

НЕОПОИСК

# Возможности Неопоиска как дискавери-сервиса: узнать, обнаружить, увидеть

Докладчик:

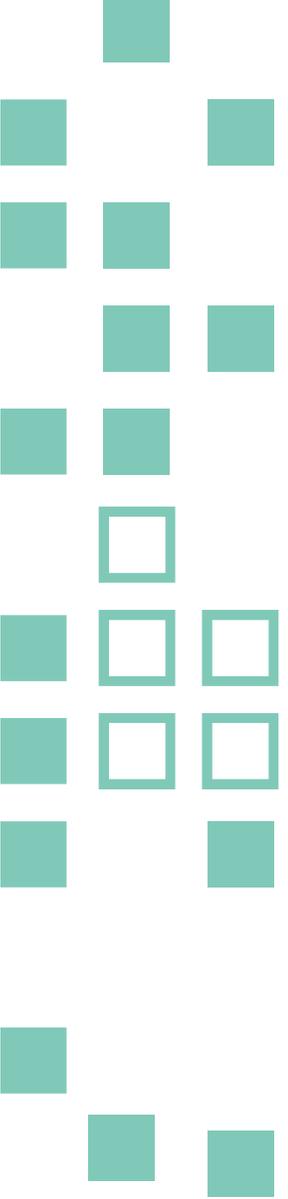
**Альбина Николаевна Нестерова**

Генеральный директор ООО «Неопоиск»

Июнь 2024 г.

## Предпосылки к развитию информационного поиска (IR)

- По экспоненте растет объем информации
- Происходят значительные изменения в предметных областях: IT, генетика, робототехника и пр., появляются новые термины, а классификаторы за ними не успевают
- Большинство дисциплин в университетах - вариативные
- Открытия случаются на стыке дисциплин
- Пользователь становится ленивым
- Можно не заботиться о формулировках - сервисы подбирают синонимы даже к сложным терминам
- Стандартный (каталожный) поиск ищет по точному совпадению в метаданных или ключевых словах - этого недостаточно для решения пользовательских задач



## Современный поиск = Информационный поиск (IR)

**Обработка естественного языка (NLP)** - подраздел компьютерных наук, который занимается искусственным интеллектом (AI); позволяет компьютерам понимать и обрабатывать человеческий язык.

### Фазы обработки естественного языка (NLP):

- I. морфологическая обработка: слова и их формы
- II. синтаксический анализ: связь между словами
- III. семантический анализ: значение слова
- IV. прагматический анализ: условия употребления сл



Важны “ближайшие соседи”, т.н. контекст.

## Семантический поиск

НЕОПОИСК

[О сервисе](#) | [Каталог](#) | [Мой неопоиск](#) | [Смылова Александра](#) ?

**Диапазон дат** v

от  до

1 год | 3 года | 5 лет

**Доступность публикации** >

**Тип публикации** v

- статья
- книга
- глава книги
- сборник материалов конференций
- препринт
- [загрузить ещё...](#)

**Базы данных** >

**Платформы** >

**Квартили SJR** >

**Категории ВАК** >

**Специальность ВАК** >

Применить фильтры (1)

Очистить

Все поля – Аппроксимационные алгоритмы, выпуклое программирование ✎

Публикации  
215
Диссертации  
1000
Патенты  
961

Сохранить запрос
Добавить в подборку
Экспорт
Облако ключевых слов
Аналитика

Найдено: 215    Показывать по: 10 20 50 100

**Курс методов оптимизации**  
Книга | 2011 • Физматлит  
Тестовый доступ | Авторы: Сухарев Алексей Григорьевич, Тимохов Александр Васильевич, Федоров Вячеслав Васильевич  
Книга может служить также введением в **выпуклый** анализ и теорию условий оптимальности в экстремальных задачах. <...>

**Методы оптимизации**  
Книга | 2021 • РИОР  
Полный доступ | Авторы: Аттетков Александр Владимирович, Зарубин Владимир Степанович, Канатников Анатолий Николаевич  
Среди задач нелинейного **программирования** выделяется достаточно широкий класс задач **выпуклого программирования**. <...> Какой вид целевая функция имеет в задаче **выпуклого программирования**? <...> При решении задачи **выпуклого программирования**, в которой **выпуклая** функция минимизируется на **выпуклом** множестве, достаточно найти любую точку ее локального минимума, так как **выпуклая** функция в любой точке локального минимума достигает ...
Ещё

