

Формирование текстового корпуса для автоматического извлечения биографических фактов из русскоязычного текста

А.В. Глазкова

Аннотация—Для решения задач компьютерной лингвистики и машинного обучения, связанных с обработкой естественного языка (natural language processing), исследователи часто пользуются текстовыми корпусами. Текстовые корпуса представляют собой специально собранные коллекции документов, оснащенных текстовой разметкой, содержащей морфологическую, синтаксическую, семантическую или другую необходимую исследователю информацию. Характер информации, представленной в корпусе, как и тип включенных в него текстов, определяются целью и задачами конкретного исследования. Данные, полученные из текстовых корпусов, используются в машинном обучении с учителем в качестве обучающих и тестовых выборок для построения классификаторов текстов, написанных на естественном языке. В данной статье представлен инструмент для построения корпуса биографических текстов на русском языке. Процесс построения корпуса включает в себя две стадии: сбор текстов и их разметку. В первую очередь мы отобрали тексты, подходящие для разметки. В корпус были включены тексты, размещенные в свободном доступе в открытой энциклопедии «Википедия». Тексты были собраны при помощи автоматического парсера, разработанного с использованием свободно распространяемых библиотек для языка программирования Python. Вторая стадия подразумевала семантическую разметку предложений текста, выделение в тексте биографических фактов и отнесение их к одному из заранее заданных типов. Семантическая разметка выполнялась в полуавтоматическом режиме. В работе описаны особенности создания корпуса биографических текстов, принятая в исследовании таксономия биографических фактов, программная реализация инструмента для сбора и разметки текстов, представление текстов в корпусе, а также характеристики созданного корпуса.

Ключевые слова— Инструменты обработки данных, корпусная лингвистика, извлечение фактов, обработка естественного языка, Википедия.

I. ВВЕДЕНИЕ

Решение задач, связанных с извлечением информации из текстов, часто требует наличия специально подготовленных текстовых коллекций, собранных,

обработанных и размеченных в соответствии со спецификой решаемой задачи. Такие коллекции называются текстовыми корпусами [1].

Текстовые корпуса широко используются в компьютерной лингвистике. Как правило, они снабжаются текстовой разметкой, представляющей собой специальные метки, содержащие синтаксическую, семантическую или иную информацию. Тип информации, содержащейся в текстовой разметке, обусловлен задачами конкретного исследования. Текстовая разметка дает возможность применять технологии информационного поиска к текстам, составляющим корпус, обучать модели на данных корпуса и оценивать качество методов обработки текстов. Чтобы принести максимальную пользу научному сообществу, текстовый корпус должен быть представлен в виде, удобном для компьютерной обработки. Он должен быть доступен не только для просмотра через пользовательский интерфейс, но и для загрузки на персональный компьютер и последующей обработки [2-3].

Построение полностью размеченного текстового корпуса представляет собой довольно сложный процесс, требующий усилий многих людей. По этой причине созданием крупных текстовых корпусов занимаются обычно исследовательские группы в специализированных институтах. Однажды созданный, такой корпус может быть использован многими учеными для решения различных прикладных и научных задач [4-5]. Пути использования готового текстового корпуса могут быть самыми разными, в том числе такими, которые не рассматривались специально его создателями. Несмотря на широкую применимость универсальных текстовых корпусов, исследователи часто сталкиваются с необходимостью создания более специальных коллекций для решения каких-либо частных задач [6-12]. Так, решение проблемы извлечения биографических фактов требует построения специального корпуса биографических текстов, размеченных в соответствии с предложенной исследователями таксономией биографических фактов. В настоящий момент для русского языка нет единого корпуса биографических текстов.

Данная статья демонстрирует текущие результаты в разработке методики и инструмента для формирования текстового корпуса для автоматического извлечения биографических текстов из естественно-языковых

Статья получена 30 декабря 2018. Рекомендована организационным комитетом III Международной научной конференции «Конвергентные когнитивно-информационные технологии».

Глазкова А.В. - ФГАОУ ВО «Тюменский государственный университет» (email: anya_kr@aol.com).

