



JOINT INSTITUTE  
FOR NUCLEAR RESEARCH

**ОБЪЕДИНЕННЫЙ ИНСТИТУТ  
ЯДЕРНЫХ ИССЛЕДОВАНИЙ**



**ЛАБОРАТОРИЯ  
ИНФОРМАЦИОННЫХ  
ТЕХНОЛОГИЙ**  
имени М.Г. Мещерякова

Доклад соискателя на должность младшего научного сотрудника  
ЛИТ им. М.Г. Мещерякова на тему:

**Интеллектуальный анализ научно-технической информации на  
примере патентной документации**

Соискатель: Зрелова Дарья Петровна, м.н.с., ЛИТ, НТО ВКиРИС, Сектор №1  
распределенных систем

**Дубна, 2025 г.**

# Патентный анализ. Актуальность

## Патентный анализ

1

### **Подтверждение практической значимости научных исследований**

Патенты позволяют определить научные направления, обладающие **практической значимостью**, поскольку они фиксируют технологические решения, нашедшие применение в реальных условиях и подтвержденные правовой защитой.

2

### **Выявление технологических трендов**

Патентные данные позволяют отслеживать развитие технологий в различных отраслях, определять перспективные направления и прогнозировать будущие изменения.

3

### **Оценка конкурентной среды**

Анализ патентов конкурентов помогает понять их стратегии, выявить сильные и слабые стороны, а также определить потенциальные угрозы.

4

### **Стандартизация патентных данных**

Единые подходы к регистрации и представлению информации об интеллектуальной собственности позволяют проводить сравнительный анализ между странами.

5

### **Эффективное управление затратами на R&D**

Патентный анализ позволяет выявить «белые пятна» в технологиях, что помогает сосредоточить ресурсы на наиболее перспективных и менее конкурентных областях.

6

### **Поддержка государственной политики**

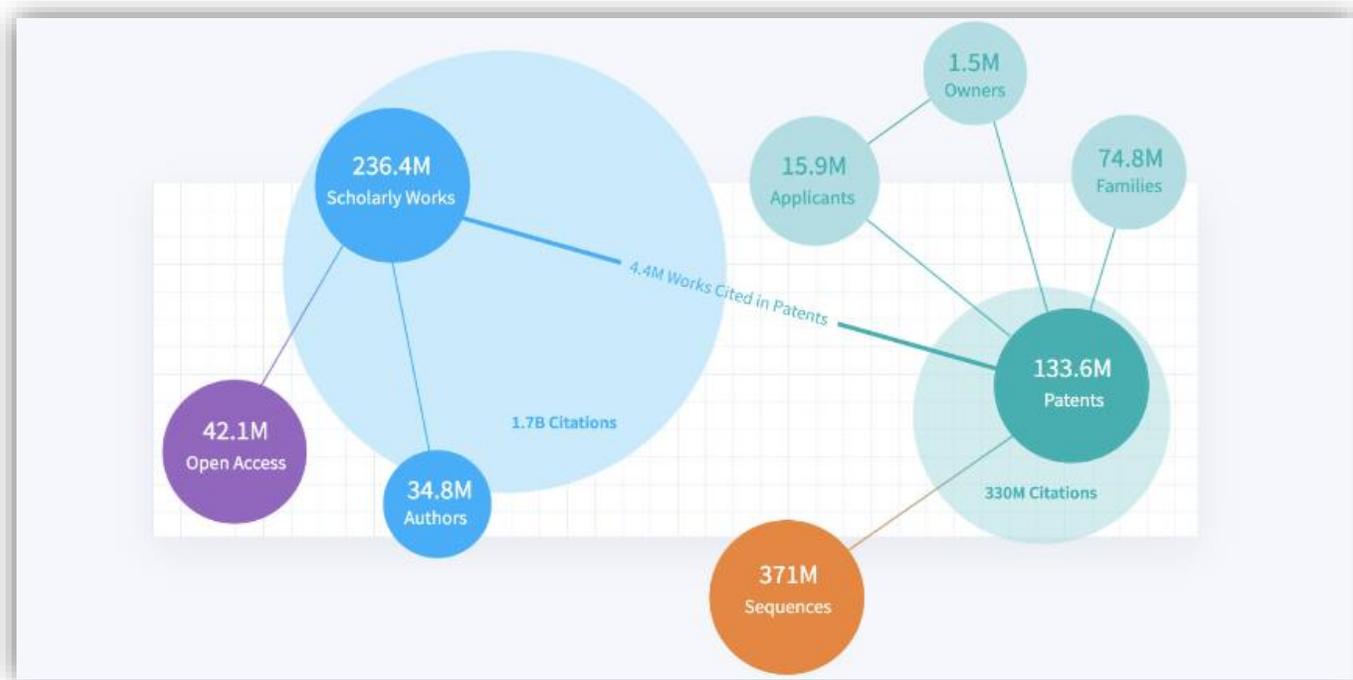
На уровне государства анализ патентов помогает формировать стратегии развития науки и технологий, а также оценивать инновационный потенциал страны.

# Методология построения патентного ландшафта



**! Патентный ландшафт** представляет собой аналитический инструмент, позволяющий выявлять актуальные технологические тренды, оценивать конкурентную среду и идентифицировать перспективные направления для технологического развития.

# Критерии выбора базы данных [Lens.org](https://www.lens.org) для сбора данных



**LENS.ORG**  
Solving The Problem Of Problem Solving™

**Рис. 1.** Патенты, научные работы, их взаимосвязи – количественные показатели базы данных Lens.org

**Источник изображения:** Официальный сайт Lens.org (<https://www.lens.org/lens/collective-action/participate#leapInitiative>).

## Особенности:

Объединяет **патенты** и **научные публикации** в единой платформе.

Позволяет анализировать связи между исследованиями и технологиями.

## Глобальный охват:

Интеграция данных из ведущих патентных ведомств (WIPO, USPTO, EPO) и научных репозиториях.

Доступ к информации из более чем 100 стран.

## Открытость и доступность:

Бесплатный доступ к данным.

Удобный интерфейс для поиска и анализа.

## Актуальность информации:

Регулярное обновление базы.

Данные за последние 30+ лет.

## Функциональность:

Возможность семантического поиска.

\*Поддержка API

\*Ограничения: лимиты на экспорт и запросы через API.

# Алгоритм сбора информации о единичном патенте

Информационный ресурс: **Lens.org**

Поисковый запрос: «robotics AND (intelligent OR quantum)»

Разработан специализированный сборщик для сбора патентной документации с сайта Lens.org.



Инструмент автоматически извлекает данные, обходя ограничение платформы на загрузку 1000 документов за раз.



Полученные данные о патентах организуются в формате JSON для удобного анализа и обработки.

