

ИНТЕРАКТИВНАЯ БАЗА ОНОМАСТИЧЕСКИХ ДАННЫХ «ИМЕНА РУССКОЙ РАВНИНЫ»

О.М. Сергеева, науч. сотр. Института русского языка им. В.В. Виноградова РАН, руководитель проекта РФФИ, канд. филол. наук, *etymol@mail.ru*

Рецензент: А.В. Колмогорова, ФГАОУ ВО «Национальный исследовательский университет «Высшая школа экономики», д-р филол. наук, *nastiakol@mail.ru*

Статья представляет собой описание целей и задач создания, основных характеристик пользовательского интерфейса и принципов работы базы ономастических данных (данных об именах собственных) «Имена Русской равнины» [8], созданной в рамках проекта реконструкции древнейшего фонда личных имен Русской равнины в интересах популяризации научных результатов, полученных в ходе осуществления указанного проекта.

Ключевые слова: личные имена собственные, база данных, пользовательский интерфейс, языкознание, этимология, антропонимика, прикладное программное обеспечение.

INTERACTIVE DATABASE OF ONOMASTIC DATA «NAMES OF THE RUSSIAN PLAIN»

O.M. Sergeeva, Researcher, V.V. Vinogradov Russian Language Institute RAS, RFBR project coordinator, Doctor of Philology, *etymol@mail.ru*

The article is a description of the goals and objectives of creation, the main characteristics of the user interface and the principles of operation of the database of onomastic data (data on proper names) «Names of the Russian Plain» [8], created as part of the reconstruction project of the oldest fund of personal names of the Russian Plain in the interests of popularization of scientific results obtained during the implementation of this project.

Keywords: personal proper names, database, user interface, linguistics, etymology, anthroponymy, applied software.

Проблематика исследования

Традиционные (исконные, т. е. не связанные по происхождению с позднейшими конфессиональными и иными заимствованными культурными пластами) личные имена представителей этносов, населявших Русскую равнину на рубеже 1 тыс. до н. э. – 1 тыс. н. э., нуждаются в сводном системном исследовании на фоне неудовлетворенного научного и общественного запроса источников, которые давали бы исчерпывающее представление о накопленном массиве этимологических и историко-культурных данных по антропонимным фондам (фондам личных имен) народов и языков указанного географического ареала. Задача создания единого корпуса таких данных не может сводиться к простой компиляции ранее опубликованных материалов, так как ономастиконы (именники, т. е. совокупные фонды личных имен) различных этносов реконструированы и изучены неравномерно в зависимости от объективной степени сохранности первоисточников сведений о них, многие требуют продолжения исследований (историко-лингвистических/этимологических, историко-культурных, историко-географических) для расширения и углубления представлений об их составе. Сказанное определяет актуальность исследовательского проекта, имевшего своей целью представить в единой организованной форме исторические именники этносов,

населявших Русскую равнину в I тыс. до н. э. — I тыс. н. э.: произвести атрибуцию и систематизацию данных, составляющих антропонимическое наследие указанных этносов, по единым методологическим принципам и отобразить результаты в виде электронной интерактивной базы данных, позволяющей ознакомиться с ними даже неподготовленному пользователю.

Постановка задач исследования

Важнейшей составляющей исторического языкознания, самым тесным образом связанной с историей общества и этнографией, является исследование ономастики, а в его составе — изучение происхождения личных имен собственных, в частности древних (не употребительных в настоящее время, но бывших в употреблении в отдаленные во времени исторические периоды). В то же время личные имена, их происхождение и значения производящих нарицательных основ в их составе — предмет значительного интереса со стороны широкого круга лиц, не являющихся специалистами в историческом языкознании, что естественным образом обусловлено неотъемлемым присутствием проблематики, связанной с личными именами и их присвоением, в жизни любого представителя языкового коллектива. В России отмечается повышенный общественный спрос на сводные (обобщающие) ресурсы ономастической информации. Источники, способные удовлетворить этот запрос, существенно ограничены по охвату материала и/или труднодоступны.

