

Анализ трёх аспектов современных стран по Викиданным: возраст стран, популярные формы правления и этнохоронимы

Andrew Krizhanovsky^a, Elizaveta Smykova^b

^a*Institute of Applied Mathematical Research of the Karelian Research Centre of the Russian Academy of Sciences*

^b*Petrozavodsk State University*

Abstract

Эта глава посвящена исследованию стран на основе базы знаний международного проекта Викиданные. С помощью SPARQL-запросов, вычисляемых на объектах типа "страна" в Викиданных, получены: выведен список всех ныне существующих стран, перечень стран, упорядоченных по дате создания, список этнохоронимов стран, пузырьковая диаграмма с формами правления стран и граф соседних стран. Кроме того, сделаны относительно полноты Викиданных по данной теме.

Примечание

Статья распространяется по лицензии Creative Commons Attribution-ShareAlike. Материалы этой статьи использованы в главе курса Викиверситета "Программирование Викиданных"[1]. Иллюстрации загружены на Викисклад. Над статьёй в 2017 году работали Крижановский А. А., Смыкова Е. М.

1. Экземпляры объекта Страна

Построим список всех стран.

- Объект: страна (Q6256),
- Свойство: экземпляр (P31).

```
1 #List of 'instances of' "country"
2 SELECT ?lang ?langLabel
3 WHERE
4 {
5     ?lang wdt:P31 wd:Q6256.
6     SERVICE wikibase:label { bd:serviceParam wikibase:language "en" }
7 }
```

1: Список стран

SPARQL-запрос, 198 записей

”+”: Примерами наиболее полных и проработанных стран на Викиданных являются: Соединённые Штаты Америки, Канада, Испания.

”-”: Почти пустыми и малоинформативными странами оказались: Сахарская Арабская Демократическая Республика, Приднестровская Молдавская Республика, Косово.

2. Возраст стран

Построим список стран, отсортированных по дате основания страны (первом упоминании о стране).

- Объект: страна (Q6256),
- Свойство: дата основания (P571).

```
1 #List of 'instances of' countries sorted by inception"
2 SELECT ?country ?countryLabel ?inception
3 WHERE
4 {
5     ?country wdt:P31 wd:Q6256.          #country
6     ?country wdt:P571 ?inception .    #inception of country
7
8     SERVICE wikibase:label { bd:serviceParam wikibase:language "en" }
9 }
10 ORDER BY (?inception)
```

2: Список стран упорядоченных по дате создания

SPARQL-запрос, 112 записей.

В результате выполнения запроса получен список стран с датами их создания. Например, Абхазия – 1 января 0786, Россия – 1 января 0862, Косово – 17 февраля 2008, Южный Судан – 9 июля 2011.

Годы, в которые было создано наибольшее количество стран – 1991 (17 стран), 1812 (6 стран) и 1918 (5 стран).

2.1. Полнота Викиданных

Проанализируем полноту Викиданных.

- По данным ”Общероссийского классификатора стран мира” [2] за 2016 год на земле существует 251 страна.
- В этой задаче не учитываются древние, уже не существующие государства (например: Ассирия (Q41137)), поскольку они являются экземпляром не объекта ”country”, а объекта ”former country” (бывшие страны). Отметим, что количество бывших стран на порядок больше существующих ныне стран (см. SPARQL-запрос, возвращающий более двух тысяч таких стран).

- По данным категории "Алфавитный список стран и территорий" Русской Википедии существует 252 страны. (В "Общероссийском классификаторе стран мира" недостает Косово)
- По данным категории "List of sovereign states" Английской Википедии существует 206 стран.

Не всегда можно точно указать дату основания страны по разным причинам: отсутствие, недостаток или противоречие письменных источников. Например, основание Древнерусского государства связывают с призыванием варяжского князя Рюрика в 862 году, но точной даты нет (объект Russia (Q159)). Так же некоторым современным странам предшествовал ряд других и дату образования какого из них считать за дату создания современной страны – это вопрос открытый (например, Монголия (Q711)).