**Теория и моделирование рынка на основе векторной оптимизации**  
Книга | 2009 • УНИВЕРСИТЕТСКАЯ КНИГА  
Вне подписок | Авторы: Машунин Юрий Константинович  
Задача векторной оптимизации Представим векторную задачу оптимизации в соответствии с [29, 30]. рассматривается **выпуклая** векторная задача математического **программирования** с неоднородными критериями.  $\text{opt } F(X) = \{ \max F_1(X) = \{fk(X), k = 1 \dots K \}, (9.1.1) \min F_2(X) = \{fk(X), k = K + 1 \dots n\} \}, (9.1.2) G(X) \leq B, (9.1.3) X \geq 0, (9.1.4) \text{ где } X = \{x_j, j = 1, N\} - \text{вектор вещественных переменных, то есть вектор из } N\text{-мерного евклидова пространства } RN \text{ (обозначение } j = 1, N \text{ эквивалентно } j = 1, \dots, N) \dots$ 
Ещё

**Хачиян Л. Г. Избранные труды**  
Книга | 2009 • МЦНМО  
Вне подписок | Авторы: Л.Г. Хачиян  
Во втором разделе «Сложность **выпуклого программирования**» публикуется докторская диссертация Л. Г. Хачияна [16], в которой анализируется сложность задач полиномиального математического **программирования** в точной и приближенной постановке, показана фундаментальная роль предположения о выпуклости задач, построены полиномиальные **алгоритмы** для линейного и выпуклого квадратичного программирования), а также получены общие результаты о сложности задач вещественного и/или ...
Ещё

## Семантический поиск по принципу стандартного

### Поиск по фразе из нескольких редких слов

### Поиск по короткой и (или) общеупотребимой фразе

Все поля – биохимия полости рта ✎ **Публикации** 981 **Диссертации** 999 **Патенты** 959

Сохранить запрос    Добавить в подборку    Экспорт    Облако ключевых слов    Аналитика

Найдено: 981    Показывать по: 10 20 50 100

**Биохимия полости рта**  
 Книга | 2012 • СпецЛит  
 ☹ **Вне подписок** | Авторы: Данилова Л. А., Чайка Н. А.  
 Биохимия полости рта . <...> Настоящее пособие рассматривает основные вопросы биохимии, касающиеся спецкурса «Биохимия полости рта» по специальности «Стоматология» в соответствии с государственным стандартом обучения студентов. <...>

**Практикум по биологической химии и биохимии полости рта**  
 Книга | 2016 • СибГМУ  
**Тестовый доступ** | Авторы: О. А. Тимин, Т. В. Жаворонок  
 Практикум подготовлен в рамках дисциплины «Биологическая химия – биохимия полости рта» в соответствии с Федеральным государственным образовательным стандартом высшего образования и предназначен для студентов, обучающихся по основной профессиональной образовательной программе высшего образования – программе специалитета по специальности «Стоматология». <...> Практикум нацелен на освоение наиболее важных разделов биологической химии с акцентированием внимания на основах биохимии ...

**Биологическая химия и биохимия полости рта. Ситуационные задачи и задания**  
 Книга | 2023 • ГЭОТАР-МЕДИА  
 ☹ **Вне подписок** | Авторы: Глухов А.И.  
 Сборник задач и упражнений по биологической химии и биохимии полости рта написан преподавателями кафедры биологической химии Первого МГМУ им. <...>

Ещё

Заголовок – {Экономика предприятия} ✎ **Публикации** 87 **Диссертации** 7 **Патенты** 4

Сохранить запрос    Добавить в подборку    Экспорт    Облако ключевых слов    Аналитика

