

От догматического метода к  
векторному  
(векторному анализу)



# О чём доклад: кратко



1. Несколько слов о догматическом (формально-юридическом) методе анализа
  2. От дескриптивной к корпусной лингвистике
  3. Векторная лингвистика
  4. Векторная юриспруденция
-

# 1. С. С. Алексеев



- «догматический метод связан с использованием правил логики и языка, поскольку бытие права неотделимо от текстов законов»
- Алексеев С. С. О составе гражданского правонарушения // Известия высших учебных заведений. Правоведение, 1958. № 1. С. 47-53.



# 1. С. С. Алексеев



- О «потенции в науке права»
  - О необходимости «приглядеться к праву с тех же общенаучных позиций, которые характерны для естественных и технических наук»
  - Алексеев С. С. Восхождение к праву. Поиски и решения. М.: Издательство НОРМА, 2001. 752 с.
-

# 1. Советские и российские правоведы: анализ права (текстов)



Основные методы анализа :

- «в виде формально юридического метода, юридино-технических разработок»
- Тарасов Н. Н. Методологические проблемы юридической науки. Екатеринбург: Изд-во Гуманит. Ун-та, 2001. 364 с.
- А также: формального, логического, а, фактически, ДОГМАТИЧЕСКОГО метода.



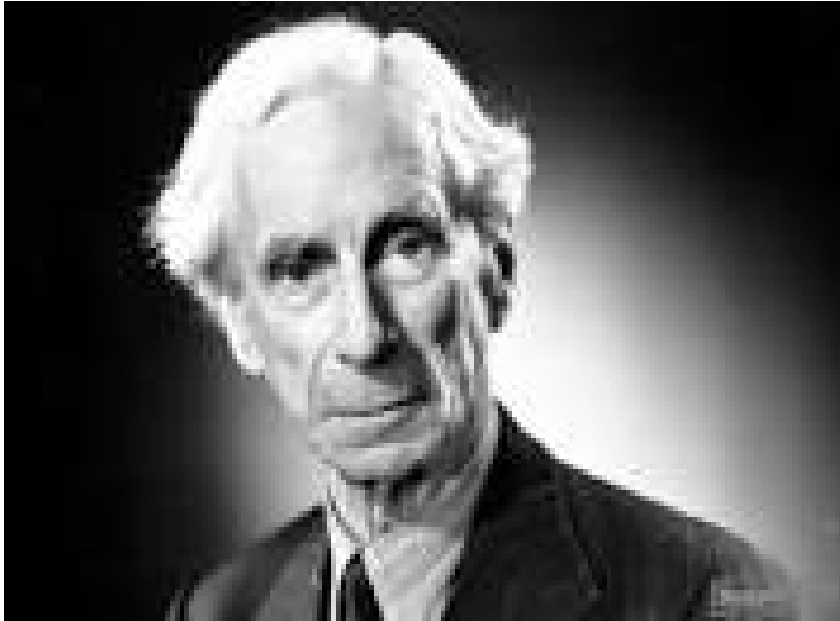
# 1. И. А. Кузьмин о формально-юридическом методе



- форма, конструкция, содержание «формальной системы или исчисления»,
- а также «система подходов, принципов и приемов познания».
- Кузьмин И. А. Формально-юридический метод в составе методологии компаративистского исследования // Право. Журнал Высшей школы экономики. 2017. № 2. С. 18-29.

# 1. Бертран Рассел

18.05.1872-02.02.1970



Математика связана с науками, логика с греками.... логика есть юность математики, а математика зрелость логики.

Рассел Б. Введение в математическую философию. - Новосибирск. 2007. С. 211.