2.2. Страны с незаполненной датой основания

Выведем список стран с пустым свойством "дата основания":

- Объект: страна (Q6256),
- Свойство: дата основания (P571).

```

1 #List of 'instances of' "countries without an inception"
2 SELECT ?country ?countryLabel
3 WHERE
4 {
5   ?country wdt:P31 wd:Q6256.          #country
6
7   MINUS { ?country wdt:P571 [] } . #inception of country is empty
8   SERVICE wikibase:label { bd:serviceParam wikibase:language "en" }
9 }

```

3: Список стран с пустым свойством "дата создания"

SPARQL-запрос, 100 записей.

Итак, на 6 марта 2017 года Викиданные содержат 100 из 198 записей о ныне существующих странах с неизвестным годом основания страны.

3. Этнохоронимы на русском языке

Этнохороним – название жителей определённой местности соотнесённое с топонимом. Например, Россия – россияне, россиянин, россиянка, Чехия – чехи, чех, чешка.

Построим список стран у которых есть этнохоронимы на русском языке.

- Объект: страна (Q6256),
- Свойство: этнохороним (P1549).

```

1 #List of countries with demonyms in Russian
2 SELECT ?country ?countryLabel
3 WHERE
4 {
5     ?country wdt:P31 wd:Q6256. #country
6     ?country wdt:P1549 ?demonym . #demonym
7
8     FILTER((LANG(?demonym)) = "ru")
9
10    SERVICE wikibase:label { bd:serviceParam wikibase:language "ru" }
11 }
12 GROUP BY ?country ?countryLabel

```

4: Список стран с этнохоронимами на русском языке

SPARQL-запрос, 28 записей.

На 19 марта 2017 года Викиданные содежали 28 из 198 стран с заполненными этнохоронимами на русском языке.

3.1. Список этнохоронимов

Выведем список всех этнохоронимом на русском языке.

```

1 #List of demonyms in Russian
2 SELECT ?country ?countryLabel ?demonym
3 WHERE
4 {
5     ?country wdt:P31 wd:Q6256. #country
6     ?country wdt:P1549 ?demonym. #demonym
7
8     FILTER((LANG(?demonym)) = "ru")
9
10    SERVICE wikibase:label { bd:serviceParam wikibase:language "ru" }
11 }

```

5: Список этнохоронимов на русском языке

SPARQL-запрос, 83 записей.

На 19 марта 2017 года Викиданные содежат 83 заполненных этнохоронима.

3.2. Страны с незаполненными этнохоронимами

Построим список стран, у которых нет этнохоронимов на русском языке.

```

1 #List of countries without demonyms in Russian
2 SELECT ?country ?countryLabel
3 WHERE
4 {
5     ?country wdt:P31 wd:Q6256. # country
6
7     MINUS {
8         ?country wdt:P1549 ?demonym. # except with demonyms
9         FILTER((LANG(?demonym)) = "ru") # in Russian
10    }

```

```

11 SERVICE wikibase:label { bd:serviceParam wikibase:language "ru" }
12 }
13 GROUP BY ?country ?countryLabel

```

6: Список стран у которых нет этнохоронимов на русском языке

SPARQL-запрос, 170 записей.

На 19 марта 2017 года Викиданные содержат 170 из 198 стран с незаполненными этнохоронимами.

После заполнения авторами статьи этнохоронимов, число стран, у которых их нет на русском языке стало 69, а самих этнохоронимов на русском языке теперь 276 (на 20 марта 2017 года).

3.3. Количество заполненных этнохоронимов у стран

Выведем список стран, упорядоченный по количеству заполненных в Викиданных этнохоронимов.

```

1 #\selectlanguage{russian} \selectlanguage{english}ount of demonyms in
   countries
2 SELECT ?country ?countryLabel (count(*) as ?count)
3 WHERE
4 {
5     ?country wdt:P31 wd:Q6256. # country
6     ?country wdt:P1549 ?demonym. # demonyms
7     SERVICE wikibase:label { bd:serviceParam wikibase:language "ru" }
8 }
9 GROUP BY ?country ?countryLabel
10 ORDER BY DESC(?count)

```

7: Страны упорядоченные по количеству заполненных этнохоронимов

SPARQL-запрос, 199 записей.

Наибольшее число этнохоронимов у Соединённых Штатов Америки (41 этнохороним), затем идут Великобритания (40), Германия (40), Канада (36) и Россия (34).

