

Даты основания и география академических институтов

Andrew Krizhanovsky^a, Egor Horolskiy^b

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

^b*Petrozavodsk State University*

Abstract

Статья посвящена исследованию объекта Викиданных – “Академическое учебное заведение”. С помощью SPARQL-запросов к Викиданным были решены следующие задачи:

Построение списка всех академических учебных заведений;

Построение пузырьковой диаграммы, показывающей соотношение языков, на которых "говорят" различные академические учебные заведения;

Отображение на карте академических учебных заведений во всем мире.

В ходе работы была дополнена информация у академических учебных заведений на Викиданных, были заполнены географические координаты. Сделаны выводы о полноте и неполноте данной информации, представленной в Википедии и в Викиданных.

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., Хорольский Е. Н.

1. Экземпляры объекта “Академическое учебное заведение”

Используем:

- Объект: академическое учебное заведение (Q4671277),

- Свойство: instance of (P31).

Построим список всех академических учебных заведений.

```
111 #added 2017-02
122 #List of 'instances of' "academic institution"
133 SELECT ?academic_institution ?academic_institutionLabel
144 WHERE
155 {
166     ?academic_institution wdt:P31 wd:Q4671277.
177     SERVICE wikibase:label { bd:serviceParam wikibase:language "en" }
188 }
```

Листинг 1: Экземпляры объекта "академическое учебное заведение"

19 SPARQL-запрос, 3892 записей.

20 “+”: Наиболее полными и проработанными академическими учебными заведения-
21 ми на Викиданных являются: Балашовский филиал Саратовского государственного
22 университета, Московский государственный институт культуры, Самарская государ-
23 ственная областная академия.

24 “-”: Почти пустыми и малоинформативными академическими учебными заведени-
25 ями оказались: Новороссийский политехнический институт, Ростовский художествен-
26 ный колледж, Омский государственный институт сервиса.

27 2. Дата основания академических учебных заведений по всему миру

28 Построим гистограмму академических учебных заведений разных стран мира, от-
29 сортированных по дате основания (Рис. 1).

30 Используем:

- 31 • Объект: академическое учебное заведение (Q4671277).
- 32 • Свойство: дата основания (P571).

```
331 #defaultView: BarChart
342 SELECT DISTINCT (SAMPLE(?year) AS ?year) (COUNT(?year) AS ?count) (SAMPLE(?
35   _academicLabel) AS ?_academicLabel)
363 WHERE {
374   ?object (wdt:P31) wd:Q4671277 .
385   BIND(str(YEAR(?_inception)) AS ?year)
396   ?object wdt:P571 ?_inception .
407   ?object wdt:P17 ?_country .
418   ?_country rdfs:label ?_academicLabel .
429   FILTER((LANG(?_academicLabel)) = "en")
430 }
441 GROUP BY ?_inception ?_country
452 ORDER BY ?year ?_inception
```

Листинг 2: Дата основания академических учебных заведений по всему миру

46 SPARQL-запрос, 381 запись.

47 По данным гистограммы можно определить, что наибольшее количество академи-
48 ческих учебных заведений было основано в 2002 году.

49 3. Академические учебные заведения России

50 Строим SPARQL-запрос, чтобы узнать, о скольких отечественных академических
51 учебных заведениях есть информация в Викиданных. Используем:

- 52 • Объект: академическое учебное заведение (Q4671277).
- 53 • Свойство: экземпляр (P31).

