 2161 пятой части лингвистических характеристик 216 – «главный ОЛЛО»; если значение не является «сказуемым», то такой ОЛЛО 91 (синтаксически главной ЛЛЕ элемента 91) получает значение 2161 пятой части лингвистических характеристик 216 – «зависимый ОЛЛО». В итоге формируют значения 2161 пятой части лингвистических характеристик 216 для всех элементов 91 (ОЛЛО).

[137] Сформированные по итогам этапа 3021 значения 2161 (синтаксическая роль синтаксически главной ЛЛЕ элемента 91) пятой части лингвистических характеристик 216 элементов 21, составляющих ОЛЛО 91, пригодной для преобразования структуры данных, содержащей ОЛЛО, вносят в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2031, формируя в итоге БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 3021. При этом все части (211, 214, 232, 215 и 216) лингвистических характеристик текстовых элементов 21 лингвистического предложения 11 и их значения (2111, 2141, 2321, 2151 и 2161) формируют уникальные лингвистические признаки элементов 91.

[138] На фиг. 30, в качестве примера, но не ограничения, изображена общая структура сформированной восьмой структуры данных СМД, являющейся первой структурой данных СМД для настоящего способа преобразования 300. Восьмая структура данных СМД представляет собой СМД, содержащий элементы 12 восьмой структуры данных СМД, которые представляют собой исходные простые лингво-логические конструкции

(ИП ЛЛК 12) лингвистического предложения 11 и идентификационные данные ИП ЛЛК, представляющие собой для каждой ИП ЛЛК, в качестве примера, но не ограничения, значение 121 элемента 12 восьмой структуры данных и порядковые номера текстовых элементов 122, составляющих ИП ЛЛК (значение и порядковые номера ТЭ

5 элемента 12). У ИП ЛЛК 12 лингвистического предложения 11 отсутствуют характеризующие их уникальные наименования (УН), имеющие практическое использование. В структуре данных элементы 12, в качестве примера, но не ограничения, могут именоваться как «ИП ЛЛК-1», «ИП ЛЛК-2», «ИП ЛЛК-3», «ИП ЛЛК-n», где $n \geq 1$ - порядковый номер элемента в лингвистическом предложении. ИП ЛЛК

10 предложения – это пограничная сущность между лингвистикой и логикой. С лингвистической точки зрения ИП ЛЛК – это синтаксическая единица вида «синтаксическая конструкция», сформированная из ОЛЛО путем объединения ОЛЛО по признаку наличия между ОЛЛО прямой синтаксической подчинительной или сочинительной связи. Упомянутая прямая синтаксическая подчинительная или

15 сочинительная связь означает наличие между элементами различных ОЛЛО (между ЛЛЕ различных ОЛЛО) синтаксической подчинительной или сочинительной связи. С точки зрения синтаксиса ИП ЛЛК – это отдельное исходное простое предложение или исходное простое предложение в составе исходного сложного предложения. При этом исходное простое предложение в ИП ЛЛК может содержать однородные члены, то

20 есть слова, между которыми имеется синтаксическая сочинительная связь. С логической точки зрения ИП ЛЛК – это логическая сущность, выражающая простое суждение (некванторное суждение), представляющее собой первичную логическую конструкцию мышления, с помощью которой формируется и передается мысль о том, что нечто (предикат суждения) утверждается или опровергается о предмете суждения (субъект

25 суждения). При этом нечто (то, что утверждается или опровергается о предмете суждения) может быть дополнительно логически структурировано. Например, не ограничиваясь, предикат суждения может быть разделен на три логически отделимые сущности – действие предиката суждения, объект предиката суждения и обстоятельства предиката суждения. Упомянутые для примера элементы простого суждения могут

30 быть отождествлены с отдельными УН ОЛЛО (ОЛЛО уникального наименования), из которых и формируется ИП ЛЛК. Кванторность простого суждения определяется его «количественной характеристикой», то есть наличием или отсутствием однозначности в простом суждении. По количественной характеристике (кванторности) все простые суждения можно разделить на две категории: однозначные простые

35 суждения и неоднозначные простые суждения. Однозначные простые суждения – это кванторные простые суждения, в которых показатель кванторности всегда определен и равен единице, то есть каждый элемент простого суждения не имеет вариативности, всегда однозначен. Однозначность простого суждения проявляется посредством нахождения в составе простого суждения элементов простого суждения одного

40 уникального наименования (УН) в единичном количестве (например, не ограничиваясь, одного субъекта суждения, объекта предиката суждения, действия предиката суждения или одного однородного обстоятельства предиката суждения). Примером однозначного простого суждения может быть следующее суждение: «Товар должен быть передан покупателю продавцом». Данное простое суждение не допускает никакой

45 неоднозначности - только «товар», только «должен быть передан», только «покупателю» и только «продавцом». Неоднозначные простые суждения – это кванторные простые суждения, в которых показатель кванторности как минимум у одного элемента простого суждения уникального наименования (УН) всегда определен и всегда больше единицы,

то есть элемент простого суждения в неоднозначном простом суждении всегда имеет вариативности, всегда неоднозначен. Неоднозначность простого суждения проявляется посредством нахождения в составе простого суждения нескольких элементов простого суждения одного уникального наименования (УН) (например, не ограничиваясь, 5 нескольких субъектов суждения и(или) объектов предиката суждения и(или) действий предиката суждения и(или) нескольких однородных обстоятельств предиката суждения). Примером неоднозначного простого суждения может быть следующее суждение: «Товар должен быть передан покупателю продавцом или менеджером компании». Данное простое суждение допускает неоднозначность ввиду наличия в нем двух элементов 10 простого суждения с одинаковыми УН - субъектов суждения «продавец» и «менеджер компании», к каждому из которых можно применить предикат суждения «должны передать товар покупателю». Если простые суждения, находящиеся в некоем массиве простых суждений (отдельной, обособленной группе простых суждений), имеют гарантировано одну и ту же количественную характеристику (показатель кванторности 15 равен единице или больше единицы), то вне зависимости от вида кванторности простого суждения (однозначное или неоднозначное) простые суждения такого массива являются «кванторными». Если простые суждения, находящиеся в некоем массиве простых суждений (отдельной, обособленной группе простых суждений), не имеют гарантировано одну и ту же количественную характеристику (показатель кванторности простых 20 суждений может быть как равен единице, так и быть большим единицы), то простые суждения такого массива именуется «некванторными» (без определенной кванторности). В связи с вышеизложенным ИП ЛЛК, формируемая из ОЛЛО исходного предложения, между которыми имеется синтаксическая сочинительная связь, может быть классифицирована только как «некванторное» простое суждение, поскольку ничем не 25 гарантируется наличие одной и той же количественной характеристики у всех простых суждений, напрямую (без какой-либо дополнительной обработки) формируемых из произвольного исходного лингвистического предложения.

[139] Формирование элементов 12 восьмой структуры данных СМД, представляющих собой исходные простые лингво-логические конструкции (ИП ЛЛК) в ходе этапа 3022 30 производят на основе результатов идентификации значений 2161 пятой части лингвистических характеристик 216 элементов 21, составляющих ОЛЛО 91. Для формирования элемента 12 восьмой структуры данных СМД необходимо осуществить следующие действия: на первом этапе выбирают главный ОЛЛО 91; на втором этапе идентифицируют все подчиненные ему зависимые ОЛЛО 91 с помощью значений 2161 35 пятой части лингвистических характеристик 216 элементов 21, составляющих ОЛЛО 91; на третьем этапе среди идентифицированных зависимых ОЛЛО 91 идентифицируют такие ОЛЛО 91, которые имеют с выбранным ранее главным ОЛЛО 91 прямую подчинительную связь; на четвертом этапе формируют предварительный элемент 12 восьмой структуры данных СМД, являющийся исходной простой лингво-логической 40 конструкцией (ИП ЛЛК), путем объединения выбранного «главного ОЛЛО» 91 и всех идентифицированных «зависимых ОЛЛО» 91, имеющих с выбранным главным ОЛЛО 91 прямые подчинительные связи; на пятом этапе идентифицируют сочиненные ОЛЛО среди уже идентифицированных зависимых ОЛЛО 91 с помощью значений 2131 первой 45 части лингвистических характеристик 213 элементов 21, составляющих ОЛЛО 91. Для завершения формирования элемента 12 формируют запрос в БДЛП, формируемую в рамках этапа 3021 для проверки наличия признаков связи между объединяемыми ОЛЛО, и, при наличии таких признаков, получают текстовый элемент (текстовые элементы), идентифицированный (идентифицированные) как одно из (или несколько из, или

комбинация из): знак препинания, союз или союзное слово, расположенные между объединяемыми ОЛЛО.

[140] Идентификацию значения и порядковых номеров ТЭ элемента 12 восьмой структуры данных СМД, производят в рамках этапа 3022 путем отождествления значения 121 элемента 12 со значениями 911 объединяемых элементов 91 и значениями 211 текстовых элементов, являющихся признаками связи согласно своим уникальным лингвистическим признакам, идентифицированных между объединяемыми элементами 91, а порядковых номеров 122 текстовых элементов 21, составляющих элемент 12 с порядковыми номерами 912 текстовых элементов, составляющих объединяемые элементы 91 и порядковыми номерами 21 текстовых элементов, являющихся признаками связи, идентифицированных между объединяемыми элементами 91.

[141] Формирование восьмой структуры данных СМД в ходе этапа 3022 производят путем объединения в одной структуре данных элементов 12 восьмой структуры данных СМД и их идентификационных данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются.

[142] Идентификацию лингвистических характеристик ТЭ, составляющих элементы 12 (ИП ЛЛК) восьмой структуры данных СМД, и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 3021, состоящего из идентификационных данных ТЭ, составляющих элемент 12 (ИП ЛЛК), и получении значений (2131, 2141, 2231, 2151, 2161) всех частей лингвистических характеристик (213, 214, 223, 215, 216) текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 12 (ИП ЛЛК). При этом, как было описано ранее, лингвистическими признаками элементов 12 (ИП ЛЛК) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 12 (ИП ЛЛК).

[143] На фиг. 31, в качестве примера, но не ограничения, изображена общая схема выполнения этапов этапа 303 формирования девятой структуры данных СМД, являющейся второй структурой данных для рассматриваемого способа 300. Этап 303 характеризуется: выполнением этапа 3031 идентификации видов элементов 12 восьмой структуры данных СМД, указывающих на наличие в ИП ЛЛК сочиненных ОЛЛО, на котором идентифицируют первые и вторые виды элементов 12 восьмой структуры данных СМД, а также идентифицируют непреобразованные элементы 13 девятой структуры данных СМД и их идентификационные данные, представляющие собой для каждого непреобразованного элемента 13, в качестве примера, но не ограничения, значение 131 непреобразованного элемента 13 девятой структуры данных СМД и порядковые 132 номера ТЭ лингвистического предложения 11, составляющих элемент 13 (значение и порядковые номера ТЭ непреобразованного элемента 13); выполнением этапа 3032 формирования преобразованных элементов 13 девятой структуры данных СМД, на котором формируют преобразованные элементы 13 девятой структуры данных СМД из элементов 12 второго вида, а также идентификационные данные преобразованных элементов 13, представляющие собой для каждого преобразованного элемента 13, в качестве примера, но не ограничения, значение 131 преобразованного элемента 13 девятой структуры данных СМД и порядковые 132 номера ТЭ лингвистического предложения 11, составляющих элемент 13 (значение и порядковые номера ТЭ преобразованного элемента 13); выполнением этапа 3033 формирования девятой структуры данных СМД, на котором формируют девятую структуру данных СМД из преобразованных элементов 13 и непреобразованных элементов 13 девятой структуры данных СМД.

[144] На фиг. 32, в качестве примера, но не ограничения, изображена общая структура сформированной девятой структуры данных СМД, являющейся второй структурой данных СМД для настоящего способа преобразования 300. Девятая структура данных СМД представляет собой СМД, содержащий элементы 13, которые представляют собой
 5 простые лингво-логические конструкции (ПЛЛК) лингвистического предложения 11, и идентификационные данные ПЛЛК, представляющие собой для каждой ПЛЛК, в качестве примера, но не ограничения, значение 131 элемента 13 девятой структуры данных и порядковые номера 132 текстовых элементов, составляющих ПЛЛК. Элементы
 10 13 девятой структуры данных СМД подразделяются по принципу их формирования на непреобразованные элементы 13 и преобразованные элементы 13. У ПЛЛК 13 лингвистического предложения 11 отсутствуют характеризующие их уникальные наименования (УН), имеющие практическое использование. В структуре данных элементы 13, в качестве примера, но не ограничения, могут именоваться как «ПЛЛК1», «ПЛЛК2», «ПЛЛК3», «ПЛЛК n », где $n \geq 1$ - порядковый номер элемента в
 15 лингвистическом предложении. ПЛЛК лингвистического предложения 11 – это пограничная сущность между лингвистикой и логикой. С лингвистической точки зрения, ПЛЛК – это синтаксическая единица вида «синтаксическая конструкция», сформированная из ИП ЛЛК путем гетерогенизации ИП ЛЛК. С точки зрения синтаксиса ПЛЛК – это отдельное простое предложение или простое предложение в
 20 составе исходного сложного предложения, которое не содержит однородных членов (слов между которыми имеется синтаксическая сочинительная связь). С логической точки зрения ПЛЛК – это логическая сущность, выражающая простое суждение (кванторное суждение), представляющее собой первичную логическую конструкцию мышления, с помощью которой формируется и передается мысль о том, что нечто
 25 (предикат суждения) утверждается или опровергается о предмете суждения (субъект суждения). При этом нечто (то, что утверждается или опровергается о предмете суждения) может быть дополнительно логически структурировано. Например, не ограничиваясь, предикат суждения может быть разделен на три логически отделимые сущности - действие предиката суждения, объект предиката суждения и обстоятельства
 30 предиката суждения. Упомянутые для примера элементы простого суждения могут быть отождествлены с отдельными видами ОЛЛО, из которых и формируется ПЛЛК. Отличие ПЛЛК от ИП ЛЛК в кванторности простого суждения. Если ИП ЛЛК допускает возможность наличия неоднозначности в простом суждении, то ПЛЛК имеет показатель кванторности равный единице, что полностью отождествляет ПЛЛК с
 35 абсолютно однозначным простым суждением. В составе ПЛЛК не может быть нескольких элементов простого суждения одного уникального наименования (УН). Например, не ограничиваясь, в составе ПЛЛК не может быть нескольких субъектов суждения, и (или) объектов предиката суждения, и (или) действий предиката суждения, и (или) однородных обстоятельств предиката суждения. В ПЛЛК не может быть никакой
 40 неоднозначности ни в субъекте суждения, ни в предикате суждения, вне зависимости от состава и конструкции исходного лингвистического предложения.

