JOINT INSTITUTE
FOR NUCLEAR RESEARCH

ОБЪЕДИНЕННЫЙ ИНСТИТУТ
ЯДЕРНЫХ ИССЛЕДОВАНИЙ

Intellectual analysis of patent materials on the example of quantum and intelligent robotics

Authors: *Zrelova D.P., Zrelov P.V., Korenkov V.V., Sokolov I.D.*

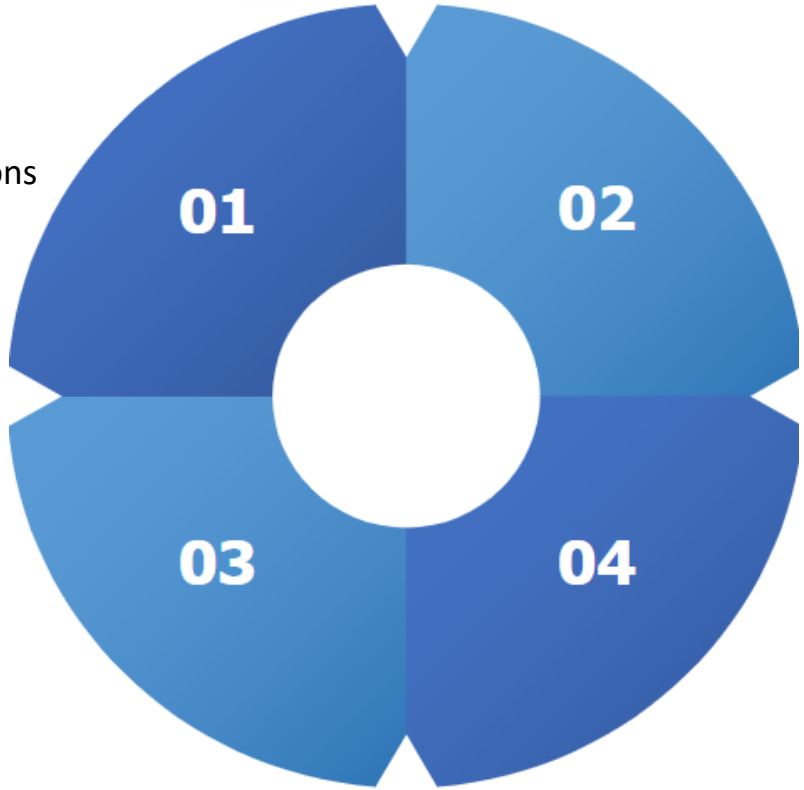
Laboratory of Information Technologies named after M.G. Meshcheryakov
Joint Institute for Nuclear Research (JINR)

Contact: *zrelova@jinr.ru*

Dubna, 2025

Patent analysis: relevance and applications

Validation of the practical relevance of scientific research
Patents help identify research areas with **practical relevance**, as they document technological solutions applied in real-world settings and backed by legal protection.



Competitive landscape assessment
Analyzing competitors' patents reveals their **strategies**, identifies **strengths and weaknesses**, and detects **potential threats**.

Identification of technology trends
Patent data enables tracking of **technology advancements** across industries, identifying **promising areas** and forecasting future developments.

R&D cost optimization
Patent analysis identifies **underdeveloped technology areas**, helping direct funding to **high-potential fields with low competition**.

Step-by-step patent research

Analyze emerging trends in quantum computing and AI-driven robotics (2013–2024).

1. **Patent databases:** WIPO, USPTO, EPO, etc.
2. **Open repositories:** Google Patents, Lens.org

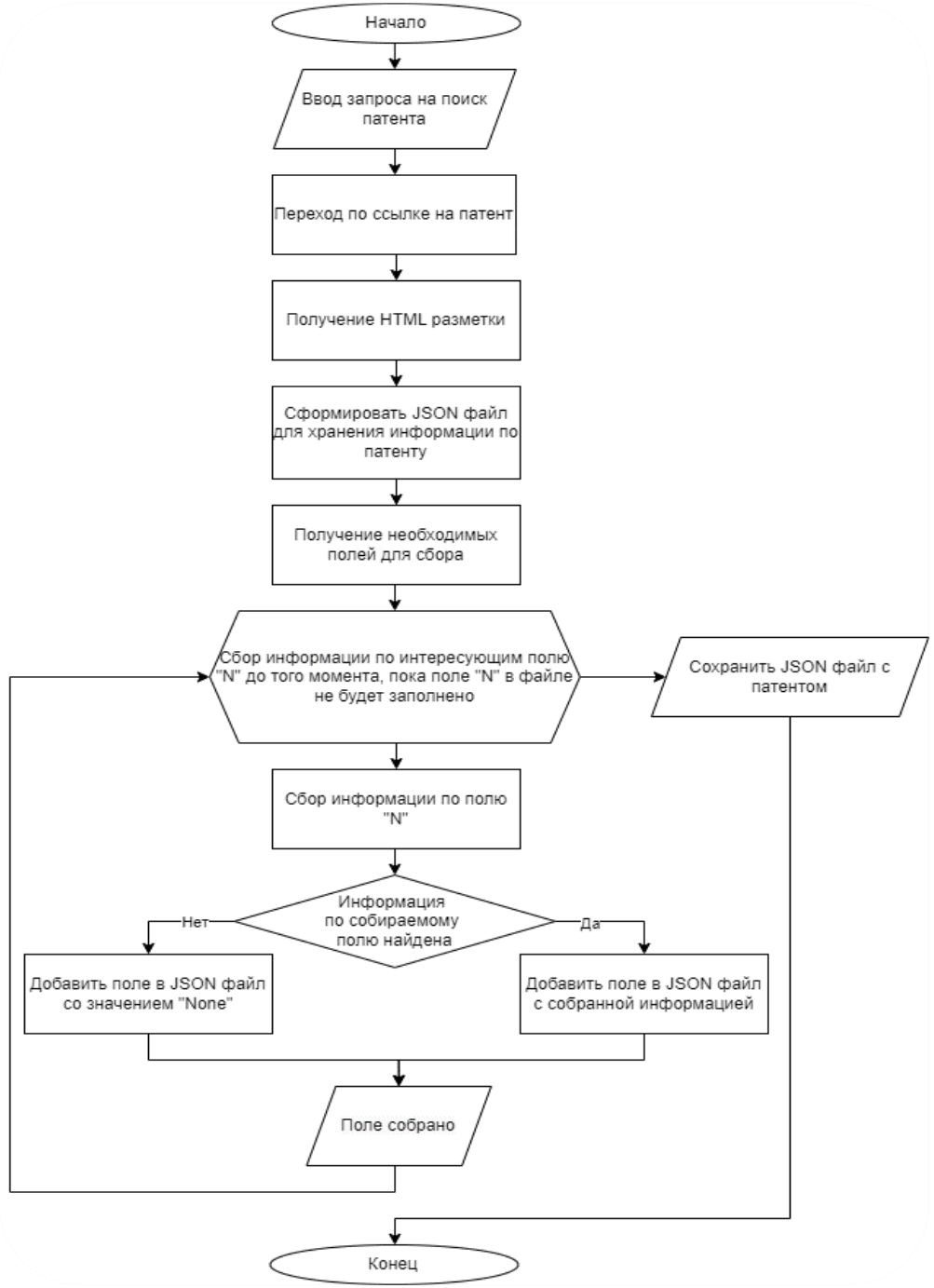
Keyword-based patent search:

1. *"Quantum software engineering"*
2. *"Quantum intelligent control"*
3. *"Intelligent cognitive control"*

1. **Statistical analysis** (number of patents by year, distribution by countries and companies, etc.).
2. **Semantic analysis** (clustering patents by technological fields, identifying key terms and their relationships).

Data visualization: Report preparation with charts, diagrams, maps, etc.