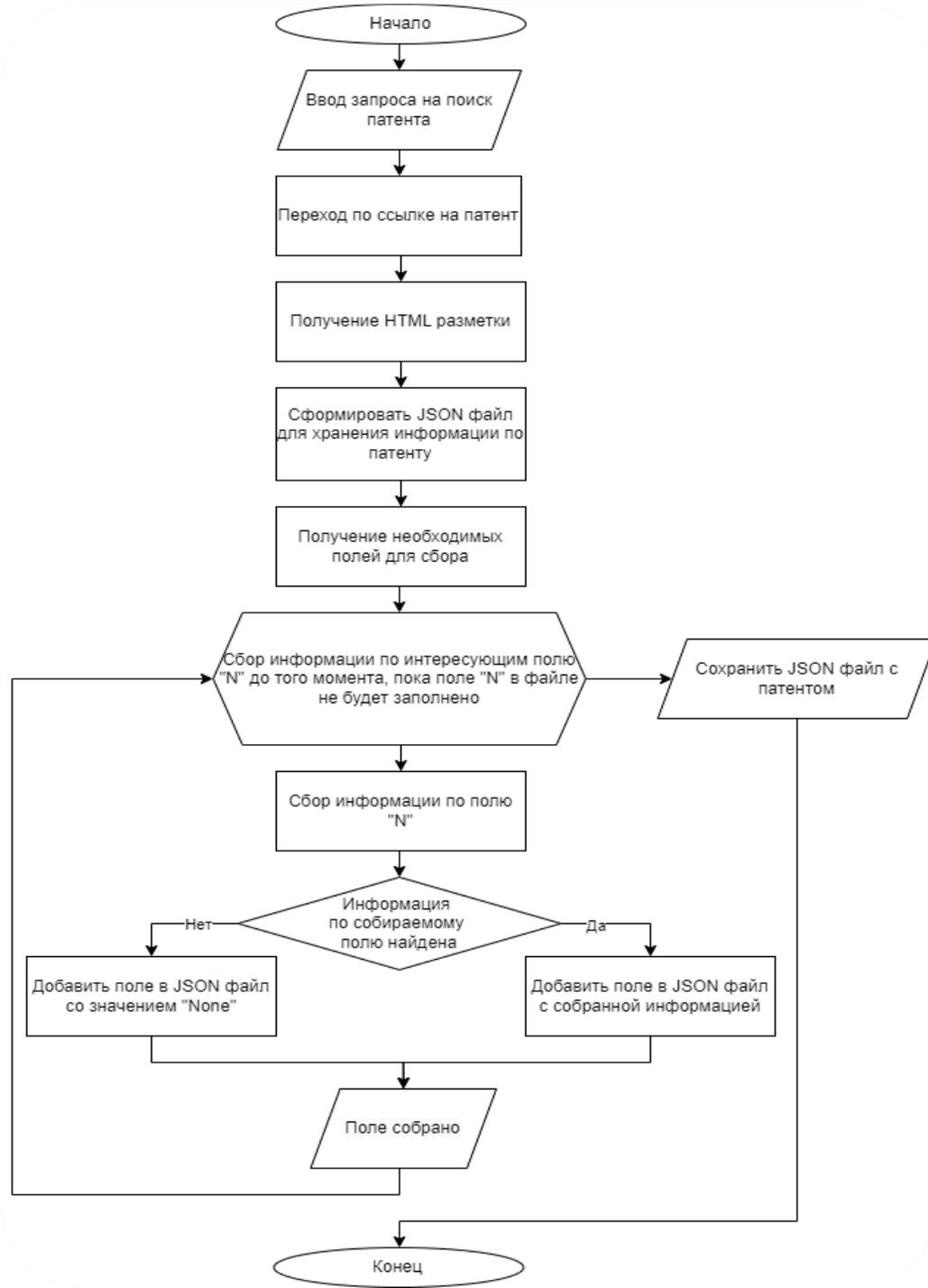


Рис.2. Блок-схема алгоритм сбора информации о единичном патенте

# Структура единичного файла патента

---

Поле	Значение
<b>title</b>	Название патента.
<b>url</b>	URL-адрес патента.
<b>abstract</b>	Краткое описание содержания патента.
<b>patent_id</b>	Информация о стране и номере патента.
<b>patent_application</b>	Детали заявки на патент, включая семейство, юрисдикции, статус, номер, даты подачи, публикации, приоритета и предоставления, а также информацию о заявителях, изобретателях и владельцах.
<b>classifications</b>	CPC и IPC классификации патента.
<b>other_links</b>	Ссылки на документацию патента и другие ресурсы.
<b>cited_by</b>	Ссылки на другие патенты и работы, которые цитируют данный патент.
<b>cited_patent</b>	Ссылки на список патентов, которые были цитированы в данном патенте.
<b>cites works</b>	Ссылки на научные публикации, которые были цитированы в данном патенте.

# Визуализация результатов

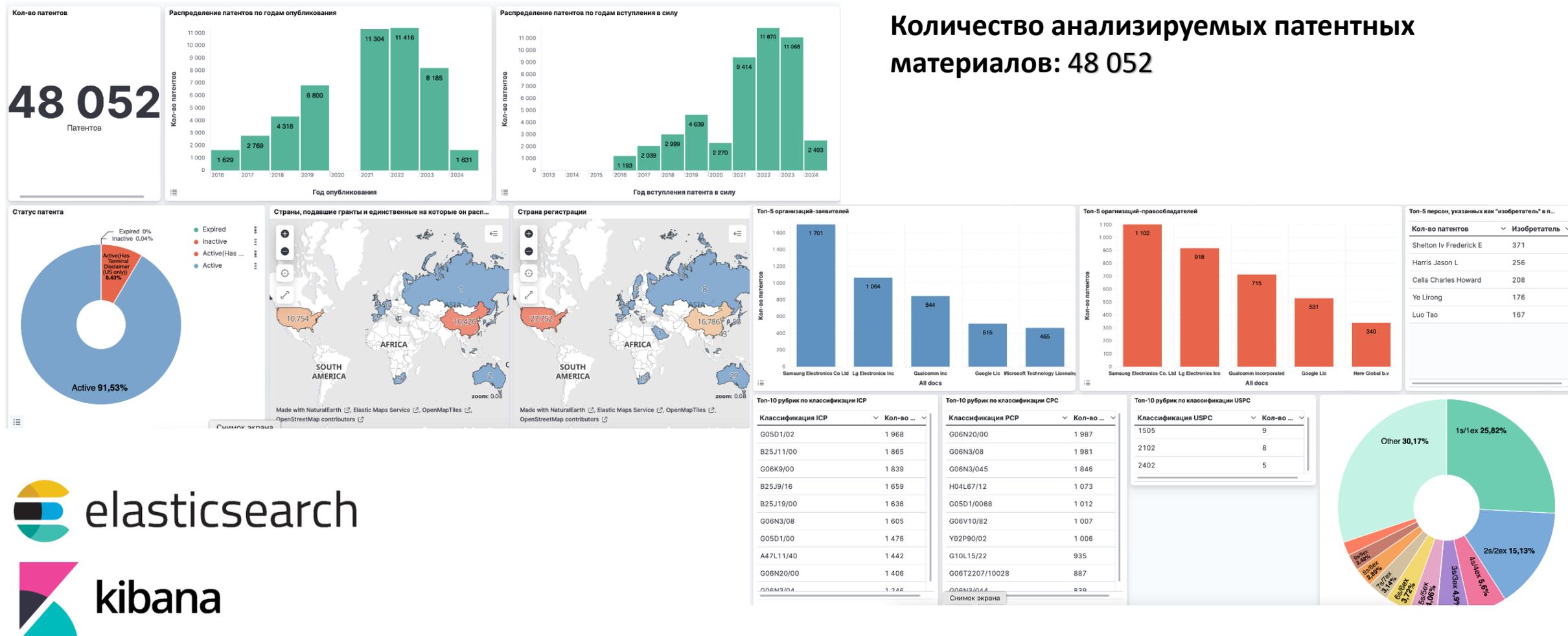


Рис. 3. Интерактивная аналитическая панель патентных данных, реализованная при помощи Kibana