текстов, написанных на русском языке. Корпус должен содержать биографические тексты, посвященные различным личностям, которые занимались различными видами основной деятельности: наукой, культурой, спортом, политикой, предпринимательством и так далее. Тексты, включенные в корпус, должны иметь специфическую текстовую разметку. Необходимо выделить предложения, содержащие биографические факты и определить тип факта, присутствующего в данном предложении.

В настоящее время существуют несколько текстовых корпусов для русского языка, находящихся в свободном доступе. В рамках проекта «Open Corpora» [13] создается крупный текстовый корпус для решения общих исследовательских задач. Несколько недавних работ посвящены разработке русскоязычных текстовых корпусов для решения более узких задач, связанных с обработкой естественного языка. Так, в статье [14] представлены инструмент для сбора и анализа коротких текстов и корпус коротких текстов, собранных на основе сообщений в Twitter. Работы [15-16] посвящены созданию корпуса диалектных текстов. В статье [17] представлен электронный корпус диалектных текстов.

II. ТИПИЗАЦИЯ БИОГРАФИЧЕСКИХ ФАКТОВ

а. Таксономия биографических текстов

Структура биографических текстов может быть различной. Как правило, биография – это не несвязное множество биографических фактов, а осмысленный текст. В работе [18] приводятся различные модели написания биографических текстов:

- 1) хронологическая модель (описание ключевых жизненных событий от рождения до смерти человека);
- 2) функциональная модель (описание жизни, основной акцент в котором сделан на профессиональной деятельности);
- 3) психологическая модель (убеждения личности, страхи, приоритеты);
- 4) социологическая модель (социальные обстоятельства, определившие ход жизни человека);
- 5) культурологическая модель (описание личности в социокультурном контексте).

Говоря о биографических текстах, мы в своем исследовании имеем в виду в первую очередь тексты, относящиеся к хронологической модели. Для создания корпуса биографических текстов мы использовали биографические статьи, размещенные в свободном доступе в онлайн-энциклопедии «Википедия» [19]. Эти тексты описывают основные личные и социально значимые события в жизни персоналий. Описание событий в текстах располагается в большинстве случаев в хронологическом порядке, часто с упоминанием дат, мест действия и связанных с событием личностей.

б. Таксономия биографических фактов

Биография может содержать различную информацию о личной и профессиональной жизни человека. Также в биографическом тексте может присутствовать информация, не относящаяся к биографической. Таким образом, текст представляет собой набор предложений,

где каждое предложение можно представить в виде:

$$S = (B, N),$$

где B – множество биографических фактов, N – множество фактов, не относящихся к биографическим. Под фактом в данном контексте подразумевается любое утверждение (в лингвистическом смысле), которое в прямом виде содержится в тексте.

В нашей работе мы используем следующую таксономию фактов, основанную на типизации, принятой в работе [20]:

- 1) не биографический факт;
- 2) события:
 - i) личные события;
 - ii) профессиональные события (встречи, награды и так далее);
- 3) неизменяемые личные характеристики:
 - i) рождение (место и дата рождения);
 - ii) смерть (место и дата смерти);
 - iii) национальность;
 - iv) информация о родителях;
- 4) изменяемые личные характеристики:
 - i) место работы (службы);
 - ii) образование;
 - iii) семья (женитьба, замужество, дети и т.д.);
 - iv) место жительства (пребывания);
 - v) род занятий (должность);
- 5) прочие биографические факты.

Таким образом, биографические факты могут быть разделены на две категории – личная и профессиональная информация – и представлены в виде иерархической структуры (рисунок 1).



Рис. 1. Таксономия биографических фактов

III. ПОСТРОЕНИЕ ТЕКСТОВОГО КОРПУСА

а. Описание текстов

Для включения в корпус мы собрали 200 текстов биографических статей, размещенных в свободной энциклопедии «Википедия». Тексты содержат биографии личностей, живших или живущих в 20-21 веках, чья основная деятельность связана с одним из следующих направлений:

- 1) военные и работники силовых структур;
- 2) деятели культуры и искусства;
- 3) деятели науки, техники и образования;
- 4) политики и общественные деятели;
- 5) предприниматели и менеджеры;

б) религиозные деятели.

Количественные характеристики текстовой коллекции представлены в таблице 1.

