

Цифровая гуманитаристика
Петербургский кампус ВШЭ

Алексей Куприянов

Национальный исследовательский университет «Высшая школа экономики»
(С. Петербург)

«Практики Digital Humanities: методы и инструменты»
7–8 ноября 2019 (НИУ ВШЭ, Пермь)

ДН в Петербургском кампусе ВШЭ

Люди и направления работы

- ▶ Кирилл Маслинский <https://www.hse.ru/staff/kmaslinsky> :
Text mining
- ▶ Татьяна Шерстинова
<https://www.hse.ru/en/org/persons/209332050> : Speech
Corpus
- ▶ Евгений Казарцев
<https://www.hse.ru/org/persons/106252141> : Стиховедение
- ▶ ...

Shameless self-promotion

Направления работы

- ▶ Просопография (совместно с Кириллом Маслинским)
- ▶ Исторические GIS (совместно с Александрой Бекасовой)
- ▶ Distant reading (совместно с Максимом Деминым)
- ▶ OCR Tesseract training (совместно с Кириллом Маслинским)

Просопография

Университеты Российской империи

Объем: на 2017-06-15

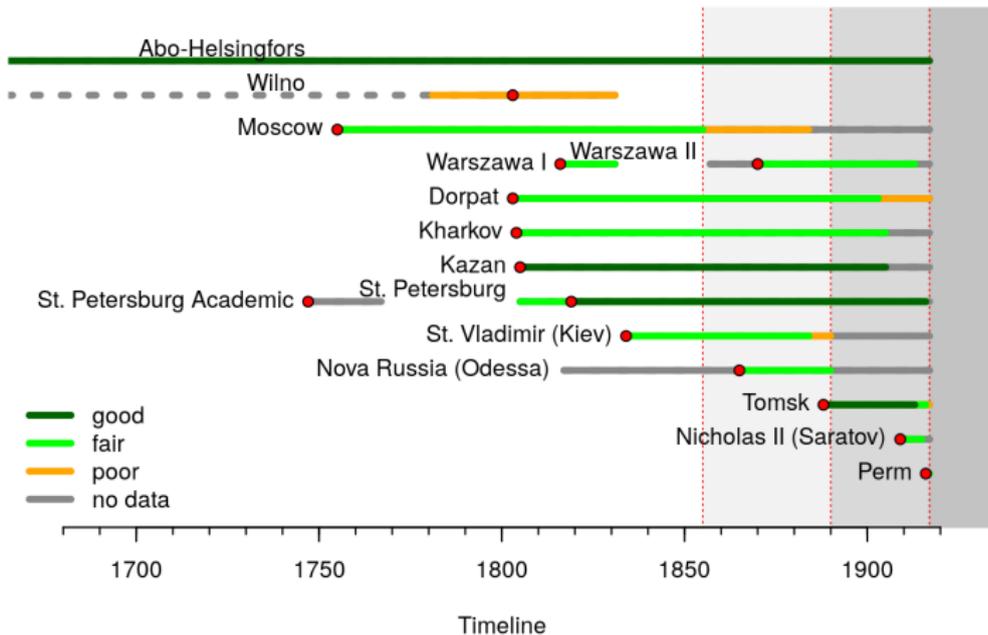
- ▶ 7 университетов (из 15).
- ▶ Около 3800 имен.
- ▶ Менее 7000 строк.

Объем: на 2018-10-05

- ▶ 14 университетов (из 15).
- ▶ 5286 имен (из возм. 6000–7000) [NB! профессора: 2091].
- ▶ 11005 строк.

Просопография

Университеты Российской империи



Просопография

Университеты Российской империи

Источники

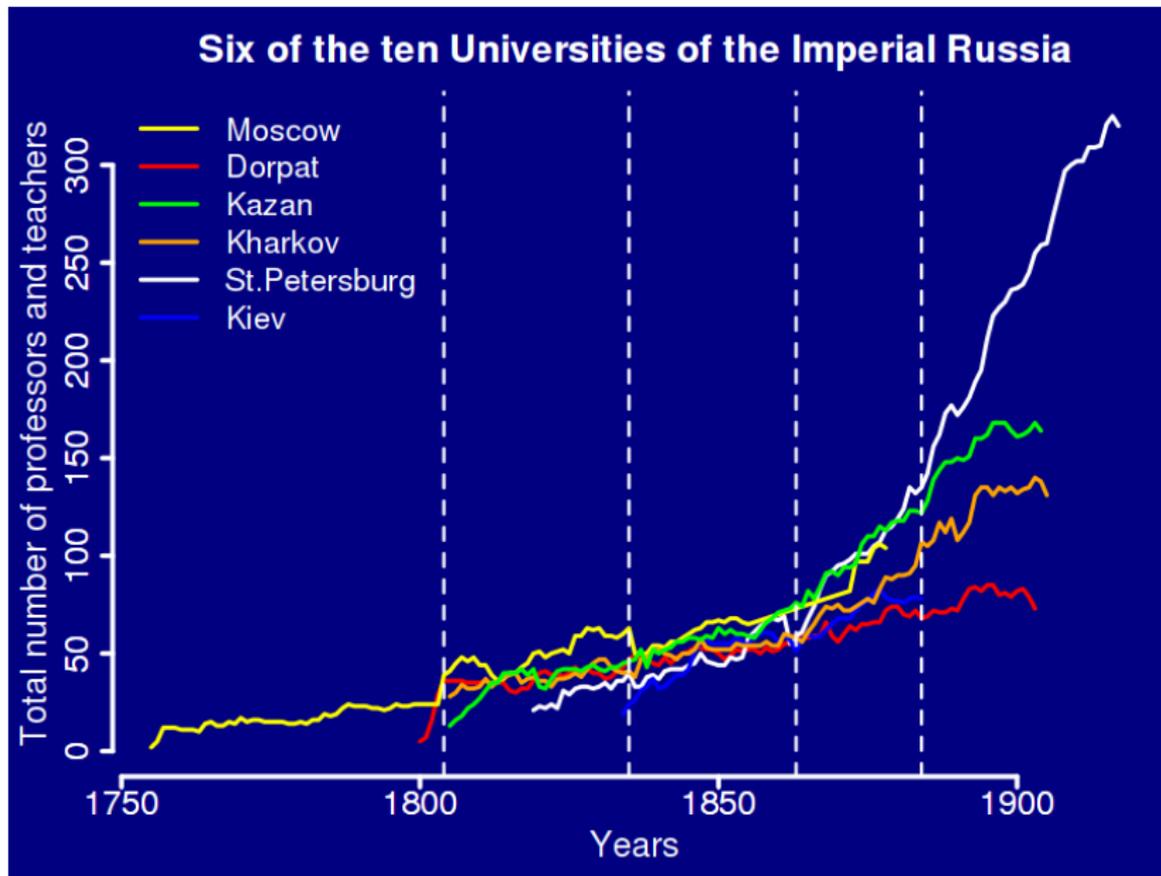
- ▶ Юбилейные биографические словари.
- ▶ Современные биографические словари.
- ▶ Опубликованные источники: отчеты, академические списки, объявления о лекциях.
- ▶ Крайне редко: архивы.