Намерение восполнить этот пробел легло в основу замысла исследовательского проекта «Древнейшие имена собственные, исконные для этнокультурных общностей Русской равнины I тыс. до н. э. — I тыс. н. э.: этимология и системное представление», поддержанного грантом РФФИ № 20-012-00060. Идея проекта принадлежала канд. филол. наук, научному сотруднику ИРЯ им. В.В. Виноградова РАН А.К. Шапошникову, в основу реализации проекта в значительной мере легли собранные им лингвистические данные, предварительное теоретическое осмысление которых отражено в таких его публикациях, как [4–6]. Дальнейший сбор, научный (этимологический) анализ и техническая обработка данных проведены канд. филол. наук О.М. Сергеевой и канд. филол. наук И.А. Горбушиной, а также научными сотрудниками ИРЯ им. В.В. Виноградова РАН. В течение 2020–2022 гг. в рамках данного проекта создан сводный каталог антропонимических данных и реализовано общедоступное представление полученных сведений в электронном виде.

Задачи исследования были сформулированы следующим образом:

1) теоретическое осмысление и постановка научных проблем изучения имен собственных этнического происхождения историко-культурного ареала Русской равнины (ориентировочно между 30 и 60 град. вост. долготы, 45 и 70 град. сев. широты, что соответствует историческому ареалу расселения изучаемых этнических групп);

2) сбор исторического ономастического материала из очерченного региона (выборка имен собственных из эпиграфических и исторических источников, извлечение из этнонимии и проприальной топонимии всех видов);

3) классификация (языковая и культурная атрибуция имен) собранного ономастического материала (мужских и женских имен собственных) по принадлежности к определенным этнокультурным общностям, а именно: киммерийской (VIII–VII вв. до н. э.), скифской (VII в. до н. э. — III в. н. э.), савроматской (VII–II вв. до н. э.), сармато-туранской (аланской, ясской) (II в. до н. э. — XV в. н. э.), эллинской (VI в. до н. э. — VI в. н. э.), римской (I в. до н. э. — III в. н. э.), анто-славянской (II–VII вв.) и восточно-славянской (VIII–X вв.), готской (вост.-германской) (III–X вв.) и варяжской (сев.-германской/др.-скандинавской) (IX–X вв.), угорской (древненевенгерской) (V–IX вв.) и финской (IX–X вв.), хазарско-печенежской (древнетюркской) (VI–X вв.), черкесской (абазинско-адыгской) (VIII–X вв.);

4) фонологическое, морфологическое и лексико-семантическое описание ономастического материала;

5) этимология имен собственных разных этнокультурных общностей на основе сопоставления с ближайшими родственными формами известных языковых групп и семей;

6) определение социумов – создателей и распространителей традиционных имен собственных;

7) выявление преемственности культур и ономастических констант в очерченном историко-культурном ареале;

8) представление полученных результатов в виде интерактивной базы данных для удаленного наполнения (редактирования) и просмотра широким кругом пользователей.

Составленный в рамках проекта компендиум личных имен сформирован посредством анализа и категоризации ономастических данных, подтверждением и по необходимости уточнением этноязыковой атрибуции систематизируемых единиц, в том числе выявлением и по возможности устранением пробелов в знаниях о происхождении тех или иных онимов (имен собственных), что предполагает оригинальные (историко-)этимологические разыскания.

Результаты систематизации данных получили воплощение в форме сплошного сводного электронного каталога антропонимического материала, доступного онлайн [URL: <https://www.antronims.ru> (дата обращения: 03.05.2023)] [8].

Особенности программной реализации интерактивной базы данных

Структура базы ономастических данных определена параметрами, позволяющими отразить следующие сведения, приведенные в таблице.