## 2. Дескриптивная (корпусная) ЛИНГВИСТИКА



- Дескриптивная лингвистика, начало положено Бертраном Расселом.
- Новое наименование (с внедрением компьютеров) и новое направление — корпусная лингвистика.
- Корпусная лингвистика «раздел прикладной лингвистики, занимающийся разработкой общих принципов построения и использования лингвистических корпусов при помощи компьютеров».
- Щипицина Л. Ю. Информационные технологии в лингвистике. М.: ФЛИНТА: Наука, 2013. 128 с.
- 





## 2. Корпусная лингвистика - учебная дисциплина



- Корпусная лингвистика развивается уже и как учебная дисциплина, обретая математическую платформу .
  - Корпусы - «тексты + их разметка» Щипицина Л. Ю. Информационные технологии в лингвистике. М.: ФЛИНТА: Наука, 2013. 128 с.
  - 
  - Захаров В. П., Богданова С.Ю. Корпусная лингвистика: Учебник для студентов направления «Лингвистика». 2-е изд., перераб. и дополн., – СПб.: СПбГУ. РИО. Филологический факультет, 2013. 148 с.
  - Батура, Т. В. Математическая лингвистика и автоматическая обработка текстов: учеб.пособие / Т. В. Батура. Новосибирск: РИЦ НГУ, 2016. 166 с.
-

# 2. Национальный корпус русского языка — [www.ruscorpora.ru](http://www.ruscorpora.ru)



Avast SafeZone | Информационное право | Информационное право | национальный корпус ru | Поиск в корпусе. Национ | instruction-main.pdf

ruscorpora.ru/search-main.html

**НАЦИОНАЛЬНЫЙ КОРПУС РУССКОГО ЯЗЫКА**

[главная](#) | [Основной корпус](#) | [инструкция](#) | [задать подкорпус](#) | [сбросить подкорпус](#) | [English](#)

**основной** | Поиск точных форм

– корпус | Слово или фраза

– биграммы |

– триграммы | **передача информации**

– 4-граммы |

– 5-граммы

**синтаксический** | Лексико-грамматический поиск

**газетный** | Слово     | Грамм. признаки  [выбрать](#) | Семант. признаки  [выбрать](#)

**параллельный** |

**обучающий** | Доп. признаки  [выбрать](#) | Словообразование [выбрать](#) |  1-е знач.  др. знач.  фильтр 1  фильтр 2

**диалектный** |

**поэтический** | Расстояние: от  до

**устный** | Слово     | Грамм. признаки  [выбрать](#) | Семант. признаки  [выбрать](#)

**акцентологический** |

**мультимедийный** | Доп. признаки  [выбрать](#) | Словообразование [выбрать](#) |  1-е знач.  др. знач.  фильтр 1  фильтр 2

**мультипарк** |

**исторический** |

**использование корпуса** | Работы в 2015—2016 г. выполнены при поддержке РГНФ, проект № 15-04-12018 «Развитие специализированных модулей НКРЯ».

Национальный корпус русского языка © 2003–2018 | Поиск осуществляется системой [Яндекс.Сервер](#)

# 2. Национальный корпус русского языка — [www.ruscorpora.ru](http://www.ruscorpora.ru)



Avast SafeZone | Информационное | Информационное | национальный ко | Поиск в корпусе. | Национальный ко | Национальный ко | Национальный ко | instruction-main.pd

search1.ruscorpora.ru/search.xml

**НАЦИОНАЛЬНЫЙ КОРПУС РУССКОГО ЯЗЫКА**

Результаты поиска в основном корпусе [перейти на страницу поиска](#) [сбросить подкорпус](#) [выбрать подкорпус](#) [версия с ударениями](#) [настройки](#) [формат KWIC](#) [English](#)

Объем всего корпуса: 115 645 документов, 23 803 881 предложение, 283 431 966 слов.

Поиск ведётся по пользовательскому подкорпусу объемом 751 документ, 78 758 предложений, 1 355 827 слов.

Найдено 751 документ, 3 004 вхождения.

[Распределение по годам](#) [Статистика](#)

Поискать в других корпусах: [акцентологическом](#), [газетном](#), [диалектном](#), [мультимедийном](#), [обучающем](#), [параллельном](#), [поэтическом](#), [синтаксическом](#), [устном](#).