Строка данных (упрощенно)

- ▶ University – Unique ID – Position – Start Date – End Date – A_1 – ... – A_n

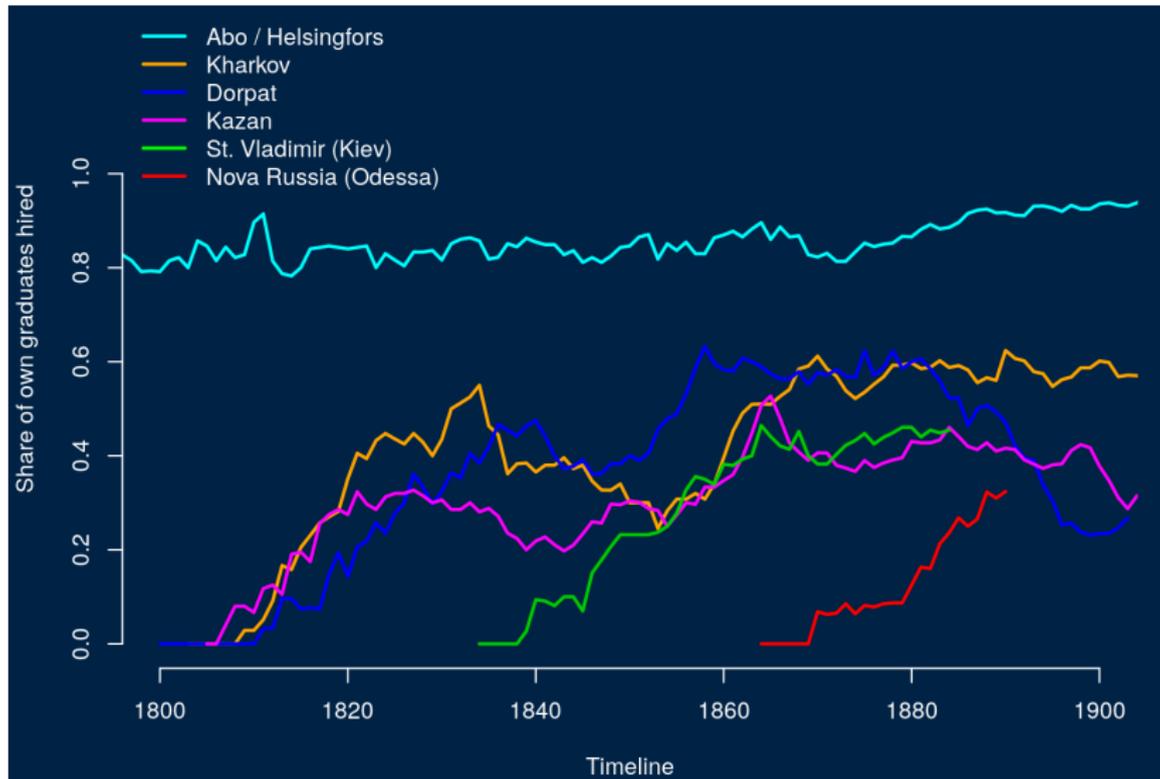
Просопография

Университеты Российской империи



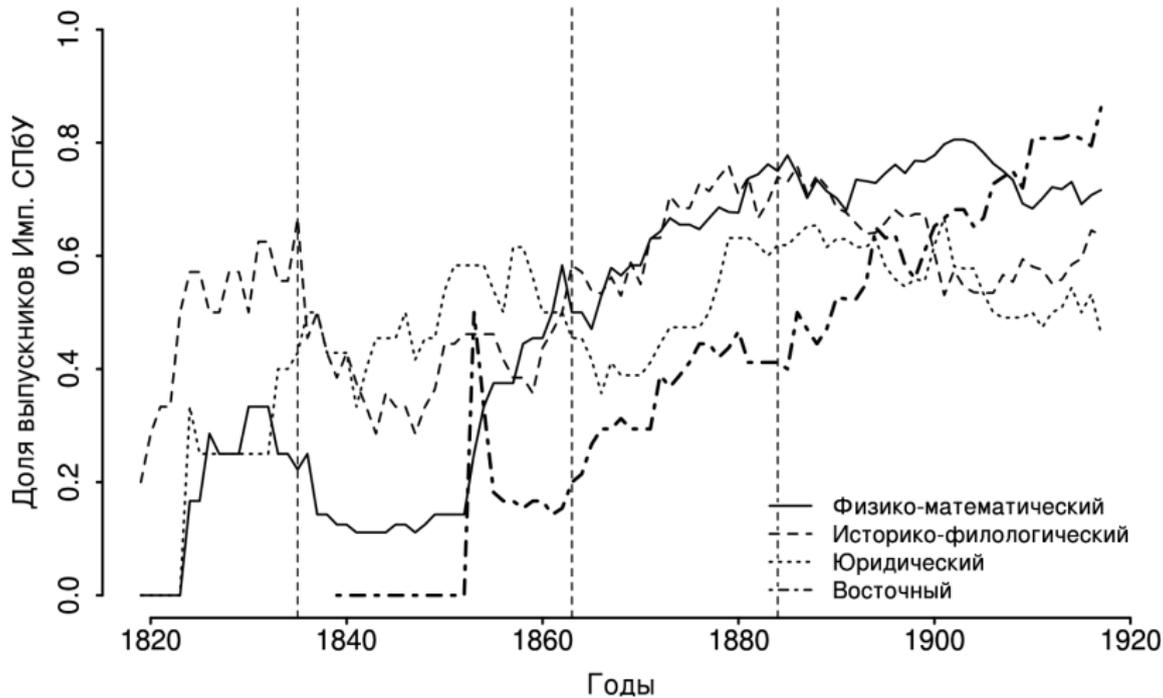
Просопография

Университеты Российской империи



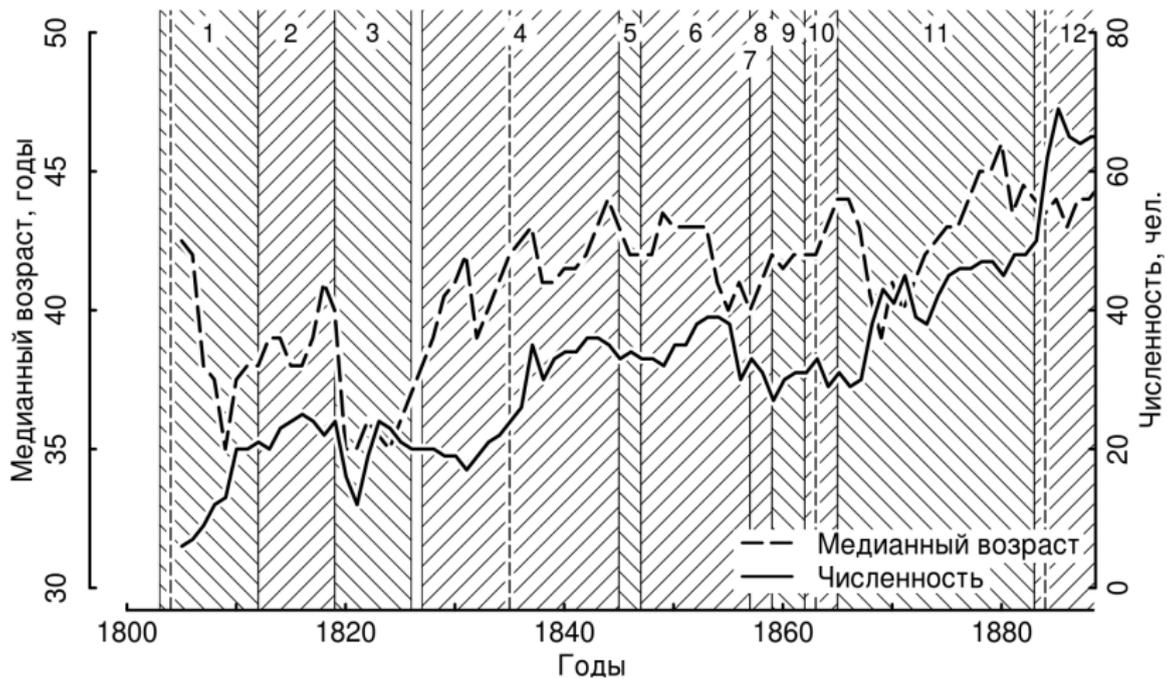
Просопография

Университеты Российской империи



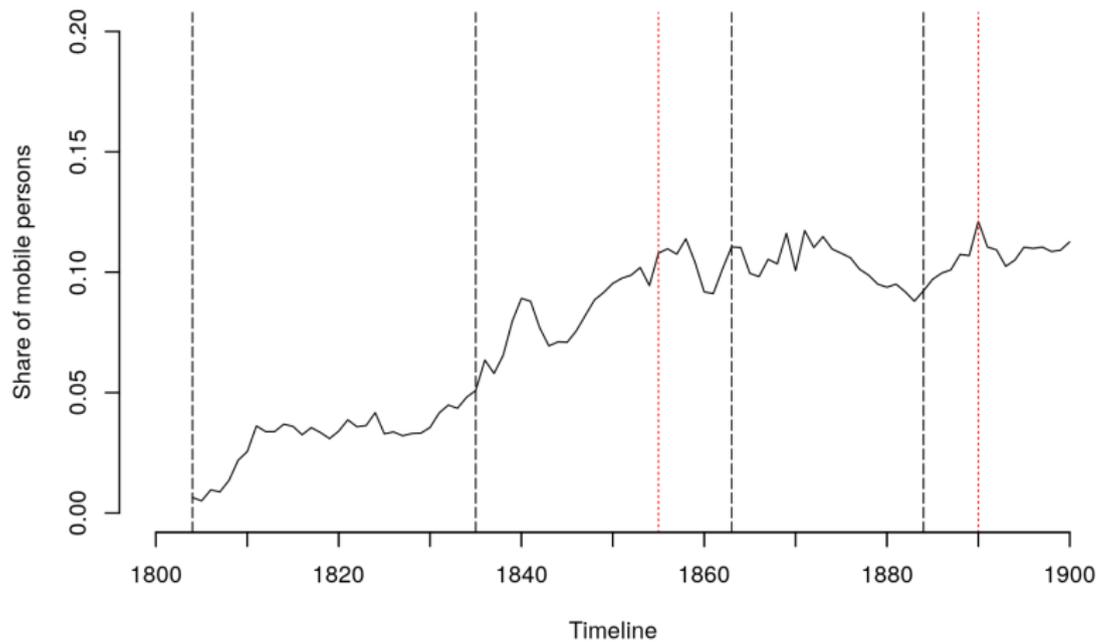
Просопография

Университеты Российской империи



Просопография

