

Исследование возможностей открытых библиографических платформ для решения наукометрических задач

Гуреев В.Н., Мазов Н.А.

Институт нефтегазовой геологии и
геофизики СО РАН,
ГПНТБ СО РАН

ГЗ ГПНТБ СО РАН
(122040600059-7);
ИНГГ СО РАН
(FWZZ-2022-0028)



Сочи-2023



ПОИСКОВЫЕ и наукометрические задачи

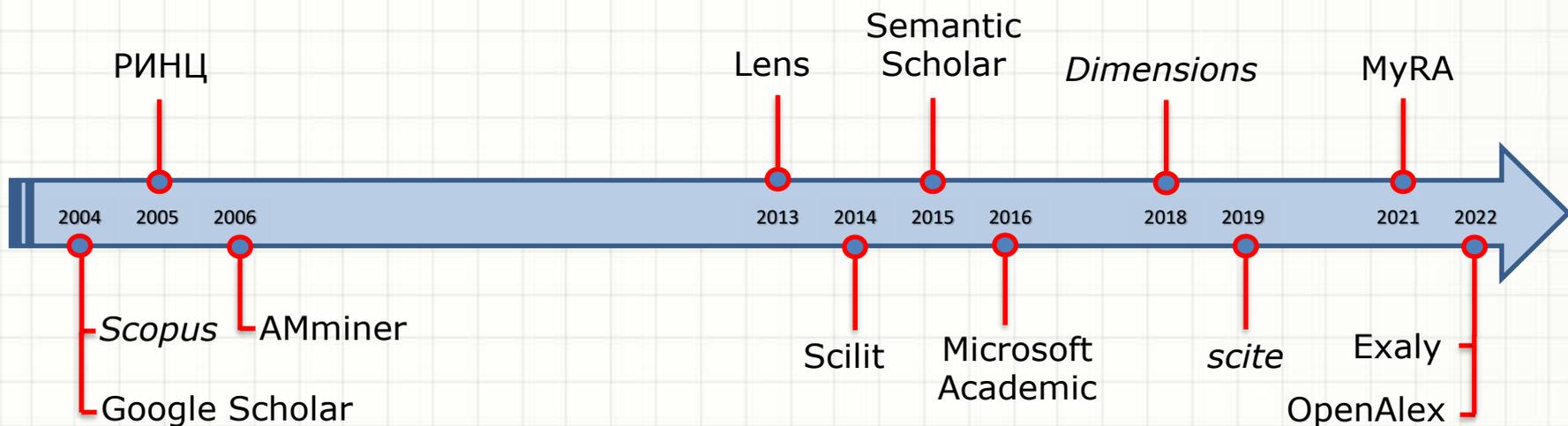
Контролирующие инстанции	Руководство организаций	Научные журналы	Научные сотрудники	Научные библиотеки
<ul style="list-style-type: none"> оценка организаций оценка диссоветов оценка заявок на научные проекты и их результатов 	<ul style="list-style-type: none"> приём на работу карьерное продвижение премирование, ПРНД аттестация сотрудников 	<ul style="list-style-type: none"> поиск рецензентов определение конкурентов верификация ссылок при рецензировании 	<ul style="list-style-type: none"> поиск литературы поиск коллег для совместных работ определение конкурентов поиск журналов для рукописей 	<ul style="list-style-type: none"> информационные справки ведение внутренних библиографических баз данных и репозиториев корректировка сведений о публикациях организации во внешних системах
		<ul style="list-style-type: none"> определение рейтинга 	<ul style="list-style-type: none"> оценка собственной продуктивности 	

С начала 2010-х гг. система управления в российской науке была в значительной мере ориентирована на зарубежные библиографические системы





Движение за открытый доступ к данным в науке затронуло не только научные публикации и журналы, но и библиографические системы





- Появляются полностью открытые системы, составляющие конкуренцию основным коммерческим базам данных
- Коммерческие системы с каждым годом выставляют в общий доступ всё больше метаданных
- Расширяется круг индексируемых документов
- Меняются источники наполнения баз данных
- Расширяется набор решаемых задач