[145] Идентификацию видов элементов 12 восьмой структуры данных СМД в рамках этапа 3031 производят путем анализа значений лингвистических характеристик текстовых элементов 12 (ИП ЛЛК) с целью идентификации синтаксических
 45 сочинительных связей в ИП ЛЛК 12. Синтаксическую сочинительную связь между ОЛЛО (элемент 91) в ИП ЛЛК 12 идентифицируют в случае наличия у синтаксически главных ЛЛЕ таких ОЛЛО одного и того же синтаксического родителя (синтаксически главного слова, у которого с такой ЛЛЕ имеется прямая синтаксическая подчинительная

связь). Дополнительными условиями может быть наличие у главных ЛЛЕ таких ОЛЛО одной и той же синтаксической роли, а также наличие признака связи между такими ЛЛЕ, а именно - запятой или сочинительного союза. При идентификации синтаксической сочинительной связи между главными ЛЛЕ двух или более ОЛЛО такие ИП ЛЛК

5 идентифицируют как второй вид элемента 12 (ИП ЛЛК) восьмой структуры данных СМД. Все иные элементы 12 (ИП ЛЛК) в которых не идентифицирована синтаксическая сочинительная связь у главной ЛЛЕ идентифицируют как первый вид элемента 12 (ИП ЛЛК) восьмой структуры данных СМД. При этом идентифицированный первый вид

10 элементом 12 девятой структуры данных СМД. Формирование преобразованных элементов девятой структуры данных СМД (преобразованных элементов 13, полученных из элементов 12 второго вида) на этапе 3032 производят на основании идентифицированной на этапе 3031 синтаксической сочинительной связи между

15 элементами 91 (ОЛЛО) элемента 12 (ИП ЛЛК) путем преобразования элемента 12 второго вида. Смысл преобразования элемента 12 второго вида на этапе 3032 состоит в гетерогенизации ИП ЛЛК 12, то есть в устранении синтаксических однородностей в ИП ЛЛК 12 путем формирования из элемента 12 второго вида множества новых ПЛЛК, идентичных по однородностям элементу 12 первого вида. При этом каждая новая

20 ПЛЛК упомянутого множества может содержать только один из упомянутых ОЛЛО 91 с однородными членами (если элемент 12 второго вида содержит один ряд однородных членов) или только одно уникальное сочетание упомянутых ОЛЛО 91 с однородными членами разных рядов однородных членов (если элемент 12 второго вида содержит несколько рядов ОЛЛО 91 с однородными членами). Множество ПЛЛК (элементов 13), сформированных из элемента 12 второго вида являются гетерогенными

25 ПЛЛК, то есть сформированными из одной ИП ЛЛК 12 второго вида путем ее гетерогенного преобразования (гетерогенизации). Процесс гетерогенизации ИП ЛЛК 12 второго вида может быть проведен, в качестве примера, но не ограничения, следующим образом: на первом этапе выявляют ряды однородных членов в элементе

30 12 второго вида; в ходе выявления рядов однородных членов выявляют сами ряды однородных членов и порядковые номера однородных членов в ряду, а также признаки связи каждого однородного члена; на втором этапе выявляют цепочки однородных членов для каждого выявленного ОЛЛО, содержащего ЛЛЕ-ОЧ (ОЛЛО-ОЧ) элемента 12 второго вида; цепочка однородных членов (цепочка ОЧ) - это ОЛЛО-ОЧ и зависимая цепочка однородных членов (цепочка синтаксически связанных ОЛЛО, начиная от

35 первого синтаксического потомка однородных членов (синтаксически зависимого ОЛЛО) и продолжая по цепочке синтаксической зависимости (по синтаксической подчинительной связи) до другого ОЛЛО-ОЧ или до последнего ОЛЛО в элементе 12 второго вида; идентифицированные цепочки ОЧ классифицируют по факту наличия синтаксического потомка ЛЛЕ в зависимой цепочке ОЧ, который не входит в зависимую

40 цепочку ОЧ; те цепочки ОЧ, которые имеют зависимые цепочки ОЧ без синтаксического потомка за рамками зависимой цепочки однородных членов считаются крайними цепочками однородных членов (крайними ЦОЧ), а те цепочки ОЧ, которые имеют зависимые цепочки ОЧ с синтаксическим потомком за рамками зависимой цепочки ОЧ считаются внутренними цепочками однородных членов (внутренними цепочками ОЧ);

45 на третьем этапе выявляют ОЛЛО элемента 12 второго вида, не входящие в цепочки ОЧ (ОЛЛО вне цепочек ОЧ); на четвертом этапе формируют сменные зависимые цепочки (сменные ЗЦ) элемента 12 второго вида; сменная ЗЦ - это основа преобразованных элементов 13, состоящая из одной крайней цепочки ОЧ и внутренних

цепочек ОЧ, если такие внутренние цепочки ОЧ по цепочке непрерывной синтаксической подчинительной связи имеются между крайней цепочкой ОЧ и ОЛЛО вне цепочек ОЧ; на пятом этапе формируют гетерогенные ПЛЛК (преобразованные элементы 13); гетерогенная ПЛЛК формируется из одной сменной ЗЦ и ОЛЛО вне цепочки ОЧ (если 5 такие ОЛЛО есть); количество сменных ЗЦ элемента 12 второго вида определяет количество гетерогенных ПЛЛК (преобразованных элементов 13), сформированных от одного элемента 12 второго вида.

[146] Идентификация видов элементов 12 восьмой структуры данных СМД в рамках 10 этапа 3031 не изменяет значение 121 элемента 12 и не изменяет порядковые номера текстовых элементов 21 лингвистического предложения 11, составляющих элемент 12 (ИП ЛЛК). Идентификацию значения и порядковых номеров ТЭ непреобразованного 15 элемента 13 (ПЛЛК) девятой структуры данных СМД в рамках этапа 3031 производят следующим образом: значения 211 текстовых элементов 21 лингвистического предложения 11, составляющих элемент 12 первого вида восьмой структуры данных СМД идентифицируют как значение 131 непреобразованного элемента 13 (ПЛЛК), а 20 порядковые номера 212 упомянутых текстовых элементов 21 лингвистического предложения 11, составляющих элемент 12 из которого состоит непреобразованный элемент 13 идентифицируют как порядковые номера 132 текстовых элементов, составляющих непреобразованный элемент 13 (ПЛЛК). Идентификацию значения и 25 порядковых номеров ТЭ преобразованного элемента 13 (гетерогенный ПЛЛК) девятой структуры данных СМД в рамках этапа 3032 производят следующим образом: значения 211 текстовых элементов 21 лингвистического предложения 11, составляющих элемент 12 второго вида восьмой структуры данных СМД, за исключением значений 211 текстовых элементов 21 лингвистического предложения 11, составляющих отдельные 30 части элемента 12 второго вида (ОЛЛО-ОЧ), которые были удалены при гетерогенизации, и признаков связи удаленных ОЛЛО-ОЧ (синтаксически подчиненные главным ЛЛЕ удаленным ОЛЛО-ОЧ сочинительные союзы или знаки препинания) идентифицируют как значение 131 преобразованного элемента 13 (ПЛЛК), а порядковые 35 номера 212 текстовых элементов 21 лингвистического предложения 11, составляющих элемент 12 второго вида восьмой структуры данных СМД, за исключением порядковых номеров 212 текстовых элементов 21 лингвистического предложения 11, составляющих отдельные части элемента 12 второго вида (ОЛЛО-ОЧ), которые были удалены при гетерогенизации, и признаков связи удаленных ОЛЛО-ОЧ (синтаксически подчиненные 40 главным ЛЛЕ удаленным ОЛЛО-ОЧ сочинительные союзы или знаки препинания) идентифицируют как порядковые номера 132 текстовых элементов, составляющих преобразованный элемент 13 (гетерогенный ПЛЛК).

[147] Формирование девятой структуры данных СМД в рамках этапа 3033 производят 45 путем объединения ранее идентифицированных на этапах 3031 и 3032 элементов 13 девятой структуры данных СМД и их идентификационных данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются. Первым из объединяемых элементов девятой структуры данных СМД является идентифицированный на этапе 3031 элемент 12 первого вида, являющийся непреобразованным элементом 13, то есть ПЛЛК 13, совпадающий с ИП ЛЛК 12 без синтаксической сочинительной связи. Вторыми из объединяемых элементов девятой 50 структуры данных СМД являются преобразованные элементы 13, сформированные на этапе 3032 как множество преобразованных элементов 13, являющихся гетерогенными ПЛЛК 13, сформированными из элемента 12 второго вида (ИП ЛЛК с синтаксической сочинительной связью) путем гетерогенизации ИП ЛЛК второго

вида.

[148] Идентификацию лингвистических характеристик ТЭ, составляющих
 5 непреобразованные или преобразованные элементы 13 (ПЛЛК) девятой структуры
 данных СМД, и их значений при необходимости производят путем организации запроса
 в БДЛП, формируемую в рамках этапа 3021, состоящего из идентификационных данных
 ТЭ, составляющих непреобразованный или преобразованный элемент 13, и получении
 значений (2131, 2141, 2231, 2151, 2161) всех частей лингвистических характеристик (213,
 214, 223, 215, 216) текстовых элементов 21 лингвистического предложения 11, из которых
 10 состоит непреобразованный или преобразованный элемент 13 (ПЛЛК или гетерогенная
 ПЛЛК). При этом, как было описано ранее, лингвистическими признаками
 непреобразованных и преобразованных элементов 13 (ПЛЛК или гетерогенная ПЛЛК)
 являются как минимум морфологические, синтаксические и семантические
 характеристики текстовых элементов 21 лингвистического предложения 11, из которых
 состоят элементы 13 (ПЛЛК и гетерогенные ПЛЛК).

[149] На фиг. 33, в качестве примера, но не ограничения, изображена общая схема
 15 выполнения этапов этапа 304 формирования десятой структуры данных СМД
 являющейся третьей структурой данных для рассматриваемого способа 300. Этап 304
 характеризуется: выполнением этапа 3041 формирования значений 2171 шестой части
 лингвистических характеристик 217 элементов 21, составляющих элементы 13 девятой
 20 структуры данных СМД, на котором формируют значения 2171 шестой части
 лингвистических характеристик 217 элементов 21, составляющих элементы 13, и вносят
 полученные сведения в БДЛП текстовых элементов 21 лингвистического предложения
 11, формируемую в рамках этапа 3021, формируя в итоге БДЛП текстовых элементов
 21 лингвистического предложения 11, формируемую в рамках этапа 3041; выполнением
 25 этапа 3042 идентификации видов элементов 13 девятой структуры данных СМД,
 указывающих на наличие сочиненных ПЛЛК, на котором идентифицируют первые и
 последующие виды элементов 13 девятой структуры данных СМД; выполнением этапа
 3043 идентификации элементов 14 как составных частей элемента 15 десятой структуры
 30 данных СМД, на котором идентифицируют элементы 14 десятой структуры данных
 СМД, а также идентификационные данные элементов 14, представляющие собой для
 каждого элемента 14, в качестве примера, но не ограничения, значение 141 элемента
 14 десятой структуры данных СМД и порядковые номера 142 ТЭ лингвистического
 предложения 11, составляющих элемент 14; и формируют десятую структуру данных
 СМД путем объединения элементов 14 как составных частей единственного элемента
 35 15 десятой структуры данных СМД.

[150] На фиг. 34, в качестве примера, но не ограничения, изображена общая структура
 формируемой в рамках этапа 3041 базы данных лингвистических признаков (БДЛП),
 являющейся БДЛП текстовых элементов 21 предложения 11, содержащихся в элементах
 13 девятой структуры данных СМД. БДЛП, формируемая в рамках этапа 3041,
 40 отличается от БДЛП, формируемой в рамках этапа 3021, наличием шестой части
 лингвистических характеристик 217 и их значений 2171 текстовых элементов 21,
 содержащихся в элементах 13, устанавливающей синтаксическую роль ПЛЛК (элементов
 13 девятой структуры данных СМД). Синтаксическая роль ПЛЛК может иметь значения:
 главная ПЛЛК или зависимая ПЛЛК. Поскольку с лингвистической точки зрения
 45 ПЛЛК – синтаксическая конструкция, то естественно, что в одном предложении одни
 синтаксические конструкции относительно других синтаксических конструкций при
 наличии между ними синтаксической подчинительной связи являются синтаксически
 главными или синтаксически зависимыми. По этому принципу одни ПЛЛК могут быть

синтаксически главными (главными ПЛЛК), а другие – синтаксически зависимыми (зависимыми ПЛЛК).

[151] Формирование шестой части лингвистических характеристик 217 и их значений 2171 для текстовых элементов 21 лингвистического предложения 11, составляющих ПЛЛК 13 производят путем классификации ПЛЛК 13 по их синтаксической роли в ходе этапа 3041. Синтаксическую роль ПЛЛК отождествляют с синтаксической ролью главного ЛЛЕ в главном ОЛЛО 91 в ПЛЛК 13. В соответствии со значением 2171 шестой части лингвистических характеристик 217 элементов 21, составляющих ПЛЛК 13, подразделяют все ПЛЛК 13 на главные ПЛЛК 13 и зависимые ПЛЛК 13.

Формирование шестой части лингвистических характеристик 217 и их значений 2171 производят путем анализа существующих лингвистических признаков элементов 21, составляющих ПЛЛК 13. В ходе такого анализа, в качестве примера, но не ограничения, могут производиться следующие действия: на первом этапе во всех ПЛЛК 13 идентифицируют синтаксически главную ЛЛЕ синтаксически главной ОЛЛО 91; на втором этапе среди лингвистических признаков упомянутых главных ЛЛЕ упомянутого ПЛЛК 13 идентифицируют лингвистическую характеристику, отвечающую за синтаксическую роль упомянутой синтаксически главной ЛЛЕ элемента 13; на третьем этапе проверяют значение идентифицированной лингвистической характеристики, отвечающей за синтаксическую роль упомянутой синтаксически главной ЛЛЕ; если значение (синтаксическая роль упомянутой синтаксически главной ЛЛЕ элемента 13) является «сказуемым», то такой ПЛЛК 13 (синтаксически главная ЛЛЕ синтаксически главного элемента 91) предоставляют значение 2171 шестой части лингвистических характеристик 217– «главная ПЛЛК»; если значение не является «сказуемым», то такой ПЛЛК 13 предоставляют значение 2171 шестой части лингвистических характеристик 217 – «зависимая ПЛЛК». В итоге формируют значения 2171 шестой части лингвистических характеристик 217 для всех элементов ПЛЛК 13.

[152] Сформированные по итогам этапа 3041 значения 2171 (синтаксическая роль синтаксически главной ЛЛЕ синтаксически главного ОЛЛО элемента 13) шестой части лингвистических характеристик 217 элементов 21, составляющих ПЛЛК 13 девятой структуры данных СМД вносят в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 3041. При этом все части (211, 214, 232, 215, 216 и 217) лингвистических характеристик текстовых элементов 21 лингвистического предложения 11, входящих в ПЛЛК 13 и их значения (2111, 2141, 2321, 2151, 2161 и 2171) формируют уникальные лингвистические признаки элементов 13.