Таблица 1. Количественные характеристики текстовой коллекции

Характеристика	Значение
Среднее количество символов в текстах	1704
Среднее количество слов в текстах	225
Среднее количество предложений в текстах	19

В корпусе собранная коллекция была размечена в соответствии с таксономией биографических фактов, представленной выше.

б. Программная реализация

Процесс создания текстового корпуса включает в себя два этапа: сбор коллекции текстов и ее разметка (рисунок 2).



Рис. 2. Основные этапы создания текстового корпуса

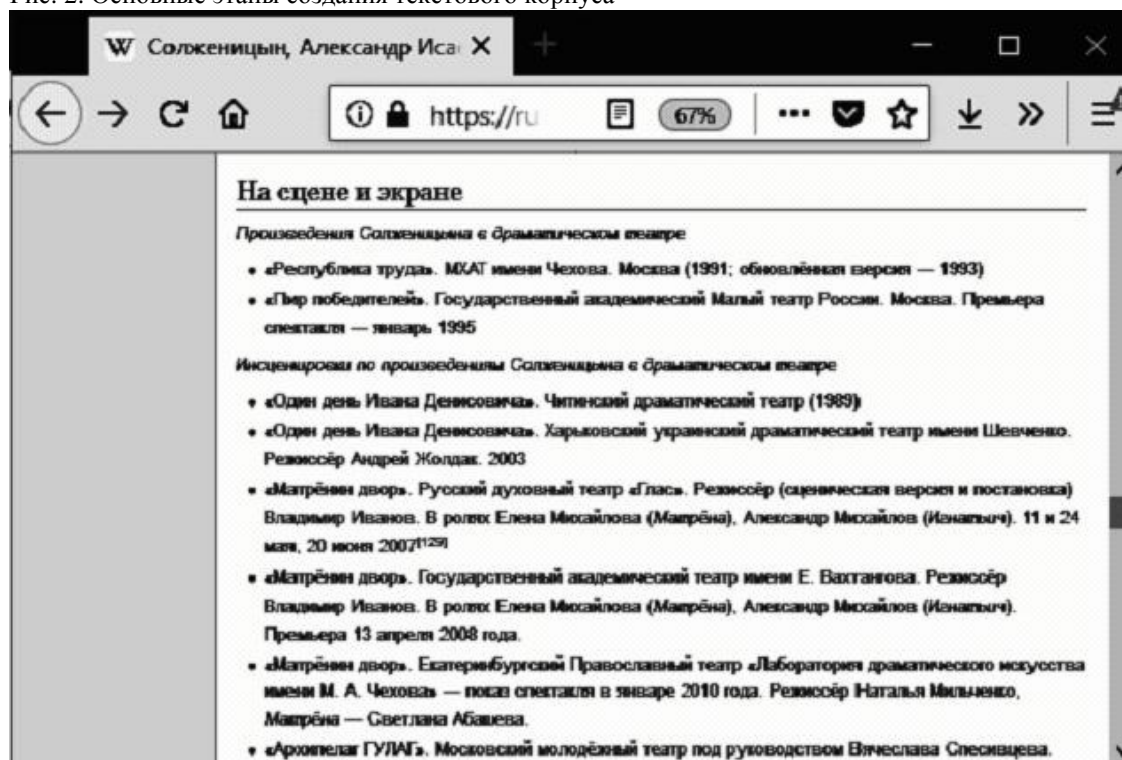


Рис. 3. Пример фрагмента биографической статьи, исключенного во время предобработки

Для проведения разметки тексты были разбиты на предложения. Каждое предложение было отнесено к одному из классов в соответствии с принятой таксономией фактов. Текстовая разметка проводилась в

Сбор текстов осуществлялся в автоматическом режиме с помощью открытых библиотек языка Python [21, 22]. В первую очередь мы получили тексты размещенных в «Википедии» статей из нескольких разноплановых категорий, например, «Кавалеры Ордена Андрея Первозванного», «Ученые России», «Художники-абстракционисты России» и т.д. Далее тексты были сохранены в формате .txt.

Предобработка и выбор текстов для корпуса проводились следующим образом:

1) были удалены короткие тексты, содержащие только годы жизни человека и список его должностей;

2) в коллекцию текстов, пригодных для разметки, были включены только биографии людей, живших не раньше 20 века;

3) из текстов были удалены все разделы, кроме раздела «Биография». Это связано с тем, что биографические статьи в «Википедии» содержат, как правило, перечни наград, научных трудов, произведений и прочие разделы, неудобные для проведения разметки (рисунок 3).