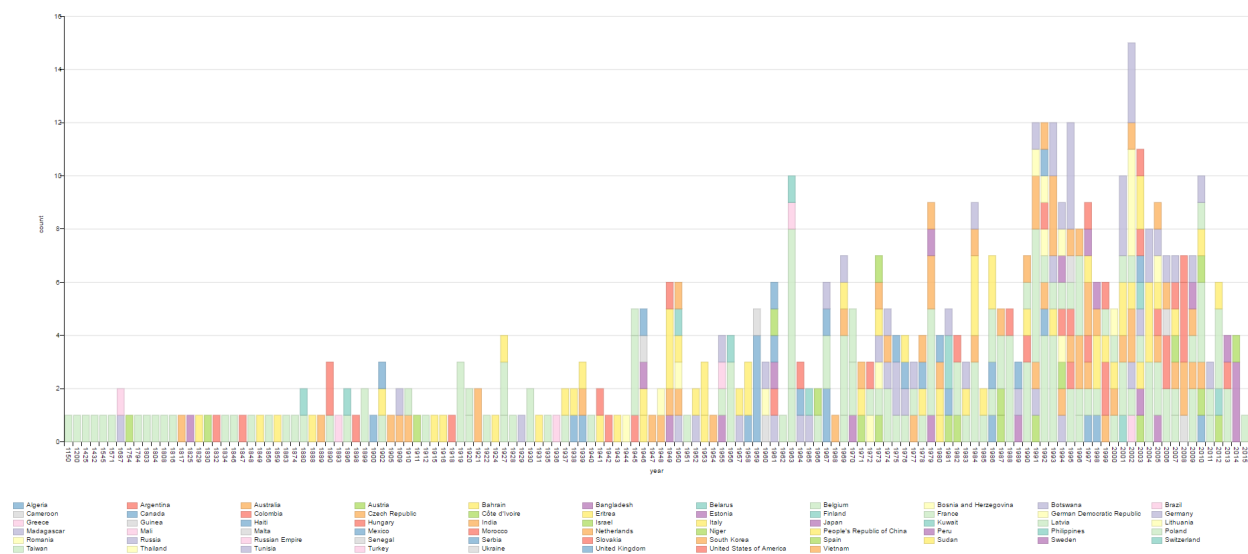


Рис. 1: Гистограмма, отображающая дату основания и количество академических учебных заведений в разных странах

```

54.1 #added 2017-05
55.2 #Number of academic institutions in Russia, USSR and Russian empire.
56.3 SELECT ?academic_institution ?academic_institutionLabel ?country WHERE {
57.4   ?academic_institution wdt:P31 wd:Q4671277.
58.5   { ?academic_institution wdt:P17 wd:Q15180 } UNION
59.6   { ?academic_institution wdt:P17 wd:Q159 } UNION
60.7   { ?academic_institution wdt:P17 wd:Q34266 }
61.8   ?academic_institution rdfs:label ?item_label.
62.9   SERVICE wikibase:label { bd:serviceParam wikibase:language "ru". }
63.0   FILTER((LANG(?item_label)) = "ru")
64.1 }

```

Листинг 3: Список академических учебных заведений России СССР и Российской Империи

65 SPARQL запрос, 24 записи.

66 Категория Academic institutions in Russia Английской Википедии содержит инфор-
67 мацию о нескольких десятках академических учебных заведений, находящихся в Рос-
68 сии. Русская Википедия вообще не имеет страниц, как-либо связанных с понятием
69 академическое учебное заведение. Таким образом, большая часть информации об ака-
70 демических учебных заведений России в представлена в Английской Википедии.

71 4. Академические учебные заведения на карте мира

72 Отображаем на карте академические учебные заведения и высшие учебные заведе-
73 ния по всему миру.

74 Используем:

- 75 • Объект: академическое учебное заведение (Q4671277), высшее учебное заведение
76 (Q38723).
- 77 • Свойство: географические координаты (P625), экземпляр (P31).

```

78 1 #List of 'coordinate location' "academic institution and higher educational
79     institution"
80 2 SELECT ?academic_institution ?academic_institution_description ?coord
81 3 WHERE
82 4 {
83 5     { ?academic_institution wdt:P31 wd:Q4671277; wdt:P625 ?coord } UNION
84 6     { ?academic_institution wdt:P31 wd:Q38723; wdt:P625 ?coord }
85 7     SERVICE wikibase:label { bd:serviceParam wikibase:language "en" }
86 8 }

```

Листинг 4: Список академических учебных заведений и высших учебных заведений по всему миру

87 SPARQL-запрос, 1395 записей.

88 4.1. Полнота Викиданных: академические учебные заведения во всем мире

89 Академическое учебное заведение - согласно английской Википедии - по окончанию
90 обучения позволяет человеку получить определенную учёную степень или квалифика-
91 цию выпускника. Таким образом, все высшие учебные заведения мы можем считать
92 академическими учебными заведениями.

93 Рассмотрим статистику на примере университетов:

94 По данным Webometrics Ranking of World Universities на Земле существует более
95 19000 университетов.

96 По данным категории “Университеты по алфавиту” Русской Википедии существует
97 более 2500 университетов.

98 Как видно из SPARQL-запроса (листинг 3), лишь 1395 вузов имеют заполненное
99 свойство географические координаты (P625). Причина в том, что малое количество
100 объектов Викиданных - около 7% - заполнены корректно и полноценно.

101 4.2. Список академических учебных заведений с пустым свойством географические 102 координаты

103 Выведем список академических учебных и высших учебных заведений заведений с
104 пустым свойством географические координаты (P625):

105 Используем:

- 106 • Объект: академическое учебное заведение (Q4671277), высшее учебное заведение
107 (Q38723).
- 108 • Свойство: географические координаты (P625), экземпляр (P31).

```

109 1 #List of 'instances of' "academic_institution without a coordinate location"
110 2 SELECT ?academic_institution ?academic_institution_label ?
111      academic_institution_description ?coord
112 3 WHERE
113 4 {
114 5     {?academic_institution wdt:P31 wd:Q4671277} UNION #academic_institution
115 6     {?academic_institution wdt:P31 wd:Q38723}
116 7
117 8     MINUS { ?academic_institution wdt:P625 [] } . #coordinate location of
118 9     academic_institution
119 0     SERVICE wikibase:label { bd:serviceParam wikibase:language "en" }
120 0 }

```

Листинг 5: Список академических учебных заведений с пустым свойством географические координаты

121 SPARQL-запрос, 3113 записей.

122 4.3. Неполнота Викиданных: академические учебные заведения во всем мире



Рис. 2: Объекты с заполненным свойством географические координаты, версия 1.0

123 Видно (листинг 3), что только у 1395 академических учебных заведений заполне-
124 но свойство coordinate location. Давайте немного улучшим данное положение вещей.
125 Данные видны на рис. 2 (версии 1.0).

126 У ста объектов было заполнено свойство “географические координаты”. В результате
127 чего выполнение скрипта (листинг 3) возвращает 1495 записей. Данные видны на рис. 3
128 (версии 1.1).



Рис. 3: Объекты с заполненным свойством географические координаты, версия 1.1

- 129 [1] https://en.wikiversity.org/wiki/Research_in_programming_Wikidata/Academic_institution,
130 Research in programming Wikidata/Academic institution, 2017.
- 131 [2] <http://www.webometrics.info/en/world>, Webometrics Ranking of World Universities, 2017.