[153] На фиг. 35, в качестве примера, но не ограничения, изображена общая структура сформированной десятой структуры данных СМД, являющейся третьей структурой данных СМД для настоящего способа преобразования 300. Десятая структура данных СМД представляет собой СМД, содержащий элемент 15, который представляет собой исходную сложную лингво-логическую конструкцию (ИС ЛЛК) лингвистического предложения 11, и идентификационные данные элемента 15, представляющие собой, в качестве примера, но не ограничения, значение 151 элемента 15 десятой структуры данных и порядковые номера 152 текстовых элементов, составляющих элемент 15. У исходной сложной ЛЛК 15 лингвистического предложения 11 отсутствует характеризующее ее уникальное наименование (УН), имеющее практическое использование. ИС ЛЛК предложения 11 – это пограничная сущность между лингвистикой и логикой. С лингвистической точки зрения, ИС ЛЛК – это синтаксическая единица вида «синтаксическая конструкция», сформированная из ПЛЛК путем объединения ПЛЛК по признаку наличия между ПЛЛК прямой синтаксической

подчинительной и сочинительной связей. ИС ЛЛК состоит из того количества ПЛЛК, которое идентифицировано в исходном предложении. Упомянутая прямая синтаксическая подчинительная связь означает наличие между элементами различных ПЛЛК (между ЛЛЕ различных ПЛЛК) синтаксической подчинительной связи. Прямая синтаксическая сочинительная связь означает наличие между элементами различных ПЛЛК (между ЛЛЕ различных ПЛЛК) синтаксической сочинительной связи. С точки зрения синтаксиса ИС ЛЛК – это преобразованное исходное предложение без искажения его смысла. Суть преобразования исходного лингвистического предложения состоит в том, что все примененные лингвистические приемы (осложнения), направленные на упрощение лингвистических форм и структур мыслимых автором утверждений (отрицаний) в тексте лингвистического предложения (например, не ограничиваясь, использование однородных членов, причастных (деепричастных) оборотов) отменяют. Вместо исходного лингвистического предложения формируют массив простых предложений, которые синтаксически и логически связаны друг с другом. Упомянутые связи проясняют, а затем тем или иным образом регистрируют. Потребность такого преобразования актуальна потому, что одновременно с упрощением лингвистических форм и структуры лингвистического предложения применение упомянутых лингвистических приемов приводит к сложности и неоднозначности логической структуры исходного лингвистического предложения, к сложности и неоднозначности восприятия смысла исходного предложения пользователем, а также к большой технической сложности корректного интеллектуального машинного анализа содержания исходного лингвистического предложения. Преобразованное таким образом исходное предложение представляет собой (при условии применения в исходном лингвистическом предложении упомянутых приемов), по сути, сложное предложение, состоящее из преобразованных простых предложений в количестве, превышающем количество исходных простых предложений в исходном лингвистическом предложении. Однако в отличие от исходного лингвистического предложения все преобразованные простые предложения в ИС ЛЛК представляют собой однозначные простые суждения (мыслимые утверждения или отрицания). Регистрация элемента 15 (ИС ЛЛК) в форме структуры данных (массива преобразованных ПЛЛК (элементов 14)) позволяет выявить многоуровневую синтаксическую структуру элемента 15 (ИС ЛЛК) и связи между элементами этой структуры. Уровни ИС ЛЛК отражают структуру синтаксической подчинительной связи между структурированными ПЛЛК (СПЛЛК) как составными частями ИС ЛЛК и обусловлены выявленными синтаксическими связями между составными частями элемента 15 (элементами 14 лингвистического предложения 11.) Элемент 15 (ИС ЛЛК) формируется из упомянутых структурированных ПЛЛК (СПЛЛК) – элементов 14. В обязательном порядке ИС ЛЛК содержит только элемент 14 (СПЛЛК) первого уровня (элементы 13 первого вида). Наличие иных элементов 14 (СПЛЛК) второго и последующих уровней (элементов 13 второго и последующих видов) обусловлено лингвистическими осложнениями исходного лингвистического предложения. К СПЛЛК первого уровня (элементам 13 первого вида) относят ПЛЛК, которые не являются подчиненными простыми предложениями. К СПЛЛК второго уровня (элементам 13 второго вида) относят ПЛЛК, которые являются подчиненными простыми предложениями, при условии, что главными для них являются СПЛЛК первого уровня. К СПЛЛК третьего уровня (элементам 13 третьего вида) относят ПЛЛК, которые являются подчиненными простыми предложениями при условии, что главными для них являются СПЛЛК второго уровня. Эти утверждения таким образом справедливы и для СПЛЛК последующих уровней (четвертого, пятого, шестого и так

далее), и для специалиста в данной области техники должна быть очевидна возможность существования таких последующих уровней. Синтаксическая структура ИС ЛЛК 15, кроме синтаксически подчиненных СПЛЛК, может содержать и синтаксические сочиненные СПЛЛК. При этом все СПЛЛК первого уровня (элементы 13 первого вида) обязательно являются сочиненными друг с другом ПЛЛК. Начиная со СПЛЛК второго уровня (элементов 13 второго и последующих видов), СПЛЛК одного уровня могут как иметь, так и не иметь синтаксических сочинительных связей друг с другом.

Синтаксическую сочинительную связь между СПЛЛК в ИС ЛЛК идентифицируют в случае наличия у синтаксически главных слов (ЛЛЕ) синтаксически главных ОЛЛО таких СПЛЛК одного и того же синтаксического родителя (синтаксически главного слова, у которого с таким словом (ЛЛЕ) имеется прямая синтаксическая подчинительная связь). Дополнительными условиями может быть наличие у главных слов (ЛЛЕ) таких СПЛЛК одной и той же синтаксической роли, а также наличие признака связи между такими словами (ЛЛЕ), а именно - запятой или сочинительного союза. ИС ЛЛК 15 в отличие от упомянутых ранее ИП ЛЛК 12 и ПЛЛК 13 может быть только в единственном числе - из одного исходного лингвистического предложения можно сформировать только одну ИС ЛЛК 15. С логической точки зрения, ИС ЛЛК 15 – это логическая сущность, выражающая сложное суждение (некванторное суждение), представляющая собой логическую конструкцию, элементами которой являются простые суждения. Кванторность сложного суждения определяется его количественной характеристикой, то есть наличием или отсутствием однозначности в сложном суждении. По количественной характеристике (кванторности) все сложные суждения можно разделить на две категории: однозначные сложные суждения и неоднозначные сложные суждения. Однозначные сложные суждения – это кванторные сложные суждения, в которых показатель кванторности всегда определен и равен единице, то есть каждый элемент сложного суждения не имеет вариативности, всегда однозначен. Однозначность сложного суждения проявляется посредством нахождения в составе сложного суждения элементов, не имеющих возможность быть истолкованными как варианты утверждений или отрицаний, либо как варианты обусловленности одного и того же утверждения или отрицания. Примером однозначного сложного суждения может быть следующее суждение: «Если покупатель оплатил товар, то товар должен быть передан покупателю продавцом». Данное сложное суждение не допускает никакой вариативности - простое суждение «товар должен быть передан покупателю продавцом» обусловлено единственным простым суждением «если покупатель оплатил товар». Неоднозначные сложные суждения – это кванторные сложные суждения, в которых показатель кванторности всегда определен и всегда больше единицы, то есть, как минимум, один элемент сложного суждения имеет вариативность. Неоднозначность сложного суждения проявляется посредством нахождения в составе сложного суждения элементов, имеющих возможность быть истолкованными как варианты утверждений или отрицаний, либо как варианты обусловленности одного и того же утверждения или отрицания. Примером неоднозначного сложного суждения может быть следующее суждение: «Если покупатель оплатил товар, то товар должен быть передан покупателю продавцом или товар должен быть передан покупателю иным должностным лицом компании». Данное сложное суждение допускает вариативность утверждения - или «товар должен быть передан покупателю продавцом», или «товар должен быть передан покупателю иным должностным лицом компании». Сложное суждение состоит из простых суждений, имеющих исходные (установленные в исходном лингвистическом предложении) синтаксические связи между собой. В связи с тем, что такими связями могут быть как

синтаксические подчинительные связи, так и синтаксические сочинительные связи, нельзя гарантировать отсутствие в сложном суждении вариативности утверждений или отрицаний, либо вариативности обусловленностей утверждений (отрицаний). В связи с вышеизложенным ИС ЛЛК является сложным суждением с неопределенной

5 количественной характеристикой (показатель кванторности может быть как равен единице, так и быть больше единицы), и такое сложное суждение является «некванторным». С аналитической точки зрения ИС ЛЛК – это информационно

10 наполненная схема, проясняющая логическую конструкцию предложения (логические связи между отдельными простыми суждениями, заложенными в исходном лингвистическом предложении). В связи с этим ИС ЛЛК предпочтительно

15 демонстрируется не в лингвистической текстовой формы (в виде сложного предложения, состоящего из множества простых предложений без сочинительной связи и причастных (деепричастных) оборотов, в котором может иметься как подчинительная, так и сочинительная связь между упомянутыми простыми предложениями), а в табличной

20 или схематичной форме. Упомянутая табличная или схематичная форма демонстрации позволяет быстро и точно уяснить и проанализировать все используемые основные логические сущности (СПЛЛК) и все логические связи между упомянутыми логическими сущностями исходного лингвистического предложения. В частности, не ограничиваясь, такое развернутое и одновременно структурированное представление исходного

25 лингвистического предложения (логически связанное множество простых предложений без однородностей) позволяет повысить эффективность обработки лингвистического предложения 11 машинными средствами, поскольку вместо одного обрабатываемого предложения со сложной структурой появляется возможность обрабатывать множество простых предложений с идентифицированной логической связью между элементами

[154] Идентификацию видов элементов 13 девятой структуры данных СМД в ходе этапа 3042 производят на основе результатов идентификации значений 2171 шестой

30 части лингвистических характеристик 217 текстовых элементов 21, составляющих ПЛЛК 13. Для идентификации видов ПЛЛК 13 необходимо, в качестве примера, но не ограничения, осуществить следующие действия: на первом этапе из всех ПЛЛК 13

35 выбирают такие ПЛЛК, в которых главная ЛЛЕ главного ОЛЛО элемента 13 (ПЛЛК) не имеет синтаксического родителя; такая ПЛЛК будет являться элементом 13 первого вида - обязательным элементом десятой структуры данных СМД; на втором этапе из оставшихся ПЛЛК 13, вид которых еще не идентифицирован, выбирают такие ПЛЛК,

40 в которых главная ЛЛЕ главного ОЛЛО элемента 13 (ПЛЛК) имеет синтаксического родителя в ПЛЛК идентифицированной как элемент 13 первого вида; такая ПЛЛК будет являться элементом 13 второго вида – не обязательным элементом десятой структуры данных СМД; на третьем и последующих этапах из оставшихся ПЛЛК 13, вид которых еще не идентифицирован, выбирают такие ПЛЛК, в которых главная ЛЛЕ

45 главного ОЛЛО элемента 13 (ПЛЛК) имеет синтаксического родителя в ПЛЛК, идентифицированной как элемент 13 второго или последующих видов; такая ПЛЛК будет являться элементом 13 третьего или последующих видов – не обязательным элементом десятой структуры данных СМД. Наличие таких ПЛЛК обусловлено только синтаксической сложностью исходного лингвистического предложения. При этом

дополнительно проводят этап идентификации сочинительных связей между элементами 13 (ПЛЛК). ПЛЛК 13, у которых отсутствует синтаксическая подчинительная связь, но идентифицирована синтаксическая сочинительная связь, относят к тому же виду, что и ПЛЛК 13, с которым она сочинена. Идентификацию элементов 14 (составных

частей элемента 15) десятой структуры данных СМД в ходе этапа 3043 производят путем отождествления с элементами 13, идентифицированными на этапе 3042 как определенные виды элемента 13. При этом элементы 13 первого вида идентифицируют как элемент 14 первого уровня, элементы 13 второго вида идентифицируют как элемент 14 второго уровня и так далее, элементы 13 третьего и последующих видов идентифицируют как элемент 14 третьего и, соответственно, последующих уровней.

[155] Идентификация видов элементов 13 девятой структуры данных СМД в рамках этапа 3042 не изменяет значение 131 элемента 13 и не изменяет порядковые номера текстовых элементов 21 лингвистического предложения 11, составляющих элемент 13 (ПЛЛК). Идентификацию значения и порядковых номеров ТЭ элемента 14 (СПЛЛК) десятой структуры данных СМД в рамках этапа 3043 производят следующим образом: значения 211 текстовых элементов 21 лингвистического предложения 11, составляющих элемент 13, идентифицируют как значение 141 элемента 14 (СПЛЛК), а порядковые номера) 212 упомянутых текстовых элементов 21 лингвистического предложения 11, составляющих элемент 13, идентифицируют как порядковые номера 142 текстовых элементов, составляющих элемент 14 (СПЛЛК). Идентификацию значения и порядковых номеров ТЭ элемента 15 (ИС ЛЛК) десятой структуры данных СМД в рамках этапа 3043 производят следующим образом: значения 211 текстовых элементов 21 лингвистического предложения 11, составляющих все элементы 14, идентифицируют как значение 151 элемента 15 (ИС ЛЛК), а порядковые номера 212 (упомянутых текстовых элементов 21 лингвистического предложения 11, составляющих все элементы 14, идентифицируют как порядковые номера 152 текстовых элементов, составляющих элемент 15 (ИС ЛЛК).

[156] Формирование десятой структуры данных СМД в ходе этапа 3043 производят путем объединения в одной структуре данных элементов 14 десятой структуры данных СМД, а также их идентификационных данных. При этом объединяемые в одной структуре данных элементы 14, являясь составными частями элемента 15, формируют таким образом элемент 15, который является основным элементом десятой структуры данных СМД. Объединение в одной структуре данных упомянутых элементов и их идентификационных данных производится по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются. В итоге десятая структура данных СМД представляет собой массив (перечень) структурированных ПЛЛК (СПЛЛК) - элементов 14 первого и последующих уровней, расположенных в перечне СПЛЛК с учетом возрастания уровня элемента 14 (и указанием этого уровня), а также с учетом возрастания порядкового номера текстового элемента признака связи элемента 14 (и указанием этого порядкового номера), а также порядкового номера текстового элемента синтаксического родителя элемента 14 (и указанием этого порядкового номера). При этом первой частью ИС ЛЛК и первым порядковым номером в упомянутом перечне структурированных ПЛЛК будет СПЛЛК 14 без синтаксического родителя и без признака связи. Элементы 14 десятой структуры данных СМД используют для построения схемы или таблицы логических связей, визуальной отображающей логическую конструкцию исходного лингвистического предложения (логические связи между отдельными простыми суждениями, заложенными в исходном лингвистическом предложении). В частности, не ограничиваясь, как это было указано ранее со ссылкой на элементы 15, такое развернутое и одновременно структурированное представление исходного лингвистического предложения (логически связанное множество простых предложений без однородностей) позволяет повысить эффективность обработки лингвистического предложения 11 машинными средствами, поскольку вместо одного

обрабатываемого предложения со сложной структурой появляется возможность обрабатывать множество простых предложений с идентифицированной логической связью между элементами этого множества.

[157] Идентификацию лингвистических характеристик ТЭ, составляющих элемент 5 15 (ИС ЛЛК) десятой структуры данных СМД, и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 3041, состоящего из идентификационных данных ТЭ, составляющих элемент 15 (ИС ЛЛК), и получения значений (2131, 2141, 2231, 2151, 2161, 2171) всех частей лингвистических характеристик (213, 214, 223, 215, 216, 217) текстовых элементов 21 лингвистического 10 предложения 11, из которых состоит элемент 15 (ИС ЛЛК). При этом, как было описано ранее, лингвистическими признаками элемента 15 (ИС ЛЛК) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 15 (ИС ЛЛК).

[158] На фиг. 36, в качестве примера, но не ограничения, изображена общая схема 15 выполнения этапов этапа 305 формирования одиннадцатой структуры данных СМД, являющейся третьей структурой данных СМД для рассматриваемого способа 300. Этап 305 характеризуется: выполнением этапа 3051 идентификации видов элемента 15 десятой структуры данных СМД, указывающих на наличие в ИС ЛЛК сочиненных ПЛЛК, на 20 котором идентифицируют первый и второй виды элемента 15 десятой структуры данных СМД, а также идентификации непреобразованных элементов 16 одиннадцатой структуры данных СМД и их идентификационных данных, представляющих собой для каждого непреобразованного элемента 16, в качестве примера, но не ограничения, значение 161 25 непреобразованного элемента 16 одиннадцатой структуры данных СМД и порядковые номера 162 ТЭ лингвистического предложения 11, составляющих непреобразованный элемент 16; выполнением этапа 3052 формирования преобразованных элементов 16 30 одиннадцатой структуры данных СМД, на котором формируют преобразованные элементы 16 одиннадцатой структуры данных СМД из элементов 15 второго вида, а также идентификационные данные преобразованных элементов 16, представляющие собой для каждого преобразованного элемента 16, в качестве примера, но не 35 ограничения, значение 161 преобразованного элемента 16 одиннадцатой структуры данных СМД и порядковые номера 162 ТЭ лингвистического предложения 11, составляющих преобразованные элементы 16; выполнением этапа 3053 формирования 40 одиннадцатой структуры данных СМД, на котором формируют одиннадцатую структуру данных СМД из преобразованных элементов 16 одиннадцатой структуры данных СМД или из непреобразованного элемента 16 одиннадцатой структуры данных СМД.