Страницы: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [следующая страница](#)

1. [Г. Шаров. Федеральный закон «Об адвокатской деятельности и адвокатуре в Российской Федерации» о страховании риска ответственности адвоката \(1 часть\) \(2004\) // «Адвокат», 2004.12.01](#) [омонимия снята] [Все примеры \(0\)](#)
2. [Надежда Семенов. Служебная записка о хранении оргтехники на кафедре \(2003\)](#) [омонимия снята] [Все примеры \(0\)](#)
3. [Должностная инструкция заместителя руководителя структурного подразделения \(2003\)](#) [омонимия снята] [Все примеры \(0\)](#)
4. [Договор с колледжем \(2002\)](#) [омонимия снята] [Все примеры \(0\)](#)
5. [Приказ \(2002\)](#) [омонимия снята] [Все примеры \(0\)](#)
6. [Текущая информация о проектах законодательных и других нормативных актов \(1999\) // «Адвокат», 1999.01.18](#) [омонимия снята] [Все примеры \(0\)](#)

Windows | 12:47 | 19.05.2018

## 2. Корпус текстов законов



Общее количество документов  
опубликованной версии набора  
данных — 409 тыс.,

объем архивных файлов — 1.8  
Гб,

в распакованном виде — 10.7  
Гб

[https://github.com/irlcode/  
RusLawOD](https://github.com/irlcode/RusLawOD)

Савельев Д. А. О создании и  
перспективах использования  
корпуса текстов российских  
правовых актов как набора  
открытых данных // Право.

Журнал Высшей школы  
экономики. 2018. № 1. С. 26–  
44.



## 2. <https://github.com/irlcode/RusLawOD>



Яндекс x GitHub - irlcode/RusLawOD: Rus x +

github.com/irlcode/RusLawOD

Why GitHub? Enterprise Explore Marketplace Pricing Search Sign in Sign up

irlcode / RusLawOD Watch 4 Star 6 Fork 1

Code Issues 0 Pull requests 0 Projects 0 Insights

**Join GitHub today**  
GitHub is home to over 28 million developers working together to host and review code, manage projects, and build software together.  
[Sign up](#)

Russian Law as Open Data

34 commits 1 branch 0 releases 1 contributor View license

Branch: master New pull request Find file Clone or download

denissaveliev Update README.md Latest commit 5076050 20 days ago

use_examples	Added an example of how to make a dataframe from it.	11 months ago
LICENSE.txt	Create LICENSE.txt	a year ago
NEWS.md	Update NEWS.md	2 months ago
README.md	Update README.md	20 days ago
README_RUS.md	Update README_RUS.md	20 days ago
data.txt	added new links to data.txt	2 months ago
description_RUS.txt	Add files via upload	a year ago
md5sum.txt	updated md5sums	2 months ago
most_complicated_sentences.txt	Create most_complicated_sentences.txt	2 months ago

README.md

**RusLawOD : Russian Law Open Data**



### 3. Векторная лингвистика



Ещё одно направление в компьютерной лингвистике, которое основано на расчётах значений семантических соответствий терминов (векторов).

Векторный анализ позволяет представить семантические значения в виде чисел, которые можно сравнивать.

Аналогично тому, как набирается абонент по записи в телефонной «книге» смартфона.



# 3. Векторный корпус русского языка

<https://rusvectors.org> - есть ролик



The image shows a screenshot of a web browser with two tabs open. The top tab is a GitHub repository for 'irlcode/RusLawOD'. The bottom tab is the website 'rusvectors.org/ru/'.

The website 'rusvectors.org/ru/' has a blue navigation bar with the following items: RusVectōrēs, Похожие слова, Визуализации, Калькулятор, Различные операции, Модели, О проекте, Контакты, EN/RU.

A notification banner at the top of the page reads: "18/01/2019 — Выложены новые модели к новому 2019 году; подробное описание изменений."

## RusVectōrēs: семантические модели для русского языка

*'You shall know a word by the company it keeps.'* (Firth 1957).

Введите слово, чтобы получить список из 10 его ближайших семантических ассоциатов (квази-синонимов). Будет использована модель, обученная на Википедии и Национальном корпусе русского языка; другие модели вы можете найти на вкладке [Похожие слова](#).