Fig. 1. Flowchart of the single-patent data collection algorithm



Patent data collection algorithm

1

Data source:

Patent database: LENS.ORG

2

Selection criteria:

- Full-text access to WIPO/USPTO/EPO patents
- Integration with scientific publications
- API for automated data collection

! *Limitation:* 1000 documents/query quota

3

Sampling parameters:

Search query:

"robotics AND (intelligent OR quantum)"

→ **48,052** patents (Lens.org, 2013-2024)

4

Automated processing:

- Custom Python script with:
  - ➢ API limit handling
  - ➢ HTML metadata parsing
  - ➢ Data structure normalization

Structured representation of a single patent document

Field	Description
title	Patent title.
url	Patent URL.
abstract	Summary of the patent's content.
patent_id	Country code and patent number.
patent_application	Application details including: family, jurisdictions, status, application/publication/priority/grant dates, applicants, inventors, and assignees.
classifications	*CPC and IPC patent classifications.
other_links	Links to patent documentation and related resources.
cited_by	References to other patents and works citing this patent.
cited_patent	List of patents cited within this patent.
cites_works	References to scholarly publications cited in this patent.

Patent classification systems

IPC (International Patent Classification)
Developed by WIPO (World Intellectual Property Organization). Provides a standardized hierarchical system for categorizing inventions by technological fields.

CPC (Cooperative Patent Classification)
An enhanced version of IPC, jointly maintained by the EPO (European Patent Office) and USPTO (U.S. Patent and Trademark Office). Features more granular subcategories for emerging technologies.

! *In this research, IPC was used as the primary classification system due to its universal adoption and broader applicability.*

Patent classification systems comparison

System	IPC (WIPO)	CPC (EPO/USPTO)
Number of codes	70,000 codes	250,000+ codes
Year introduced	Since 1971	Since 2013
Coverage	Global (all countries)	EU/U.S. focus

Visual analytics dashboard

Analyzed patents: **48,052**

elasticsearch

kibana

Fig. 2. Interactive patent analytics dashboard implemented using Kibana

Research findings

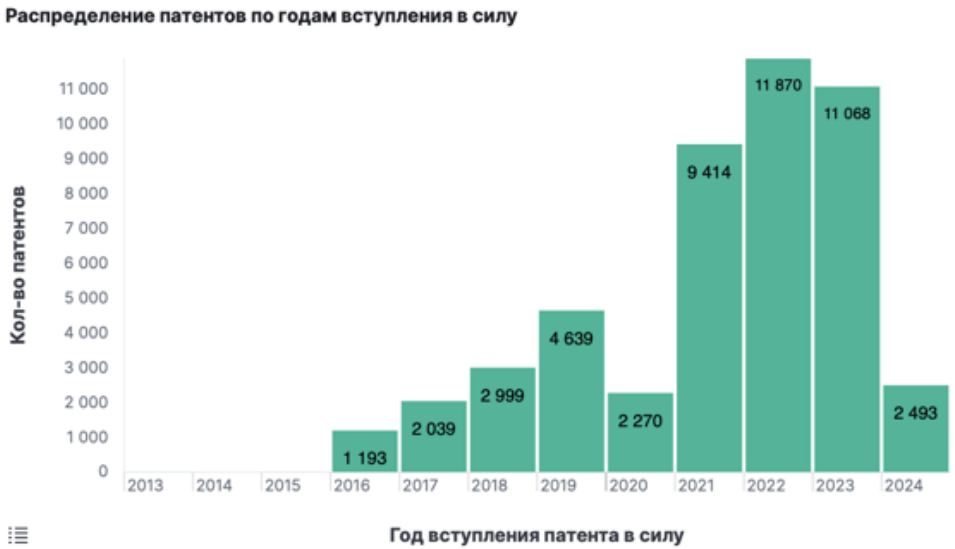


Fig. 3. Patents granted per year

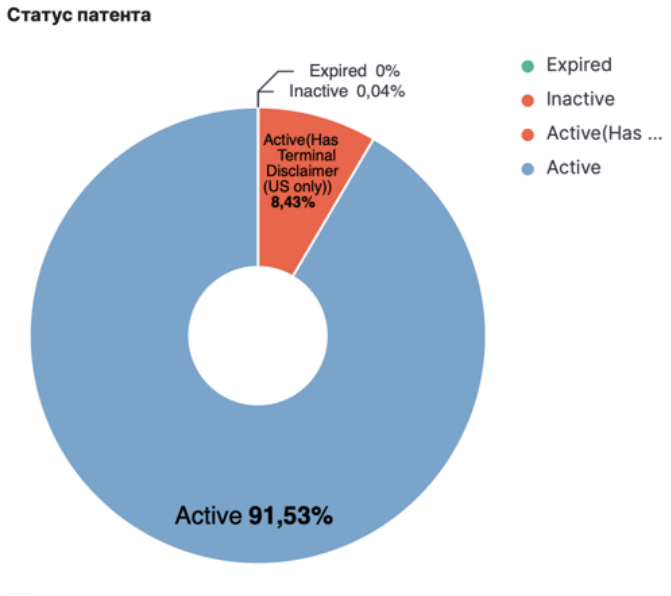


Fig. 4. Patent legal status

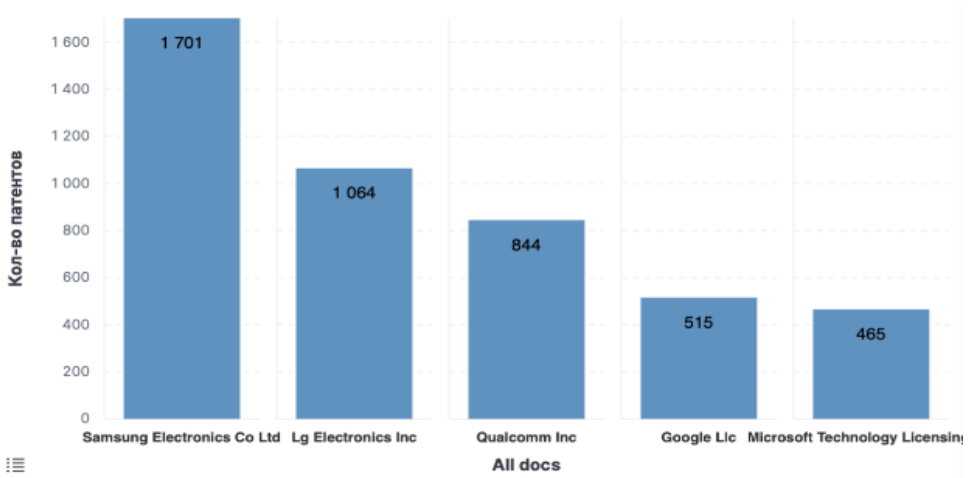


Fig. 5. Top 5 applicant organizations

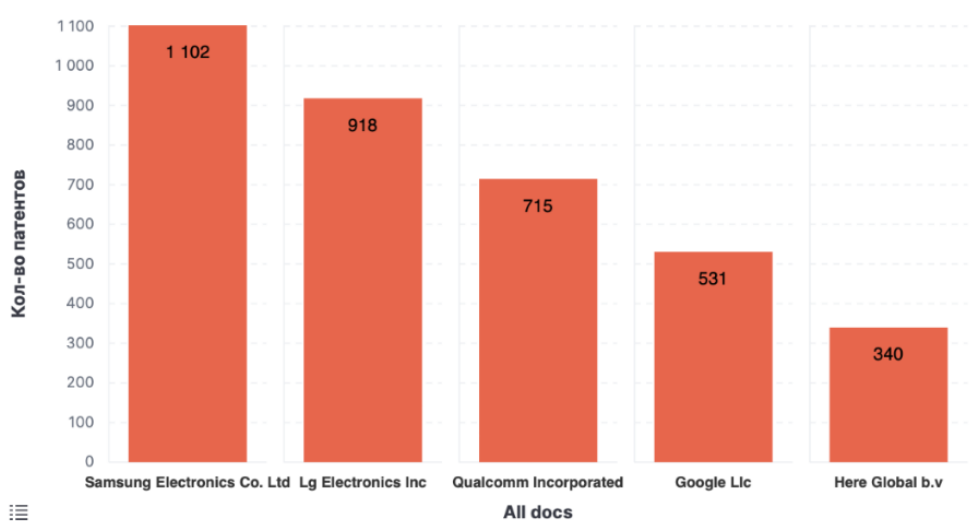


Fig. 6. Top 5 patent owners

Analysis of patent data using the IPC classification system

IPC Classification	Description	Number of patents
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**