[159] На фиг. 37, в качестве примера, но не ограничения, изображена общая структура сформированной одиннадцатой структуры данных СМД, являющейся четвертой 45 структурой данных СМД для настоящего способа преобразования 300. Одиннадцатая структура данных СМД представляет собой СМД, содержащий элементы 16 одиннадцатой структуры данных СМД, которые представляют собой сложные лингво-логические конструкции (СЛЛК) лингвистического предложения 11 и идентификационные данные ЛогО, представляющие собой для каждого элемента 16, в качестве примера, но не ограничения, значение 161 элемента 16 одиннадцатой 50 структуры данных СМД и порядковые номера 162 текстовых элементов, составляющих элемент 16. У СЛЛК лингвистического предложения 11 отсутствуют какие-либо характеризующие ее уникальные наименования (УН), имеющие практическое использование. В структуре данных элементы 16, в качестве примера, но не ограничения, могут именоваться как «СЛЛК1», «СЛЛК2», «СЛЛК3», «СЛЛК n », где $n \geq 1$ -

порядковый номер элемента в лингвистическом предложении. СЛЛК лингвистического предложения 11 – это пограничная сущность между лингвистикой и логикой. С лингвистической точки зрения, СЛЛК – это синтаксическая единица вида «синтаксическая конструкция», сформированная из ИС ЛЛК путем гетерогенизации ИС ЛЛК. С точки зрения синтаксиса СЛЛК – это отдельное простое предложение (если исходное лингвистическое предложение является простым предложением без синтаксических осложнений (однородных членов, оборотов и так далее)) или сложноподчиненное предложение, которое не содержит однородных членов (слов между которыми имеется синтаксическая сочинительная связь). С логической точки зрения, СЛЛК – это логическая сущность, выражающая сложное суждение (кванторное суждение), представляющее собой завершённую логическую конструкцию мышления, с помощью которой формируется и передается мысль о том, что нечто (предикат суждения) утверждается или опровергается о предмете суждения (субъект суждения). Отличие СЛЛК от ИС ЛЛК в кванторности сложного суждения. Если ИС ЛЛК допускает возможность наличия неоднозначности в сложном суждении, то СЛЛК (имеющая показатель кванторности, равный единице) является абсолютно однозначным сложным суждением. В составе СЛЛК не может быть нескольких однородных (сочиненных) простых суждений. В СЛЛК не может быть никакой неоднозначности ни при утверждении (опровержении) чего-либо, ни при обусловленности утверждения (опровержения), ни при обусловленности элементов утверждения (опровержения), вне зависимости от состава и конструкции исходного лингвистического предложения.

[160] Идентификацию вида ИС ЛЛК 15 в рамках этапа 3051 производят путем анализа значений лингвистических характеристик текстовых элементов, составляющих ИС ЛЛК 15, с целью идентификации синтаксических сочинительных связей между СПЛЛК 14 в ИС ЛЛК 15. Синтаксическую сочинительную связь между СПЛЛК 14 идентифицируют в случае наличия у синтаксически главных ЛЛЕ синтаксически главных ОЛЛО таких СПЛЛК одного и того же синтаксического родителя (синтаксически главной ЛЛЕ другой СПЛЛК, у которой со словом (ЛЛЕ) в сочиненных СПЛЛК имеется прямая синтаксическая подчинительная связь). Дополнительными условиями может быть наличие у главных ЛЛЕ главных ОЛЛО упомянутых сочиненных СПЛЛК одной и той же синтаксической роли, а также наличие признака связи между упомянутыми сочиненными ЛЛЕ, а именно - запятой или сочинительного союза. При идентификации синтаксической сочинительной связи между главными ЛЛЕ главных ОЛЛО двух или более СПЛЛК такую ИС ЛЛК идентифицируют как второй вид элемента 15. ИС ЛЛК 15, в которой не идентифицируют синтаксической сочинительной связи у главного ЛЛЕ главного ОЛЛО в СПЛЛК идентифицируют как первый вид элемента 15. При этом идентифицированный первый вид элемента 15 десятой структуры данных СМД отождествляют с преобразованным элементом 16 одиннадцатой структуры данных СМД (преобразованных элементов 16, полученных из элементов 15 второго вида) на этапе 3052 производят на основании идентифицированной на этапе 3051 синтаксической сочинительной связи между СПЛЛК путем преобразования ИС ЛЛК второго вида. Смысл преобразования ИС ЛЛК второго вида на этапе 3052 состоит в гетерогенизации ИС ЛЛК, то есть в устранении синтаксических однородностей в ИС ЛЛК путем формирования из элемента 15 второго вида множества новых ИС ЛЛК, идентичных по однородностям элементу 15 первого вида. При этом каждая новая ИС ЛЛК упомянутого множества может содержать только одну из упомянутых СПЛЛК с однородными членами (если ИС ЛЛК второго вида содержит один ряд однородных

членов) или только одно уникальное сочетание упомянутых СПЛЛК с однородными членами разных рядов однородных членов (если ИС ЛЛК второго вида содержит несколько рядов СПЛЛК с однородными членами). Множество ИС ЛЛК, сформированных из ИС ЛЛК (элемент 15) второго вида являются гетерогенными СЛЛК (преобразованными элементами 16), то есть сформированными из одной ИС ЛЛК второго вида путем ее гетерогенного преобразования (гетерогенизации). Процесс гетерогенизации ИС ЛЛК (элемента 15) второго вида может быть проведен, в качестве примера, но не ограничения, следующим образом: на первом этапе выявляют ряды однородных членов в элементе 15 второго вида; в ходе выявления рядов однородных членов выявляют сами ряды однородных членов и порядковые номера однородных членов в ряду, а также признаки связи каждого однородного члена; на втором этапе выявляют цепочки однородных членов для каждого выявленного СПЛЛК, содержащего ЛЛЕ-ОЧ (СПЛЛК-ОЧ) элемента 15 второго вида; цепочка однородных членов (цепочка ОЧ) - это СПЛЛК-ОЧ и зависимая цепочка однородных членов (цепочка синтаксически связанных СПЛЛК, начиная от первого синтаксического потомка однородных членов (синтаксически зависимой СПЛЛК) и продолжая по цепочке синтаксической зависимости (по синтаксической подчинительной связи) до другой СПЛЛК-ОЧ или до последней ПЛЛК в элементе 15 второго вида; идентифицированные цепочки ОЧ классифицируются по факту наличия синтаксического потомка ЛЛЕ в зависимой цепочке ОЧ, который не входит в зависимую цепочку ОЧ; те цепочки ОЧ, которые имеют зависимые цепочки ОЧ без синтаксического потомка за рамками зависимой цепочки однородных членов, считаются крайними цепочками однородных членов (крайними цепочками ОЧ), а те цепочки ОЧ, которые имеют зависимые цепочки ОЧ с синтаксическим потомком за рамками зависимой цепочки ОЧ, считаются внутренними цепочками однородных членов (внутренними цепочками ОЧ); на третьем этапе выявляют ПЛЛК элемента 15 второго вида, не входящие в цепочки однородных членов (ПЛЛК вне цепочек ОЧ); на четвертом этапе формируют сменные зависимые цепочки (сменные ЗЦ) элемента 15 второго вида; сменная ЗЦ - это основа преобразованных элементов 16, состоящая из одной крайней цепочки ОЧ и внутренних цепочек ОЧ, если такие внутренние цепочки ОЧ по цепочке непрерывной синтаксической подчинительной связи имеются между крайней цепочкой ОЧ и ПЛЛК вне цепочек ОЧ; на пятом этапе формируют гетерогенные СЛЛК (элементы 16); гетерогенная СЛЛК формируется из одной сменной ЗЦ и ПЛЛК вне цепочки ОЧ (если такие ПЛЛК есть); количество сменных ЗЦ элемента 15 второго вида определяет количество гетерогенных СЛЛК, сформированных от одного элемента 15 второго вида.

[161] Идентификация видов элементов 15 десятой структуры данных СМД в рамках этапа 3051 не изменяет значение 151 элемента 15 и не изменяет порядковые номера текстовых элементов 21 лингвистического предложения 11, составляющих элемент 15 (ИС ЛЛК). Идентификацию значения и порядковых номеров ТЭ непреобразованного элемента 16 (СЛЛК) одиннадцатой структуры данных СМД в рамках этапа 3051 производят следующим образом: значения 211 текстовых элементов 21 лингвистического предложения 11, составляющих элемент 15 первого вида десятой структуры данных СМД, идентифицируют как значение 161 непреобразованного элемента 16 (СЛЛК), а порядковые номера 152 упомянутых текстовых элементов 21 лингвистического предложения 11, составляющих элемент 15 первого вида, из которого состоит непреобразованный элемент 16, идентифицируют как порядковые номера 162 текстовых элементов, составляющих непреобразованный элемент 16 (СЛЛК). Идентификацию значения и порядковых номеров ТЭ преобразованного элемента 16 (гетерогенной

СЛЛК) одиннадцатой структуры данных СМД в рамках этапа 3052 производят следующим образом: значения 211 текстовых элементов 21 лингвистического предложения 11, составляющих элемент 15 второго вида десятой структуры данных СМД, за исключением значений 211 текстовых элементов 21 лингвистического предложения 11, составляющих отдельные части элемента 15 второго вида (СПЛЛК-ОЧ), которые были удалены при гетерогенизации, и признаков связи удаленных СПЛЛК-ОЧ (синтаксически подчиненные удаленным СПЛЛК-ОЧ сочинительные союзы или знаки препинания), идентифицируют как значение 161 преобразованного элемента 16 (СЛЛК), а порядковые номера 212 текстовых элементов 21 лингвистического предложения 11, составляющих элемент 15 второго вида шестой структуры данных СМД, за исключением порядковых номеров 212 текстовых элементов 21 лингвистического предложения 11, составляющих отдельные части элемента 15 второго вида (СПЛЛК-ОЧ), которые были удалены при гетерогенизации, и признаков связи удаленных СПЛЛК-ОЧ (синтаксически подчиненные удаленным СПЛЛК-ОЧ сочинительные союзы или знаки препинания) идентифицируют как порядковые номера 162 текстовых элементов, составляющих преобразованный элемент 16 (гетерогенная СЛЛК).

[162] Формирование одиннадцатой структуры данных СМД в рамках этапа 3053 производят путем объединения в одной структуре данных, идентифицированных на этапах 3051 и 3052 элементов 16 одиннадцатой структуры данных СМД, а также их идентификационных данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются. Первым из объединяемых элементов одиннадцатой структуры данных СМД является идентифицированный на этапе 3051 элемент 15 первого вида, являющийся непреобразованным элементом 16, то есть СЛЛК 16, совпадающей с ИС ЛЛК 15 без синтаксической сочинительной связи. Вторым из объединяемых элементов одиннадцатой структуры данных СМД являются преобразованные элементы 16, сформированные на этапе 3052 как множество преобразованных элементов 16, являющихся гетерогенными СЛЛК, сформированными из элемента 15 второго вида (ИС ЛЛК с синтаксической сочинительной связью) путем гетерогенизации ИС ЛЛК второго вида.

[163] Идентификацию лингвистических характеристик ТЭ, составляющих элемент 16 (СЛЛК) одиннадцатой структуры данных СМД, и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 3041, состоящего из идентификационных данных ТЭ, составляющих элемент 16 (СЛЛК), и получении значений (2131, 2141, 2231, 2151, 2161, 2171) всех частей лингвистических характеристик (213, 214, 223, 215, 216, 217) текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 16 (СЛЛК). При этом, как было описано ранее, лингвистическими признаками элемента 16 (СЛЛК) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 16 (СЛЛК).

[164] Вместе с тем, после завершения описанного ранее этапа 301, в целях обеспечения альтернативного повышения точности последующего поиска в структурированном массиве данных (СМД), альтернативно становится возможным осуществить дальнейшее преобразование СМД, содержащего, по меньшей мере, основные лингво-логические объекты (ОЛЛО) лингвистического предложения и их идентификационные данные. При этом специалисту в данной области техники должно быть очевидно, что описанные далее принципы и способы по завершении описанного ранее этапа 301 способа 300 могут осуществляться как одновременно (параллельно) с описанными ранее этапами

301-305 способа 300, так и не одновременно (альтернативно или последовательно, в том числе вне зависимости от порядка следования), а также, что отдельные принципы и способы отдельных описываемых далее этапов могут реализовываться аналогичным образом.

5 [165] На фиг. 38, в качестве примера, но не ограничения, изображена общая схема выполнения этапов заявленного способа 400 преобразования СМД, содержащего, по меньшей мере, основные лингво-логические объекты (ОЛЛО) лингвистического предложения и идентификационные данные ОЛЛО, являющегося исходной структурой данных для рассматриваемого способа 400. Заявленный способ 400 преобразования
10 СМД, содержащего, по меньшей мере, ОЛЛО лингвистического предложения и их идентификационные данные, характеризуется: выполнением этапа 301 идентификации пригодной для преобразования структуры данных, содержащей ОЛЛО, на котором идентифицируют структуру данных СМД, содержащую элементы упомянутой пригодной для преобразования структуры данных, содержащей ОЛЛО, являющейся исходной
15 структурой данных СМД в рамках рассматриваемого способа 400, причем упомянутые элементы пригодной для преобразования структуры данных представляют собой основные лингво-логические объекты (ОЛЛО) лингвистического предложения и идентификационные данные ОЛЛО; выполнением этапа 402 формирования двенадцатой структуры данных СМД, на котором формируют двенадцатую структуру данных СМД,
20 являющуюся первой структурой данных СМД для рассматриваемого способа 400, содержащую элементы упомянутой двенадцатой структуры данных СМД, причем упомянутые элементы двенадцатой структуры данных СМД представляют собой основные лингво-логические элементы (ОЛЛЭ) лингвистического предложения сформированные путем разделения ОЛЛО на семантические части ОЛЛО, а также
25 представляют собой идентификационные данные ОЛЛЭ, представляющие собой для каждого ОЛЛЭ, в качестве примера, но не ограничения: значение ОЛЛЭ и порядковый (порядковые) номер (номера) ТЭ лингвистического предложения, составляющих ОЛЛЭ; выполнением этапа 403 формирования тринадцатой структуры данных СМД, на котором формируют тринадцатую структуру данных СМД, являющуюся итоговой структурой
30 данных СМД для рассматриваемого способа 400, содержащую элементы упомянутой тринадцатой структуры данных СМД, причем упомянутые элементы тринадцатой структуры данных СМД представляют собой исходные лингво-логические элементы (ИЛЛЭ) лингвистического предложения, сформированные путем разделения ОЛЛЭ на семантические части ОЛЛЭ и формирования из них исходных элементов, а также
35 представляют собой идентификационные данные ИЛЛЭ, представляющие собой для каждого ИЛЛЭ, в качестве примера, но не ограничения: значение ИЛЛЭ и порядковый (порядковые) номер (номера) ТЭ лингвистического предложения, составляющих ИЛЛЭ.

[166] Как было указано ранее со ссылкой на фиг. 26, в качестве примера, но не
40 ограничения, изображена общая схема выполнения этапа 301 идентификации структуры данных СМД, на котором идентифицируют структуру данных СМД, являющуюся исходной структурой данных для рассматриваемого способа 400, элементы 91 которого представляют собой ОЛЛО лингвистического предложения 11, а также
идентификационные данные элементов 91, представляющие собой для каждого элемента
45 91, в качестве примера, но не ограничения, значение 911 элемента 91 и порядковый (порядковые) номер (номера) 912 ТЭ лингвистического предложения 11, составляющего (составляющих) элемент 91. Описанный таким образом со ссылкой на фиг. 26 процесс далее подробно не описывается. По результатам выполнения этапа 301 идентифицируют

структуру данных, показанную на фиг. 27, состав которой, описанный со ссылкой на фиг. 27, соответственно далее подробно не описывается.

[167] На фиг. 39, в качестве примера, но не ограничения, изображена общая схема выполнения этапов этапа 402 формирования двенадцатой структуры данных СМД, являющейся первой структурой данных для рассматриваемого способа 400. Этап 402 характеризуется: выполнением этапа 4021 формирования значений 2181 седьмой части лингвистических характеристик 218 элементов 21, составляющих элементы 91 пригодной для преобразования структуры данных, содержащей ОЛЛО на котором формирует значения 2181 седьмой части лингвистических характеристик 218 элементов 21, составляющих элементы 91 и вносят полученные сведения в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 3041, формируя в итоге БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 4021; выполнением этапа 4022 формирования семантических частей элементов 91 пригодной для преобразования структуры данных, содержащей ОЛЛО, на котором формируют первую и последующие (вторую, третью и так далее по порядку) семантические части ОЛЛО, проясняющие смысловую структуру ОЛЛО; выполнением этапа 4023 формирования двенадцатой структуры данных СМД, на котором идентифицируют элементы 17 двенадцатой структуры данных СМД, а также идентификационные данные элементов 17, представляющие собой для каждого элемента 17, в качестве примера, но не ограничения, значение 171 элемента 17 двенадцатой структуры данных СМД и порядковый (порядковые) номер (номера) 172 ТЭ лингвистического предложения 11, составляющего (составляющих) элемент 17, и формируют двенадцатую структуру данных СМД.

