

О стратегии развития Национальной научно-  
информационной инфраструктуры

ГПНТБ  
РОССИИ

# Ресурсы и услуги ГПНТБ России для нужд школьного образования. Растим нового исследователя

**Карауш Александр Сергеевич,**

Генеральный директор,  
Государственная публичная  
научно-техническая библиотека России, к.т.н,  
Вице-президент, Национальная библиотечная ассоциация  
«Библиотеки будущего»

# План сообщения

1. Ресурсы и информация о ресурсах
2. Новейшие имеющиеся разработки в современных условиях
3. Четвертая парадигма науки
4. Создание российской научно-информационной инфраструктуры
5. Существующие инструменты научных коммуникаций
6. Н<sup>2</sup>И<sup>2</sup> – цели, задачи, принципы
7. Заключение

Национальная научно-информационная инфраструктура (Н<sup>2</sup>И<sup>2</sup>) - совокупность объектов, которые предоставляют исследовательскому сообществу ресурсы и услуги, связанные с поиском, сбором, изучением, обработкой, анализом, синтезом и распространением научной информации.

# Ресурсы и информация о ресурсах

1. Специализированный экспертный фонд всех типов и видов по науке, технике и смежным дисциплинам (9 млн.).
2. Удаленные электронные ресурсы (Национальная подписка, базы данных собственной генерации, электронно-библиотечные системы...).
3. Информационные системы построения свертки документов и индексов поиска, построенные на технологиях разработки самой популярной в РФ АБИС «ИРБИС64».
4. Системы изучения и формирования ландшафтов знаний (ГРНТИ, УДК, ББК, ...)

# Новейшие имеющиеся разработки в современных условиях

1. Хранение документов (электронные форматы хранения).
2. Поиск документов (поисковые системы).
3. Анализ информационных источников (электронные энциклопедии).
4. Рекомендательные функции (развитие навыков на основе информационных источников) – электронные системы «вопрос-ответ».
5. Воспитательные функции (оцифровка).
6. Онтологии.

**Культуры  
исследова-  
тельской  
деятельности**



**Навыков  
научного  
поиска**



**Доступа  
к легитимным  
источникам  
актуальной  
информации**

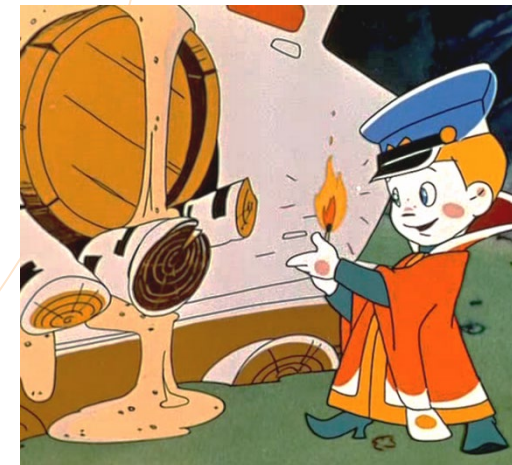
# Отсутствие у педагогов и обучающихся

**Навыков  
исследования  
на основе  
имеющихся  
данных**

**Навыка  
предвари-  
тельного  
анализа  
данных**



**Данных  
о предыдущих  
работах**



# Четвертая парадигма науки меняет научные коммуникации

# Смена научной парадигмы

Жизненный цикл исследования неизменен:



Но меняются инструменты исследователей:

## 1. Экспериментальная наука – последние тысячелетия

Описание природных явлений

## 2. Теоретическая наука – последние века

Законы Ньютона, уравнения Максвелла...

## 3. Вычислительная наука – последние десятилетия

Компьютерное моделирование сложных явлений

## 4. Data-driven Science – СЕГОДНЯ

Изучение наборов данных из разных источников:

- Научное оборудование
- Результаты моделирования
- Сети измерительных устройств

**Отставание в адаптации к парадигме Data-driven Science приводит к:**

1. Низкой производительности
2. Низкой конкурентоспособности
3. Технологическому отставанию

# Парадигма Data-driven Science меняет...

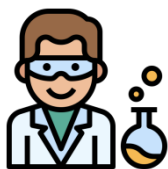
## Систему научных коммуникаций

### Публикация

- название, авторы
- аннотация
- ключевые слова

Текст

Исследователь



Объекты научной коммуникации

Субъект научной коммуникации

### Публикация

- название, авторы
- аннотация
- ключевые слова

Текст +



методы  
объекты  
свойства  
эффекты  
модели

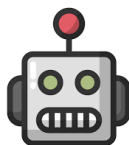
Новые объекты:  
Датасеты,  
3D-модели...