полуавтоматическом режиме. Каждое предложение может быть отнесено к одному или двум типам фактом. Инструмент для текстовой разметки реализован как графическое приложение .NET Framework на языке C# 6.0 [23].

с. Представление текстов

Размеченный корпус биографических текстов состоит

из следующих элементов.

1. Тексты, представленные в формате .xml. Пример размеченного текста представлен на рисунке 4.

2. Файл описания корпуса в формате .csv, в котором содержится основная информация о текстах. Структура файла описана в таблице 2.

Таблица 2. Структура файла описания

Столбец	Назначение
Person	Имя личности, которой посвящен текст
Path	Путь к файлу, содержащему размеченный текст
Birth	Год рождения
Death	Год смерти
Occupation	Направление основной деятельности (см. п. «Описание текстов»)

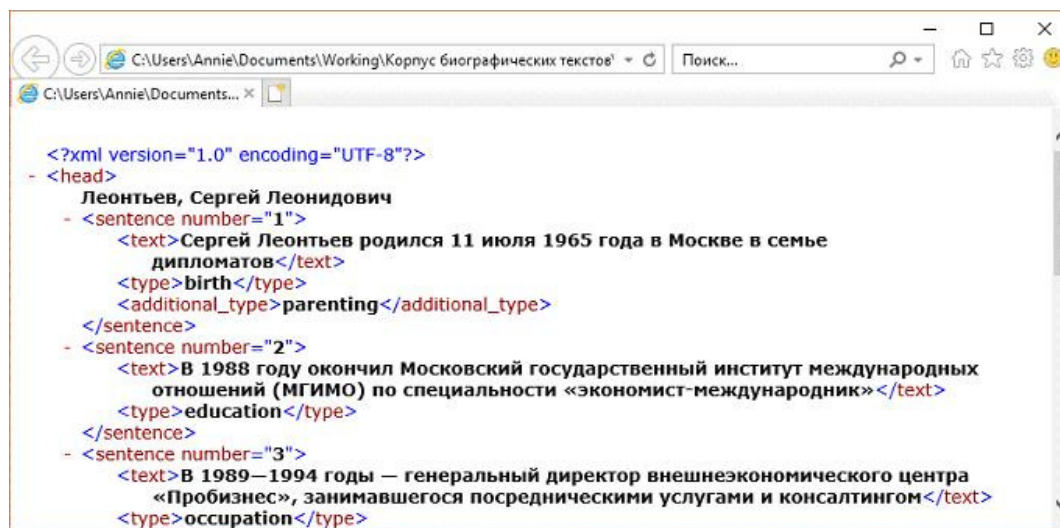


Рис. 4. Пример размеченного текста

IV ХАРАКТЕРИСТИКИ КОРПУСА

Чтобы оценить сбалансированность созданного корпуса биографических текстов, было рассчитано процентное соотношение типов фактов среди текстов всего корпуса и текстов трех крупнейших тематических групп: «Деятели науки, техники и образования», «Деятели культуры и искусства» и «Военные и работники силовых структур». Если процентное соотношение типов фактов в данных категориях и во всем корпусе является примерно одинаковым, можно предположить, что преобладание того или иного типа фактов не зависит от тематики текстов. В таком случае можно говорить о сбалансированности корпуса в том смысле, что наличие слов, специфических для текстов определенной тематики, не окажет существенного влияния на обучение классификатора биографических

фактов на основе данных корпуса. Это говорит о возможности применения методов машинного обучения и автоматической обработки текстов к созданному корпусу.

Результаты проверки сбалансированности корпуса представлены в таблице 3. Отклонение было рассчитано как разница между процентными показателями между текстами всего корпуса и текстами конкретной категории. Поскольку отклонения для всех типов фактов и категорий не превышают 5%, может быть сделан вывод о том, что в целом тематическая группа текста не определяет наличие или отсутствие в нем биографических фактов определенного типа.

Общие данные о структуре текстового корпуса проиллюстрированы на рисунке 5.