База данных	Тип данных	Число записей	Число цитирований	Число авторов	Активные индексируемые журналы	Основные источники наполнения
Web of Science	журнальные статьи (≈86 млн в Web of Science CC), патенты (109 млн), наборы данных (14,5 млн)	196 млн	1,9 млрд	н/д	34 тыс., из них 21,5 тыс. в WoS CC	издатели, отбираемые экспертами Clarivate
Scopus	журнальные статьи, материалы конференций, монографии	84 млн	1,8 млрд	17,6 млн	28 тыс.	издатели, отбираемые экспертами Elsevier
Google Scholar	научные статьи, монографии, материалы конференций, диссертации	389 млн	н/д	н/д	н/д	издатели репозитории веб-сканирование
РИНЦ	научные статьи, монографии, материалы конференций, диссертации, патенты	41,3 млн	620 млн	1 млн	75,5 тыс., из них 5,6 тыс. в РИНЦ	издатели прочие базы данных
AMiner	журнальные статьи, материалы конференций	316 млн	1,1 млрд	130 млн	н/д	веб-сканирование
Lens	научные публикации (255 млн), патенты (146 млн)	401 млн	1,9 млрд ²	н/д	51 тыс.	Microsoft Academic Crossref PubMed ORCID OpenAlex
Scilit	научные статьи	155 млн	1,2 млрд	16 млн	114 тыс.	издатели Crossref PubMed репозитории ORCID
Semantic Scholar	н/д	211 млн	2,4 млрд	80 млн	н/д	Crossref издатели PubMed репозитории веб-сканирование
Microsoft Academic	все типы публикаций	225 млн	2 млрд	н/д	н/д	веб-сканирование
Dimensions	научные публикации (135 млн), патенты (151 млн), наборы данных (12 млн), гранты (6 млн), книги (1,6 млн)	305 млн	1,7 млрд	313 млн	137 тыс.	Crossref PubMed ORCID
scite	научные публикации	180 млн	1,2 млрд	260 млн	н/д	Crossref PubMed
Web of Science My Research Assistant (MyRA)	преимущественно журнальные публикации	≈15 млн	н/д	–	21,5 тыс.	издатели
OpenAlex	статьи, книги, наборы данных, диссертации	240 млн	н/д	213 млн	118 тыс.	Microsoft Academic Crossref PubMed репозитории ORCID веб-сканирование
Exaly	журнальные публикации	143 млн	н/д	4,9 млн	80 тыс.	н/д

Основные требования к ББД для решения наукометрических задач

- расширенный поиск, создание глобальных и сводных запросов
- наличие авторских профилей
- наличие профилей организаций
- полнота метаданных
- полнота выгрузки
- поиск по пристатейной литературе
- расширенная классификация

База данных	Веб-интерфейс / API	Полнота данных	Широкий набор фильтров	Полнота выгрузки	Глобальный / сводный запрос	Профиль автора	Профиль организации	Поиск по ссылкам	Рубрикатор, число рубрик
Web of Science	+ / +	+	+	+	+	+	+	+	Категории WoS, 21 – верхний уровень; 254 – нижний уровень; 2500 тематик
Scopus	+ / +	+	+	+	+	+	+	+	ASJC, 4 – верхний уровень; 30 – средний уровень; 330 – нижний уровень
Google Scholar	+ / –	–	–	–	–	+	–	–	н/д
РИНЦ	+ / +	+	+	–	–	+	+	+	ГРНТИ, 100 – верхний уровень; >800 – средний уровень; >7000 – нижний уровень
AMiner	+ / +	–	–	–	–	–	+	–	8,7 млн концепций, выделенных на основе машинного обучения
Lens	+ / +	+	+	+	+	–	+	+	см. OpenAlex
Scilit	+ / +	–	+	–	–	–	+	–	н/д
scite	+ / +	–	+	+	–	+	н/д	–	см. OpenAlex
Semantic Scholar	+ / +	–	–	–	–	+	–	–	24
Microsoft Academic	+ / +	неприменимо							см. OpenAlex
Dimensions	+ / +	+	+	+	+	+	+ / –	+	ANZSRC, 236; несколько млн концепций, выделенных на основе машинного обучения
Web of Science My Research Assistant (MyRA)	+ / –	–	+	–	–	–	–	–	Категории WoS, 21 – верхний уровень; 254 – нижний уровень
OpenAlex	– / +	+	+	+	+	–	+	+	19 – верхний уровень; 65 тыс. тематик на основе машинного обучения
Exaly	+ / –	–	+	–	–	–	+	–	8 – верхний уровень, 93 – средний уровень, 272 – нижний уровень; 69 тыс. тематик на основе автоматической обработки естественного языка



с 2018

<https://www.dimensions.ai/>

компания Digital Science
(Великобритания)



с 2013

<https://www.lens.org/>

компания Cambia (Австралия)