The IPC patent analysis revealed that the key domains in intelligent robotics are ***autonomous transportation systems, specialized robotic solutions, and AI-based pattern recognition technologies.***

Patent abstracts analysis

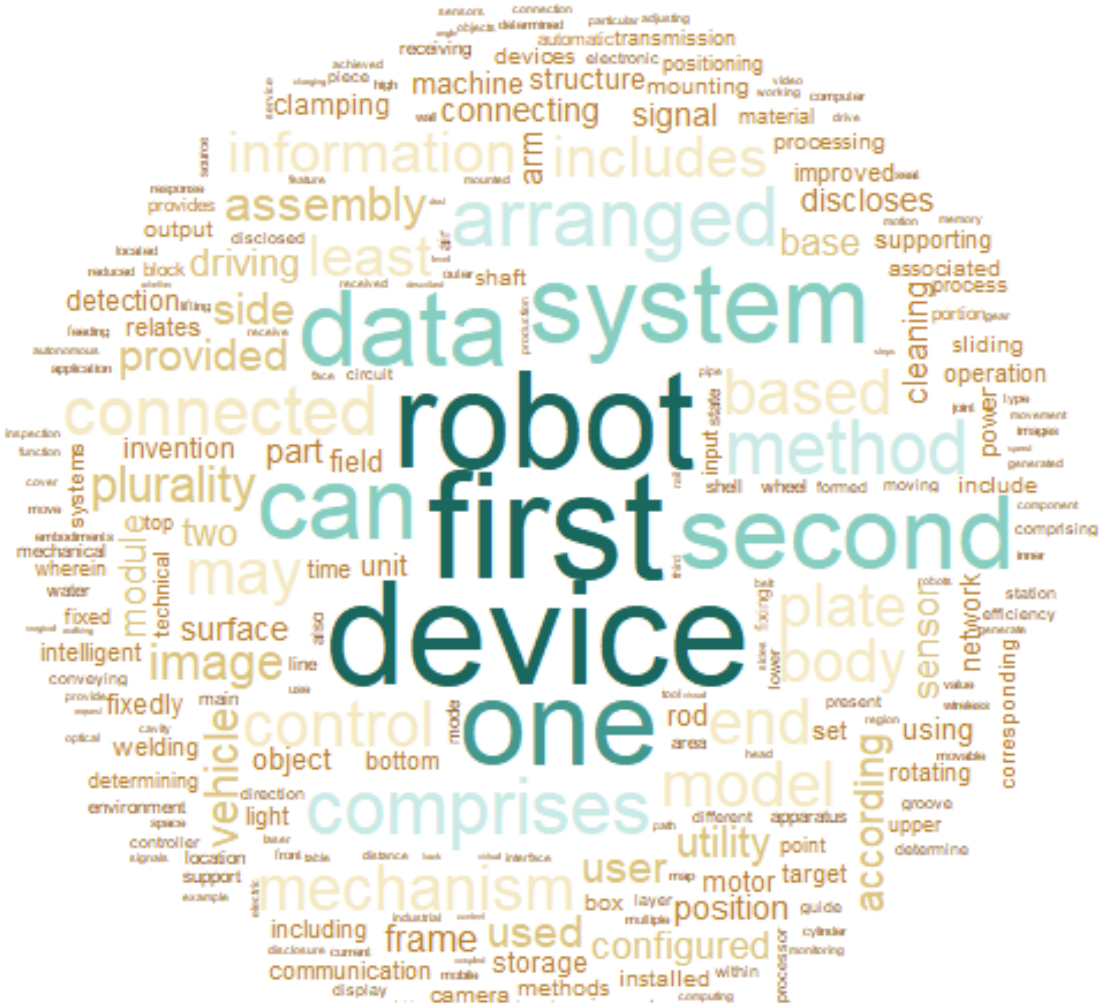


Fig. 7. Word cloud generated from patent abstracts

Out of 48,052 collected patents, **2,576 (5.4%) lacked abstracts**. There were **45,474 patents with abstracts**, of which **45,222 were in English** and only **252 in other languages (0.6%)**. The abstracts were used to **create a document corpus** and establish the **corpus vocabulary**.

Figure 7 shows the corresponding **word cloud** generated from this corpus.

Programming language:

Development environment:

Statistical analysis

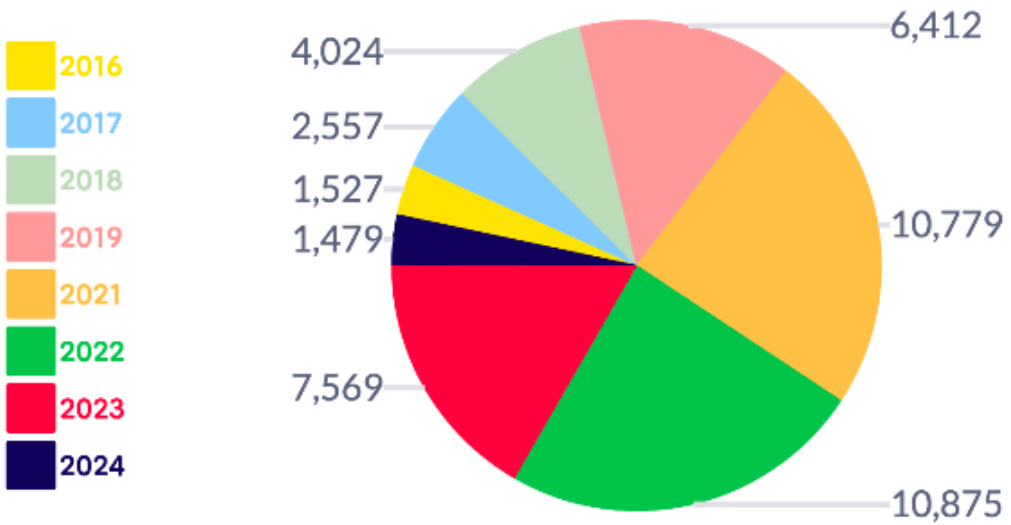


Fig. 8. Distribution of patents by year

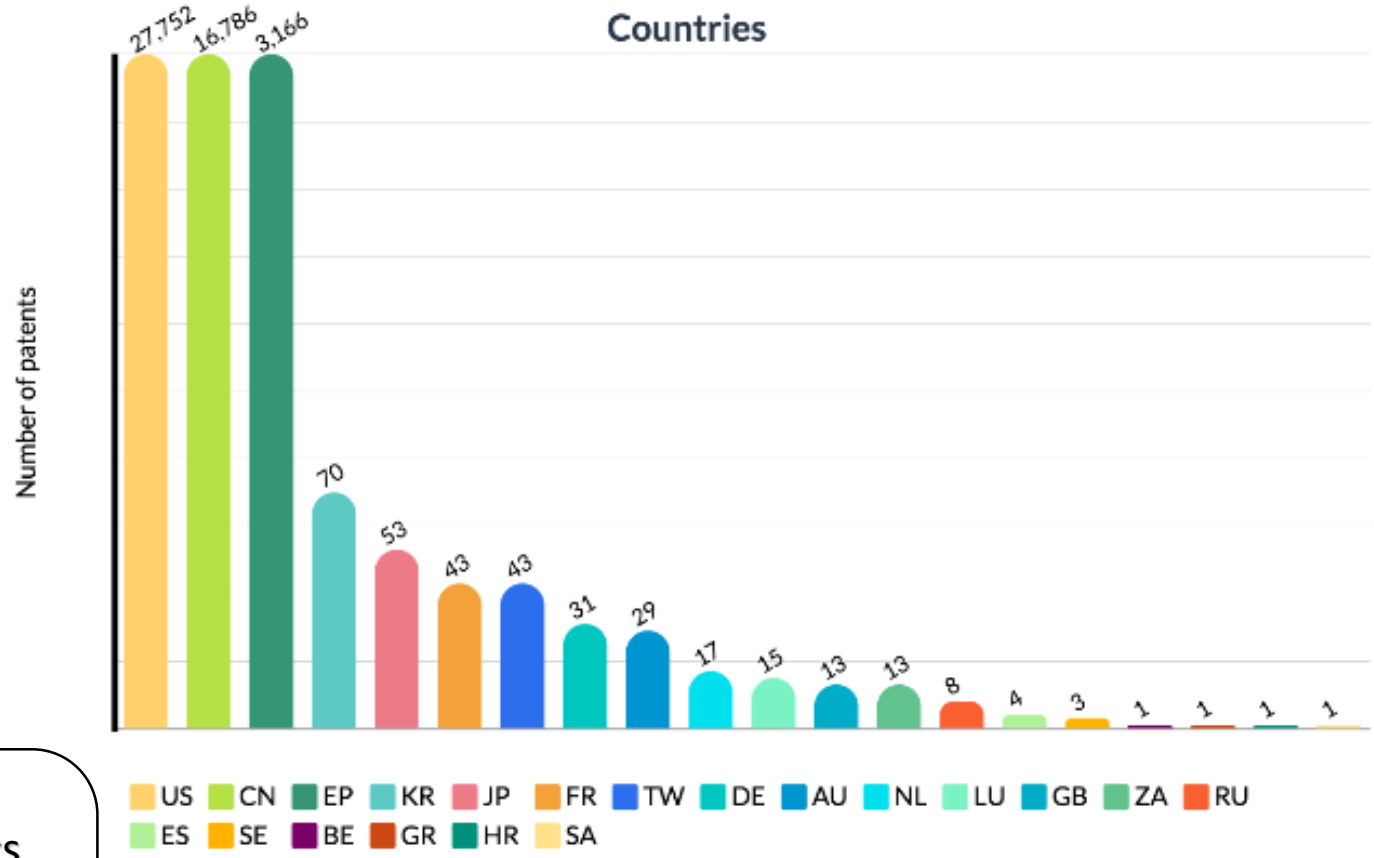
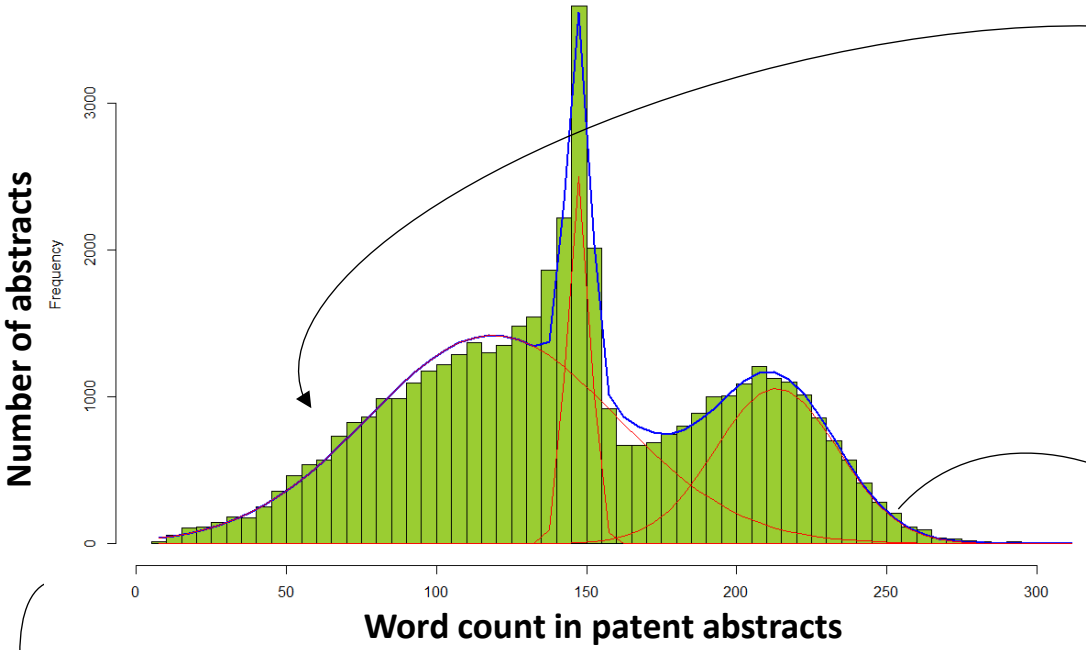


Fig. 9. Geographical distribution of patents

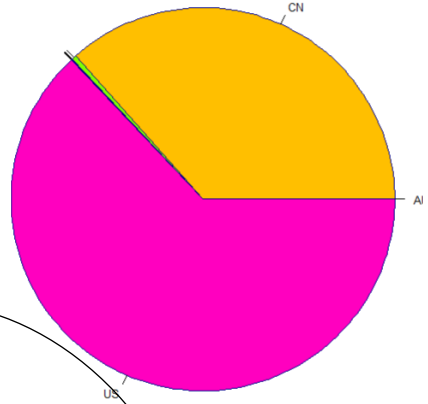
Leaders in intelligent robotics patents:

United States – 27,752 patents
China – 16,786 patents

Statistical analysis

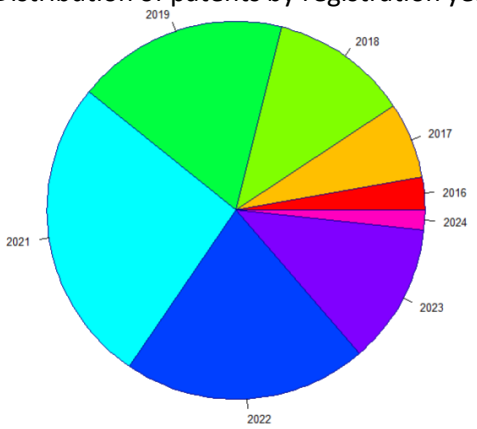


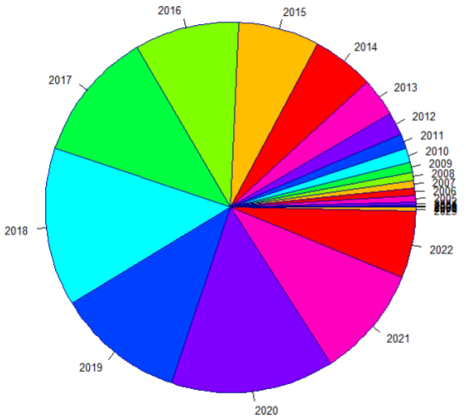
Patents with short abstracts **(≤130 words)**



Distribution of abstracts by country

Distribution of patents by registration year





Patent distribution by priority date

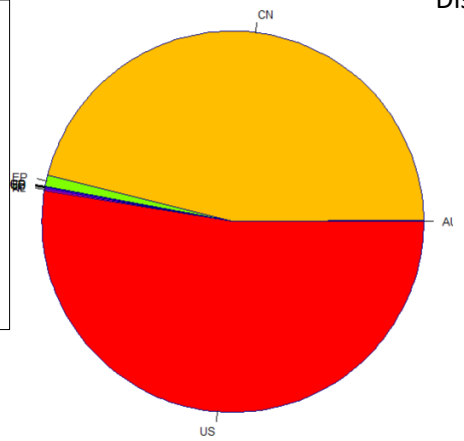
Patent abstract corpus statistics:

Total patents in corpus: 45,222

Abstract word count:

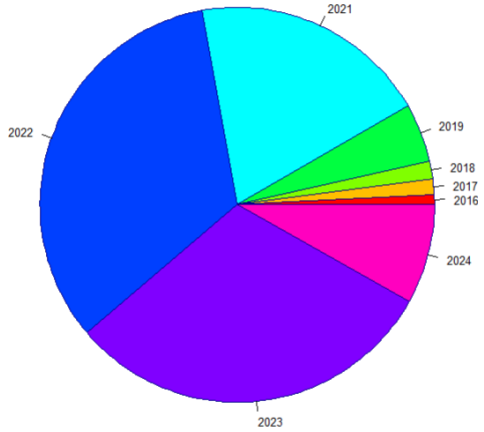
- **Minimum:** 8 words
- **Maximum:** 422 words
- **Mean:** 143 words

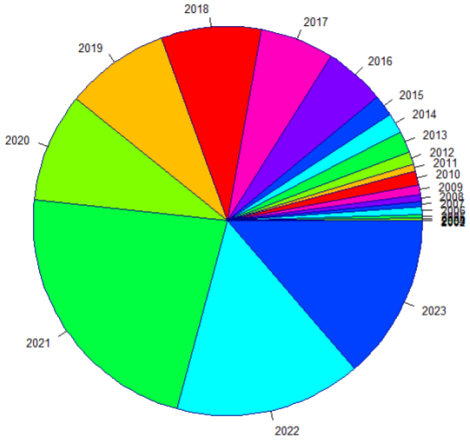
Patents with long abstracts **(≥ 210 words)**



Distribution of abstracts by country

Distribution of patents by registration year





Patent distribution by priority date

Note:

Percentile distribution of abstract lengths:

25th percentile: ≤108 words

75th percentile: ≤189 words"

→2021–2022 saw a significant rise in annotation volume (words), tied to earliest priority (2021) and granted dates (2022), with Chinese patents contributing disproportionately.

Abstracts

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 t
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	104348381	The invention discloses a full automatic underground minir craper and the built in automatic pilot system set in the main body The built in automatic pilot system includes Sensor module is set on each part of the main body to sense the ex
10	I634885B	A method for controlling a walking assistant apparatus incl rmation associated with gait of the user detecting a torque applied to a torque sensor estimating a speed of the user based on the information calculating a compliant motion spe
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	I672586B	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	I671049B	An automatic device for brewing beverages at lea
15	I665609B	A resident activity recognition method is provide
16	I721390B	A monitoring system a base station and a contro
17	I736861B	A monitoring system and a control method there

Text preprocessing

removing special characters/ punctuation/ numbers, trimming spaces, lowercasing, stopword removal, tokenization, lemmatization, etc.

Vocabulary

Word2Vec →

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15
	All	All	All	All	All	All	All	All	All	All	All	All	All	All	All
Inter	0.8467912078	-0.816917539	-0.72379023	-0.25842783	0.087512195	-1.1665312052	-2.254925013	0.779150724	0.62747121	0.29496697	1.44522905	-0.907707512	-1.25861299	0.3945505619	-9.119184e-01
wounded	-1.0311551094	-0.228087634	-0.55628121	0.31987944	0			-0.635232270	-1.71995795	-0.60550171	-1.16129661	1.415262103	-2.27859855	0.7135583758	-8.340638e-02
blackboard	-0.1097669750	-1.367889762	-0.90172982	0.61814988	0			-1.156199336	1.61438584	-1.21474671	-0.65962642	-1.543899417	-0.93764180	-1.0874022245	-6.831360e-01
digestion	0.1478852928	-0.712823987	-0.94042146	-0.64542764	0			-0.030062983	0.11346254	-1.96688700	0.36359704	-0.335189670	-2.10549736	-1.5351772308	-6.469876e-01
assemblie	-1.0125856400	-1.537259340	0.35351068	0.96359092	-1.059786201	1.2245568037	0.588716567	1.536285281	-0.12341926	1.62443423	-0.95465887	-0.567693472	-0.12422322	-1.1459354162	-1.109656e-01

Embeddings

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
I634885B	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
I672586B	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
I671049B	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
I665609B	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
I721390B	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
I736861B	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
I722414B	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

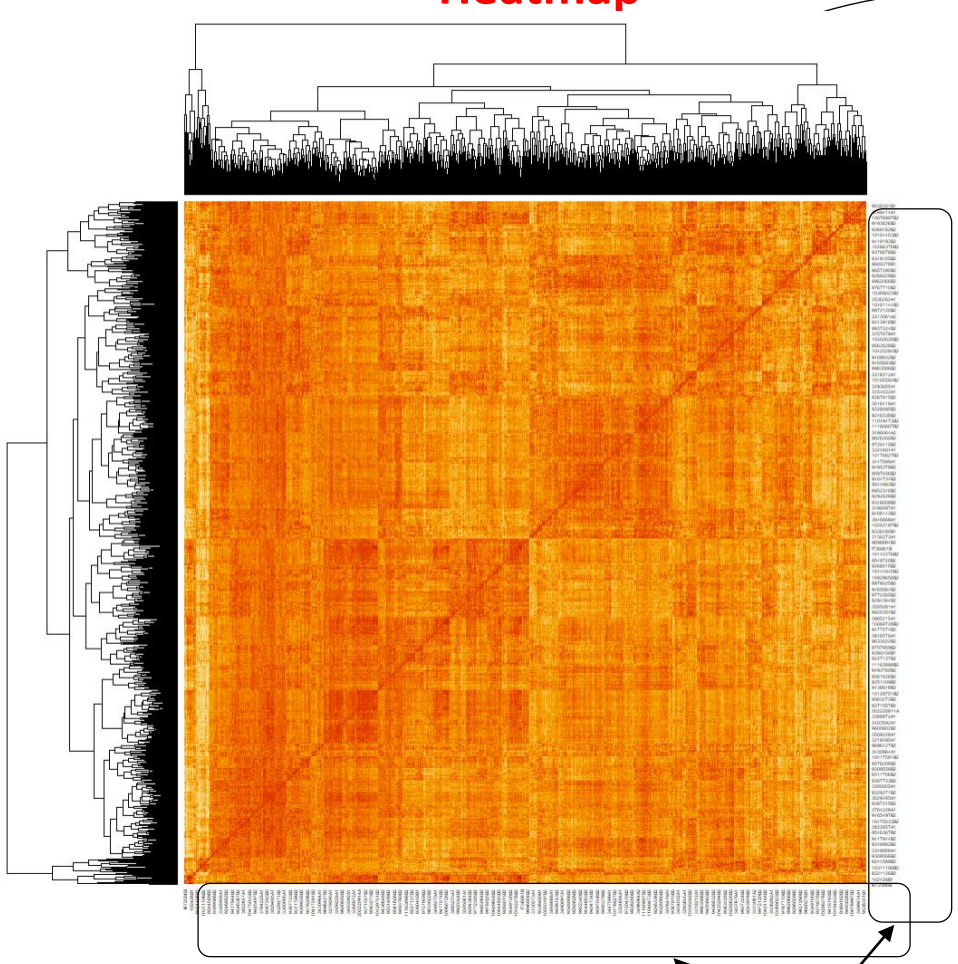
Showing 1 to 18 of 45,222 entries, 15 total columns

-4.684337e-01
-1.001097e-01
-2.008249e+00
4.779609e-01
-7.121503e-01
-1.844755e-01
8.320908e-01
5.945995e-01
2.316403e-01
-6.131788e-01
-9.721525e-01
7.468807e-02
-2.497472e-01