Университеты Российской империи



Матрица мобильности (ряд: *из*; столбец: *в*)

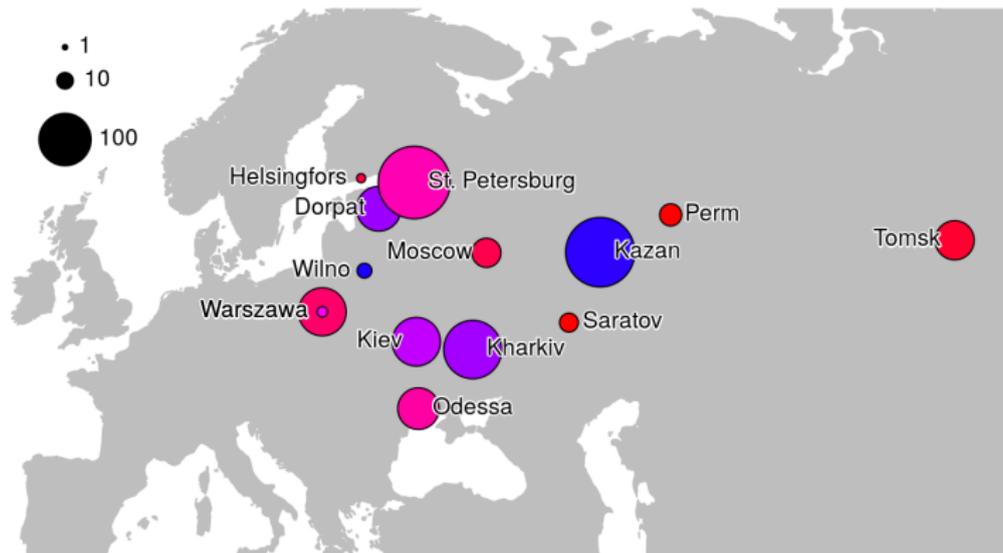
	AHE	WLN	WR1	WR2	DPT	KAZ	MOS	NIK	NVR	PRM	SPB	USV	TOM	KHR
AHE	1	.	.
WLN	.	.	2	.	.	.	1	1	.	2
WR1	.	1	1	.	.	.
WR2	3	7	.	.	3	.	11	2	3	4
DPT	1	.	.	5	.	8	1	.	1	1	15	5	.	4
KAZ	.	1	.	12	13	.	9	8	10	.	27	8	18	18
MOS	1	3	4	2	.	1
NIK
NVR	.	.	.	5	3	5	11	1	1	2
PRM
SPB	1	.	.	15	8	14	3	.	7	16	.	6	9	8
USV	.	.	.	1	.	6	5	.	5	.	13	.	3	13
TOM	.	.	.	1	.	4	.	.	.	1	8	.	.	3
KHR	.	.	.	12	4	5	1	5	9	.	14	14	5	.