Найдено: 87    Показывать по: 10 20 50 100

**Экономика предприятия**  
 Книга | 2019 • ЮНИТИ-ДАНА  
**Полный доступ** | Авторы: Горфинкель Владимир Яковлевич

**Экономика предприятия (фирмы)**  
 Книга | 2009 • Научно-издательский центр ИНФРА-М  
**Полный доступ** | Авторы: Волков Ольгерд Иванович, Девяткин Олег Владимирович

**Экономика предприятия**  
 Книга | 2022 • Казанский национальный исследовательский технологический университет  
**Полный доступ** | Авторы: Гусарова Ирина Алексеевна, Николаева Ксения Владимировна, Пантелеева Юлия Владимировна

**Экономика предприятия**  
 Книга | 2017 • НАУЧНЫЙ КОНСУЛЬТАНТ  
**Полный доступ** | Авторы: Бигачева Елена Николаевна, Горлевская Людмила Эдуардовна

**Экономика предприятия. Ч. 2**  
 Книга | 2008 • Высшая школа  
**Полный доступ** | Авторы: Головачев Александр Степанович

С применением поисковых операторов и полей расширенного поиска

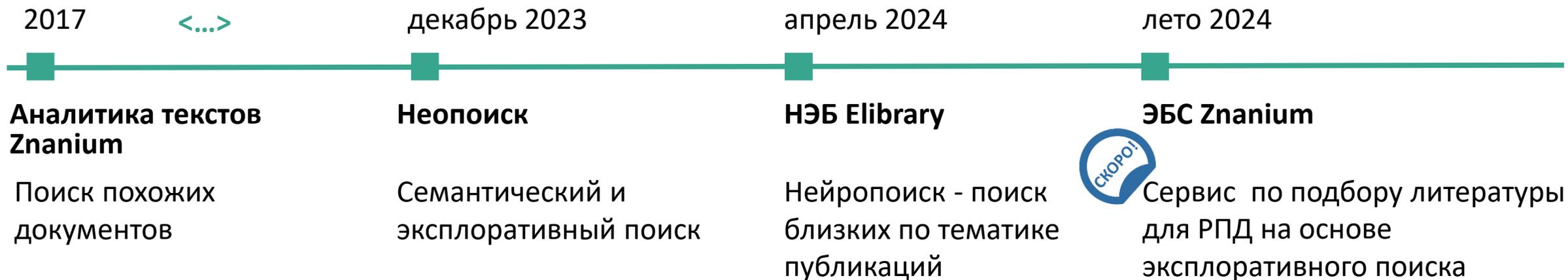
## Эксплоративный поиск

### Поиск документов по документам



Исследователь еще не разбирается в теме, не знает терминов и ключевых слов. У него есть только фрагмент текста или эталонный документ, и он “говорит” системе: “Найди всё, похожее на ЭТО”

*Примеры реализации эксплоративного поиска*



## Эксплоративный поиск

Добавить в подборку
Найти похожие
Экспорт
Облако ключевых слов

**Гроаем глубокое обучение с подкреплением. - (Серия «Библиотека программиста»)**

Книга | ПИТЕР • 2023 • Страницы: 464  
 Полный доступ • Авторы: Моралес Мигель • ISBN: 978-5-4461-3944-6

Мы учимся, взаимодействуя с окружающей средой, и получаемые вознаграждения и наказания определяют наше поведение в будущем. Глубокое обучение с подкреплением привносит этот естественный процесс в искусственный интеллект и предполагает анализ результатов для выявления наиболее эффективных путей движения вперед. Агенты глубокого обучения с подкреплением могут способствовать успеху маркетинговых кампаний, прогнозировать рост акций и побеждать гроссмейстеров в Го и шахматах. Давайте научимся создавать системы глубокого обучения на примере увлекательных упражнений, сопровождаемых кодом на Python с подробными комментариями и понятными объяснениями. Вы увидите, как работают алгоритмы, и научитесь создавать собственных агентов глубокого обучения с подкреплением, используя оценочную обратную связь. <br>16+ (В соответствии с Федеральным законом от 29 декабря 2010 г. № 436-ФЗ.)