[168] На фиг. 40, в качестве примера, но не ограничения, изображена общая структура формируемой в рамках этапа 4021 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 предложения 11, содержащихся в элементах 91 пригодной для преобразования структуры данных, содержащей ОЛЛО, формируемой в рамках этапа 4021. БДЛП, формируемая в рамках этапа 4021, отличается от БДЛП, формируемой в рамках этапа 3021, наличием значений 2181 седьмой части лингвистических характеристик 218, указывающих главное слово семантической части элемента 91 (ОЛЛО). Причина выявления главного слова семантической части ОЛЛО состоит в том, что, используя выявленное главное слово семантической части ОЛЛО и идентифицированные ранее синтаксические связи между словами (ЛЛЕ) в ОЛЛО, можно разделить ОЛЛО на отдельные семантические части. Семантические части ОЛЛО проясняют смысловую структуру ОЛЛО, в которой отдельные части выполняют свои ограниченные смысловые роли в ОЛЛО. Семантические части в ОЛЛО позволяют выделить в ОЛЛО интегрированные в нем доминирующую и уточняющие сущности окружающего мира (объекты и действия) для корректной идентификации их смысловых ролей и, как следствие, смысла ОЛЛО. Выделение отдельных семантических частей в ОЛЛО позволяет установить их уникальные логические роли в ОЛЛО, что позволяет хранить эти части раздельно без потери их смысловой логической роли в предложении. Такое раздельное хранение позволяет осуществлять быстрый и точный поиск семантических частей ОЛЛО, а также отдельно от других частей лингвистического предложения производить разного рода их анализ и обработку.

[169] Формирование седьмой части лингвистических характеристик 218 и их значений 2181 в ходе этапа 4021 производят на основании комплексного анализа значений лингвистических характеристик текстовых элементов 21, входящих в элементы 91. В ходе такого анализа у всех текстовых элементов 21, составляющих элемент 91 проверяют

наличие признака главного слова семантической части элемента 91. Признаком главного слова семантической части элемента 91 может быть, в качестве примера, но не ограничения, наличие у текстового элемента (ТЭ) следующих значений лингвистических характеристик: ТЭ является главным словом ОЛЛО или ТЭ имеет зависимый ТЭ в форме предлога. Предварительно могут быть заданы иные собственные критерии идентификации синтаксической части ОЛЛО. Например, не ограничиваясь, может быть уточнен перечень частей речи, которые указывают на главное слово синтаксической части ОЛЛО. При наличии у ТЭ указанных значений лингвистических характеристик формируют соответствующие значения 2181 седьмой части лингвистических характеристик 218 - «главное слово семантической части ОЛЛО».

[170] Сформированные по итогам этапа 4021 значения 2181 (главное слово семантической части элемента 91) седьмой части лингвистических характеристик 218 элементов 21, составляющих ОЛЛО 91 пригодной для преобразования структуры данных, содержащей ОЛЛО вносят в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 2031, формируя в итоге БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 4021. При этом все части (211, 214, 232, 215, 216 и 218) лингвистических характеристик текстовых элементов 21 лингвистического предложения 11 и их значения (2111, 2141, 2321, 2151, 2161 и 2181) формируют уникальные лингвистические признаки элементов 91 для альтернативного способа повышения точности поиска в структурированном массиве данных.

[171] На фиг. 41, в качестве примера, но не ограничения, изображена общая структура сформированной двенадцатой структуры данных СМД являющаяся первой для настоящего способа преобразования 400. Двенадцатая структура данных СМД представляет собой СМД, содержащий элементы 17 двенадцатой структуры данных СМД, которые представляют собой основные лингво-логические элементы (ОЛЛЭ) лингвистического предложения 11 и идентификационные данные ОЛЛЭ, представляющие собой, в качестве примера, но не ограничения, значения 171 элементов 17 двенадцатой структуры данных и порядковые номера 172 текстовых элементов, составляющих ОЛЛЭ. У ОЛЛЭ 12 лингвистического предложения 11 отсутствуют характеризующие их уникальные наименования (УН), имеющие практическое использование. В структуре данных элементы 12, в качестве примера, но не ограничения, могут именоваться как «ОЛЛЭ n.m», где $n \geq 1$ указывает на порядковый номер ОЛЛО в лингвистическом предложении, а $m \geq 1$ указывает на синтаксический уровень ОЛЛЭ в ОЛЛО, начиная с первого. ОЛЛЭ предложения 11 – это пограничная сущность между лингвистикой и логикой. С лингвистической точки зрения, ОЛЛЭ - это синтаксическая единица вида «синтаксическая конструкция», сформированная из ОЛЛО 91 путем выделения в ОЛЛО 91 отдельных синтаксических частей по признаку наличия главных слов синтаксических частей в ОЛЛО 91. Упомянутые главные слова синтаксической части ОЛЛО 91 представляют собой такие части речи как существительное, местоимение или глагол (в любой грамматической форме), которые либо являются синтаксически главным словом ОЛЛО 91, либо содержат прямого синтаксического потомка (слово, с которым имеется прямая синтаксическая подчинительная связь) с синтаксической ролью «предлог». Предварительно могут быть заданы иные критерии идентификации синтаксической части ОЛЛО 91. Например, не ограничиваясь, возможно задать критерии в виде конкретного перечня предлогов, наличие которых указывают на главное слово синтаксической части ОЛЛО 91. С точки зрения синтаксиса ОЛЛЭ 17 – это слова и особые словосочетания внутри ОЛЛО 91. В зависимости от сложности лингвистической

структуры ОЛЛО 91 в нем можно выделить один или несколько ОЛЛЭ 17. С логической точки зрения, ОЛЛЭ 17 – это логическая сущность, являющаяся логически отделимой частью элемента простого суждения. При этом такие логически отделимые части по своей структуре и логической роли ничем не отличаются от ОЛЛО 91 по той причине, что, так же, как и ОЛЛО 91, они представляют мыслимый (семантический) образ некого предмета или действия окружающего мира). Причина существования семантических частей образа ОЛЛО 91 в том, что не всякий образ ОЛЛО 91 можно лингвистически реализовать, используя лишь один составной элемент (группу слов без предлога между словами). Например, не ограничиваясь, мыслимый образ ОЛЛО 91 о праве на что-либо (например, не ограничиваясь, «право потребителя на возврат товара») невозможно реализовать одним составным элементом (одной группой слов без предлога между словами)). В упомянутом примере лишь два составных элемента - первый - «право потребителя», и второй - «возврат товара», - формируют мыслимый образ элемента простого суждения. Составной элемент ОЛЛО 91 (семантическая часть ОЛЛО) именуется термином ОЛЛО 91. С семантической точки зрения термины ОЛЛО 91, формирующие ОЛЛО 91 неодинаковы, то есть имеют различные семантические роли. Синтаксически главные термины ОЛЛО 91 имеют максимальную семантическую роль, то есть являются терминами ОЛЛО 91 первой семантической части. Термин ОЛЛО 91 первой семантической части определяет смысловой вектор ОЛЛО 91. Например, не ограничиваясь, в примере «право потребителя на возврат товара» смысловую суть образа ОЛЛО 91 формирует именно составной элемент «право потребителя», а не составной элемент «возврат товара». Термины ОЛЛО второй и последующих семантических частей, являются синтаксически зависимыми терминами ОЛЛО 91, и их присутствие в ОЛЛО 91 необходимо только для уточнения мыслимого образа ОЛЛО 91. Чем сложнее мыслимый образ ОЛЛО 91, тем большего количества терминов ОЛЛО 91 может потребовать уточнение его мыслимого образа. Структурирование терминов ОЛЛО 91 как семантических частей ОЛЛО позволяет правильно воспринимать семантические роли отдельных терминов в ОЛЛО 91. Разделение таким образом ОЛЛО 91 позволяет производить интеллектуальный поиск информации с повышенной точностью, так как при наличии более структурированного (по логической роли слов) массива данных возможно более точное описание искомых объектов. В результате машинной обработки такого массива данных повышается точность поиска и сокращается время поиска,

[172] Формирование частей (семантических частей) элементов 91 пригодной для преобразования структуры данных, содержащей ОЛЛО, на этапе 4022 производят путем выполнения следующих действий: на первом этапе от идентифицированного главного слова семантической части ОЛЛО 91 идентифицируют все синтаксически подчиненные слова, вплоть до слова, которое само является главным словом семантической части ОЛЛО 91, но не включая данное слово; на втором этапе главное слово семантической части ОЛЛО 91 и все упомянутые синтаксически подчиненные слова, в том числе последовательно подчиненные, формируют семантическую часть ОЛЛО 91. Порядковый номер семантической части в ОЛЛО, указывает на ее семантическую роль в ОЛЛО и формируется в соответствии с порядковым номером текстового элемента (ТЭ) главного слова семантической части ОЛЛО 91. Семантическая часть ОЛЛО 91 с наименьшим порядковым номером ТЭ главного слова является первой семантической частью ОЛЛО 91. Семантическая часть ОЛЛО 91, имеющая ближайший к нему номер ТЭ главного слова семантической части является второй семантической частью ОЛЛО 91, и так далее. Порядковый номер ОЛЛЭ указывается в наименовании

ОЛЛЭ, например, но не ограничиваясь - «ОЛЛЭ 1.2». Такое наименование элемента 17 говорит о том, что он является второй семантической частью в первом ОЛЛО (по синтаксическому уровню «.2»). Идентификацию элементов 17 двенадцатой структуры данных СМД в ходе этапа 4023 производят на основе этапа 4022 формирования семантических частей элементов 91 пригодной для преобразования структуры данных, содержащей ОЛЛО. Все сформированные семантические части элементов 91 отождествляют с элементами 17 двенадцатой структуры данных СМД.

[173] Идентификацию значения и порядковых номеров ТЭ элемента 17 двенадцатой структуры данных СМД, производят в рамках этапа 4023 путем отождествления значения 171 элемента 17 со значениями 911 идентифицированной семантической части элемента 91, а порядковых номеров 172 текстовых элементов 21, составляющих элемент 17 путем отождествления с порядковыми номерами 912 текстовых элементов, составляющих идентифицированную семантическую часть элемента 91.

[174] Формирование двенадцатой структуры данных СМД в ходе этапа 4023 производят путем объединения в одной структуре данных элементов 17 двенадцатой структуры данных СМД и их идентификационных данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются.

[175] Идентификацию лингвистических характеристик ТЭ, составляющих элементы 17 (ОЛЛЭ) двенадцатой структуры данных СМД и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 4021, состоящего из идентификационных данных ТЭ, составляющих элемент 17 (ОЛЛЭ), и получении значений (2131, 2141, 2231; 2151, 2161, 2181) всех частей лингвистических характеристик (213, 214, 223, 215, 216, 218) текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 17 (ОЛЛЭ). При этом, как было описано ранее, лингвистическими признаками элементов 17 (ОЛЛЭ) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 17 (ОЛЛЭ).

[176] На фиг. 42, в качестве примера, но не ограничения, изображена общая схема выполнения этапов этапа 403 формирования тринадцатой структуры данных СМД, являющейся второй структурой данных для рассматриваемого способа 400. Этап 403 характеризуется выполнением этапа 4031 формирования значений 2191 восьмой части лингвистических характеристик 219 элементов 21, составляющих элементы 17 двенадцатой структуры данных СМД, на котором формируют значения 2191 восьмой части лингвистических характеристик 219 элементов 21, составляющих элементы 91, и вносят полученные сведения в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 4021, формируя в итоге БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 4031; выполнением этапа 4032 формирования семантических частей элементов 17 двенадцатой структуры данных СМД, на котором формируют первую и последующие (вторую, третью и так далее по порядку) семантические части ОЛЛЭ, проясняющие смысловую структуру ОЛЛЭ; выполнением этапа 4033 формирования элементов 18 тринадцатой структуры данных СМД, на котором формируют элементы 18 тринадцатой структуры данных СМД и идентификационные данные элементов 18, представляющие собой, в качестве примера, но не ограничения, значения 181 элементов 18 тринадцатой структуры данных СМД, порядковые номера 182 ТЭ лингвистического предложения 11, составляющих элементы 18, а также идентифицируют первые и вторые виды элементов 18 тринадцатой структуры данных СМД и формируют тринадцатую структуру данных

СМД.

[177] На фиг. 43, в качестве примера, но не ограничения, изображена общая структура формируемой в рамках этапа 4031 базы данных лингвистических признаков (БДЛП), являющейся БДЛП текстовых элементов 21 предложения 11, содержащихся в элементах 5 17 двенадцатой структуры данных СМД. БДЛП, формируемая в рамках этапа 4031, отличается от БДЛП, формируемой в рамках этапа 4021, наличием значений 2191 восьмой части лингвистических характеристик 219, указывающих главное слово семантической части элемента 17 (ОЛЛЭ). Причина выявления главного слова семантической части ОЛЛЭ состоит в том, что, используя выявленное главное слово 10 семантической части ОЛЛЭ и идентифицированные ранее синтаксические связи между словами (ЛЛЕ) в ОЛЛЭ можно разделить ОЛЛЭ на отдельные семантические части. Семантические части ОЛЛЭ проясняют смысловую структуру ОЛЛЭ, в которой отдельные части выполняют свои ограниченные смысловые роли в ОЛЛЭ. Семантические части в ОЛЛЭ позволяют выделить в ОЛЛЭ интегрированные в нем 15 доминирующую и уточняющие сущности окружающего мира (объекты и действия) для корректной идентификации их смысловых ролей и в итоге смысла ОЛЛЭ. Выделение отдельных семантических частей в ОЛЛЭ позволяет установить их уникальные логические роли в ОЛЛЭ, что позволяет хранить эти части отдельно без потери их смысловой логической роли в предложении. Такое раздельное хранение позволяет 20 осуществлять быстрый и точный поиск семантических частей ОЛЛЭ, а также отдельно от других частей лингвистического предложения производить разного рода их анализ и обработку.

[178] Формирование восьмой части лингвистических характеристик 219 и их значений 2191 в ходе этапа 4031 производят на основании комплексного анализа значений 25 лингвистических характеристик текстовых элементов 21, входящих в элементы 17. В ходе такого анализа у всех текстовых элементов 21, составляющих элемент 17 проверяют наличие признака главного слова семантической части элемента 17. Признаком главного слова семантической части элемента 17 может быть, в качестве примера, но не ограничения, наличие у текстового элемента (ТЭ) следующих значений лингвистических 30 характеристик: ТЭ является главным словом ОЛЛЭ или ТЭ имеет форму существительного, местоимения или глагола (в любой грамматической форме). Предварительно могут быть заданы иные собственные критерии идентификации синтаксической части ОЛЛЭ. Например, не ограничиваясь, может быть уточнен перечень частей речи, которые указывают на главное слово синтаксической части ОЛЛЭ. При 35 наличии у ТЭ указанных значений лингвистических характеристик формируют соответствующие значения 2191 восьмой части лингвистических характеристик 219 - «главное слово семантической части ОЛЛЭ».

[179] Сформированные по итогам этапа 4031 значения 2191 (главное слово семантической части элемента 17) восьмой части лингвистических характеристик 219 40 элементов 21, составляющих ОЛЛЭ 17 двенадцатой структуры данных СМД вносят в БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 4021, формируя в итоге БДЛП текстовых элементов 21 лингвистического предложения 11, формируемую в рамках этапа 4031. При этом все части (211, 214, 232, 215, 216, 218, 219) лингвистических характеристик текстовых элементов 21 45 лингвистического предложения 11 и их значения (2111, 2141, 2321, 2151, 2161, 2181, 2191) формируют уникальные лингвистические признаки элементов 17 для альтернативного способа повышения точности поиска в структурированном массиве данных.