4. Формы правления стран

Построим пузырьковую диаграмму форм правления стран.

- Объект: страна (Q6256),
- Свойство: форма правления (P122).

```

1 #basic form of government ranking
2 #defaultView:BubbleChart
3 SELECT ?bfog ?form (count(*) as ?count)
4 WHERE
5 {
6     ?country wdt:P31 wd:Q6256. #country
7     ?country wdt:P122 ?bfog . #basic form of government

```

```

8   OPTIONAL {
9   ?bfog rdfs:label ?form
10  filter (lang(?form) = "ru")
11  }
12  }
13  GROUP BY ?bfog ?form
14  ORDER BY DESC(?count) ASC(?form)

```

8: Пузырьковая диаграмма форм правления стран

SPARQL-запрос, 30 записей.

В результате выполнения запроса 8 мы получаем пузырьковую диаграмму с наиболее распространенными формами правления в странах. Основные формы правления стран: республика (в 20 странах), конституционная монархия (в 18 странах), федеративная республика (в 18 странах), парламентская республика (в 17 странах) и президентская республика (в 12 странах). 1

5. Соседние страны

Построим граф соседних стран.

- Объект: страна (Q6256),
- Свойство: имеет границы с (P47).

```

1 #neighboring countries graph
2 #defaultView:Graph
3 SELECT ?country ?countryLabel ?sharesBorderWith ?sharesBorderWithLabel
4 WHERE
5 {
6   ?country wdt:P31 wd:Q6256. #countries
7
8   SERVICE wikibase:label { bd:serviceParam wikibase:language "ru" }
9   OPTIONAL { ?country wdt:P47 ?sharesBorderWith . } #shares border with
10 }

```

SPARQL-запрос, 787 записей.

В результате выполнения запроса мы получаем граф 2 с 787 ребрами, где ребро – это соседство между двумя странами. Граф представляет из себя несколько связанных компонент, так как есть островные страны, у которых нет соседей (например, Маврикий, Мальдивы, Мадагаскар).

- [1] Country. Research in programming Wikidata [Programmirovaniye Vikidannykh] // Russian Wikiversity, https://ru.wikiversity.org/wiki/%D0%9F%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B8%D1%80%D0%BE%D0%B2%D0%B0%D0%BD%D0%B8%D0%B5_%D0%92%D0%B8%D0%BA%D0%B8%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85/%D0%A1%D1%82%D1%80%D0%B0%D0%BD%D1%8B, accessed 23 Feb 2017 (2017).
- [2] Russian classification of countries of the world], <http://classifikators.ru/oksm>, accessed 14 May 2017 (2016).



Figure 1: Пузырьковая диаграмма форм правления стран

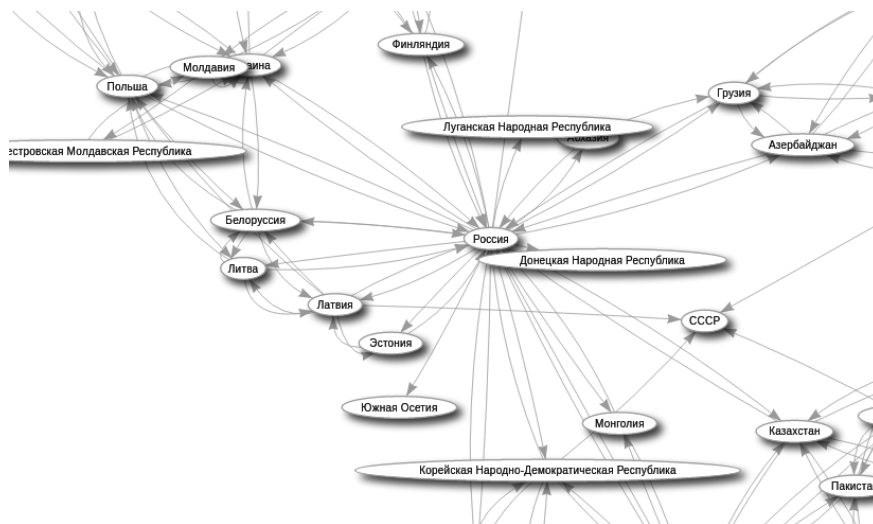


Figure 2: Граф соседних стран, в центре Россия