№	Наименование поля	
1	Имя собственное	Лемма, т. е. сам антропоним, выступающий предметом описания, в той форме, в которой он зафиксирован в источниках (см. п. 7 таблицы).
2	Транслитерация латиницей	Передача антропонима средствами латинской графики, в случае если система письменности языка-первоисточника использует графику, отличную от латинской. Для имен с языком происхождения «праславянский» применяется система транслитерации, идентичная используемой словарным коллективом «Этимологического словаря славянских языков» [7].
3	Транслитерация кириллицей	Передача антропонима средствами кириллической графики. Введена с целью дать пользователям-неспециалистам представление о лингвистически корректном способе адаптации описываемого имени современным русским языком.
4	Пол	Устанавливаемый в соответствии с выявляемыми морфологическими показателями рода предполагаемый пол носителя имени: мужской (М) либо женский (Ж).
5	Язык происхождения	Естественный язык, чьи носители составляли этнокультурную общность, в пределах которой функционировал описываемый антропоним в I тыс. до н. э. – I тыс. н. э.
6	Ономастический ареал	Территория, на которой встречается данный антропоним или ряд антропонимов; географическая локализация антропонима.
7	Источник	Библиографическая отсылка к научному труду или корпусу, послужившему источником включаемых в каталог сведений. Например, для фракийских данных это преимущественно [1], для древненовгородских [2, 3] и т. д., однако источниками могут выступать и исследовательские работы. Расшифровка всех библиографических сокращений присутствует на вкладке «Список литературы», в самом каталоге предусмотрены всплывающие подсказки с отображением полных библиографических данных источника при наведении.

№	Наименование поля	
8	Этимология	Лингвистические данные о происхождении имени, приводимые в соответствии с методологией сравнительно-исторического языкознания, применяемой для реконструкции предьстории единиц морфологического и лексического уровня языковой структуры. Источником выступает весь объем доступной научной литературы по компаративистской и/или этимологической проблематике, цитируемые работы включены в «Список литературы». По необходимости предложены критические уточнения к приводимым гипотезам, в некоторых случаях – оригинальные трактовки с соответствующими обоснованиями.
9	Прочее	Дополнительные сведения об антропониме, существенные для понимания предьстории бытования (практического функционирования) его в качестве личного имени собственного: контекст засвидетельствования, сохранившаяся/подлежащая восстановлению социолингвистическая информация и т. п.

Дополнительно в структуре хранения данных предусмотрены следующие функции: для полей 4–9 реализована функция пометы «Сомнительное», проставляемой в случае недостаточной верифицированности сведений, внесенных в соответствующее поле (если верификация затруднена в силу объективных причин методологического характера), и на практике отображается в форме особого выделения цветом соответствующих полей. Здесь следует оговорить, что самый характер материала с неизбежностью предполагает известную долю неустранимых неточностей и пробелов в данных, связанных с такими факторами, как:

- палеографически обусловленные разночтения при фиксации (записи) имен в используемых источниках;

- объективное отсутствие убедительных этимологических решений, имеющее причиной, как правило, затрудненность идентификации в составе имени формантов, отождествимых с какими-либо известными корневыми и иными морфемами имен нарицательных соответствующих языков.

При просмотре данных базы пользователем поддерживается функция поиска по всем полям. В качестве поискового запроса может быть задан, в том числе, произвольный фрагмент текста (набор символов), что позволяет, к примеру, получить в качестве вывода поиска выборку всех имен, содержащих в себе определенный формант (так, по запросу /без/ будут выведены все имена, в составе которых имеется последовательность /-без-/).

Интерактивная база ономастических данных реализована в наиболее распространенной архитектуре «клиент – сервер», когда серверная часть программного обеспечения «берет на себя» все функции по хранению данных, обработке запросов пользователя и передаче их результатов по Интернету на компьютер-клиент. При таком подходе обеспечивается доступность данных базы самому широкому кругу пользователей, так как от них требуется лишь наличие компьютера или смартфона с установленным Web-браузером любого производителя (см., напр., [9]).