Характеристики выгрузки данных

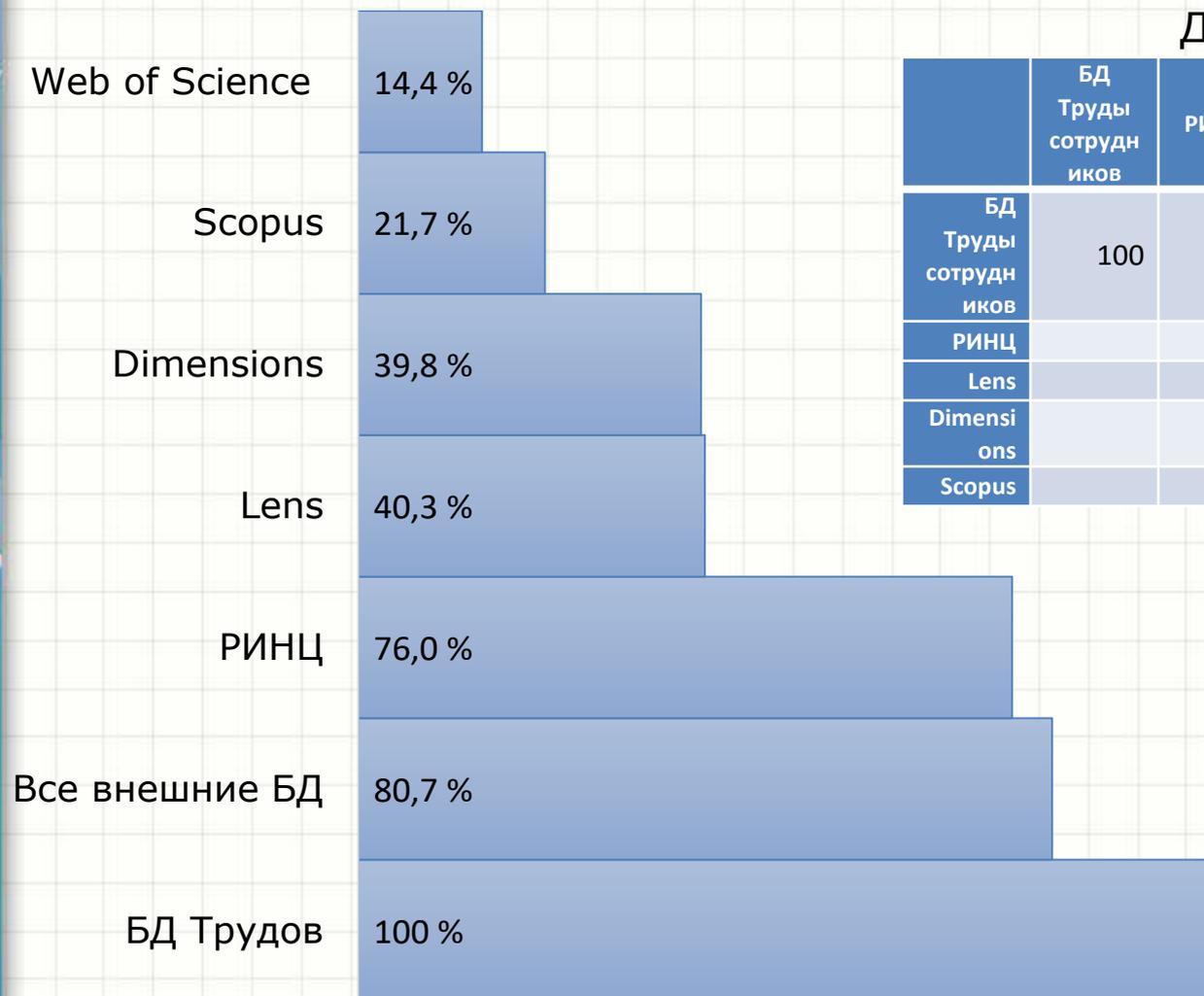
	Выбор форматов выгрузки	Выбор полей для выгрузки	Число полей для выгрузки	Максимальное число выгружаемых записей
Dimensions	+	-	16 полей (бесплатная версия) 53 поля (платная версия)	500 (бесплатная версия) 5000 (платная версия)
Lens	+	+	32 поля	1000
РИНЦ	-	-	99 полей	неограниченно в рамках подписки на Science Index
Web of Science	+	+	4 категории, 29 полей	1000
Scopus	+	+	5 категорий, 30 полей	20000

Использование Dimensions / Lens в научных организациях

В ИНГГ СО РАН в 2023 г. весь технологический комплекс по сбору библиографической информации о публикациях организации для библиометрических расчетов был переориентирован со Scopus и Web of Science на Dimensions и Lens.