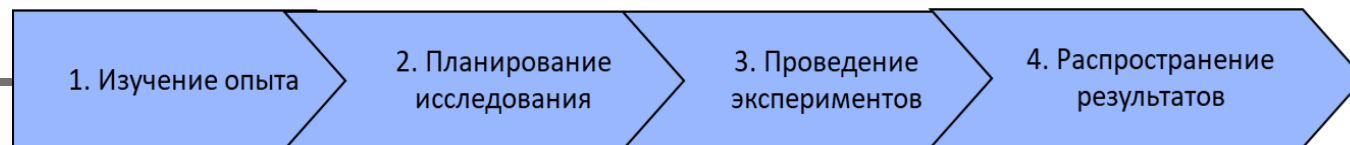
Исследователь



Цифровой  
помощник



## Инструменты жизненного цикла исследований



Научно-информационная инфраструктура



# Пример из настоящего: Подготовка научных обзоров

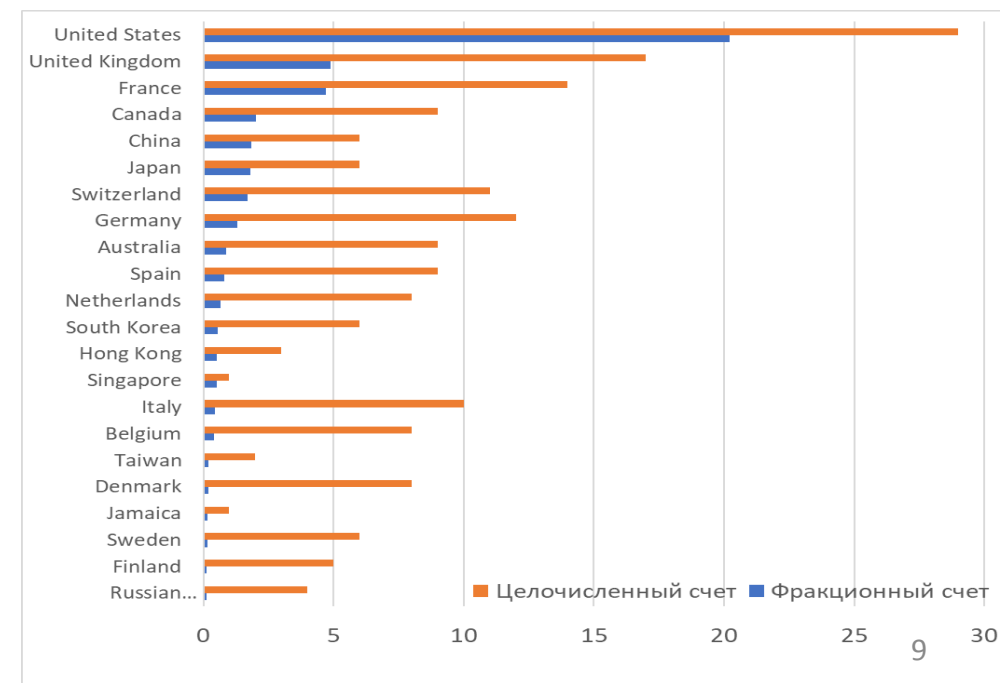
Обзор – весомый элемент системы разделения труда исследователей:

- 1. Роль.** Систематизируют знания о научном ландшафте.
- 2. Объем.** Занимают в различных областях наук до 16% публикаций.
- 3. Трудоемкость.** 10%-45% времени исследователей – работа с научной информацией
- 4. Массовость.** Частный случай обзора – вводная часть статьи.
- 5. Авторитетность.** Чаше цитируются.
- 6. Инструмент мониторинга научной повестки.**

В России культура подготовки обзоров отстаёт:

1. Доля публикаций России в 2020 году (Scopus):  
все публикации: **3,6%** (8-е место)  
только обзоры: **2,3%** (14-ое место)  
только обзоры по медицине: **1,7%** (17-ое место)

2. Присутствие стран в высоко цитируемых обзорах:



# Пример из настоящего: Подготовка научных обзоров



## ТРУДНОСТИ

Ограничения эксперта:

- биологические
- социально-политические
- личностные

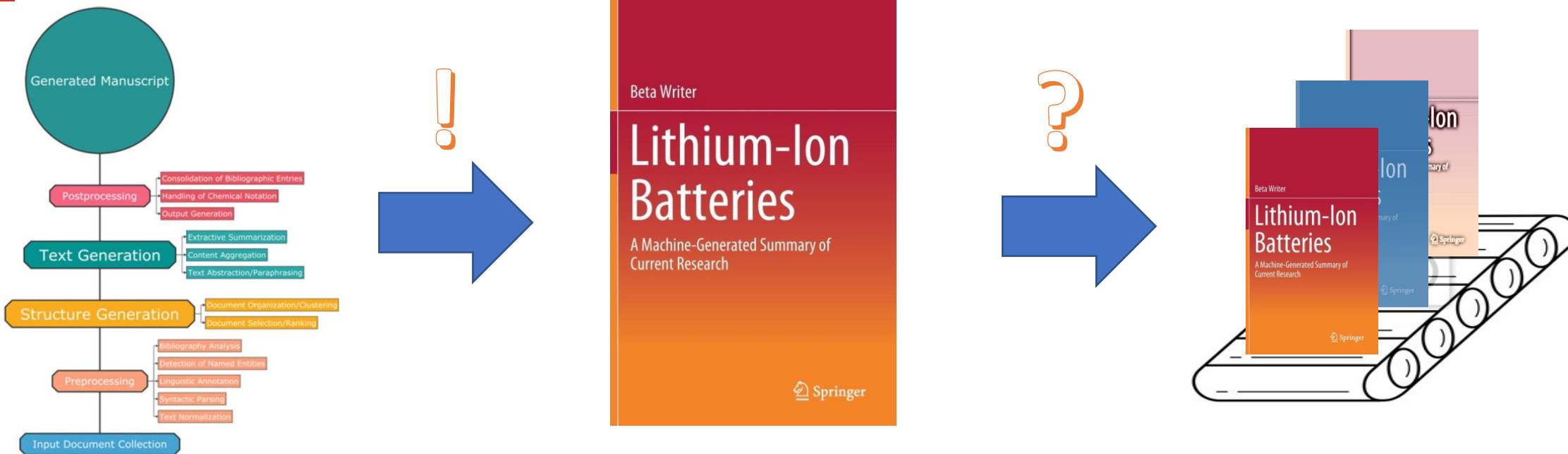


## Необходимы:

1. Инфраструктура доступа к НИ
2. Методы обработки НИ
3. Рынок услуг в области научной информации (НИ)