В качестве программного обеспечения, решающего вопросы хранения ономастических данных, использована свободно распространяемая реляционная система управления базами данных MySQL (версия 5.7, релиз 5.7.32 от 19.10.2020 [10]). Выбор СУБД MySQL основан на том, что MySQL является решением для малых и средних приложений [11], к числу которых относятся базы ономастических данных с удаленным доступом. Обычно MySQL используется в качестве сервера, к которому обращаются локальные или удаленные клиенты, однако в дистрибутив входит библиотека внутреннего сервера, позволяющая включать MySQL и в автономные программы.

Гибкость СУБД MySQL обеспечивается поддержкой большого количества типов таблиц: пользователи могут выбрать как таблицы типа MyISAM, поддерживающие полнотекстовый поиск, так и таблицы InnoDB, поддерживающие транзакции на уровне отдельных записей. Более того, СУБД MySQL поставляется со специальным типом таблиц EXAMPLE, демонстрирующим принципы создания новых типов таблиц. Благодаря открытой архитектуре и GPL-лицензированию в СУБД MySQL постоянно появляются новые типы таблиц [11]. В MySQL, релиз 5.7.32 (19.10.2020, общедоступная версия), внесены следующие изменения [10]:

- улучшена проверка привилегий для представлений (ошибка № 31304432).
- исправлены ошибки, обнаруженные при эксплуатации предыдущих релизов, в количестве 23 единицы.

В качестве языка программирования был выбран PHP (релиз 8.0) – С-подобный скриптовый язык [11] общего назначения, интенсивно применяемый для разработки веб-приложений. В настоящее время PHP поддерживается подавляющим большинством хостинг-провайдеров и является одним из лидеров среди языков, применяющихся для создания динамических веб-сайтов [12]. Язык PHP и его интерпретатор (Zend Engine) разрабатываются группой энтузиастов в рамках проекта с открытым кодом.

При программировании прикладного программного обеспечения были использованы JavaScript – библиотека jQuery и Datables, а также объектная модель документа DOM (от англ. Document Object Model).

DOM – это не зависящий от платформы и языка программный интерфейс [13], позволяющий программам и скриптам получить доступ к содержимому HTML-, XHTML- и XML-документов, а также изменять содержимое, структуру и оформление таких документов.

jQuery – небольшая библиотека JavaScript [14], включающая обширный функционал, который позволяет упростить управление DOM, обработку событий, анимацию и реализацию, в том числе и с помощью jquery datatables – плагина для дополнения любой HTML-таблицы элементами управления.

Известную техническую трудность представляла собой поддержка шрифтов, необходимых для корректного отображения введенных данных в соответствии с принятыми для передачи текстов на древних языках Евразии орфографическими конвенциями, отчасти по той причине, что архив данных, легший в основу каталога, по времени составления предшествует появлению новейших unicode-совместимых шрифтов. Для снятия этой трудности предусмотрена страница «Шрифты», с которой можно скачать установочные файлы использованных в архиве шрифтов, например ZRCola, который, несмотря на устаревание, до сих пор широко используется в рабочей практике этимологами-славистами в связи с исключительной полнотой набора символов.

Электронный каталог рассчитан на постепенное и постоянное пополнение новыми данными с возможностью корректировки внесенных сведений, в том числе с учетом вновь публикуемого ономастического материала и этимологических исследований.

Разработка адресуется специалистам-гуманитариям (языковедам, этнографам, историкам), а также широкому кругу пользователей, углубленно интересующихся предысторией фондов личных имен древнейших народов Евразии.

Статья написана при поддержке гранта РФФИ в рамках реализации проекта № 20-012-00060 «Древнейшие имена собственные, исконные для этнокультурных общностей Русской равнины I тыс. до н. э. – I тыс. н. э.: этимология и системное представление».

Список литературы

1. Detschew D. Die thrakischen Sprachreste. Wien: im Kommission bei Rudolf M. Rohrer, 1957. 595 S.
2. Васильев В.Л. Славянские топонимические древности Новгородской земли. М.: Рукописные памятники Древней Руси, 2012. 816 с.