- I. Определены профили авторов и организации в обеих системах.
- II. Построены запросы на поиск новых публикаций.
- III. Установлено оповещение о новых публикациях.
- IV. Подготовлена программа адаптации выгружаемых метаданных для институциональной базы данных трудов сотрудников.
- V. По DOI найдены все публикации ИНГГ и репроспективно проставлены ID публикаций из Dimensions и Lens

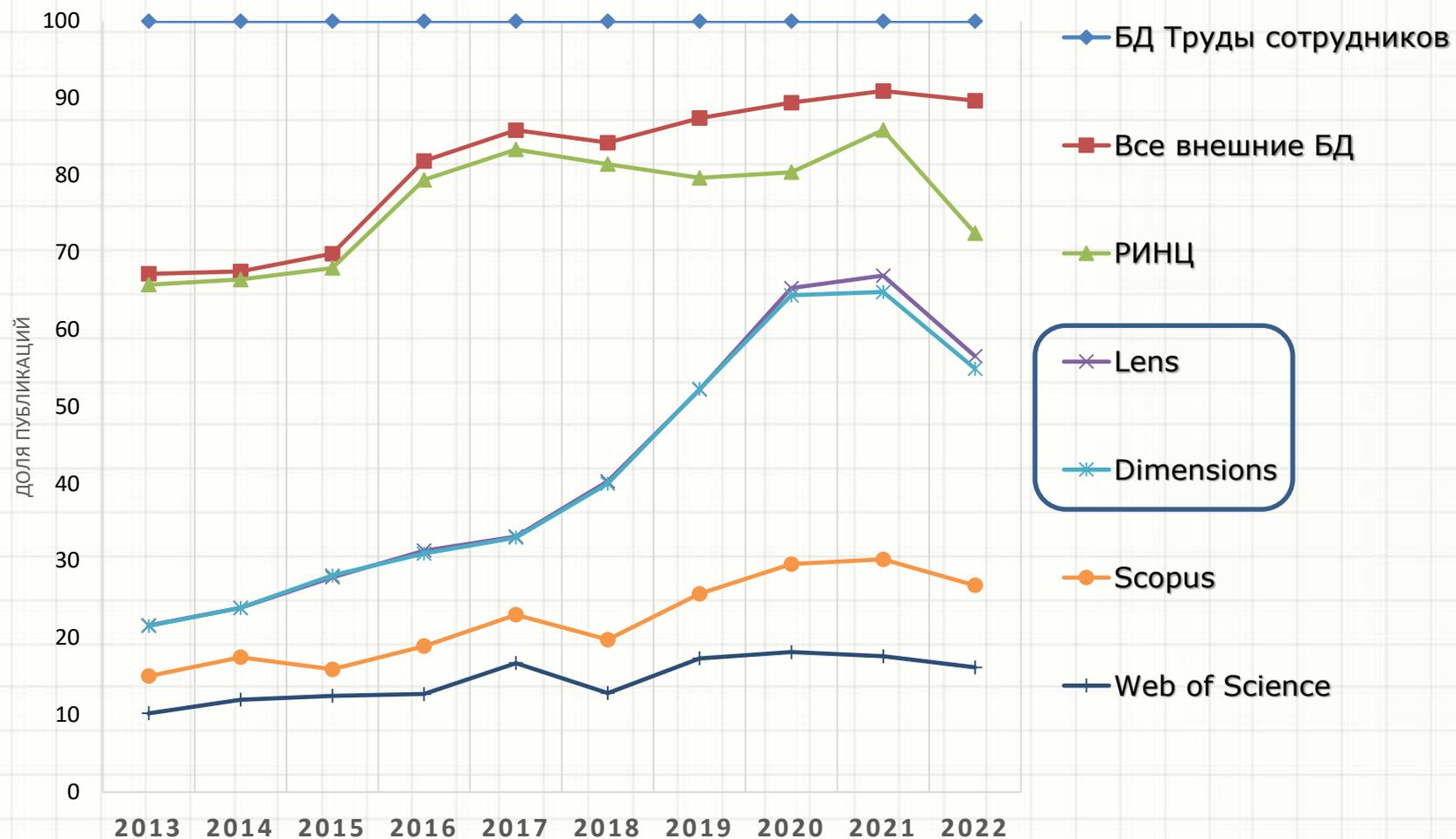
Охват публикаций организации во внешних БД