[180] На фиг. 44, в качестве примера, но не ограничения, изображена общая структура

сформированной тринадцатой структуры данных СМД, являющейся второй структурой данных СМД для настоящего способа преобразования 400. Тринадцатая структура данных СМД представляет собой СМД, содержащий элементы 18 первого и второго вида тринадцатой структуры данных СМД, которые представляют собой исходные

5 лингво-логические элементы (ИЛЛЭ) лингвистического предложения 11 и идентификационные данные ИЛЛЭ, представляющие собой для каждого ИЛЛЭ, в качестве примера, но не ограничения: значение 181 элемента 18 шестой структуры данных СМД, порядковый (порядковые) номер (номера) 182 ТЭ лингвистического предложения 11, составляющих элементы 18. У ИЛЛЭ 18 лингвистического предложения

10 11 отсутствуют характеризующие их уникальные наименования (УН), имеющие практическое использование. В структуре данных элементы 18, в качестве примера, но не ограничения, могут именоваться как «ИЛЛЭ n.m.k», где $n \geq 1$ указывает на порядковый номер ОЛЛО в лингвистическом предложении, а $m \geq 1$ указывает на синтаксический уровень ОЛЛЭ в ОЛЛО, начиная с первого, а $k \geq 1$ указывает на синтаксический уровень ИЛЛЭ в ОЛЛЭ, начиная с первого. ИЛЛЭ 18 предложения 11

15 – это пограничная сущность между лингвистикой и логикой. С лингвистической точки зрения ИЛЛЭ 18 – это первичная синтаксическая единица, являющаяся словом. С логической точки зрения ИЛЛЭ 18 – это логическая сущность, являющаяся логически отделимой частью термина элемента простого суждения (термина ОЛЛО). При этом

20 такие логически отделимые части по своей структуре и логической роли ничем не отличаются от ОЛЛЭ 17 по той причине, что так же, как и ОЛЛЭ 17, они представляют собой мыслимый (семантический) образ о некоем предмете или действии окружающего мира). Причина существования синтаксических частей ОЛЛЭ 17 в том, что не всякий мыслимый образ ОЛЛЭ 17 можно представить, используя лишь один «исходный

25 элемент» (слово, выражающее объект или действие). Например, не ограничиваясь, мыслимый образ ОЛЛЭ 17 о правах кого-либо, например, не ограничиваясь, «права потребителя», невозможно реализовать одним исходным элементом. В упомянутом примере лишь два исходных элемента – первый - «права», и второй - «потребителя», - формируют мыслимый образ термина элемента простого суждения (термина ОЛЛО).

30 Исходный элемент ОЛЛЭ 17 именуется понятием термина ОЛЛО (ПТ ОЛЛО) или элементов 18 (ИЛЛЭ). С семантической точки зрения ПТ ОЛЛО (ИЛЛЭ 18), формирующие термин ОЛЛО (ОЛЛЭ 17) неравнозначны, то есть имеют различные семантические роли. Синтаксически главные ПТ ОЛЛО имеют максимальную семантическую роль, то есть являются ПТ ОЛЛО первой семантической части. ПТ

35 ОЛЛО первой семантической части определяют смысловой вектор термина ОЛЛО (ОЛЛЭ 17). Например, в упомянутом примере «права потребителя» суть образа термина ОЛЛО (ОЛЛЭ 17) - именно «права», а не «потребитель». ПТ ОЛЛО второй и последующих семантических частей (синтаксически зависимые ПТ ОЛЛО) в термине ОЛЛО (ОЛЛЭ 17) необходимы для уточнения смысла термина ОЛЛО (ОЛЛЭ 17). Чем

40 сложнее мыслимый образ термина ОЛЛО (ОЛЛЭ 17), тем большего количества ПТ ОЛЛО (ИЛЛЭ 18) он может потребовать. Структурирование ПТ ОЛЛО (ИЛЛЭ 18) как семантических частей ОЛЛЭ 17 позволяет правильно воспринимать семантические роли отдельных ПТ ОЛЛО (ИЛЛЭ 18) в терминах ОЛЛО (ОЛЛЭ 17). Элемент 18 (ИЛЛЭ) формируют из ОЛЛЭ 17 путем выделения в ОЛЛЭ 17 отдельных синтаксических

45 частей по признаку наличия главных слов синтаксических частей ОЛЛЭ 17 и последующего структурирования текстовых элементов синтаксической части. В зависимости от сложности лингвистической структуры ОЛЛЭ 17 в нем можно выделить одну или несколько синтаксических частей (ИЛЛЭ 18). Синтаксическая часть (ИЛЛЭ

18) ОЛЛЭ 17 может представлять собой только одно упомянутое главное слово или несколько синтаксически связанных слов (значимых частей речи), синтаксически главным из которых является упомянутое главное слово. Таким образом главное слово синтаксической части (ИЛЛЭ 18) ОЛЛЭ 17 отождествляется с элементом 18 первого вида, а все синтаксически зависимые от него слова (значимые части речи), если такие имеются, отождествляются с элементом 18 второго вида. С логической точки зрения исходные лингво-логические элементы (элементы 18) первого вида являются «понятиями» - то есть объектами или действиями окружающего мира, а элементы 18 второго вида – «признаками понятия», то есть некими семантическими характеристиками упомянутых объектов и действий.

[181] Формирование частей (семантических частей) элементов 17 двенадцатой структуры данных СМД, на этапе 4032 производят путем выполнения следующих действий: на первом этапе от идентифицированного главного слова элемента 17 (ОЛЛЭ) идентифицируют все синтаксически подчиненные слова, вплоть до слова, которое само является главным словом семантической части элемента 17 (ОЛЛЭ), но не включая данное слово; на втором этапе главное слово семантической части элемента 17 (ОЛЛЭ) и все упомянутые синтаксически подчиненные слова, в том числе последовательно подчиненные, формируют семантическую часть элемента 17 (ОЛЛЭ). Порядковый номер семантической части в ОЛЛЭ, указывает на ее семантическую роль в ОЛЛЭ, и его формируют в соответствии с номером текстового элемента (ТЭ) главного слова семантической части ОЛЛЭ 17. Семантическая часть ОЛЛЭ 17 с наименьшим номером ТЭ главного слова является первой семантической частью элемента 17 (ОЛЛЭ). Семантическая часть ОЛЛЭ 17, имеющая ближайший к нему номер ТЭ главного слова семантической части является второй семантической частью ОЛЛЭ 17, и так далее. Порядковый номер ИЛЛЭ указывает в наименовании ОЛЛЭ, например, но не ограничиваясь - «ИЛЛЭ 1.2.1». Такое наименование элемента 18 говорит о том, что он является первой семантической частью во втором ОЛЛЭ первого ОЛЛО. Формирование элементов 18 тринадцатой структуры данных СМД в ходе этапа 4033 производят на основе результатов этапа 4032 формирования семантических частей элементов 17 двенадцатой структуры данных СМД следующим образом: на первом этапе все значимые части речи в семантических частях элементов 17 отождествляют с элементами 18 тринадцатой структуры данных СМД; на втором этапе упомянутые идентифицированные главные слова семантических частей элементов 17 отождествляют с элементами 18 первого вида; на третьем этапе все остальные элементы 18 данной семантической части, кроме элемента 18 первого вида отождествляют с элементами 18 второго вида.

[182] Идентификация значения и порядкового (порядковых) номера (номеров) ТЭ элемента 18 двенадцатой структуры данных СМД, производится в рамках этапа 4033 путем отождествления значения 181 элемента 18 со значением (значениями) 211 текстового (текстовых) элемента (элементов) 21 лингвистического предложения 11, составляющего (составляющих) элемент 18, а порядковый (порядковые) номер (номера) 182 текстового (текстовых) элемента (элементов) 21, составляющего (составляющих) элемент 18 путем отождествления с порядковым (порядковыми) номером (номерами) 212 текстового (текстовых) элемента (элементов) 21, составляющего (составляющих) элемент 18. Идентификация видов элементов 18 двенадцатой структуры данных СМД в рамках этапа 4031 не изменяет значение 181 элемента 18 и не изменяет порядковый (порядковые) номер (номера) ТЭ 21 лингвистического предложения 11, составляющего (составляющих) элемент 18 (ИЛЛЭ).

[183] Формирование тринадцатой структуры данных СМД в ходе этапа 4033 производят путем объединения в одной структуре данных элементов 18 первого и второго вида тринадцатой структуры данных СМД, а также их идентификационных данных по известным из уровня техники принципам и способам, которые, соответственно, далее подробно не описываются.

[184] Идентификацию лингвистических характеристик ТЭ, составляющих элементы 18 (ИЛЛЭ) тринадцатой структуры данных СМД и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа 4031, состоящего из идентификационных данных ТЭ, составляющих элемент 18 (ИЛЛЭ), и получения значений (2131, 2141, 2231; 2151, 2161, 2181, 2191) всех частей лингвистических характеристик (213, 214, 223, 215, 216, 218, 219) текстовых элементов 21 лингвистического предложения 11, из которых состоит элемент 18 (ИЛЛЭ). При этом, как было описано ранее, лингвистическими признаками элементов 18 (ИЛЛЭ) являются как минимум морфологические, синтаксические и семантические характеристики текстовых элементов 21 лингвистического предложения 11, из которых состоят элементы 18 (ИЛЛЭ).

На фиг. 45, в качестве примера, но не ограничения, проиллюстрирована примерная схема системы 500 преобразования структурированного массива данных, которая в предпочтительном варианте реализации содержит, по меньшей мере, одно или более компьютерных устройств 501 преобразования структурированного массива данных, содержащих, по меньшей мере, один или более процессоров 5011 и память 5012.

Упомянутые устройства 501 преобразования структурированного массива данных могут представлять собой, но не ограничиваться: персональный компьютер, портативный компьютер, планшетный компьютер, карманный компьютер, смартфон, фаблет и тому подобное. Память (машиночитаемый носитель данных) 5012 устройства 501 преобразования структурированного массива данных, содержит код программы, который при выполнении побуждает упомянутые один или более процессоры 5011 упомянутого устройства 501 выполнять действия описанных ранее способов преобразования структурированного массива данных. В некоторых случаях компьютерное устройство 501 может представлять собой серверное компьютерное устройство, связанное с пользовательским компьютерным устройством, выполненным с возможностью передачи серверному компьютерному устройству 501 команды или команд, побуждающих процессор или процессоры 5011 серверного компьютерного устройства выполнять код программы, который при выполнении процессором или процессорами серверного компьютерного устройства 5011 побуждает процессор или процессоры 5011 серверного компьютерного устройства выполнять действия какого-либо из описанных ранее способов преобразования структурированного массива данных. Пользовательское компьютерное устройство 502 может представлять собой, но не ограничиваться: персональный компьютер, портативный компьютер, планшетный компьютер, карманный компьютер, смартфон, фаблет, тонкий клиент и тому подобное.

Пользовательское компьютерное устройство 502 может быть связано с серверным компьютерным устройством 501 посредством проводного или беспроводного соединения. Упомянутая память 5012 компьютерного устройства 501 (серверного компьютерного устройства 501) содержит подлежащие преобразованию один или несколько структурированных массивов данных, содержащих, по меньшей мере, лингвистическое предложение, а также может содержать любую из описанных ранее исходных структур данных для какого-либо из описанных ранее способов преобразования структурированного массива данных. Более того, подлежащие преобразованию один или несколько структурированных массивов данных могут

являться загружаемыми и храниться, в частности, в базе данных 503 системы преобразования структурированного массива данных. В качестве примера, но не ограничения, машиночитаемый носитель данных (память 5012) может включать в себя оперативную память (RAM); постоянное запоминающее устройство (ROM);
5 электрически-стираемое программируемое постоянное запоминающее устройство (EEPROM); флэш-память или другие технологии памяти; CDRом, цифровой универсальный диск (DVD) или другие оптические или голографические носители данных; магнитные кассеты, магнитную пленку, запоминающее устройство на магнитных дисках или другие магнитные запоминающие устройства, несущие волны
10 или другой носитель данных, который может быть использован для кодирования требуемой информации, и к которому может быть осуществлен доступ посредством устройства 501. Память включает в себя носитель данных на основе запоминающего устройства компьютера в форме энергозависимой или энергонезависимой памяти, или их комбинации. Примерные аппаратные устройства включают в себя твердотельную
15 память, накопители на жестких дисках, накопители на оптических дисках и так далее. В памяти хранится примерная среда, в которой при помощи компьютерных команд или кодов, хранящихся в памяти устройства, может быть осуществлена процедура преобразования структурированного массива данных. Устройство содержит один или более процессоров 5011, которые предназначены для выполнения компьютерных
20 команд или кодов, хранящихся в памяти устройства с целью обеспечения выполнения процедуры преобразования структурированного массива данных. Компьютерные команды или коды, хранящиеся в памяти, предназначены для выполнения преобразования структурированного массива данных. Система 500 также может включать в себя базу данных (БД) 503. БД 503 может представлять собой, но не
25 ограничиваясь: иерархическую БД, сетевую БД, реляционную БД, объектную БД, объектно-ориентированную БД, объектно-реляционную БД, пространственную БД, комбинацию перечисленных двух и более БД, и тому подобное. БД 503 хранит данные в памяти, которая может представлять собой, но не ограничиваясь: постоянное
30 запоминающее устройство (ROM), электрически-стираемое программируемое постоянное запоминающее устройство (EEPROM), флэш-память, CDRом, цифровой универсальный диск (DVD) или другие оптические или голографические носители данных; магнитные кассеты, магнитную пленку, запоминающее устройство на магнитных дисках или другие магнитные запоминающие устройства, несущие волны
35 или другой носитель данных, который может быть использован для хранения требуемой информации, и к которому может быть осуществлен доступ посредством устройства 501 преобразования структурированного массива данных. БД 503 служит для хранения данных, представляющих собой, по меньшей мере, команды для выполнения этапов описанных ранее способов преобразования структурированного массива данных; подлежащие преобразованию один или несколько структурированных массивов данных,
40 содержащих, по меньшей мере, лингвистическое предложение, или одну из описанных ранее исходных для какого-либо способа преобразования структур данных, которые могут быть загружены в память 5012 устройства 501 преобразования структурированного массива данных; и других данных, необходимых для функционирования системы. Примерная система 500 преобразования
45 структурированного массива данных дополнительно может содержать серверное компьютерное устройство 501, которое помимо описанных ранее функций, сохраняет и содействует манипуляции компьютерными командами или кодами, ранее описанными в данном документе, которые, соответственно, дополнительно не описываются.

Серверное компьютерное устройство 501, помимо описанных ранее функций, может обеспечивать регулирование обменом данных в системе 500 преобразования структурированного массива данных, а также обеспечивает обработку данных при условии подключения к нему одного или более чем одного пользовательских компьютерных устройств 502. В этом случае все вычислительные мощности, необходимые для обеспечения выполнения процедуры преобразования структурированного массива данных, расположены на серверном компьютерном устройстве 501. Система 500 так же может содержать одну или более сетей 504 передачи данных. Сети 204 передачи данных могут включать в себя, но не ограничиваться, одну или более локальных сетей (LAN) и/или глобальных сетей (WAN), или могут представлять собой информационно-телекоммуникационную сеть Интернет, или Интранет, или виртуальную частную сеть (VPN), или их комбинацию, и тому подобное. Серверное компьютерное устройство 501 также имеет возможность обеспечивать виртуальную вычислительную среду (Virtual Machine) для обеспечения взаимодействия между пользовательским компьютерным устройством 502 и БД 503. Сеть 504 служит для обеспечения взаимодействия между компьютерным устройством 501, базой данных 503 и пользовательским компьютерным устройством 502 системы 500 преобразования структурированного массива данных. При этом пользовательское компьютерное устройство 502 может быть связано с серверным компьютерным устройством 501 напрямую, используя известные из уровня техники проводные и беспроводные способы и методы связи, которые, соответственно, далее не подробно не описываются.