3. Тупи́ков Н.М. Словарь древнерусских личных собственных имен. СПб.: Тип. И.Н. Скороходова, 1903. 863 с.
4. Шапошников А.К. Праславянская этимология ст.-чеш. антропонима Kroměžír в свете древненовгородских антропонимических свидетельств // *Studia etymologica Brunensia*. Brno, 2017. С. 39–42.
5. Шапошников А.К. Древнерусские сложные имена младшей серии на праславянском фоне: этимологический аспект // Ономастика Поволжья: мат-лы 16 Международной научной конференции, посв. 50-летнему юбилею Первой Поволжской ономастической конференции и памяти ее организатора В.А. Никонова (Ульяновск, 20–23 сентября 2017 г.) / под ред. С.В. Рябушкиной, В.И. Супруна, Е.В. Захаровой, Е.Ф. Галушко. В 2 т. Т. 2. Ульяновск: УлГПУ им. И. Н. Ульянова, 2017. С. 177–184.
6. Шапошников А.К. Сравнение именных традиций болгаро-антского и ильменско-словенского диалектных ареалов // Слѣдовати достоинь. Доклади от международната ономастична конференция «Антропоними и антропонимни изследвания в началото на XXI век», посветена на 100-годишнината от рождението на проф. д.ф.н. Йордан Заимов (1921–1987), 20–22 април 2021 г., гр. София. Анна Чолева-Димитрова, Мая Влахова-Ангелова, Надежда Данчева (отг. редактори). София: Издателство на БАН «Проф. Марин Дринов», 2021. С. 107–121.
7. Этимологический словарь славянских языков. Праславянский лексический фонд. Вып. 1–42. М.: Наука, 1974–2021.
8. URL: <https://www.anthroponyms.ru> (дата обращения: 03.05.2023).
9. URL: <http://java.sun.com/developer/Books/jdbc/ch07.pdf> (дата обращения: 03.05.2023).
10. Сообщение о выпуске релиза MySQL 5.7.32 от 19.10.2020. URL: <https://dev.mysql.com/doc/relnotes/mysql/5.7/en/news-5-7-32.html> (дата обращения: 14.12.2022).
11. Суэринг С., Конверс Т., Парк Дж. PHP и MySQL. Библия программиста = PHP 6 and MySQL 6 Bible. 2-е издание. М.: Диалектика, 2010. 912 с.
12. URL: TIOBE Index for December 2022 (англ.) (дата обращения: 03.05.2023).
13. URL: <https://dom.spec.whatwg.org> (дата обращения: 03.05.2023).
14. URL: <https://www.internet-technologies.ru/articles/kak-s-datatables-sozdat-tablicu.html> (дата обращения: 03.05.2023).