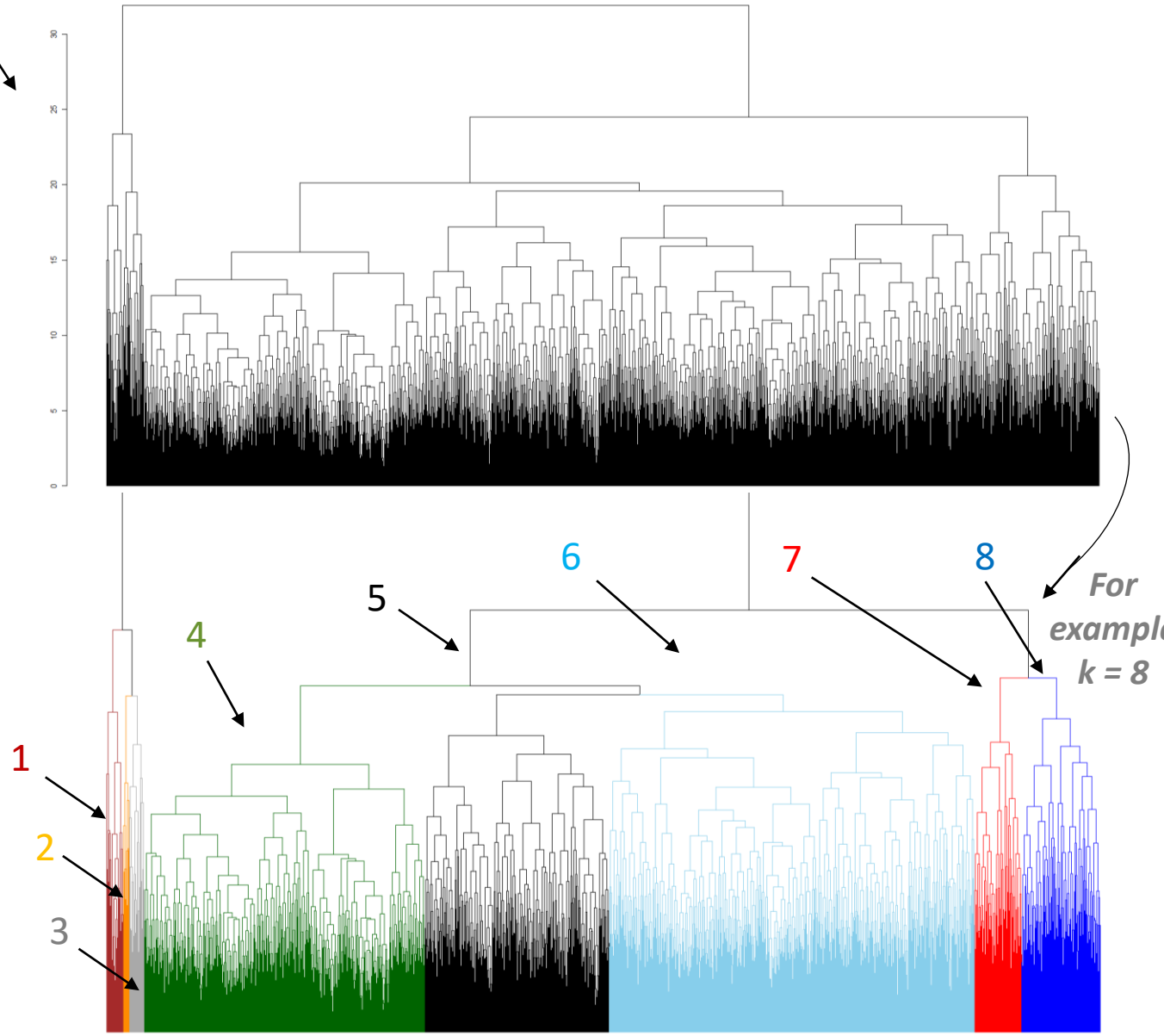
The heatmap visually represents semantic proximity between patents (annotations) through color intensity gradients.

Heatmap

Visualization demonstrates patterns across **2,000** patent documents



Dendrogram of agglomerative clustering

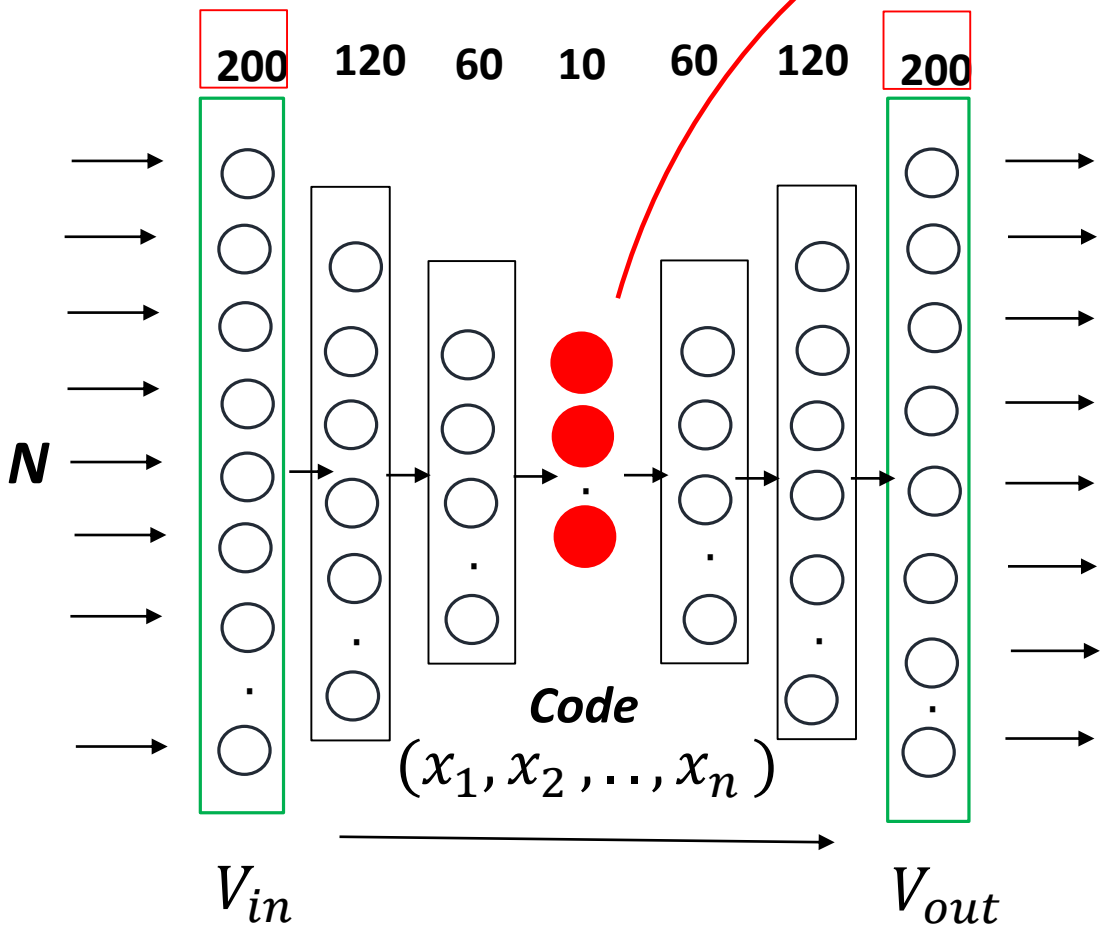


Each cluster is characterized by a distinct set of keywords defining its unique technological domain (topic)

Patent ID

The number of groups is determined based on the estimated number of clusters **(k)**