# Результаты анализа

Распределение патентов по годам вступления в силу

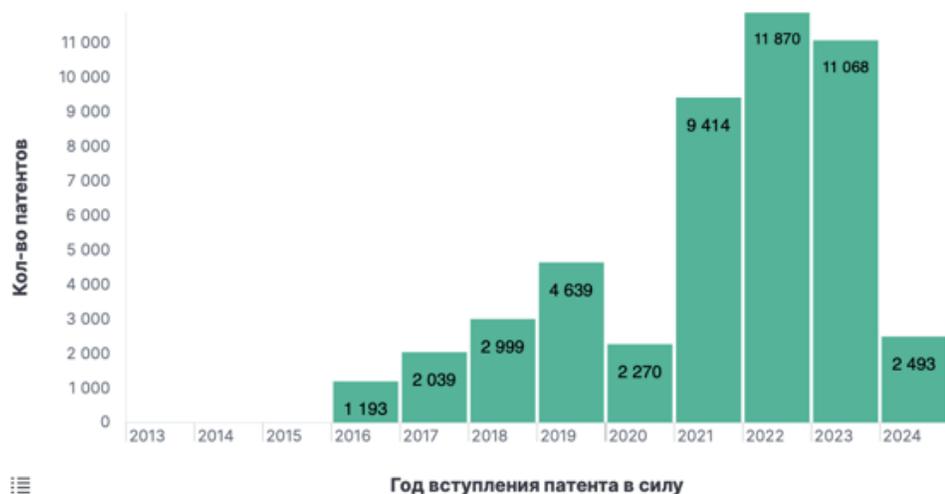


Рис.4. Распределение патентов по годам вступления в силу

Статус патента

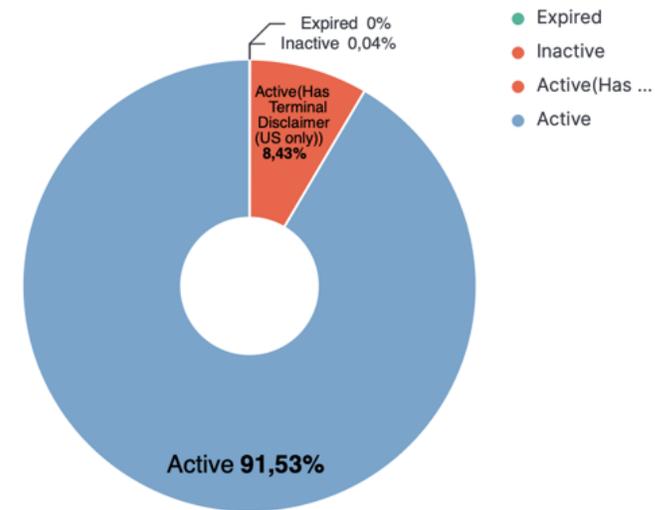


Рис.5. Юридический статус патента

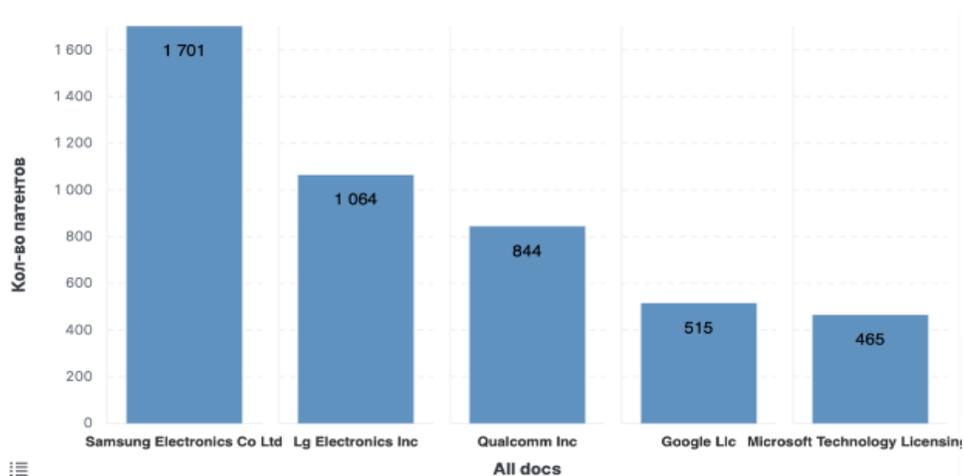


Рис.6. Топ-5 организаций-заявителей

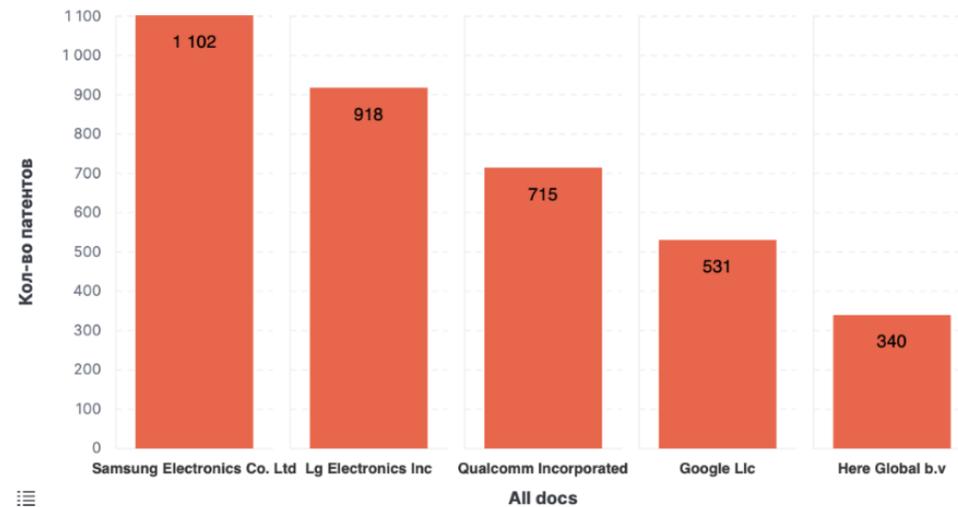
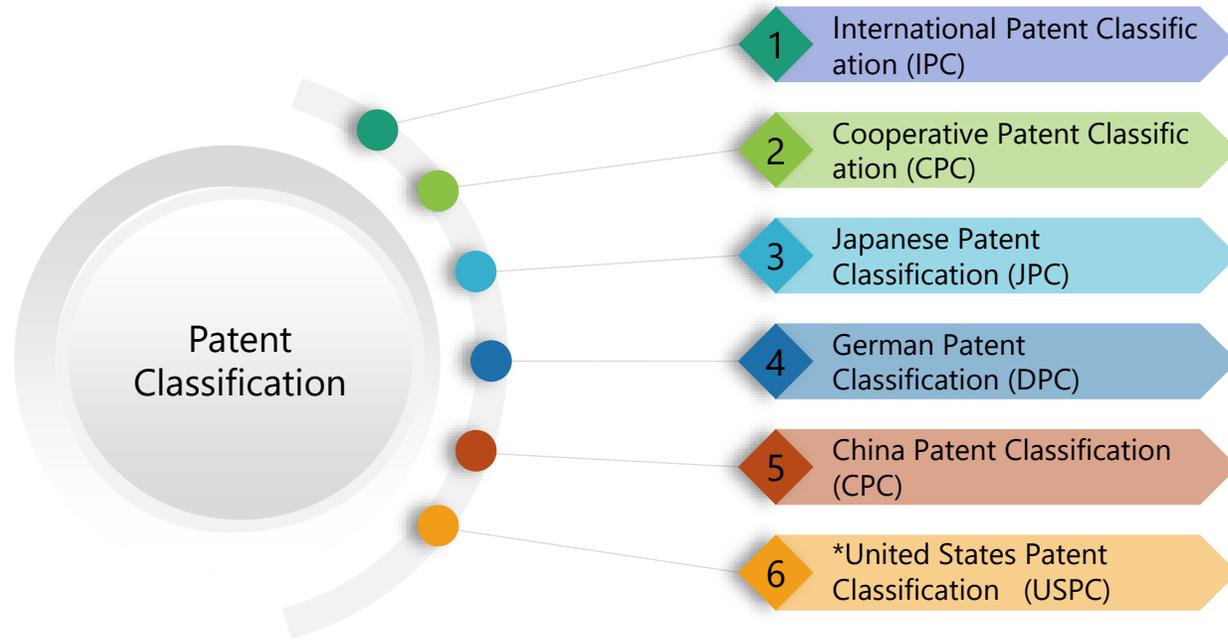