Сводные данные о мобильности

UNIVERSITY	OUT	IN	FLOW
AHE	1	2	3
WLN	6	2	8
WR1	2	2	4
WR2	33	51	84
DPT	41	32	73
KAZ	124	52	176
MOS	11	20	31
NIK	0	13	13
NVR	28	35	63
PRM	0	18	18
SPB	87	104	191
USV	46	40	86
TOM	17	39	56
KHR	69	55	124

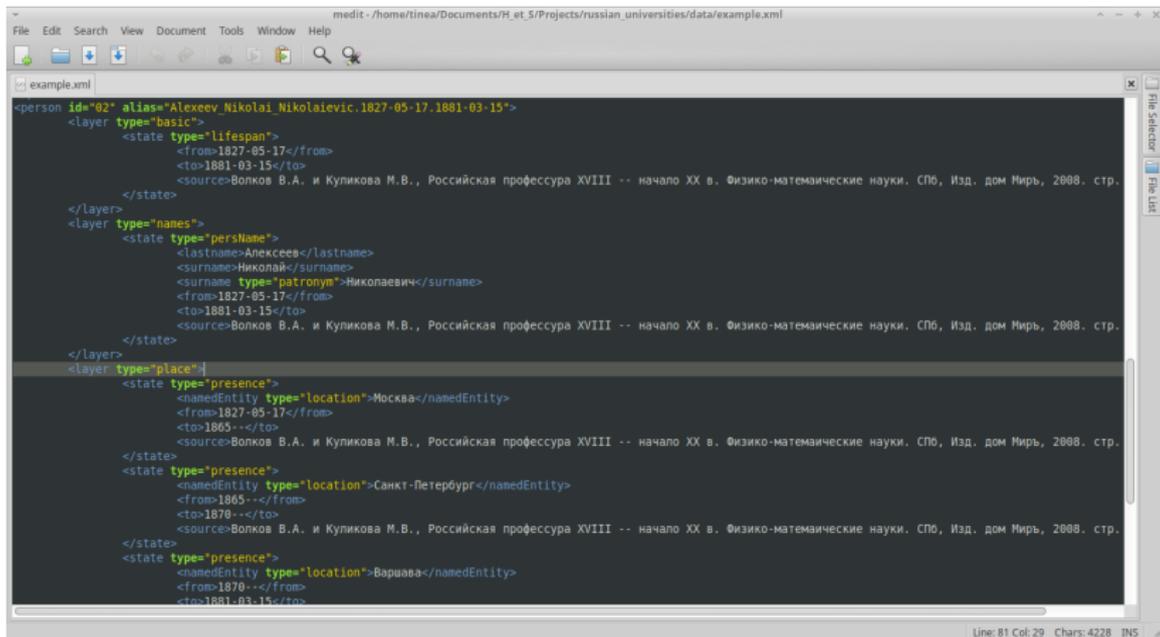
Просопография

Университеты Российской империи



Просопография

Стандарт представления биографий



```
medit - /home/tinear/Documents/H_et_5/Projects/russian_universities/data/example.xml
File Edit Search View Document Tools Window Help
example.xml
<person id="02" alias="Alexeev Nikolai Nikolaevich.1827-05-17.1881-03-15">
  <layer type="basic">
    <state type="lifespan">
      <from>1827-05-17</from>
      <to>1881-03-15</to>
      <source>Волков В.А. и Куликова М.В., Российская профессура XVIII -- начало XX в. Физико-математические науки. СПб, Изд. дом Мирь, 2008. стр.
    </state>
  </layer>
  <layer type="names">
    <state type="persName">
      <lastname>Алексеев</lastname>
      <surname>Николай</surname>
      <surname type="patronym">Николаевич</surname>
      <from>1827-05-17</from>
      <to>1881-03-15</to>
      <source>Волков В.А. и Куликова М.В., Российская профессура XVIII -- начало XX в. Физико-математические науки. СПб, Изд. дом Мирь, 2008. стр.
    </state>
  </layer>
  <layer type="place">
    <state type="presence">
      <namedEntity type="location">Москва</namedEntity>
      <from>1827-05-17</from>
      <to>1865</to>
      <source>Волков В.А. и Куликова М.В., Российская профессура XVIII -- начало XX в. Физико-математические науки. СПб, Изд. дом Мирь, 2008. стр.
    </state>
    <state type="presence">
      <namedEntity type="location">Санкт-Петербург</namedEntity>
      <from>1865</from>
      <to>1870</to>
      <source>Волков В.А. и Куликова М.В., Российская профессура XVIII -- начало XX в. Физико-математические науки. СПб, Изд. дом Мирь, 2008. стр.
    </state>
    <state type="presence">
      <namedEntity type="location">Варшава</namedEntity>
      <from>1870</from>
      <to>1881-03-15</to>
    </state>
  </layer>
</person>
```

Line: 81 Col: 29 Chars: 4228 INS

Исторические GIS

Коммуникационная инфраструктура Российской империи

Источники

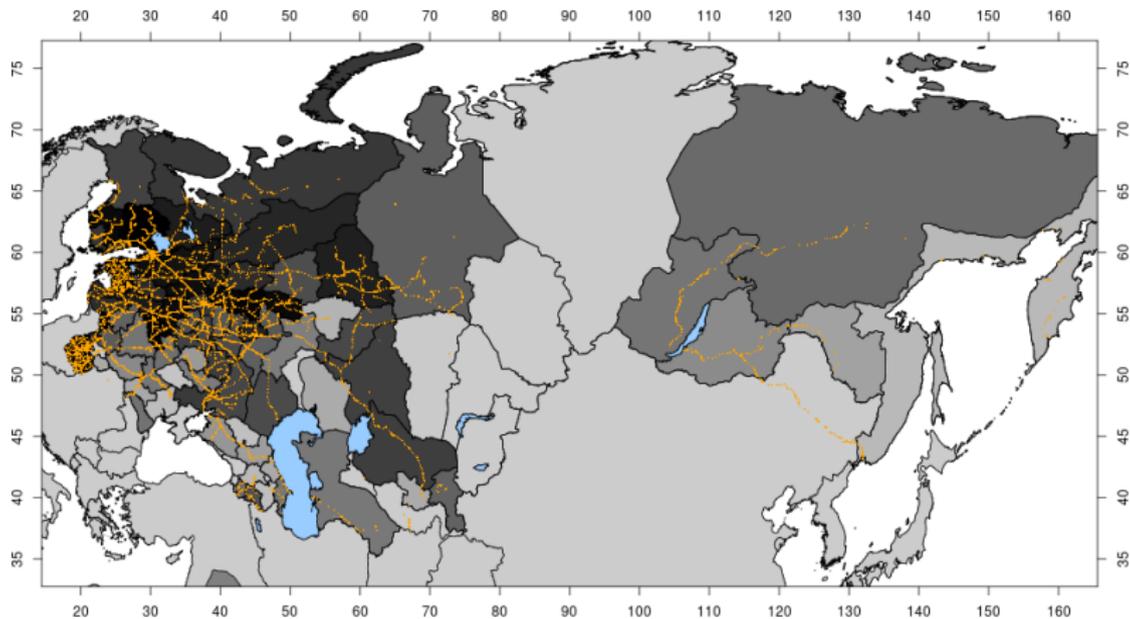
- ▶ Почтовые дорожники. Сейчас в работе 1915 г.