Ссылки на публикации:  
 Полный доступ [↗ ibooks.ru](#)

+ Регулировка степени «похожести»

**Поиск похожих:** Гроаем глубокое обучение с подкреплением. - (Серия «Библиотека программиста»)

**Публикации**  
172

**Диссертации**  
3

**Патенты**  
39

Добавить в подборку
Экспорт
Облако ключевых слов
Аналитика

Найдено: 172      Показывать по: 10 20 50 100

- Глубокое обучение с подкреплением на Python. OpenAI Gym и TensorFlow для профи**  
 Книга | 2020 • ПИТЕР  
 Полный доступ | Авторы: Равичандиран Судхарсан
- Обучение с подкреплением для реальных задач: Пер. с англ.**  
 Книга | 2022 • БХВ-Петербург  
⊖ Вне подписок | Авторы: Ф. Уиндер
- Обучение с подкреплением: введение**  
 Книга | 2020 • ДМК Пресс  
 Полный доступ | Авторы: Барто Эндрю Дж., Саттон Ричард С., Слинкин Алексей Александрович
- Обучение с подкреплением на PyTorch: сборник рецептов**  
 Книга | 2020 • ДМК Пресс  
 Полный доступ | Авторы: Лю Юси (Хэйден), Слинкин Алексей Александрович
- Глубокое обучение с подкреплением. AlphaGo и другие технологии**  
 Книга | 2021 • ПИТЕР  
 Полный доступ | Авторы: Максим Лапань

# Схема индексной базы

**30 +**  
млн публикаций  
на старте

**390 тыс.**  
Диссертации

**7.2 млн**  
Патенты

**Клиентский контент**

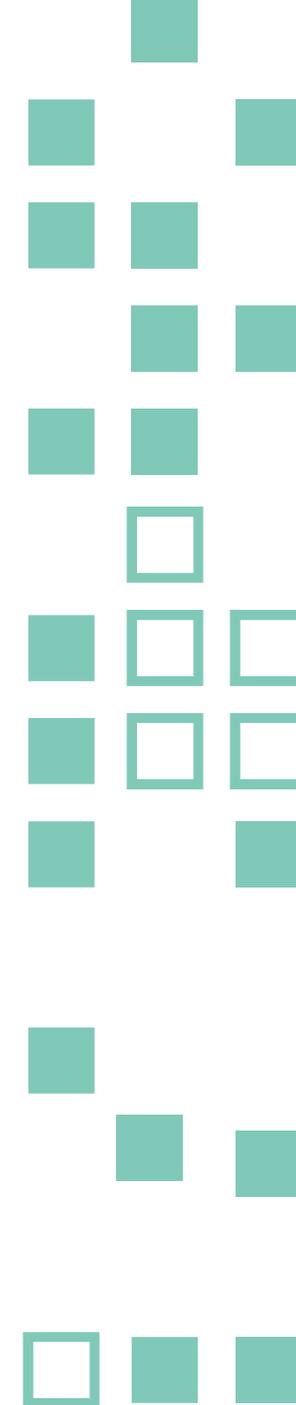
<b>Индивидуально</b> Каталог библиотеки	<b>Индивидуально</b> Репозиторий организации
--	---

**Базовый подписной контент**

<b>1.6 млн</b> Российские провайдеры	<b>11.5 млн</b> Зарубежные провайдеры
---	--

**Ресурсы открытого доступа**

<b>&gt; 10 млн</b> Open Access платформы	<b>≈ 1 млн</b> Открытые репозитории ВУЗов РФ и СНГ
---	---



## Наполнение индексной базы

### Российские провайдеры:

Elibrary

Znanium

ibooks

Консультант студента

Консультант врача

BookUp

*и еще 11 платформ*

**+ коллекции других российских провайдеров**

### Зарубежные провайдеры:

Cambridge University Press

Wiley

Oxford University Press

Springer

Royal Society of Chemistry

ScienceDirect

Emerald Publishing

World Scientific Publishing

British Medical Journals

Sage

American Physical Society

EDP Sciences

Institute of Electrical and Electronics Engineers

SAE International

*и еще 28 платформ*

**+ коллекции других зарубежных провайдеров**

### Ресурсы открытого доступа:

Pubmed

Hindawi

КиберЛенинка

### Препринты

arXiv

### Авторефераты диссертаций

ВАК и РГБ

### Патенты:

ФИПС

USPTO

Евразийское патентное общество

### Репозитории ведущих вузов

30+ репозиториев ВУЗов РФ и стран СНГ

**Электронный каталог библиотеки**

■ Общедоступный по подписке контент

■ Индивидуальные настройки



# Дедубликация: обогащение метаданных и поиска

НЕОПОИСК

[О сервисе](#)
[Каталог](#)
[Мой неописк](#)
[Смылова Александра](#)