[Найти похожие слова!](#)

Below the search area, there are two blurred boxes with text: "Семантические аналоги для **мор...**" and "НКРЯ и Wikipedia".

# 3. Векторный корпус русского языка



## Вычисление семантической близости

Введите через пробел 2 слова, чтобы вычислить их семантическое сходство. Можно также ввести несколько пар, разделяя их запятыми, как в примере.

справедливость\_NOUN правосудие\_NOUN

Выберите модель:

Araneum fastText  Новостной корпус  НКРЯ  Тайга  НКРЯ и Wikipedia

Вычислить семантическую близость!

Пары слов	Косинусная близость (НКРЯ и Wikipedia)
<i>справедливость правосудие</i>	0.492

## История запросов близости

<i>справедливость право</i>	0.349
-----------------------------	-------

Слова, выделенные **зеленым**, являются высокочастотными (доля в корпусе выше 0.00001); слова, выделенные **красным**, являются низкочастотными (доля в корпусе ниже 0.0000005).



# 3. Векторный корпус русского языка



## Вычисление семантической близости

Введите через пробел 2 слова, чтобы вычислить их семантическое сходство. Можно также ввести несколько пар, разделяя их запятыми, как в примере.

справедливость\_NOUN право\_NOUN

Выберите модель:

Araneum fastText  Новостной корпус  НКРЯ  Тайга  НКРЯ и Wikipedia

Вычислить семантическую близость!

Пары слов

Косинусная близость  
(НКРЯ и Wikipedia)

*справедливость право*

0.349

Слова, выделенные **зеленым**, являются высокочастотными (доля в корпусе выше 0.00001); слова, выделенные **красным**, являются низкочастотными (доля в корпусе ниже 0.0000005).

### 3. «Передавать» vs «Распространять»



- 1. передавать – сообщать 0.440
  - 2. передавать – пересылать 0.424
  - 3. передавать – воспроизводить 0.406
  - 4. передавать – переводить 0.402
  - 5. передавать – информировать 0.401
  - 6. передавать – хранить 0.395
  - 7. передавать – извещать 0.378
  - 8. передавать – уведомлять 0.371
  - 9. передавать – пересказывать 0.353
  - 10. передавать – разглашать 0.338
  - 11. передавать – создавать 0.334
  - 12. передавать – получать 0.327
  - 13. передавать – доводить 0.327
  - 14. передавать – получать 0.327
  - 15. передавать – рассказывать 0.301
  - 16. передавать – показывать 0.296
  - 17. передавать – транслировать 0.291
  - 18. передавать – распространять 0.283
- Волков Ю. В. Передача информации как правоотношение / Законность в современном обществе: Сборник статей Международной научно - практической конференции (13 апреля 2017 г., г. Оренбург). В 2-х частях. Ч.1. – Уфа: АЭТЕРНА, 2017. С. 69-71.
  - Волков Ю. В. Анализ юридического термина «передача информации» / Приемственность и новации в юридической науке: Сборник статей международной научно-практической конференции (г. Самара, 04 марта 2018 года). – Стерлитамак: АМИ, 2018. – С. 29-32.



3. Подбор синонимов (может выполняться по словарям, по информационным сетевым ресурсам, а также произвольно - экспертом)

- Для термина «право» - 22
- Для термина «справедливость» - 61



# 3. Векторный корпус русского языка



Найти похожие слова!

## Семантические аналоги для *справедливость* (NOUN)

### НКРЯ и Wikipedia

1. справедливости 0.68
2. несправедливость 0.57
3. гуманность 0.56
4. честность 0.56
5. **здравость** 0.56
6. человеколюбие 0.56
7. беспристрастие 0.55
8. законность 0.54
9. **равенство** 0.54
10. истина 0.54



Слова, выделенные **зеленым**, являются высокочастотными (доля в корпусе выше 0.00001); слова, выделенные **красным**, являются низкочастотными (доля в корпусе ниже 0.0000005).

право\_NOUN



Выберите модель:

Araneum fastText  Новостной корпус  НКРЯ  Тайга  НКРЯ и Wikipedia

Показывать только:

Прилагательные  Существительные  Глаголы  Имена собственные  Наречия  Все части речи  Часть речи запроса

Найти похожие слова!