Структура и объем

- ▶ Станции (10473 не уникальных имени [будет больше], 5446 из них с геопозициями [позиционировано $\approx 50\%$])
- ▶ Сегменты (4702 уникальных сегмента [будет больше])
- ▶ Маршруты (1024 маршрута [будет больше])
- ▶ Набор оригинальных shapfiles для стран Евразии и Северной Африки и губерний Российской империи на 1905 г. [будет расширяться]

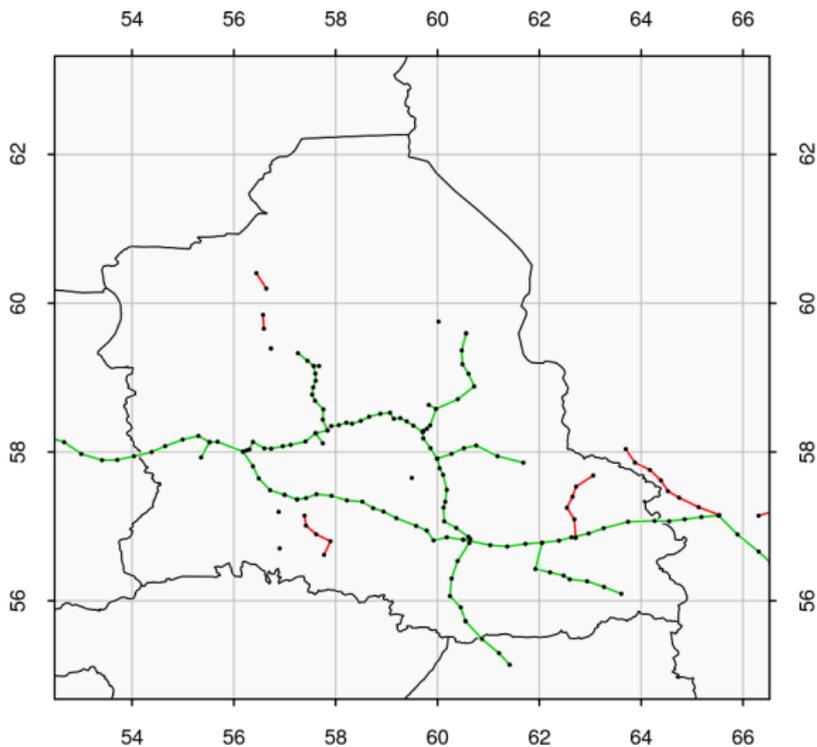
Исторические GIS

Коммуникационная инфраструктура Российской империи



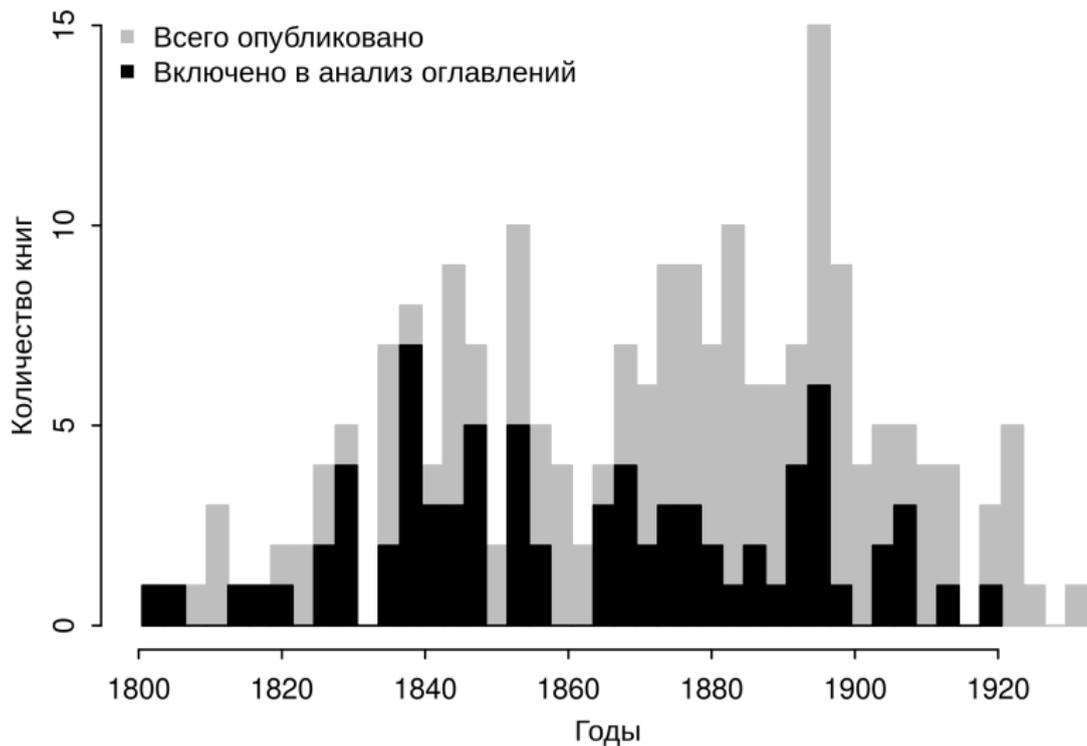
Исторические GIS

Коммуникационная инфраструктура Российской империи



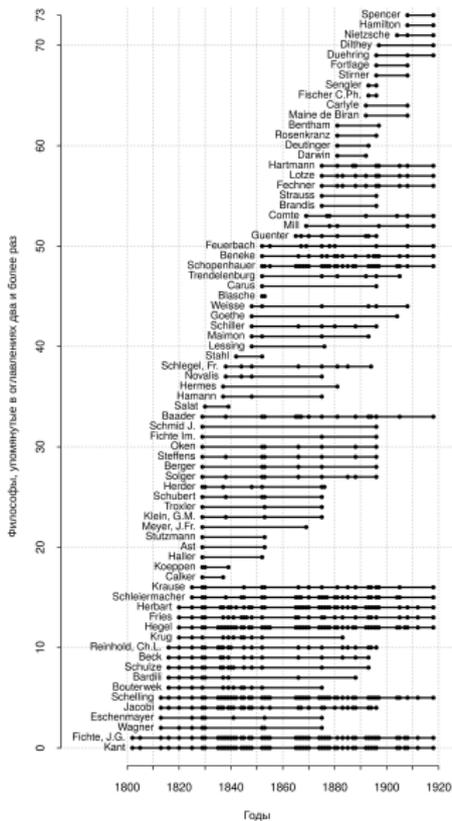
Distant reading

Немецкая философия XIX в.