References

1. Detschew D. (1957) *Die thrakischen Sprachreste*. Wien: im Kommission bei Rudolf M. Rohrer. 595 p.
2. Vasiliev V.L. (2012) *Slavyanskije toponimicheskie drevnosti Novgorodskoy zemli* [Slavic toponymic antiquities of the Novgorod land] *Rukopisnye pamyatniki Drevney Rusi* [Handwritten monuments of Ancient Russia]. Moscow. 816 p.
3. Tupikov N.M. (1903) *Slovar' drevnerusskikh lichnykh sobstvennykh imen* [Dictionary of Old Russian personal proper names] *Tip. I.N. Skorokhodova* [Typogr. I.N. Skorokhodova]. St. Petersburg. 863 p.
4. Shaposhnikov A.K. (2017) *Praslavyanskaya etimologiya st.-chesh* [Proto-Slavic etymology of art.-chesh] *Antroponima Kroměžír v svete drevnenovgorodskikh antroponimicheskikh svidetel'stv* [The anthroponym Kroměžír in the light of ancient Novgorod anthroponymic evidence] *Studia etymologica Brunensia* [Studia etymologica Brunensia]. Brno. P. 39–42.
5. Shaposhnikov A.K. (2017) *Drevnerusskie slozhnye imena mladshey serii na praslavyanskom fone: etimologicheskij aspekt* [Old Russian compound names of the younger series on the Proto-Slavic background: etymological aspect] *Onomastika Povolzh'ya: mat-ly 16 Mezhdunarodnoy nauchnoy konferentsii, posv. 50-letnemu yubileyu Pervoy Povolzhskoy onomasticheskoy konferentsii i pamyati ee organizatora V.A. Nikonova (Ul'yanovsk, 20–23 sentyabrya 2017 g.) Pod red. S.V. Ryabushkinoy, V.I. Supruna, E.V. Zakharovoy, E.F. Galushko. V 2t. T. 2. UIGPU im. I.N. Ul'yanova* [Onomastics of the Volga Region: materials of the 16th International Scientific Conference dedicated to the 50th anniversary of the First Volga Onomastic Conference and the memory of its organizer V.A. Nikonov (Ulyanovsk, September 20–23, 2017) Edited by S.V. Ryabushkina, V.I. Suprun, E.V. Zakharova, E.F. Galushko. In 2 vols. Vol. 2. UIGPU named after I.N. Ulyanov. Ulyanovsk. P. 177–184.
6. Shaposhnikov A.K. (2021) *Sravnenie imennykh traditsiy bolgaro-antskogo i il'mensko-slovenskogo dialektnykh arealov* [Comparison of the nominal traditions of the Bulgarian-Ant and Ilmen-Slovene dialect

areas] *Slavovati dostoit*. *Dokladi ot mezhdunarodnata onomastichna konferentsiya «Antroponimi i antro-ponimni izsledvaniya v nachaloto na XXI vek», posvetena na 100-godishninata ot rozhdenieto na prof. d.f.n. Yordan Zaimov (1921–1987), 20–22 april 2021 g., gr. Sofiya. Anna Choleva-Dimitrova, Maya Vlahova-Angelova, Nadezhda Dancheva (otg. redaktori). Izdatelstvo na BAN «Prof. Marin Drinov* [Slavovati dostoit. The report from the international onomastic conference «Anthroponymy and anthroponymy from the beginning to the XXI century», illuminated for the 100th year of the birth of Professor Yordan Zaimov (1921–1987), April 20–22, 2021. Anna Choleva-Dimitrova, Maya Vlahova-Angelova, Nadezhda Dancheva (editor’s note). Publishing house on BAN «Prof. Marine Drinov». Sofia. P. 107–121.

7. (1974–2021) *Etimologicheskiy slovar’ slavyanskikh yazykov* [Etymological dictionary of Slavic languages] *Praslavyanskiy leksicheskiy fond* [Proto-Slavic lexical fund]. Issue 1-42] *Nauka* [Nauka]. Moscow.

8. Available at: <https://www.antroponims.ru> (accessed: 03.05.2023).

9. Available at: <http://java.sun.com/developer/Books/jdbc/ch07.pdf> (date of application: 03.05.2023).

10. *Soobshchenie o vypuske reliza MySQL 5.7.32 ot 19.10.2020* [Message about the release of MySQL release 5.7.32 from 19.10.2020]. Available at: <https://dev.mysql.com/doc/relnotes/mysql/5.7/en/news-5-7-32.html> (date of application: 14.12.2022).

11. Suering S., Converse T., Park J. (2010) *Bibliya programmista = PHP 6 and MySQL 6 Bible. 2-e izdanie* [Programmer’s Bible = PHP 6 and MySQL 6 Bible. 2nd edition] *Dialektika* [Dialectics]. Moscow. 912 p.

12. Available at: TIOBE Index for December 2022 (date of application: 03.05.2023).

13. Available at: <https://dom.spec.whatwg.org> (date of application: 03.05.2023).

14. Available at: <https://www.internet-technologies.ru/articles/kak-s-datatables-sozdat-tablicu.html> (date of application: 03.05.2023).