Таблица 3. Процентное соотношение типов биографических фактов

Тип биографического факта	Соотношение типов фактов (%)				Отклонение (%)		
	Весь корпус	«Деятели науки, техники и образования»	«Деятели культуры и искусства»	«Военные и работники силовых структур»	«Деятели науки, техники и образования»	«Деятели культуры и искусства»	«Военные и работники силовых структур»
Рождение	5,23	5,75	5,00	4,52	-0,52	0,23	0,71
Смерть	4,17	3,17	4,95	7,34	1,00	-0,78	-3,17
Информация о	3,99	3,56	4,23	1,45	0,43	-0,24	2,54

родителях							
Семья	2,15	3,25	3,00	0,95	-1,10	-0,85	1,20
Национальность	1,51	1,51	1,23	2,34	0,00	0,28	-0,83
Личные события	6,46	5,78	9,43	7,21	0,68	-2,97	-0,75
Профессиональные события	13,80	15,13	9,45	10,80	-1,33	4,35	3,00
Образование	14,67	15,24	14,33	18,21	-0,57	0,34	-3,54
Место жительства	4,68	7,68	6,35	5,58	-3,00	-1,67	-0,90
Род занятий	34,94	32,13	37,55	35,94	2,81	-2,61	-1,00
Место работы	4,49	3,68	2,13	4,10	0,81	2,36	0,39
Прочее	3,90	3,12	2,35	1,56	0,78	1,55	2,34