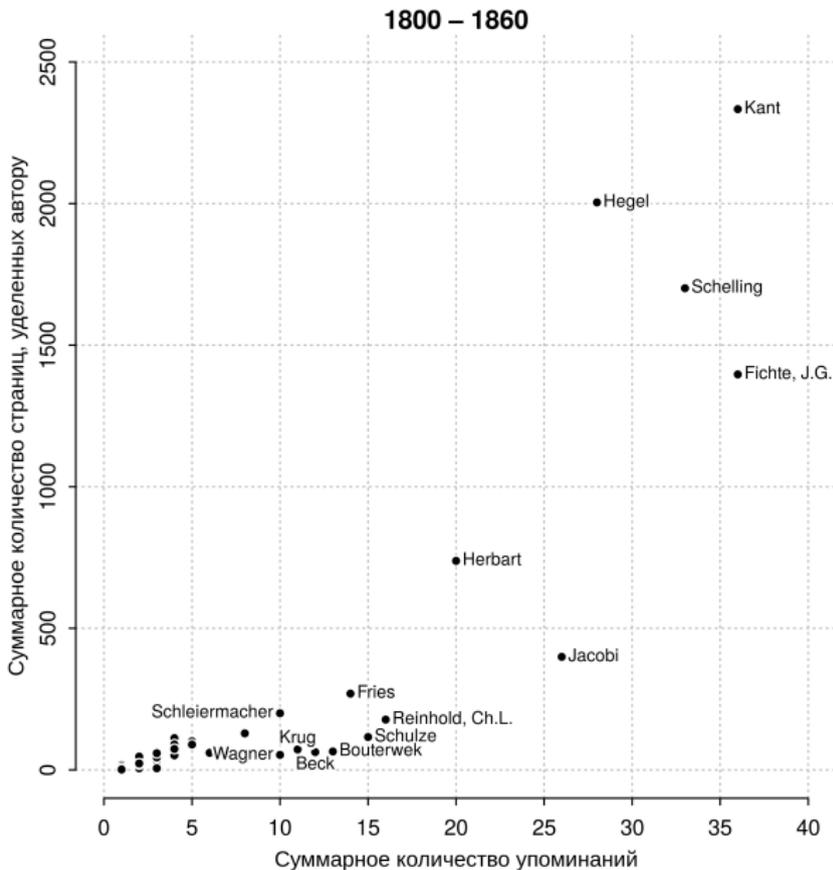
Distant reading

Немецкая философия XIX в.



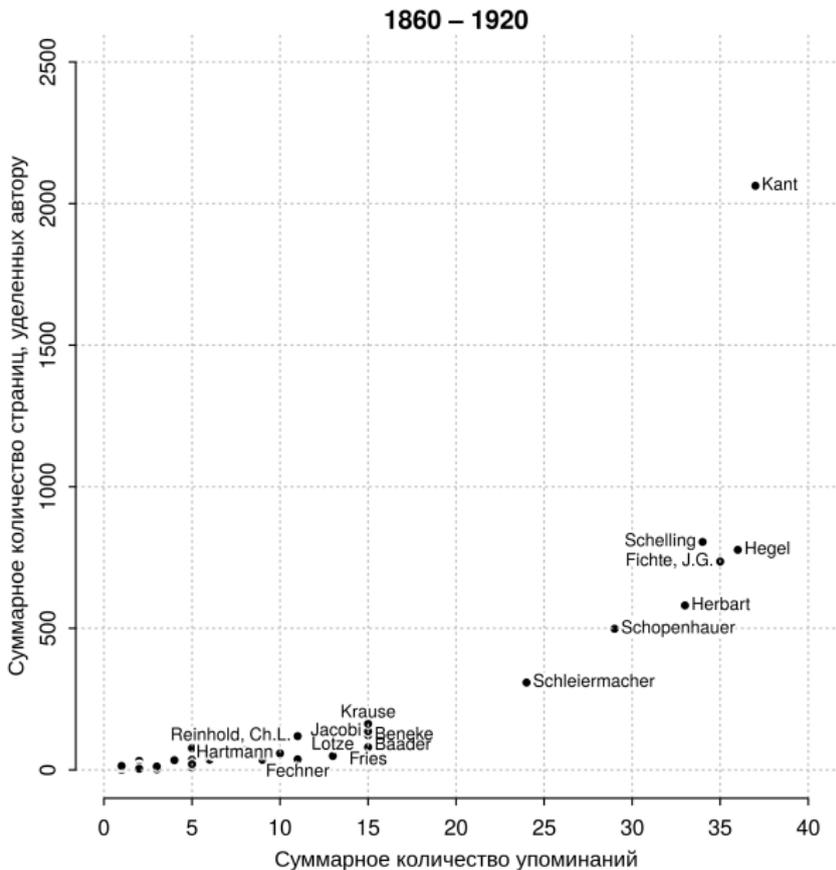
Distant reading

Немецкая философия XIX в.



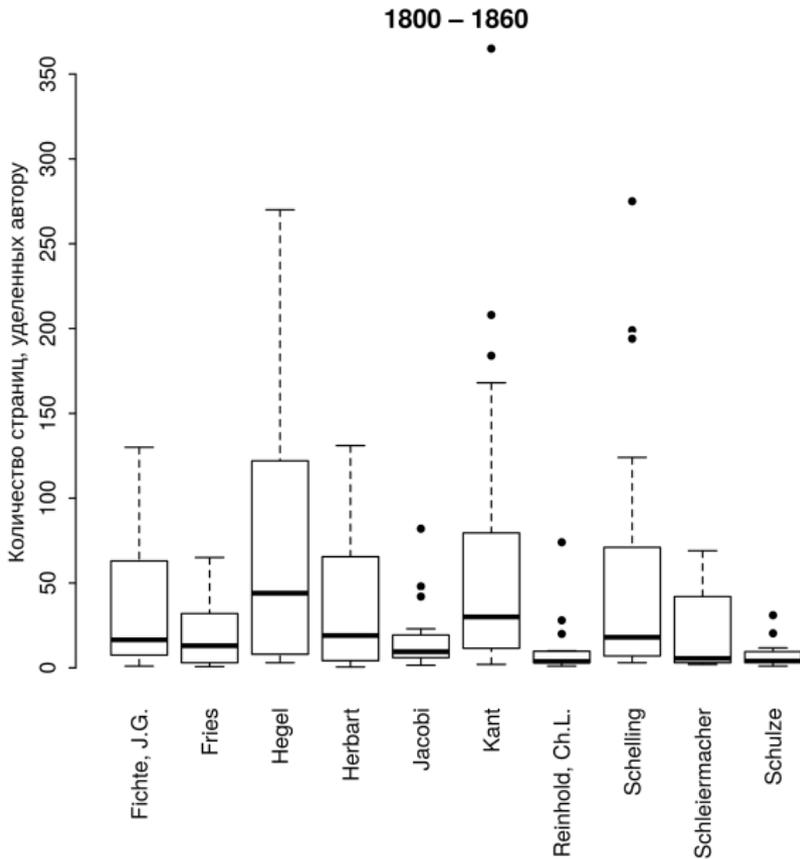
Distant reading

Немецкая философия XIX в.