Рис.7. Топ-5 организаций-правообладателей

## Системы патентной классификации



\*Система классификации, использовавшаяся до 2015 года в США, заменена на CPC

## Пример структуры классификационного кода IPC:

- | — **G** — Секция (Физика)
- |   └─ **G21** — Класс (Ядерная физика, ядерная техника)
- |       └─ **G21C** — Подкласс (Ядерные реакторы)
- |           └─ **G21C 7/00** — Группа (Управление ядерной реакцией)
- |               └─ **G21C 7/36** — Подгруппа (Управление ядерной реакцией – схемы управления)

## Основные секции Международной патентной классификации (IPC)

Секция	Описание
<b>A</b>	Удовлетворение жизненных потребностей человека
<b>B</b>	Различные технологические процессы
<b>C</b>	Химия; Metallургия
<b>D</b>	Текстиль; бумага
<b>E</b>	Строительство и горное дело
<b>F</b>	Машиностроение; Освещение; Отопление; Оружие и боеприпасы; Взрывные работы
<b>G</b>	Физика
<b>H</b>	Электричество

### **\*IPC (Международная патентная классификация):**

Универсальная иерархическая система, разработанная *Всемирной организацией интеллектуальной собственности*, применяется **по всему миру** для классификации патентов по областям техники.

# Результаты анализа

Анализ патентных данных по классификации IPC выявил 5 ключевых направлений:

Классификация IPC	Описание	Количество патентов
G05D 1/02	Control of position, course, altitude, or attitude of land, water, air, or space vehicles, e.g., automatic pilot.	1986
B25J 11/00	Manipulators adapted for specific uses (e.g., handling radioactive materials, servicing nuclear reactors).	1865
G06K 9/00	Methods or arrangements for reading or recognising printed or written characters or for recognising patterns.	1839
B25J 9/16	Programme-controlled manipulators (e.g., robots) with means for control of the manipulator by the programme.	1659
B25J 19/00	Accessories fitted to manipulators, e.g., for monitoring, for viewing; Safety devices combined with or specially adapted for use in connection with manipulators.	1636

Анализ патентных данных по классификации IPC выявил, что наиболее активные направления связаны с **автоматизацией управления транспортными средствами, робототехникой для специализированных задач и технологиями распознавания образов.**



# Результаты анализа

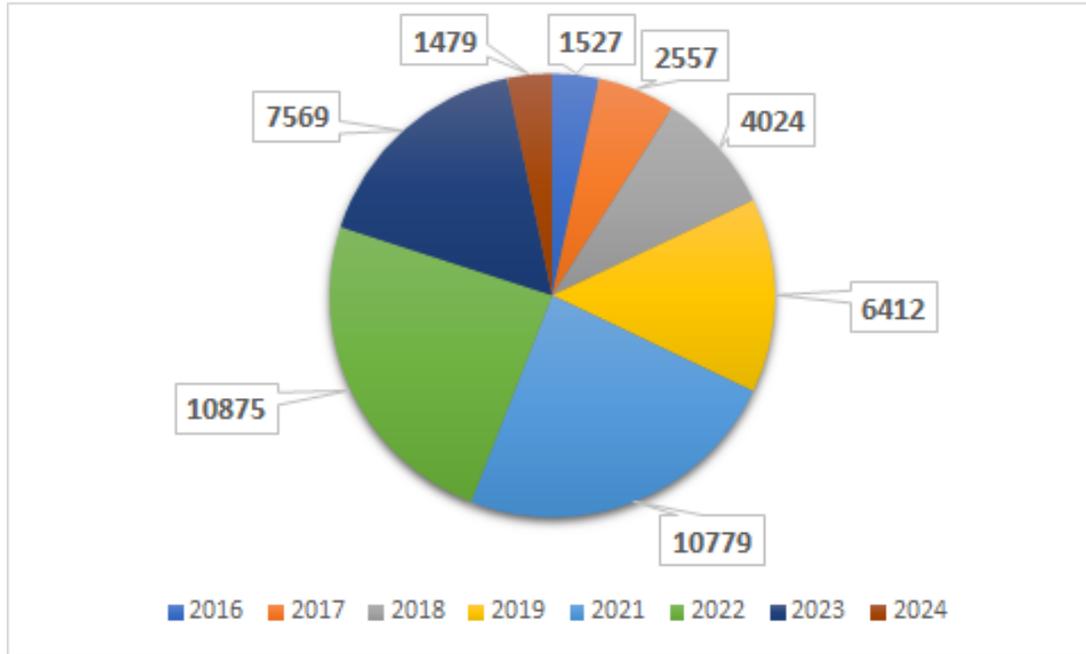
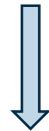


Рис.9. Распределение числа патентов по годам



Год	2016	2017	2018	2019	2021	2022	2023	2024
Патентов	1527	2557	4024	6412	10779	10875	7569	1479

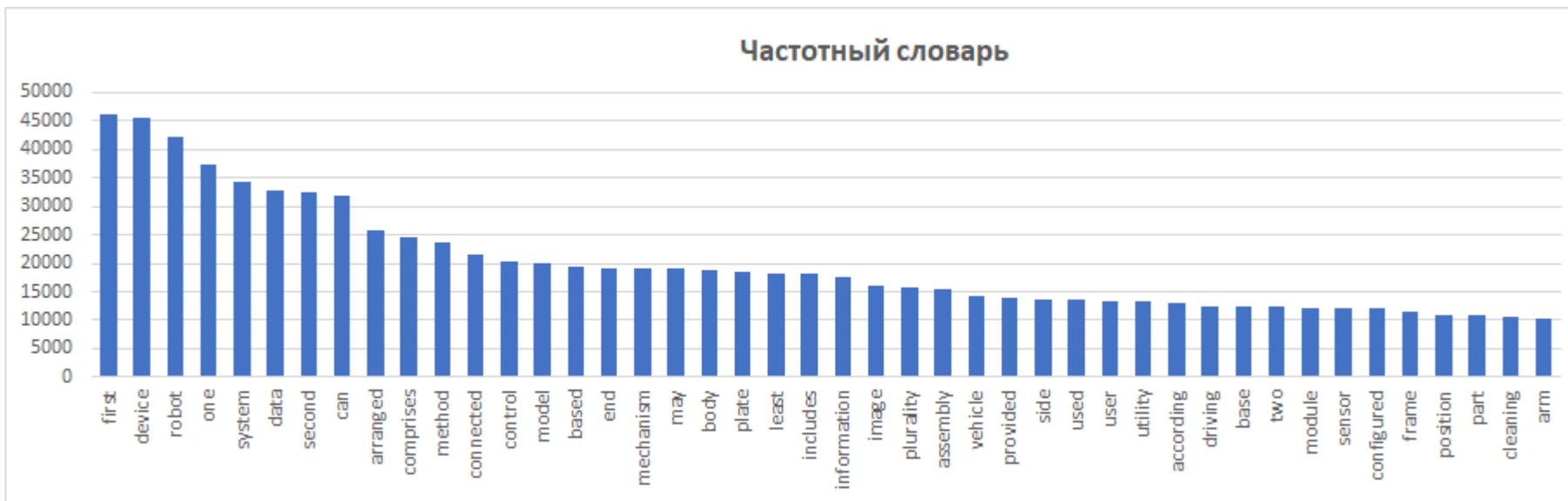


Рис.10. Распределение числа патентов по странам



USA	China	EP	KR	JP	FR	TW	DE	AU	NL	LU	GB	ZA	RU	ES	SE	BE	GR	HR	SA
27752	16786	3166	70	53	43	43	31	29	17	15	13	13	8	4	3	1	1	1	1