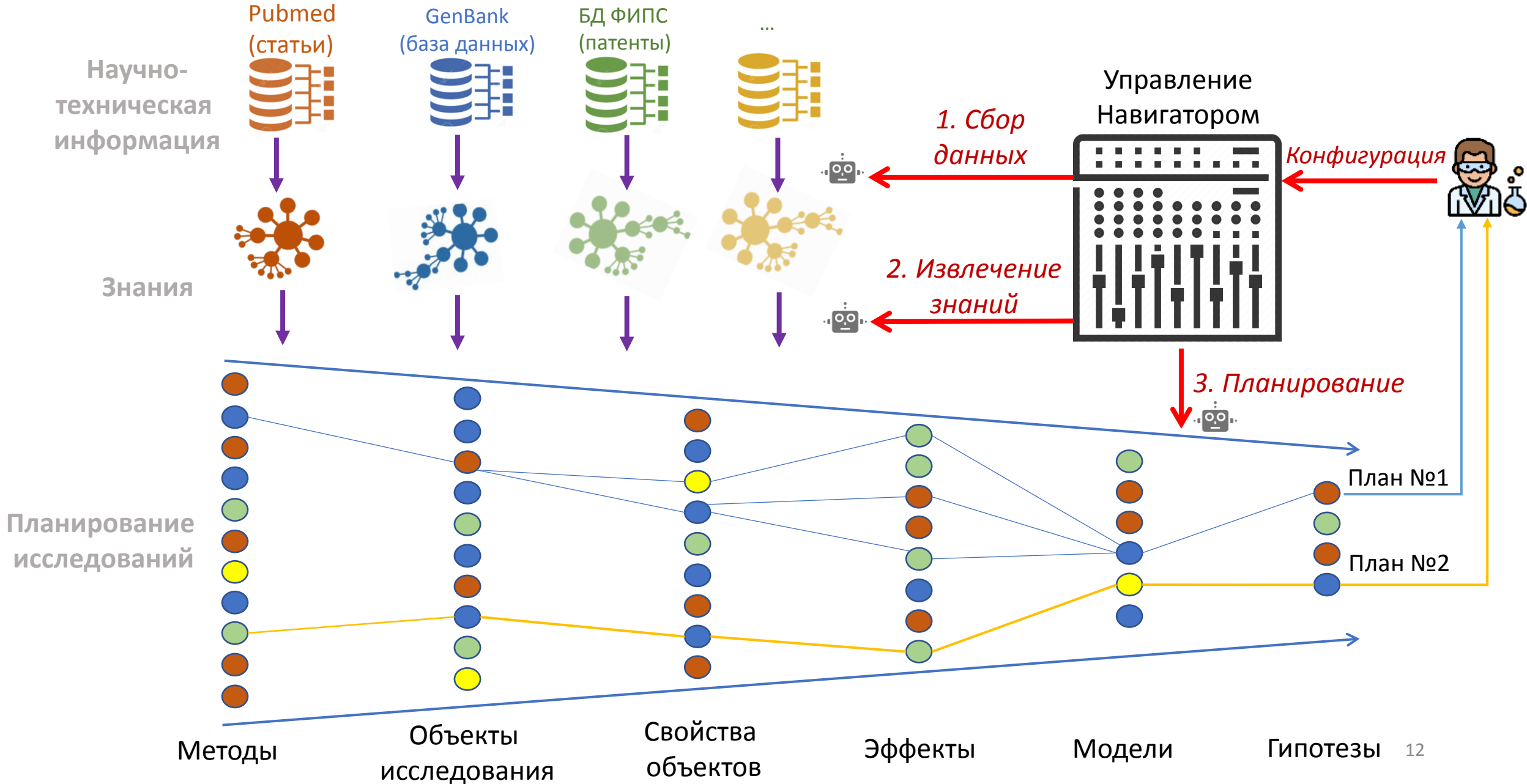
Систематизация научного знания – традиционный способ преодоления информационной перегрузки

# Пример из настоящего: Подготовка научных обзоров



В 2019 г. опубликован первый научный обзор, подготовленный методами искусственного интеллекта. Возможен ли конвейер научных обзоров?

# Пример из будущего: Навигатор для планирования исследований



# Попытки создания российской научно-информационной инфраструктуры

# ГСНТИ РФ: государственная система научно-технической информации



<b>Конечные клиенты</b>	Школьник / студент (будущий исследователь)	Исследователь	Коммерческий заказчик	Государственный заказчик	Административный работник НО
<b>Внутренние клиенты</b>	ФОИВы	РОИВы	Вузы	НИИ	Фонды
<b>Потребности</b>	Определить проблематику, цели и задачи исследования	Найти выход на заказчика, понять технические требования	Найти источник средств для финансирования	Провести исследование и получить результаты	Урегулировать вопросы интеллектуальной собственности
<b>Каналы услуг</b>	Системы науч. цитирования	Сайты научных журналов	Электронная почта	ЕПГУ	Коммерческие каналы