Autoencoder



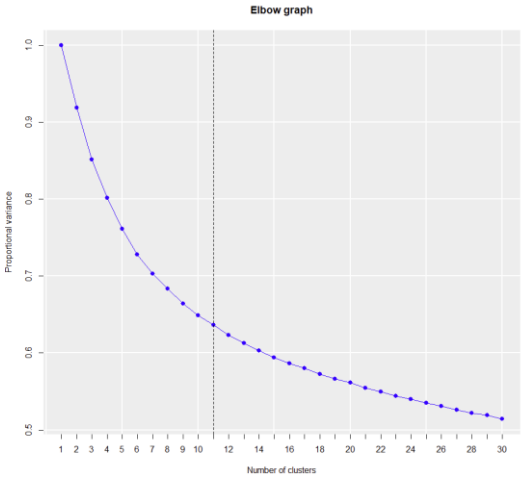
In general case:

1) ***N*** ~ *1000* ***n*** ~ *50 – 100*
2) ***N*** ~ *100* ***n*** ~ *10 – 50*

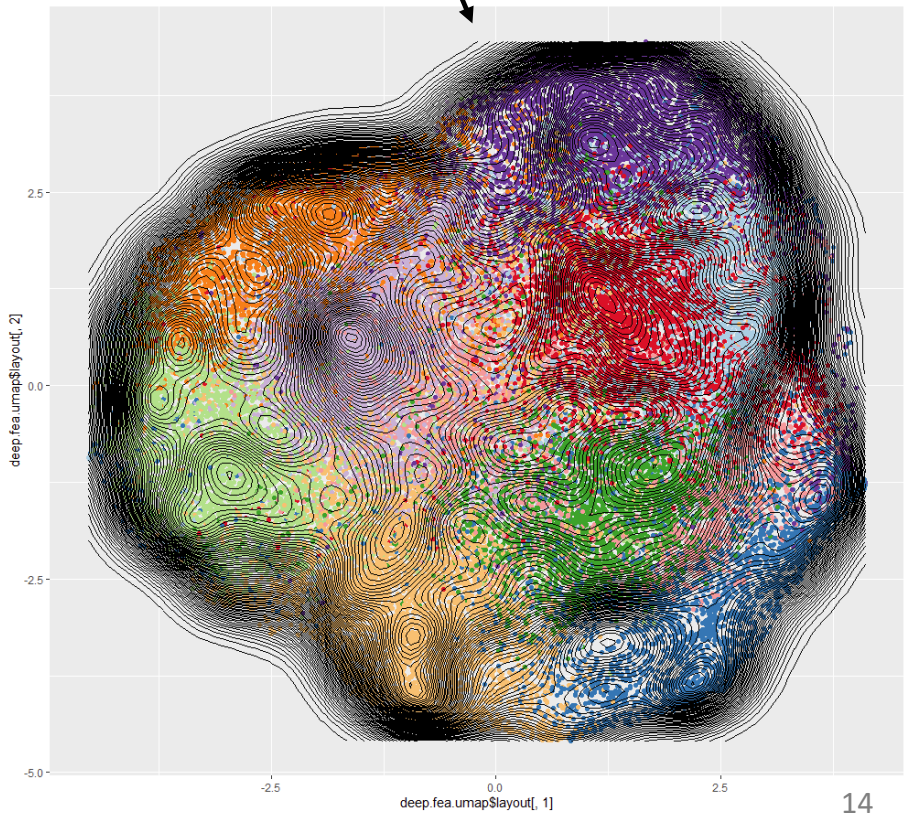
Estimation of optimum number of clusters by elbow method**

Clasterization in 10D-space by k-means*

Visual "control" by projecting in 2D or 3D-space by means of umap***

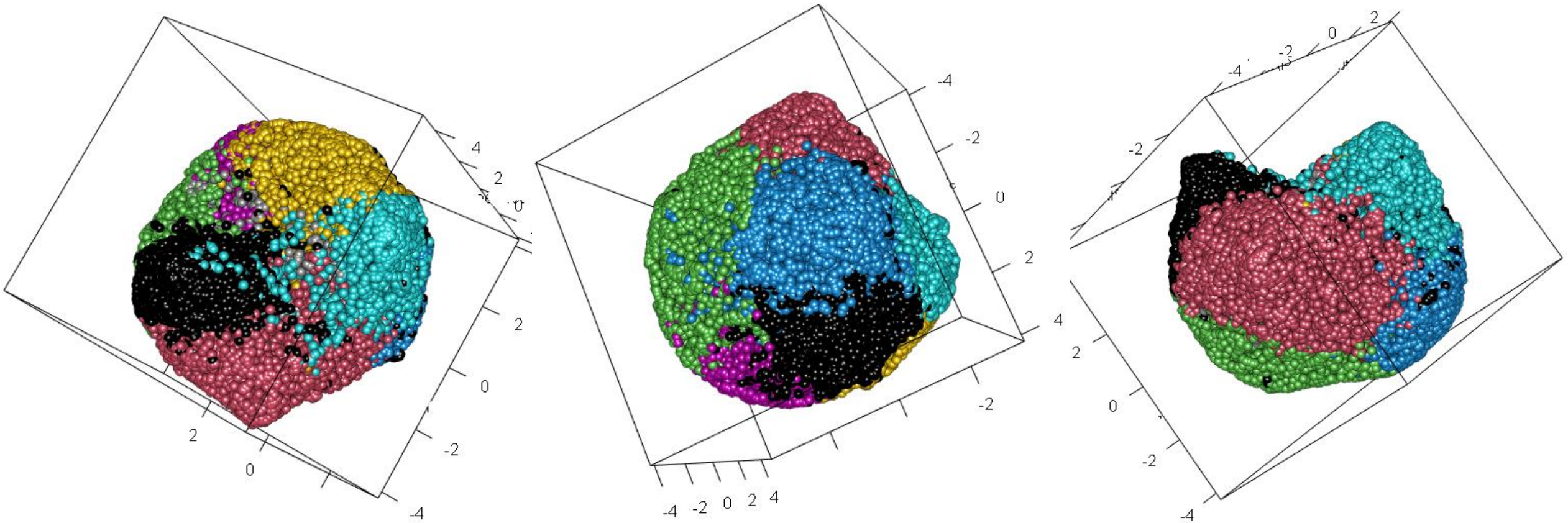


Different embedding dimensions were considered:
N = 15, 50, and 200



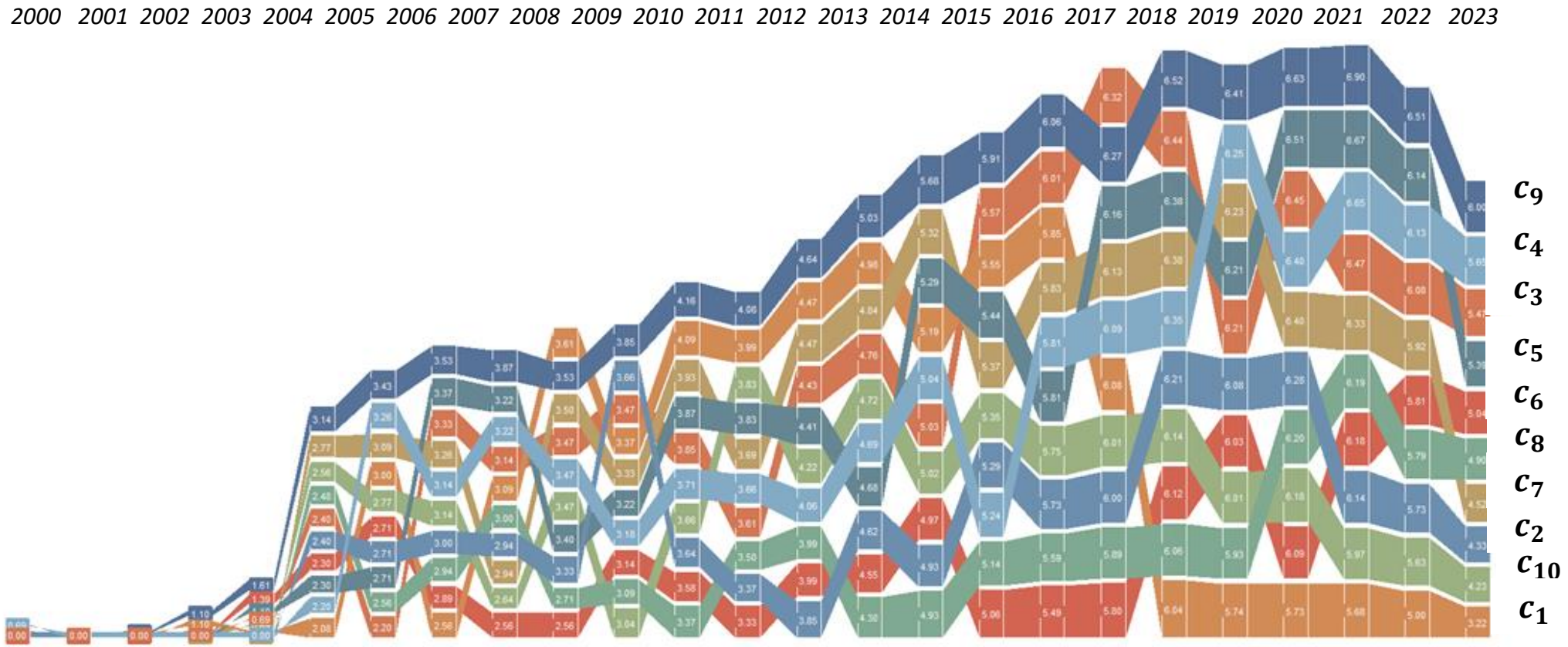
**K-means, DBSCAN, decision trees…*
*** elbow, silhouette…*
****umap, t-SNE, PCA, SVP…*