# Статистический анализ

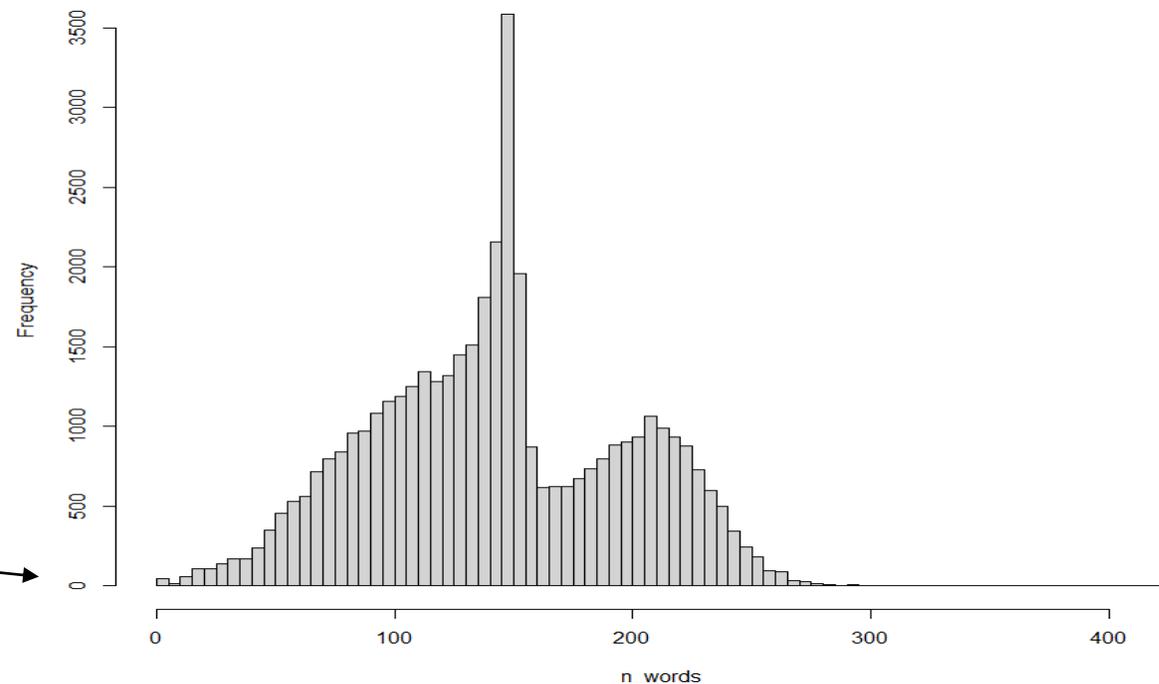


Общее количество слов в словаре: 17 211

Корпус содержит 45 222 патентов  
Мин. количество слов в аннотации патента – 8,  
Макс. количество слов в аннотации – 422,  
Среднее значение – 143,  
Медиана – 142,  
1 квартиль (25%) – 108,  
3 квартиль (75%) – 189

## Примечание

25% аннотаций патентов состоит из  $\leq 108$  слов.  
75% аннотаций патентов состоит из  $\leq 189$  слов.



# Abstracts

Пред-обработка текста:

Удаление служебных символов и знаков препинания, чисел, лишних пробелов, приведение к нижнему регистру, удаление стоп-слов, токенизация, лемматизация и т.д.

1	9430610B2	The present invention relates to pathogen detection and identification by use of DNA resequencing microarrays The present invention also provides resequencing microarray chips for differential diagnosis and serotyping of pathogens present in a biological sample The present inve
2	2341140B1	Disclosed is a method for obtaining a bifunctional complex comprising a molecule linked to a single stranded identifier oligonucleotide wherein a nascent bifunctional complex comprising a chemical reaction site and a priming site for enzymatic addition of a tag is a reacted at the c
3	2336315B1	Disclosed is a method for obtaining a bifunctional complex comprising a molecule linked to a single stranded identifier oligonucleotide wherein a nascent bifunctional complex comprising a chemical reaction site and a priming site for enzymatic addition of a tag is a reacted at the c
4	101997B1	The present invention discloses an anchor supporting intelligent device belonging to the field of electromechanical devices for fully mechanized working faces The device includes four parts of net supporting systems anchoring systems ground supporting systems and power and ti
5	102027B1	The present disclosure provides an intelligent obstacle avoidance method and system of a human robot safe interaction oriented robotic arm The robot monitors an identification marker on the obstacle in real time through an identification marker visual positioning method to obtai
6	102384B1	A flexible end effector of intelligent harvesting robot is disclosed including an end effector body and further a flexible clamping claw The flexible clamping claw is successively provided with a straight fixed section a front inclined section an arc shaped section a rear inclined sectio
7	102392B1	The invention discloses an ultrasonic visualization Da Vinci intelligent robot surgical system is composed of the bedside mechanical arm surgical assembly video imaging system and doctor controlled assembly the bedside mechanical arm surgical assembly is composed of the opera
8	102426B1	The invention discloses a smart agriculture system based on big data which comprises an acquisition module a data center a management module and an execution module The LoRa communication mode is used in the acquisition module data center management module and e
9	1043483B1	The invention discloses a full automatic underground miner
10	1634885B	A method for controlling a walking assistant apparatus incl
11	201918271A	The invention relates to a magnetic intelligent combination toy interactive to Internet of Things IoT The toy comprises a body and a plurality of external parts magnetically assembled with the body The body is respectively provided with a magnetic sensor at a position where the ex
12	1672586B	Disclosed is a power saving system and power saving method for an intelligent robot including a central processing unit a first device group and a second device group When a voltage level of the battery is changed to a second voltage level from the first voltage level the central p
13	201925032A	An unmanned aerial vehicle UAV including a mai
14	1671049B	An automatic device for brewing beverages at lea
15	1665609B	A resident activity recognition method is providec
16	1721390B	A monitoring system a base station and a contrc
17	1736861B	A monitoring system and a control method there