## Семантические аналоги для **право** (NOUN)

### НКРЯ и Wikipedia

1. **права** 0.72
2. **прав** 0.65
3. **правами** 0.58
4. **привилегия** 0.56
5. **правомочие** 0.53
6. **праву** 0.53
7. **правоспособность** 0.52
8. **правах** 0.52

# Вычисление семантической близости



Введите через пробел 2 слова, чтобы вычислить их семантическое сходство. Можно также ввести несколько пар, разделяя их запятыми, как в примере.

справедливость\_NOUN право\_NOUN

Выберите модель:

Araneum fastText  Новостной корпус  НКРЯ  Тайга  НКРЯ и Wikipedia

Вычислить семантическую близость!

Пары слов

Косинусная близость

(НКРЯ и Wikipedia)

*справедливость право*

0.349

Слова, выделенные **зеленым**, являются высокочастотными (доля в корпусе выше 0.00001); слова, выделенные **красным**, являются низкочастотными (доля в корпусе ниже 0.0000005).

Вычислить семантическую близость!



Пары слов

Косинусная близость  
(НКРЯ и Wikipedia)

*справедливость* *совестливость*

0.476

## История запросов близости

*справедливость право*

0.349

*справедливость правосудие*

0.492

*справедливость возможность*

0.130

*справедливость власть*

0.274

*справедливость воля*

0.350

*справедливость честность*

0.559

Слова, выделенные **зеленым**, являются высокочастотными (доля в корпусе выше 0.00001); слова, выделенные **красным**, являются низкочастотными (доля в корпусе ниже 0.0000005).

*справедливость законность*

0.544



## История запросов близости

*справедливость правосудие*

0.492

*справедливость возможность*

0.130

*справедливость власть*

0.274

*справедливость воля*

0.350

*справедливость честность*

0.559

*справедливость совесть*

0.476

*справедливость порядочность*

0.520

*справедливость правда*

0.475

*справедливость верность*

0.425

Слова, выделенные **зеленым**, являются высокочастотными (доля в корпусе выше 0.00001); слова, выделенные **красным**, являются низкочастотными (доля в корпусе ниже 0.0000005).



### 3. Векторная (семантическая) близость к термину «справедливость»



- Честность 0,559
- Законность 0,544
- Порядочность 0,520
- .....
- Право 0,349



## 4. Векторная юриспруденция



- Анализ текстов (законы, монографии, статьи): на предмет семантического заимствования; сравнительный анализ/
- Подготовка текстов (законов, нормативных актов, инструкций) на основе заданной алгоритмической модели.
- Приведение в соответствие (консолидация) текстов законов на основе заданных условий (принципов, векторов, текстов).



## 4. Векторная юриспруденция



- Правовая сравнительная экспертиза на основе векторного анализа  
(текст решения оценивается на основе эталона - закона)
- Робот-юрист принимающий решения по типовым делам
- Административное и судебное производство сокращается многократно

## 4. Векторная юриспруденция

### С. С. Алексеев об учебных курсах



- Учебные курсы могут быть «сконструированы» из разнородного материала, относящегося к различным наукам»
- Алексеев С. С. Проблемы теории права: Курс лекций в двух томах. Том второй. Свердловск, СЮИ. 1973. 401 с.

## 4. Векторная юриспруденция



- На основе СУБД — цитатники :
  - 1) по правовым темам;
  - 2) по авторам;
  - 3) по иным критериям.
- Студенты
  - Аспиранты
  - Преподаватели-авторы статей
  - Редакторы
  - Иные субъекты



# Контакты



- Волков Юрий Викторович
- Доцент, к.ю.н.,
- Уральский государственный юридический университет, кафедра информационного права
- [yurii.volkov@usla.ru](mailto:yurii.volkov@usla.ru)
- [www.telecomlaw.ru/about/CV-VYV.pdf](http://www.telecomlaw.ru/about/CV-VYV.pdf)





Спасибо за внимание !