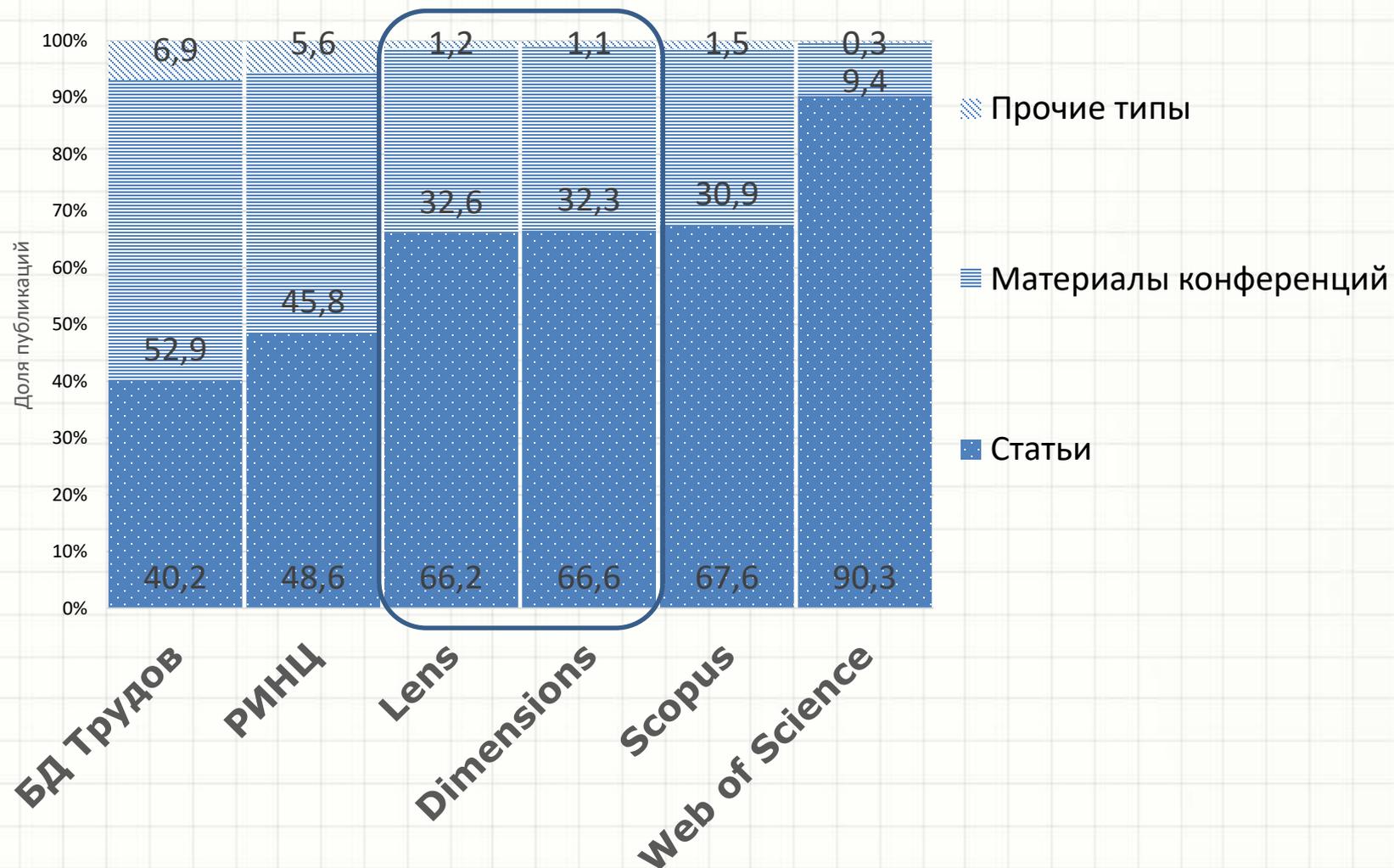
Доля пересечений

	БД Труды сотрудников	РИНЦ	Lens	Dimensions	Scopus	Web of Science
БД Труды сотрудников	100	76,0	40,3	39,8	21,7	14,4
РИНЦ			35,9	35,8	19,0	12,6
Lens				39,7	20,1	13,8
Dimensions					20,1	13,8
Scopus						12,9