**Функц. области и сервисы**

Значимый сервис

<p><b>Управление научным проектом</b></p> <ul style="list-style-type: none"> <li>Виртуальный проектный офис</li> <li>Конструктор плана проекта</li> <li>Калькулятор экономики проекта</li> <li>Сервис поиска зеркальных лабораторий</li> <li>Помощник по составлению отчетности для грантополучателя / исполнителя</li> <li>Сервис поиска и подбора команды (HR-сервис)</li> <li>Маркетплейс ЦКП / УНУ / расходных материалов</li> <li>Помощник по формированию, согласованию и подписанию договоров</li> <li>Формирование real-time отчетности по проекту</li> <li>Дэш борд статуса выполнения работ для рук. проектов</li> </ul>	<p><b>Личный кабинет исследователя</b></p> <p>Управление РИД</p> <ul style="list-style-type: none"> <li>Поисковая платформа патентной информации</li> <li>Сервис «Патентные пулы»</li> <li>Сервис оборота прав</li> <li>Экспертно-аналитическое сопровождение проектов, патентоспособность решения</li> <li>Сервис учета «роялти»</li> <li>Сервис субсидирования зарубежного патентования</li> </ul>	<p><b>Исследовательская деятельность</b></p> <p>Научный Data-hub</p> <ul style="list-style-type: none"> <li>Сервис доступа к большим данным</li> <li>Сервис доступа к виртуальным вычислениям</li> <li>Сервис доступа к биобанкам открытого типа</li> <li>Платформа опросов верифицированных пользователей</li> <li>Помощник по анализу тем исследований</li> <li>Помощник по семантическому анализу гипотез</li> </ul>	<p><b>Коммерциализация</b></p> <p>Научная биржа</p> <ul style="list-style-type: none"> <li>Платформа НкС (НИОКР как сервис)</li> <li>Сервис формализации требований и сопровождения переговоров</li> <li>Сервис визуализации дорожных карт</li> <li>Инструмент для самооценки компетенций исследователя</li> <li>Маркетплейс экспертов по технологическому аудиту</li> <li>Сервис проверки бизнес-идеи</li> <li>Консультации по масштабированию бизнеса</li> <li>Сервис одного окна в ЦТТ</li> <li>Помощник по оценке соответствия обязательным требованиям к бизнесу</li> <li>Поиск промышленного подрядчика на изготовление опытного образца</li> <li>Помощник по оформлению ТЗ на опытный образец</li> </ul>
--	--	---	---

**НАВИГАТОР ДЛЯ ПЛАНИРОВАНИЯ = ИССЛЕДОВАНИЙ**

<p><b>Просветительская деятельность</b></p> <ul style="list-style-type: none"> <li>Портал популяризации науки (календарь мероприятий, записи лекций, научно-популярные видео и публикации)</li> <li>...</li> </ul>	<p><b>Регулирование науки</b></p> <ul style="list-style-type: none"> <li>Сквозная прослеживаемость (Мониторинг сферы науки)</li> <li>Статистическая отчетность</li> <li>Контроль гос. заданий</li> </ul>	<p><b>Управление НО</b></p> <ul style="list-style-type: none"> <li>Сервис управления НО (вкл. калькулятор эк. показателей, распределение задач и отслеживание статусов)</li> <li>Контрактная служба научной организации</li> </ul>	<p><b>Поддержка и мотивация</b></p> <p>Портал инструментов развития</p> <p>Площ адка-агрегатор инструментов развития</p> <ul style="list-style-type: none"> <li>Поддержка по финансовым и правовым вопросам</li> <li>Первичный помощник по выбору инструмента финансирования</li> <li>Помощник по оформлению заявок</li> <li>Формирование смарт-контрактов</li> <li>Агрегатор научных новостей</li> <li>Навигатор крупных научных проектов (предложения для участия)</li> <li>Категоризация исследований</li> <li>Премирование ученых</li> </ul>
--	--	--	--

**Данные**

Реестр патентов	Реестр публ. патентных заявок	Реестр журналов	Реестр публикаций	Реестр научных организаций	Реестр типовых договоров	Каталог ЦКП	Реестр оборудования	Ролевой состав команд	Реестр обязательных требований к бизнесу
Реестр 3D-моделей	База диссертаций	Реестр грантов	Цифровой профиль исследователя	Цифровой профиль научной организации	Формы и шаблоны	Каталог УНУ	Реестр расходных материалов	База научных текстов (для ML)	Финансовая отчетность НО

## Некоторые проблемы российской научно-информационной инфраструктуры

1. Недоступность или низкое качество метаданных.
2. Доступ к полным текстам публикаций затруднен (при этом есть иллюзия их доступности).
3. Результаты научных экспериментов (в т. ч. дорогостоящих), на основании которых публикуются научные работы, остаются недоступными.
4. Ряд профинансированных проектов в области развития научно-информационной инфраструктуры не дали ожидаемого эффекта.
5. Отсутствие стратегии цифровизации науки.