Distant reading

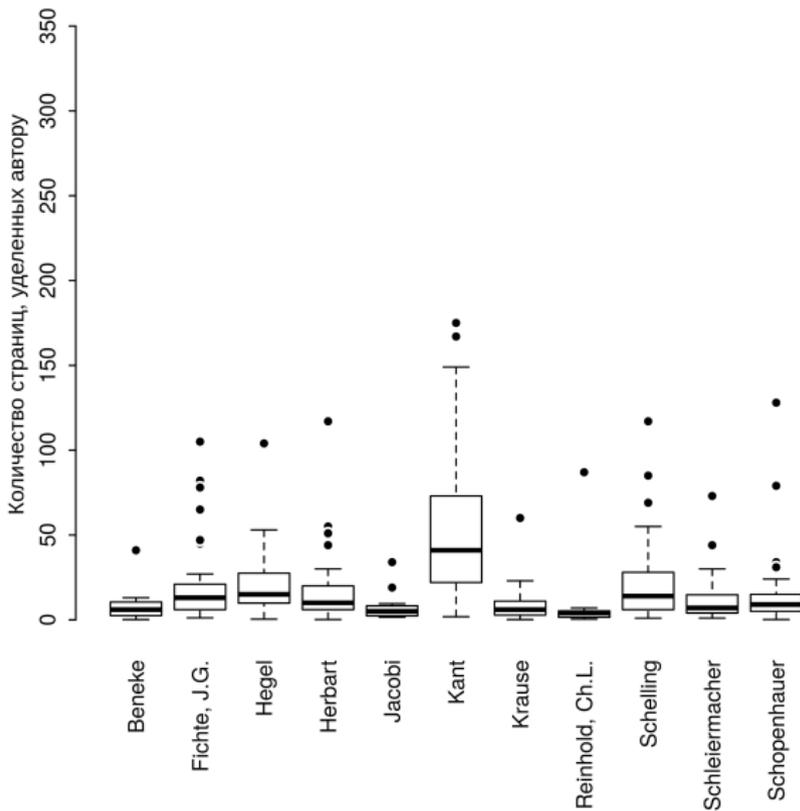
Немецкая философия XIX в.



Distant reading

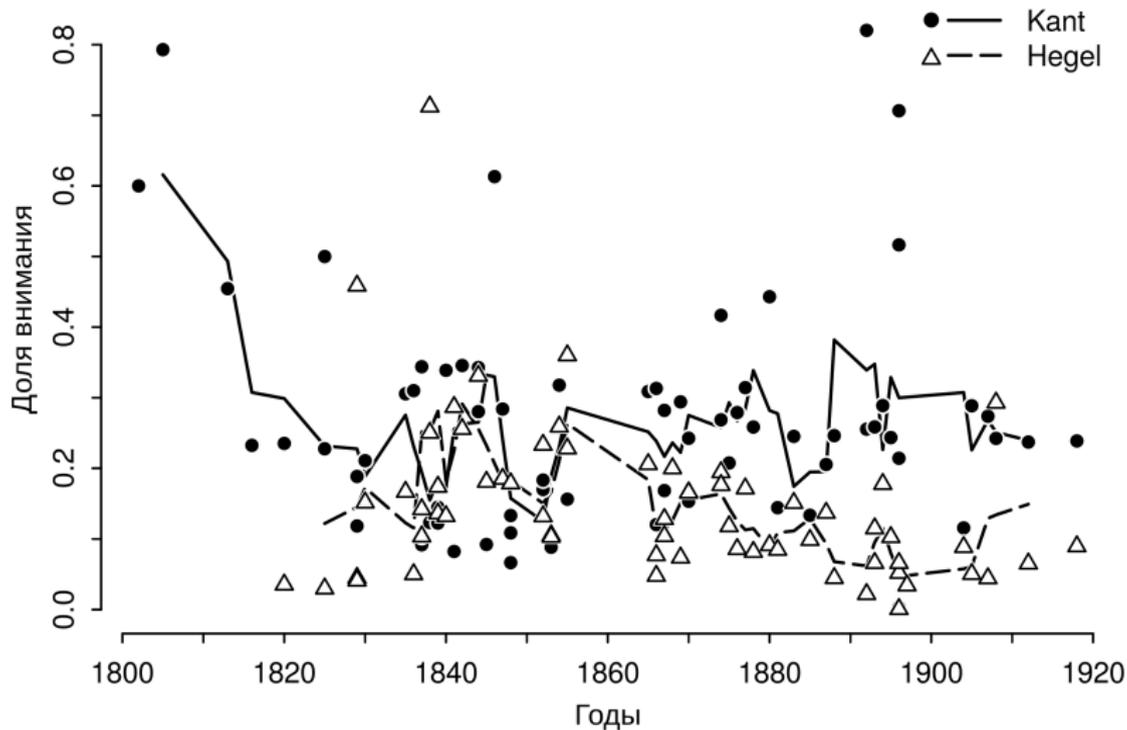
Немецкая философия XIX в.

1860 – 1920



Distant reading

Немецкая философия XIX в.