## Vocabulary

Inter  
wounded  
blackboard  
digestion  
assemblie

## Embeddings

Word2Vec →

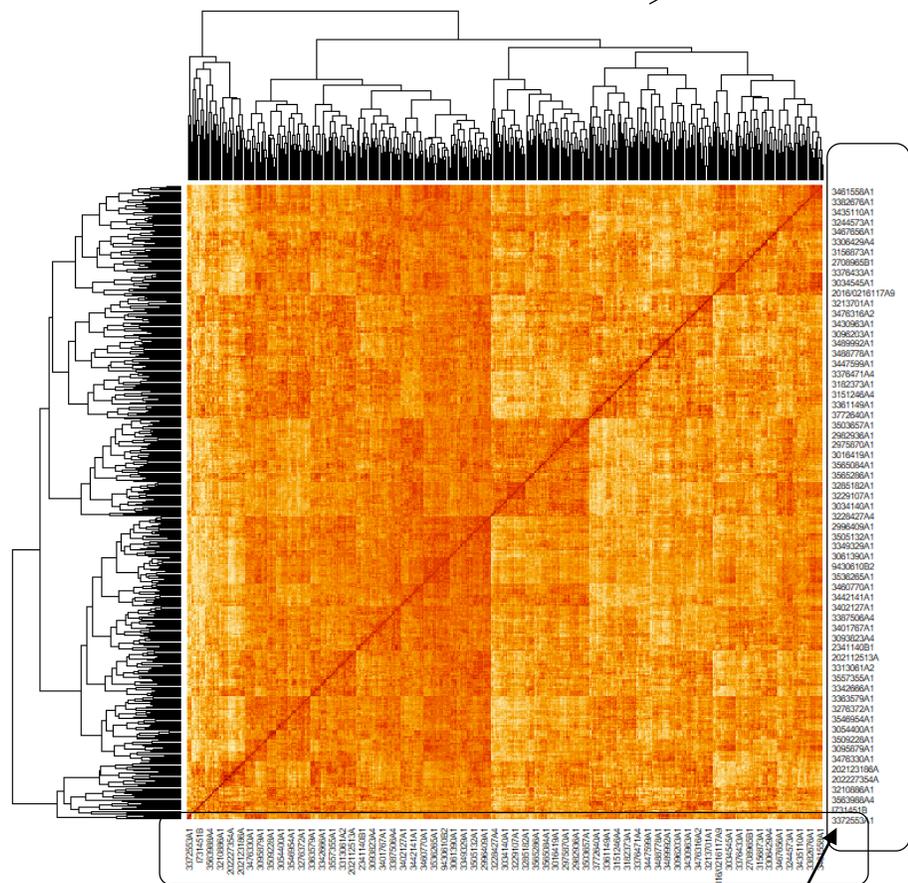
Patent ID

Doc2Vec →

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15
9430610B2	0.275044411	-0.1660728008	-0.3547606766	0.624883175	-2.560036e-01	0.1331577301	-0.044364754	-0.071094655	0.063182421	0.315132678	-0.0682417229	0.080530979	0.048146412	-0.0285942443	-0.415411085
2341140B1	0.305862397	0.2828143537	-0.2483878434	0.600026250	-1.535672e-01	0.1847874671	-0.230152890	-0.180341020	0.092610881	0.383559227	-0.2661279142	0.176091045	0.061994120	-0.0005126456	-0.011981815
2336315B1	0.290316612	0.2591560483	-0.5001454949	0.557459533	-1.196200e-01	0.1075287685	-0.307921380	-0.052677263	-0.033705097	0.289060503	-0.2224047333	0.147058874	0.086190239	0.0160935484	0.026760526
101997B1	0.194547445	0.2714209259	0.1732926071	0.084368050	-5.951387e-01	0.2433299571	-0.070777006	0.076652415	0.529698133	-0.048028775	0.1167405173	-0.263501972	-0.182259366	-0.0878196210	-0.142021224
102027B1	0.019675719	-0.3144090176	-0.1782065332	0.357603997	-3.042061e-01	-0.0473996662	-0.118336767	-0.110348620	-0.032437641	0.449981153	0.3003595471	-0.293948054	-0.280761600	-0.0991601273	-0.389020264
102384B1	0.308348775	0.2521142066	0.1731423587	0.523677588	4.787703e-02	0.0653617308	-0.409541070	-0.099844292	0.130483344	-0.063801296	-0.0752497092	-0.431218952	0.246373251	-0.2786344886	-0.042896569
102392B1	0.081127651	-0.1200899109	-0.4807970524	0.266935319	-2.367833e-01	-0.1448665261	-0.132827953	-0.441822290	0.133470207	-0.064239964	-0.0146643491	-0.434822470	0.243557781	-0.0635000169	-0.334804505
102426B1	0.676775217	-0.1758346856	0.1606598645	0.023210820	-4.520454e-02	-0.0430560596	0.247368217	0.154359415	0.208318427	0.294573188	0.2885634601	-0.180019498	0.024538716	-0.1795592308	-0.342202187
1043483B1	0.302884042	-0.3951968253	0.0949143097	-0.211052448	-4.503415e-01	-0.2392228693	0.204721868	-0.296042114	0.463807911	-0.010266399	0.0091215260	-0.049379230	-0.009440910	-0.1289344579	-0.272917420
1634885B	0.517093003	-0.2436286211	-0.0053204866	0.156904548	-3.952959e-01	-0.2578389645	-0.258720249	-0.056656111	0.051016334	0.480178028	-0.0631802529	-0.246746048	-0.180916160	-0.0905402079	-0.129419982
201918271A	0.011045066	0.0964362100	-0.1250425875	0.090719402	-4.371949e-01	-0.0028237943	0.121664666	-0.279871732	0.281507909	0.091705523	0.3440404534	0.341133684	-0.168605715	-0.5055361390	-0.275766641
1672586B	0.741644979	-0.2092611492	0.0730692968	0.179622531	1.176336e-01	-0.0118545741	0.151817828	0.086390167	0.429745317	-0.110287182	-0.0786225945	0.055642456	-0.058571845	-0.2435238212	-0.234994918
201925032A	0.454962283	-0.1116551161	0.2941091359	-0.003218368	-2.007181e-01	0.0432289168	-0.460949242	-0.186077923	0.440030813	-0.171984062	-0.3652243614	-0.069633923	0.131468609	-0.1613646150	0.003206463
1671049B	0.257539034	0.0282233041	-0.0656224862	0.273874432	-4.285491e-01	0.3885388374	0.329341710	0.022238882	0.292975128	-0.108692080	-0.0605783388	-0.032102998	0.360026836	-0.1670017987	-0.387464374
1665609B	0.452474803	-0.4227519929	0.0435468294	0.341543913	1.132706e-01	-0.1525760889	-0.424899757	-0.070452720	-0.024231747	0.382951528	-0.0241236519	-0.051269423	0.288727611	0.2047673166	-0.025352269
1721390B	0.461585164	-0.0696945935	0.1703672260	0.292726338	-4.329652e-01	-0.1825168282	-0.305188090	-0.100585148	-0.018920647	0.277299374	-0.1183644459	-0.029401701	-0.154287905	-0.4744006991	0.048491284
1736861B	0.492585331	0.0042352509	-0.0665183142	0.044850267	-4.816509e-01	-0.3466610014	0.183146879	-0.394340038	0.030702632	0.090213224	-0.1869543791	-0.165582135	-0.200030878	-0.2983322144	0.096084945
1722414B	0.590247393	-0.1432181001	-0.4883844554	0.079746835	-2.255344e-01	0.0546547174	0.049394492	0.098751068	0.347472966	-0.365992337	-0.1089598015	0.072353326	-0.110902503	-0.0722668022	-0.175659433

Тепловая карта показывает семантическую «близость» патентов (аннотаций) с помощью интенсивности цвета

Тепловая карта



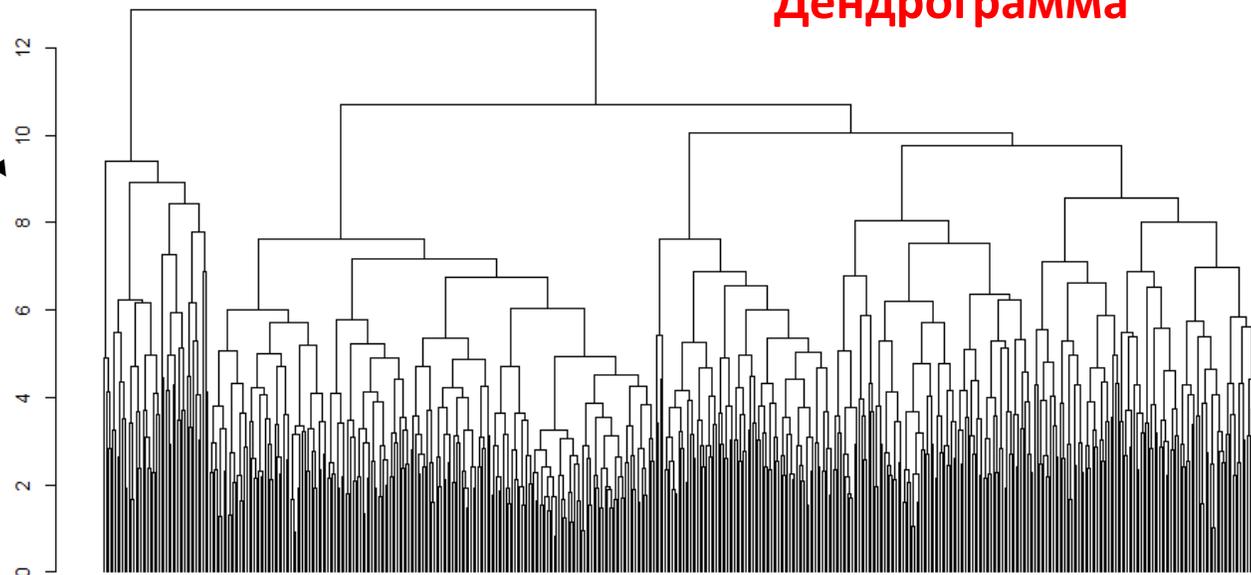
Для примера представлены диаграммы для 500 патентов