(57) Формула изобретения

1. Выполняемый процессором компьютерного устройства способ преобразования структурированного массива данных (СМД), содержащего, по меньшей мере, лингвистическое предложение, характеризующийся выполнением этапов:

А) формирования первой структуры данных, на котором формируют первую структуру данных СМД, содержащую элементы упомянутой первой структуры данных, причем упомянутые элементы первой структуры данных представляют собой текстовые элементы (ТЭ) лингвистического предложения, а также идентификационные данные текстовых элементов (идентификационные данные ТЭ), представляющие собой для каждого ТЭ, по меньшей мере: значение текстового элемента (значение ТЭ) и порядковый номер текстового элемента (порядковый номер ТЭ) в лингвистическом предложении;

Б) формирования базы данных лингвистических признаков (БДЛП), на котором выявляют лингвистические признаки текстовых элементов (лингвистические признаки ТЭ) лингвистического предложения, из которых формируют базу данных, представляющую собой БДЛП текстовых элементов лингвистического предложения;

В) формирования второй структуры данных, на котором формируют вторую структуру данных СМД, содержащую элементы упомянутой второй структуры данных, причем упомянутые элементы второй структуры данных представляют собой синтаксические единицы (СЕ) лингвистического предложения, сформированные на основании сведений из БДЛП текстовых элементов лингвистического предложения, а также представляют собой идентификационные данные синтаксических единиц (идентификационные данные СЕ), представляющие собой для каждой СЕ, по меньшей мере: значение синтаксической единицы (значение СЕ) и порядковый (порядковые) номер (номера) ТЭ лингвистического предложения, составляющих СЕ.

2. Способ по п. 1, характеризующийся тем, что в рамках этапа А) выполняют способ,

характеризующийся выполнением этапов:

А1) идентификации исходной структуры данных СМД, на котором идентифицируют элементы исходной структуры данных СМД, являющиеся лингвистическими предложениями;

5 А2) идентификации элементов первой структуры данных СМД, на котором идентифицируют элементы первой структуры данных СМД, являющиеся текстовыми элементами (ТЭ) лингвистического предложения, а также идентификационные данные ТЭ, представляющие собой для каждого ТЭ, по меньшей мере: значение ТЭ и порядковый номер ТЭ, и формируют первую структуру данных СМД.

10 3. Способ по п. 2, характеризующийся тем, что в рамках этапа А1) обеспечивают классификацию элементов, из которых состоит исходная структура данных, как лингвистических предложений.

4. Способ по п. 2, характеризующийся тем, что в рамках этапа А2) идентифицируют элементы первой структуры данных СМД, являющиеся текстовыми элементами (ТЭ) лингвистического предложения, путем выявления отдельных слов или групп слов, цифр (чисел) или индексов, причем цифры (числа) или индексы не разделены пробелом, а также знаков препинания.

5. Способ по п. 4, характеризующийся тем, что последний знак препинания в лингвистическом предложении не учитывается при идентификации и не рассматривается
20 в качестве текстового элемента лингвистического предложения.

6. Способ по п. 2, характеризующийся тем, что в рамках этапа А2) идентификацию значения текстового элемента производят путем регистрации символов (букв, цифр и (или) знаков препинания), из которых состоит текстовый элемент, а идентификацию порядкового номера текстового элемента производят путем расчета местоположения
25 текстового элемента в лингвистическом предложении.

7. Способ по п. 6, характеризующийся тем, что первый текстовый элемент в лингвистическом предложении получает порядковый номер «1», а все последующие текстовые элементы получают порядковый номер, больший на единицу, чем порядковый номер предыдущего текстового элемента.

30 8. Способ по п. 1, характеризующийся тем, что этап Б) характеризуется выполнением этапов:

Б1) формирования лингвистических признаков текстовых элементов лингвистического предложения, на котором для лингвистического анализа текстового элемента предоставляют идентификационные данные элемента (значение и порядковый номер
35 текстового элемента) и получают лингвистические характеристики текстовых элементов лингвистического предложения, а также значения упомянутых лингвистических характеристик 2131;

Б2) формирования базы данных лингвистических признаков (БДЛП), на котором формируют БДЛП текстовых элементов лингвистического предложения;

40 причем лингвистическим признаком текстового элемента являются все полученные для текстового элемента в ходе выполнения этапа Б1) лингвистические характеристики, обладающие значениями лингвистических характеристик.

9. Способ по п. 8, характеризующийся тем, что в рамках этапа Б1) лингвистические характеристики и их значения формируют путем комплексного лингвистического
45 анализа каждого текстового элемента лингвистического предложения, представляющего, в качестве примера, но не ограничения, комплексный анализ текстового элемента на основе местонахождения текстового элемента в структуре лингвистического предложения, его значения, вида, классификации его мыслимого образа и анализа его

связей с другими текстовыми элементами в лингвистическом предложении, после чего производят формирование основных лингвистических характеристик.

10. Способ по п. 8, характеризующийся тем, что формирование БДЛП осуществляют путем внесения сформированных в рамках этапа В1) лингвистических характеристик и их значений в БДЛП в виде перечня первой части лингвистических характеристик со значениями этих характеристик.

11. Способ по п. 1, характеризующийся тем, что этап В) характеризуется выполнением этапов:

В1) идентификации и формирования первых элементов второй структуры данных СМД, на котором идентифицируют и формируют первые элементы второй структуры данных СМД, являющиеся первыми синтаксическими единицами, а также идентификационные данные первых синтаксических единиц, представляющие собой для каждой первой синтаксической единицы, по меньшей мере, значение первой синтаксической единицы и порядковый (порядковые) номер (номера) текстового элемента или текстовых элементов лингвистического предложения, составляющих первую синтаксическую единицу;

В2) формирования вторых элементов второй структуры данных СМД, на котором формируют вторые элементы второй структуры данных СМД, являющиеся вторыми синтаксическими единицами, а также идентификационные данные вторых синтаксических единиц, представляющие собой для каждой такой синтаксической единицы, по меньшей мере, значение второй синтаксической единицы и порядковые номера текстовых элементов лингвистического предложения, составляющих вторую синтаксическую единицу;

В3) формирования второй структуры данных СМД, на котором формируют вторую структуру данных СМД путем объединения первых синтаксических единиц, вторых синтаксических единиц и их идентификационных данных.

12. Способ по п. 11, характеризующийся тем, что в рамках этапа В1) идентификацию или формирование первых синтаксических единиц производят путем комплексного анализа значений лингвистических характеристик текстовых элементов лингвистического предложения, причем идентификации подвергаются первые синтаксические единицы, состоящие из одного текстового элемента, а формированию подвергаются первые синтаксические единицы, состоящие из нескольких текстовых элементов, при этом идентификация или формирование осуществляется путем отождествления первых синтаксических единиц с актуальными исходными синтаксическими объектами, представляющими собой перечень исходных синтаксических объектов, предварительно устанавливаемый и записываемый в первую пользовательскую базу данных.

13. Способ по п. 11, характеризующийся тем, что в рамках этапа В2) формирование вторых синтаксических единиц из двух и более первых синтаксических единиц производят путем комплексного анализа значений лингвистических характеристик текстовых элементов, являющихся первыми синтаксическими единицами, при выявлении лингвистических характеристик текстовых элементов, составляющих первые синтаксические единицы, которые соответствуют описанию синтаксической конструкции, содержащемуся в первой пользовательской базе данных.

14. Способ по п. 11, характеризующийся тем, что в рамках этапа В1) идентификацию значения и порядкового (порядковых) номера (номеров) ТЭ первых синтаксических единиц производят следующим образом: значение (значения) текстового (текстовых) элемента (элементов) лингвистического предложения, составляющего (составляющих) первую синтаксическую единицу, идентифицируют как значение элемента первой

синтаксической единицы, а порядковый (порядковые) номер (номера) упомянутого (упомянутых) текстового (текстовых) элемента (элементов) лингвистического предложения, составляющего (составляющих) первую синтаксическую единицу, идентифицируют как порядковый (порядковые) номер (номера) текстового (текстовых) элемента (элементов) 312, составляющего (составляющих) первую синтаксическую единицу.

15. Способ по п. 11, характеризующийся тем, что в рамках этапа В2) идентификацию значения и порядковых номеров ТЭ вторых синтаксических единиц производят следующим образом: значения текстовых элементов лингвистического предложения, составляющих вторую синтаксическую единицу, идентифицируют как значение второй синтаксической единицы, а порядковые номера упомянутых текстовых элементов лингвистического предложения, составляющих вторую синтаксическую единицу, идентифицируют как порядковые номера текстовых элементов, составляющих вторую синтаксическую единицу.

16. Способ по п. 11, характеризующийся тем, что в рамках этапа В3) формирование второй структуры данных СМД производят путем объединения в одной структуре данных первых синтаксических единиц и вторых синтаксических единиц.

17. Способ по п. 12, характеризующийся тем, что в рамках этапа В1) идентификацию лингвистических характеристик ТЭ, составляющих первые синтаксические единицы, и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа Б), состоящего из идентификационных данных ТЭ, составляющих первую синтаксическую единицу, и получения значений лингвистических характеристик текстовых элементов лингвистического предложения, из которых состоит первая синтаксическая единица.

18. Способ по п. 13, характеризующийся тем, что в рамках этапа В2) идентификацию лингвистических характеристик ТЭ, составляющих вторые синтаксические единицы, и их значений при необходимости производят путем организации запроса в БДЛП, формируемую в рамках этапа Б), состоящего из идентификационных данных ТЭ, составляющих вторую синтаксическую единицу, и получения значений лингвистических характеристик текстовых элементов лингвистического предложения, из которых состоит вторая синтаксическая единица.

19. Компьютерное устройство для преобразования структурированного массива данных, содержащего, по меньшей мере, лингвистическое предложение, содержащее, по меньшей мере:

процессор компьютерного устройства, и память, содержащую код программы, который при выполнении процессором компьютерного устройства побуждает процессор компьютерного устройства выполнять действия способа по любому из пп. 1-18.

20. Система для преобразования структурированного массива данных, содержащего, по меньшей мере, лингвистическое предложение, содержащая, по меньшей мере: серверное компьютерное устройство, являющееся компьютерным устройством по п. 19, и

пользовательское компьютерное устройство, выполненное с возможностью передачи серверному компьютерному устройству команды или команд, побуждающих процессор серверного компьютерного устройства выполнять код программы, который при выполнении процессором серверного компьютерного устройства побуждает процессор серверного компьютерного устройства выполнять действия способа по любому из пп. 1-18.

21. Машиночитаемый носитель данных, содержащий код программы, который при выполнении процессором компьютерного устройства побуждает процессор компьютерного устройства выполнять действия способа по любому из пп. 1-18.