# Н<sup>2</sup>И<sup>2</sup> – цели, задачи, принципы

# Цель и задачи программы

## ЦЕЛЬ. Развитие инфраструктуры рынка НИ в соответствии с парадигмой Data-driven Science

**Задача 1.** Создание научного задела для разработки высокотехнологичных продуктов

Метаданные

Методы ИИ

Онтологии

Методы NLP

**Задача 2.** Создание распределенной инфраструктуры для сбора, хранения, обработки, обмена и доступа к массивам НИ в машиночитаемой форме

Системы

Стандарты

Сервисы

Протоколы

**Задача 3.** Создание системы мотивации для всех пользователей инфраструктуры

Целевая аудитория

Потребности

Кадры

- Барьеры**
1. Сложность извлечения знаний из данных и текстов.
  2. Разнородные подходы к систематизации знаний.
  3. Распределенность источников.

1. Организации «сидят» на данных.
2. Авторское право.
3. Рассогласованность форматов.
4. Большой объем.
5. Организационная разобщенность.

1. Незрелость рынка.
2. Устаревшая СРТ.
3. Недостаток квалифицированных специалистов.

# Принцип приращения ценности научной информации



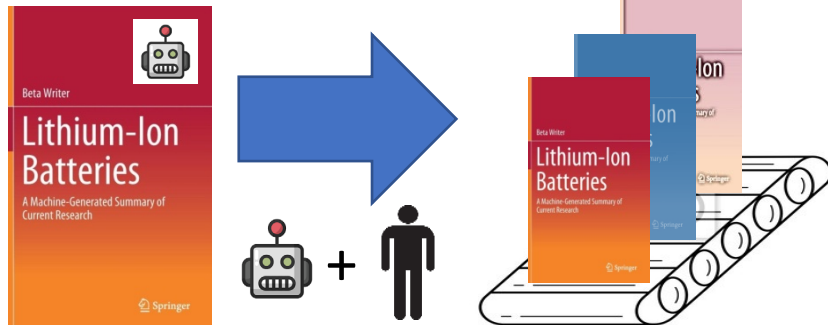
Виды продуктов	Уровни обработки научной информации	Новые свойства	Барьеры
4. Интеллектуальные сервисы	Анализ, обобщение и генерация знаний	Машинное создание производных (вторичных) знаний	Научный фронтир
3. Базы знаний	Извлечение и сопоставление знаний	Машинное «понимание» знаний	Междисциплинарная рассогласованность, сложность извлечения знаний
2. Каталоги научных публикаций	Агрегирование и структурирование массивов научной информации	Потоковая (машинная) обработка, поиск	Большой объем, рассогласованность форматов
1. Доступ к публикациям	Хранение и предоставление доступа	Сохранность, доступность	Организационная разобщенность, «монополия» на данные