Добавить в подборку
Найти похожие
Экспорт
Облако ключевых слов

## Базы данных: Работа с распределенными базами данных и файловыми системами на примере MongoDB и HDFS с использованием Node.js, Express.js, Apache Spark и Scala

Книга | Научно-издательский центр ИНФРА-М • 2023 • Страницы: 235  
Полный доступ • Авторы: Мартишин Сергей Анатольевич, Симонов Владимир Львович, Храпченко Марина Валерьевна • ИНСТИТУТ СИСТЕМНОГО ПРОГРАММИРОВАНИЯ  
РОССИЙСКОЙ АКАДЕМИИ НАУК; Российский государственный социальный университет • ISBN: 978-5-16-015133-5 • EISBN: 978-5-16-107636-1 • МОСКВА

Предназначено для изучения и создания информационных систем и приложений с использованием распределенных баз данных и файловых систем. Для создания информационных систем на основе распределенных баз данных используется серверная часть стека программного обеспечения MEAN, включающая в себя СУБД NoSQL-типа MongoDB, платформу Node и фреймворк Express. Рассмотрены также основы администрирования и репликации в СУБД MongoDB. В качестве инструмента моделирования объектов баз данных используется Mongoose.js. С целью создания фронтенд частей сайтов и веб-приложений изучается применение свободного программного обеспечения Bootstrap, HTML5 Boilerplate, Initializr. Приводится детальный процесс разработки и тестирования API REST-приложения на основе модели MVC. В результате обучающиеся будут способны создавать полноценную распределенную информационную систему на основе современных СУБД NoSQL-типа. Для создания приложений также используются распределенные хранилища данных (больших массивов неструктурированной информации) на примере Hadoop Apache с инструментарием Spark и Scala. Предназначено для бакалавров, магистрантов, аспирантов и специалистов в области проектирования распределенных информационных систем на основе современных СУБД NoSQL-типа, создания мобильных и интернет-приложений. Также полезно для лиц, занимающихся самообразованием.

**Ссылки на публикации:**  
📖 Печатное издание [srv.library.itmo.ru](http://srv.library.itmo.ru)  
🔗 Вне-подписок [znanium.ru](http://znanium.ru)

Ключевые слова (авторские)

apache spark, express.js, hdfs, mongodb, node.js, scala, базы данных, распределенные базы данных, файловые системы

Тематики:

**ГРНТИ:** 20.23 Информационный поиск, 50.41 Программное обеспечение вычислительных машин, комплексов и сетей

Информация из  
**znanium**  
электронная библиотечная система

# Взаимодействие с поставщиками контента

Лицензионный договор



Соглашение



Издатель

Владелец платформы

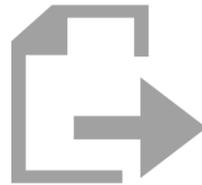
Неоиск



Передает оригинал-макет в PDF



Преобразует PDF в txt



Передает txt в Неоиск



Проводит индексацию txt и удаляет файлы из буфера



Размещает согласованный вариант “видимости” платформы

## Сотрудничество с Поставщиками контента

### Возможности сотрудничества:

#### 1. Передача по API

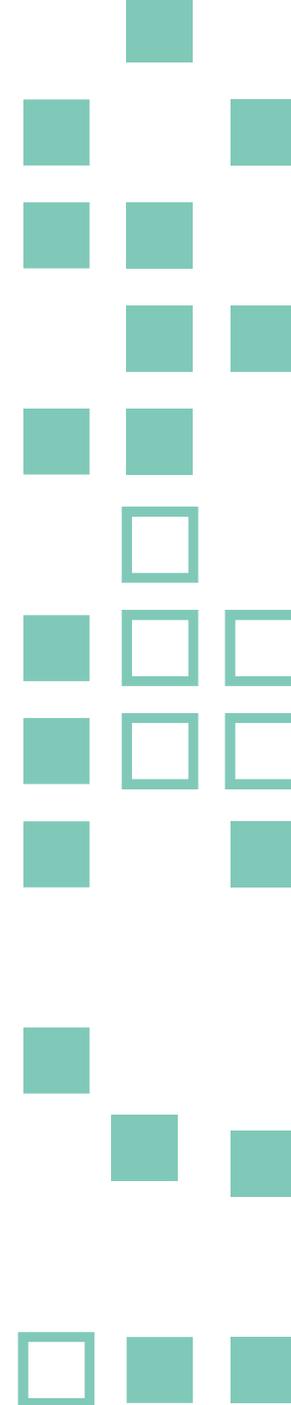
- Весь фонд - Метаданные - Полные тексты – Состав подписок в разрезе абонентов
- Метаданные - Полные тексты – Состав подписок в разрезе абонентов
- Метаданные – Состав подписок в разрезе абонентов

#### 2. Передача выгрузки в формате Rusmarc

- Весь фонд - Метаданные – Состав подписок в разрезе абонентов
- Метаданные – Состав подписок в разрезе абонентов

### Преимущества сотрудничества

- + «Неопоиск» становится «рекламной площадкой» для Поставщика за счет привлечения дополнительного трафика
- + повышение статистики чтения документов - читать полный текст пользователи «Неопоиска» переходят в ридер на сайте Поставщика
- + при наличии полного текста в выдаче выводятся те документы, у которых метаданные не соответствуют поисковому запросу, но есть совпадения в тексте
- + автоматическое добавление новинок из подписки Абонента в поисковый массив



## To discover - узнать, обнаружить, увидеть

### Что увидеть:

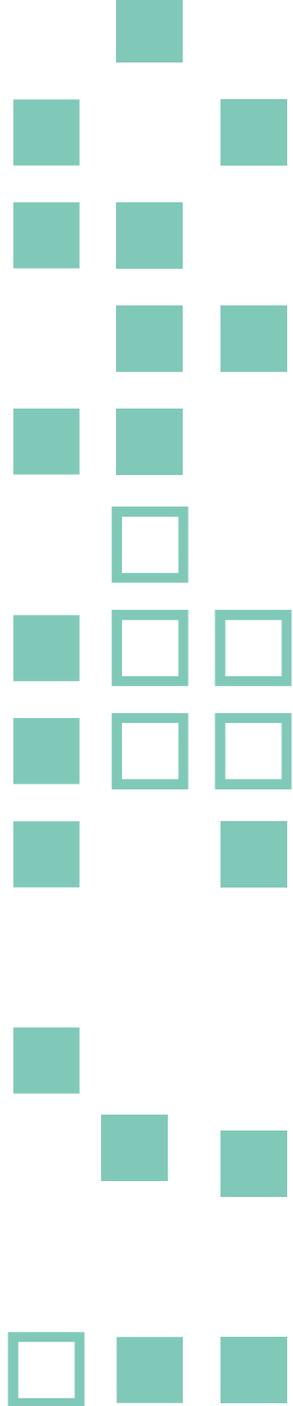
- **пользовательское намерение (интент)** (с учетом уточнений по неограниченному количеству полей расширенного поиска, фасетным фильтрам, поисковым операторам)
- **связанность** пользовательского намерения и текстов публикаций

### Где обнаружить:

- в метаданных и **в проиндексированных полных текстах** для поиска по ним с применением современных алгоритмов
- **в полных текстах печатных публикаций** (при наличии электронных копий даже на других платформах и даже не в подписке)

### Что узнать:

**метаданные для обогащения публикаций** с других платформ (аффилиации, ключевые слова, рубрики ГРНТИ и пр.)



НЕОПОИСК

# Приглашаем к сотрудничеству!

**Нестерова Альбина Николаевна**

Генеральный директор

Тел: +7 (495) 859-48-60

Email: [albina\\_nik@infra-m.ru](mailto:albina_nik@infra-m.ru)