Каждый кластер характеризуется набором ключевых слов, которые определяют отдельную технологическую тематику (тему)

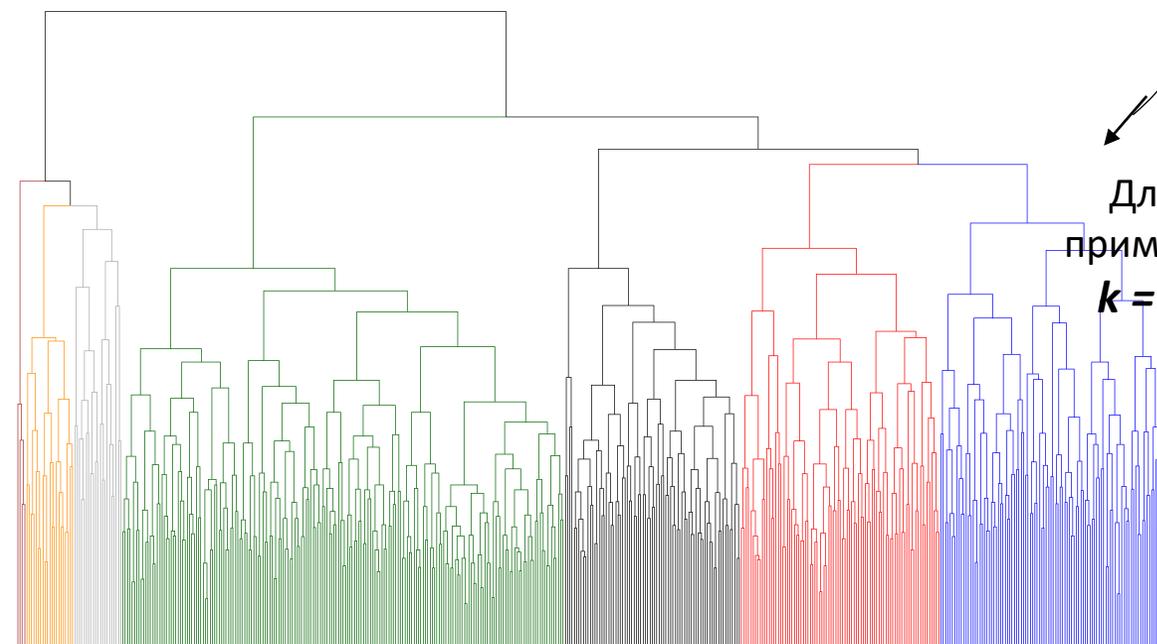
Patent ID

Задается количество групп по предполагаемому количеству кластеров ( $k$ )

Дендрограмма



Для примера  $k = 7$



## Основные результаты работы

- 1) Разработан специализированный алгоритм для сбора патентной документации из информационного источника Lens.org и осуществлена его программная реализация. Этот инструмент позволяет преодолеть ограничения бесплатной версии API информационного ресурса Lens.org.
- 2) Разработано типовое представление информации об единичном патентном материале, которое направлено на создание единообразной структуры данных, содержащей ключевую информацию о патенте. Разработанная структура JSON-файла обеспечивает легкость в обработке, удобство интеграции с другими системами и возможность эффективного анализа патентной информации.
- 3) Был построен патентный ландшафт на основе платформы Elasticsearch. Этот анализ выявил ключевые технологические тренды в таких областях, как «квантовая программная инженерия», «квантовое интеллектуальное управление» и «интеллектуальное когнитивное управление».
- 4) Был создан корпус текстов аннотаций патентов и его словарь, на основе которых обучены векторные нейросетевые модели Word2Vec и Doc2Vec. На основе векторов семантического пространства рассчитана близость документов корпуса (каждого с каждым). С помощью агломеративной кластеризации (дендрограмма) показана принципиальная возможность определения групп схожих по смыслу документов.

# Научная деятельность

**Публикации:** PIN – 21 (из них: ПИНЦ – 18; Web of Science – 5; Scopus – 1; Входит в перечень ВАК – 7).

**Участие в конференциях, проектах:**

Конференции за 5 лет – 2 (соавтор, слушатель)

**Интеллектуальная собственность – 1:**

Свидетельство о государственной регистрации программы для ЭВМ (№ 2024690979, 18.12.2024) **«Аналитический фреймворк обработки и представления научно-технической информации»** (выдано Федеральной службой по интеллектуальной собственности, патентам и товарным знакам (Роспатент). Правообладатель: Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский ядерный университет «МИФИ».

<p>РОССИЙСКАЯ ФЕДЕРАЦИЯ</p>  <p>ФЕДЕРАЛЬНАЯ СЛУЖБА ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ</p> <p><b>RU2024690979</b></p> <p><b>ГОСУДАРСТВЕННАЯ РЕГИСТРАЦИЯ ПРОГРАММЫ ДЛЯ ЭВМ</b></p>	
<p>Номер регистрации (свидетельства): 2024690979</p> <p>Дата регистрации: 18.12.2024</p> <p>Номер и дата поступления заявки: 2024689516 04.12.2024</p> <p>Дата публикации и номер бюллетеня: 18.12.2024 Бюл. № 12</p> <p>Контактные реквизиты: 115409, Москва, Каширское ш., 31, НИЯУ МИФИ, ОУИС, info@mephi.ru</p>	<p>Автор(ы): Артамонов Алексей Анатольевич (RU), Тукумбетова Руфина Рашитовна (RU), Соколов Иван Дмитриевич (RU), Зрелова Дарья Петровна (RU), Коренькова Татьяна Вячеславовна (RU), Хвостова Мария Олеговна (RU), Вуйкович Анастасия Драгомировна (RU), Антонов Евгений Вячеславович (RU), Улизко Михаил Сергеевич (RU), Черкасский Андрей Игоревич (RU)</p> <p>Правообладатель(и): федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский ядерный университет «МИФИ» (RU)</p>
<p>Название программы для ЭВМ: «Аналитический фреймворк обработки и представления научно-технической информации»</p> <p>Реферат: Программа предназначена для обработки и загрузки научно-технической информации в нереляционную базу данных Elasticsearch. Обработка данных позволяет: извлекать из текста химические элементы и соединения; извлекать из файлов (PDF, HTML, XML) таблицы, изображения и данные о статье; извлекать из текста ключевые слова/словосочетания, извлекать физические величины с единицами измерения; получать рубрику первого уровня по рубрикам dimensions (ANZSRC - Australian and New Zealand Standard Research Classification); обрабатывать аффилиации научных организаций (извлекать страну и данные об организации). Программа создана в рамках гранта №70-2023-001309 от 27.12.2023. Тип ЭВМ: IBM PC-совмест. ПК; ОС: Linux.</p> <p>Язык программирования: Python 3.12.7</p> <p>Объем программы для ЭВМ: 900 МБ</p>	

## Планы на 2025-2028 гг.

---

1. **Разработка методологии определения трендов технологического развития.**
2. **Анализ активных патентов.** Проведение анализа **активных патентов** с целью выявления актуальных технологических трендов, исключая заявки на стадии рассмотрения и неактивные патенты.
3. **Семантический анализа.** Проведение **семантического анализа** с акцентом на **научные публикации**, цитируемые в патентах, поскольку они содержат детальное описание технических решений, не раскрытое в патентных формулах.
4. **Анализ полных текстов и цитирований.** Реализация глубокого семантического анализа **полных текстов патентов** и их **цитирований** с использованием **больших языковых моделей (LLM)**, что позволит выявить взаимосвязи между патентами и научными работами.
5. Разработка методологии для создания платформы комплексного анализа научных публикаций и патентов, включая: определение архитектуры и функциональных модулей платформы, интеграцию инструментов для семантического анализа, визуализации данных и прогнозирования трендов, обеспечение совместимости с различными источниками данных (патентные базы, научные репозитории).