# Ориентация на новое поколение продуктов рынка НИ

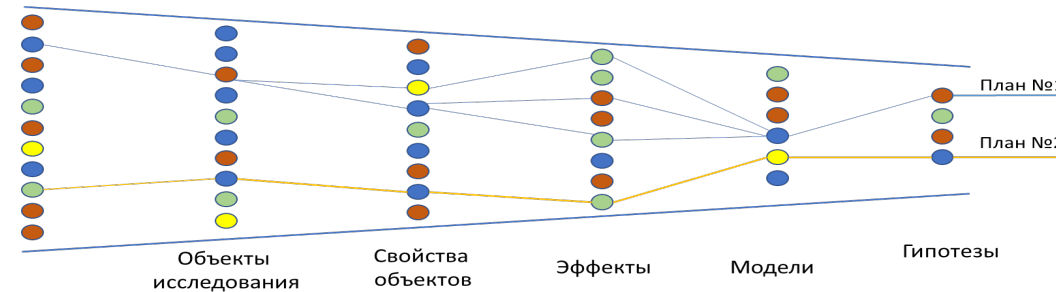


## Сервис подготовки научных обзоров на основе алгоритмов ИИ

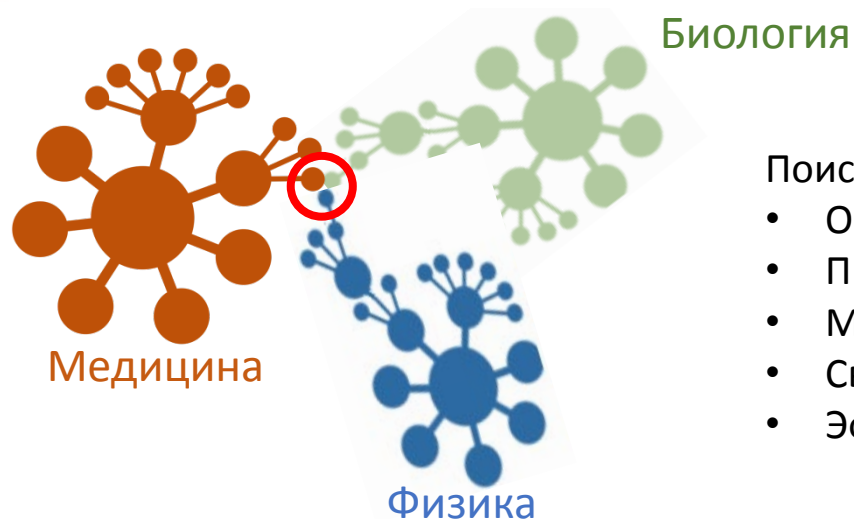
Первый обзор, написанный ИИ (2019, Springer)



## Навигатор для планирования исследований



## Сервис поиска близких исследований в других дисциплинах

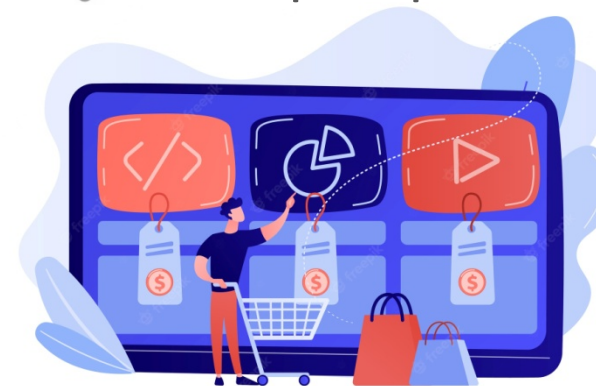


Поиск по:

- Объектам
- Проблемам
- Методам
- Свойствам
- Эффектам



## Маркетплейс научных инфосервисов



- Информационный поиск
- Систематические обзоры
- Мета-анализы
- Патентный поиск

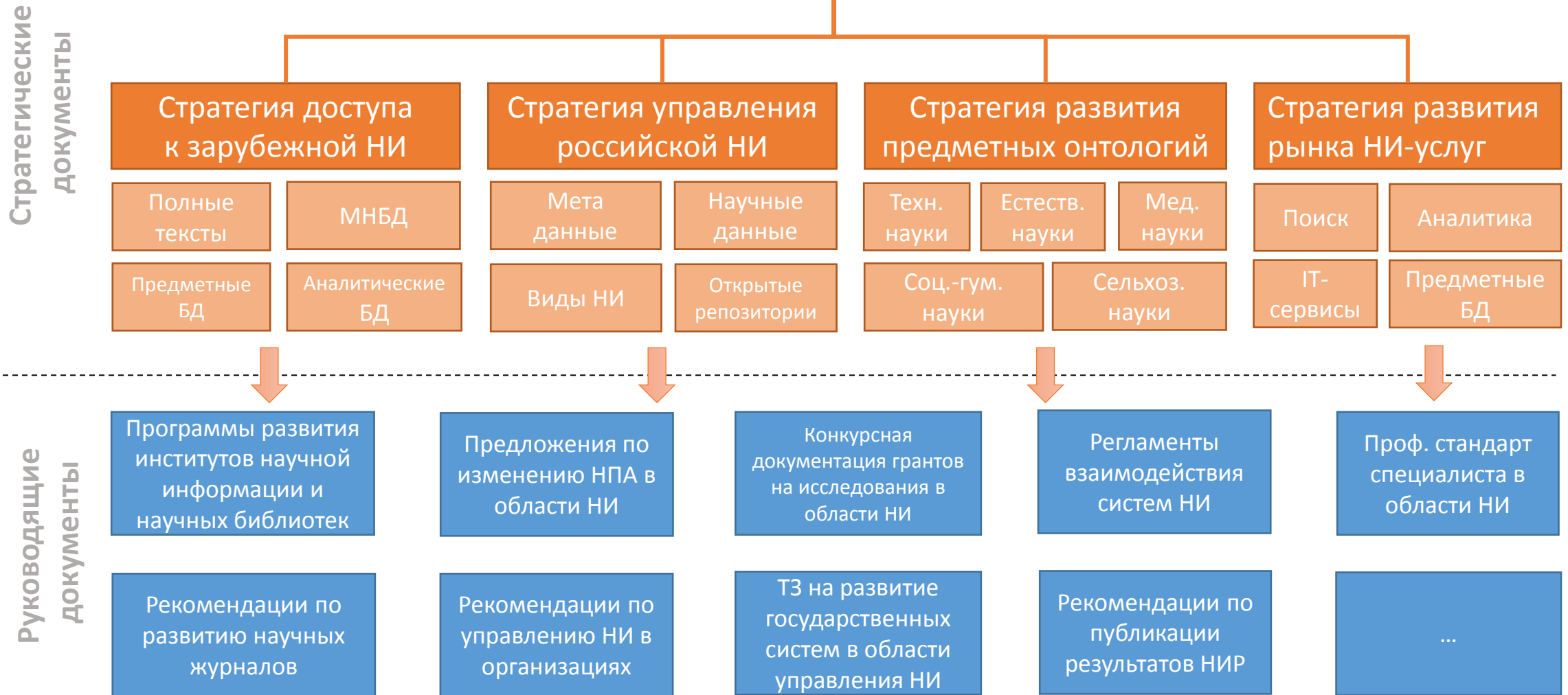
## Другие принципы:

1. Единые правила машиночитаемой НИ
2. Безбарьерный доступ к массивам НИ
3. Открытость рынка решений и ИТ разработок
4. Глобальная конкурентоспособность
5. Impact investing

## Механизмы реализации

- Стратегия развития национальной научно-информационной инфраструктуры
- Консорциум по развитию национальной научно-информационной инфраструктуры
- Восстановление функций систематизации научно-информационного пространства восстановить в существующих центрах научной информации
- Определение центров предметных компетенций,
- Целевые конкурсы Минобрнауки, РФФИ, РФФИТ
- Подготовка научных инфодайверов

# Концепция развития национальной научно-информационной инфраструктуры



# Тематические блоки программы

Задача 1.  
Научный задел

## 1. Систематизация

- Анализ и развитие предметных онтологий
- Сопоставление онтологий смежных предметных областей
- Выделение научных терминов в тексте

## 2. Извлечение знаний

- Обработка текстов на ЕЯ (NLP)
- Семантический анализ текстов
- Формирование логических выводов
- Генерация аналитических текстов

Задача 2.  
Инфраструктура

## 3. Ресурсы (источники НИ)

- Обязательный экземпляр
- Национальная подписка
- Требования к научным издателям
- Развитие репозиторий
- Открытые данные

## 4. Каталогизация

- Системы идентификации сущностей
- Системы классификации
- Системы федеративного поиска
- Протоколы и форматы обмена документами
- Управление научными данными

Задача 3.  
Мотивация

## 5. Продукты

- Конвейер обзоров
- Ассистент эксперта
- Ассистент патентоведа
- Предметные базы данных
- Маркетплейс научных инфосервисов

## 6. Распространение

- Мероприятия по внедрению практик регулярной подготовки обзоров в организациях и лабораториях
- Курсы обучения библиотечных специалистов
- Хакатоны

# Заключение



## Почему этого нельзя не делать

Информационная инфраструктура науки =  
транспортная инфраструктура региона

С плохой инфраструктурой как-то ездить будут, но:

- Потери времени
- Дополнительные расходы
- Низкое качество жизнедеятельности
- Низкая привлекательность для внешних акторов
- Демотивация потребителей инфраструктуры к её использованию

Потери от плохой инфраструктуры велики  
и при этом трудно измеримы.



## Так! Ну ладно! А что же делать!?

- Интерес к исследовательской и научно-популярной тематике в противовес «рекламной» продукции
- Возрождать клубы научно-популярной направленности (фантастика, исследования необычных явлений и пр.)
- Формирование исторического ландшафта отечественного научного опыта (история науки)
- Научное краеведение

# СПАСИБО ЗА ВНИМАНИЕ!

**Карауш Александр Сергеевич**

a@karaush.ru / ask@gpntb.ru

t.me/akaraush

Национальная научно-информационная инфраструктура: проблемы, задачи и перспективы / А. Е. Гуськов [и др.] // Управление наукой и наукометрия. 2022. Т. 17, № 3. С. 380-407. DOI: <https://doi.org/10.33873/2686-6706.2022.17-3.380-407>