Example of visual "control" in 3D



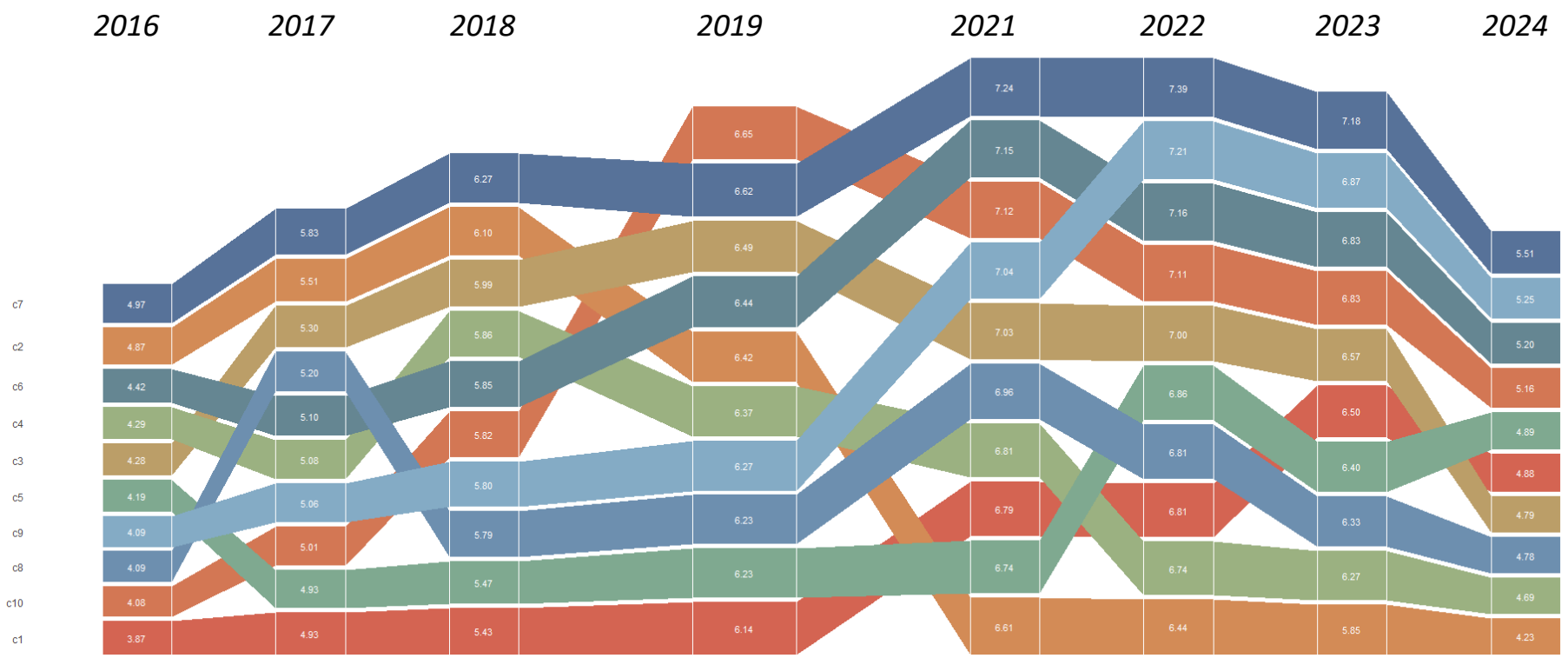
10 “Topics” (keywords)

The distribution of the number of patents for the years 1999-2023, with the earliest priority for 10 clusters (on a logarithmic scale)



clusters	1996	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
c1	0	1	0	0	2	4	10	15	18	13	13	23	36	28	54	95	144	158	242	330	455	417	441	484	333	154
c2	0	2	0	0	3	2	8	20	13	22	37	29	60	54	87	146	179	257	348	438	418	312	309	294	148	25
c3	1	0	1	0	1	0	16	22	26	19	33	28	51	40	87	127	204	214	342	458	589	508	600	560	373	92
c4	0	1	0	0	1	2	13	16	23	14	32	21	39	46	68	112	152	210	313	407	463	408	481	392	279	69
c5	0	0	1	0	0	1	12	13	19	20	15	22	29	33	54	74	139	170	267	363	429	378	495	488	328	134
c6	0	1	0	0	1	3	10	15	29	25	30	25	48	46	82	108	198	231	333	472	592	497	672	787	462	220
c7	0	0	1	2	3	5	23	31	34	48	34	47	64	59	104	153	292	368	430	526	679	610	760	997	671	405
c8	0	0	1	0	1	1	11	15	20	19	28	39	38	29	47	101	139	199	309	403	497	436	536	463	308	76
c9	0	2	0	1	1	1	9	26	23	25	32	24	41	39	58	109	154	189	333	440	574	518	603	770	460	285
c10	0	1	0	0	0	2	11	9	28	23	32	32	47	37	84	117	153	262	407	555	625	496	635	646	435	238

The distribution of the number of patents granted over the years 2016-2024, for 10 different clusters (on a log scale)



clusters	2016	2017	2018	2019	2021	2022	2023	2024
c1	48	139	228	462	893	903	666	131
c2	130	246	445	612	739	624	346	69
c3	72	200	401	659	1126	1101	713	120
c4	73	160	352	582	910	846	530	109
c5	66	139	238	509	849	950	600	133
c6	83	164	348	626	1278	1281	926	181
c7	144	340	528	752	1395	1624	1316	247
c8	60	182	326	510	1056	904	559	119
c9	60	158	331	531	1136	1350	960	191
c10	59	150	338	776	1233	1219	925	175

Conclusions and future plans

- Developed a specialized algorithm for automated patent data collection from **Lens.org**, overcoming free-tier API limitations.
- Created a **standardized JSON structure** for uniform representation of patent metadata and content.
- Using the **ElasticSearch** system, applicants and owners were identified, main technological trends through the IPC were determined, and an express analysis of the dynamics of emergence and development of new technologies, patents, and solutions in the selected fields was conducted.
- Semantic analysis identified **10 clusters** of key terms. These clusters will be further transformed into distinct thematic topics.

Plans:

Integration of the selected methodology into a decision support system for specialists and researchers to identify promising areas for patenting and filing applications for inventions in understudied fields that have the potential to become future technological trends.

Thank you for your attention!