---

**СПАСИБО ЗА ВНИМАНИЕ!**

# ПРИЛОЖЕНИЕ

# Обзор информационных ресурсов патентной информации

Информационный ресурс	Описание	Наличие API	Платный / Бесплатный
Российские информационные ресурсы патентной документации			
Федеральный институт промышленной собственности (ФИПС)	Федеральный институт промышленной собственности (ФИПС) предоставляет доступ к патентным базам данных.	Да (ограниченный доступ)	Бесплатный
eLibrary.ru	Российский информационный портал, предоставляющий доступ к научным публикациям, включая патентные базы данных.	Да (ограниченный доступ)	Платный
Patent.ru	Онлайн-платформа, предоставляющая доступ к патентным базам данных и правовым документам.	Нет	Платный
КонсультантПлюс	КонсультантПлюс — сервис для доступа к правовой и патентной информации.	Нет	Платный
Гарант	Гарант — информационная система, предоставляющая доступ к патентным базам данных и законодательству.	Нет	Платный

# Обзор информационных ресурсов патентной информации

Информационный ресурс	Описание	Наличие API	Платный / Бесплатный
Зарубежные информационные ресурсы патентной документации			
Lens.org	Открытая платформа для поиска патентов и научных публикаций.	Да (ограниченный доступ)	Бесплатный
Google Patents	Платформа для поиска патентов по всему миру, включая множество национальных патентных баз данных.	Да (ограниченный доступ)	Бесплатный
Espacenet	Патентный поисковый сервис Европейского патентного ведомства (EPO).	Нет	Бесплатный
USPTO	Официальный сайт Патентного ведомства США для поиска патентов и товарных знаков.	Да (ограниченный доступ)	Бесплатный
PATENTSCOPE	Платформа Всемирной организации интеллектуальной собственности (ВОИС) для поиска международных патентов.	Да (ограниченный доступ)	Бесплатный

# Обзор информационных ресурсов патентной информации

Информационный ресурс	Описание	Наличие API	Платный / Бесплатный
Зарубежные информационные ресурсы патентной документации			
Derwent Innovations Index	Коммерческая база данных для поиска патентов, предоставляющая углубленный анализ.	Да (ограниченный доступ)	Платный
Orbit Intelligence	Платная платформа для профессионалов, предлагающая анализ патентов и других интеллектуальных активов.	Да (ограниченный доступ)	Платный
PatSnap	Инновационная платформа для анализа патентов и управления интеллектуальной собственностью.	Да (ограниченный доступ)	Платный
LexisNexis TotalPatent One	Всеобъемлющая база данных патентов с продвинутыми инструментами поиска и анализа.	Да (ограниченный доступ)	Платный
Questel	Информационно-аналитическая платформа для патентов и интеллектуальной собственности.	Да (ограниченный доступ)	Платный
Derwent Innovations Index	Коммерческая база данных для поиска патентов, предоставляющая углубленный анализ.	Да (ограниченный доступ)	Платный

# Обзор существующих систем патентной классификации

Патентная классификация	Описание	Применение
Международная патентная классификация (IPC)	Стандартная система классификации патентов, используемая во всем мире для упрощения поиска патентов по технике.	Широко используется патентными ведомствами и изобретателями для поиска и классификации патентов.
Кооперативная патентная классификация (CPC)	Объединенная классификационная система, разработанная Европейским патентным ведомством (EPO) и Патентным ведомством США (USPTO), основанная на МПК.	Улучшенная система для более точного поиска и анализа патентов в Европе и США.
Japanese Patent Classification (JPC)	Национальная система классификации, используемая Патентным ведомством Японии (JPO).	Используется для поиска и классификации патентов в Японии.
German Patent Classification (DPC)	Национальная система классификации, используемая Германским патентным и торговым ведомством (DPMA).	Используется для поиска и классификации патентов в Германии.
China Patent Classification (CPC)	Национальная система классификации, используемая Китайским патентным ведомством (CNIPA).	Используется для поиска и классификации патентов в Китае.
United States Patent Classification (USPC)	Система классификации, использовавшаяся до 2015 года в США, заменена на CPC.	Использовалась для классификации патентов в США до 2015 года.

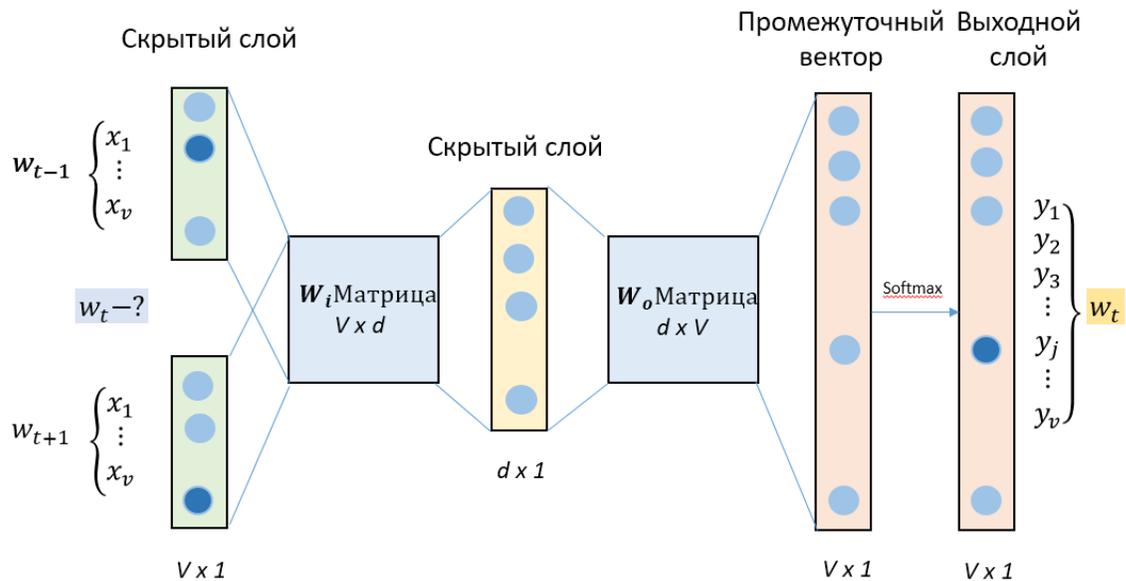


Рис.11. Упрощенная схема нейросетевой модели

### Word2Vec (CBOW)

**Word2Vec** — это алгоритм машинного обучения, разработанный для преобразования слов в векторные представления. Используется для моделирования семантики слов на основе их контекста.

#### Основные архитектуры :

**1.CBOW (Continuous Bag of Words):** Предсказывает текущее слово по контексту (окружающим словам).

**2.Skip-Gram:** Предсказывает контекст по текущему слову.

Таблица 5. Предсказание обученной модели: 5 наиболее близких слов к слову

«robot»

№	Предсказанное слово	Близость	Ранг
1	machine	0.9109197	1
2	equipment	0.9083496	2
3	workstation	0.9005070	3
4	trolley	0.8956949	4
5	AGV	0.8898239	5

Таблица 6. Предсказание обученной модели: 5 наиболее близких слов к слову

«communication»

№	Предсказанное слово	Близость	Ранг
1	wireless	0.9712374	1
2	communications	0.9584386	2
3	repeater	0.9452377	3
4	WiFi	0.9410444	4
5	Bluetooth	0.9323486	5