Tesseract training

Дореформенная орфография

The screenshot shows the Tesseract GUI interface. The main window displays a document page with the following text:

2 СЛОВЕСНОСТЬ

составляет высокое отличие человека
отъ всѣхъ земнородныхъ существъ и
принадлежитъ всѣмъ людямъ: потому
каждый народъ говоритъ известнымъ
языкомъ, своимъ собственнымъ либо пе-
ренятымъ отъ другихъ народовъ, и
каждому народу вродено имѣть свою
Словесность. Но какъ раскрытіе души

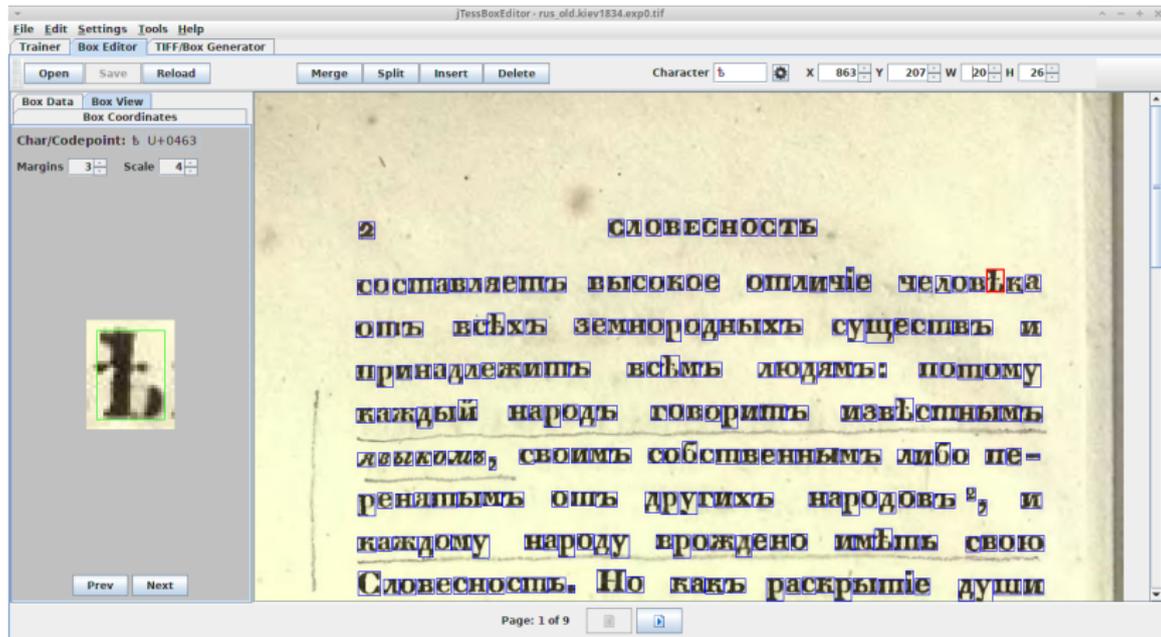
On the left side, there is a 'Box Data' table with the following columns: Char, X, Y, Width, Height. The table contains 26 rows of data, corresponding to the characters in the document text.

Char	X	Y	Width	Height	
13	c	121	219	16	17
14	o	141	218	16	17
15	c	163	218	15	17
16	т	181	217	30	17
17	a	213	217	15	17
18	в	230	217	20	16
19	л	252	217	15	16
20	я	272	216	16	17
21	e	293	215	15	18
22	т	311	215	30	18
23	ь	343	216	24	17
24	в	393	215	17	17
25	ы	416	215	26	17
26	c	445	214	17	18
27	o	465	215	17	17
28	к	486	214	19	17
29	o	509	213	18	18
30	e	531	213	16	19
31	o	580	213	19	19
32	т	601	213	30	18
33	л	634	214	17	18
34	и	654	215	20	18
35	ч	677	214	18	18
36	і	697	204	9	28
37	e	709	213	17	19
38	ч	759	214	20	18
39	e	782	215	16	18
40	л	802	214	18	20
41	o	822	215	18	18
42	в	842	216	18	16
43	к	863	208	20	26
44	к	886	216	19	18
45	a	908	213	18	19

Page: 1 of 9

Tesseract training

Дореформенная орфография



Цифровая гуманитария

Проблемы развития

- ▶ Отсутствие стабильности (кратковременные контракты – короче срока реализации инфраструктурных проектов, хаотические изменения образовательных программ в отсутствие ясных стратегических целей, отношение к преподавателям и исследователям как к легко заменимым винтикам)
- ▶ Проблемы идентичности (не все идентифицируют себя с Digital Humanities, нет удачного русского эквивалента [предложение: цифровая гуманитария])
- ▶ Организационная слабость движения (дисциплины в учебных планах, периодическое издание, исследовательские подразделения [?])
- ▶ Сложности выстраивания равноправного партнерства

Цифровая гуманитарика

Проблемы развития

Сотрудники - Куприянов Алексей Валерьевич — Национальный исследовательский университет «Высшая школа экономики» - Mozilla Firefox

Сотрудники - Куприянов А... X +

https://www.hse.ru/org/persons/202735

Национальный исследовательский университет «Высшая школа экономики» — О Выше — Преподаватели и сотрудники — Куприянов Алексей Валерьевич

Кампус в Москве • RU EN

Куприянов Алексей Валерьевич

Доцент: НИУ ВШЭ в Санкт-Петербурге / Санкт-Петербургская школа социальных наук и востоковедения / Департамент социологии

*Начал работать в НИУ ВШЭ в 2003 году.
Научно-педагогический стаж: 22 года.*

Домашняя страница Преподавание Публикации и исследования Опыт работы Прочее

[В новостях](#) [В СМИ](#)

Образование, учёные степени

2005 Кандидат биологических наук. Институт истории естествознания и техники им. С.И. Вавилова РАН, специальность 07.00.10 «История науки и техники», тема диссертации: Развитие принципов систематики в XVII–XVIII вв.

2005 Аспирантура: Институт истории естествознания и техники им. С.И. Вавилова РАН, специальность «кандидат биологических наук»

1992 Специалитер: Санкт-Петербургский государственный университет, специальность «биология»



Владение языками
английский
немецкий

Контакты
Телефон:
+7 (812) 560-04-43
Электронная почта:
akupriyanov@hse.ru,
alexei.koupriyanov@gmail.com
Адрес: г. Санкт-Петербург,
Набережная Канала Грибоедова
123, каб. 212
Время консультаций: пт, 13:40 –
15:10

Цифровая гуманитарика

Перспективы

- ▶ Поддержание разнообразия (не замыкаться на text-mining и корпусной лингвистике)
- ▶ Больше инфраструктурных проектов
- ▶ Движение в сторону FOSS и Open Data
- ▶ Вкладываться в развитие методологии
- ▶ Продвижение идентификационных маркеров, журнал, учебные дисциплины, учебник

Appendices

Abbreviations used

AHE : Royal Academy of Åbo / Alexander I University in Helsingfors

WLN : Principal School of the Gr. Duchy of Lithuania / Wilno Univ.

WR1 : Royal University of Warsaw (1816–1831)

WR2 : University of Warsaw (1870 on)

DPT : Dorpat University

KAZ : Kazan University

MOS : Moscow University

NIK : Nicholas I University in Saratov

NVR : Nova Russia University in Odessa

PRM : Perm branch of the St. Petersburg University / Perm Univ.

SPB : Pedagogical Institute / St. Petersburg University

USV : St. Vladimir university in Kiev

TOM : Tomsk University

KHR : Kharkiv University

Appendices

Author's e-mail and websites

`alexei.kouprianov@gmail.com`

`https://hse-spb.academia.edu/AlexeiKouprianov`

`https://github.com/alexei-kouprianov/`

`https://www.hse.ru/org/persons/202735`