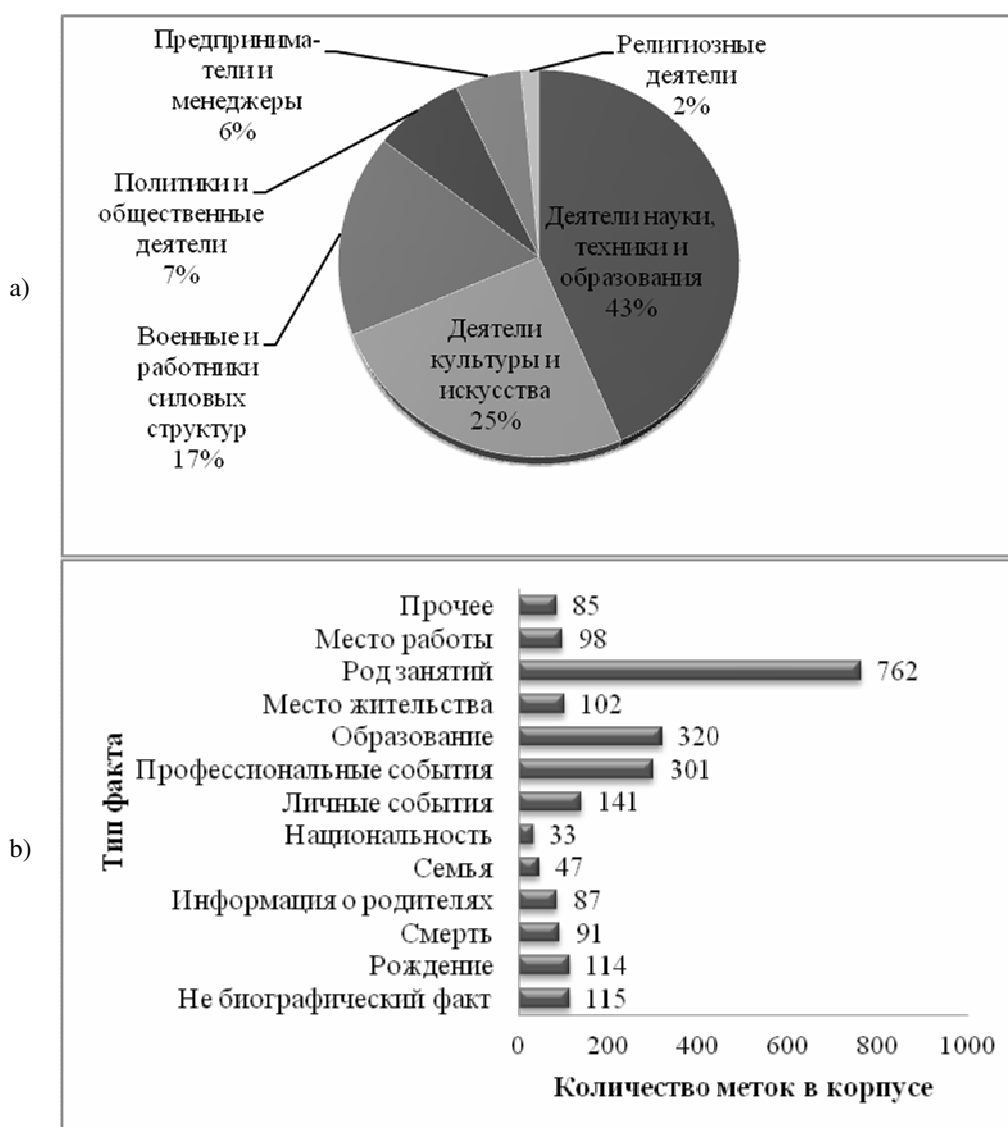


Рис. 5. Структура текстового корпуса: а) тематика текстов; б) представленность типов биографических фактов (каждое предложение может быть соотнесено с 2 типами фактов)

V ЗАКЛЮЧЕНИЕ

В результате работы был создан корпус биографических текстов. В статье описан процесс построения, а также основные количественные характеристики составленного корпуса.

Тексты, входящие в корпус, разделены на предложения, которые размечены в соответствии с таксономией биографических фактов, описанной в

данной статье. Во время работы над корпусом был разработан инструмент для создания и разметки текстовых корпусов. В настоящий момент производится проверка составленного корпуса. После ее завершения корпус будет доступен для скачивания по ссылке [25].

В перспективе планируется решение следующих задач:

1) доработка существующего корпуса (добавление текстов, проверка корпуса);

2) проведение экспериментов по извлечению биографических фактов из текстов на естественном языке, а также построение и обучение классификатора биографических фактов на основе данных корпуса;

3) доработка инструмента разметки текстовых

корпусов.

БЛАГОДАРНОСТИ

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-37-00272 «Автоматизированное извлечение биографических фактов из текстов на естественном языке».