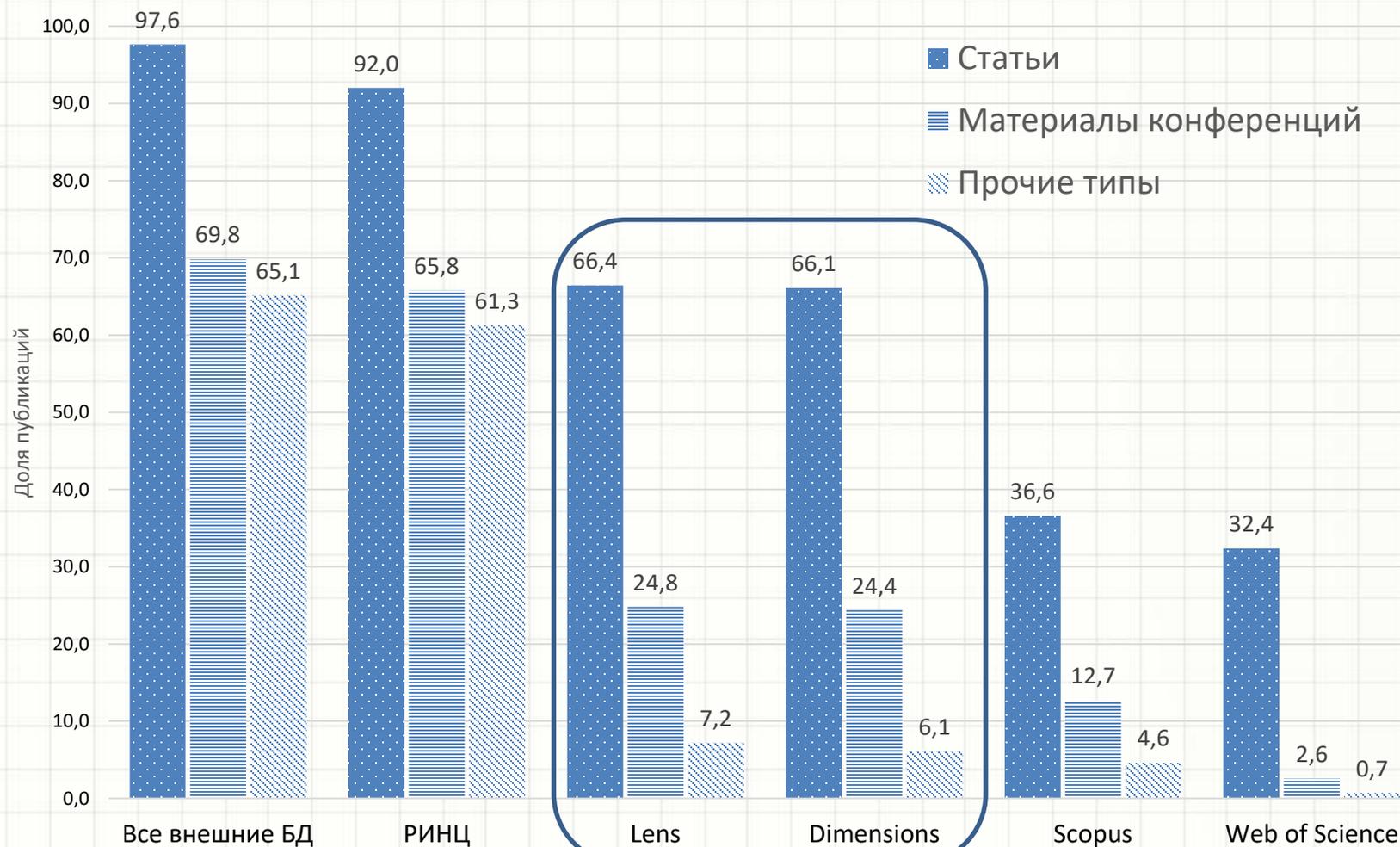
Охват публикаций организации по годам



Распределение по типам документов

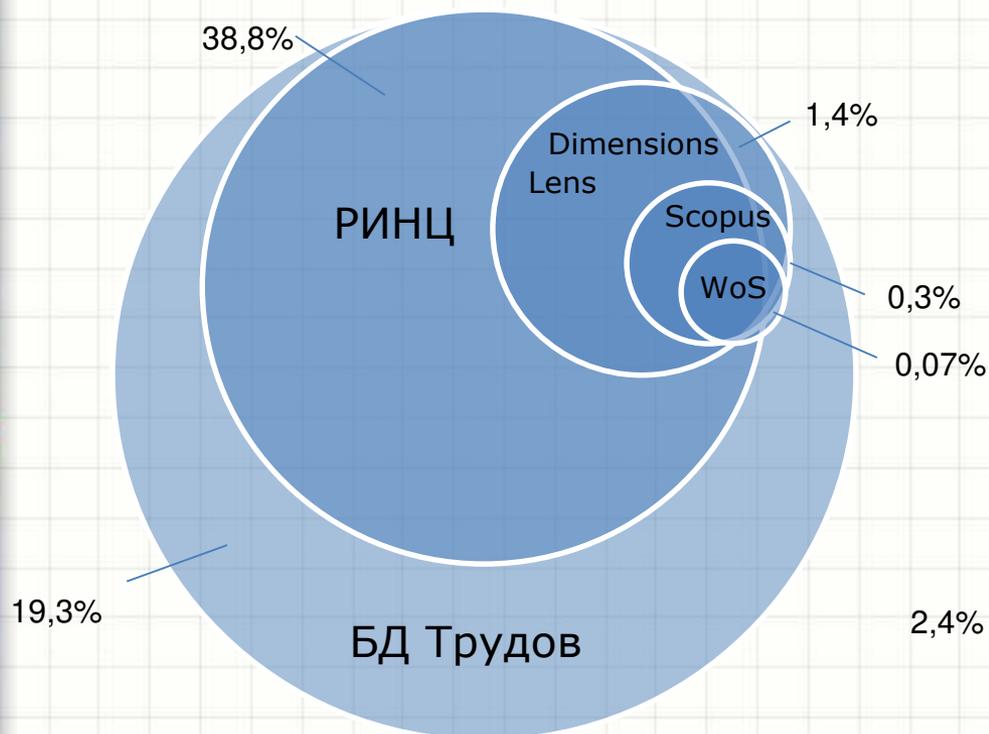


Распределение по типам документов

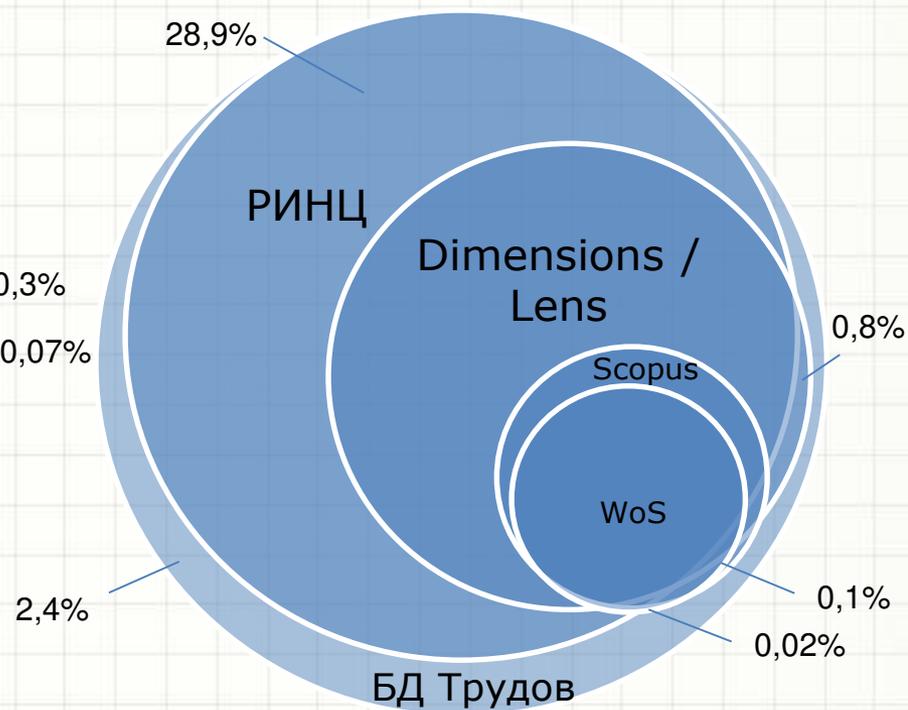


Пересечения и доля уникальной индексации

по всем документам



по статьям



I. Запрос



II. Оповещение

app.dimensions.ai/details/publication/pub/115482671

Dimensions

Publication Chapter

Parallel Non-Conforming Finite Element Technique for Mathematical Simulation of Fluid Flow in Multiscale Porous Media

In Vladimir Jordan, Ilya Tikhonov, and Yury A. Dzyuba

Xenon hydrate formation in water-in-oil emulsion: Investigation with the radiographic method

Authors: Sergey I. Markov, Sergey I. Markov, Anastasiya Yu. Kutshcheva, Natalya B. Iliash, Nikoloz

Abstract: We consider the three-dimensional framework of the Neuber physical process with the model, computational aid