5

10

15

20

25

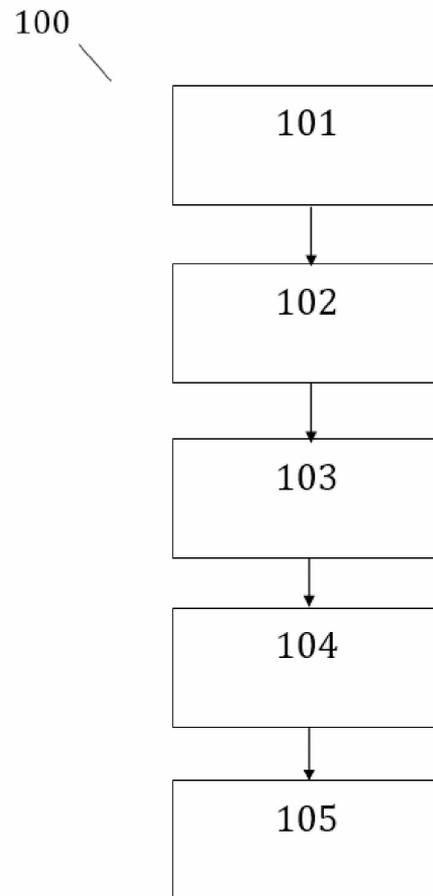
30

35

40

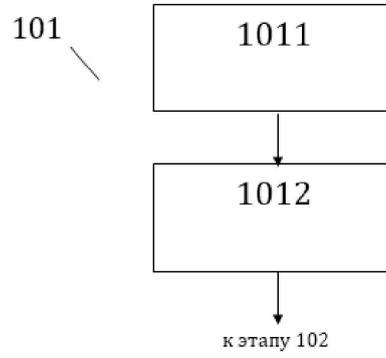
45

1

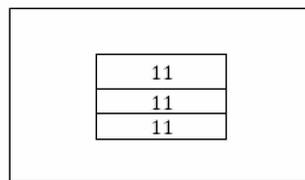


Фиг. 1

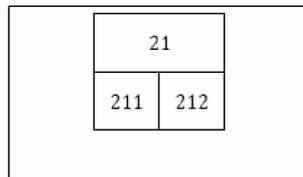
2



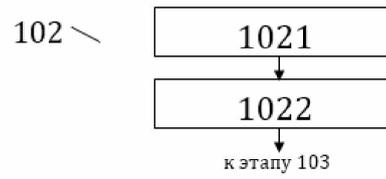
Фиг. 2



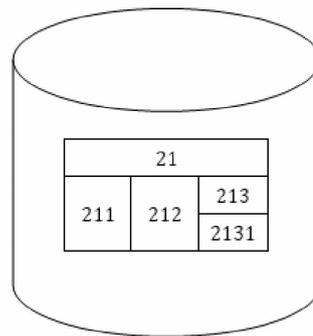
Фиг. 3



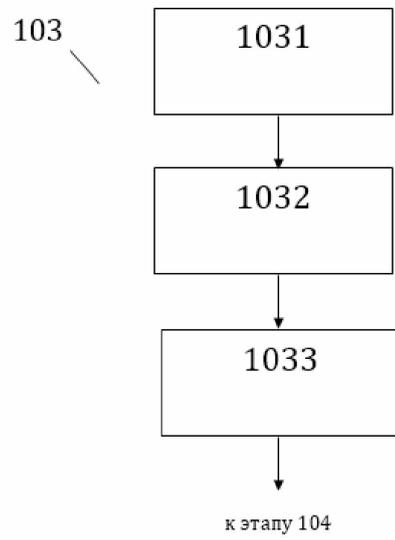
Фиг. 4



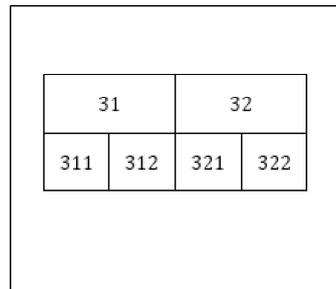
Фиг. 5



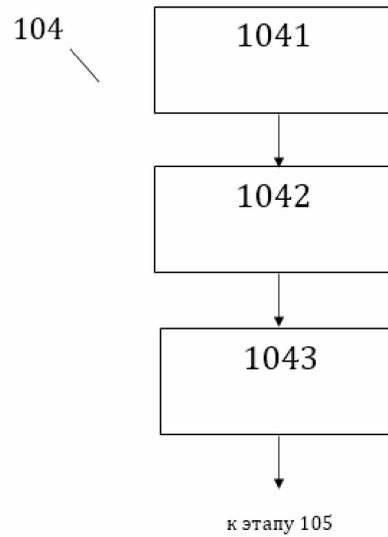
Фиг. 6



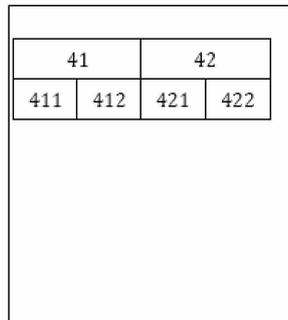
Фиг. 7



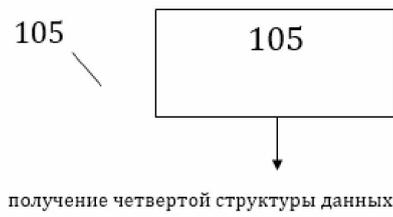
Фиг. 8



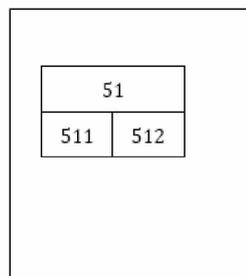
Фиг. 9



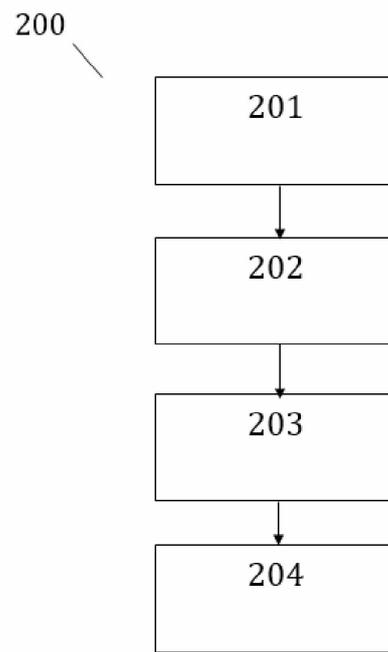
Фиг. 10



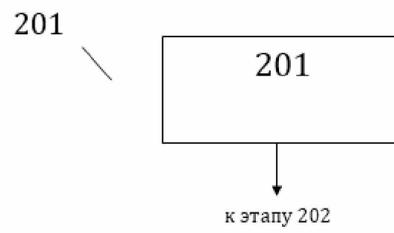
Фиг. 11



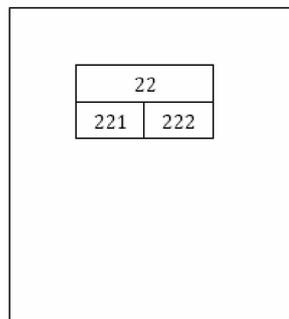
Фиг. 12



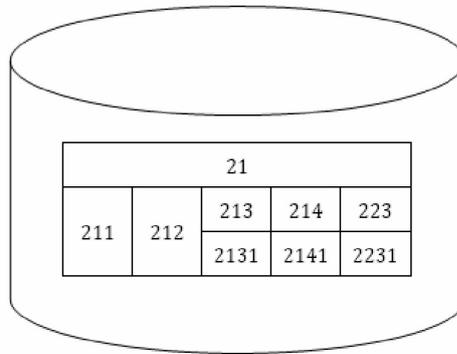
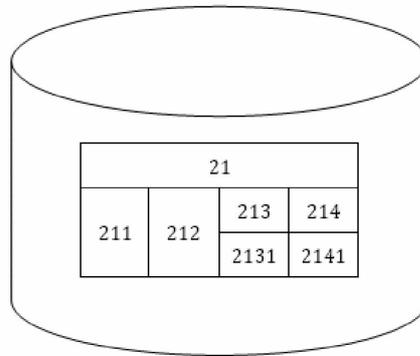
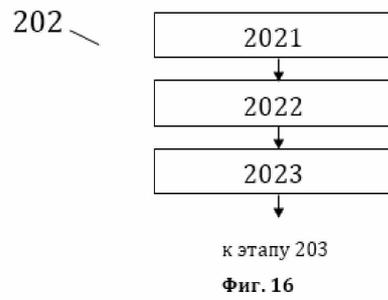
Фиг. 13

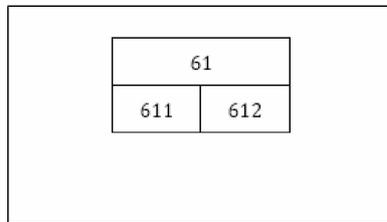


Фиг. 14

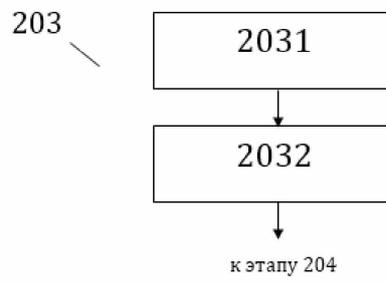


Фиг. 15

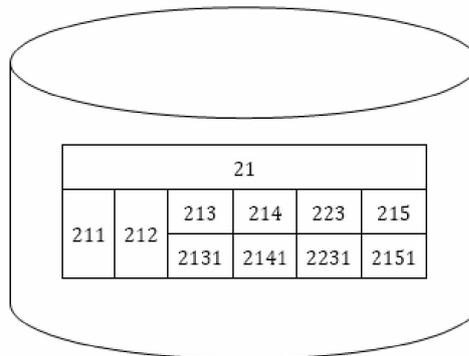




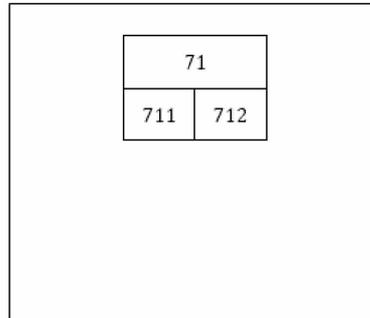
Фиг. 19



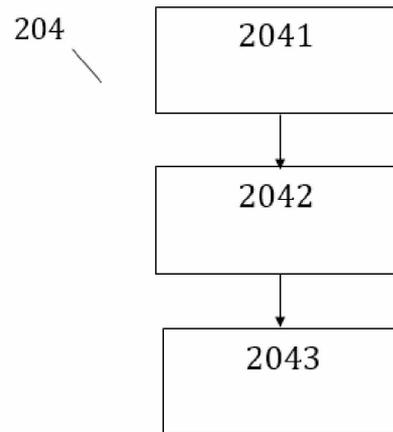
Фиг. 20



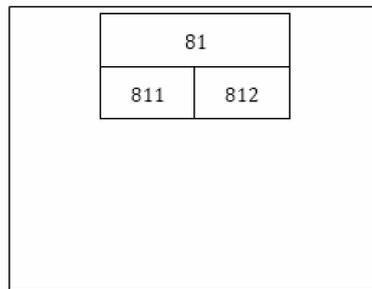
Фиг. 21



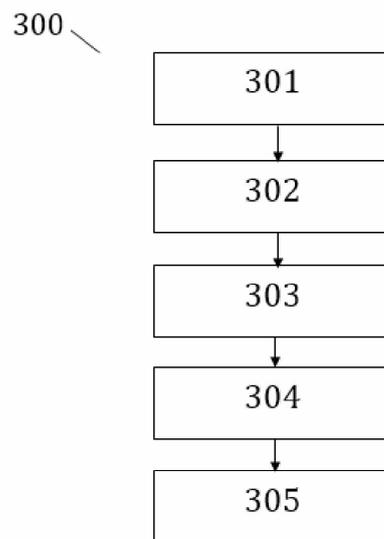
Фиг. 22



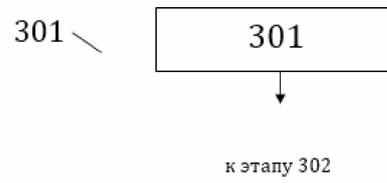
Фиг. 23



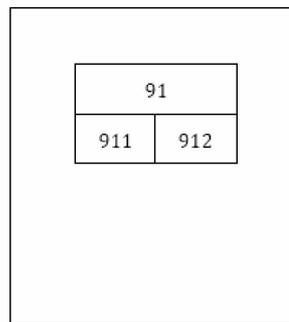
Фиг. 24



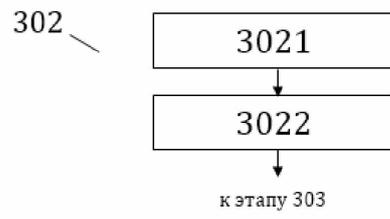
Фиг. 25



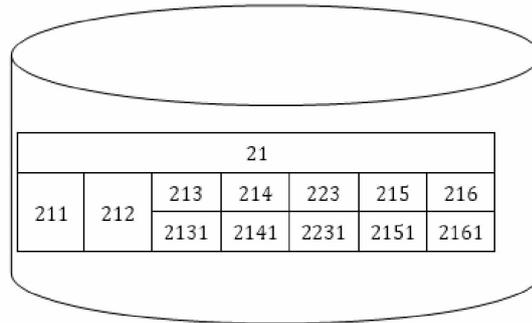
Фиг. 26



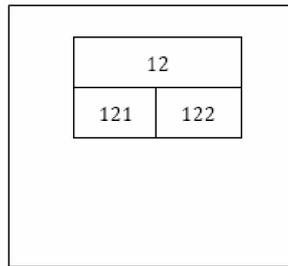
Фиг. 27



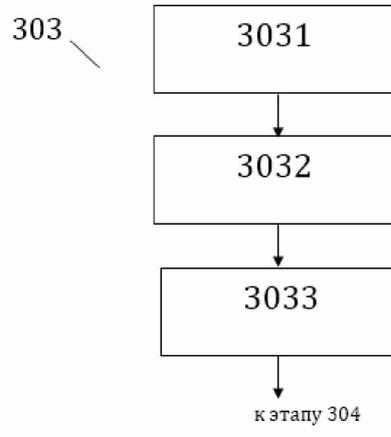
Фиг. 28



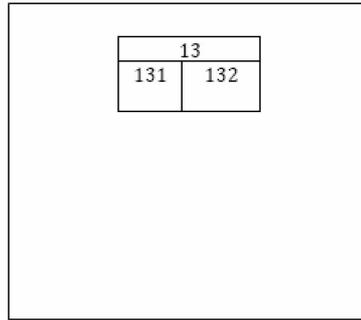
Фиг. 29



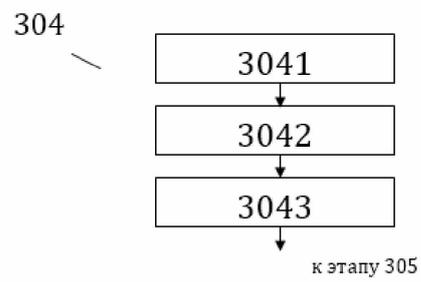
Фиг.30



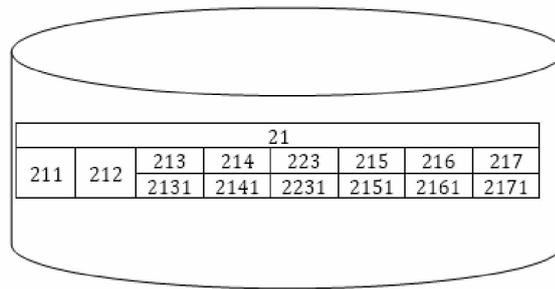
Фиг. 31



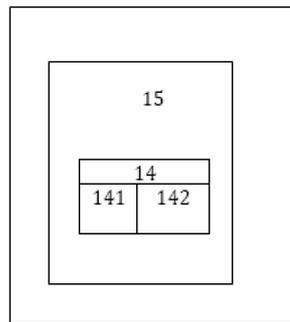
Фиг. 32



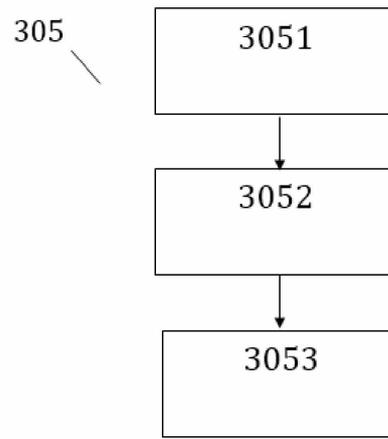
Фиг. 33



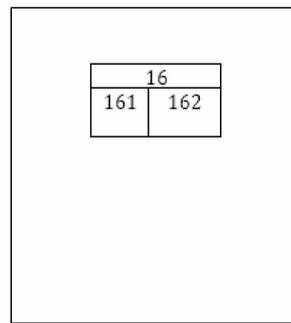
Фиг. 34



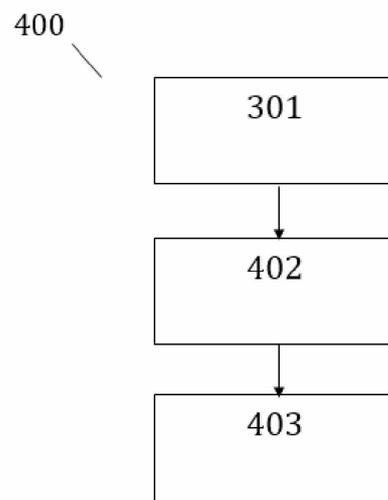
Фиг. 35



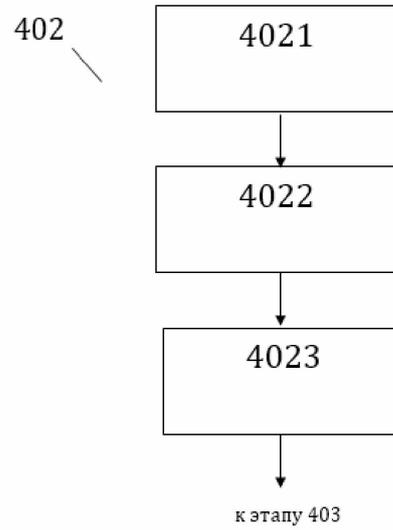
Фиг. 36



Фиг. 37



Фиг. 38



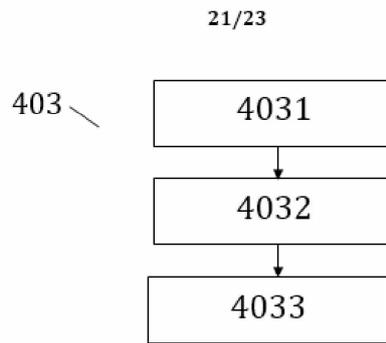
Фиг. 39

21								
212	213	213	214	223	215	216	217	218
		2131	2141	2231	2151	2161	2171	2181

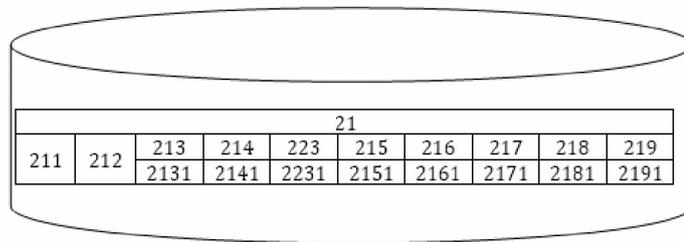
Фиг. 40

17	
171	172

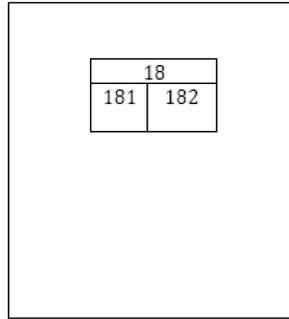
Фиг. 41



Фиг. 42

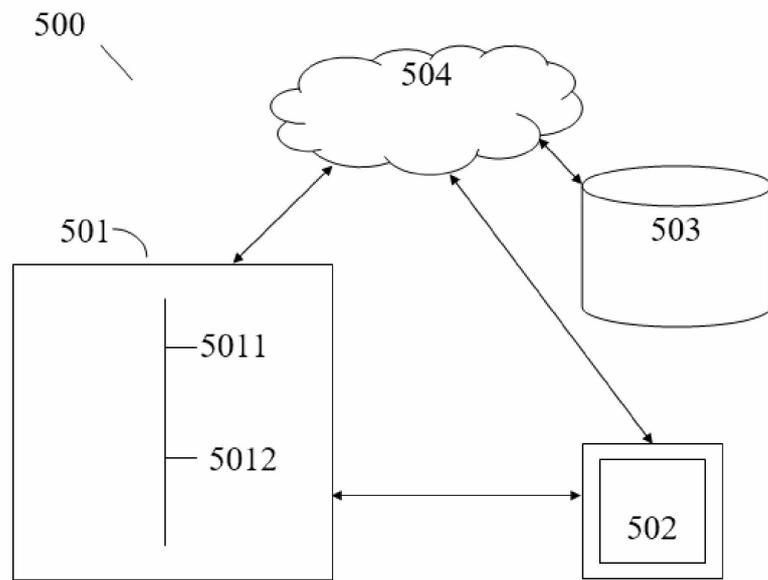


Фиг. 43



18	
181	182

Фиг. 44



Фиг. 45