БИБЛИОГРАФИЯ

- [1] Meyers A. Corpus Linguistics for NLP, New York University, URL: <https://cs.nyu.edu/courses/spring18/CSCI-UA.0480-009/lecture7-corpus.pdf> (дата обращения: 14.06.2018).
- [2] Хохлова М. В. Обзор больших русскоязычных корпусов текстов // Компьютерная лингвистика и вычислительные онтологии. Сборник научных статей. Труды XIX Международной объединенной научной конференции. – Санкт-Петербург, 2016. – С. 74-77.
- [3] Khokhlova M. Large Corpora and Frequency Nouns // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2016”. – Moscow, 2016. – С. 224-238.
- [4] Shang J. et al. Automated phrase mining from massive text corpora //IEEE Transactions on Knowledge and Data Engineering. – 2018.
- [5] Roll U., Correia R. A., Berger Tal O. Using machine learning to disentangle homonyms in large text corpora //Conservation Biology. – 2018. – Т. 32. – №. 3. – С. 716-724.
- [6] Campillos L., Deléger L., Grouin C., Hamon T., Ligozat A.-L., Névéal A. A French clinical corpus with comprehensive semantic annotations: development of the Medical Entity and Relation LIMS annotated Text corpus (MERLOT) // Language Resources and Evaluation. – 2018. – Vol. 52(2). – С. 571-601.
- [7] Uhrig P., Evert S., Proisl T. Collocation Candidate Extraction from Dependency-Annotated Corpora: Exploring Differences across Parsers and Dependency Annotation Schemes //Lexical Collocation Analysis. – Springer, Cham, 2018. – С. 111-140.
- [8] Jia C. et al. Concept decompositions for short text clustering by identifying word communities //Pattern Recognition. – 2018. – Т. 76. – С. 691-703.
- [9] Sameen S. et al. Measuring Short Text Reuse for the Urdu Language //IEEE Access. – 2018. – Т. 6. – С. 7412-7421.
- [10] Sojka C., Liška M., Růžička M. Building Corpora of Technical Texts: Approaches and Tools // Fifth Workshop on Recent Advances in Slavonic Natural Languages Processing, RASLAN. – Brno, 2011. – С. 71-82.
- [11] Litvinova T., Zagorovskaya O., Litvinova O. Russian text corpora for deception detection studies // International Journal of Open Information Technologies. – 2017. - Vol. 5, № 11. – С. 58-63.
- [12] Zevakhina N., Dzhakupova S. Russian metalinguistic comparatives: a functional perspective // Working papers by NRU HSE. Series WP BRP "Linguistics". – 2015. – № 39.
- [13] Open Corpora, URL: opencorpora.org. Дата обращения: 14.06.2018.
- [14] Rubtsova Yu. Constructing a corpus for sentiment classification training // Software & Systems. – 2014. – n Vol. 1. – С. 7-78.
- [15] Резанова З. И. Лингвистический корпус "Томский региональный текст": типологически релевантные параметры сбалансированности и репрезентативности // Вестник Томского государственного университета. Филология. – 2015. – Vol. 1(33). – С. 38-50.
- [16] Резанова З. И., Веснина Г. Ю. Подкорпус русской речи билингвов лингвистического корпуса "Томский региональный текст": принципы разметки и метаразметки корпуса // Вопросы лексикографии. – 2016. – Vol. 1(9). – С. 29-39.
- [17] Драчева Ю. Электронный корпус диалектных текстов в аспекте изучения динамики культурных концептов (на примере мультимедийного корпуса вологодских текстов) // Contemporary Современная русская лексикология, лексикография и лингвогеография. – 2014. – С. 114-121.
- [18] Медведева Е.В. Классификация биографий как один из методов биографики в контексте исследований библиотечной отрасли // Вестник Томского государственного университета. Культурология и искусствоведение. – 2016. – Vol. 2(22). – С. 198-205.
- [19] Wikipedia, URL: ru.wikipedia.org (дата обращения: 17.03.2018).
- [20] da Costa dias Soares S.-F. Extraction of Biographical Information from Wikipedia Texts. – Lisbon, 2011.
- [21] Python 3.6.0., URL: <https://www.python.org/downloads/release/python-360/> (дата обращения: 14.06.2018).
- [22] Wikipedia 1.4.0, URL: <https://pypi.org/project/wikipedia/> (дата обращения: 14.06.2018).
- [23] .NET, URL: <https://www.microsoft.com/net4> (дата обращения: 14.06.2018).
- [24] Zakharov V. Evaluation of Internet corpora of Russian // Proceedings of the International Conference “Corpus linguistics-2015”. – St. Petersburg, 2015. – С. 219–229.
- [25] Корпус биографических текстов, URL <https://sites.google.com/site/utcorpus/> (дата обращения: 01.07.2018).

Building a text corpus for automatic biographical facts extraction from Russian texts

A.V. Glazkova

Abstract— The tasks of computer linguistics and machine learning related to natural language processing (NLP) often require the use of text corpora. Text corpora are specially prepared collection of documents equipped with text markup containing morphological, syntactic, semantic or other information. The data received from the text corpora is used in supervised machine learning for building classifiers of texts written in natural language and in other tasks associated with natural language processing and computer linguistics. The specificity of the information presented in the corpus, as well as the type of texts, is determined by the aim and tasks of the particular study. This article presents a tool for building a corpus of biographical texts in Russian. The process of building a text corpus includes two stages: the collection of texts and their markup. At the first stage we collected texts suitable for markup. Thus, we included in the corpus biographical articles placed in Wikipedia in free access. For this purpose, we developed an automatic parser based on open Python libraries. The second stage is the semantic markup of the text sentences and the selection of biographical facts. This stage took place in a semi-automatic mode. The article describes the features of the process of building the corpus of biographical facts, taxonomy of biographical facts using in our work, software implementation for text collecting and markup, text representation in the corpus and the characteristics of the prepared corpus.