III. Выгрузка

Rank	Publicator	DOI	PMID	PMCID	Title	Abstract	Acknowledgements	Fur
751	pub.11548	10.1007/978-3-031-23744-7_6			Parallel Non-Conforming Finite Element Technique for Mathematical Simulation of Fluid Flow in Multiscale Porous Media			
	Source title	Anthology	MeSH term	Publication Year	Publication Volume	Issue	Page	
	Communications in Computational Chemistry	High-Performance Computing	2023-01-2	2023	2023-01-2	2023	1733	72-82
	Open Access	Publication	Authors	Authors (R)	Corresponding Authors	Author Affiliations	Times Cited	Recent Citations
	Closed	Chapter	Markov, S I	Markov, S I	Markov, S I	Markov, S I	0	0

Сочи-2023

V. Загрузка и ссылка на Dimensions / Lens

Труды сотрудников ИНГГ СО РАН (№ 028708)
 Азаров А.В., Патутин А.В., Сердюков А.С.
 О форме трещин гидроразрыва в окрестности сопряжения скважины с боковым стволом // *Физико-технические проблемы разработки полезных ископаемых*. - 2022. - № 5. - С. 49-62. - DOI: 10.15372/ptpr20220505. - EDN: HKWJWQ.

Труды сотрудников ИНГГ СО РАН (№ 028844)
 Александров А.В., Дородницын Л.В., Колухин Д.Р.
 Исследование моделей эволюции турбулентного поля скоростей на основе анализа автокорреляций // *Препринты ИПМ им. М.В. Келдыша*. - 2022. - № 93. - С. 1-18. - DOI: 10.20948/prepr-2022-93. - EDN: XLWBUU.

Труды сотрудников ИНГГ СО РАН (№ 029448)
 Артамонова В.С., Бортинкова С.В.
 Биогеохимическая характеристика корнеобитаемого слоя травянистых растений на рекультивированных участках техногенных отходов // *Вестник Пермского университета. Серия: Биология*. - 2022. - № 2. - С. 155-163. - DOI: 10.17072/1994-9952-2022-2-155-163. - EDN: BRZZBV.



IV. Обработка

200 *aSaline lakes of Northern Kazakhstan: Geochemical correlations of elements and controls on their accumulation in water and bottom sediments

206 *a107

210 *a2919

215 *a8-18

215 *ahttps://app.dimensions.ai/details/publication/pub.115216471

531 *aapplied Geochemistry

620 37 Earth Sciences

620 3703 Geochemistry

620 3705 Geology

620 3709 Physical Geography and Environmental Geoscience

660 This paper is a review of the saline endorheic lakes of the Northern Kazakhstan, with the provided data on chemical composition of their waters and sediments, diversity/uniformity, mineral precipitation sequences and brine evolution. Study saline lakes are located in the Ishim steppe geographical unit of the southern western Siberia and are ancient closed basins. The lakes are of ClNa and ClMg chemical types (the most saline Zhalauly lake (428g/L) is ClMg type), their pH values vary from 6.8 (Kallbek lake) to 9.1 (Shureksor lake). The common trend for the lakes chemical state is a low calcium concentration with high values of sodium, i.e. calciummagnesiumsodium (with three exceptions). The Na/Cl, Cl/SO4 and Mg/Ca ratios were calculated for the solutions from the lakes in order to reveal their primary trends of geochemical evolution. The geochemical indexes of enrichment and pollution (EF and Igeo) of the bottom sediments show that enrichment is low or absent in comparison with the geochemical background for the lakes of western Siberia. We imply four main stages of lake evolution indicated by precipitation of different secondary minerals: calcite and Mg-carbonates, gypsum, halite, thenardite and bischofite. We suggest the onset of crystallization of each mineral as a tipping point of the water evolution that resulted in the complete removal of a certain chemical element consistent with the solubility constant of corresponding mineral.

700 *aKokopkova N.N. *bSobolev Institute of Geology and Mineralogy, Siberian branch of Russian Academy of Sciences, Prospekt Akademika Koptuyga, 3, Novosibirsk, 630090, Russia

На текущий момент открытые международные ББД позволяют решать большинство наукометрических задач, которые прежде решались с помощью коммерческих систем.

Наборы метаданных перестают быть уникальным продуктом и переходят в открытый доступ.

На первое место выходят аналитические возможности баз данных:

- отслеживание этапов и формирование цикла научного исследования
- построение сетей сотрудничества и выявление влияния исследований
- определение точности и смысла цитирований
- определение веса цитирований
- уточнение тематической классификации

Гуреев В.Н., Мазов Н.А.
Возрастание роли открытых
библиографических данных в
условиях ограничения доступа к
коммерческим информационным
системам // Управление наукой:
теория и практика. – 2023. – Т. 2