Keywords - Data processing tools, corpus linguistics, fact extraction, natural language processing, Wikipedia

REFERENCES

- [1] Meyers A. Corpus Linguistics for NLP, New York University, URL: <https://cs.nyu.edu/courses/spring18/CSCI-UA.0480-009/lecture7-corpus.pdf>. Date of access: 14.06.2018.
- [2] Khokhlova M. A survey of Large Russian Corpora // Computer linguistics and computing ontologies. Proceedings of the XIX International Joint Scientific Conference. – Saint-Petersburg, 2016. – P. 74-77.
- [3] Khokhlova M. Large Corpora and Frequency Nouns // Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialogue 2016”. – Moscow, 2016. – P. 224-238.
- [4] Shang J. et al. Automated phrase mining from massive text corpora //IEEE Transactions on Knowledge and Data Engineering. – 2018.
- [5] Roll U., Correia R. A., Berger Tal O. Using machine learning to disentangle homonyms in large text corpora //Conservation Biology. – 2018. – Vol. 32. – №. 3. – P. 716-724.
- [6] Campillos L., Deléger L., Grouin C., Hamon T., Ligozat A.-L., Névéal A. A French clinical corpus with comprehensive semantic annotations: development of the Medical Entity and Relation LIMS annotated Text corpus (MERLOT) // Language Resources and Evaluation. – 2018. – Vol. 52(2). – P. 571-601.
- [7] Uhrig P., Evert S., Proisl T. Collocation Candidate Extraction from Dependency-Annotated Corpora: Exploring Differences across Parsers and Dependency Annotation Schemes //Lexical Collocation Analysis. – Springer, Cham, 2018. – P. 111-140.
- [8] Jia C. et al. Concept decompositions for short text clustering by identifying word communities //Pattern Recognition. – 2018. – Vol. 76. – P. 691-703.
- [9] Sameen S. et al. Measuring Short Text Reuse for the Urdu Language //IEEE Access. – 2018. – Vol. 6. – P. 7412-7421.
- [10] Sojka P., Liška M., Růžička M. Building Corpora of Technical Texts: Approaches and Tools // Fifth Workshop on Recent Advances in Slavonic Natural Languages Processing, RASLAN. – Brno, 2011. – P. 71-82.
- [11] Litvinova T., Zagorovskaya O., Litvinova O. Russian text corpora for deception detection studies // International Journal of Open Information Technologies. – 2017. - Vol. 5, № 11. – P. 58-63.
- [12] Zevakhina N., Dzhakupova S. Russian metalinguistic comparatives: a functional perspective // Working papers by NRU HSE. Series WP BRP "Linguistics". – 2015. – № 39.
- [13] Open Corpora, URL: opencorpora.org. Date of access: 14.06.2018.
- [14] Rubtsova Yu. Constructing a corpus for sentiment classification training // Software & Systems. – 2014. – n Vol. 1. – P. 7-78.
- [15] Rezanova Z. Linguistic corpus "Tomsk regional text": concept and structure // Tomsk State University Journal of Philology. – 2015. – Vol. 1(33). – P. 38-50.
- [16] Rezanova Z., Vesnina G. Meta-data and annotation design of the Russian-speaking bilinguals speech subcorpus in the structure of the Tomsk Regional Corpus // Voprosy Leksikografii Russian Journal of Lexicography. – 2016. – Vol. 1(9). – P. 29-39. DOI: 10.17223/22274200/9/3.
- [17] Dracheva Yu. Electronic body of dialective texts in the aspect of studying the dynamics of cultural concepts (on the example of the multimedia case of Vologda texts) // Contemporary Russian lexicology, lexicography and linvoeography. – 2014. – P. 114-121.
- [18] Medvedeva E. Classification biographies as one of the biographies research methods in the context of library branch // Tomsk State University Journal of Cultural Studies and Art History. – 2016. – Vol. 2(22). – P. 198-205.
- [19] Wikipedia, URL: ru.wikipedia.org. Date of access: 17.03.2018.
- [20] da Costa dias Soares S.-F. Extraction of Biographical Information from Wikipedia Texts. – Lisbon, 2011.
- [21] Python 3.6.0., URL: <https://www.python.org/downloads/release/python-360/>. Date of access: 14.06.2018.
- [22] Wikipedia 1.4.0, URL: <https://pypi.org/project/wikipedia/>. Date of access: 14.06.2018.
- [23] .NET, URL: <https://www.microsoft.com/net4>. Date of access: 14.06.2018.
- [24] Zakharov V. Evaluation of Internet corpora of Russian // Proceedings of the International Conference “Corpus linguistics-2015”. – St. Petersburg, 2015. – P. 219–229.
- [25] Corpus of biographical texts, URL <https://sites.google.com/site/utrcorpus/>. Date of access: 01.07.2018.